

**PAVEMENT DETERIORATION PREDICTION MODEL AND PROJECT  
SELECTION FOR KENTUCKY HIGHWAYS**

**SUBMISSION DATE**

August 13, 2014

**WORD COUNT:**

4,883

**AUTHORS CONTACTS**

Guangyang Xu

Lihui Bai

Zihui Sun

Speed School of Engineering

University of Louisville

Tracy Nowaczyk

Chad Shive

Jon Wilcoxson

---

Operations and Pavement Management Branch

Kentucky Transportation Cabinet

## **ABSTRACT**

Pavement deterioration is an important factor in evaluating and prioritizing pavement management and preservation (PMP) projects. The primary goal of this paper is to provide quality predictive functions from multiple linear regression (MLR) models that can be adopted by Kentucky Transportation Cabinet (KYTC). Furthermore, the paper proposes to use a decision analysis procedure, i.e., an analytic hierarchy process (AHP), in developing a composite pavement distress index for KYTC to prioritize and select PMP projects. Such a prioritization of candidate PMP projects is based on 11 different distress indices. Numerical results show that the MLR models provide relatively high R squared values of approximately 0.8. In addition, preliminary study shows that the proposed AHP-based project selection method overcomes the drawback of KYTC's current rating and selection system for overemphasizing the international roughness index (IRI) among all distress indices.

## **1. INTRODUCTION**

### **1.1. Background and Problem Statement**

Pavement management and preservation (PMP) is important to our nation's highway infrastructure development. Currently, the Operations & Pavement Management Branch of Kentucky Transportation Cabinet (KYTC) performs annual pavement condition evaluations and ride quality measurements. Then, pavement rehabilitation and resurfacing treatments and priority rankings are recommended based on pavement conditions, traffic considerations and district recommendations. Recently, KYTC has increasingly realized the need for an integrated and rigorous framework to prioritize alternative PMP projects to assist in final selection of projects to be undertaken in each budget cycle. This framework first requires the knowledge of predicted pavement performance for the near future over the given time horizon. The pavement performance herein consists of as many as 11 individual distress indices (e.g., wheel path cracking, out of section, raveling), each to be predicted for prioritization purpose. Once these 11 individual distress indices are predicted, a formal procedure is needed to create a composite or overall condition index. Hence, each project or road segment is associated with a single composite condition index; and it is based on these composite indices that the integrated project selection approach will prioritize alternative projects.

One focus of the paper is the modeling of pavement deterioration for KY interstate and parkway system by the means of statistical methods. First, the road condition data for KY interstate and parkways, focusing on asphalt concrete (AC) roadways during the past ten years will be analyzed. Second, the organized data will then be used to develop predictive models for studying the pavement deterioration for KY interstate and parkways. The predictive model uses inputs such as cracking index, pavement age and average daily traffic (ADT) in previous years for a certain road segment to predict the pavement performance (e.g., cracking index) in the future years. Particularly, multiple linear regressions (MLR) models are developed.

Another focus of the paper is the development of an objective priority rating method that can be integrated with the above-mentioned distress index prediction models. The decision analysis procedure of analytic hierarchy process (AHP) is not only a rather rigorous mathematical framework but also widely adopted by decision makers in real world. Organizations such as U.S. DOE, U.S. DOD, NASA, Xerox and IBM have all successfully used AHP in their efforts of making choice, prioritization, evaluation, resource allocation, among others. In this project,

Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

based on interviews from a KYTC expert panel, weights for 11 individual distress indices are calculated and validated. Then, using the predicted individual distress indices and weights calculated by the proposed AHP method, each project receives a single overall priority score, which allows KYTC to finally select PMP projects to be undertaken in the planning period.

## 1.2. Literature Review

In recent years, predicting the expected pavement deterioration has been the focus of many works (Attoh-Okine, 1994) using traffic and time-related models, interactive time, traffic, or distress models. To date, approaches used in forecasting the pavement condition have included: regression models, artificial neural network, empirical model, mechanistic models and deterministic and probabilistic models. Within these approaches, regression analysis is used by researchers as a traditional method to predict pavement deterioration rate (see, e.g., Carey and Irick, 1960). Particularly, Isa and Hwa (2005) propose a study which established a simple, practical pavement performance model for network level of the Malaysian Federal road. They also conducted the statistical analysis by the means of multiple linear regressions to test and examine the data as well as to develop the model. In their model, relevant variables such as pavement condition, pavement strength, traffic loading and pavement age are used as input data to predict the target variable rutting depth. Similarly, Kim and Kim (2006) develop the asphalt pavement performance prediction models for the state highways and the interstate highways with the applications of simple and multiple regression analysis methods. Kim and Kim's model uses the pavement condition evaluation system (PACES) data rating. They develop three different models using one, two and three input variables (AADT, Pavement Age, the interaction) in applying linear and multiple regression analysis to predict the value of each road segment. In conclusion, AADT, pavement age, and the interaction between AADT and age have significant impact on the PACES rating.

More recently, the development of artificial intelligence and machine learning has spurred new methods such as artificial neuron networks (ANN). For example, Yang and Lu (2003) develop a pavement performance model applying neural network algorithm for the Florida Department of Transportation (FDOT) pavement management system. In their model, a three-layered neural network with one output neuron is chosen to be the architecture, back-propagation (BP) method applied to be the learning method, at last the sigmoid function was employed as the neuron activation function. Similarly, Lou and Gunaratne (2001) also developed multiyear back-propagation neural network (BPNN) models for Florida's highway network to forecast accurately the short-term time variation of cracking index (CI). On the other hand, Huang and Moore (1997) tasked by the Kansas Department of Transportation, use multiple-linear regression and two ANN structures to predict the probability of roughness deterioration level for asphalt pavements. In their ANN model, the input layer consists of 17 independent variables such as cumulative traffic expressed in 80kN equivalent single axle loads (ESAL), layer thickness and back calculated modulus values, and soil support values. The multiple linear regression models in Huang and Moore (1997) experience success rates ranging from 70% to 90%. Between the two ANN models, the average success rate ranges from 70% to 93%, slightly higher than that of the two regression models.

In contrast to the literature on pavement deterioration modeling, studies on prioritization and selection of pavement-related projects are rather scant. Fwa and Chan (1993) use two different priority-setting schemes to test the feasibility of using neural network models. One is a linear

Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

rating function, and the other is a nonlinear rating function. In both functions, priority rating is expressed as a function of normalized scores for factors such as highway functional class, skid resistance, width of crack, length of crack, pavement serviceability and rut depth. Fwa and Chan (1993) show that the ANN model can accurately predict the pavement priority rating governed by both the linear and nonlinear equations. Finally, in current practices at KYTC (KYTC, 2006), a rating function similar to the above linear rating function is used. Details of the KYTC rating function are discussed in Section 3. Using decision analysis, the current paper proposes a more objective procedure using AHP to determine the coefficients or weights used in the linear or nonlinear equations for priority rating.

The remainder of the paper is organized as follows. Section 2 introduces the data processing procedures. Section 3 presents the design of multiple linear regression models and the AHP method for priority rating. Section 4 discusses the results for the MLR prediction models. In addition, the results for the current KYTC priority rating system and the proposed AHP-based rating method are also discussed. Finally, Section 5 concludes the paper with future directions.

## **2. DATA PROCESSING**

### **2.1. Data Collection**

Each spring the Operations and Pavement Management Branch of KYTC performs pavement condition evaluations of all interstates and parkways. These evaluations are used to document roadway deterioration, recommend pavement rehabilitation treatments, and prioritize projects. The road condition evaluation data maintained by KYTC consists of asphalt concrete (AC), Portland cement concrete and composite pavements varying in condition, age, and performance. In general, pavement distresses are classified and rated according to type, severity, and extent. Ideally, the extent of each severity of each type of distress would be measured and recorded using finite values.

According to the KYTC pavement distress identification manual (2009), a demerit point system (0-9) is used to measure pavement distress. Pavements with highest demerit point, i.e., a score of 9, exhibit greatest deterioration, while those with lowest demerit point, i.e., a score of 0, exhibit none deterioration. There are five different types of flexible cracking: wheel path cracking (extent and severity), raveling (extent and severity), other cracking (extent and severity), out of section (extent and severity) and appearance of pavement. Thus, in the present paper, nine variables are considered by all regression models.

### **2.2. Data Processing and Analysis**

Data preprocessing consists of removing incomplete data entries, rounding on the road segment mileage marks, removing incompatible data entries and calculating pavement age according to most recent maintenance year.

Within the data set, each data sample (row) has several physical attributes, such as year of evaluation, route ID, lane direction, evaluation start and end points, pavement type. In order to use the cracking indices of the current year to predict cracking index of the next year, it is necessary to have at least cracking indices on two consecutive years for each road segment. However, it remains a challenge to identify the “same” road segment across different years in the

Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

original database. It is discovered that the start and end points of road segments varies slightly from year to another. For example, in 2004, Route BG9002 has start and end points at 9.523 and 16.54 miles, respectively, while in 2003 the same road segment has start and end points as 10.172 and 16.54 miles, respectively. In order to mitigate this type of discrepancy and reduce the number of road segments with only one year of data (thus including more sample data with multiple years of performance in our model), all start\_point and end\_point are rounded to nearest integers. This has significantly increased the sample size by 100 data points, with final sample size being 4,586.

Another great challenge in data preprocessing is to calculate the age of each pavement segment. KYTC has kept information (e.g., year of project, type of rehabilitation) on all construction/maintenance projects done to each interstate and highway road segments in the database. It is desirable to calculate the age of each road segment as the difference between the current year and the most recent maintenance year. However, in some rows, the construction information was recorded incorrectly for unknown reason. Thus, the following two rules are used to decide if the given data is incompatible with respect to age calculation:

- 1) If a road segment is recorded with rehabilitation in the evaluation year but its corresponding cracking indices are not significantly reduced in the following years, the sample data point is regarded as incompatible.
- 2) If a road segment is not recorded with rehabilitation in the evaluation year but its corresponding cracking indices are significantly reduced in the following years, the sample data point is regarded as incompatible.

After incompatible segments are identified, their ages are calculated based on whether or not there is a significant decrease in all cracking indices. For example, if all cracking indices of one road segment in a specific year are all reduced from a relatively high value to 0, then it is considered there is a major rehabilitation done in this year and the pavement age for the current year is reset to zero. On the contrary, there is also a large amount of road segments having a reduced cracking index for the next year even without rehabilitation. All these type of incompatible data points are removed. After this adjustments to the incompatible data points, the sample size is reduces to 3,124, of which only 1,289 are asphalt pavement type.

In the final data set with 1,289 samples, the following 12 road condition attributes (listed in Table 1) are the input variables in all the regression models. For each regression model, the single target variable would represent one of the nine distress indices (i.e., WPC\_EXT) for next year. Thus, the dimensionality of the model is 13 (nine cracking indices, pavement age, ADT, IRI and the target variable WPC\_EXT<sub>t+1</sub>). The final 1,289 data samples are used to develop the MLR models.

Table 1. Road Condition Attributes

WPC_EXT = extent of wheel path cracking	WPC_SEV = severity of wheel path cracking
RF_EXT = extent of raveling	RF_SEV = severity of raveling
OC_EXT = extent of other cracking	OC_SEV = severity of other cracking
OS_EXT = extent of out of section	OS_SEV = severity of out of section
APPEAR = appearance	IRI = international roughness index
ADT = average daily traffic	AGE = pavement age

JS = joint separation	
-----------------------	--

### 3. METHODOLOGY

#### 3.1. Deterioration Predictive Models using Multiple Linear Regressions

To accommodate the size and the randomness of the available data, SAS Enterprise Miner 12.1 (SAS Institute, 1989) is chosen as a tool to assist developing and validation of the models. In data mining, a strategy for assessing the quality of model generalization is to partition the data source. A portion of the data, called the training data set, is used for preliminary model fitting. The rest is reserved for empirical validation and is often split into two parts: validation data and test data. The validation data set is used to prevent a modeling node from overfitting the training data and to compare models. The test data set is used for a final assessment of the model. For this paper, 50% of the data is allocated to training, 25% to validation and the rest 25% to testing.

In scientific as well as practical studies, the change of data scale is often done through variable transformation. In SAS Enterprise Miner, the Transform Variables node is often used to make variables better suited for logistic regression models and neural networks. In our data sets, the ADT is originally in the range of thousands to hundreds of thousands, which is a vastly larger range than that of the target variable. Therefore, the ADT variable is transformed using LG10 function. After this transformation, the new range of LG10\_ADT is from 3.62 to 5.17. Finally, the stepwise regression is chosen as the selection model.

#### 3.2. Pavement Projects Selection using Decision Analysis

Once the individual distress indices are predicted, they need to be transformed to a single measure representing the overall condition of the pavement for project prioritization. The current rating system for interstate and parkways at KYTC simply takes the scale points for all nine distress indices along with the scale point for international roughness index (IRI) and adds them together. In other words, each road to be considered for treatment receives a total score calculated as follows:

$$\begin{aligned} \text{Total Score} = & WPC\_EXT + WPC\_SEV + RF\_EXT + RF\_SEV \\ & + OC\_ECT + OC\_SEV + OS\_EXT \\ & + OS\_SEV + APPEAR + JS + SIRI, \end{aligned}$$

where SIRI represents the scaled point for IRI based a pre-defined lookup table converting the original IRI to a value between 0 and 38.

However, the experts of KYTC have concerns that the IRI receives too much weight in the overall score while recent studies suggest that the impact of roughness index on pavement life may not be significant. To this end, this paper proposes a new and rather objective method of reconciling

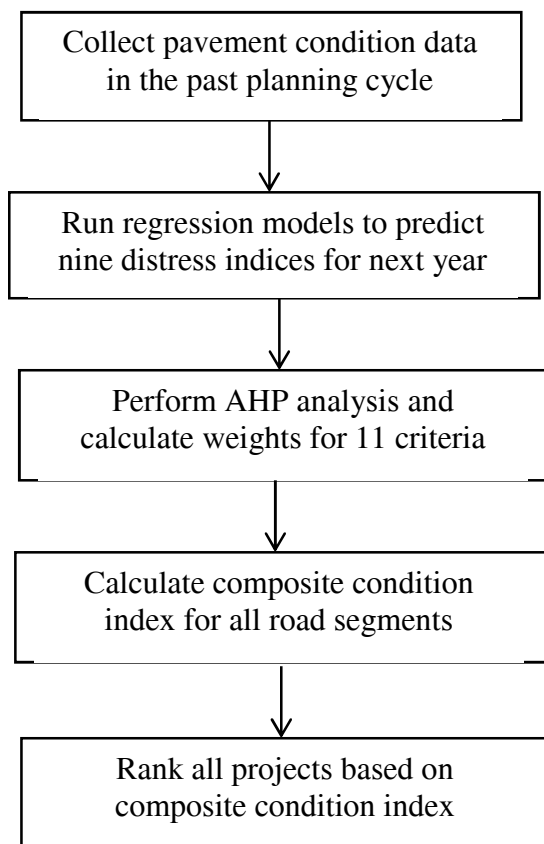


Figure 1. The Integrated Approach

Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

various indices using AHP. The AHP is a structured technique for organizing and analyzing complex decisions, based on mathematics and psychology. In practice, AHP has been used by companies and organizations including Intel, Apple, NASA and Xerox (Forman and Gass, 2001) to make decisions on choice, prioritization, resource allocation, among others.

The goal for this new method is to decide the weights for WPC\_EXT, WPC\_SEV, RF\_EXT, RF\_SEV, OC\_EXT, OC\_SEV, OS\_EXT, OS\_SEV, APPEAR, JS and IRI in prioritizing PMP projects. Thus, in the AHP there are 11 criteria and 55 pairwise comparisons. The experts are interviewed for relative importance ranging from 1 to 9 between each pairs of criteria. For example, when rating the relative importance of WPC\_EXT over RF\_EXT, expert would give a score of “1” indicating equal importance, or “9” indicating extreme importance, or “5” indicating strong importance. Then the AHP method is applied for developing a composite distress index consisting of nine individual indices and a roughness index.

Figure 1 illustrates the integrated approach for pavement projects selection using AHP and deterioration prediction models.

## **4. PRELIMINARY RESULTS**

### **4.1. Deterioration Prediction Models**

For the deterioration prediction model, we have created four different scenarios described below. These scenarios are different in that they use different subsets of the entire 12 input variables or factors when predicting each distress index. For example, S1 uses all 12 input variables in the regression models while S3 uses only one input variable. Note that the final regression equations for S1 and S2 may only involve some of the designated input variables because others are statistically insignificant.

Scenario 1 (S1). There are a total of 12 input variables, i.e., ADT, age, IRI, APPEAR, WPC\_EXT, WPC\_SEV, RF\_EXT, RF\_SEV, OC\_EXT, OC\_SEV, OS\_EXT, and OS\_SEV.

Scenario 2 (S2). A subset of the entire 12 input variables is used based on recommendations from KYTC experts. For each distress index to be predicted, KYTC expert would recommend a set of input variables that they think would affect the deterioration of the concerned index. Compared to Scenario 1, this scenario intends to have simplified model with fewer input variables. For each distress index, the following section will identify the select input variables suggested by KYTC experts. All models derived under Scenario 2 are referred to as “Select” models subsequently.

Scenario 3 (S3). Only use the target variable from the previous year as input variable to predict this variable in next year. For example, under this scenario, one would predict the severity of “Other Cracking” (or OC\_SEV) for next year using OC\_SEV for this year as the only input variable.

Scenario 4 (S4). Only use the pavement age as input variable to predict any distress index for next year

Note that among the four experimental scenarios, scenario 1 is intended to be inclusive and comprehensive, but may produce a complex prediction model that is impractical to implement in

Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

real life. Scenarios 3 and 4 are intended to be simple, but may produce high prediction errors. Scenario 2 is a compromise case that is likely to produce fairly accurate prediction using reasonably simple models.

Furthermore, in the first two scenarios linear, 2<sup>nd</sup> order polynomial and 3<sup>rd</sup> order polynomial MLR models are applied. For the other two scenarios, only the 3<sup>rd</sup> order polynomial model due to the fact that there is only one input variable is applied. The average squared error (ASE) of training, validation and testing data sets, and the R squared value are used in evaluating the results of developed models.

The results are compared in Table 2 (see details in Xu et al., 2014). In particular, models “Polynomial 3”, “Polynomial 2” and “Linear” all belong to Scenario 1 using all 12 input variables. Models “Polynomial SLCT 3” and “Polynomial SLCT 2” and “Linear SLCT” all belong to Scenario 2, in which experts from KYTC has identified Age, ADT, WPC\_EXT, and WPC\_SEV to be the four input variables. Finally, “WPC\_EXT 3” and “Age\_3” represent Scenarios 3 and 4, respectively.

It can be observed that except the model with only pavement age as the input variable, the other seven models can achieve promising R<sup>2</sup> values and relatively low ASE. For example the cubic polynomial model with all 12 input variables can achieve the highest R<sup>2</sup> (0.8848) and the smallest training ASE (0.9582) among all eight models. Even the linear model with four selected (SLCT) input variables can achieve an R<sup>2</sup> of 0.8710 and the average ASE of 1.1037. The regression function of each model can be found below Table 2.

Table 2. Linear Regression Models for WPC\_EXT

Scenarios	Model	ASE Training	ASE Validation	ASE Testing	ASE Average	R Square
S 1	Polynomial 3	0.9582	1.3173	0.9147	1.0634	0.8848
	Polynomial 2	0.9911	1.2983	0.8888	1.0594	0.8809
	Linear	1.0388	1.2964	0.9214	1.0855	0.8751
S 2	Polynomial SLCT 3	0.9803	1.2561	0.8956	1.0440	0.8821
	Polynomial SLCT 2	1.0335	1.2919	0.8860	1.0705	0.8757
	Linear SLCT	1.0726	1.3123	0.9262	1.1037	0.8710
S 3	WPC_EXT 3	1.0701	1.3420	0.8998	1.1039	0.8713
S 4	Age 3	4.0492	3.9725	4.1682	4.0633	0.5132

Polynomial3:

$$x_1: \text{apperarence}, x_2: \text{age}, x_3: \text{WPC\_EXT}, x_4: \text{LG10(ADT)}, x_5: \text{OC\_SEV}, x_6: \text{R\_F\_EXT}, x_7: \text{R\_F\_SEV}$$

$$\text{WPC\_EXT}_{(t+1)} = 0.2113 + 0.6897x_1 + 0.0585x_2 + 0.9701x_3 - 0.2445x_1^2 + 0.1476x_4x_5 - 0.1476x_5x_3 + 0.0299x_1x_5x_6 - 0.00353x_2x_4x_7 + 0.0103x_5x_3^2$$

Polynomial 2:  $x_1: \text{apperarence}, x_2: \text{WPC\_EXT}, x_3: \text{LG10(ADT)}, x_4: \text{OC\_SEV}$

$$\text{WPC\_EXT}_{(t+1)} = 0.3854 + 0.7615x_1 + 0.9275x_2 - 0.2136x_1^2 + 0.1068x_3x_4 - 0.0624x_3x_2$$

Linear:  $x_1: \text{apperarence}, x_2: \text{age}, x_3: \text{OC\_SEV}, x_4: \text{WPC\_EXT}$

$$\text{WPC\_EXT}_{(t+1)} = 0.4636 + 0.3265x_1 + 0.0302x_2 + 0.1863x_3 + 0.8212x_4$$



Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

WPC\_EXT 3:  $x_1$ : WPC\_EXT

$$WPC\_EXT_{(t+1)} = 0.5323 + 1.46x_1 - 0.1083x_1^2 + 0.0058x_1^3$$

Age 3:  $x_1$ : age

$$WPC\_EXT_{(t+1)} = -0.3547 + 0.6292x_1 - 0.0146x_1^2$$

Polynomial SLCT 3:  $x_1$ : age,  $x_2$ : WPC\_EXT,  $x_3$ : LG10(ADT),  $x_4$ : WPC\_SEV

$$WPC\_EXT_{(t+1)} = -0.0198 + 0.2691x_1 + 0.3664x_2 + 0.1693x_3x_2 - 0.0263x_2^2 + 0.000298x_1^3 - 0.00453x_1^2x_3 + 0.000474x_1^2x_4$$

Polynomial SLCT 2:  $x_1$ : age,  $x_2$ : WPC\_EXT

$$WPC\_EXT_{(t+1)} = 0.2604 + 0.1039x_1 + 1.1009x_2 - 0.00276x_1^2 - 0.0259x_2^2$$

Linear SLCT:  $x_1$ : age,  $x_2$ : WPC\_EXT,  $x_3$ : WPC\_SEV

$$WPC\_EXT_{(t+1)} = 0.493 + 0.0449x_1 + 0.8612x_2 + 0.1x_3$$

From the above eight regression functions, the linear model with selected four variables (among which three are eventually chosen in the regression model) can achieve relatively low ASE and high  $R^2$ , also the regression function is relatively simple. Thus, it is finally recommended for adoption by KYTC.

Similar regression analysis, model validation and benchmarking are conducted for all eight remaining indices. The  $R^2$  values of all the MLR models are larger than 0.8 except OS\_SEV. Meanwhile, the average ASE from the MLR models is fairly small ranging from 0.1 to 0.5. Below is the summary of final recommend regression models.

- $WPC\_EXT_{(t+1)} = 0.493 + 0.0449x_1 + 0.8612x_2 + 0.1x_3$ ,

where  $x_1$ : age,  $x_2$ : WPC\_EXT,  $x_3$ : WPC\_SEV

- $WPC\_SEV_{(t+1)} = 0.4495 + 0.034x_1 + 0.2705x_2 + 0.0883x_3 + 0.807x_4$

where  $x_1$ : age,  $x_2$ : OS\_SEV,  $x_3$ : WPC\_EXT,  $x_4$ : WPC\_SEV

- $RF\_EXT_{(t+1)} = 0.3436 + 0.0257x_1 + 0.7064x_2 + 0.1216x_3 + 0.0358x_4$ ,

where  $x_1$ : age,  $x_2$ : RF\_EXT,  $x_3$ : RF\_SEV,  $x_4$ : WPC\_EXT

- $RF\_SEV_{(t+1)} = 0.4331 + 0.0207x_1 + 0.2442x_2 + 0.5657x_3 + 0.0459x_4$

where  $x_1$ : age,  $x_2$ : RF\_EXT,  $x_3$ : RF\_SEV,  $x_4$ : WPC\_EXT

- $OC\_EXT_{(t+1)} = 0.3156 + 0.0178x_1 + 0.8588x_2 + 0.0575$ ,

where  $x_1$ : age,  $x_2$ : OC\_EXT,  $x_3$ : WPC\_EXT

- $OC\_SEV_{(t+1)} = 0.2037 + 0.0162x_1 + 0.1186x_2 + 0.7265x_3 + 0.0621x_4$

where  $x_1$ : age,  $x_2$ : OC\_EXT,  $x_3$ : OC\_SEV,  $x_4$ : RF\_SEV

- $OS\_EXT_{(t+1)} = -0.0751 + 0.0794x_1 + 0.9372x_2 + 0.0297x_3x_4 - 0.2177x_1x_3 + 0.051x_5x_6 - 0.0252x_3^2x_5 + 0.0857x_3^2x_1 - 0.051x_3^2x_7$ ,

where  $x_1$ : OC\_SEV,  $x_2$ : OS\_EXT,  $x_3$ : appearance,  $x_4$ : LG10\_ADT,  $x_5$ : OC\_EXT,

$x_6$ : WPC\_SEV,  $x_7$ : OS\_SEV

- $OS\_SEV_{(t+1)} = -0.00473 + 0.615x_1 + 0.1332x_2 + 0.0317x_3x_4 - 0.3157x_4x_5 + 0.0307x_1x_3 + x_4^2x_6 - 0.00103x_3^2x_4 - 0.00796x_3x_4x_6 - 0.00254x_4x_7x_8 + 0.000826x_4x_6x_7 + 0.0595x_4x_5^2 + 0.0014x_2^2x_3 - 0.00000518x_7^2x_9 + 0.0104x_5x_{10}^2 + 0.0227x_8x_{10}^2 - 0.0103x_6x_9x_{10} - 0.00585x_6x_{10}x_{11} - 0.0134x_5x_9x_{11} - 0.00615x_1x_{11}^2 + 0.0119x_6^2x_9 - 0.00182x_6^3$ ,

where  $x_1$ : OS\_SEV,  $x_2$ : WPC\_SEV,  $x_3$ : age,  $x_4$ : appearance,  $x_5$ : OC\_SEV,

$x_6$ : WPC\_EXT,  $x_7$ : CUR\_IRI,  $x_8$ : OS\_EXT,  $x_9$ : RF\_EXT,  $x_{10}$ : LG10\_ADT,  $x_{11}$ : WPC\_SEV

Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

- $APPEAR_{t+1} = 0.2191 + 0.8479x_1 + 0.00932x_2 + 0.035x_3$ ,  
where  $x_1$ : appearance,  $x_2$ : age,  $x_3$ : WPC\_EXT

#### 4.2. Pavement Projects Selection Using AHP

In order to evaluate the proposed AHP-based PMP projects rating method, a pilot study using a subset of 17 road segments is conducted to compare the recommendations from the current rating system and the proposed AHP. The randomly selected 17 road segments are from the 2010 pavement condition database and they all need some level of treatment.

After based on importance matrix provided by the panel of KYTC experts, the priority weights for WPC\_EXT, WPC\_SEV, RF\_EXT, RF\_SEV, OC\_EXT, OC\_SEV, OS\_EXT, OS\_SEV, APPEAR, JS (joint section) and (adjusted)IRI are determined to be 0.0995, 0.2423, 0.0376, 0.0646, 0.0894, 0.1710, 0.0244, 0.0521, 0.0242, 0.0745 and 0.1204, respectively. Further, the consistency index (CI) of 0.0725 and consistency ratio (CR) of 0.0482.

Table 3 displays the information for all 11 individual criteria for the 17 roads, as well as the total score obtained by KYTC's current rating system. The 17 roads are arranged in descending order with respect to the total score, i.e., the road with highest priority for treatment is listed on the top. Table 4 contains similar information as does Table 3, except the last column "total score" is calculated by the proposed new rating method.

Table 3. Ranking for 17 road segments by using current rating system

Road #	WPC_EXT	WPC_SEV	RF_EXT	RF_SEV	OC_EXT	OC_SEV	OS_EXT	OS_SEV	APPEAR	JS	IRI	Total Score
4	9	9	4	4	5	4	0.5	1	3	0	85.442	49.5
16	4	6	4	4	3	3	0	0	2	4	107.111	47
15	8	7	4	4	2	3	1.5	1.5	2.5	2	72.54	41.5
5	8	6	4	4	5	2	0	0	2	0	74.989	38
6	8	7	4	3	5	2	0	0	2	0	74.021	37
14	6	5	4	4	3	4	1.5	1	2.5	3	64.355	37
26	7	4	3	2	4	2	1	1	2	1	85.144	37
7	6	8	2	2	4	3	1	1	1.5	0	80.442	36.5
29	7	5	4	3	2	3	0	0	2	2	75.578	35
17	7	7	2	2	4	2	1	1.5	2	0	69.702	33.5
13	6	4	4	4	3	2	0	0	2	3	68.606	33
3	7	4	4	4	5	1	0	0	1.5	0	69.54	31.5
27	6	5	4	5	3	2	0.5	0.5	2	1	57.011	30
9	6	4	2	2	2	1	1	0.5	2.5	5	60.772	28
25	4	3	4	3	3	2	0	0	1	2	70.552	27

20	3	3	2	2	3	2	0	0	1.5	2	80.62	26.5
30	5	4	3	3	2	3	0.5	0.5	2	2	53.003	25

Table 4. Ranking for 17 road segments by using composite cracking distress index

Road #	WPC_EXT	WPC_SEV	RF_EXT	RF_SEV	OC_EXT	OC_SEV	OS_EXT	OS_SEV	APPEAR	JS	IRI	Total Score
4	10	10	5	5	6	5	1.5	2	4	1	85.442	0.8432
15	9	8	5	5	3	4	2.5	2.5	3.5	3	72.54	0.7128
14	7	6	5	5	4	5	2.5	2	3.5	4	64.355	0.6721
7	7	9	3	3	5	4	2	2	2.5	1	80.442	0.6521
16	5	7	5	5	4	4	1	1	3	5	107.111	0.6509
6	9	8	5	4	6	3	1	1	3	1	74.021	0.6021
17	8	8	3	3	5	3	2	2.5	3	1	69.702	0.5996
5	9	7	5	5	6	3	1	1	3	1	74.989	0.5881
29	8	6	5	4	3	4	1	1	3	3	75.578	0.5524
26	8	5	4	3	5	3	2	2	3	2	85.144	0.5464
27	7	6	5	6	4	3	1.5	1.5	3	2	57.011	0.5191
13	7	5	5	5	4	3	1	1	3	4	68.606	0.5011
30	6	5	4	4	3	4	1.5	1.5	3	3	53.003	0.4743
3	8	5	5	5	6	2	1	1	2.5	1	69.54	0.4603
9	7	5	3	3	3	2	2	1.5	3.5	6	60.772	0.4512
25	5	4	5	4	4	3	1	1	2	3	70.552	0.4099
20	4	4	3	3	4	3	1	1	2.5	3	80.62	0.3854

From Tables 3 and 4, it can be observed that the top ten road segments (road #4, #5, #6, #7, #14, #15, #16, #17, #26, and #29) are exactly the same by both methods, and the only difference is the ranking order. This difference is mainly caused by the emphasis given to IRI by the current method. For example, in Table 3 compared to road #16, road #14 is more distressed in almost all categories except for a significantly lower IRI score. As a result, it is only ranked No. 6 while #16 is ranked No. 2 overall. In contrast, the AHP-based rating method successfully addresses this overemphasis on IRI, giving road #14 a rank of No. 3 and road #16 a rank of No. 6. This indicates that AHP provides a more objective weight than the current rating system. Similarly observations can be made between roads #17 and #26, in which case road #26 have relatively high IRI thus receiving a higher ranking. Similarly, Figure 2 depicts the relationship between the priority score for a road segment and IRI, in which a higher priority score gives priority for a road to receive treatment. The figure indicates that the priority score is less influenced by IRI under the AHP-based rating system. Overall, it can be concluded that AHP-based rating method overcomes the problem overemphasizing IRI among all distress indices.

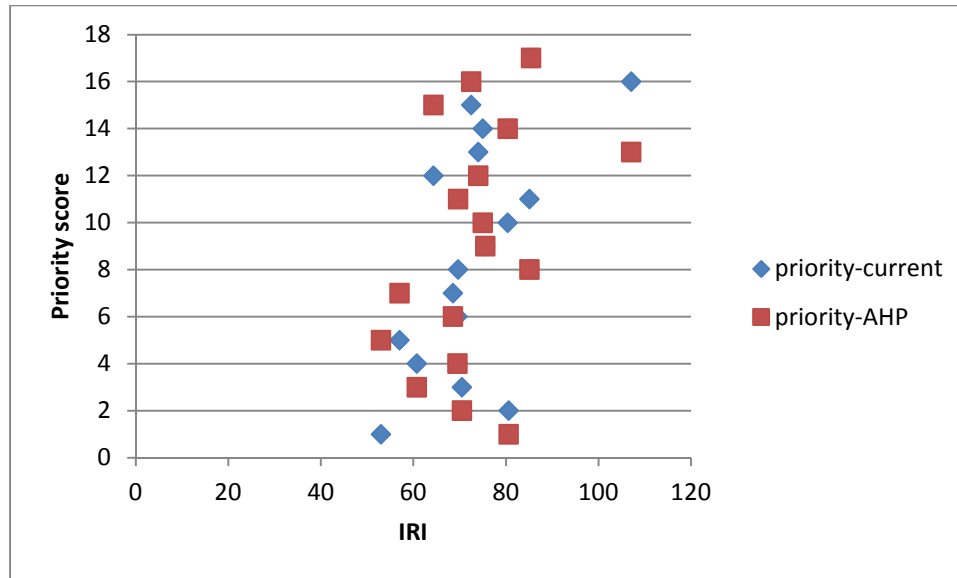


Figure 2. Priority vs. IRI

## 5. CONCLUSIONS AND FUTURE WORK

A novel integrated approach for pavement projects selection is developed consisting of two key modules: the pavement deterioration models for predicting nine individual distress indices using multiple linear regressions; and the hierarchical analytical process for developing a composite pavement condition index based on individual distress indices as well as joint separation and IRI. The developed prediction models yield an average ASE of around or less than 1.0 with R squared value around 0.8, which indicates that the regression model provide high quality prediction. The proposed AHP based project selection method is more objective and overcomes the problem of overemphasizing IRI in pavement project prioritization.

Future research includes: extending the prediction period from one year to three years in order to match with three-year budgeting cycle, study nonlinear models such as sigmoidal or power functions to predict pavement deterioration, and incorporating percentage of truck in lieu of ADT in the prediction model.

## REFERENCES

- Attoh-Okine, N. O. (1994). Predicting roughness progression in flexible pavements using artificial neural networks. In *Transportation Research Board Conference Proceedings* (Vol. 1, No. 1).
- Carey Jr, W. N., & Irick, P. E. (1960). The pavement serviceability-performance concept. *Highway Research Board Bulletin*, (250).
- Forman, E. H., & Gass, S. I. (2001). The analytic hierarchy process-an exposition. *Operations research*, 49(4), 469-486.
- Fwa, T. F., & Chan, W. T. (1993). Priority rating of highway maintenance needs by neural networks. *Journal of transportation engineering*, 119(3), 419-432.

Xu, Bai, Sun, Nowaczyk, Shive & Wilcoxson

Huang, Y., & Moore, R. K. (1997). Roughness level probability prediction using artificial neural networks. *Transportation Research Record: Journal of the Transportation Research Board*, 1592(1), 89-97.

Isa, A. H. M., Ma'soem, D. M., & Hwa, L. T. (2005). Pavement performance model for federal roads. In *Proceedings of the Eastern Asia Society for Transportation Studies* (Vol. 5, pp. 428-440).

Kim, S. H., & Kim, N. (2006). Development of performance prediction models in flexible pavement using regression analysis method. *KSCE Journal of Civil Engineering*, 10(2), 91-96.

KYTC Pavement Distress Identification Manual. (2009), <http://transportation.ky.gov/maintenance/documents/annual%20reports/pm%20field%20manual09.pdf>

Lou, Z., Gunaratne, M., Lu, J. J., & Dietrich, B. (2001). Application of neural network model to forecast short-term pavement crack condition: Florida case study. *Journal of infrastructure systems*, 7(4), 166-171.

SAS Institute Inc. (1989) *SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2*, Cary, NC: SAS Institute Inc.

Xu, G., Bai, L., & Sun, Z. (2014). Pavement Deterioration Modeling and Prediction for Kentucky Interstate and Highways. *Proceedings of the 2014 IIE Conference*.

Yang, J., Lu, J. J., & Gunaratne, M. (2003). Application of neural network models for forecasting of pavement crack index and pavement condition rating (No. Final Report).