Pavement Performance: Approaches Using Predictive Analytics

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

IOWA STATE UNIVERSITY
Institute for Transportation

About MTC

The Midwest Transportation Center (MTC) is a regional University Transportation Center (UTC) sponsored by the U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology (USDOT/OST-R). The mission of the UTC program is to advance U.S. technology and expertise in the many disciplines comprising transportation through the mechanisms of education, research, and technology transfer at university-based centers of excellence. Iowa State University, through its Institute for Transportation (InTrans), is the MTC lead institution.

About InTrans

The mission of the Institute for Transportation (InTrans) at Iowa State University is to develop and implement innovative methods, materials, and technologies for improving transportation efficiency, safety, reliability, and sustainability while improving the learning environment of students, faculty, and staff in transportation-related fields.

ISU Non-Discrimination Statement

Notice

Quality Assurance Statement

Technical Report Documentation Page

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

4. Title and Subtitle

Pavement Performance: Approaches Using Predictive Analytics

5. Report Date

March 2018

6. Performing Organization Code

7. Author(s)

William Duckworth and Ravi Nath

8. Performing Organization Report No.

9. Performing Organization Name and Address

Department of Business Intelligence & Analytics
Heider College of Business
Creighton University
2500 California Plaza
Omaha, Nebraska 68178

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

Part of DTRT13-G-UTC37

12. Sponsoring Organization Name and Address

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

13. Type of Report and Period Covered

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

Acceptable pavement condition is paramount to road safety. Using predictive analytics techniques, this project attempted to develop models that provide an assessment of pavement condition based on an array of indictors that include pavement distress, pavement type, traffic load, structural data, and pavement repair history. 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 from calendar year 2013 and consisted of nearly 4,000 observations.

Various distress indicators were used to model PCI. These distress measures quantify a variety of cracks (types of cracks, severity of cracks, and amount of cracking) as well as joint spalling (severity and amount) and the condition of previous patching (condition and amount). Twenty-three distress measures were considered as possible model inputs. In addition to distress measures, nine descriptive variables were tested as potential model inputs for improving the overall fit of the model to the data. These descriptive variables included traffic, load, speed limit, number of lanes, pavement thickness, and pavement age.

Series of multiple regression models were developed for different pavement types and for combined data (when all pavement types were aggregated). The results reveal that a number of distress variables and descriptive variables have a statistically significant relationship with PCI. The efficacies of the derived models, as measured by R^2 values, range from 44% to 86%. The results of further analyses show that the introduction of the quadratic effects of certain variables on PCI improves model efficacy. Therefore, it is concluded that linear predictive models that 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

analytics—pavement condition index—prediction—regression

18. Distribution Statement

No restrictions.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	26	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, Chair
Department of Business Intelligence & Analytics, Heider College of Business, Creighton University

Authors
William Duckworth and Ravi Nath

Sponsored by
Midwest Transportation Center and
the 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 from calendar year 2013.

The research described in this report investigated the use of various distress measures to model PCI. These distress measures quantify a variety of cracks (types of cracks, severity of cracks, and amount of cracking) as well as joint spalling (severity and amount) and the condition of previous patching (condition and amount). Twenty-three distress measures were considered as possible model inputs.

In addition to distress measures, nine descriptive variables were tested as potential model inputs for improving the overall fit of the model to the data. These descriptive variables included traffic, load, speed limits, number of lanes, pavement thickness, and pavement age.

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. Complete variable definitions are given in Appendix A.

All analyses were completed using JMP Pro software (version 12.0.1, 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 30% of the observations in the data set were for Pavement Type 1. Of the 23 distress measures, only 10 remain in the final "distress-only" model. Thirteen variables were easily eliminated from consideration based on collinearity concerns and statistical insignificance ($p>5\%$). The prediction equation is as follows:

PCI = 69.5 + 31.77*ACRACKM + 0.005*ACRACKL - 1.30*TCRACKH - 0.10*TCRACKL + 8.70*LCRACKH + 6.58*LCRACKM + 4.40*LCRACKL - 0.56*PATCHES - 21.1*ACRACK - 4.41*LCRACK

The model has an RMSE of 9.432 and an R^2 of 63.5%.

We can interpret the RMSE value as indicating that approximately 95% of all recorded PCI values (for Pavement Type 1) should fall within 18.864 (2*9.432) of the PCI predicted by this model. The R^2 indicates that approximately 63.5% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

With the exception of ACRACKL, all variables in this model are statistically significant ($p<5\%$); however, removing ACRACKL for a simpler model decreases the model's fit (RMSE increases to 10.16), so the choice was made to leave this variable in the model for the improved fit to the data.

If the modeling process is initiated with the 23 distress variables as well as 9 descriptive variables, the model fit can be improved by 7%. The prediction equation becomes the following:

PCI = 74.2 - 0.238*AGE - 0.00024*ADT + 0.0017*TRUCKS - 0.0766*TCRACK + 0.011*LCRACKL - 0.0087*LCRACK - 0.0280*DCRACKH - 0.0722*JTSPALLH - 0.0547*JTSPALLM + 0.0011*PATCHAB - 0.0016*PATCHAG - 0.2389*PATCHES

The model has an RMSE of 8.774 and an R^2 of 68.7%.

Approximately 95% of all recorded PCI values (for Pavement Type 1) should fall within 17.548 (2*8.774) of the PCI predicted by this model. The R^2 indicates that approximately 68.7% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

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

Figure 1 shows scatter plots that demonstrate how various distress measures and descriptive variables interact with PCI for Pavement Type 1. Table 1 offers summary statistics for the measures and variables for Pavement Type 1.

Figure 1. Scatter plots showing how 27 measures and variables interact with PCI for Pavement Type 1

Table 1. Summary statistics for Pavement Type 1

Distress Measure	Minimum	Q1	Median	Mean	Q3	Maximum
ACRACKH	0	0	0	0.083	0	35
ACRACKM	0	0	0	10.562	0	1441
ACRACKL	0	0	0	38.252	0	8454
ACRACK	0	0	0	16.042	0	2162
TCRACKH	0	0	0	1.786	1	401
TCRACKM	0	0	2	8.858	8	411
TCRACKL	0	0	3	27.160	19	1371
TCRACK	0	2	9	44.177	42	1996
LCRACKH	0	0	0	13.078	0	1119
LCRACKM	0	0	5	50.336	40	1816
LCRACKL	0	0	23	191.570	133	23068
LCRACK	0	5	56	293.374	275	25303
LCRACKWH	0	0	0	2.184	0	253
LCRACKWM	0	0	0	25.715	9	1490
LCRACKWL	0	0	0	137.784	57	23068
LCRACKW	0	0	6	180.813	103	25303
DCRACKH	0	0	0	18.272	15.25	232
DCRACKM	0	0	2	10.958	14	138
JTSPALLH	0	0	0	5.199	4	154
JTSPALLM	0	0	1	5.287	5	140
PATCHAB	0	0	0	204.272	52	9146
PATCHAG	0	0	0	370.735	277	20647
PATCHES	0	0	0.7	4.503	4.575	72
Descriptive Variables						
AGEIF	0	12	21	25.349	36	87
SPEED	20	55	55	55.243	65	70
PAVTHICK	3	9	10	10.001	10	19
ADT	10	4080	7950	11543.684	15600	90400
TRUCKS	66	400	767.5	1541.848	1670.5	11498
KIPSANN	468	66837.5	164950	412534.144	397080	3672775
KIPSRES	53610	835906.5	3011537.5	4142451.563	7943191	10557275
KIPSCON	166816	2607174	4707553	10285465.455	8955503	94585012
Summary						
PCI_2	7	53	66	62.179775281	74	94

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 48% of the observations in the data set were for Pavement Type 3. Of the 23 distress variables considered as potential inputs in the model for PCI, only seven variables were eliminated based on statistical significance considerations. The prediction equation is as follows:

PCI = 79.3 + 17.58*ACRACKH + 12.55*ACRACKM + 0.0004*ACRACKL - 8.37*ACRACK + 0.311*TCRACKM + 0.194*TCRACKL - 0.222*TCRACK + 11.174*LCRACKH + 8.39*LCRACKM + 5.59*LCRACKL - 5.60*LCRACK - 0.19*DCRACKH - 0.572*JTSPALLH - 0.393*JTSPALLM + 0.0018*PATCHAG - 0.48*PATCHES

The model has an RMSE of 9.845 and an R^2 of 66.9%.

Approximately 95% of all recorded PCI values (for Pavement Type 3) should fall within 19.69 (2*9.845) of the PCI predicted by this model. The R^2 indicates that approximately 66.9% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

With the exceptions of ACRACKL (p=0.3372) and JTSPALLM (p=0.0746), all variables in this model are statistically significant (p<5%); however, removing either (or both) of these variables to simplify the model decreases the model's fit (RMSE rises above 10), so both variables were left in the model in the interest better fitting the model to the data.

If all descriptive variables and distress variables are initially considered, the model fit can be improved by 16%. The prediction equation becomes the following:

PCI = 61.5 - 0.28*AGE + 0.377*SPEED + 0.212*PAVTHICK - 0.00737*TRUCKS + 0.0000295*KIPSANN - 0.00000033*KIPSRES + 7.334*ACRACKH + 5.89*ACRACKM - 0.000037*ACRACKL - 3.928*ACRACK + 4.72*TCRACKH + 3.848*TCRACKM + 2.55*TCRACKL - 2.576*TCRACK + 6.48*LCRACKH + 4.84*LCRACKM + 3.225*LCRACKL - 3.23*LCRACK - 0.212*DCRACKH - 0.1459*DCRACKM - 0.542*JTSPALLH + 0.0008*PATCHAG - 0.23*PATCHES

The model has an RMSE of 8.23 and an R^2 of 76%.

Approximately 95% of all recorded PCI values (for Pavement Type 3) should fall within 16.46 (2*8.23) of the PCI predicted by this model. The R^2 indicates that approximately 76% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

Of the 23 variables in this model, 9 are not statistically significant (p>5%); however, lingering collinearity among some of these variables may explain their apparent statistical insignificance. Another consideration is model fit. Removing any of these variables (even the statistically insignificant ones) decreases the model's fit to the data (and increases the RMSE). If a simpler model (fewer input variables) is a main goal, then statistically insignificant variables could be removed using various model selection techniques at the expense of an increased RMSE.

Figure 2 shows scatter plots that demonstrate how various distress measures and descriptive variables interact with PCI for Pavement Type 3. Table 2 offers summary statistics for the measures and variables for Pavement Type 3.

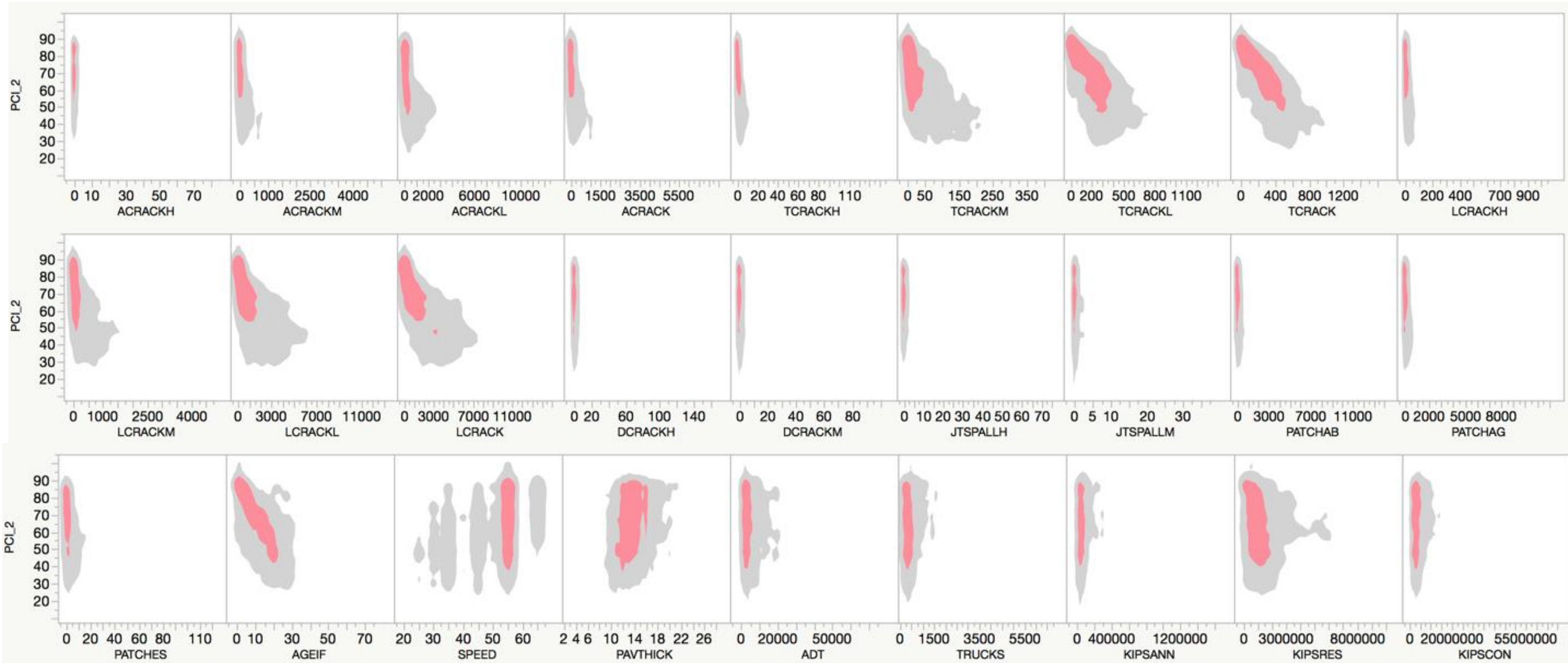

Figure 2. Scatter plots showing how 27 measures and variables interact with PCI for Pavement Type 3

Table 2. Summary statistics for Pavement Type 3

Distress Measure	Minimum	Q1	Median	Mean	Q3	Maximum
ACRACKH	0	0	0	0.611	0	87
ACRACKM	0	0	0	142.424	84	5231
ACRACKL	0	0	32	548.646	457	13057
ACRACK	0	0	0	214.998	126	7917
TCRACKH	0	0	0	2.293	2	150
TCRACKM	0	2	21	48.659	71	446
TCRACKL	0	77	217	238.094	346	1450
TCRACK	0	97	284	315.877	467.75	1756
LCRACKH	0	0	0	16.584	11	1135
LCRACKM	0	2	72	296.299	405	5127
LCRACKL	0	114.25	843.5	1542.732	2352.25	13729
LCRACK	0	172	1270.5	2020.541	3209	15838
LCRACKWH	0	0	0	7.780	5	1489
LCRACKWM	0	0	43	204.666	238	4805
LCRACKWL	0	0	16	800.832	965.25	13729
LCRACKW	0	24	278	1123.577	1514.75	15838
DCRACKH	0	0	0	1.772	0	172
DCRACKM	0	0	0	1.148	0	109
JTSPALLH	0	0	0	0.441	0	79
JTSPALLM	0	0	0	0.442	0	41
PATCHAB	0	0	0	152.126	38	14672
PATCHAG	0	0	0	175.753	104	12481
PATCHES	0	0	0.7	4.003	3.1	124.4
Descriptive Variables						
AGEIF	0	5	13	13.671	20	84
SPEED	20	45	55	51.341	55	70
PAVTHICK	3	12	14	13.878	15	29
ADT	380	2120	3470	5911.157	7300	82800
TRUCKS	45	248	385	532.384	617	7187
KIPSANN	5740	34177.5	54445	86587.846	92142.5	1728320
KIPSRES	6050	646069	1113795	1528863.593	1848517	10676391
KIPSCON	256205	2381347.5	3560825	4616582.191	5549852	72547153
Summary						
PCI_2	11	52	66	64.907249467	79	100

MODELING PCI FOR PAVEMENT TYPE 4

Pavement Type 4 is asphalt cement (AC). About 12% of the observations in the data set were for Pavement Type 4. Of the 23 distress variables considered as potential inputs in the model for PCI, 12 variables were eliminated based on statistical significance considerations. The prediction equation is as follows:

PCI = 81.1 + 13.4*ACRACKH + 9.86*ACRACKM + 0.00012*ACRACKL - 0.735*TCRACKH - 0.0416*TCRACKL + 11.73*LCRACKH + 8.66*LCRACKM + 5.77*LCRACKL - 0.4685*PATCHES - 6.577*ACRACK - 5.775*LCRACK

The model has an RMSE of 8.357 and an R^2 of 76.8%.

Approximately 95% of all recorded PCI values (for Pavement Type 4) should fall within 16.714 (2*8.357) of the PCI predicted by this model. The R^2 indicates that approximately 76.8% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

When 14 of the 32 potential input variables considered (23 distress and 9 descriptive) drop out of the model, the RMSE shows a 20% improvement. The prediction equation becomes the following:

PCI = 60.4 - 0.55*AGE + 0.416*SPEED - 0.0003*ADT + 0.0000004*KIPSRES + 0.00035*ACRACKL - 0.004*ACRACK - 7.52*TCRACKH - 4.678*TCRACKM - 3.193*TCRACKL + 3.16*TCRACK + 6.5488*LCRACKH + 4.828*LCRACKM + 3.215*LCRACKL - 3.22*LCRACK - 3.45*DCRACKM + 7.189*JTSPALLM + 0.00117*PATCHAG - 0.4555*PATCHES

The model has an RMSE of 6.659 and an R^2 of 86%.

Approximately 95% of all recorded PCI values (for Pavement Type 4) should fall within 13.318 (2*6.659) of the PCI predicted by this model. The R^2 indicates that approximately 86% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

Of the 18 variables in this model, only 3 are statistically significant (although collinearity may be obscuring the statistical significance of some of the other variables). AGE, SPEED, and ACRACK are statistically significant ($p<5\%$). However, removing any of the statistically insignificant variables comes at the cost of reduced model fit (increased RMSE).

Figure 3 shows scatter plots that demonstrate how various distress measures and descriptive variables interact with PCI for Pavement Type 4. Table 3 offers summary statistics for the measures and variables for Pavement Type 4.

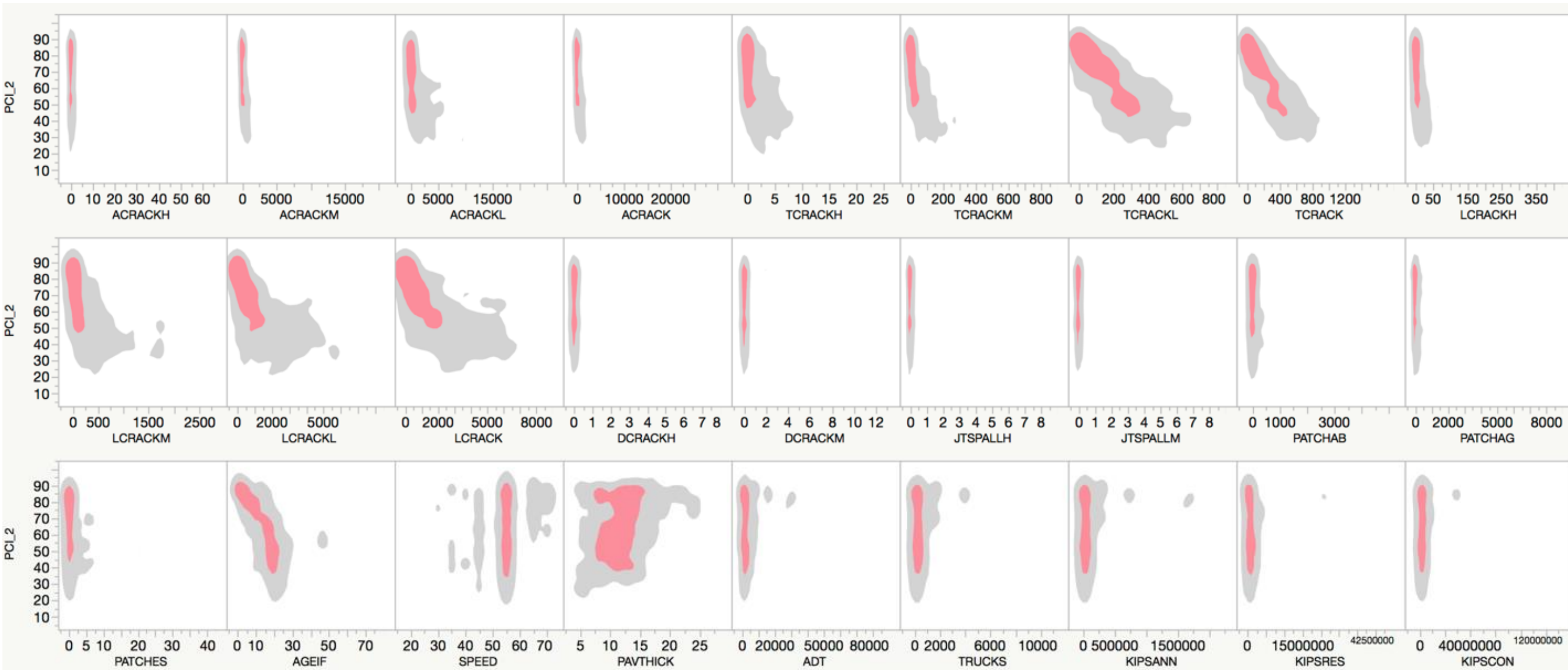

Figure 3. Scatter plots showing how 27 measures and variables interact with PCI for Pavement Type 4

Table 3. Summary statistics for Pavement Type 4

Distress Measure	Minimum	Q1	Median	Mean	Q3	Maximum
ACRACKH	0	0	0	0.470	0	69
ACRACKM	0	0	15	292.247	181	20225
ACRACKL	0	1	127	1204.459	1332.75	26111
ACRACK	0	0	23	439.470	271.5	30476
TCRACKH	0	0	0	1.855	2	27
TCRACKM	0	2	14	47.108	67	900
TCRACKL	0	39.75	176	201.429	314.25	862
TCRACK	0	45	222.5	276.013	432.5	1800
LCRACKH	0	0	0	12.545	11	436
LCRACKM	0	1.75	83.5	280.727	375.75	2925
LCRACKL	0	81.25	656.5	1276.693	2025	8569
LCRACK	0	115.5	965.5	1723.076	2766.25	9433
LCRACKWH	0	0	0	8.848	5	436
LCRACKWM	0	1	47	289.703	344.5	4345
LCRACKWL	0	0	16	736.327	942.5	5776
LCRACKW	0	24	369.5	1188.777	1618.75	9117
DCRACKH	0	0	0	0.118	0	8
DCRACKM	0	0	0	0.153	0	13
JTSPALLH	0	0	0	0.055	0	9
JTSPALLM	0	0	0	0.077	0	9
PATCHAB	0	0	0	111.626	0	5083
PATCHAG	0	0	0	71.426	0	9195
PATCHES	0	0	0	1.573	0.5	43.2
Descriptive Variables						
AGEIF	0	7	16	15.022	20	83
SPEED	20	55	55	54.589	55	70
PAVTHICK	3	9	12	11.974	14	29
ADT	370	1140	1880	4511.020	3735	90400
TRUCKS	10	150.75	232	791.282	424	11344
KIPSANN	1760	21287.5	32190	149744.087	63900	2265580
KIPSRES	8090	340746.5	621288.5	1829317.535	1120843.25	40805229
KIPSCON	304281	1055645	1591946.5	5537606.476	2606758.75	111423400
Summary						
PCI_2	9	50	66	64.91991342	82	100

MODELING PCI – ALTERNATIVE CONSIDERATIONS

While modeling PCI separately for different pavement types has some benefit in terms of model fit, modeling PCI for all pavement types combined may be a reasonable and interesting alternative. If we begin the modeling process with just the distress variables, the resulting prediction model includes 10 distress variables in the following equation:

PCI = 73.1 + 7.458*ACRACKH + 5.313*ACRACKM - 0.748*TCRACKH - 0.01124*TCRACKL + 4.98*LCRACKH + 3.742*LCRACKM + 2.49*LCRACKL - 0.515*PATCHES - 3.546*ACRACK - 2.50*LCRACK

The model has an RMSE of 12.145 and an R^2 of 44%.

Approximately 95% of all recorded PCI values (across all pavement types) should fall within 24.29 (2*12.145) of the PCI predicted by this model. The R^2 indicates that approximately 44% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

If all 32 variables (23 distress and 9 descriptive) are considered, model fit improves by 27%. The resulting prediction equation with 22 input variables is as follows:

PCI = 66.46 - 0.348*AGE + 0.2634*SPEED + 0.1215*PAVTHICK - 0.0002*ADT - 0.00000006*KIPSRES + 0.00000014*KIPSCON + 3.725*ACRACKH + 2.65*ACRACKM - 1.77*ACRACK + 0.17*TCRACKM + 0.127*TCRACKL - 0.144*TCRACK + 2.335*LCRACKH + 1.75*LCRACKM + 1.1614*LCRACKL - 1.164*LCRACK - 0.157*DCRACKM - 0.232*JTSPALLH - 0.22*JTSPALLM - 0.0004*PATCHAB - 0.00075*PATCHAG - 0.155*PATCHES

The model has an RMSE of 8.8 and an R^2 of 70%.

Approximately 95% of all recorded PCI values (across all pavement types) should fall within 17.6 (2*8.8) of the PCI predicted by this model. The R^2 indicates that approximately 70% of the observation-to-observation variability in recorded PCI values can be accounted for by this model.

The simplification of having one model that applies to all pavement types might outweigh the model fit gains of having separate models for Pavement Types 1, 3, and 4 (and no models for less represented pavement types like 2A, 2B, 3A, and 3B).

Another alternative consideration is to analyze the residuals from these models. Residual analyses indicate that polynomial forms of the AGE variable may improve the model fit. For example, adding AGE^2 and AGE^3 to the model above improves model fit by 5% (yielding a new RMSE of 8.4) and increases R^2 to 72.5%, and both polynomial forms of AGE are statistically significant ($p<5\%$) without introducing any new collinearity among input variables.

Input from the Iowa DOT may shed light on these and other considerations.

Figure 4 shows scatter plots that demonstrate how various distress measures and descriptive variables interact with PCI for all pavement types. Table 4 offers summary statistics for the measures and variables for all pavement types.

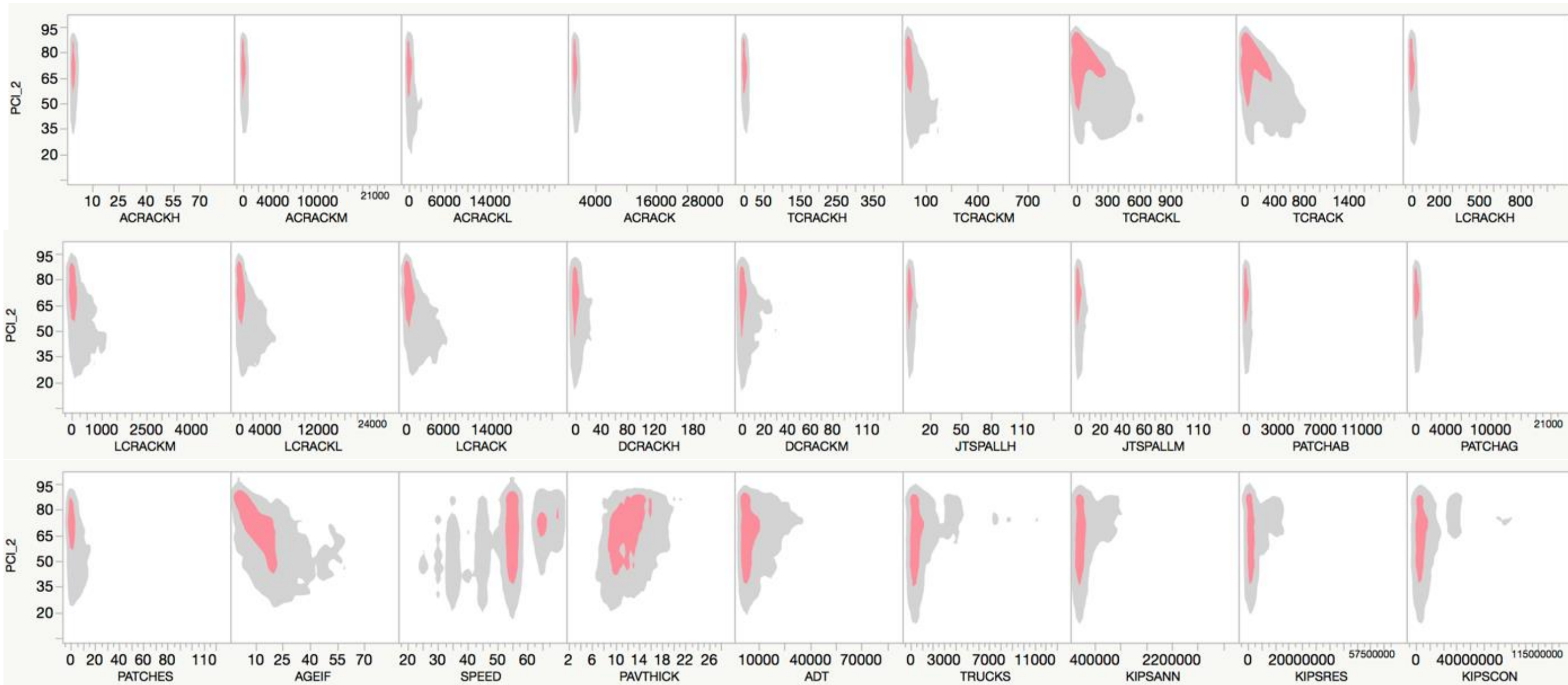

Figure 4. Scatter plots showing how 27 measures and variables interact with PCI for all pavement types

Table 4. Summary statistics for all pavement types

Distress Measure	Minimum	Q1	Median	Mean	Q3	Maximum
ACRACKH	0	0	0	0.384	0	87
ACRACKM	0	0	0	114.536	30	20225
ACRACKL	0	0	0	406.560	152	26111
ACRACK	0	0	0	172.677	45	30476
TCRACKH	0	0	0	1.948	2	401
TCRACKM	0	1	7	33.655	41	900
TCRACKL	0	3	86	159.251	266	1450
TCRACK	0	9	116	213.821	351	1996
LCRACKH	0	0	0	14.139	8	1135
LCRACKM	0	0	27	210.161	211.5	5127
LCRACKL	0	15	277	1028.513	1383	23068
LCRACK	0	33.5	423	1372.208	2030.5	25303
LCRACKWH	0	0	0	5.758	2	1489
LCRACKWM	0	0	11	150.083	125.5	4805
LCRACKWL	0	0	7	562.777	358.5	23068
LCRACKW	0	0	104	799.571	820.5	25303
DCRACKH	0	0	0	6.886	0	232
DCRACKM	0	0	0	4.188	1	138
JTSPALLH	0	0	0	1.922	0	154
JTSPALLM	0	0	0	1.961	0	140
PATCHAB	0	0	0	156.524	35	14672
PATCHAG	0	0	0	232.611	139	20647
PATCHES	0	0	0.5	3.781	3.1	124.4
Descriptive Variables						
AGEIF	0	7	14	17.269	23	87
SPEED	20	55	55	54.139	55	70
PAVTHICK	3	10	12	12.461	14	29
ADT	10	2280	4850	8861.325	11400	90400
TRUCKS	10	265	476.5	1173.424	1081.75	13577
KIPSANN	468	38322.5	73495	259963.699	235850	3672775
KIPSRES	6050	614058.25	1168366.5	2846334.201	2251214.5	55397869
KIPSCON	166816	2268359.25	3814345	8754232.384	7381336.5	111491975
Summary						
PCI_2	7	53	68	64.804291624	77	100

APPENDIX A. VARIABLE DEFINITIONS

All data are from the year 2013 and were collected and provided by the Iowa Department of Transportation (DOT).

The distress variables considered as potential input variables in this research are as follows:

- ACRACKH, ACRACKM, and ACRACKL – alligator cracking severity (high, moderate, or low, respectively) measured in square feet per mile
- TCRACKH, TCRACKM, and TCRACKL – transverse cracking severity (high, moderate, or low, respectively) measured in count per mile
- LCRACKH, LCRACKM, and LCRACKL – longitudinal cracking severity (high, moderate, or low, respectively) measured in feet per mile
- LCRACKWH, LCRACKWM, and LCRACKWL – longcrack wheelpath cracking severity (high, moderate, or low, respectively) measured in feet per mile
- DCRACKH and DCRACKM – count of joints per mile with durability cracking severity (high or moderate, respectively)
- JTSPALLH and JTSPALLM – count of joints per mile with spalling severity (high or moderate, respectively)
- PATCHAB and PATCHAG – moderate severity patch condition (bad or good, respectively) measured in square feet per mile
- PATCHES – count of patches per mile

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

- AGE – years since construction or resurfacing
- SPEED – speed limit in miles per hour
- PAVTHICK – pavement thickness in inches
- LANES – number of lanes
- ADT – average daily traffic as a count per day
- TRUCKS – number of trucks per day
- KIPSANN – annual 18 kips measured in esals
- KIPSRES – accumulated kips since resurfacing measured in kips
- KIPSCON – accumulated kips since construction measured in kips

Table 5 provides a condensed overview of the input variables chosen for each modeling scenario, along with the RMSE and R^2 for each model.

Table 5. Overview of variables and measures for each modeling scenario

Variables and Measures	PAVTYP=1		PAVTYP=3		PAVTYP=4		PAVTYP=ALL	
	Distress	**Combined**	**Distress**	**Combined**	**Distress**	**Combined**	**Distress**	**Combined**
AGEIF		X		X		X		X
SPEED				X		X		X
PAVTHICK				X				X
LANES								
ADT		X				X		X
TRUCKS		X		X				
KIPSANN				X				
KIPSRES				X		X		X
KIPSCON								X
ACRACKH			X	X	X		X	X
ACRACKM	X		X	X	X		X	X
ACRACKL	X		X	X	X	X		
ACRACK	X		X	X	X	X	X	X
TCRACKH	X			X	X	X	X	
TCRACKM			X	X		X		X
TCRACKL	X		X	X	X	X	X	X
TCRACK		X	X	X		X		X
LCRACKH	X		X	X	X	X	X	X
LCRACKM	X		X	X	X	X	X	X
LCRACKL	X	X	X	X	X	X	X	X
LCRACK	X	X	X	X	X	X	X	X
LCRACKWH								
LCRACKWM								
LCRACKWL								
LCRACKW								
DCRACKH		X	X	X				
DCRACKM				X		X		X
JTSPALLH		X	X	X				X
JTSPALLM		X	X			X		X
PATCHAB		X						X
PATCHAG		X	X	X		X		X
PATCHES	X	X	X	X	X	X	X	X
R^2	**0.634679**	**0.687179**	**0.669072**	**0.761375**	**0.767951**	**0.862331**	**0.440022**	**0.698637**
RMSE	**9.432009**	**8.774102**	**9.84496**	**8.229689**	**8.357131**	**6.659208**	**12.14548**	**8.802193**