Report MPM-16

Final Report

26-1107-0097-001

Investigation of the Asphalt Pavement Analyzer (APA) Testing Program in Nebraska

Yong-Rak Kim, Ph.D.
Associate Professor
Department of Civil Engineering
University of Nebraska-Lincoln

Minki Hong
Graduate Research Assistant

2008

Nebraska Transportation Center
262 WHIT
2200 Vine Street
Lincoln, NE 68583-0851
(402) 472-1975

"This report was funded in part through grant[s] from the Federal Highway Administration [and Federal Transit Administration], U.S. Department of Transportation. The views and opinions of the authors [or agency] expressed herein do not necessarily state or reflect those of the U. S. Department of Transportation."

Investigation of the Asphalt Pavement Analyzer (APA) Testing Program in Nebraska

Yong-Rak Kim, Ph.D.
Associate Professor
Department of Civil Engineering
University of Nebraska-Lincoln

Minki Hong
Graduate Student
Department of Civil Engineering
University of Nebraska-Lincoln

A Report on Research Sponsored by

Mid-America Transportation Center

University of Nebraska-Lincoln

March 2008

Technical Report Documentation Page

1. Report No MPM-16	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Investigation of the Asphalt Pavement Analyzer (APA) Testing Program in Nebraska		5. Report Date March 2008
		6. Performing Organization Code
7. Author/s Yong-Rak Kim and Minki Hong		8. Performing Organization Report No. MPM-16
9. Performing Organization Name and Address University of Nebraska-Lincoln (Department of Civil Engineering) 2200 Vine St. 362M Whittier Research Center Lincoln, NE 68583-0856		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. 26-1107-0097-001
12. Sponsoring Organization Name and Address Nebraska Department of Roads (NDOR) 1400 Highway 2, PO Box 94759 Lincoln, NE 68509		13. Type of Report and Period Covered July 2006-March 2008
		14. Sponsoring Agency Code MATC TRB RiP No. 11916
15. Supplementary Notes		

16. Abstract

The asphalt pavement analyzer (APA) has been widely used to evaluate hot-mix asphalt (HMA) rutting potential in mix design and quality control-quality assurance (QC-QA) applications, because the APA testing and its data analyses are relatively simple, rapid, and easy. However, as demonstrated in many studies and also experienced by the state of Nebraska, APA testing is in question due to its high testing variability and a lack of sufficient correlation with actual filed performance. The primary objective of this research was to find critical materials and/or mixture design factors affecting APA test results so as to eventually improve the current APA testing program in Nebraska. In addition to that, development of models to predict APA rut performance with given properties of HMA mixture ingredients and mixture design characteristics were also attempted. To find variables affecting APA rut results and the extent of these variables, *SP-4* mixture data from Nebraska and HMA mixture data from Kentucky were statistically analyzed using the multiple linear regression method considering six factors (binder PG, aggregate gradation, nominal maximum aggregate size, aggregate angularity, air voids in mixture, and asphalt content in mixture) as probable candidates for significantly affecting APA rut results. For a detailed characterization of gradation effects, three indicators (gradation density, fineness modulus, and restricted zone) were considered, and each of them was used for each statistical analysis. Results from analyses demonstrated that the binder PG was the only variable that always shows significant impact on APA rut results, which is in good agreement with other studies. Predicting models developed through the results of multiple linear regression analysis and the artificial neural network technique presented a relatively low level of model adequacy which can be observed by the coefficients of determination and cross-plots between predicted APA rut values and the measured APA rut data. More data would be helpful to confirm the findings from this research and also to develop a better prediction model.

17. Key Words Asphalt Pavement Analyzer (APA), Rutting, Superpave, Hot-Mix Asphalt (HMA), Pavement			18. Distribution Statement
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. Of Pages 47	22. Price

Table of Contents

Chapter 1 Introduction viii
1.1 Research Objectives 2
1.2 Research Scope 2
1.3 Organization of the Report 2
Chapter 2 Literature Review 4
Chapter 3 Research Methodology 12
3.1 Multiple Linear Regression Analysis 13
3.2 Artificial Neural Network Approach 18
3.3 APA Data for Analyses 20
Chapter 4 Analysis Results and Discussion 25
4.1 Analysis of Significance 25
4.2 Development of Prediction Models 31
4.2.1 Prediction Models from the Multiple Linear Regression Analysis 31
4.2.2 Prediction Models from the Artificial Neural Network 35
Chapter 5 Concluding Remarks 39
5.1 Conclusions 39
5.2 Recommended Further Studies 40
5.3 NDOR Implementation Plan 40
References 42
Appendix A 44
Appendix B 47

List of Figures

Figure 2.1 Effect of nominal maximum aggregate size on rutting 6
Figure 2.2 Relationship between binder PG and rutting at different asphalt content 7
Figure 2.3 PURWheel rut depths at different FAA values 8
Figure 3.1 Basic structure of the artificial neural network 19
Figure 3.2 Factors selected for analyses 20
Figure 3.3 A gradation curve with its density characteristics on the 0.45 power chart 22
Figure 3.4 Data sheet developed for statistical analyses 24
Figure 4.1 Variation of *p*-value with different FAD in gradation 28
Figure 4.2 A gradation curve with FAD value greater than 27 28
Figure 4.3 Variation of *p*-value with different CAD in gradation 29
Figure 4.4 Four gradation curves with CAD value greater than 25 30
Figure 4.5 Cross-plots between measured and predicted using model 1 34
Figure 4.6 Cross-plots between measured and predicted using model 2 34
Figure 4.7 Cross-plots between measured and predicted using model 3 35
Figure 4.8 Process of the artificial neural network modeling 36
Figure 4.9 Cross-plots between measured and predicted (training stage) 37
Figure 4.10 Cross-plots between measured and predicted (validation stage) 37
Figure 4.11 Cross-plots between measured and predicted (testing stage) 38
Figure 4.12 Multiple linear regression vs. artificial neural network 38

List of Tables

Table 2.1 Average rut depths for WesTrack sections 6
Table 2.2 Analysis of variance for rut depths of mixes 9
Table 2.3 Statistical results 10
Table 3.1 ANOVA table from the multiple linear regression analysis 13
Table 3.2 ANOVA table resulting from a real multiple regression analysis 15
Table 3.3 Parameter estimate from the multiple linear regression analysis 17
Table 3.4 R^2 and *adj*. R^2 values from the multiple linear regression analysis 18
Table 4.1 Statistical analysis results 26
Table 4.2 Prediction models from the multiple linear regression analysis 32

List of Abbreviations

Accelerated Loading Facility (ALF)
Analysis of Variance (ANOVA)
Asphalt Pavement Analyzer (APA)
Coarse Aggregate Density (CAD)
Fine Aggregate Angularity (FAA)
Fine Aggregate Density (FAD)
Fineness Modulus (FM)
Hot Mix Asphalt (HMA)
Mean Square Due to Regression (MSR)
Mean Square Error (MSE)
National Cooperative Highway Research Program (NCHRP)
Nebraska Department of Roads (NDOR)
Nominal Maximum Aggregate Size (NMAS)
Performance Grade (PG)
Quality Control-Quality Assurance (QC-QA)
Restricted Zone (RZ)
Strategic Highway Research Program (SHRP)
Voids in Mineral Aggregates (VMA)

Acknowledgments

The authors thank the Nebraska Department of Roads (NDOR) for the financial support needed to complete this study. In particular, the authors thank NDOR Technical Advisory Committee (TAC) members for their technical support and invaluable comments.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Abstract

The asphalt pavement analyzer (APA) has been widely used to evaluate hot-mix asphalt (HMA) rutting potential in mix design and quality control-quality assurance (QC-QA) applications, because the APA testing and its data analyses are relatively simple, rapid, and easy. However, as demonstrated in many studies and also experienced by the state of Nebraska, APA testing is in question due to its high testing variability and a lack of sufficient correlation with actual filed performance. The primary objective of this research was to find critical materials and/or mixture design factors affecting APA test results so as to eventually improve the current APA testing program in Nebraska. In addition to that, development of models to predict APA rut performance with given properties of HMA mixture ingredients and mixture design characteristics were also attempted. To find variables affecting APA rut results and the extent of these variables, *SP-4* mixture data from Nebraska and HMA mixture data from Kentucky were statistically analyzed using the multiple linear regression method considering six factors (binder PG, aggregate gradation, nominal maximum aggregate size, aggregate angularity, air voids in mixture, and asphalt content in mixture) as probable candidates for significantly affecting APA rut results. For a detailed characterization of gradation effects, three indicators (gradation density, fineness modulus, and restricted zone) were considered, and each of them was used for each statistical analysis. Results from analyses demonstrated that the binder PG was the only variable that always shows significant impact on APA rut results, which is in good agreement with other studies. Predicting models developed through the results of multiple linear regression analysis and the artificial neural network technique presented a relatively low level of model adequacy which can be observed by the coefficients of determination and cross-plots between

predicted APA rut values and the measured APA rut data. More data would be helpful to confirm the findings from this research and also to develop a better prediction model.

Chapter 1 Introduction

The asphalt pavement analyzer (APA) has been widely used in many states as a straightforward method to evaluate hot-mix asphalt (HMA) rutting potential in mix design and quality control and quality assurance (QC/QA) applications. The APA is more advantageous than other testing methods that have been proposed from Strategic Highway Research Program (SHRP) studies, because the APA testing and resulting data analyses are relatively simple, rapid, and easy to perform. However, as demonstrated in many studies including the national study, NCHRP 9-17 (Kandhal and Cooley 2003), the use of APA testing is being re-evaluated due to its high testing variability and a lack of sufficient correlation with actual field performance. Nebraska Department of Roads (NDOR) has employed the APA testing for several years as a supplemental tool to validate and evaluate rutting potential of Superpave mixes paved in Nebraska. NDOR typically found that the APA testing data were not sufficiently reliable to judge rutting characteristics of a Superpave mix because of the high testing variability and poor correlations with actual field performance. In many cases, APA rut depth monitored from a specific Superpave mix was not consistent with rut depths from other mixes within the same mix design criteria. Due to this fact, most state highway agencies have tried to find problems and solutions associated with current APA techniques, so that they can reach the level of confidence needed to utilize the simple APA testing to accept or reject HMA mixtures. Nebraska has also accumulated, but not yet fully investigated, APA testing data of each different mixture that has been paved and is in service. Therefore, there is a pressing need for careful investigations of APA testing and resulting data to better understand why the high testing variability from specimens within the same mix design criteria has been observed.

1.1 Research Objectives

The primary objective of this research was to seek better understanding and potential improvements of the current APA testing program incorporated with Superpave specifications implemented in Nebraska. A comprehensive literature review including careful investigations of APA data available from Nebraska (approximately four-year data) and other states was conducted to find critical factors affecting APA test results and to monitor sensitivity of APA results with mix design variables. This can minimize currently observed high testing variability. In addition, development of models to predict APA rut performance with given properties of HMA mixture ingredients and mixture design characteristics was targeted.

1.2 Research Scope

To meet the research goals, this report was performed in three phases. Phase 1 consisted of literature survey to review significant findings from other studies investigating the variability of APA test results and sensitivity of APA test results to mixture characteristics. Based on findings from the literature review, phase 2 employed statistical approaches to determine which factors of mixture characteristics affect the APA rut performance with a high level of sensitivity. APA test data from Nebraska and another state, Kentucky, were obtained and used to conduct the statistical sensitivity analysis. In phase 3, prediction models were developed using multiple linear regression analysis and artificial neural network technique. Predicted and measured values were compared with both methods.

1.3 Organization of the Report

This report is composed of five chapters. Following this introduction, chapter 2 briefly summarizes findings from several other studies investigating the effects of HMA mixture and material characteristics on APA rut test results. In chapter 3, detailed descriptions of APA data acquisition and implementation to conduct the statistical analysis, which has been employed for this study, are presented. Chapter 4 presents statistical analysis results of significant factors that affect APA rut results and the prediction models developed through the multiple linear regression analysis and the artificial neural network technique. Finally, chapter 5 provides a summary of study findings, recommended future research, and implementation plans for the Nebraska Department of Roads.

Chapter 2 Literature Review

To seek a better understanding of the current APA testing program which is being questioned due to its high testing variability, literature reviews have been conducted by primarily targeting investigation of the effects of HMA mixture and material characteristics on rut test results from APA, as well as other traditional wheel-loading testers. This chapter briefly introduces some significant findings from several studies where the relationship between HMA rutting performance and related materials and/or mixture design factors has been investigated.

Kandhal and Cooley (2003) performed a National Cooperative Highway Research Program (NCHRP) study, NCHRP 9-17 project, with two objectives: to identify test conditions within the APA that produced results most related to field rutting performance and to validate the proposed APA test method as an appropriate QC/QA rut predicting tool. They selected ten HMA mixes of known rutting performance to determine the combination of testing conditions for the APA that best predicted field rutting. These ten mixes were selected from three full-scale pavement research projects: WesTrack (Nevada), the Minnesota Road Research (MnRoad), and the FHWA Accelerated Loading Facility (ALF) at Turner-Fairbank Highway Research Center (Virginia). Numerous APA tests were performed under different testing conditions by varying specimen geometry, APA loading hose diameter, test temperature, and air void content in mixtures. APA test results were then analyzed and correlated to field performance data. Statistical analysis results demonstrated that five percent air voids were more closely related to field rutting performance than seven percent air voids, and specimens tested at a temperature corresponding to the high temperature of the standard performance grade (PG) better predicted field rutting performance than at 6 ^{o}C higher than the high temperature of the standard PG. Loading hose diameter and the specimen geometry (cylinder vs. beam) did not show any significant dependency on APA rutting performance. Kandhal and Cooley (2003) also tried to validate the proposed APA test method, and they concluded that laboratory rut depths measured by the APA generally showed good correlations on an individual project basis. However, APA results clearly depended on geographic locations and traffic level.

Related to this research, some useful findings from the NCHRP study were gathered and are presented in table 2.1 and figure 2.1. Table 2.1 shows WesTrack results indicating that the high value of air voids (7%) was generally more susceptible to rutting than low air voids (4%). It is obvious that the testing temperature affects APA rut performance. Specimens were better rut-resistant at the lower temperature (64 ^{o}C) than at the temperature corresponding to the 6 ^{o}C higher (i.e., 70 ^{o}C) of the high temperature of the standard PG. Figure 2.1 presents the effect of nominal maximum aggregate size (NMAS) on rut depths. The figure demonstrates that the larger NMAS mixtures (such as the one with 37.5 mm) had a lower laboratory rut depth than the mixtures designed with a smaller NMAS (19 mm). Therefore, the use of larger NMAS can reduce rut susceptibility.

Table 2.1 Average rut depths for WesTrack sections (NCHRP 508 2003)

Air Voids	Test Temp, °C	Hose Diameter	Specimen Type	Rut Depth @ 10,000 cycles, mm		
				Section 15	Section 19	Section 24
7%	64 (PG)	Standard	Cylinder	10.13	6.12	8.17
			Beam	8.52	9.14	8.70
		Larger	Cylinder	8.30	4.19	4.05
			Beam	6.46	7.99	6.39
	70 (PG+6)	Standard	Cylinder	12.78	8.61	10.56
			Beam	12.65	11.14	11.69
		Larger	Cylinder	12.01	7.22	4.57
			Beam	9.33	6.21	8.40
4%*	64 (PG)	Standard	Cylinder	6.64	7.76	8.27
			Beam	6.08	10.84	13.33
		Larger	Cylinder	6.36	5.35	7.54
			Beam	5.10	5.76	7.08
	70 (PG+6)	Standard	Cylinder	8.81	8.80	9.29
			Beam	11.57	13.07	14.88
		Larger	Cylinder	7.83	6.10	5.47
			Beam	5.64	6.87	5.24

*Beam samples compacted to 5.0 ± 0.5% air voids.

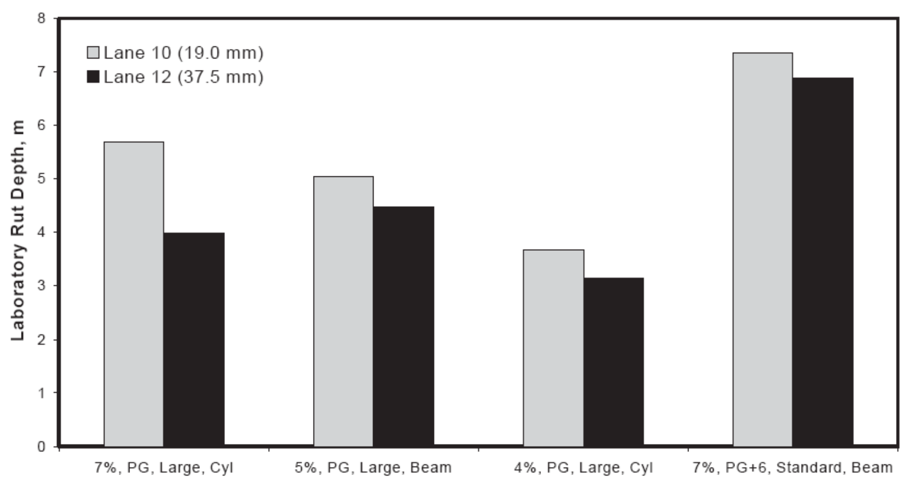
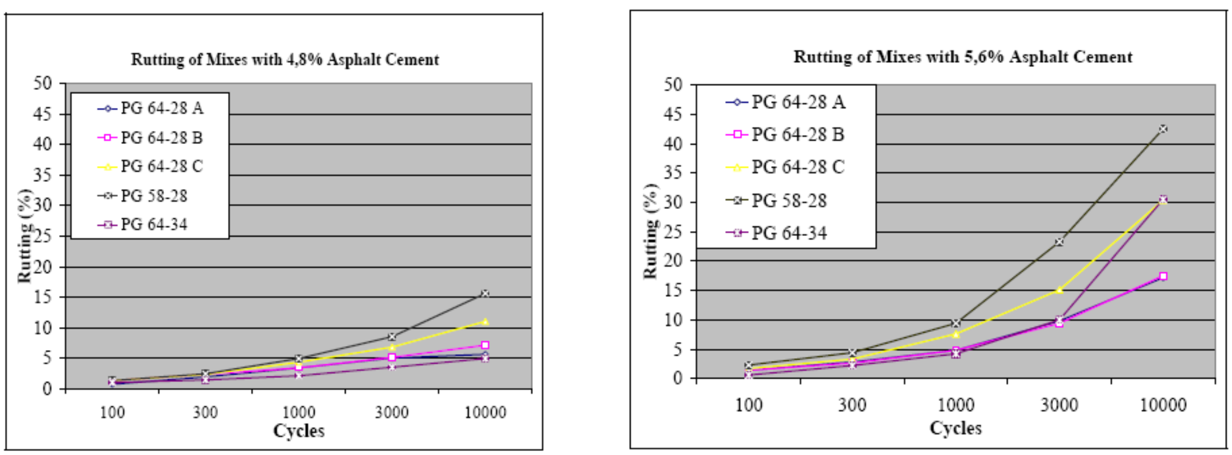


Figure 2.1 Effect of nominal maximum aggregate size on rutting (NCHRP 508 2003)

Uzarowski et al. (2004) tested accelerated performance of Canadian asphalt mixes using three different wheel rut testers: French laboratory tester, Hamburg wheel tester, and APA. They controlled asphalt content, binder PG, and wheel cycles on each testing. Despite the different methods, the test results revealed similar patterns. First of all, every test showed good correlation between field observations and testing results. Also, the high level of asphalt content (5.6%) was more susceptible to rutting performance than the low asphalt contents (4.8%). Figure 2.2 shows how the modified binder resisted rutting better than neat binder using the French laboratory wheel-loading tester.



(a) 4.8% asphalt content (b) 5.6% asphalt content

Figure 2.2 Relationship between binder PG and rutting at different asphalt content (Uzarowski et al. 2004)

Cross and Purcell (2001) investigated effects of fine aggregate angularity (FAA) on voids in mineral aggregates (VMA) and rutting in Kansas HMA mixtures. For the evaluation, two gradations (coarse and fine) of 100% crushed limestone were used. To change the FAA of the mix, natural sand and chat were used instead of crushed limestone. Among the three materials, limestone and natural sand showed an increasing trend of FAA with increasing mixture VMA. However, chat did not show the same increasing pattern. Increasing the FAA resulted in less rutting in the limestone and natural sand mixtures, but there was no clear relationship between the FAA of the chat mixes and rutting.

Lee et al. (1999) found a relationship between FAA and asphalt mixture performance. They used PURWheel designed by Purdue University to evaluate HMA rut potential at different FAA values. They indicated that the specimen with the high value of FAA showed less susceptibility than the low value in rutting. The mixture with a FAA value of 45 performed better than the others.

Stiady et al. (2001) studied effects of aggregate properties (NMAS), coarse aggregate type (granite and limestone), fine aggregate angularity, and gradation types using PURWheel. As shown in figure 2.3, there was a significant relationship between FAA and permanent deformation, but a FAA value too high (greater than 45) did not show better performance in the mixtures. They also concluded that a NMAS of 9.5 mm and 19 mm had no difference statistically.

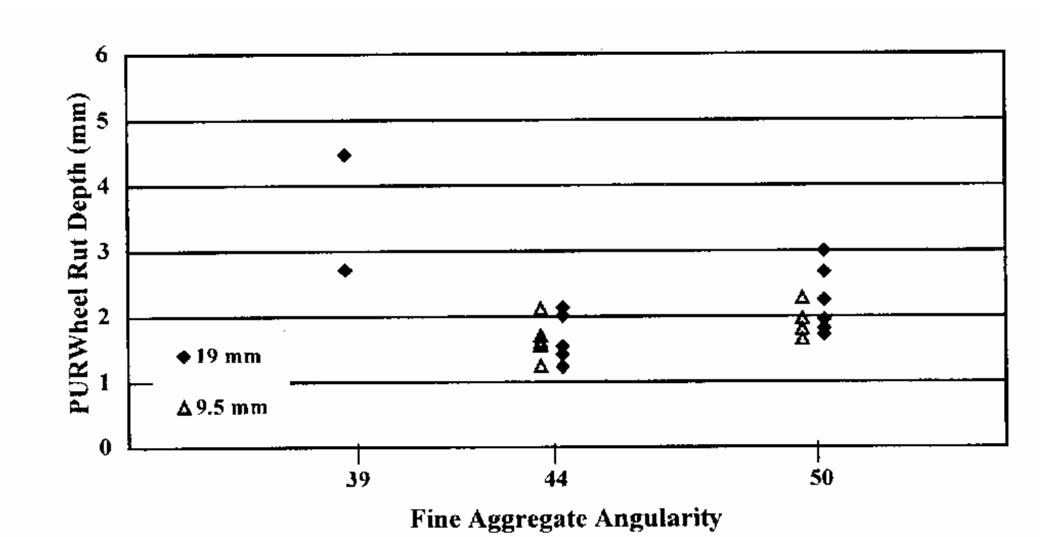


Figure 2.3 PURWheel rut depths at different FAA values (Stiady et al. 2001)

Kandhal and Mallick (2001) evaluated APA testing for HMA mixture design by using the test data and statistical method. The focus of their study was to find the effect of mix gradations on HMA rutting performance. They used three aggregates (granite, limestone, and gravel) and three types of aggregate gradations: above-restricted zone (RZ), through-RZ, and below-RZ. Permanent deformation was significantly affected by the gradation and the type of the aggregate as shown in table 2.1. Analysis of variance (ANOVA) indicates the significant effect of aggregate type, gradation, and course type as well as interaction of aggregate and gradation.

Table 2.2 Analysis of variance for rut depths of mixes (Kandhal and Mallick 2001)

Source	DF	Mean Square	F Value	Pr>F
Aggregate	2	51.59	45.64	0.0001
Asphalt	1	34.96	30.94	0.0001
Gradation	2	24.35	21.54	0.0001
Course	1	57.56	50.92	0.0001
Aggregate*Gradation	4	33.79	29.90	0.0001

Tarefder et al. (2003) attempted to identify the most significant factors that have been known to affect rut potential of HMA mixtures using APA testing. They tested three sets (set A, B, and C) with seven factors (binder PG, temperature, load, hose pressure, asphalt content, moisture in test specimen, and type of specimen), six factors (gradation, temperature, load, hose pressure, asphalt content, and moisture in test specimen), and five factors (gradation, temperature, load, hose pressure, and moisture in test specimens), and analyzed them using the statistical method. Table 2.3 summarizes the results of ANOVA indicating that the binder's PG, testing temperature, moisture in specimen, and aggregate gradation were commonly observed factors affecting mixtures' APA rutting performance significantly. Wheel load, asphalt content and loading hose pressure were less significant.

Table 2.3 Statistical results (Tarefder et al. 2003)

Test	Factor	$df_x = (n_x - 1)$	SS_x	Variance, $V_x = \frac{SS_x}{df_x}$	$F_x = \frac{V_x}{V_{er}}$ (Statistics)	$F_{Table} = F(1,3)_{0.05}$	SS_x'	% Contribution, P_x
Set A	PG	1	20.9	20.7	102.6	10.1	20.3	58.5
	Specimen	1	7.3	7.3	36.1	10.1	6.7	19.4
	Temperature	1	2.9	2.9	14.1	10.1	2.3	6.5
	Moisture	1	1.9	1.9	9.5	10.1	1.3	3.8
	Error	3	0.6	0.2				
	Total	7	34.6	5.0				88.2
Set B	Temperature	1	7.4	7.4	21.5	10.1	7.1	25.2
	Gradation	1	7.0	7.0	20.4	10.1	6.7	23.8
	AC	1	5.0	5.0	14.4	10.1	4.6	16.4
	Moisture	1	3.8	3.8	11.0	10.1	3.4	12.2
	Error	2	0.7	0.4				
	Total	6	28.1	5.6				77.6
Set C	Temperature	1	28.9	28.8	22.1	18.5	27.6	47.6
	Gradation	1	18.0	18.0	13.8	18.5	16.7	28.8
	Moisture	1	5.5	5.5	4.2	18.5	4.1	7.2
	Error	2	2.6	1.3				
	Total	5	57.9	11.6				83.6

Mohammad et al. (2001) evaluated aggregate contributions to rutting susceptibility of asphalt mixtures. Three types of mixtures were used in the research: SMA, CMHB, and dense-graded wearing course. Three types of aggregate (siliceous limestone, sandstone, and novaculite) were used in SMA, and crushed limestone was used in the other mixtures. PG 70-22M modified binder was used in all the mixtures. SMA made of sandstone aggregate was the best-performing mixture, and the dense-graded and the CMHB generally showed better performance in rutting performance than the other SMA mixtures.

More recently, Shu et al. (2006) investigated the effects of coarse aggregate angularity (CAA) and binder PG grade on rutting performance of HMA mixtures. APA was used to evaluate the rut depth of mixtures. Two types of binder performance grade (64-22 and 76-22) with varying CAA values were investigated. The test results showed that CAA significantly affected rutting performance of HMA mixtures when the binder grade was critical to the environment.

Chapter 3 Research Methodology

To accomplish the research objectives, the statistical method based on the multiple linear regression analysis was selected for this study. NDOR has performed APA testing for four years and accumulated testing data. Using APA data, it was possible to identify materials and/or mixture design factors affecting APA rutting and the extent to which each factor affects APA rutting in HMA through statistical analysis. Among many advantages of the use of statistical approaches, one is easy adaptation of the same approach to other available data. A successfully developed statistical approach for a set of data from the state of Nebraska can be directly applied to data obtained from another state, Kentucky. Another advantage of the statistical approach is that this method requires much less time and costs than other methods such as laboratory testing.

The multiple linear regression analysis selected to identify factors significantly affecting APA rut results can also provide prediction models relating the APA rut depth to materials and/or mixture design factors considered. The APA rut values predicted by the multiple linear regression technique were then compared to values from another technique, the artificial neural network, which has been widely employed in developing prediction models with many variables. This chapter briefly explains the statistical method (multiple linear regression analysis) and the artificial neural network technique that were employed for this research. After the brief introduction to the multiple linear regression analysis and the artificial neural network technique, target APA data selected for this study are presented. The type of Superpave mix and materials and/or mixture design factors to be considered for the analyses are determined and also presented in this chapter.

3.1 Multiple Linear Regression Analysis

Linear regression is defined as the method which finds the statistical model that defines the experimental data (Draper and Smith 1998). The model consists of one independent variable and one dependent variable in simple linear regression. Simple linear regression analysis is a method to find a relationship between these two variables. The multiple linear regression analysis is used when experimental data has several independent variables. The general purpose of multiple linear regression analysis is to investigate the linear relationship between several independent (or predictor) variables and a dependent (or response) variable.

The regression result is presented through the ANOVA table. ANOVA is a tool expressing test results based on the F-ratio which is defined as a test of standard deviation of populations. A typical format and entities in an ANOVA table from the multiple linear regression analysis with n number of data and p number of independent variables in the model is presented in table 3.1.

Table 3.1 ANOVA table from the multiple linear regression analysis

Source	Degree of Freedom (DF)	Sum of Squares (SS)	Mean Square (MS)	F-Ratio
Regression Model	p	SSR^1	$MSR^4 = SSR/p$	MSR/MSE
Error	$n-p-1$	SSE^2	$MSE^5 = SSE/(n-p-1)$	
Total	$n-1$	$SSTO^3$		

Note:
SSR^1 = regression sum of squares
SSE^2 = error sum of squares
$SSTO^3$ = total sum of squares
MSR^4 = mean square due to regression, and
MSE^5 = mean square due to error.

The components in the regression sum-of-squares (*SSR*, *SSE*, and *SSTO*) in the third column can be defined as follows:

$$SSR = \sum_{i=1}^{n} \left(\hat{Y}_i - \overline{Y}\right)^2 \tag{3.1}$$

$$SSE = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 = \sum_{i=1}^{n} e_i^2 \tag{3.2}$$

$$SSTO = \sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)^2 \tag{3.3}$$

where,

Y_i = observed values,

$\hat{Y}_i$ = fitted values,

$\overline{Y}$ = mean of fitted values, and

e_i = residuals.

The mean square due to error (MSE) and the mean square due to regression (MSR) are given in the fourth column of the ANOVA table. The *F*-ratio in the fifth column is simply calculated by dividing MSR by MSE, and provides a statistic for testing whether or not the independent variables explain some of the variation in the response variable (dependent variable). The significance of the test results is justified by comparing the *F*-ratio to a significance level (α), which is typically equal to either 0.01, 0.05, or 0.10. The significance level is decided by the statistician who performs the analysis. When the calculated *F*-ratio is equal or greater than $F(\alpha; p, n\text{-}p\text{-}1)$, it means that there is at least one independent variable that explains the variation in the dependent variable. Alternatively, many statistical software programs calculate the *p*-value = Probability $[F(p, n\text{-}p\text{-}1) \geq F]$ where *F* is the calculated *F*-ratio. If this *p*-value is small (less than the significance level α), one can conclude that with the data there is sufficient evidence to say that at least one independent variable contributes to the variation in the dependent variable. Thus, the resulting model from the multiple linear regression analysis is considered a significant model where a meaningful relationship between a dependent variable and independent variables exists. As an example, an ANOVA table resulting from a real multiple regression analysis was produced and is shown in table 3.2. With α level of 0.05 selected, the ANOVA results indicate that there exists at least one independent variable that contributes to variation in the dependent variable. There is a significant relationship between variables, because the *p*-value (i.e., Pr > F, as presented in the table) is less than the specified significance level (α value), 0.05.

Table 3.2 ANOVA table resulting from a real multiple regression analysis

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	0.00019450	0.00002161	2.39	0.0188
Error	81	0.00073324	0.00000905		
Corrected Total	90	0.00092774			

If the testing analysis is significant, a multiple linear regression model relating variables can be produced. Equation 3.4 (below) is the typical form of the model produced from the multiple linear regression analysis:

$$Y_i = \beta_o + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip} + \varepsilon_i \qquad i = 1,2,\ldots,n \qquad (3.4)$$

where,

Y_i = dependent variable,

β_o = intercept,

β_p = parameters (coefficients) of independent variables,

X_{ip} = independent variables,

p = total number of independent variables,

n = total number of data, and

ε_i = error.

Table 3.3 presents typical results from the multiple linear regression analysis. The table shows the parameter estimate (β_p) of each variable (X_{ip}) and its level of significance based on the t-ratio values. Similar to the F-ratio, the t-ratio is used to assess the significance of individual regression coefficient (β_p) multiplied to each variable in the model. In the case of t-tests, statistical software calculates the p-value = 2*Probability [$t(n\text{-}p\text{-}1) \geq |t|$] where t is the computed value of the t-statistic, with the significance level, α. If the p-value is less than or equal to α, one can conclude that the corresponding independent variable (X_{ip}) has a significant impact on the response. These t-tests are also called partial t-tests, since they assess the partial (or additional) significance of the variables X_{ip}, over and above the impact of all other variables in the model. The sign of each parameter estimate in table 3.3 indicates the trend of relationship between Y_i and X_{ip}. A positive sign in a parameter estimate infers proportionality between the dependent variable and the independent variable corresponding to the parameter estimate. The parameters shown in table 3.3 identify the prediction model (eq. 3.4).

Table 3.3 Parameter estimate from the multiple linear regression analysis

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	0.05868	0.02352	2.49	0.0146
NMA	NMA	1	-0.00312	0.00133	-2.34	0.0216
PG	PG	1	-0.00191	0.00059569	-3.20	0.0020
airvoid	airvoid	1	-0.00090993	0.00060056	-1.52	0.1336
Binder	Binder	1	-0.00071804	0.00103	-0.70	0.4867
CAA1	CAA1	1	0.00004542	0.00014198	0.32	0.7499
CAA2	CAA2	1	-0.00002329	0.00010893	-0.21	0.8312
FAA	FAA	1	-0.00095162	0.00056299	-1.69	0.0948
Coarse	Coarse	1	0.00025703	0.00012084	2.13	0.0365
Fine	Fine	1	-0.00021523	0.00010560	-2.04	0.0448

Regression analysis typically provides a measure of the strength of the relationship between dependent and independent variables. One measure that has been widely used to quantify the strength of the relationship is called the coefficient of correlation, *r*-value. The r-value lies between -1 and +1, therefore R^2-value (called coefficient of determination) is more frequently used to give the proportion of the total variability in the dependent variable that is accounted for by the independent variables. The r^2 can be calculated by the following expression:

$$R^2 = 1 - \frac{SSE}{SSTO} \tag{3.5}$$

If the R^2 is equal to 1.0, all measured values are predicted by the regression model. In other words, the developed model explains the relationship among variables perfectly. On the other hand, R^2 value of zero indicates that no measured data agrees with the prediction model.

Since the R^2 value usually can be made larger by including a large number of predictor variables, it is sometimes suggested that a modified measure be used that adjusts for the number of independent variables in the model. The adjusted coefficient of determination, denoted by *adj. R,*^{2} modifies R^2 value by dividing each sum of squares by its associated degrees of freedom:

$$adj.R^2 = 1 - \left(\frac{n-1}{n-p-1}\right)\frac{SSE}{SSTO} \tag{3.6}$$

As the number of independent variables increases, the value of R^2 also increases. However, the *adj.* R^2 may actually become smaller when another independent variable is introduced into the model, because any decrease in *SSE* may be more than that which is offset by the loss of a degree of freedom in the denominator (*n-p*) in equation 3.6. Therefore, the *adj.* R^2 has been known as a better indicator than the R^2 value to measure the strength of the relationship between variables. As an example, the values of R^2 and *adj.* R^2 from the multiple linear regression analysis are presented in table 3.4.

Table 3.4 R^2 and *adj.* R^2 values from the multiple linear regression analysis

Root MSE	0.00301	R-Square	0.2096
Dependent Mean	0.00255	Adj R-Sq	0.1218
Coeff Var	118.12363		

3.2 Artificial Neural Network Approach

The artificial neural network is a kind of statistical, mathematical, or computational methodology. It was developed from inspiration of biological neurons in the human brain.

McCulloch and Pits (1943) first introduced the concept of artificial neurons. The neural network concept gained popularity after the development of inexpensive computer emulation of neural networks. Currently, various fields such as finance, medicine, environmental science, and transportation engineering often use the artificial neural network. The advantages of the artificial neural network approach are that it detects trends from complicated data and it can also do prediction and forecasting. Figure 3.1 shows the basic structure of the artificial neural network approach. The components of the neural network are input variables, one or more output variables, and one or more hidden layers relating input and output variables through networking. The input variables are transformed by a special function such as a logistic or sigmoidal function to account for nonlinearity in the model.

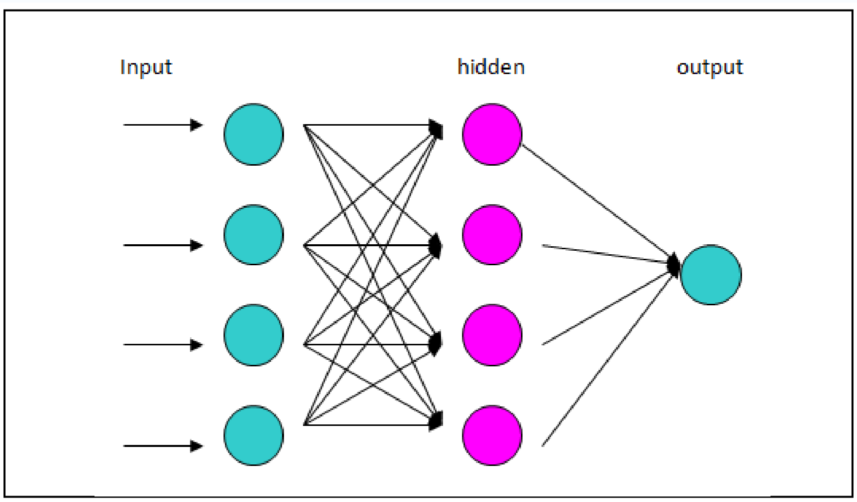


Figure 3.1 Basic structure of the artificial neural network

Among several artificial neural network algorithms, the back-propagation algorithm was adopted in this study because it has been widely used for prediction. The back-propagation algorithm was first introduced by Rumelhart et al. (1986). The back-propagation paradigm usually uses a sigmoidal function for transformation from linearity to nonlinearity.

3.3 APA Data for Analyses

To accomplish the research objectives, the APA data of *SP-4* mix type were chosen for two reasons: 1) the number of *SP-4* data was the highest in the NDOR APA database (total 91 *SP-4* mix APA data), and 2) *SP-4* mix type was one of the primary mixes frequently paved in Nebraska. With the *SP-4* mix APA database, materials and/or mixture variables to be considered as candidates affecting APA rut results needed to be determined. Based on the literature review, six factors (performance grade of binder, aggregate gradation, nominal maximum aggregate size, aggregate angularity, air voids, and asphalt content) were selected, as illustrated in figure 3.2. PG of binder represents the binder mechanical property which clearly affects mixture rut potential. Size and shape factors of aggregates such as the gradation, angularity, and the NMAS were included in the analysis. For the mixture side, two variables (air voids and asphalt content) were selected as primary factors because they are crucial indicators identifying mixture volumetric characteristics and were also expected to affect APA rut depth.

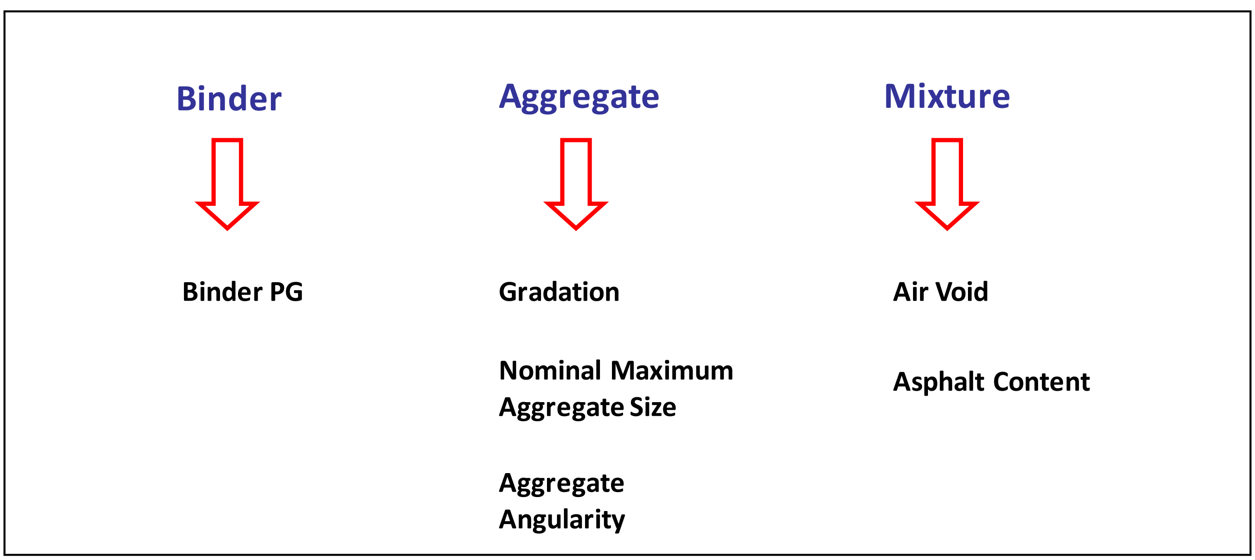


Figure 3.2 Factors selected for analyses

From the APA test database of the *SP-4* mix, it was observed that two nominal maximum aggregate sizes (0.375 inch and 0.5 inch) and three binder performance grades (PG 64-22, 64-28, and 70-28) were used. For a more detailed analysis, aggregate angularity factor was categorized into three variables: coarse aggregate angularity value with one or more fractured faces (denoted by CAA1), coarse aggregate angularity value with two or more fractured faces (denoted by CAA2), and FAA. In the case of aggregate gradation, the gradation factor needed to be quantified in numbers to be implemented in the statistical analyses (multiple linear regression analysis and artificial neural network for this study). In an attempt to quantify the characteristics of gradation, three alternatives were attempted. First, the gradation curve of each mixture was plotted on the 0.45 power chart as shown in figure 3.3, and then the gradation curve was divided into a coarse aggregate part and a fine aggregate part based on sieve No. 4 (4.75 mm mesh size, which corresponds to 2.016 on the 0.45 power chart). The gradation curve was compared to the maximum density line, which is represented by a straight line on the 0.45 power chart as illustrated in figure 3.3. Then, the areas formed between the gradation curve and the maximum density line were calculated to quantify density characteristics (coarse aggregate density signified by CAD and fine aggregate density signified by FAD) of the gradation. In other words, when the gradation curve is closer to the maximum density line, which is an indication of denser mix, the calculated area becomes smaller.

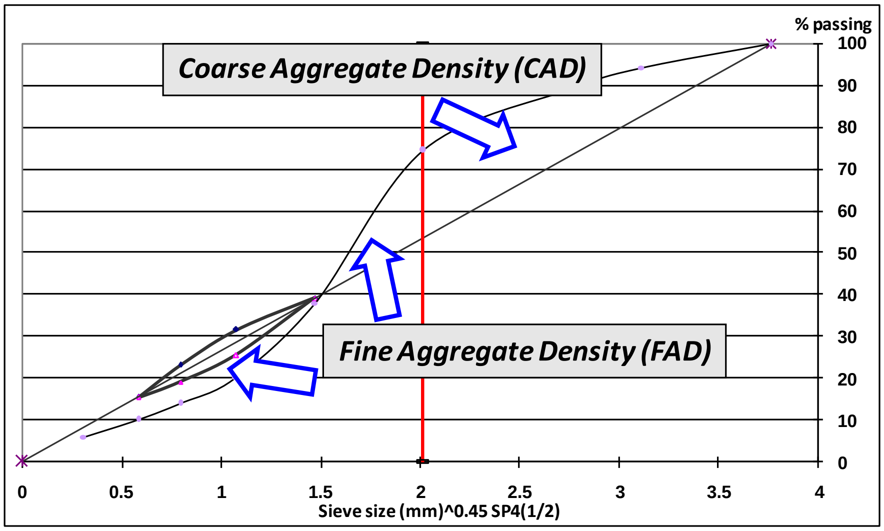


Figure 3.3 A gradation curve with its density characteristics on the 0.45 power chart

The second alternative selected to represent the gradation characteristics was the fineness modulus (signified by FM). The fineness modulus is defined as an empirical factor obtained by adding the total percentages of a sample of the aggregate retained on each of a specified series of sieves and dividing that sum by 100 (Mamlouk and Zaniewski 2006). As the value of the fineness modulus increases, the amount of coarse aggregate increases: in other words, a greater fineness modulus of an aggregate blend means a coarser mix. Therefore, the value of fineness modulus can potentially be a good indicator that represents gradation characteristics of the mix.

The third alternative for representing the gradation characteristics was the use of the RZ. The RZ forms a band residing along the maximum density gradation between an intermediate sieve and the 0.3 mm sieve, as shown in figure 3.3. There is a common belief in the asphalt community that a humped gradation indicates an over-sanded mixture, which often results in compaction problems during construction and reduced resistance to rutting. However, the concept of the restricted zone related to mixture rutting potential has been almost discarded from the Superpave specification today as many studies (Watson et al. 1997; Hand and Epps 2001; Kandhal and Mallick 2001; Kandhal and Cooley 2002; Sebaaly et al. 2004) demonstrated no clear relationship between the restricted zone and mixture rutting potential. Nevertheless, the effect of the restricted zone as an indicator representing gradation characteristics was considered in this study.

With all independent variables selected, a data sheet for statistical analyses was developed and is shown in figure 3.4 for illustrative purposes. The data sheet, including all samples (total 91 samples) from Nebraska can be seen in appendix A. Similar to the Nebraska data sheet, data sheets for the state of Kentucky were also generated and are attached in appendix B. Figure 3.4 presents specific values of independent variables of each mixture and its APA rut result (shown in the last column in the figure) as the dependent variable. In the case of the variable, NMAS, the numbers 1 and 2 were used to represent 0.375 inch and 0.5 inch NMAS, respectively for the purpose of statistical analyses. Similarly, the numbers 1, 2, and 3 were assigned to binder PG (1 for PG 64-22, 2 for PG 64-28, and 3 for PG 70-28) and to the RZ (1 for above-RZ, 2 for through-RZ, and 3 for below-RZ), respectively. For other variables, real experimental values were used. Instead of using APA rut depth, the rut ratio was calculated by dividing total rut depth by the number of loading cycles. This is because the APA test automatically stopped when the wheel loading reached 8,000 cycles before 12 mm rut depth or when the total rut depth exceeded 12 mm. To provide an identical measure of mixture rut potential for both cases, the rut ratio was calculated and used. One more thing to be noted from the figure is that only one alternative among three of the gradation factors was used for analyses. For example, if the concept of gradation densities (CAD and FAD) was used, the other two gradation-related variables (FM and RZ) were excluded in the analyses.

No.	NMAS	PG	Airvoid	Binder	CAA1	CAA2	FAA	CAD	FAD	FM	RZ	Rutratio
1	2	2	4.2	5.9	94	90	45.7	20.273	11.9	4.28	2	0.0002
2	2	1	3.9	5.14	89	86	45.4	15.941	8.533	4.477	2	0.0002
3	2	2	4.8	5.7	99	94	50.1	24.814	13.903	4.114	1	0.0003
4	1	2	4	6.83	95	90	46.9	10.02	10.707	3.813	2	0.0003
						⋮						
89	2	2	3.6	5.41	88	80	45.1	24.382	15.456	4.123	2	0.012
90	2	1	5	5.56	91	82	45.5	21.819	12.669	4.441	2	0.013
91	2	2	4	5.11	95	89	45.2	19.451	9.234	4.498	2	0.0181

Figure 3.4 Data sheet developed for statistical analyses

Chapter 4 Analysis Results and Discussion

To find variables affecting APA rut results and the extent of these variables, *SP-4* APA data from NDOR and data from another state (Kentucky) were analyzed through the Statistical Analysis Software (SAS) using multiple linear regression analysis. In addition to the analysis of significance, prediction models were developed using the results of multiple linear regression analysis and the artificial neural network technique. The predicted rutting values from the multiple linear regression analysis were compared to the values from the artificial neural network.

4.1 Analysis of Significance

Table 4.1 presents statistical analysis results. As can be seen, three cases were considered using the Nebraska data due to the use of different gradation-related alternatives (gradation density signified by CAD and FAD, fineness modulus [FM], and restricted zone [RZ]). For the data from Kentucky, the gradation effects were investigated by using the gradation density indicators (CAD and FAD) only. Table 4.1 shows the overall significance of the test results justified by the *F*-ratio and *p*-value of each case. By comparing the *p*-value (i.e., Pr > *F*) to a given α value (0.01, 0.05, or 0.10), one can decide if the resulting model from the multiple linear regression analysis is considered as a model where there is at least one independent variable that affects the variation of the dependent variable (i.e., APA rut results). Table 4.1 also presents the significance of each individual regression coefficient and its parameter estimate by providing the *t*–test results, which are useful for assessing the significance of each independent variable in the model. As mentioned previously, if the *p*-value (i.e., Pr > |*t*|) is less than the significance level (α value) specified, then the independent variable considered is a significant factor affecting the APA rut depth. To maintain consistency of the analysis, the same value of α level (0.05) was applied to all cases.

Table 4.1 Statistical analysis results

	Nebraska Data						Kentucky Data	
	Case 1		Case 2		Case 3		Case 4	
	F-ratio = 2.39 **Pr > *F*** = 0.0188		***F***-ratio = 2.05 **Pr > *F*** = 0.0506		***F***-ratio = 2.02 **Pr > *F*** = 0.0539		***F***-ratio = 6.60 **Pr > *F*** = 0.0037	
Variables	Parameter Estimate	**Pr > |*t*|**	Parameter Estimate	**Pr > |*t*|**	Parameter Estimate	**Pr > |*t*|**	Parameter Estimate	**Pr > |*t*|**
Intercept	0.05868	0.0146	0.03953	0.1140	0.05570	0.0290	0.03820	0.0808
NMAS	-0.00312	0.0216	-0.00138	0.1182	-0.00106	0.1752	-0.000053	0.8164
PG	-0.00191	0.0020	-0.00164	0.0069	-0.00165	0.0066	-0.000258	0.0005
% Air	-0.00091	0.1336	-0.000887	0.1522	-0.00096	0.1212	-0.000281	0.6337
% Binder	-0.00072	0.4867	-0.000779	0.4619	-0.00127	0.1985	0.000068	0.7084
CAA1	0.00005	0.7499	0.000070	0.6289	0.000061	0.6748	-0.000372	0.0735
CAA2	-0.00002	0.8312	-0.000078	0.4783	-0.00007	0.5263	***	***
FAA	-0.00095	0.0948	-0.00648	0.2484	-0.00078	0.1820	0.000017	0.7804
CAD	0.000257	0.0365	***	***	***	***	0.000008	0.6337
FAD	-0.000215	0.0448	***	***	***	***	0.000005	0.8225
FM	***	***	0.00146	0.3016	***	***	***	***
RZ	***	***	***	***	-0.00073	0.3469	***	***

For case 1, test statistics infer that there existed a meaningful relationship between the APA rut results and at least one independent variable, since the *p*-value (Pr > *F*) was less than the specified α value (0.05). Among the nine variables (excluding the intercept), four variables (NMAS, PG, CAD, and FAD) were found to be significant at the α value of 0.05. By analyzing the parameter estimate of the four significant variables, the NMAS had the largest effect on rutting because the absolute quantity of parameter estimate of the NMAS was the largest. The negative sign of the parameter estimate indicates that 0.5 inch was less susceptible than 0.375 inch of NMAS in rutting performance. The second largest variable affecting APA rutting was binder performance grade (PG), which was -0.00191 in the analysis. As with the NMAS, the value of the parameter estimate of PG had a negative sense. Among the three types of PG (62.12, 64-28 and 70-28) used in the analysis, 70-28 was the best in rutting performance followed by 64-28 and 64-22, respectively. Gradation also affected APA rutting. When compared with the parameter estimate of NMAS and PG, coefficients to gradation were relatively small values

(CAD of 0.000257 and FAD of -0.000215), suggesting that the gradation affects rutting performance in a less significant way. In the case of CAD, the parameter estimate was in positive sign. The more the value of CAD increased, the more that APA rut depth increased. FAD had a negative value, which implied that the more fine aggregate density decreased, the more the depth of permanent deformation increased.

To investigate the gradation effects in a more detailed way, the same multiple linear regression analysis was repeated with different data sets, in which samples in each data set were categorized by their FAD values. Figure 4.1 presents the analysis results, which clearly demonstrated that as the FAD increased, the *p*-value tended to decrease. It can be inferred that APA rut depth was more likely affected by gradation characteristics such as the FAD. In other words, if the FAD values of mixtures are less than approximately 27 (as shown in fig. 4.1), APA rut results among mixtures were expected to produce the same value statistically when the significance level is 0.05. In fact, there was only one gradation presented in figure 4.2 where the FAD value was greater than 27.

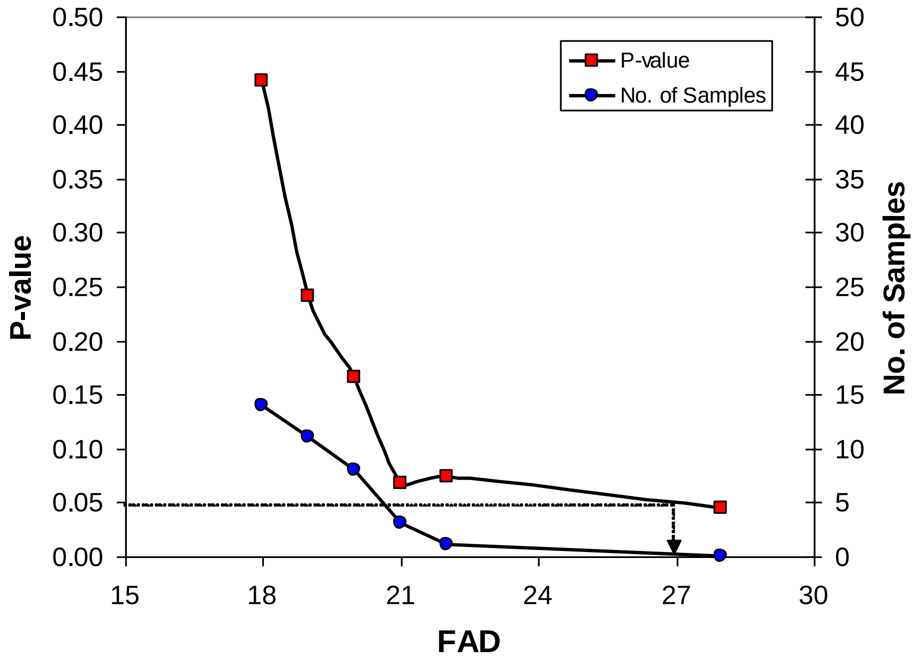


Figure 4.1 Variation of *p*-value with different FAD in gradation

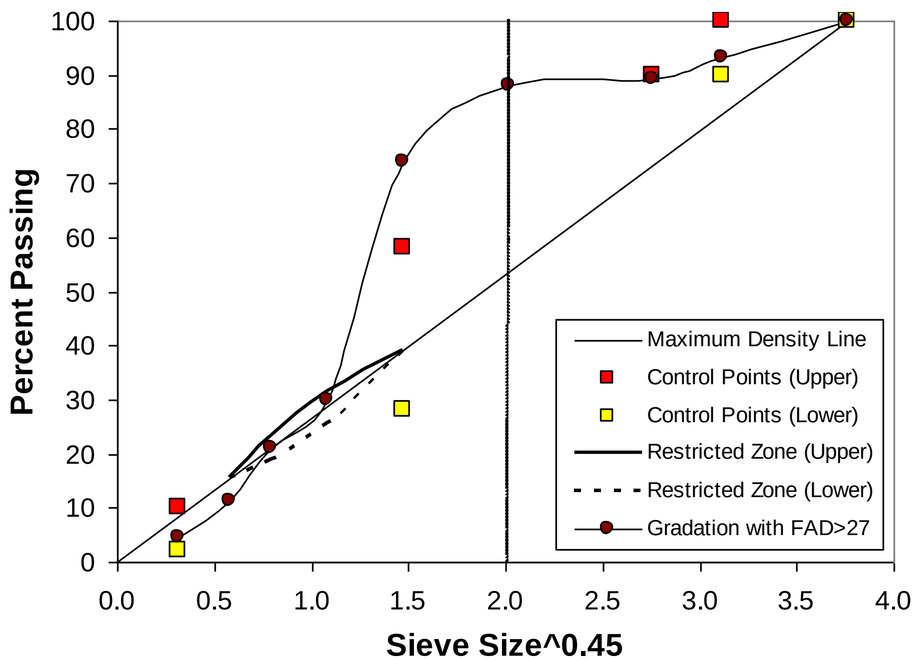


Figure 4.2 A gradation curve with FAD value greater than 27

Similar to the case of FAD, the multiple linear regression analysis was also repeated for different data sets grouped by varying CAD values. As expected and clearly shown in figure 4.3, the *p*-value decreased as the CAD of mixtures increased. From the figure, it can be concluded that mixtures with the CAD value less than approximately 25 are likely producing the same APA rut results at the significance level of 0.05, if other variables of the mixture remain constant. Among the total 91 samples, only four gradations experienced the CAD value greater than 25. The four gradations are plotted in figure 4.4. As demonstrated in figures 4.1 through 4.4, the majority of mixtures were designed with the FAD and CAD less than their critical values; therefore the effect of gradation may be trivial in practice even if the statistical analysis produced the significance of gradation characteristics based on its density.

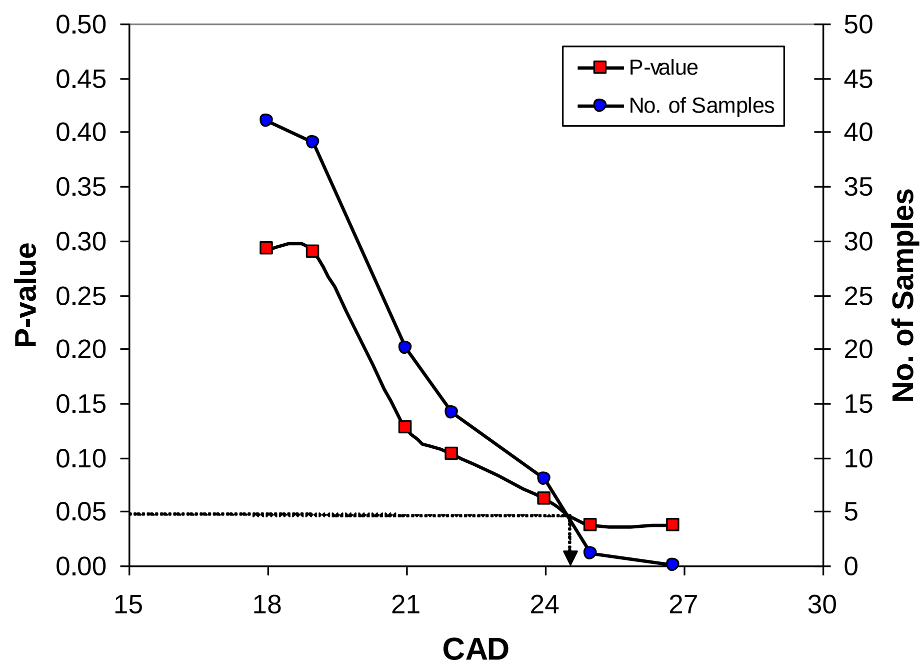


Figure 4.3 Variation of *p*-value with different CAD in gradation

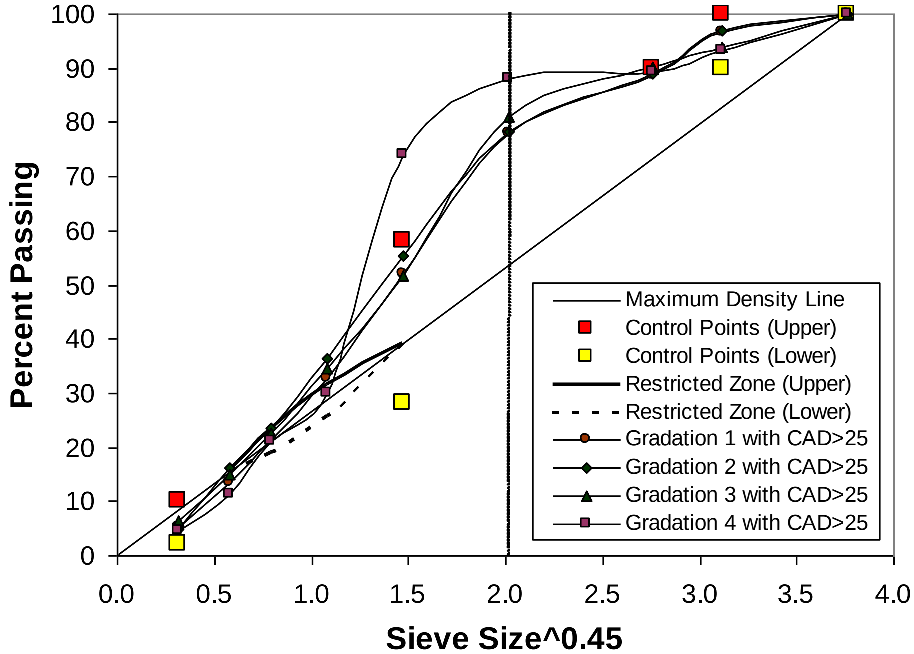


Figure 4.4 Four gradation curves with CAD value greater than 25

In case 2 using the fineness modulus instead of gradation densities, the *p*-value from the *F*-test was 0.0506, which is slightly greater than but very close to the α value (0.05). Among the independent variables included, the fineness modulus did not show significance (Pr > |t|: 0.3016), and binder PG was the only significant variable (Pr > |t|: 0.0069) found from this analysis.

In case 3 based on the concept of restricted zone, the statistical analysis results were similar to the results from case 2 that are presented in table 4.1. The *p*-value (Pr > F) was 0.0539, which is greater than but still close to the specified significance level ($\alpha = 0.05$). Among the variables included in the model, only binder PG showed its significance in the APA rutting. The restricted zone did not show any significant effects on rutting potential, which is in good agreement with general findings from many other studies (Watson et al. 1997; Kandhal and

Mallick, 2001; Hand and Epps, 2001; Kandhal and Cooley, 2002; Sebaaly et al. 2004) that have demonstrated the insignificance of the restricted zone to the HMA rutting potential.

As mentioned earlier, APA data sets were also obtained from Kentucky in an attempt to compare analysis results from the state of Nebraska to the results from a different state. Table 4.1 includes the analysis results (case 4) using the Kentucky data (21 samples total). There were three binder performance grades (PG 64-22, 70-22 and 76-22), and the values of CAA1 and CAA2 were identical with the range of 98-100. Due to the redundancy in CAA, the CAA2 was excluded in the analysis. Test statistics shown in table 4.1 suggest that there exists a meaningful relationship between the APA rut results and at least one independent variable, since the *p*-value (0.0037) is less than the α value (0.05). Among the eight independent variables, binder PG was the only significant variable (Pr > |*t*|: 0.0005), which was the same result found in cases 2 and 3. Binder PG produced a negative effect in performance as observed from all previous cases (case 1 to 3) using Nebraska data.

4.2 Development of Prediction Models

Prediction models were developed using the results of the multiple linear regression analysis and the artificial neural network technique. The predicted APA rutting values from the multiple linear regression models were compared to the values from the artificial neural network.

4.2.1 Prediction Models from the Multiple Linear Regression Analysis

As discussed in chapter 3, a multiple linear regression model (eq. 3.4) relates variables by providing the parameter estimate of each independent variable. The strength of relationship between the dependent variable and independent variables is measured by the R^2 and/or *adj.* R^2 values, which are defined by equations 3.5 and 3.6 in the previous chapter. For the development of multiple regression models to predict APA rut results with given materials and mixture design variables, Nebraska data were used, and a total of nine independent variables (NMAS, binder PG, % air voids, % binder content, CAA1, CAA2, FAA, CAD, and FAD) were considered for the model. Table 4.2 presents resulting model parameters and coefficients of determination (R^2 and *adj.* R^2) of each model.

Table 4.2 Prediction models from the multiple linear regression analysis

	Nebraska Data					
	Model 1		Model 2		Model 3	
	F*-ratio** = 2.39 **Pr > *F = 0.0188 R^2 = 0.21 ***Adj.*** R^2 = 0.12		***F*-ratio** = 4.41 **Pr > *F*** = 0.0001 R^2 = 0.33 ***Adj.*** R^2 = 0.25		***F*-ratio** = 3.26 **Pr > *F*** = 0.0020 R^2 = 0.27 ***Adj.*** R^2 = 0.19	
Variables	Parameter Estimate	**Pr > |*t*|**	Parameter Estimate	**Pr > |*t*|**	Parameter Estimate	**Pr > |*t*|**
Intercept	0.05868	0.0146	22.01186	0.0033	0.04856	0.0097
NMAS	-0.00312	0.0216	-0.96337	0.0218	-0.00290	0.0063
PG	-0.00191	0.0020	-0.77936	< 0.0001	-0.00192	< 0.0001
% Air	-0.00091	0.1336	-0.33716	0.0729	-0.00055	0.2440
% Binder	-0.00072	0.4867	-0.34853	0.2756	-0.00081	0.3102
CAA1	0.00005	0.7499	0.00276	0.9500	0.000001	0.9928
CAA2	-0.00002	0.8312	-0.00512	0.8794	0.000016	0.8499
FAA	-0.00095	0.0948	-0.48916	0.0062	-0.000740	0.0946
CAD	0.000257	0.0365	0.05383	0.1532	0.000178	0.0622
FAD	-0.000215	0.0448	-0.05232	0.1127	-0.000153	0.0673

The first model (model 1) was developed using all 91 data. As shown in table 4.2, the value of the R^2 and the *adj.* R^2 was 0.21 and 0.12, respectively. Several attempts were made to look for more appropriate models that produce a higher correlation between variables, such as data transformation and the diagnostics of outliers. Model 2 was developed through the data transformation by taking a natural log function (*LN* function) on the response variable (*Y*). Simple transformations of the response variable (dependent variable), the predictor variables (independent variables), or of both, are often used to make the regression model more appropriate to the data. In investigating several different transformation functions, such as 1/*Y*, log Y, $\sqrt{Y}$, and other powers of Y, the following transformed model was selected because the model produced the best performance.

$$LN(Y_i) = \beta_o + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip} + \varepsilon_i \qquad i = 1,2,\ldots,n \tag{4.1}$$

As a result of the data transformation, the adequacy of the model improved with higher values of the R^2 (0.33) and the *adj.* R^2 (0.25). Another method which typically improves the adequacy of the fitted model is checking the presence of outliers and removing them from the data. Outliers are extreme observations that are numerically distant from the rest of the data. When there are outliers in a data set, a statistical analysis will return values that do not represent the overall data. Among various diagnostics for outliers, the Cook's D technique was employed in this study. The Cook's D technique measures precision of estimation and detects influential points that pull regression towards their direction. Through this analysis, two observations (sample nos. 89 and 91) were found as influential points. After two observations were removed from the data set, the multiple linear regression analysis was performed and resulted in R^2 (0.27) and the *adj.* R^2 (0.19). Figures 4.5 to 4.7 present cross-plots between measured APA rut results and the predicted values from each model. As the coefficients of determination increased, cross-plots were closer to the line of equality.

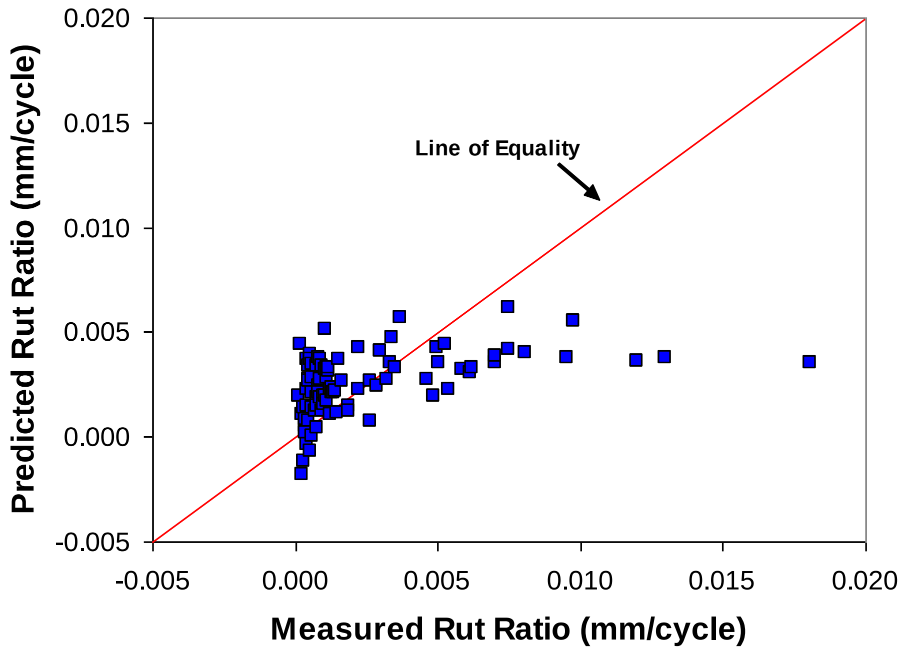


Figure 4.5 Cross-plots between measured and predicted using model 1

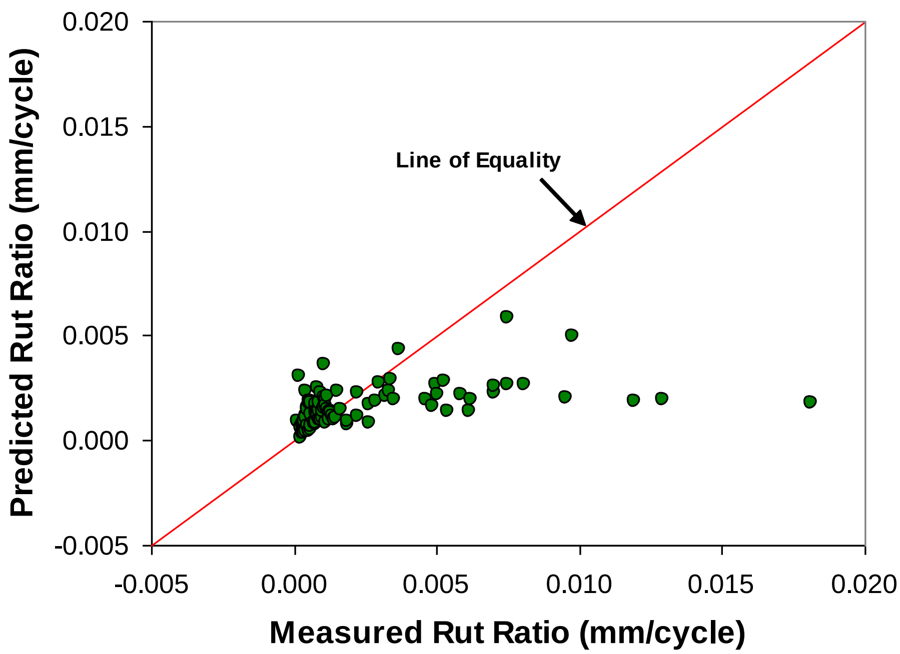


Figure 4.6 Cross-plots between measured and predicted using model 2

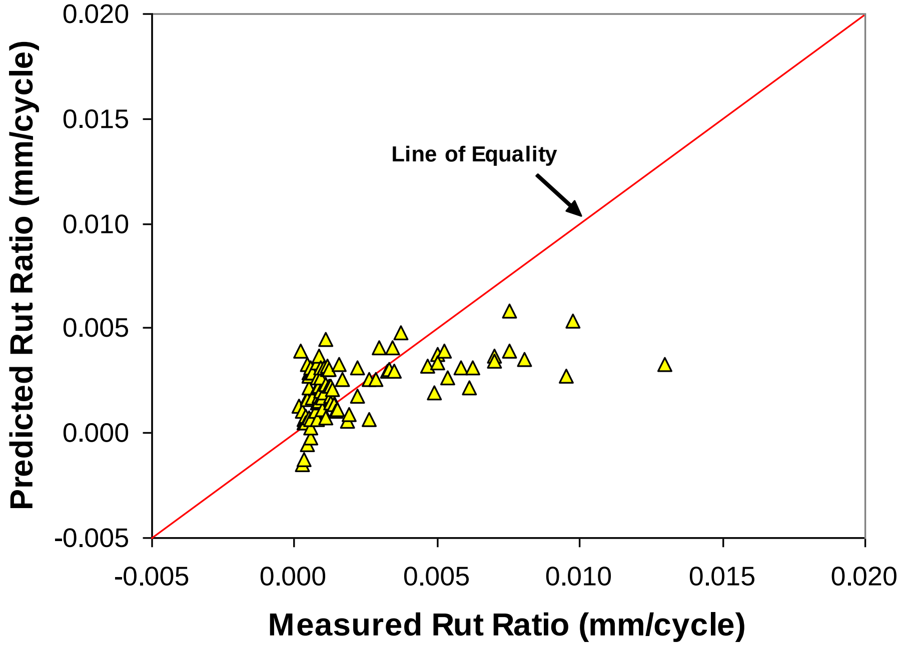


Figure 4.7 Cross-plots between measured and predicted using model 3

4.2.2 Prediction Models from the Artificial Neural Network

As illustrated in figure 4.8, the artificial neural network consists of three steps (Lacroix et al. 2008). The first step is the training stage, the second step is the validation stage, and the last step is the testing stage. The first and second stages are the network developing steps, and the last stage is to predict the response variable based on the network developed.

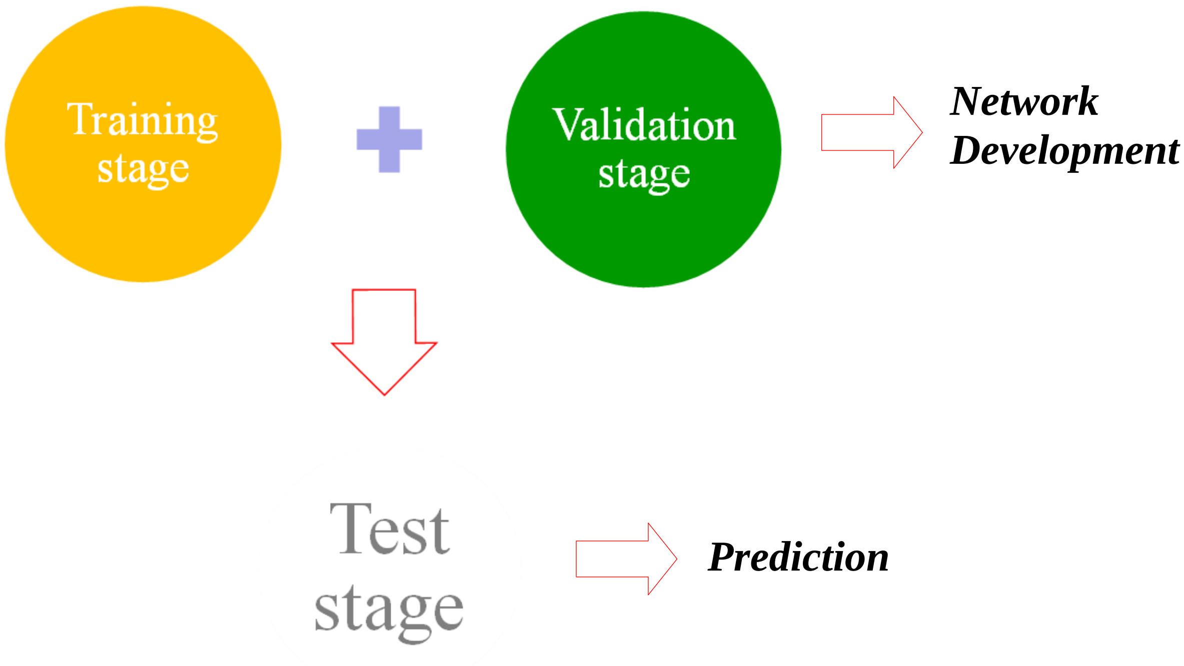


Figure 4.8 Process of the artificial neural network modeling

For the artificial neural network modeling, the same set of data (91 samples of the *SP-4* mix) was used to compare predictions from the artificial neural network modeling to the predictions obtained from the multiple regression modeling (model 1 in table 4.2). With the 91 samples, 56 data were used for training, 15 data were used for validation, and 20 data were used for testing. Cross-plots relating the measured APA rut results to the predictions from each stage are shown in figures 4.9 to 4.11. Figure 4.12 was also developed to compare a predicting power between two methods (the artificial neural network modeling vs. the multiple linear regression analysis) by plotting both predictions to the measured APA data on the same graph. Even if there is no equivalent indicator that can be used to quantitatively estimate the predicting power between two modeling approaches, it can be inferred that there is no large difference between two methods. Neither method provided a high level of model accuracy, which might be due to a lack of data involved.

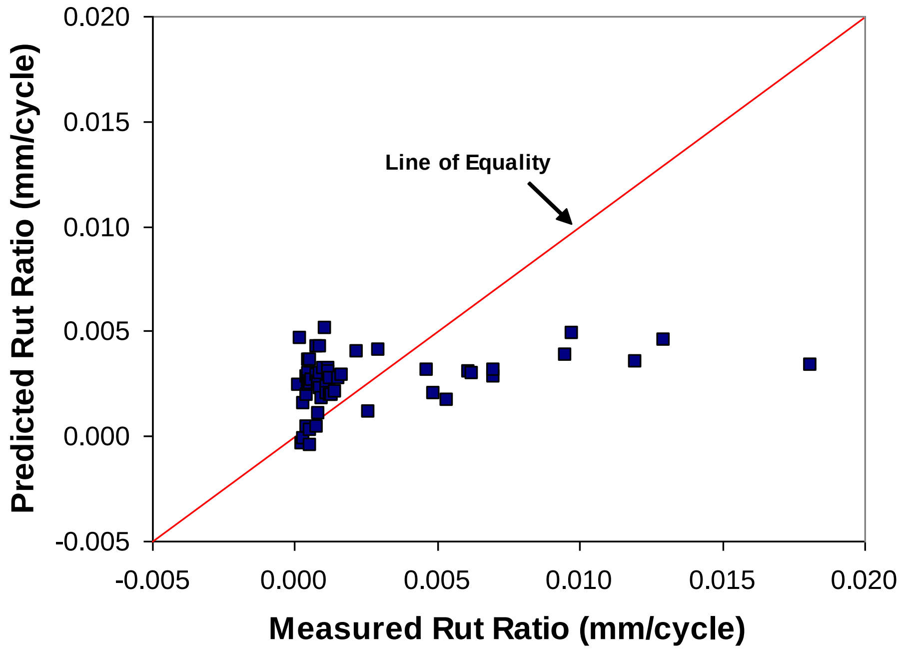


Figure 4.9 Cross-plots between measured and predicted (training stage)

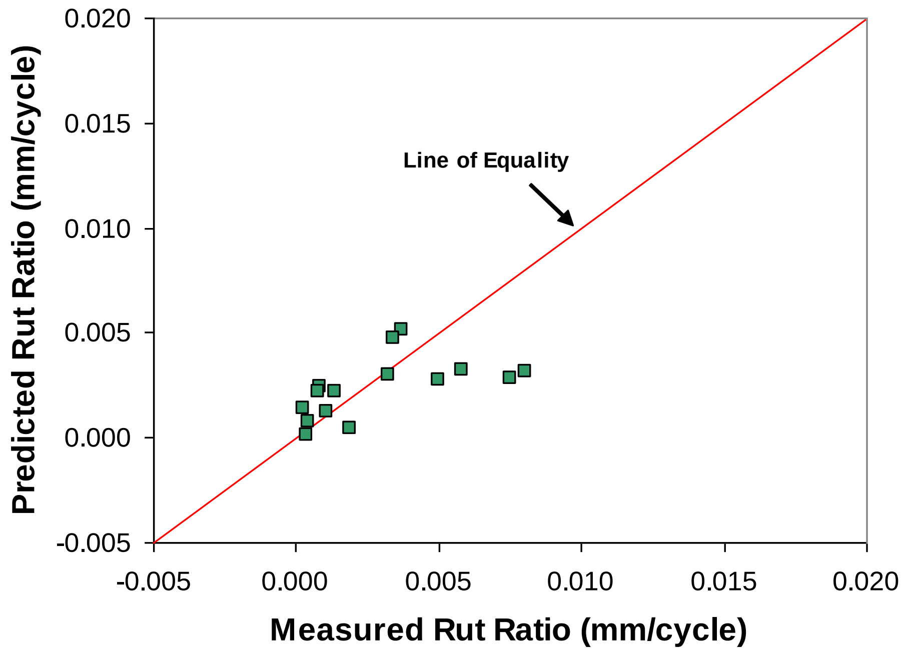


Figure 4.10 Cross-plots between measured and predicted (validation stage)

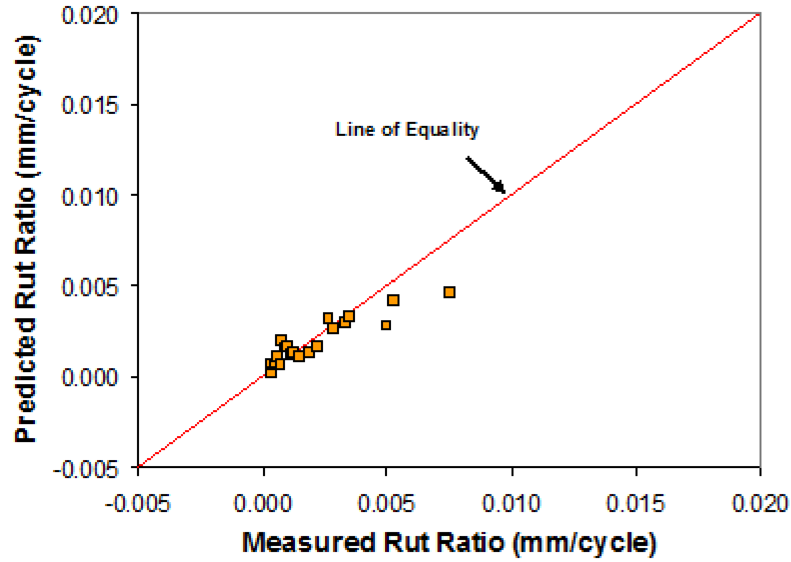


Figure 4.11 Cross-plots between measured and predicted (testing stage)

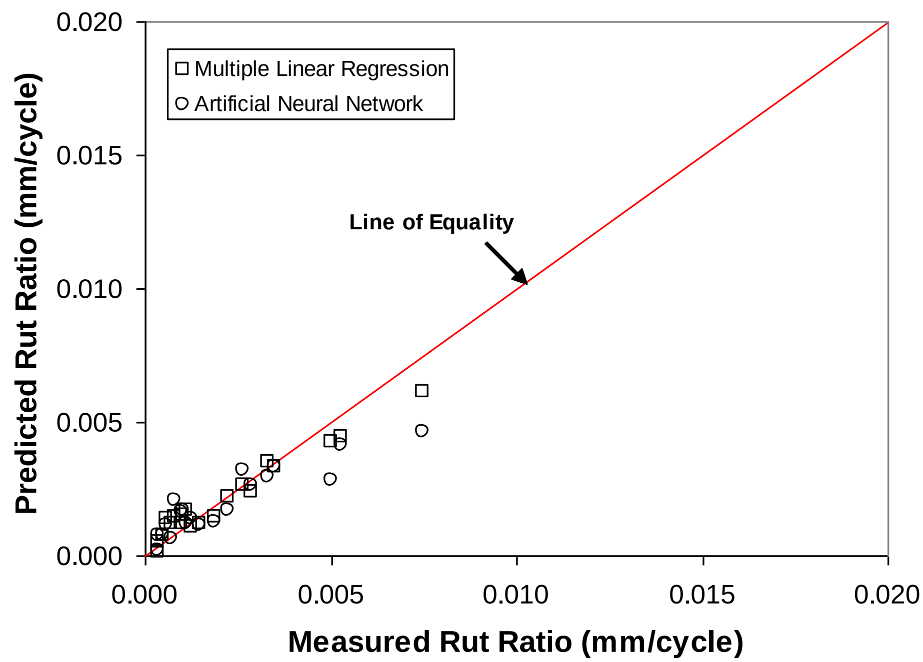


Figure 4.12 Multiple linear regression vs. artificial neural network

Chapter 5 Concluding Remarks

Based on this study, the following conclusions can be drawn, and follow-up studies are suggested:

5.1 Conclusions

- To find variables affecting APA rut results and the extent of these variables, a total of 91 *SP-4* mixture data from Nebraska and 21 samples from Kentucky were statistically analyzed using the multiple linear regression method. Based on literature review, six factors (binder PG, aggregate gradation, nominal maximum aggregate size, aggregate angularity, air voids in mixture, and asphalt content in mixture) were considered as probable candidates for significantly affecting APA rut results. For characterizing gradation effects, three indicators (gradation density, fineness modulus, and restricted zone) were considered and each of them was used for each statistical analysis.
- A common variable found from repeated multiple regression analyses by merely varying gradation indicators included was the binder PG. The binder PG was the variable that always showed significant impact on APA rut results from both Nebraska mixtures and Kentucky data.
- In the case of considering gradation effects by including gradation density factors, aggregate gradation was a significant factor together with the binder PG and the nominal maximum aggregate size. However, the effect of gradation is trivial in actual practice, because the significance of aggregate gradation resulting from the statistical analysis was due to a small number of mixtures in the data set.

- The trivial effect of aggregate gradation was verified by considering different gradation indicators (fineness modulus or restricted zone). Neither one showed any significance. Only binder PG showed its significance in the APA rutting.
- Predicting models were also developed using the results of multiple linear regression analysis and the artificial neural network technique. Adequacy of the models was investigated by observing coefficients of determination and cross-plotting predicted APA rut values to the measured APA rut data. Both methods generally did not provide a high level of model adequacy, which might be from a lack of data involved.

5.2 Recommended Further Studies

- A total of 91 *SP-4* data were used for the statistical analyses in this study. Even if the 91 data produced outcomes that were expected and were good agreements with other studies, more data would be helpful to derive better understanding. With more APA data gathered, the analyses can be conducted again.
- APA test results have generally shown poor correlations with actual field performance. Further research investigating the correlations using Nebraska data and finding out any significant factors of APA tests and results to the field rutting performance would be recommended. Similar statistical analyses employed for this study can be conducted.

5.3 NDOR Implementation Plan

The findings of this research project complement the past findings of NDOR laboratory personnel following limited years of APA testing of various asphalt mixtures in-house. Although currently a common mixture characteristic has not been identified to be capable of predicting in-field performance using APA, NDOR will continue to perform APA testing in 2008. NDOR has also agreed to become involved in an “Aggregate Imaging System” study sponsored by FHWA. This new study has the potential to provide aggregate surface texture information that may bring more meaning to the APA research findings.

References

Cross, S. A. and E. M. Purcell. 2001. "Effects of fine aggregate angularity on VMA and rutting of Kansas HMA mixtures." In T.D. White, S. R. Johnson, and J. J. Yzenas (editors), *Aggregate contribution to Hot Mix Asphalt (HMA) performance, ASTM STP 1412*, West Conshohocken, PA, American Society for Testing and Materials.

Draper, N. R. and H. Smith. 1998. *Applied regression analysis.* New York: John Wiley & Sons, Inc.

Hand, A. J., and A. L. Epps. 2001. "Impact of gradation relative to Superpave restricted zone on Hot-Mix Asphalt performance." *Transportation Research Record*, 1767: 158-166.

Kandhal, P. S., and R. B. Mallick. 2001. "Effect of mix gradation on rutting potential of dense-graded asphalt mixtures." *Transportation Research Record*, 1767: 146-151.

Kandhal, P. S. and L. A. Cooley, Jr. 2002. "Coarse- versus fine-graded Superpave mixtures – Comparative evaluation of resistance to rutting." *Transportation Research Record,* 1789: 216-224.

Kandhal, P. S. and L. A. Cooley, Jr. 2003. "Accelerated laboratory rutting tests: Evaluation of the asphalt pavement analyzer." *NCHRP Report No. 508,* National Research Council, Washington, D.C.

Lacroix, A., Y. R. Kim, and S. R. Ranjithan. 2008. "Backcalculation of the dynamic modulus from the resilient modulus of asphalt concrete using an artificial neural network." *Proceedings of the Transportation Research Board Annual Meeting*, Washington, D.C.

Lee, C., T. D. White, and T. R. West. 2000. "Effect of fine aggregate angularity on asphalt mixture performance." *Final Report FHWA/IN/JTRP-98/20*, Purdue University, West Lafayette, IN.

Mamlouk, M. S. and J. P. Zaniewski. 2006. *Materials for civil and construction engineers.* Upper Saddle River, NJ: Prentice Hall.

McCulloch, W. S. and W. Pitts. 1943. "A logical calculus of the ideas immanent in nervous activity." *Bulletin of Mathematical Biology*, 5: 115-133.

Mohammad, L. N., B. Huang, and Z. Z. Tan. 2001. "Evaluation of aggregate contribution to rutting susceptibility of asphalt mixture." American Society for Testing and Materials, West Conshohocken, PA.

Rumelhart, D. E., and J. L. McClelland. 1986. *Parallel distributed processing: Explorations in the microstructures of cognition.* Cambridge, MA: MIT Press / Bradford Books.

Sebaaly P. E., A. J. Hand, W. M. McNamara, D. Weitzel, and J. A. Epps. 2004. “Filed and laboratory performance of Superpave mixtures in Nevada.” *Transportation Research Record*, 1891: 76-84.

Shu, X., B. Huang, X. Chen, and L. Robison. 2006. “Effect of coarse aggregate angularity on rutting performance of HMA.” *Pavement mechanics and performance,* ASCE Geotechnical Special Publication No. 154: 126-133.

Stiady, J., A. Hand, and T. White. 2001. “Quantifying contributions of aggregate characteristics to HMA performance using PURWheel laboratory tracking device.” In T.D. White, S. R. Johnson, and J. J. Yzenas (editors), *Aggregate contribution to Hot Mix Asphalt (HMA) performance, ASTM STP 1412*, West Conshohocken, PA, American Society for Testing and Materials.

Tarefder, R. A., M. Zaman, and K. Hobson. 2003. “A laboratory and statistical evaluation of factors affecting rutting.” *International Journal of Pavement Engineering*, 4, no. 1: 59-68.

Uzarowski, L., M. Paradis, and P. Lum. 2004. “Accelerated performance testing of Canadian asphalt mixes using three different wheel rut testers.” *Proceedings of the 2004 Annual Conference of the Transportation Association of Canada*, Quebec City, Quebec.

Watson, D. E., A. Johnson, and D. Jared. 1997. “The Superpave gradation restricted zone and performance testing with Georgia loaded wheel tester.” *Transportation Research Record*, 1583: 106-111.

Appendix A

Table A.1 Datasheet of Nebraska *SP-4* mixtures

No.	NMAS	PG	Air void	% Binder	CAA1	CAA2	FAA	CAD	FAD	FM	RZ	Rut ratio
1	2	2	4.2	5.9	94	90	45.7	20.273	11.9	4.28	2	0.000156
2	2	1	3.9	5.14	89	86	45.4	15.941	8.533	4.477	2	0.000222
3	2	2	4.8	5.7	99	94	50.1	24.814	13.903	4.114	1	0.000251
4	1	2	4	6.83	95	90	46.9	10.02	10.707	3.813	2	0.00028
5	2	3	4.2	5.15	91	86	45.2	20.061	9.591	4.501	2	0.000302
6	2	3	4.5	5.6	94	90	45.3	13.094	10.39	4.751	3	0.000324
7	2	2	4.7	5.1	93	86	45.4	26.771	28.263	3.884	2	0.000353
8	2	2	4.1	5.43	98	98	45	18.883	20.515	4.08	2	0.000378
9	1	2	4.3	6.79	93	91	48	12.111	10.908	3.689	1	0.00038
10	2	2	3.8	5.51	94	90	45.5	19.554	13.403	4.235	2	0.00043
11	2	3	5	5.22	92	88	45.3	16.594	9.496	4.4	2	0.000434
12	1	2	3.8	5.36	94	83	45.3	9.002	10.706	4.218	3	0.00044
13	1	2	4.8	6.13	95	88	45.2	14.26	20.737	3.518	2	0.000459
14	2	2	3.7	5.46	93	80	45.4	24.99	16.418	4.046	2	0.000504
15	1	2	5.4	5.3	96	85	45.4	10.679	10.317	3.858	2	0.000509
16	1	2	5.4	5.3	96	85	45.4	10.679	10.579	3.867	2	0.000509
17	1	3	5.4	5.3	96	85	45.4	10.679	10.579	3.867	2	0.000509
18	2	2	4	5.06	96	87	45.3	22.11	10.74	4.278	2	0.000529
19	1	3	5.2	5.83	98	94	45.5	13.699	19.188	3.542	2	0.000549
20	2	2	4	5.28	97	95	45.2	20.588	9.061	4.468	2	0.00056
21	1	2	4.4	6.41	91	81	46.9	13.484	19.575	3.542	2	0.00058
22	1	2	3.9	5.99	98	95	45.3	13.743	14.581	3.74	2	0.000581
23	2	3	4.3	5.45	95	94	45.3	22.925	11.111	4.25	2	0.00059
24	1	2	4.3	5.35	89	85	45.5	7.769	9.089	4.178	2	0.000606
25	2	2	4.1	5.17	87	82	45.5	15.83	8.696	4.495	2	0.000635
26	2	2	4.1	5.19	99	96	45.6	18.48	16.737	4.194	2	0.000718
27	2	2	5.3	5.4	99	96	45	17.517	16.083	4.342	2	0.000774
28	1	2	5	5.67	85	81	45.2	12.231	15.685	3.795	2	0.000775
29	2	2	4.2	5.04	99	94	45.4	20.319	9.794	4.48	2	0.000799
30	2	1	6	5.21	96	94	45.3	19.329	9.437	4.632	3	0.000808
31	2	2	6	5.21	96	94	45.3	19.329	9.437	4.632	3	0.000808
32	1	2	3.9	6.31	86	81	45.5	14.42	16.993	3.592	2	0.000837
33	1	2	3.9	6.31	86	81	45.5	14.42	17.027	3.591	2	0.000837
34	2	1	4.2	5.04	93	84	45.1	24.096	21.629	3.96	2	0.000843
35	2	2	4.1	5.2	96	95	45.2	22.908	18.909	4.037	2	0.000848

Table A.1 (cont'd.) Datasheet of Nebraska *SP-4* mixtures

No.	NMAS	PG	Air void	% Binder	CAA1	CAA2	FAA	CAD	FAD	FM	RZ	Rut ratio
36	2	2	5.2	5.42	94	88	45.3	20.232	11.326	4.313	2	0.000876
37	2	2	4	5.41	85	80	45.3	24.295	16.26	4.074	2	0.000906
38	2	2	3.7	5.7	89	86	45.2	21.71	9.824	4.318	2	0.000909
39	1	2	3.5	5.6	92	88	45.1	9.968	13.865	3.809	2	0.000932
40	2	2	3.9	5.4	92	91	45.4	12.149	9.208	4.775	3	0.00097
41	2	2	4.6	5.2	92	89	45.4	22.208	15.653	4.185	2	0.000997
42	1	2	4.2	5.32	87	81	45.5	5.912	13.352	4.418	3	0.00101
43	2	1	4.1	5.1	95	91	45.8	12.599	8.159	4.739	3	0.001059
44	1	2	4.5	5.04	99	92	45.2	3.363	10.645	4.37	3	0.001075
45	2	1	4	5.24	89	87	45.2	20.7	10.831	4.456	2	0.001093
46	2	3	3.8	5.12	96	89	45.1	19.311	10.417	4.307	2	0.001108
47	1	2	3.9	5.42	96	88	45.5	9.002	10.759	4.222	3	0.001119
48	1	1	4.3	6.4	97	96	45.9	14.127	20.838	3.474	2	0.001154
49	1	2	4	5.68	86	83	45.1	8.992	10.605	4.059	2	0.001196
50	2	2	4	5.2	95	83	45.5	20.727	11.112	4.462	2	0.001214
51	2	2	4.3	5.2	85	80	45.3	20.041	12.807	4.31	2	0.001255
52	1	2	5	6.29	92	89	45.2	9.454	14.536	3.79	2	0.001267
53	1	2	4.1	6.3	92	87	45.2	13.718	17.691	3.624	2	0.001286
54	2	1	4.6	5.86	94	90	45.4	20.875	20.479	4.053	2	0.00133
55	2	2	4.2	5.54	92	81	45.1	19.427	14.43	4.277	2	0.001363
56	2	3	3.7	5.12	90	82	45.2	21.088	9.708	4.378	2	0.001425
57	2	2	4	5.25	94	93	45.1	14.672	13.961	4.264	2	0.001497
58	2	1	3.8	5.48	95	91	45.2	21.679	19.623	4.032	2	0.001548
59	1	2	4.2	5.8	87	81	45.9	11.725	11.509	4.126	3	0.001658
60	2	3	3.7	5.23	90	86	45.5	22.422	12.541	4.46	2	0.001879
61	2	2	3.8	5.34	96	89	45.6	20.875	20.479	4.053	2	0.001915
62	2	2	3.5	5.13	93	87	45.2	20.728	9.392	4.415	2	0.002222
63	2	2	4.6	5.56	96	93	45	24.882	18.575	3.993	1	0.00223
64	2	1	4.5	5.6	94	90	45.3	13.094	10.39	4.751	3	0.002627
65	2	2	4.5	5.6	94	90	45.3	13.094	10.39	4.751	3	0.002627
66	2	1	4.5	5.24	94	92	45	12.945	13.6	4.337	2	0.002861
67	1	1	5.1	6	91	83	45.1	15.58	18.348	3.577	2	0.002977
68	1	2	4.4	5.1	85	81	45.1	7.537	10.89	4.221	2	0.003246
69	1	2	3.5	5.4	93	81	45.3	7.079	10.372	4.223	3	0.003329
70	2	1	4.7	4.82	91	80	45.5	21.366	11.786	4.331	2	0.003428
71	1	2	4.2	6.12	88	83	45.1	13.208	12.819	3.908	2	0.003503
72	2	1	4.6	4.6	91	80	45.1	21.697	10.571	4.36	2	0.003707

Table A.1 (cont'd.) Datasheet of Nebraska *SP-4* mixtures

No.	NMAS	PG	Air void	% Binder	CAA1	CAA2	FAA	CAD	FAD	FM	RZ	Rut ratio
73	1	1	5.2	6.32	90	87	45.2	12.367	17.933	3.703	2	0.004646
74	2	1	3.7	6	93	89	45.1	11.039	13.934	4.992	3	0.004876
75	1	2	3.5	5.56	99	95	45.2	10.327	10.54	4.15	3	0.005015
76	1	2	4.1	5.31	99	93	45.3	6.851	7.884	4.13	2	0.005025
77	2	1	4.3	5.47	95	92	45	20.212	13.118	4.205	2	0.005272
78	1	2	5	5.39	96	94	45.5	10.107	11.797	4.18	2	0.005374
79	1	2	3.9	5.41	87	83	45.3	9.002	10.664	4.234	3	0.005859
80	2	2	4.2	5.31	96	87	45.5	20.676	10.563	4.543	2	0.006128
81	1	2	4.4	5.63	94	89	45.2	9.573	9.728	4.169	3	0.006239
82	1	2	4.2	5.29	92	90	45.2	10.33	10.98	4.111	2	0.007023
83	1	2	3.9	5.32	92	90	45.1	9.889	10.246	4.11	2	0.007023
84	1	1	4.1	5.13	97	91	45.1	12.335	12.384	3.949	2	0.007506
85	1	2	4.1	5.13	97	94	45.1	12.335	12.384	3.949	2	0.007506
86	2	1	4.2	5.04	93	84	45.1	24.096	21.629	3.96	2	0.00806
87	2	2	3.5	5.49	88	84	45	22.798	13.005	4.471	2	0.009532
88	1	1	4.1	5.53	90	87	45.1	13.112	13.753	3.883	2	0.009756
89	2	2	3.6	5.41	88	80	45.1	24.382	15.456	4.123	2	0.011962
90	2	1	5	5.56	91	82	45.5	21.819	12.669	4.441	2	0.012956
91	2	2	4	5.11	95	89	45.2	19.451	9.234	4.498	2	0.018115

Appendix B

Table B.1 Datasheet of Kentucky mixtures

NO.	NMAS	PG	Air void	% Binder	CAA1	CAA2	FAA	CAD	FAD	Rut ratio
1	1	1	3.8	6.3	100	100	45	4.73103	11.1229	0.00082
2	1	1	4	6.2	98	98	46	3.93326	9.11786	0.00116
3	1	1	4	6.2	98	98	46	3.93326	9.11786	0.00165
4	1	2	4.2	5.9	100	100	46	12.6013	13.8411	0.00046
5	1	3	4	5.4	100	100	45	3.10271	16.4924	0.00027
6	1	3	4.1	6	100	100	46	2.74071	9.78786	0.00018
7	1	3	4.2	5.4	100	100	47	3.03526	6.51035	0.00023
8	1	3	4	5.9	100	100	46	6.53026	10.0281	0.00018
9	1	3	3.9	6	100	100	46	4.85926	11.7491	0.00027
10	1	3	4.1	6.2	100	100	46	6.33526	8.1569	0.00026
11	2	1	4	5.8	99	99	48	15.6455	8.03541	0.00151
12	2	1	4.1	5.6	100	100	45	15.8132	11.4804	0.00063
13	2	1	4	5.5	100	100	45	14.3012	11.4804	0.00057
14	2	3	4.1	5.7	100	100	46	19.4815	8.05308	0.00032
15	2	3	4	5.8	100	100	45	16.8695	14.1575	0.00047
16	2	3	4.3	5.5	100	100	48	22.0185	12.8929	0.00023
17	2	3	4	5.8	99	99	48	15.6455	8.03541	0.00044
18	2	3	4	5.8	99	99	48	15.6455	8.03541	0.00057
19	2	3	4.1	6.3	100	100	45	18.4055	13.0956	0.00032
20	2	3	4.1	5.6	100	100	45	23.1715	11.5372	0.00033
21	2	3	4.1	5.6	100	100	48	13.8955	9.28995	0.00025