Statistical Challenges presented by the SPS-2 Experiment and Data

Neeraj Buch[1], Member ISCP,
Karim Chatti[1]
Syed Waqar Haider[2]
Aswani Pulipaka[2]
Rick. W. Lyles[3] and
Dennis Gilliland[4]

Abstract

The challenges presented by the Specific Pavement Studies-2 (SPS-2) experiment and data for statistical analyses are presented in this paper. Experiment-related issues included (i) wide variation in traffic levels and ages among the fourteen test sites; and (ii) unbalanced distribution of test sites among climatic zones and subgrade types. The challenges presented by the data included (i) outlier issues, (ii) quantification of performance, and (iii) the lack of measurable distresses due to the "young" age of test sections. In this paper the importance of selecting appropriate data analysis methods for obtaining reliable results is illustrated. Though most of the methods that were applied for the analysis are well-established, the choice of the methods (magnitude- versus frequency-based) was driven by the extent and occurrence of distresses. Based on the current status of the data, frequency-based methods such as Linear Discriminant Analysis (LDA) and Binary Logistic Regression (BLR) lend themselves well to explaining trends associated with occurrence of cracking; whereas a magnitude based method like analysis of variance (ANOVA) is appropriate for evaluating joint faulting and change in IRI.

Research Significance

This paper highlights the challenges presented by in-service pavement performance data for statistical analyses. The methodology that was employed to address the identified issues is described in the paper. Such issues are typical for any in-service pavement data and thus the methodology presented here will assist pavement engineers perform more appropriate analyses and obtain meaningful results. The content presented in this paper is a part of a larger NCHRP project 20-50 (10/16), *LTPP Data Analysis: Influence of Design and Construction Features on the Response and Performance of New Flexible and Rigid Pavements*. The primary objectives of the project were to determine (i)"for specific site conditions, the effect of design and construction features on pavement response and (ii) the contributions of design and construction features to achieving different levels of performance."

Overview of the SPS-2 experiment

The primary objective of the SPS-2 experiment is to determine the relative influence and long-term effectiveness of design features and site conditions on the

[1]Associate Professor, Department of Civil and Environ. Engg., Michigan State University, USA
[2]Research Assistant, Department of Civil and Environ. Engg., Michigan State University, USA
[3]Professor, Department of Civil and Environmental Engineering, Michigan State University, USA
[4]Professor, Department of Probability and Statistics, Michigan State University, USA

performance of doweled jointed plain concrete pavement (JPCP) sections with transverse joints and uniform 4.6 m (15 ft.) joint spacing (Hanna et al. 1994). As the test sections in the experiment have been monitored since construction, the experiment provides an opportunity to better estimate, the relative influence of the key elements that affect pavement performance.

The design factorial for the SPS-2 experiment is shown in Table 1. The overall experiment consists of 192 factor-level combinations comprised of eight site-related (subgrade soil type and climate) combinations and 24 pavement structure combinations (design factors). The experiment was developed such that 12 sections were built, with only half of the possible combinations of design factors, at each of the 14 sites. It was planned that "48 test sections representing all structural factor and subgrade type combinations in the experiment are to be constructed in each of the climatic zones, with 24 test sections to be constructed on fine-grained soil and 24 test sections to be constructed on coarse-grained soil" (SHRP 1990). Moreover, for each climatic zone and soil type combination, 12 sections were to be constructed at one site and the other 12 sections at the other site. The geographical distribution of the SPS-2 sites is illustrated in Figure 1.

The structural (design) factors included in the experiment are:

- drainage (presence or lack of drainage),
- base type (DGAB, LCB, and PATB),
- PCC slab thickness [203 mm (8 in.) and 279 mm (11 in.)],
- PCC flexural strength [3.8 MPa (550 psi) and 6.2 MPa (900 psi), at 14-day], and
- lane width [3.7 m (12 ft.) and 4.3 m (14 ft.)].

The site factors included in the experiment are:

- subgrade soil type (fine-grained and coarse-grained), and
- climate (wet freeze, wet no freeze, dry freeze and dry no freeze).

At each site, six sections have a target PCC slab thickness of 203 mm (8 in.) and the remaining six have a target PCC slab thickness of 279 mm (11 in.). The 76 mm (3 in.) difference between the lower and upper levels of PCC slab thickness was believed to be necessary to demonstrate the effect of PCC slab thickness and its interaction with other factors on performance (SHRP 1990). The other factors with two levels (PCC flexural strength and lane width) each have six test sections corresponding to each level. Also four test sections have dense-graded aggregate base (DGAB), four sections have lean concrete base (LCB), and the other four sections have permeable asphalt treated base (PATB) over DGAB. In-pavement drainage is provided only for the sections with PATB as the base.

Though a major factor, traffic is not addressed like other design factors in that only a lower limit was specified for traffic volume in terms of ESALs per year. A SPS-2 test site must have a minimum estimated traffic loading of 200,000 rigid ESALs per year in the design lane. Based on the average annual precipitation and the Freezing Index, the sites in the experiment have been classified into different climatic zones using the thresholds defined in the LTPP program.

Table 1. SPS-2 Experiment Design Matrix (Hanna et al. 1994)

Pavement Structure					Climatic Zones, Subgrade															
Drainage	Base Type	PCC		Lane Width (m)	WET								DRY							
		Thickness (mm)	Strength (MPa)		FREEZE				NO FREEZE				FREEZE				NO FREEZE			
					Fine		Coarse		Fine		Coarse		Fine		Coarse		Fine		Coarse	
					J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
No	Dense-graded Aggregate Base (DGAB)	203	3.8	3.7	201		201		201		201		201		201		201		201	
				4.3		213		213		213		213		213		213		213		213
			6.2	3.7		214		214		214		214		214		214		214		214
				4.3	202		202		202		202		202		202		202		202	
		279	3.8	3.7		215		215		215		215		215		215		215		215
				4.3	203		203		203		203		203		203		203		203	
			6.2	3.7	204		204		204		204		204		204		204		204	
				4.3		216		216		216		216		216		216		216		216
No	Lean Concrete Base (LCB)	203	3.8	3.7	205		205		205		205		205		205		205		205	
				4.3		217		217		217		217		217		217		217		217
			6.2	3.7		218		218		218		218		218		218		218		218
				4.3	206		206		206		206		206		206		206		206	
		279	3.8	3.7		219		219		219		219		219		219		219		219
				4.3	207		207		207		207		207		207		207		207	
			6.2	3.7	208		208		208		208		208		208		208		208	
				4.3		220		220		220		220		220		220		220		220
Yes	Permeable Asphalt-treated Base (PATB)	203	3.8	3.7	209		209		209		209		209		209		209		209	
				4.3		221		221		221		221		221		221		221		221
			6.2	3.7		222		222		222		222		222		222		222		222
				4.3	210		210		210		210		210		210		210		210	
		279	3.8	3.7		223		223		223		223		223		223		223		223
				4.3	211		211		211		211		211		211		211		211	
			6.2	3.7	212		212		212		212		212		212		212		212	
				4.3		224		224		224		224		224		224		224		224

Note: 1 inch = 25.4 mm, 1 psi = 0.0069 MPa

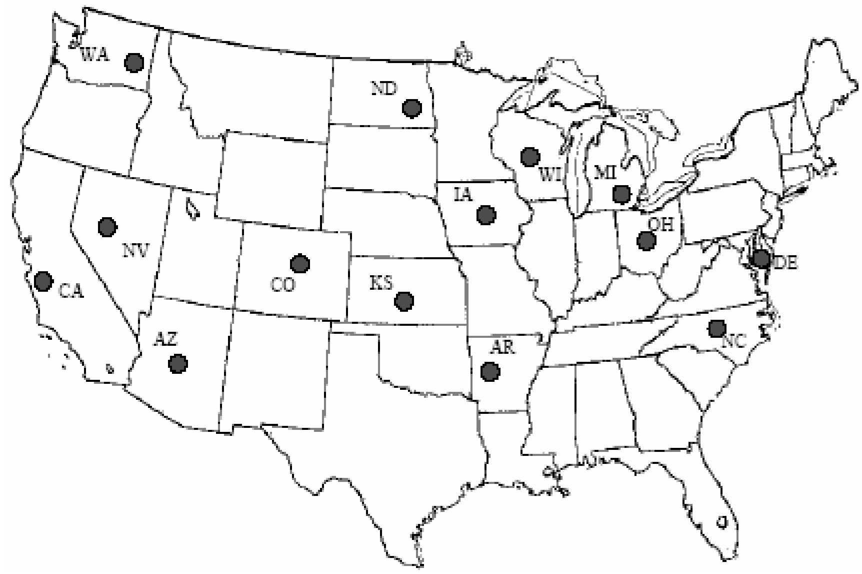


Figure 1. Geographical locations of SPS-2 sites (Chatti et al. 2005)

Experiment Design and Data Challenges

The challenges encountered during the execution of this project can be broadly classified as data- and experiment-related. Data-related issues are those that are a consequence of the extent and occurrence of distresses, and performance quantification whereas experiment-related issues are those that are due to the design of the experiment. These challenges are briefly summarized below:

- **<u>Data-related issues</u>**

  - *Outliers*: Certain test sections in the experiment have exhibited premature failure. This premature failure was primarily related to PCC mixture design issues and not due to inadequate design features. The inclusion of such data in the analysis has a potential to bias the conclusions.

  - *Quantification of performance*: The traffic opening dates for the test sites were staggered. This implies that ages of the test sites are not the same. Therefore, a direct comparison of pavement performance (area under the curve) is inappropriate. Moreover, there was a need to capture all of the time-series data in a single quantity that can best represent the performance.

  - *Extent and occurrence of distresses*: The average age of the SPS-2 test sections is 7.5 years. As a result of this, the extent and occurrence of distress is "low". For example, approximately 73% of the sections have not shown any transverse cracking. Though this indicates "good" performance, the "absence" of distress in a considerable number of sections presents challenges for statistical analysis and the interpretation of the results.

- **<u>Experiment design-related issues</u>**

  - *Traffic and age variation*: As explained in the overview of the SPS-2 experiment, only a lower limit for traffic was stipulated for a candidate site. As a result of this, a wide variation in traffic exists among SPS-2 sites. The traffic ranges from 151 KESAL/year for the WI (55) site to 2405 KESAL/year for the CA (6) site. Similarly, all sites were not constructed and opened to traffic at the same time, which resulted in a variation in age for the sites. The age of the sites ranges from 5 years for the CA (6) site to 12 years for the KS (20) site.

  - *Unbalanced site distribution*: Although the original experiment (Hanna et al. 1994) design called for four sites in each climatic zone, due to implementation issues the site distribution became unbalanced. Seven sites are located in the WF zone compared to two, three and two in WNF, DF, and DNF zones, respectively. In addition to the unbalance, the experiment does not have "true" (statistical) replication of designs. Although two sites (with 12 designs at each site) are located in the same climate-subgrade soil combination, the traffic, age, and material-related properties vary between the sites.

Data–related issues

The data-related issues include the consideration of outliers, quantification of performances, and the extent and occurrence of distress. Each is discussed in more detail below.

Outliers: Some of the test sections in the experiment have exhibited premature failure. These were primarily related to construction and PCC mixture design issues and not due to design features. The inclusion of such data in the analysis has a potential to bias the conclusions.

More specifically, the site-wise occurrence of transverse cracking in the SPS-2 test sections is shown in Figure 2. It is evident from the plot that the sections in Nevada, NV (32) have distinctly higher cracking than sections in any of the other sites. Figure 3 and Figure 4 are time-series plots for all SPS-2 sections with and without sections from NV (32). Premature cracking in sections from NV (32) is evident from the plots (see Figure 3). Review of construction reports for the NV site revealed that the concrete mixture design parameters were not in compliance with the SPS-2 guidelines and subsequently led to premature failure of PCC slabs. This is believed to be the main cause of abnormal performance of these sections. The site was thus excluded from the analysis so as to prevent bias in the results. Exclusion of sections from NV (32) site for data analysis has also contributed to unbalance in the experiment.

Quantification of Performance: For each section in the experiment, time-series performance data are available, starting from when the section was opened to traffic. The time-series performance data were not enough to perform an ANOVA based on repeated measures, which would have been more appropriate. However, all sites were opened to traffic at different points in time and, thus, a direct comparison of area under the performance curve (AREA) would not be appropriate. Thus there was a need to capture the entire set of time-series data in a single quantity that can best represent the performance.

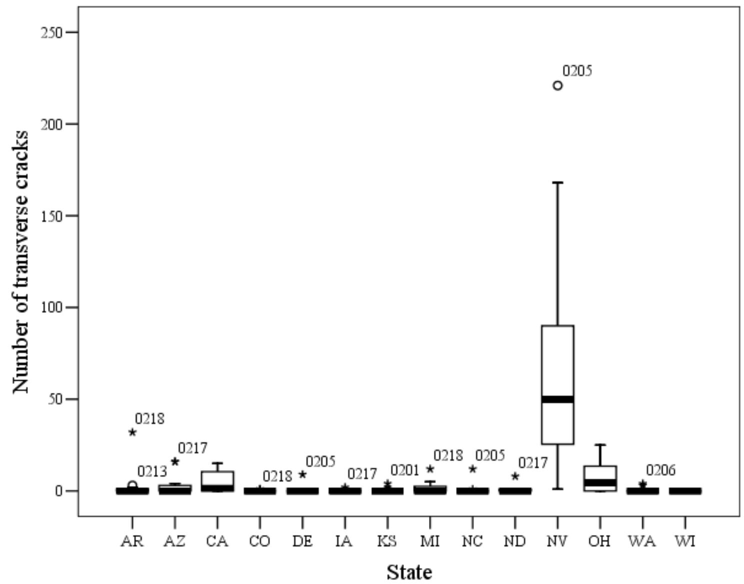


Figure 2. Site-wise occurrence of transverse cracking for SPS-2 test sections

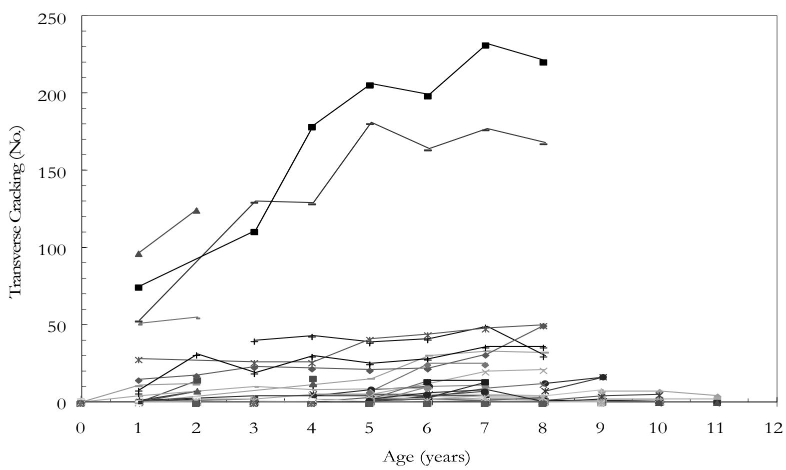


Figure 3. Transverse cracking with time - All sections

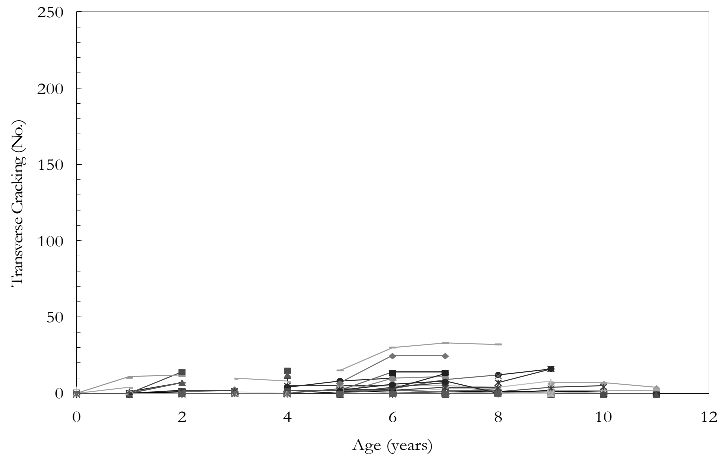


Figure 4. Transverse cracking with time - Without Nevada

Pavement performance is an accumulation of damage over time. Hence the area under a performance curve can be representative of the performance of a section. However, given the variation in age among the sites, two sections of different ages can have the same area though the older section has actually shown better performance. For example, the younger section could have significant distress over only a few years while the older one has more-or-less uniform distress over a much longer time; while the area under the curve would be the same, the older section would have exhibited much better performance. Normalizing the performance (i.e. AREA) with the age of a section is one of the potential solutions to the problem. However, for this study, to capture the time-series distress data of a section, a Performance Index (PI) as defined in equation (2) was used:

$$PI = \frac{\sum_i (t_i . y_i)}{\sum_i t_i} \qquad (2)$$

Where:

t_i = age of the section at distress measurement in year i

y_i = performance of the section at year i

Equation (2) can be further simplified to the form of a series as shown in equation (3). Only the ages at which distress measurements were taken are included in the calculation of PI. It can be seen from equation (3) that a higher weight will be given to the performance measured at later ages (as t_{i+1} is greater than t_i, and Σ t_i is constant for a given pavement section, at a given point in time). For example, for a pavement section with IRI of 1.25 m/km in year 2, 1.5 m/km in year 5 and 1.7 m/km in year 7, the PI at year 7 will be [(2*1.25)+ (5*1.5)+ (7*1.7)]/(2+5+7)= 1.6. Since PI gives greater weight to the most recent performance of a section, it may capture the most probable future trend of performance.

$$PI = \frac{t_1 . y_1}{\sum_i t_i} + \frac{t_2 . y_2}{\sum_i t_i} + \frac{t_3 . y_3}{\sum_i t_i} + ... + \frac{t_n . y_n}{\sum_i t_i} \qquad (3)$$

The following examples illustrate the advantages of PI over AREA or AREA/age. Figure 5 shows the growth in IRI of two test sections from sites in

AR (5) and KS (20). From the plot it is evident that the change in IRI associated with section 20-0209 is not significant, while the change in IRI associated with section 5-0214 is significant. Given the performance trend of section 5-0214, it can be expected that the growth in roughness will continue in the future. While the area under both the performance curves is equal, the PI correctly identifies the rougher section.

Figure 6 also shows the IRI performance of two sections from the same sites in AR (5) and KS (20) for comparison between PI and area normalized with age (i.e. AREA/age). The growth in IRI associated with section 20-0212 is less than the IRI growth associated with section 5-0214.

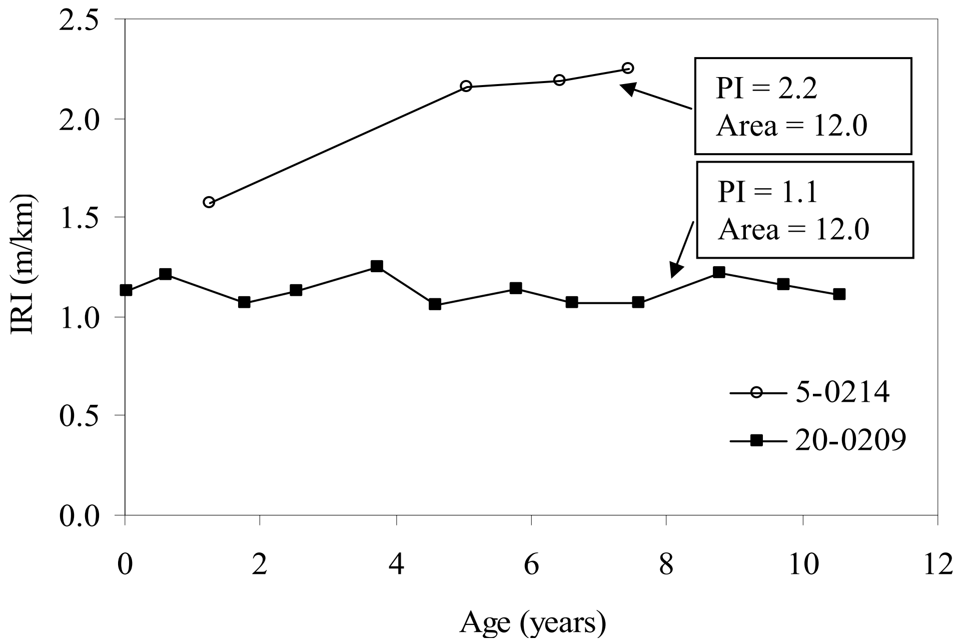


Figure 5. Comparison between PI and Area

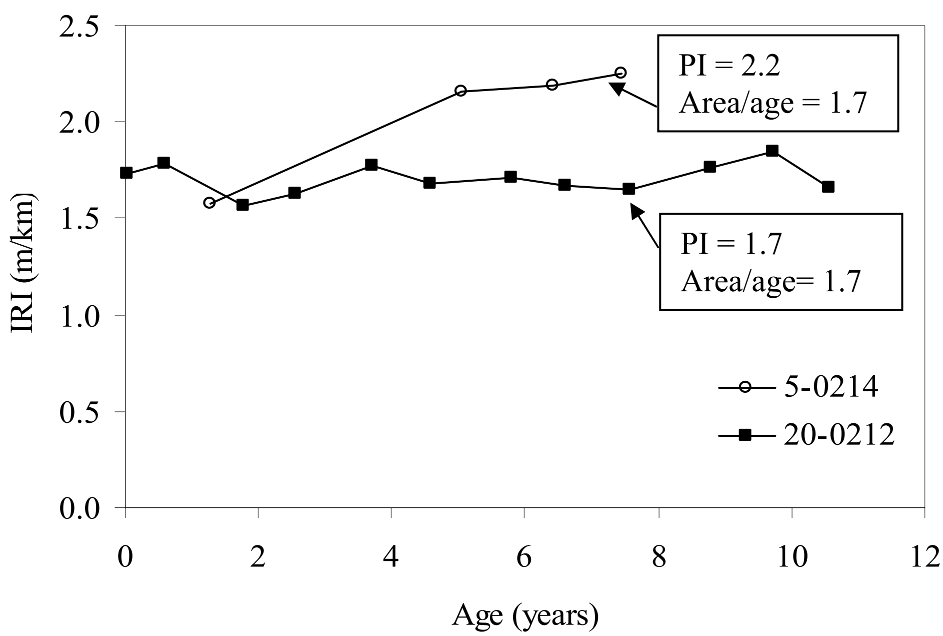


Figure 6. Comparison between PI and Area/age

While the area normalized to age for each of the sections is equal, PI could again capture this difference in performance. Similarly, Figure 7 shows a comparison of performance (transverse cracking) of two sections from sites in DE (10) and OH (39). Section 10-0205 exhibits early age cracking, while for section 39-0206 cracking initiated three years after construction. As mentioned before, PI gives more weight to performance at later ages. As a consequence of this, the early cracking in section 10-0205 that is not likely to be load-related is given lesser weight. Thus according to PI, section 39-0206 has shown inferior performance compared to section 10-0205. Area normalized to age could not detect this, as shown in the figure.

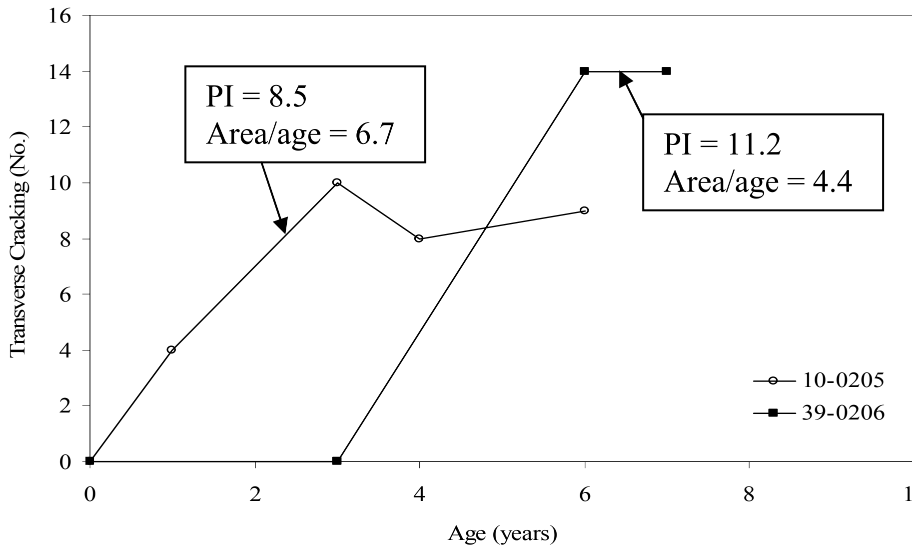


Figure 7. Comparison between PI and Area/age

Although PI appears to be a viable option it has some inherent limitations. These limitations are mainly because PI is dependent on the number and timing of the distress surveys. For two pavement sections of the same age, with one monitored each year and the other monitored every other year, the PI for the former section will be slightly lower. For the same sections if the monitoring were not performed at a regular time interval, the section with more surveys towards the later age will have a slightly higher PI. However, this limitation may not have considerable impact in the case of the SPS-2 experiment as all the pavement sections were monitored at a regular time interval of 1 to 2 years. In the SPS-2 experiment the pavement sections at different sites have different ages. Among pavement sections with different ages and similar performance, the PI of younger sections will be lower than that of older sections. To address this issue, the age of test sections was considered as a covariate in all statistical analyses of PI. This will adjust the PIs according to the age of pavement sections (Chatti et al. 2005).

Extent and Occurrence of distresses: The majority of the test sections in the SPS-2 experiment have been in service for less than 10 years. As a consequence, the extent and occurrence of distresses are “low”. Figure 8 and Figure 9 show the frequency (histogram) of occurrence of transverse and longitudinal cracking, respectively. From this histogram it is clear that the majority of the sections have not exhibited any distress. This implies that the sections are exhibiting "good" performance from an engineering standpoint. However, the non-occurrence of distress presents challenges for statistical analyses. Figures 10 and 11 illustrate the occurrence of joint faulting and roughness for the SPS-2 sections. The issue of non-occurrence was not encountered with these performance measures and magnitude-based methods were applied to these performance measures.

Traditionally, magnitude-based methods like multi-factorial ANOVA are used to establish the impact of design and site factors on pavement performance from an experiment like SPS-2. For the normal theory least squares ANOVA to be meaningful the model residuals should (i) be normally distributed and (ii) have a constant variance. For the ANOVA of occurrence of transverse cracking, these assumptions are violated as shown in Figure 12. A similar trend was observed when ANOVA was used to analyze longitudinal cracking. To address this issue, if a logarithmic transformation were applied to the data, ANOVA will be restricted only to sections that have shown cracking, as the transformation function is not defined for a value of zero. Given the "low" occurrence of cracking, this is undesirable at this point in time. The frequency-based methods are more suitable when the occurrence of the distress is considerable though the extent (magnitude and severity levels) is less, such as for cracking. Hence, Linear Discriminant Analysis (LDA) and Binary Logistic Regression (BLR), which are frequency-based methods, were used to study the effects of design and site factors on the occurrence of cracking. A detailed explanation of analyses for each performance measure can be found in the final report of NCHRP project (Chatti et al. 2005). Based on the current SPS-2 data, LDA yielded the following classification model for transverse cracking:

$$\begin{aligned} Score &= -4.8 - 0.26(\mathbf{BaseType}) + 0.62(\mathbf{PCCThickness}) + 0.45(\mathbf{Drainage}) \\ &-0.14(Zone) + 0.24(LWidth) + 1.05(SG) + 0.15(FS) \end{aligned} \quad (4)$$

Group Centroids: Transverse Crack 'No'= 0.51
Transverse Crack 'Yes'= -1.47

Where:

BaseType: '1' for DGAB or PATB, '2' for LCB.
PCCThickness: '1' for target slab thickness of 203 mm, '2' for target slab thickness of 279 mm.
Drainage: '0' for No Drainage, '1' for Drainage.
Zone: '1' for WF, '2' for WNF, '3' for DF, '4' for DNF.
LWidth: '1' for 3.7 m, '2' for 4.3 m.
SG: '1' for Fine-grained, '2' for Coarse-grained.
FS: '1' for target 14-day strength of 3.8 MPa, '2' for target 14-day strength of 6.2 MPa.

In Equation (4), the factors in bold are statistically significant in affecting the occurrence of transverse cracking. The equation can be interpreted by considering the group centroids and the particular effect each factor has on the score. For example, base type has a negative effect on the score, i.e. when base type goes from DGAB or PATB (code: 1) to LCB (code: 2), the chances of occurrence of cracking increases. Also, an increase in PCC thickness from 203 mm (code: 1) to 279 mm (code: 2) will have a positive effect on the overall score. This implies that increasing PCC slab thickness decreases the chances of cracking. The effects implied in the model are summarized in Table 2.

Table 2. Summary of results from LDA on transverse cracking

Factor	*p*-value	Effect on transverse cracking
Drainage	0.001	↑ (Presence of drainage significantly reduces the chances of occurrence of cracking.)
Target PCC thickness	0.001	↑ (Thicker PCC thickness reduces the chances of occurrence of cracking.)
Base type	0.044	↓ (Sections with LCB are more likely to exhibit transverse cracking.)

ANOVA is the magnitude-based analysis. The magnitude-based methods, although powerful, are more appropriate for analyses of distresses that have considerable occurrence and magnitude (for example, roughness). ANOVA was used in explaining the change in IRI as a function of design and site factors. Figure 13 is the residual plot from ANOVA on roughness. It can be seen that there is no serious violation of assumptions.

BLR is used often in the case where the outcome variable is discrete (e.g. dichotomous). The difference between logistic and linear regression is reflected both in the choice of a parametric model and in the assumptions. This method is based on the maximum likelihood method for determining the parameters of interest. The details of theoretical background of BLR are available in the literature (Hosmer and Stanley 2000). The interpretation of effects for various levels of the categorical (independent) variables is very convenient in terms of the odds ratio when this type of model is used. Logistic regression models are also very useful for discrimination analysis (of various groups) when categorical variables are used as independent variables. Table 3 is a summary of results from BLR on transverse cracking.

Table 3. Summary of results from BLR on transverse cracking

Experimental Factors	*p*-value (odds ratio)
Drainage	0.083 (3.4)
Base type	0.073 (4.9)
PCC thickness	0.019 (3.6)
Subgrade type	0.003 (9.5)

The following observations can be made as a result of BLR analysis on transverse cracking:

Drainage: The effect of drainage on cracking is marginally significant. Sections with drainage have lesser likelihood of cracking compared to those without drainage.

Base Type: The effect of base type on cracking is also marginally significant. The sections built on PATB have lesser likelihood of cracking than those built on *LCB.*

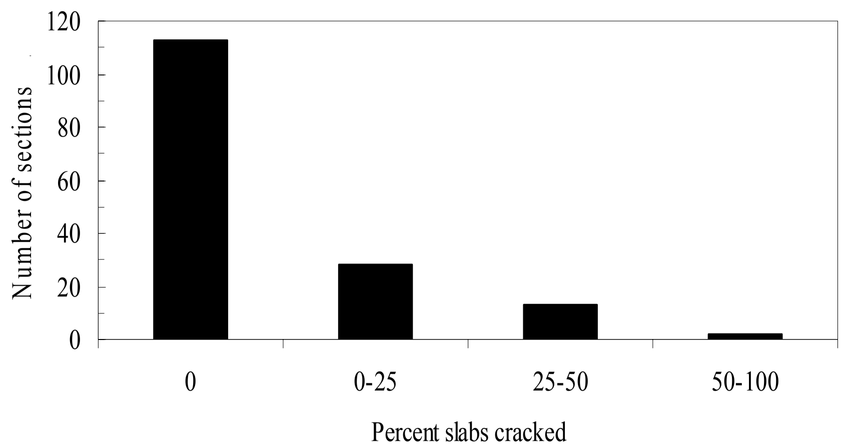


Figure 8. Occurrence of transverse cracking

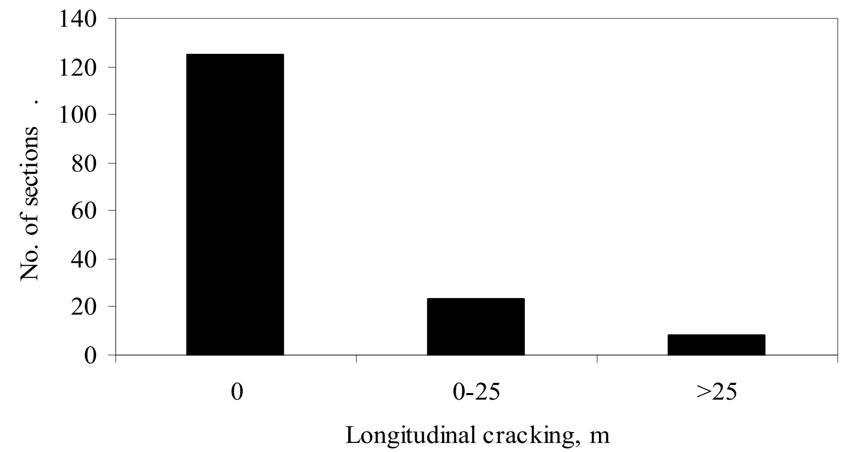


Figure 9. Occurrence of longitudinal cracking

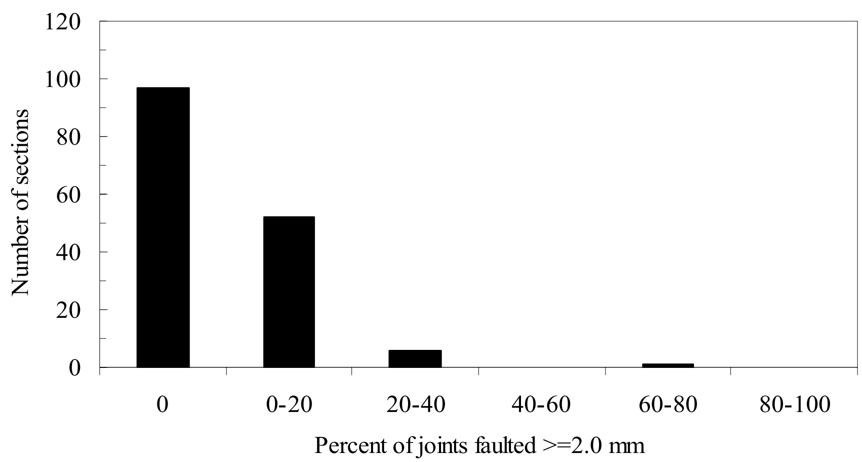


Figure 10. Occurrence of faulting

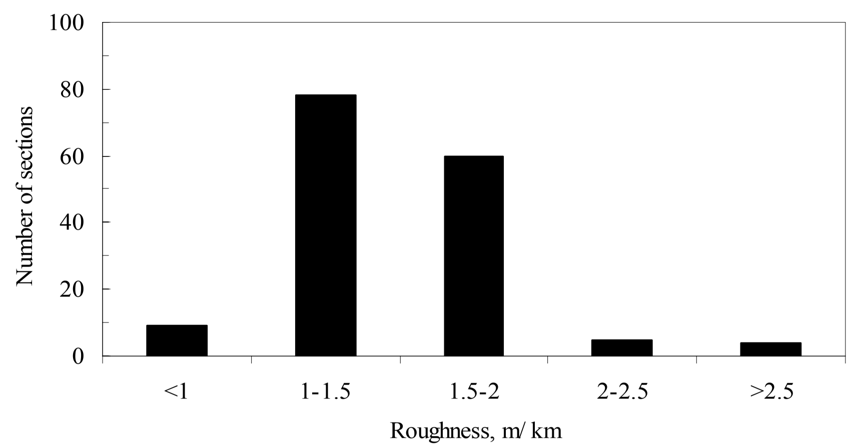


Figure 11. Distribution of roughness

PCC slab thickness: The sections built with 203 mm (8”) PCC slab have higher likelihood of cracking than the ones built with 279 mm (11”) PCC slab. This effect is statistically significant.
Subgrade: The sections built on fine subgrade soils have higher likelihood of cracking than the ones built on coarse subgrade soils. The effect is statistically significant.

In the interpretation of the results, the p-value is of primary importance as it indicates the statistical significance of the effect. When an effect is statistically significant it can be said that the effect is not due to random error (i.e. a happenstance). Hence, irrespective of the odds ratio, a higher p-value (>0.05) will make the odds ratio less reliable.

Experiment–related issues

As noted earlier, there were experiment-related issues as well. The issues that were encountered are described below, followed by how these were addressed in undertaking and interpreting the analysis.

Traffic and age variation: As explained in the overview of the SPS-2 experiment, only a lower limit for traffic was stipulated for a candidate site. As a result, wide variation in traffic exists among SPS-2 sites as shown in Figure 14. The dashed line in the figure is the stipulated traffic limit. It can be seen that a wide variation in traffic exists among the sites. Similarly, all the sites were not constructed and opened to traffic at the same time, which caused a variation in age for the sites. Figure 15 shows the age variation among the sites. The age of the sites ranges from 5 years for the CA (6) site to 12 years for the KS (20) site. These shortcomings were addressed by considering pavement age and traffic as covariates. Though the actual effects of traffic and age on pavement performance are quite complex (non-linear) and may not be adequately addressed by the covariates (linear adjustment), their inclusion in that manner helps in reducing the bias.

Unbalanced site distribution: The SPS-2 experiment, which is a fractional-factorial design, was rendered unbalanced because unequal numbers of sites were constructed in each zone-subgrade combination (Chatti et al. 2005). Table 4 shows the actual status of the SPS-2 experiment. Seven sites are located in the WF zone compared to two, three and two in the WNF, DF, and DNF zones, respectively. Furthermore, for some of the sites [CO (8), NV (32), and AR (5)] not all sections were constructed on the same subgrade soil type. The unbalance in the experiment resulted in a few or no replications for some treatments, which affects the statistical power (Kutner et al. 2005). In simple terms, the lower the power the lower is the reliability of results. As a result, the interaction effects (especially higher order interactions) of the factors may not be studied using methods like ANOVA (Haider 2005).

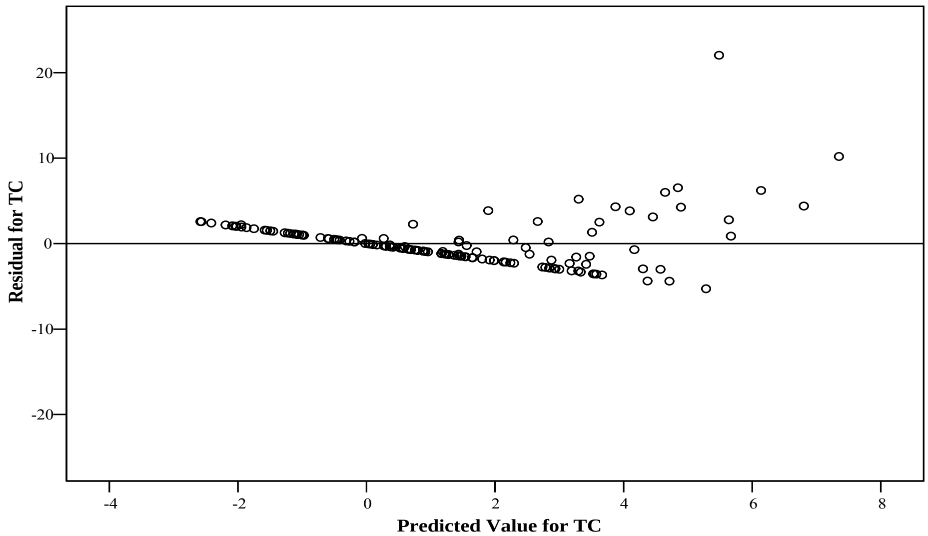


Figure 12. Residual plot for ANOVA on transverse cracking

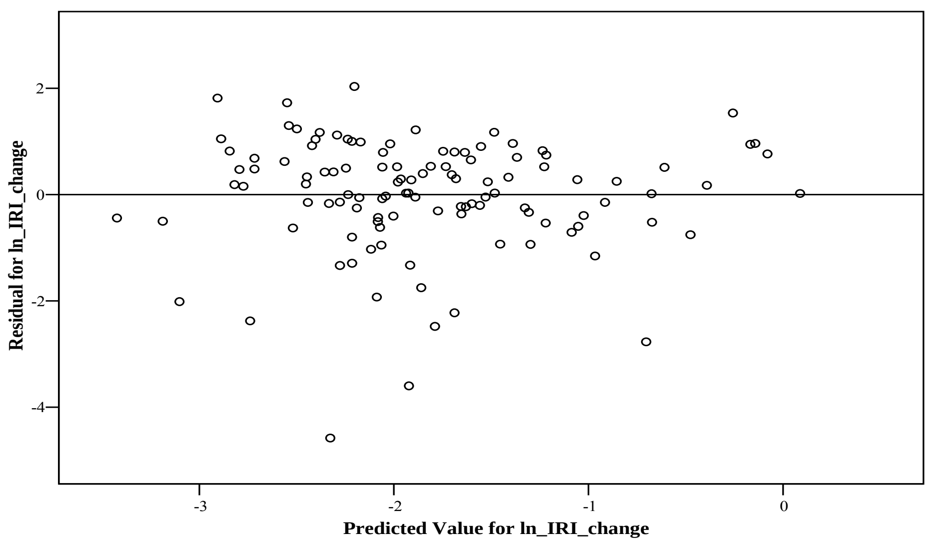


Figure 13. Residual plot for ANOVA on change in roughness

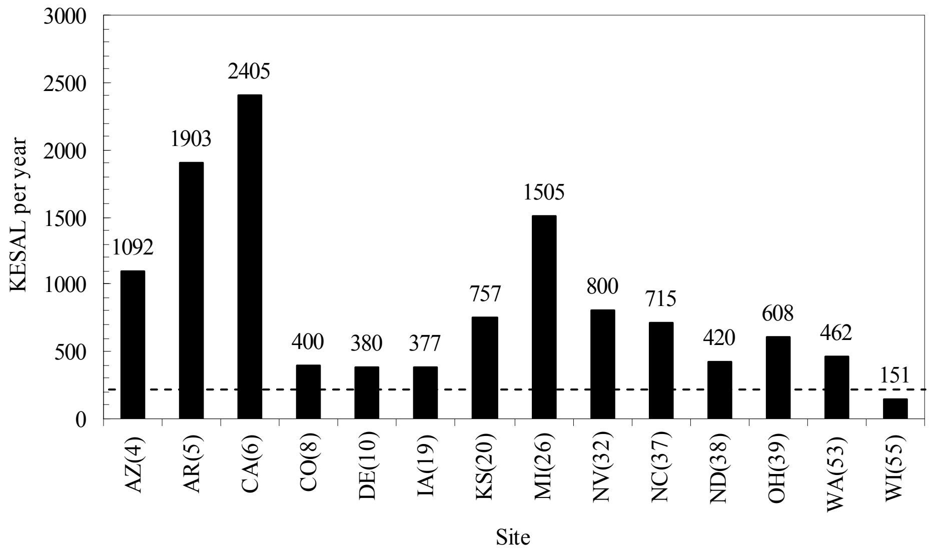


Figure 14. Traffic in KESAL per year for SPS-2 test sites

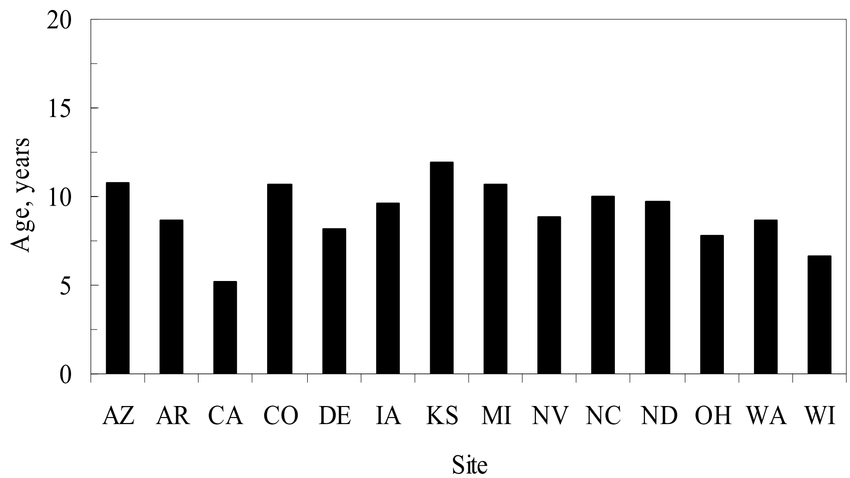


Figure 15. Age distribution of the SPS-2 sections

Table 4. Actual experiment status of the SPS-2 design (Chatti et al. 2005)

Subgrade Type	Designs	Wet		Dry		Total
		Freeze	No- Freeze	Freeze	No-Freeze	
Fine-grained	0201-0212	KS (20) OH (39)	NC (37)	WA (53) NV (32)*		9
	0213-0224	MI (26) IA (19) ND (38)		CO (8)*		
Coarse-grained	0201-0212	DE (10)			CA (6)	5
	0213-0224	WI (55)	AR (5)*		AZ (4)	
Total		7	2	3	2	14

*Two sections in NV and five sections in CO are coarse-grained while two sections in AR are fine-grained.

In addition to being unbalanced, the experiment does not have actual ("true") replication of pavement designs. Though two sites (with 12 designs at each site) are located in the same climate-subgrade soil combination, the traffic, age, and material-related properties vary between the sites.

In light of these issues, a simplified analysis based on the concept of the standard deviate was performed. This magnitude-based analysis was applied only to roughness and faulting where there was no problem with non-occurrence of distresses. The performance of designs with respect to each other was evaluated using the deviation from the mean performance, which is the standard deviate. The designs were evaluated based on selected distresses, considering one distress at a time. The standard deviate was calculated for each of the twelve designs within each site by using equation (5).

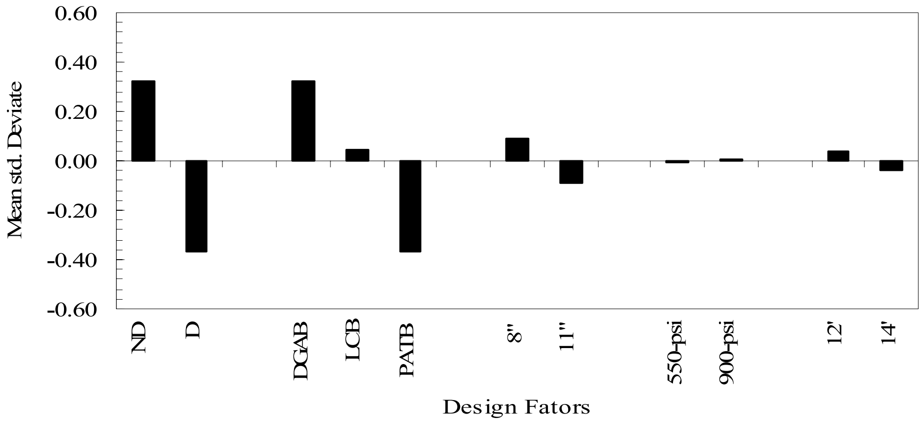
$$\text{Standard Deviate} = \frac{(\text{PI of a given design - Average PI for the given site})}{(\text{Standard Deviation of PIs for the given site})} \quad (5)$$

Since this measure was calculated individually for each section, considering one site at a time, it indicates the relative ranking of the section compared to other sections at the same site. This transformation adjusts for the variation in performance caused due to site conditions, as the sections are weighed with respect to companion sections in each site. This method is analogous to statistical "blocking" for site effects.

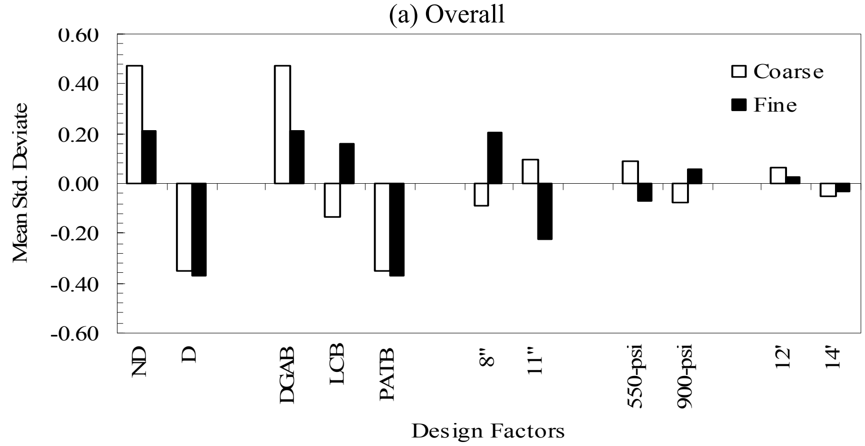
The standard deviate for a particular performance can also be used to compare the individual effects of design factors, and for this a one-way ANOVA was performed on the standard deviates of the sections. These analyses were performed on data from all sections and also on subsets of data stratified by subgrade type, climate and combinations of these two factors. This helps in identifying the effects of design factors under different site conditions. The standard deviate values of each design were averaged from the various sites to study the overall as well as the interaction effects of climatic zones and subgrade type. Figure 16 shows results from this analysis for roughness. Figure 16 (a) is for all the sections, while Figure 16 (b) and Figure 16 (c) show results from stratified data. From these figures, the effects of drainage, base type, and PCC slab thickness are evident. Sections with no drainage showed greater change in roughness as compared to those with drainage. This effect is consistent for different subgrade types [Figure 16 (b)] and in WF zone [Figure 16 (c)]. Since one-way ANOVA was employed in the analysis of standard deviates, age and traffic could not be used as covariates. However, the standard deviate as used here, does adjust for site effects, as explained above.

Conclusions

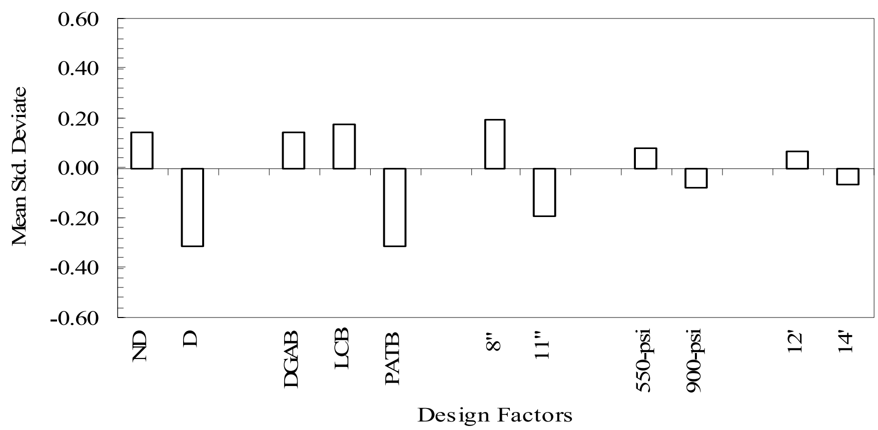
The statistical challenges presented by the SPS-2 experimental layout and data have been identified and examples of the applicability of frequency- and magnitude-based methods in evaluating the impact of design and site factors on the occurrence of pavement distress have been discussed and presented. The shortcomings of using the area under the performance curve as a measure to compare sections when pavement ages are different have also been examined. As an alternative, the performance index (PI) has been used to address issues related to time-series data. The z-deviate concept was used to block the site effects and allow for the relative comparisons of designs.



(a) Overall



(b) By subgrade type



(c) Wet Freeze (WF) zone

Figure 16. Effect of design factors on change in IRI

Acknowledgements

The research reported herein was performed as a part of NCHRP Project 20-50 (10/16), LTPP Data Analysis: Influence of Design and Construction Features on the Response and Performance of New Flexible and Rigid Pavements, by the Department of Civil and Environmental Engineering at Michigan State University. The research team acknowledges the funding provided by the National Cooperative Highway Research Program.

Disclaimer

The opinions and conclusions expressed or implied in the paper are those of the authors. They are not necessarily those of the National Cooperative Highway Research Program.

References

Chatti, K., Buch, N., Haider, S. W., Pulipaka, A., Lyles, R. W., Gilliland, D., and Desaraju, P. (2005). "Final Report NCHRP 20-50 (10/16), LTPP Data Analysis: Influence of Design and Construction Features on the Response and Performance of New Flexible and Rigid Pavements." *NCHRP 20-50 (10/16)*, Michigan State University, East Lansing.

Haider, S. W. (2005). "The use of long term pavement performance data for quantifying the relative effects of structural and environmental factors on the response and performance of new flexible pavements," Ph. D. Dissertation, Michigan State University, East Lansing.

Hanna, A. N., Tayabji., S. D., and Miller, J. S. (1994). "SHRP-LTPP Specific Pavement Studies: Five-Year Report." *SHRP-P-395*, Strategic Highway Research Program, National Research Council.

Hosmer, D. W., and Stanley, L. (2000). *Applied Logistic Regression*, John Wiley & Sons, Inc., NY.

Kutner, M. H., Christopher, J. N., John, N., and William, L. (2005). *Applied Linear Models*, McGraw-Hill, Irwin.

SHRP. (1990). "Specific Pavement Studies Experimental Design and Research Plan for Experiment SPS-2, Strategic Study of Structural Factors for Rigid Pavements." National Research Council, Washington, DC.