An Assessment of the Models to Predict Pavement Performance

Final Report
March 2018

Sponsored by
Midwest Transportation Center
U.S. Department of Transportation
Office of the Assistant Secretary for
Research and Technology

MIDWEST
TRANSPORTATION
CENTER
MTC

IOWA STATE UNIVERSITY
Institute for Transportation

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
An Assessment of the Models to Predict Pavement Performance	March 2018
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
William Duckworth, Ravi Nath, and Victor Ekpoke	
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
Department of Business Intelligence & Analytics Heider College of Business Creighton University 2500 California Plaza Omaha, Nebraska 68178	**11. Contract or Grant No.** Part of DTRT13-G-UTC37
12. Sponsoring Organization Name and Address	**13. Type of Report and Period Covered**
Midwest Transportation Center 2711 S. Loop Drive, Suite 4700 Ames, IA 50010-8664 U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE Washington, DC 20590	Final Report **14. Sponsoring Agency Code**

15. Supplementary Notes

Visit www.intrans.iastate.edu for color pdfs of this and other research reports.

16. Abstract

Data collected by the Iowa Department of Transportation (DOT) regarding road conditions across the state of Iowa were used to model pavement condition index (PCI). The data were for calendar year 2013, with the exception of updated PCI values from 2014 and 2015 and indicators of the resurfacing of road segments in 2014 and 2015. The data file provided by the Iowa DOT consisted of nearly 4,000 observations.

Eighteen different road conditions and measures were considered as possible model inputs. Of the 18 measures, 11 were used in the final prediction of PCI in 2014 and 2015 for portland cement, composite, and asphalt cement pavement types. These measures included International Roughness Index (IRI), friction value, age, average daily traffic, PCI value in 2013, number of lanes, daily temperature change, surface type, pavement thickness, speed limit, and reconstructed kips.

Series of multiple regression models were developed for the different pavement types, including aggregated pavement types with combined data. The results reveal that all 11 variables except age have a statistically significant relationship with PCI. The efficacies of the derived models, as measured by R^2 values, range from 61% to 83%. Additional analyses also show that the efficacies of the derived models, as measured by root mean square error (RMSE) values, range from 6.29 to 9.52. We can interpret the RMSE values as indicating that approximately 95% of all prediction values should fall within 12.58 and 19.04 of the PCI values predicted by the models. Therefore, it is concluded that linear predictive models, which involve distress and descriptive characteristics of road conditions, provide a reasonable basis for estimating PCI. However, these models can be further improved by examining nonlinear effects.

17. Key Words	18. Distribution Statement
analytics—pavement condition index—prediction—regression	No restrictions.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	24	NA

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

PAVEMENT PERFORMANCE: APPROACHES USING PREDICTIVE ANALYTICS

Final Report
March 2018

Principal Investigator
William Duckworth, Associate Professor
Department of Business Intelligence & Analytics, Heider College of Business, Creighton University

Co-Principal Investigator
Ravi Nath, Department Chair
Department of Business Intelligence & Analytics, Heider College of Business, Creighton University

Research Assistant
Victor Ekpoke

Authors
William Duckworth, Ravi Nath, and Victor Ekpoke

Sponsored by
Midwest Transportation Center and
U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology

A report from
Institute for Transportation
Iowa State University
2711 South Loop Drive, Suite 4700
Ames, IA 50010-8664
Phone: 515-294-8103 / Fax: 515-294-0467
www.intrans.iastate.edu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The authors would like to thank the Midwest Transportation Center and the U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology for sponsoring this research.

MODELING PAVEMENT CONDITION INDEX

Data collected by the Iowa Department of Transportation (DOT) regarding road conditions across the state of Iowa were used to model pavement condition index (PCI). All data were for calendar year 2013, with the exception of updated PCI values from 2014 and 2015 and indicators of the resurfacing of road segments in 2014 and 2015.

The research described in this report investigated the use of various road characteristics and measures to model future PCI values. Specifically, predictive models for PCI values in 2014 and 2015 were developed using only data available at the close of 2013. Eighteen different road conditions and measures were considered as possible model inputs. Of the 18 measures, 11 were used in the final prediction of PCI in 2014 and 2015 for portland cement, composite, and asphalt cement pavement types. These measures included International Roughness Index (IRI), friction value, age, average daily traffic, PCI value in 2013, number of lanes, daily temperature change, surface type, pavement thickness, speed limit, and reconstructed kips. Comprehensive descriptions of the variables are provided in Appendix A.

Analyses were conducted by pavement type for those pavement types with sufficient data (Pavement Types 1, 3, and 4 in the Iowa DOT data). An overall analysis for all pavement types combined is also presented in this report.

The data file was provided by the Iowa DOT and included nearly 4,000 observations.

All analyses were completed using JMP Pro software (version 12.1.0, 64-bit) from SAS Institute, Inc. The analysis workflow incorporated multiple regression modeling, including multicollinearity considerations and residual analyses. Variable selection techniques utilized in the analyses included stepwise regression and JMP's All Possible Models platform. Best model fit was determined by minimizing model root mean square error (RMSE).

MODELING PCI FOR PAVEMENT TYPE 1

Pavement Type 1 is portland cement (PC). About 35% of the observations in the data set were for Pavement Type 1. Of the 18 measures, only 5 remain in the final model for predicting PCI one year ahead (predicting PCI_2014). Thirteen variables were easily eliminated from consideration based on collinearity concerns and statistical insignificance ($p>5\%$). The prediction equation is as follows:

Predicted PCI_2014 = 3.65 + 0.8*PCI_2013 + 0.066*IRI_Index + 0.124*Friction_Value + 0.0001*Average_Daily_Traffic + 0.62*Number_Of_Lanes

The model has an RMSE of 6.29 and an R^2 of 77.9%.

We can interpret the RMSE value as indicating that approximately 95% of all PCI_2014 values (for Pavement Type 1) should fall within 12.58 (2*6.29) of the PCI predicted by this model. The R^2 indicates that approximately 77.9% of the observation-to-observation variability in recorded PCI_2014 values can be accounted for by this model.

All variables in this model are statistically significant ($p<5\%$).

If we turn our attention to predicting PCI two years ahead (predicting PCI_2015), then of the 18 measures, only 7 remain in the final model. As expected, the model predicting PCI two years ahead does not fit as well as the model predicting PCI one year ahead. The prediction equation is as follows:

Predicted PCI_2015 = 13.18 + 0.54*PCI_2013 + 0.02*DaysTempChange_2013 + 0.19*IRI_Index + 0.14*Friction_Value + 0.00017*Average_Daily_Traffic + 0.00000004*Reconstruct_18_KIPS + 0.9*Number_Of_Lanes

The model has an RMSE of 8.76 and an R^2 of 61.7%.

Approximately 95% of all PCI_2015 values for Pavement Type 1 should fall within 17.52 (2*8.76) of the PCI predicted by this model. The R^2 indicates that approximately 61.7% of the observation-to-observation variability in recorded PCI_2015 values can be accounted for by this model.

All variables in this model are statistically significant ($p<5\%$).

Table 1 shows the summary statistics for each of the variables used to model Pavement Type 1.

Table 1. Summary statistics for Pavement Type 1

Variable	N	Mean	Std Dev	Minimum	Maximum
Age_2013	1251	25.3493	17.5327	0.0000	87.0000
Speed_Limit	1235	55.2429	10.6600	20.0000	70.0000
Pavement_Thickness	1251	10.0008	1.4948	3.0000	19.0000
Number_Of_Lanes	1251	3.7754	1.2119	2.0000	9.0000
Average_Daily_Traffic	1166	11543.7	11981.8	10.0000	90400.0
Average_Daily_Trucks	1184	1541.85	2149.64	66.0000	11498.0
TRAFFIC DATA-Annual_18_KIPS	1244	412534	695016	468.000	3672775
Accum_KIPS_Since_Resurfacing	32	4142452	3599507	53610.0	1.06E+07
Reconstruct_18_KIPS	1171	1.03E+07	1.67E+07	166816	9.46E+07
CONDITION DATA-IRI_Index	1246	46.7608	21.9242	0.0000	84.0000
Friction_Value	921	51.1042	7.2548	22.0000	65.0000
Surface_Type	1251	73.7458	2.1371	60.0000	92.0000
Pavement_Width	1230	27.2568	7.9389	5.9000	67.9000
PCI_2013	1246	62.1798	15.9733	7.0000	94.0000
DaysTempChange_2013	1251	88.9488	32.2593	0.0000	137.000
PCI_2014	1236	65.6076	16.7997	7.0000	98.0000
PCI_2015	1196	69.3687	17.6523	11.0000	100.000

The selected scatter plots in Figure 1 show how four variables—PCI_2014, PCI_2013, Age_2013, and CONDITION DATA-IRI_Index—interact to help predict PCI_2014 for Pavement Type 1. The scatter plots in Figure 2 use the same variables to help predict PCI_2015.

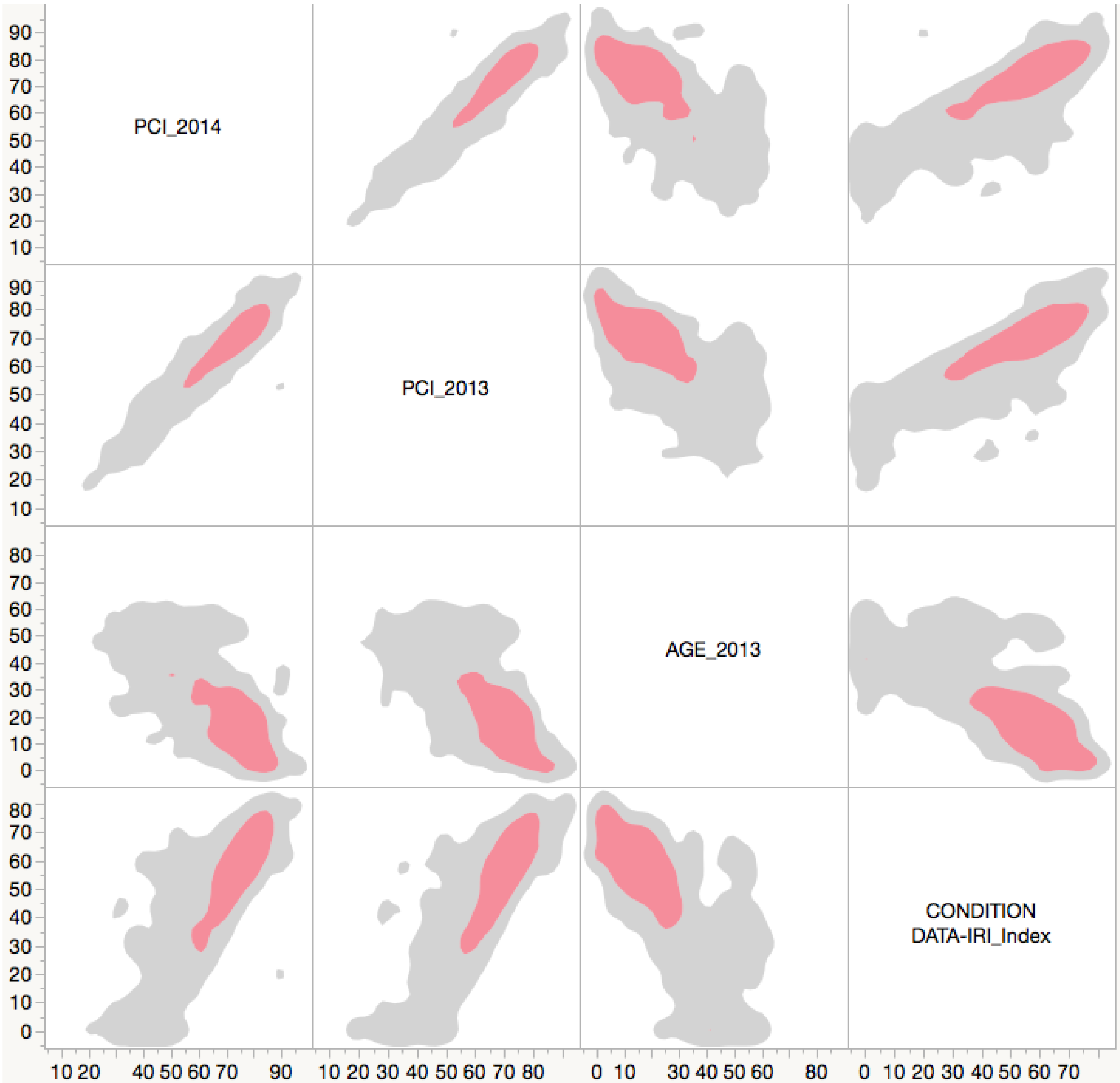

Figure 1. Selected scatter plots showing how four variables interact to predict PCI_2014 for Pavement Type 1

Figure 2. Selected scatter plots showing how four variables interact to predict PCI_2015 for Pavement Type 1

MODELING PCI FOR PAVEMENT TYPE 3

Pavement Type 3 is composite pavement, which typically indicates portland cement or continuously reinforced concrete overlaid with asphalt at some point in the life of the road. About 52% of the observations in the data set were for Pavement Type 3. Of the 18 measures, only 9 remain in the final model for predicting PCI one year ahead (predicting PCI_2014). Nine variables were easily eliminated from consideration based on collinearity concerns and statistical insignificance (p>5%). The prediction equation is as follows:

Predicted PCI_2014 = 29.52 + 0.75*PCI_2013 + 0.15*IRI_Index + 0.00015*Average_Daily_Traffic – 0.0000008*Accum_KIPS_Since_Resurfacing + 0.06*Surface_Type – 0.16*Pavement_Thickness + 1.86*(if Median=YES) – 0.06*AGE_2013 + 19.7*(if RS_in2014=YES)

The model has an RMSE of 7.35 and an R^2 of 83.1%.

We can interpret the RMSE value as indicating that approximately 95% of all PCI_2014 values for Pavement Type 3 should fall within 14.7 (2*7.35) of the PCI predicted by this model. The R^2 indicates that approximately 83.1% of the observation-to-observation variability in recorded PCI_2014 values can be accounted for by this model.

All variables in this model are statistically significant (p<5%).

If we turn our attention to predicting PCI two years ahead (predicting PCI_2015), then only 9 of the 18 measures remain in the final model. As expected, the model predicting PCI two years ahead does not fit as well as the model predicting PCI one year ahead. The prediction equation is as follows:

Predicted PCI_2015 = 51.65 + 0.7*PCI_2013 + 0.00012*Average_Daily_Traffic + 0.000008*Annual_18_KIPS – 0.0000015*Accum_KIPS_Since_Resurfacing + 0.22*Speed_Limit + 0.89*(if Median=YES) – 0.17*AGE_2013 + 17.7*(if RS_in2014=YES) + 21.3*(if RS_in2015=YES)

The model has an RMSE of 9.52 and an R^2 of 71.1%.

Approximately 95% of all PCI_2015 values for Pavement Type 3 should fall within 19.04 (2*9.52) of the PCI predicted by this model. The R^2 indicates that approximately 71.1% of the observation-to-observation variability in recorded PCI_2015 values can be accounted for by this model.

All variables in this model are statistically significant (p<5%).

Table 2 shows the summary statistics for each of the variables used to model Pavement Type 3.

Table 2. Summary statistics for Pavement Type 3

Variable	N	Mean	Std Dev	Minimum	Maximum
Age_2013	1875	13.6645	10.4561	0.0000	84.0000
Speed_Limit	1875	51.3493	9.3573	20.0000	70.0000
Pavement_Thickness	1875	13.8827	2.8245	3.0000	29.0000
Number_Of_Lanes	1875	2.7349	1.0975	2.0000	8.0000
Average_Daily_Traffic	1874	5913.57	6496.80	380.000	82800.0
Average_Daily_Trucks	1838	532.639	519.889	45.0000	7187.00
TRAFFIC DATA-Annual_18_KIPS	1875	86630.0	117258	5740.00	1728320
Accum_KIPS_Since_Resurfacing	1686	1529599	1461844	6050.00	1.07E+07
Reconstruct_18_KIPS	1860	4618742	4337785	256205	7.25E+07
CONDITION DATA-IRI_Index	1875	55.4741	20.3906	0.0000	100.000
Friction_Value	1257	51.0350	7.5218	22.0000	68.0000
Surface_Type	1875	67.1051	6.5332	60.0000	92.0000
Pavement_Width	1875	27.6139	8.0814	16.1000	71.9000
PCI_2013	1875	64.9253	16.9972	11.0000	100.000
DaysTempChange_2013	1875	91.9595	26.6472	0.0000	137.000
PCI_2014	1853	67.7312	17.9984	14.0000	97.0000
PCI_2015	1797	68.1391	17.6334	6.0000	97.0000

The selected scatter plots in Figures 3 and 4 show how four variables interact to help predict PCI_2014 and PCI_2015, respectively, for Pavement Type 3.

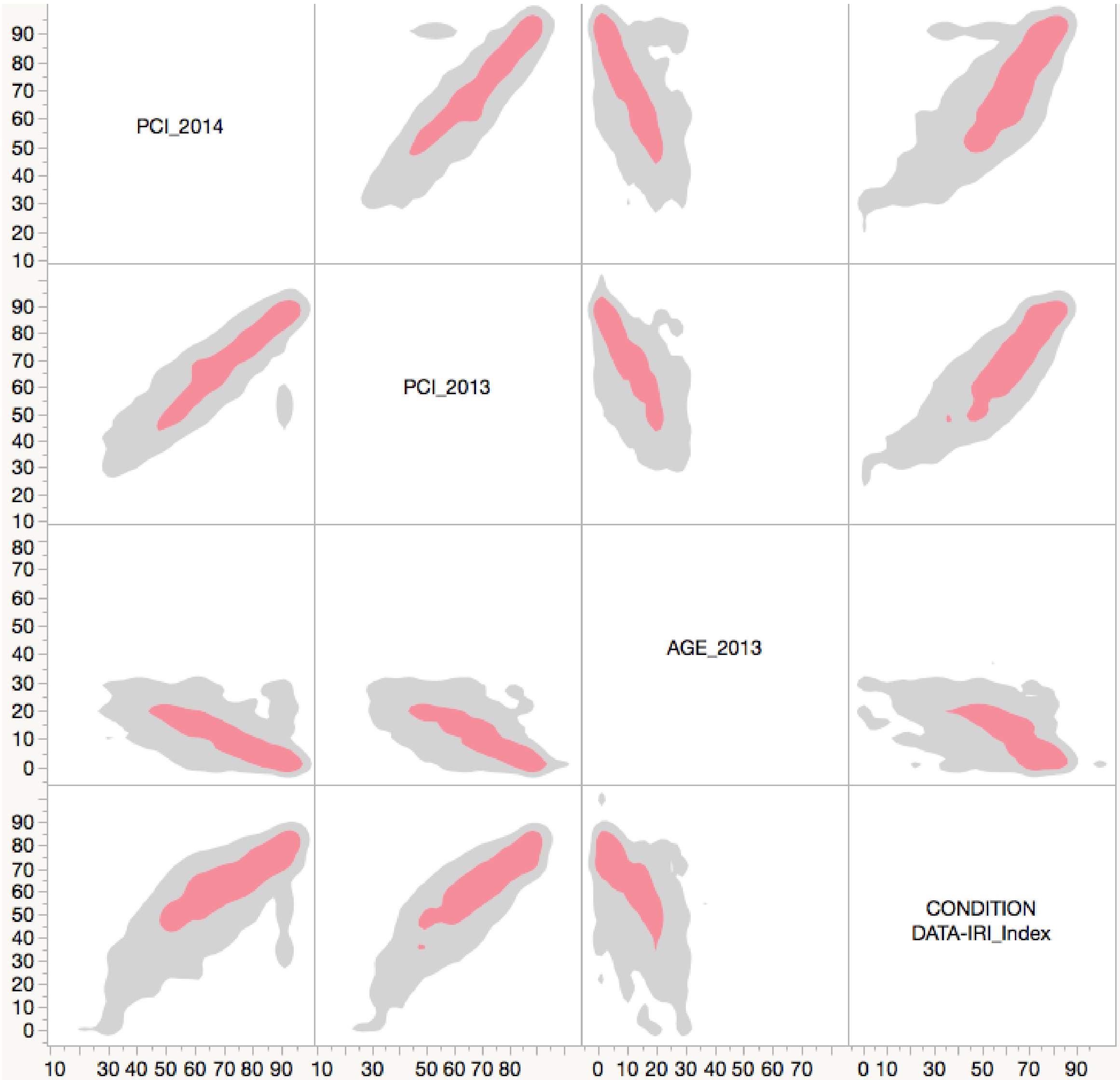

Figure 3. Selected scatter plots showing how four variables interact to predict PCI_2014 for Pavement Type 3

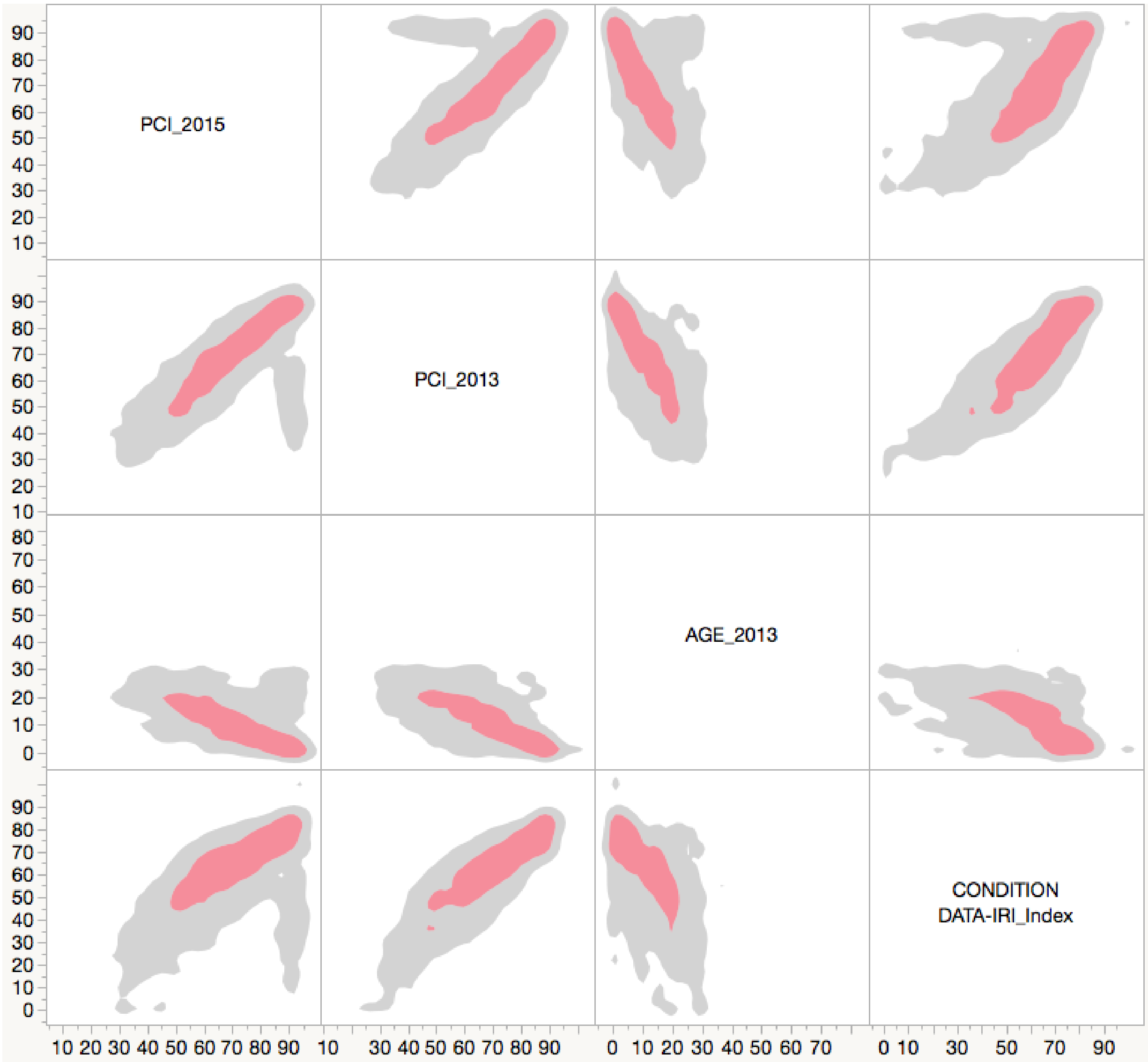

Figure 4. Selected scatter plots showing how four variables interact to predict PCI_2015 for Pavement Type 3

MODELING PCI FOR PAVEMENT TYPE 4

Pavement Type 4 is asphalt cement (AC). About 13% of the observations in the data set were for Pavement Type 4. Of the 18 measures, only 4 remain in the final model for predicting PCI one year ahead (predicting PCI_2014). Fourteen variables were easily eliminated from consideration based on collinearity concerns and statistical insignificance (p>5%). The prediction equation is as follows:

Predicted PCI_2014 = 32.24 + 0.86*PCI_2013 + 1.63*(if Median=YES) – 0.09*AGE_2013 + 18.55*(if RS_in2014=YES)

The model has an RMSE of 7.70 and an R^2 of 83.9%.

We can interpret the RMSE value as indicating that approximately 95% of all PCI_2014 values for Pavement Type 4 should fall within 15.4 (2*7.70) of the PCI predicted by this model. The R^2 indicates that approximately 83.9% of the observation-to-observation variability in recorded PCI_2014 values can be accounted for by this model.

All variables in this model except AGE_2013 (p=0.0645) are statistically significant (p<5%); however, removing AGE_2013 increases the RMSE of the model.

If we turn our attention to predicting PCI two years ahead (predicting PCI_2015), then only 11 of the 18 measures remain in the final model. As expected, the model predicting PCI two years ahead does not fit as well as the model predicting PCI one year ahead. The prediction equation is as follows:

Predicted PCI_2015 = 83.78 + 0.71*PCI_2013 + 0.14*IRI_Index - 0.18*Friction_Value - 0.004*Average_Daily_Trucks + 0.00002*Annual_18_KIPS - 0.388*Pavement_Thickness + 5.8*(if Median=YES) - 4.84*Number_Of_Lanes - 0.26*AGE_2013 + 19.25*(if RS_in2014=YES) + 19.96*(if RS_in2015=YES)

The model has an RMSE of 8.38 and an R^2 of 80.5%.

Approximately 95% of all PCI_2015 values for Pavement Type 4 should fall within 16.76 (2*8.38) of the PCI predicted by this model. The R^2 indicates that approximately 80.5% of the observation-to-observation variability in recorded PCI_2015 values can be accounted for by this model.

All variables in this model are statistically significant (p<5%).

Table 3 shows the summary statistics for each of the variables used to model Pavement Type 4.

Table 3. Summary statistics for Pavement Type 4

Variable	N	Mean	Std Dev	Minimum	Maximum
Age_2013	462	15.0216	10.7763	0.0000	83.0000
Speed_Limit	462	54.5887	7.5491	20.0000	70.0000
Pavement_Thickness	462	11.9740	4.3446	3.0000	29.0000
Number_Of_Lanes	462	2.4697	1.0194	2.0000	9.0000
Average_Daily_Traffic	461	4511.02	8426.04	370.000	90400.0
Average_Daily_Trucks	458	791.282	1845.13	10.0000	11344.0
TRAFFIC DATA-Annual_18_KIPS	462	149744	386828	1760.00	2265580
Accum_KIPS_Since_Resurfacing	340	1829318	5146355	8090.00	4.08E+07
Reconstruct_18_KIPS	458	5537606	1.45E+07	304281	1.11E+08
CONDITION DATA-IRI_Index	462	59.2446	19.8291	0.0000	100.000
Friction_Value	357	52.3417	7.6589	20.0000	70.0000
Surface_Type	462	63.6190	6.0724	30.0000	92.0000
Pavement_Width	458	25.0048	4.5177	22.0000	54.1000
PCI_2013	462	64.9199	18.4179	9.0000	100.000
DaysTempChange_2013	462	86.9264	32.5324	0.0000	137.000
PCI_2014	459	67.5447	18.6986	10.0000	97.0000
PCI_2015	453	68.4128	18.7326	22.0000	97.0000

The scatter plots in Figures 5 and 6 show how four variables interact to help predict PCI_2014 and PCI_2015, respectively, for Pavement Type 4.

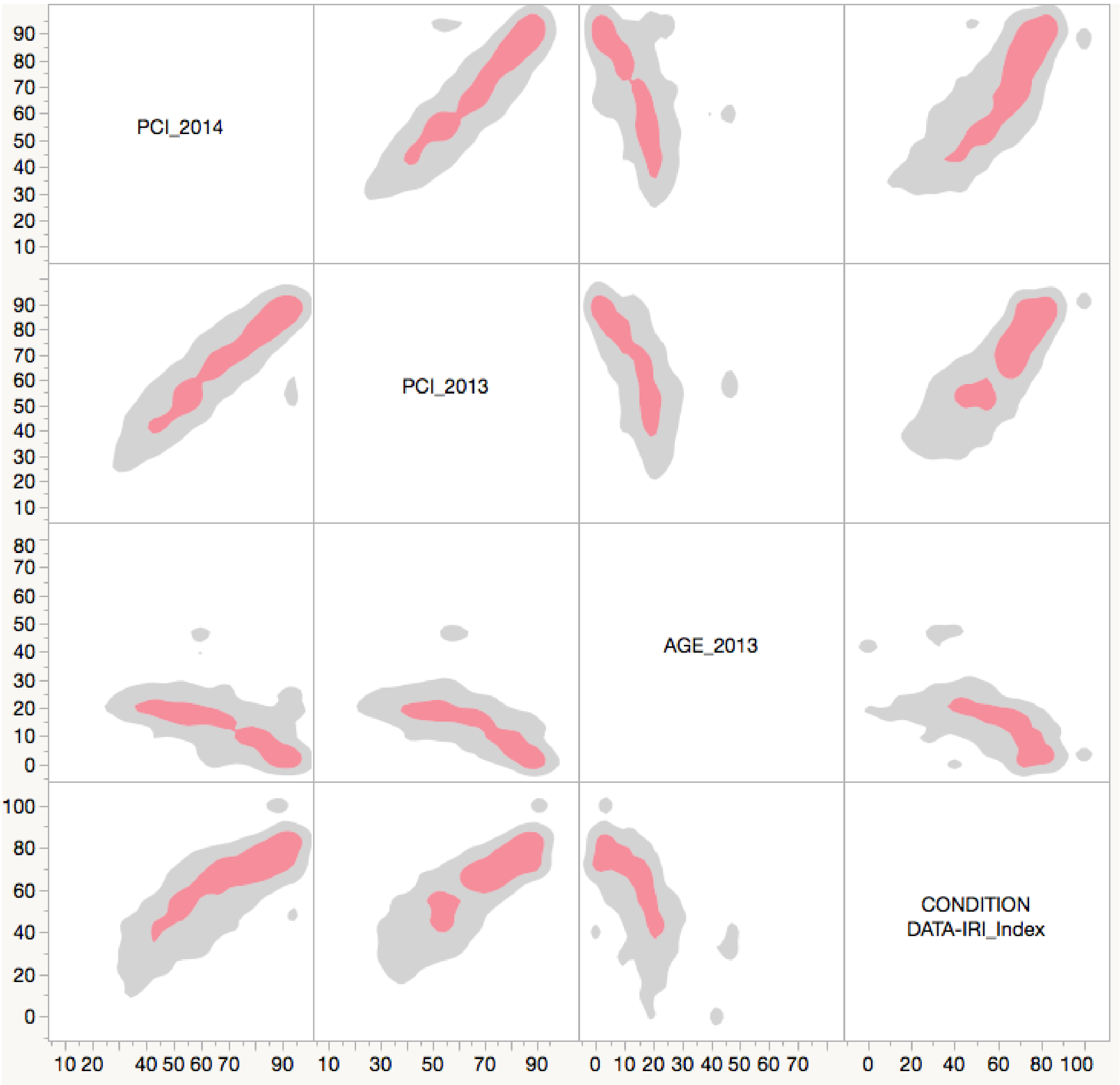

Figure 5. Selected scatter plots showing how four variables interact to predict PCI_2014 for Pavement Type 4

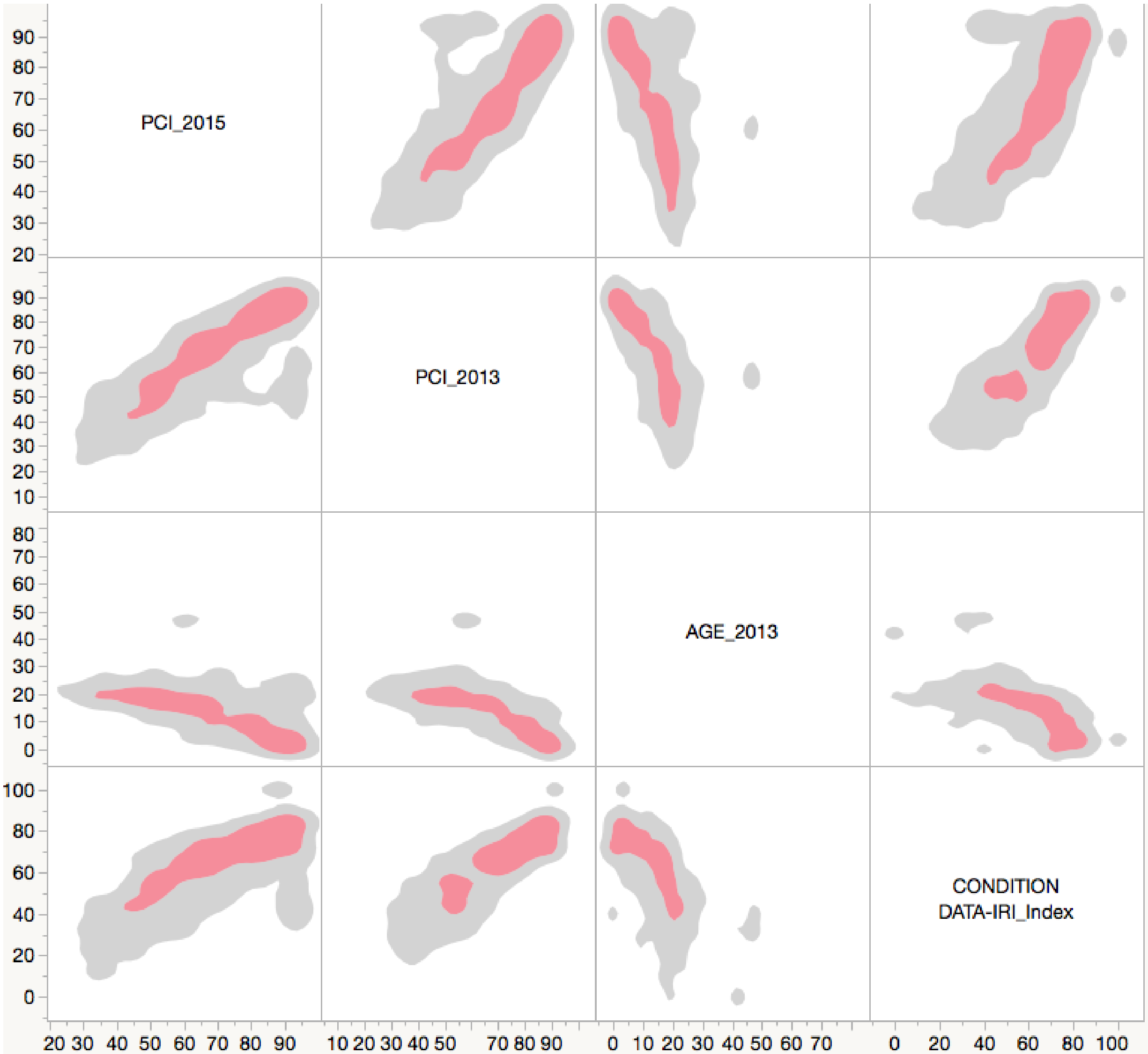

Figure 6. Selected scatter plots showing how four variables interact to predict PCI_2015 for Pavement Type 4

APPENDIX A. VARIABLE DEFINITIONS

The descriptive variables considered as potential input variables in this research were as follows:

- AGE_2013 – years since construction or resurfacing (as calculated from 2013 data)
- Speed_Limit – speed limit in miles per hour
- Pavement_Thickness – pavement thickness in inches
- Number_Of_Lanes – number of lanes
- Average_Daily_Traffic – average daily traffic as a count per day
- Average_Daily_Trucks – number of trucks per day
- Annual_18_KIPS – annual 18 kips measured in ESALs
- Accum_KIPS_Since_Resurfacing – accumulated kips since resurfacing, measured in kips
- Reconstruct_18_KIPS – accumulated kips since construction, measured in kips
- IRI_Index – International Roughness Index
- Friction_Value – friction value from 5 to 75
- Surface_Type – surface type from 0 to 96
- Pavement_Width – pavement width
- Median – YES/NO, with YES indicating the segment has a median and NO indicating the segment does not have a median
- DaysTempChange_2013 – Number of days in 2013 where the maximum temperature was greater than 32°F and the minimum temperature was less than or equal to 32°F
- RS_in2013 – YES/NO, with YES indicating the segment was resurfaced in 2013 and NO indicating the segment was not resurfaced in 2013
- RS_in2014 – YES/NO, with YES indicating the segment was resurfaced in 2014 and NO indicating the segment was not resurfaced in 2014
- RS_in2015 – YES/NO, with YES indicating the segment was resurfaced in 2015 and NO indicating the segments was not resurfaced in 2015
- PCI_2013 – pavement condition index as recorded in 2013 data