A Method to Evaluate Performance Uniformity for Highway Pavements

Matthew S. Haubrich

A report
submitted in partial fulfillment of the
requirements of the degree of

Master of Business Analytics

Iowa State University
2017

Project Committee:
Dr. Sree Nilakanta, Chair
Dr. Hyungseok David Jeong
1 Abstract

This paper describes a proposed index which may be used to evaluate the performance uniformity of highway pavements. The methodology to calculate the Pavement Uniformity Index (PUI) relies on the availability of detailed data regarding pavement distresses. Such data are commonly collected by road agencies in the USA and around the world, however due to the density of the data usually only average values for longer segments have been considered to represent pavement performance. The PUI allows for a deeper understanding of pavement performance by looking not only at the average values but also at the uniformity or consistency of distresses within each segment. In a sense, it indicates how representative the average values are of the condition of that overall segment. The PUI is built upon the statistical measure called the Coefficient of Variation, which can readily be computed by most database software. The PUI can be used in early screening of treatment alternatives, especially in giving insight into the level of investigation which might be needed to complete a treatment design.

2 Introduction

The pavement network managed by the Iowa Department of Transportation (DOT) is about 11,092 miles of Primary routes – these are the Interstate highways, US highways, and Iowa highways that serve as the backbone of the road network in Iowa. The pavements making up this network are segmented into 3,922 unique "pavement management sections" corresponding to characteristics of the pavements, and which are the basic unit of management. A pavement management section is typically defined by construction history -- each section was built at the same time and out of the same materials, and is therefore expected to develop cracks and other problems in a relatively homogeneous fashion. In reality, however, it has been observed that segments do not always act uniformly – perhaps there were variations in the soils under the segment, or possibly construction-related differences. Regardless of the reason, we have observed that there are parts of some segments that seem to deteriorate faster (perform worse) than other parts.

Since the late 1990's, the Iowa DOT has collected detailed data on the distresses (ride, cracking, rutting etc.) across the network, with each pavement section being measured every two years. Although this data has been collected and is reported to Iowa DOT for every 1/100th of a mile (52.8 feet), the data are aggregated to the pavement management section. The detail about what is happening WITHIN a segment is not considered when the data are aggregated. I, therefore, set out to develop a Pavement Uniformity Index to evaluate how consistently each segment is performing. I have done this by evaluating the variability of the measured data within that segment using the 1/100th-mile source data.

This Pavement Uniformity Index (PUI) could be used by engineers and analysts at the Iowa DOT to improve how they manage pavements. The index can help model pavement performance and improve the prediction of pavement deterioration. It could also be used to help better identify problem areas within a pavement management section. Engineers will be able to use this index to provide additional information when selecting alternatives to maintain or repair pavements. Since many road agencies are now collecting data in a similar fashion, this index could have broad applicability beyond Iowa.
2.1 Use Case

One example of a use of this tool is in the early screening of pavement treatment alternatives. Engineers will use various data sources to help determine the optimal treatment to use on any given pavement. This process often starts with a high-level look at the overall pavement condition index (PCI), which is rated on a 0-100 scale (100 being best), along with a scan of the key distress indices (roughness, cracking, rutting, faulting), which are also each measured on a 0-100 index. In this early stage of the analysis engineers currently lack an indicator of how representative the average value is of the variability within each segment. Severe problems with a small area of pavement can be “masked” by a section average that is not so bad. If engineers are aware of the variability within the segment, they can use it to determine the level of investigation that might be required in order to develop an appropriate treatment. The PUI, along with the original indices, gives a more rounded picture of how each pavement is performing. This relationship could be represented in a screening matrix such as the one shown in Figure 1.

In this matrix, the uniformity information is combined with the condition information to give additional context to the potential complexity of the design process. The uniformity index gives the engineer a better idea of the level of investigation which may be required to determine the appropriate treatment for a given pavement section. Pavements sections with low uniformity may require additional field testing during the treatment design process, while those with good uniformity may not require that same investment in pre-design investigation.

3 Background

This project is built on many years of standardized data collection, however I only found one example of any published paper working with the detailed section data in the way I am for this index. I searched the Transportation Research Board’s extensive transportation research database, known as TRID\(^1\). This database contains over 15,000 resources on the topic of pavement management, and I was able to review only a handful of the most likely from that list.

---

\(^1\) [http://www.trb.org/InformationServices/AboutTRID.aspx](http://www.trb.org/InformationServices/AboutTRID.aspx)
3.1 Colleagues

I consulted with pavement management experts within the Iowa DOT as well as at several other states, and no one was aware of any similar work. When I described my project and the intent, there was broad agreement that the resulting index may be useful for pavement engineers in the management of pavements.

One colleague at Iowa DOT is Dr. Scott Schram. Dr. Schram currently serves as the DOT’s District 4 Engineer, but previously was the Pavement Management Engineer. He holds a PhD in Civil Engineering with a specialization in pavement management. Dr. Schram is extremely familiar with the data used in this project, as he has created a number of his own one-off analyses using the detail data. He was able to provide some very useful background information and has been a great resource as I have developed my project.

We are fortunate to have other very talented and experienced engineers on staff at Iowa DOT and they also served as resources to me as the project has developed. Together we have also identified future enhancements to this work which might add even greater value.

3.2 Data Vendors

I contacted the vendors who have performed data collection for Iowa DOT. These internationally known firms own and operate highly sophisticated equipment able to collect detailed pavement data. One of the vendors pointed me toward a paper presented at a conference in 2000\(^2\). This paper has provided some ideas for an extension of my project, particularly to attempt to use the available data to perform “dynamic segmentation” using an approach called Cumulative Difference Approach (CDA). This approach requires that all data are referenced along a linear network. Although our detailed pavement data are all referenced to the Iowa DOT’s linear network, it has not been rigorously analyzed and therefore would not be suitable for this type of analysis at this time. A future enhancement might include a proof-of-concept to leverage the information available in the spatial sequencing of data.

4 Methods

My proposed index is built from detailed data which has long been available but very rarely used. With data collected every 1/100th mile across a system of nearly 11,092 miles in 3,922 management sections, there are over 1,000,000 data points to evaluate across approximately 30 attributes. This is far too much data for anyone to reasonably use in raw form, and hence the reliance on using the section averages to represent the condition of the section. Because of the large volume of data, my development work was based on a subset of the data, selected at random. But before that, my first step was to obtain the data.

---

\(^2\) DYNAMIC SEGMENTATION OF PAVEMENT SURFACE CONDITION DATA by James Kennedy, Ahmed Shalaby, and Ray Van Cauwenbergh. This paper was submitted to the 3rd Transportation Specialty Conference of the Canadian Society for Civil Engineers, held in London, Ontario in June of 2000.
4.1 DATA

I used the data integration tool FME by Safe Software, Inc.\(^3\) to extract the data for this project from Iowa DOT’s Oracle database. Database views containing the data are open to any user inside the DOT. I used FME Desktop to develop a process to extract the data (or a sample of the data, as discussed in the following section), and write it to CSV format files. Figure 2 represents the FME process used create the data files.

*Figure 2: FME Process*

![FME Process Diagram]

Because a full cycle of pavement data covers two years of data collection, and because data from each year are stored in separate tables, I had to duplicate the process to extract two CSV files (one for 2014 and one for 2015) and subsequently merge them together. This process was done using the open-source statistical software package called R.\(^4\)

Data about the pavement management sections were sourced from the 2015 pavement management snapshot spreadsheet (PMIS2015) which is available internally within the Iowa DOT. This lists the 3,922 pavement management sections along with their calculated values and other important section information (construction details, traffic volumes, location). This is one of the most common data sources for pavement management at Iowa DOT. Key data from this spreadsheet were imported into R for use in the project.

Of the three primary classes of distress evaluated for consistency by my index, two are relatively straightforward and one required some additional calculation. They are:

- **Ride** – this is measured using a worldwide standard measure known as the International Roughness Index or IRI. It measures the amount of vertical displacement (bumps) over a given distance. The higher the value, the rougher the pavement is at any given speed. The Iowa DOT, like many US road agencies, currently measures IRI in inches/mile. A relatively smooth highway pavement will have an IRI value around 60 in/mile, and very rough pavements can have IRI values over 200.

---

\(^3\) [http://www.safe.com/]

\(^4\) [http://cran.r-project.org/; http://www.rstudio.com/]
• **Rutting** – this measures the amount of depression within the roadway wheel path relative to the rest of the pavement surface. Rutting can be a safety hazard, as rainwater can pond or run within the ruts, leading to loss of traction for vehicles. For Iowa highways, rutting is typically only a problem for pavements with an asphalt (flexible) surface. Rutting is measured by finding the average depth of the ruts within any given pavement management section. The rut depth is measured (in tenths of an inch) relative to the transverse profile of the roadway. On a highway, average rutting of more than ¼ inch can create a hazardous situation for motorists. At lower speeds rutting is somewhat less dangerous.

• **Cracking** – Cracking is a generalized term for a number of different distresses. For example, cracks that are across the driving lane are known as transverse cracks. Some cracks are measured in length, while others are measured according to their surface area. At the advice of Iowa DOT engineers, I converted all cracks to an area measurement by simply multiplying those with a length times two (so a six-foot crack would represent 12 square feet). This estimates the extent of the impact a crack has on pavement (approximately 1 foot on either side of the crack). This value was then added to the sum of area for crack types reported as an area measure (e.g. fatigue cracking). The total cracked area was divided by the total surface area of each segment (assumed to be 52.8 feet long and 10 feet wide, or 528 square feet) to obtain a percent cracking measure, rounded to the nearest whole-number percent and capped at 100%.

## 4.2 Sampling
I started by selecting 10% of the pavement management sections. I had access to an Excel spreadsheet with the summary data for each of the 3,922 pavement management sections. I imported the key attributes into R and then used the sampling procedures available in the *dplyr* package to select 392 sections at random.

I then matched all detail records available for these 392 segments, giving 105,479 observations. This only included observations pertinent to the pavement evaluation; some segments were excluded including those collected on bridges or over rail crossings, as well as those that did not meet QA criteria for other reasons. The 105,479 points represents over 95% of all of the data collected within the 392 sample segments.

At various points in the process, I further sub-sampled using a similar procedure. This was primarily done in the service of visualizations, where an evaluation of 392 distributions simultaneously would not have been feasible. In these cases I used multiple seed values to select no fewer than five random samples of 40 sections, with the objective of seeing a broad cross-section of pavements in my visualizations.
4.3 EXPLORATORY ANALYSIS

I used various techniques to visualize the data. I found that box plots were very useful to help me gain insights into the distribution of the detailed data within each management section. Based on the box plots, I could see that some key factors would include age and pavement type. It was also helpful in determining that the data in each section were generally following a bell-shaped distribution, frequently with a bit of right skew. This was useful in confirming that my statistical procedures, which were based on the assumption of a normal distribution, were likely to give reasonably good results. Figure 3 is an example of one set of boxplot visualizations that were examined during the project.

I also plotted some of the summary (section-level) data in scatterplots against key attributes of interest, such as the plot shown in Figure 4. This helped me to better understand the relationships of these variables and consider whether this index was likely to work equally in all scenarios.

Figure 3: Sample of Boxplot Visualization for IRI

Figure 4: Example of Graphical Exploratory Analysis
4.4 Filtering

Based on the initial analysis and input from the pavement engineering team at Iowa DOT, I decided to focus my analysis on a subset of the segments. The following criteria were used to focus the analysis:

1. Any segment which had been constructed or rehabilitated within the past two years was removed. Sometimes data from segments with recent work can have temporal problems – it may be that the data were collected prior to the work having been done.
2. Segments identified as within city limits were excluded. Our engineers felt that this index would be most valuable as applied to rural segments, since factors such as ride and rutting are less important at lower speeds and are also more variable in urbanized areas.

We started with 11,092 miles of pavement management sections, and after eliminating miles corresponding to these criteria I was left with 7,952 miles from 2,111 segments.

4.5 Index Development

My work settled on the idea of using the Coefficient of Variation as the key building block for development of the Pavement Uniformity Index. The Coefficient of Variation is simply the standard deviation of each section divided by its mean. This gives a unit-less measure of “relative” variation that can be compared across measures. The Coefficient of Variation ($c_v$) can be calculated with the simple formula $c_v = \frac{\sigma}{\mu}$, where $\sigma$ is the standard deviation of the detailed measurements, representing each 1/100th mile, within each segment, and $\mu$ is the arithmetic mean of that segment. This value is quite straightforward to calculate using basic SQL commands and therefore should be able to be easily implemented in a production system.

Values of $c_v$ are to be computed for the major distress types (Cracking, Rutting, and Roughness). Note that rutting only applies to asphalt (flexible) surface pavements, and therefore the rutting measure is only part of the calculation for flexible pavements. Rigid pavements use only the IRI and percent cracking measures. These individual values would be useful, for example in the screening matrix shown in Figure 1 above. The values of $c_v$ for each distress could also be compared, and the average value for each section will represent the overall section’s raw index value. The raw value of the index is calculated as follows:

$$PUI_{raw} = \begin{cases} \frac{c_v[IRI] + c_v[Crack\%]}{2}, & \text{Pavement Type = Rigid} \\ \frac{c_v[IRI] + c_v[Crack\%] + c_v[RUT]}{3}, & \text{Pavement Type = Flexible} \end{cases}$$

The raw PUI values for all pavements are then evaluated, and the PUI value is determined by the decile of the empirical distribution of the PUI raw scores. For example, the largest 10% of all PUI raw values is assigned a PUI score of 10, while the lowest 10% of values is assigned a score of 1. Thus the PUI is always relative to the other pavements being evaluated.
5 Results

Referring back to Figure 1, we can now populate a similar table to see how many pavement miles fall into each category. As shown in Figure 5, the Iowa DOT has a significant number of rural highway pavement miles which could be classified in the “high risk” category. These are pavements which either have poor condition, poor uniformity, or both. Such pavements are shown in the red shaded area of this matrix and comprise approximately 1,383 miles of the system.

Those miles in the yellow-shaded region of the matrix might be considered medium-risk pavements. There are about 3,046 miles of pavement in this category. These pavements may need more investigation as various treatment alternatives are being explored.

Last we have the low-risk pavements, comprising about 3,522 miles of rural highways. These pavements may not need much in the way of treatment beyond routine maintenance, however if they do need treatment not much more than a basic set of investigation would be needed to have reasonable confidence in the design choices.

Another way to visualize the pavement uniformity is spatially. Using the geographic location of the pavement sections, I created a map with the segments colored according to their PUI value. This map is shown in Figure 6, with segments exhibiting greater uniformity shown in green and non-uniform sections shown in red. Segments with low uniformity are distributed throughout the state. In this map it is possible to see that highways in urbanized areas were excluded (for example in the Des Moines region near the center of the state). There are also a few other road segments shown without data, corresponding to segments that were under construction at the time of the data collection, or segments that had recently been rehabilitated and were therefore exempted from the analysis. This particular map was created using ArcGIS software from ESRI but any GIS package should be able to create something comparable.
6 CONCLUSION & FUTURE WORK

The index as it is currently formulated represents an important first step in the use of detailed pavement distress data to help inform how the Iowa DOT manages pavements. The PUI can be used as-is to provide useful information to the Iowa DOT relative to two important tasks:

1. Early screening and analysis of treatment alternatives.
2. Evaluating the potential accuracy of pavement performance models

Although this PUI has been developed as a prototype in the R language, it is “production ready” and DOT analysts are exploring its implementation in SQL code within the pavement management database.

Although the PUI has value in current form, this project has brought forward a number of potential enhancements which will be explored.

- Analyze the spatial component to dynamically segment the system into chunks that perform in a homogeneous fashion.
- Analyze changes in pavement uniformity over time in order to identify pavement sections which are deteriorating faster than expected.