Intern. J. Pavement Engineering, 2000, Vol. 1(1), pp. 61-85
Reprints available directly from the publisher
Photocopying permitted by license only

Variability of Film-Derived Pavement Distress Data

CHUNG-LUNG WU*, GONZALO R. RADA, JOHN S. MILLER and RAJESH K. BHANDARI

LAW PCS, 12104 Indian Creek CT., #A, Beltsville, MD 20705, USA

(Received January 22, 1999; Revised June 08, 1999)

This paper presents the findings of an effort undertaken by the Federal Highway Administration (FHWA) Long-Term Pavement Performance (LTPP) program to assess variability of distress data collected using a photographic method. The primary source of data used to characterize the film-derived distress data variability is the results of a distress survey conducted on 12 pavement test sections, including six asphalt concrete (AC) and six portland cement concrete (PCC) pavements. Distress surveys were performed by experts, individual raters, and two-person teams. A consensus survey on the pavements by a group of experts based on the photographic method was also conducted. Finally, a reference rating of the distress present on the same 12 test sections was performed using a consensus manual rating method by the same group of experts; data from these reference surveys were used as a surrogate of "ground-truth". Based on the analyses of these data, variability of the film-derived distress data was evaluated and both apparent bias and precision for the data were quantified. It was observed that data interpreted by two-person teams were more consistent and showed smaller variations.

Keywords: Variability, Distress data, Filmed-derived, PASCO, Bias and precision

INTRODUCTION

Pavement distress is the primary performance indicator used in highway and airfield pavement maintenance and rehabilitation decision making. Through the use of composite indices such as the Pavement Condition Index (PCI), distress serves as a tracking mechanism and ultimately a trigger for maintenance and rehabilitation actions. Performance models developed based on distress data are also essential to the development of any mechanistic pavement design method, such as the AASHTO 2002 design procedure currently under development.

The process of collecting pavement distress data, analyzing it, and basing decisions for maintenance and rehabilitation outlays, i.e., pavement management, is an economic essential in operating a pavement facility. Reliable forecasting must be based on knowledge of performance deterioration (increasing amounts of distress) measured periodically over time. These data, coupled with mathematical modeling techniques, provide a picture of the performance of pavement types which may be used to predict future conditions.

Accurate, reliable distress data are critical to sound modeling and forecasting. Bad data will produce bad

* Corresponding author. Phone: (301) 210–5105 Fax: (301) 210–5032 E-Mail: clwu@lawco.com

models but the most accurate models will provide poor results if the distress data are not consistent, accurate and repeatable. Traditionally, distress data collection has been conducted by individuals according to some written guidelines, including distress definitions and methods of evaluating the recorded data. This process can be labor intensive and subject to human subjectivity and errors, even though guidelines have been well established.

To reduce these problems, a decision was made at the beginning of the SHRP Long-Term Pavement Performance (LTPP) program to use PASCO 35mm black and white photography to obtain frequent, objective surveys of the LTPP pavement test sections. Manual distress surveys were only used on a limited basis as a backup to the photographic survey method until 1992, when both their intensity and coverage increased. Consequently, PASCO film is the ONLY source of distress information for the majority of LTPP test sections in the first five to seven years of the LTPP program.

Actual distress data, in terms of distress types, severity levels and quantities, are obtained through a film interpretation process conducted some time after the filming event. Distress data interpretation was based on the guidelines contained in "Distress Identification Manual for the Long-Term Pavement Performance Project" (1993), which is typically referred to as the Distress Identification Manual (DIM).

The study presented here was part of an overall effort undertaken by the Federal Highway Administration to assess the level of variability associated with distress data, both manually collected and film-derived using PASCO's Pavement Distress Analysis System (PADIAS) v4.x system. Specific objectives of the study presented in this paper were to define the bias and precision for film-derived distress data, and to recommend possible means to improve them if necessary. Studies of variability associated with LTPP manual distress data have been presented elsewhere (Rada et al., 1998).

DATA SOURCE

Data used in this paper for assessing the variability of film-derived distress data consisted of:

- Reference surveys for 12 test sections (6 AC and 6 PCC) obtained through manual data collection.
- Consensus surveys for the same 12 test sections used in reference surveys obtained through interpretation of film using the PADIAS v4.x system.
- Individual expert surveys for the same 12 test sections used in reference surveys obtained through interpretation of film using the PADIAS v4.x system.
- Individual rater surveys for six of the 12 test sections used in reference surveys (3 AC and 3 PCC) obtained through interpretation of film using the PADIAS v4.x system.
- Two-person team consensus surveys for the remaining six of the 12 test sections used in reference surveys (3 AC and 3 PCC) obtained through interpretation of film using the PADIAS v4.x system.

Reference surveys were consensus manual surveys conducted by a group of four experts. These values were used as a surrogate of "ground truth" data in this study. The same group of experts also conducted consensus film distress data interpretation using the PADIAS v4.x system, and these values were assumed to be the best possible set of values for distress data derived from the system. These four experts also independently interpreted distress data obtained from these 12 pavement sections.

Distress data reduction from film was also done by six individual raters. These individuals independently performed distress data collection from film for six of the 12 test sections. Three repeat interpretations were performed on each section by each individual rater. The same individual raters were then paired into three groups and two-person team surveys from film were performed on the six test sections not interpreted by the individual raters. Three repeat interpretations were also performed on each of these sections by the two-person teams.

STATISTICAL DISTRESS DATA ANALYSES

To assess the variability of distress data derived from film using the PADIAS v4.x system, a series of statistical analyses were performed. An analysis of variance (ANOVA) test was first performed to determine whether agreement (based on means) existed within each of these groups -- experts, individual raters and two-person teams. A student's t-test was then conducted to assess whether or not significant differences existed between the groups. In this exercise, distress data generated by the experts was first compared to that from the individual raters and then to that from the two-person teams. The last statistical analysis, a components of variance analysis, was performed to identify sources of variations within each group.

Analysis of Variance (ANOVA)

An analysis of variance (ANOVA) was performed in this study to determine whether agreement (based on means) existed within each of the groups in question -- experts, individual raters and two-person teams. The null hypothesis of equal (within group) mean was assumed and F-statistics were computed for each distress type and severity level combination. If the calculated F-statistic was greater than the critical F-value at a confidence level of 95%, then the hypothesis was rejected -- i.e., within group means were not statistically the same.

Results of the ANOVA test are presented in Tables I and II for AC and PCC pavement sections, respectively. Three symbols are used in these tables, "S", "D", and "NA". The letter "S" denotes that the mean values for a given distress type and severity level combination are the same within the group in question, while "D" indicates that at least one of the mean values within the group for a given distress type-severity level combination is different from the others. "NA" indicates that the particular distress type was not observed by any of the experts, individual raters and/or two-person teams. In addition, several cells in Table II have been left blank if the distress type in question is not applicable; e.g. Section PCC2 cannot have corner breaks as this is a CRC pavement test section.

Not surprisingly, the results presented in these tables clearly indicate that there is variance within each of the groups. For AC test sections 1, 2 and 3, the individual raters agreed with one another only 47% of the time, while the experts agreed 43% of the time -- agreement here refers to equal within group means. For PCC test sections 1, 2 and 3, individual raters agreed 60% of the time with one another, while experts agreed 52% of the time. Similar results were obtained for the remaining pavement test sections. For AC test sections 4, 5 and 6, experts agreed with one another 49% of the time, and two-person teams agreed with each other 45% of the time. For PCC pavement sections 4, 5 and 6, the within group agreement was 82% and 68% for the experts and teams, respectively. (Note: test sections 1, 2 and 3, both AC and PCC, have been separated from 4, 5 and 6 in this comparison since individual raters only looked at the first three test sections and two-person team surveys were only performed on the latter three test sections.)

For both AC and PCC pavement sections, no appreciable differences were observed between experts and individual raters, and between experts and two-person teams. In general, all groups exhibited better agreement for PCC pavements than for AC pavements.

Student t-test

A student's t-test was conducted to determine whether or not significant differences in the group means existed between the experts, individuals and two-person teams for each distress type. In this exercise, distress data generated by the experts was first compared to that from the individual raters and then to that from the two-person teams. The null hypothesis of equal (between groups) mean was assumed and t-statistics were computed for each distress type and severity level combination. If the calculated t-statistic was greater than the critical t-value at a confidence level of 95%, then the hypothesis was rejected -- i.e., means of two groups being compared were not statistically the same.

TABLE I Results of analysis of variance (ANOVA) for AC sections – within group

Distress type	*Units*	*Sev.*	*AC1*		*AC2*		*AC3*		*AC4*		*AC5*		*AC6*	
			Expert	*Indiv.*	*Expert*	*Indiv.*	*Expert*	*Indiv.*	*Expert*	*Team*	*Expert*	*Team*	*Expert*	*Team*
Fatigue cracking	Sq. M	Low	D	D	D	D	D	D	D	D	S	D	S	D
		Mod.	D	D	NA	NA	S	S	D	S	D	D	NA	NA
		High	S	S	NA	NA	NA	NA	NA	NA	S	D	NA	NA
		Total	D	D	D	D	D	D	D	D	D	D	S	D
Long. cracking NWP	Meters	Low	D	D	S	D	D	S	S	S	D	D	S	D
		Mod.	S	D	S	S	S	S	D	S	S	D	D	D
		sHigh	S	D	S	NA	NA	S	S	S	S	D	S	S
		Total	D	S	D	D	D	S	S	S	S	S	S	D
Long. cracking WP	Meters	Low	D	D	S	S	D	D	D	D	D	S	S	NA
		Mod.	D	D	S	NA	S	D	S	S	D	D	NA	NA
		High	S	S	NA	NA	S	S	NA	NA	NA	S	NA	NA
		Total	D	D	S	S	D	D	S	D	S	S	S	NA
Trans. cracking	Meters	Low	D	D	D	S	D	D	D	S	S	D	D	D
		Mod.	D	S	S	D	S	S	D	D	S	S	S	S
		High	S	S	NA	NA	S	S	S	NA	D	D	NA	NA
		Total	D	S	S	S	D	D	D	S	S	D	D	S
Trans. cracking	No.	Low	D	D	D	D	D	D	D	S	D	D	D	S
		Mod.	D	S	S	D	S	S	D	D	D	S	S	S
		High	S	S	D	D	S	S	S	D	D	D	NA	NA
		Total	D	S	D	D	D	S	D	S	D	D	D	S

Note:
S = equal group means
D = at least one different group mean
NA = no distress observed

TABLE II Results of analysis of variance (ANOVA) for AC sections – within group

Distress type	*Units*	*Sev.*	*PCC1*		*PCC2*		*PCC3*		*PCC4*		*PCC5*		*PCC6*	
			Expert	*Indiv.*	*Expert*	*Indiv.*	*Expert*	*Indiv.*	*Expert*	*Team*	*Expert*	*Team*	*Expert*	*Team*
Corner breaks	No.	Low	D	D			NA	S	D	S	D	S	NA	NA
		Mod.	D	D			NA	S	D	S	D	S	NA	NA
		High	NA	D			NA	NA	S	S	S	S	NA	NA
		Total	S	D			NA	D	S	D	D	S	NA	NA
Long. cracking	Meters	Low	S	S	S	S	S	S	S	D	S	D	S	D
		Mod.	D	S	D	D	NA	S	S	NA	D	S	NA	NA
		High	NA	S	S	D	NA	NA	S	NA	S	S	NA	NA
		Total	S	S	S	S	S	S	S	D	S	D	S	D
Trans. cracking	No.	Low	D	S	S	S	D	D	S	S	S	S	S	D
		Mod.	S	D	S	S	D	S	S	S	S	S	S	S
		High	S	S	S	NA	D	D	S	S	S	D	S	D
		Total	D	S	S	S	D	S	S	S	S	S	S	D
Trans. cracking	Meters	Low	S	S	D	S	D	D	S	S	S	S	S	D
		Mod.	D	D	S	S	D	S	S	S	S	S	D	S
		High	S	S	S	NA	D	D	S	S	S	D	S	D
		Total	D	S	D	D	D	S	S	S	S	S	S	D
Spalling of long. joints	Meters	Low	D	D	S	S	D	S	D	S	D	D	S	S
		Mod.	S	S	S	S	D	S	S	D	D	S	S	NA
		High	NA	NA	S	D	S	D	S	NA	S	S	S	S
		Total	D	D	D	D	D	D	S	D	D	D	D	S
Spalling of trans. joints	No.	Low	S	S			S	D	S	S	S	S	NA	NA
		Mod.	S	S			D	S	D	S	S	S	S	NA
		High	D	D			S	D	S	S	S	NA	D	S
		Total	S	S			S	D	S	S	S	S	S	S
Spalling of trans. joints		Low	S	S			S	D	S	S	S	S	S	S
		Mod.	S	S			D	S	S	S	S	S	S	D
		High	D	D			D	S	S	S	S	S	S	D
		Total	D	S			S	D	S	D	S	S	S	D

Note:
S = equal group means
D = at least one different group mean
NA = no distress observed

Comparisons were first made between the means from the individual experts and those from the individual raters; i.e., test sections 1, 2 and 3, for both AC and PCC pavements. Similar comparisons were then made between the means from the individual experts and two-person teams using data from test sections 4, 5 and 6, for both AC and PCC pavements.

The computed t-statistics and critical t-values are shown in Tables III and IV for AC and PCC pavements, respectively. Three symbols are used in these tables, "S", "D", and "NA". The letter "S" denotes that the between group means being compared for a given distress type and severity level combination are statistically the same, while "D" indicates that the means are not statistically the same. "NA" indicates that the particular distress type was not observed by any of the experts, individual raters and/or two-person teams. In addition, several cells in Table IV have been left blank if the distress type in question is not applicable; e.g. Section PCC2 cannot have corner breaks as this is a CRC pavement test section.

The results contained within both of these tables appear to indicate that there are no significant differences in the group means between experts and individual raters and between experts and two-person teams. However, this encouraging outcome must be tempered by the fact that the results are, to a large extent, affected by the high degree of variability associated with each data group. In several cases, differences in the means between groups were masked by the high variability.

Components of Variance Analysis

Total measurement variation for a given distress type severity level combination can be attributed to two sources, which are referred to as the *components of variance*. These two sources are the within and the between rater variation, and can be mathematically expressed as follows:

$$\sigma_t^2 = \sigma_w^2 + \sigma_a^2$$

where σ_t^2 = total measurement variation

σ_w^2= within rater variation (σ_w= within rater standard deviation)

σ_a^2 = between rater variation (σ_a= between rater standard deviation.)

The within rater component can be estimated anytime the raters repeat some or all of their measurements; this component can be viewed as the *repeatability component*. If this variance component is large, it indicates that the raters are unable to produce precise measurements and thus they need either improved measurement methods or better training in the measurement of those distresses having large variances.

The between rater component of variance is the result of different raters giving different values for a given distress type and severity level combination which is not accounted for by the within rater noise. This variance may be regarded as a *bias* due to differences in the manner in which the raters perform their work. If this variance is large, it indicates that rater training is required to improve measurement consistency between raters. In essence, this training must serve as a means for rater calibration. This training may need to be updated over time, depending on the measured between rater variance.

The PASCO/PADIAS distress data available for this study was adequate for conducting the components of variance analysis. This analysis was performed on all distress type-severity level combinations for the same 12 pavement test sections referenced earlier in this paper; i.e., six AC and six PCC test sections. Tables V and VI present part of the analysis results for AC and PCC pavement sections, respectively. The information presented in these tables includes average distress quantity, within rater standard deviation, and between rater standard deviation for experts, individual raters, and teams.

The coefficient of variation (CV), defined as standard deviation divided by mean, is a statistical term normally used for representing the relative variability associated with experimental data. However, the CV can be misleading when dealing with small magnitudes, as is often the case with distress data. A small amount of distress can inflate the CV tremendously; however, that variation in data (high CV) is not as much of a concern when dealing with small distress amounts. To overcome this deficiency, an alternative approach was used for determining the CV associated with film-derived distress data.

TABLE III Results of t-test for AC pavement sections

Distress type	*Units*	*Sev.*	*AC1*			*AC2*			*AC3*			*AC4*			*AC5*			*AC6*		
			Calcu. t-sta.	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*		*Calcu t-sta.*	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*	
Fatigue cracking	Sq. M	Low	2.51	2.09	D	1.94	2.16	S	2.77	2.05	D	3.81	2.09	D	1.59	2.16	S	1.84	2.31	S
		Mod.	0.90	2.06	S	NA	NA	NA	2.29	2.11	D	0.10	2.15	S	0.79	2.09	S	NA	NA	NA
		High	0.91	2.05	S	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.93	2.31	S	NA	NA	NA
		Total	1.57	2.08	S	1.94	2.16	S	1.82	2.05	S	3.76	2.09	D	2.42	2.10	d	1.84	2.31	s
Long. cracking NWP	Meters	Low	0.18	2.07	S	2.24	2.10	D	2.40	2.05	D	2.82	2.26	D	1.09	2.12	S	0.60	2.16	S
		Mod.	0.13	2.05	S	4.65	2.06	D	2.88	2.09	D	1.51	2.09	S	2.17	2.23	S	2.28	2.09	D
		High	1.58	2.06	S	1.00	2.20	S	2.61	2.11	D	0.82	2.09	S	1.73	2.26	S	0.35	2.20	S
		Total	0.97	2.09	S	2.83	2.10	D	3.84	2.05	D	3.06	2.31	D	3.28	2.09	D	0.59	2.15	S
Long. cracking WP	Meters	Low	2.06	2.05	D	3.28	2.07	D	1.76	2.05	S	3.89	2.11	D	2.68	2.09	D	1.29	2.20	S
		Mod.	3.09	2.06	D	1.00	2.20	S	2.77	2.10	D	0.03	2.09	S	1.10	2.23	S	NA	NA	NA
		High	1.67	2.09	S	NA	NA	NA	2.76	2.11	D	NA	NA	NA	1.20	2.31	S	NA	NA	NA
		Total	4.60	2.06	D	3.26	2.07	D	3.17	2.05	D	3.91	2.12	D	3.52	2.09	D	1.29	2.20	S
Trans. cracking	Meters	Low	0.17	2.06	S	1.03	2.16	S	0.13	2.05	S	3.13	2.09	D	0.44	2.09	S	2.05	2.18	S
		Mod.	0.17	2.06	S	0.40	2.08	S	1.85	2.05	S	2.33	2.18	D	1.52	2.12	S	0.07	2.10	S
		High	1.10	2.05	S	NA	NA	NA	2.17	2.08	D	1.00	2.20	S	0.49	2.15	S	NA	NA	NA
		Total	2.46	2.08	D	1.03	2.16	S	0.26	2.05	S	2.35	2.09	D	0.66	2.16	S	1.27	2.13	S
Trans. cracking	No.	Low	0.37	2.06	S	0.00	2.05	S	0.37	2.05	S	0.44	2.09	S	1.14	2.12	S	0.57	2.11	S D
		Mod.	1.76	2.08	S	2.10	2.16	S	0.12	2.06	S	1.73	2.26	S	2.03	2.31	S	1.18	2.20	S
		High	1.20	2.06	S	NA	NA	NA	2.40	2.09	D	1.00	2.21	S	0.26	2.16	S	NA	NA	NA
		Total	2.74	2.05	D	1.00	2.16	S	0.01	2.05	S	4.58	2.11	D	1.33	2.10	S	1.93	2.15	S

Note:
S = equal group means
D = at least one different group mean
NA = no distress observed
Sections AC1, AC2 and AC3 are expert vs. individual
Sections AC4, AC5 and AC6 are expert vs. two-person team

TABLE IV Results of t-test for PCC pavement sections

Distress type	*Units*	*Sev.*	*PCC1*			*PCC2*			*PCC3*			*PCC4*			*PCC5*			*PCC6*		
			Calcu. t-sta.	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*		*Calcu. t-sta.*	*Critical t-sta.*	
Corner breaks	No.	Low	0.45	2.05	S				1.84	2.11	S	5.20	2.09	D	0.51	2.11	S	NA	NA	NA
		Mod.	1.60	2.05	S				1.37	2.11	S	0.42	2.11	S	1.69	2.10	S	NA	NA	NA
		High	1.72	2.11	S				NA	NA	NA	0.04	2.13	S	1.63	2.26	S	NA	NA	NA
		Total	4.19	2.06	D				2.06	2.11	S	3.28	2.13	D	4.63	2.11	D	NA	NA	NA
Long. cracking	Meters	Low	0.23	2.05	S	1.52	2.07	S	1.01	2.11	S	1.52	2.20	S	1.13	2.09	S	0.28	2.09	S
		Mod.	0.31	2.09	S	2.44	2.08	D	1.00	2.11	S	1.72	2.20	S	5.29	2.12	D	NA	NA	NA
		High	1.02	2.11	S	0.11	2.06	S	NA	NA	NA	1.00	2.20	S	0.34	2.10	S	NA	NA	NA
		Total	0.03	2.11	S	1.11	2.06	S	1.01	2.11	S	1.74	2.20	S	4.04	2.10	D	0.28	2.09	S
Trans. cracking	No.	Low	1.10	2.05	S	1.52	2.07	S	2.32	2.05	D	5.37	2.18	D	0.73	2.13	S	0.58	2.16	S
		Mod.	2.99	2.05	D	1.29	2.09	S	1.04	2.10	S	2.28	2.09	D	1.18	2.09	S	3.43	2.11	D
		High	0.90	2.06	S	1.00	2.20	S	0.95	2.06	S	0.28	2.10	S	0.33	2.10	S	0.27	2.11	S
		Total	2.31	2.05	D	0.68	2.07	S	1.15	2.15	S	6.15	2.16	D	0.62	2.10	S	0.36	2.16	S
Trans. cracking	Meters	Low	0.07	2.05	S	2.13	2.06	D	2.54	2.05	D	3.35	2.20	D	0.76	2.15	S	1.58	2.09	S
		Mod.	2.89	2.05	D	1.19	2.09	S	1.53	2.11	S	2.05	2.09	S	0.47	2.10	S	4.01	2.16	D
		High	1.53	2.05	S	1.00	2.20	S	1.15	2.06	S	0.48	2.18	S	0.33	2.10	S	0.22	2.09	S
		Total	2.07	2.09	S	2.36	2.05	D	0.31	2.18	S	3.62	2.18	D	0.08	2.15	S	0.59	2.11	S
Spalling of long. joints	Meters	Low	0.68	2.11	S	7.27	2.20	D	5.24	2.20	D	1.17	2.16	S	0.39	2.13	S	4.48	2.18	D
		Mod.	1.76	2.08	S	2.10	2.16	S	0.12	2.06	S	1.73	2.26	S	2.03	2.31	S	1.18	2.20	S
		High	NA	NA	NA	0.76	2.13	S	2.88	2.07	D	1.12	2.20	S	0.59	2.23	S	0.89	2.16	S
		Total	1.06	2.08	S	7.73	2.06	D	5.43	2.18	D	0.64	2.10	S	0.82	2.15	S	5.54	2.18	D
Spalling of trans. joints	No.	Low	1.33	2.11	S				1.22	2.06	S	1.20	2.13	S	2.19	2.23	S	NA	NA	NA
		Mod.	0.71	2.06	S				2.28	2.15	D	0.45	2.18	S	0.30	2.12	S	1.77	2.20	S
		High	0.48	2.06	S				2.30	2.11	D	0.39	2.09	S	1.00	2.20	S	5.12	2.10	D
		Total	1.27	2.06	S				2.92	2.05	D	1.42	2.15	S	1.73	2.16	S	6.89	2.09	D
Spalling of trans. joints	Meters	Low	2.22	2.11	D				0.22	2.11	S	0.17	2.11	S	2.17	2.13	D	0.73	2.26	S
		Mod.	0.53	2.05	S				1.52	2.09	S	0.32	2.09	S	0.95	2.11	S	0.12	2.09	S
		High	0.19	2.07	S				1.96	2.07	S	0.35	2.10	S	1.32	2.16	S	1.04	2.09	S
		Total	1.78	2.07	S				1.88	2.11	S	0.29	2.09	S	1.67	2.11	S	0.91	2.11	S

Note:
S = equal group means
D = at least one different group mean
Sections PCC1, PCC2 and PCC3 are expert vs. individual
Sections PCC4, PCC5 and PCC6 are expert vs. two-person team

TABLE V Partial results of component of variance analysis for AC pavement sections

Distress	*Units*	*Sev.*	*Sec.*	*Expert*			*Individual*			*Team*		
				Mean	ϕ_a	ϕ_w	*Mean*	ϕ_a	ϕ_w	*Mean*	ϕ_a	ϕ_w
Fatigue cracking	Sq. Meters	Low	#1	19.7	9.3	7.1	10.7	6.9	4.9			
			#2	3.0	4.7	1.1	0.7	1.1	0.6			
			#3	60.7	25.7	11.9	30.9	28.8	17.8			
			#4	24.1	7.7	7.2				9.3	8.1	3.4
			#5	20.0	5.02	5.8				13.2	12.4	3.0
			#6	0.0	0.0	0.0				2.1	3.6	1.0
		Mod.	#1	5.7	2.6	2.4	7.6	5.6	5.8			
			#2	0.0	0.0	0.0	0.0	0.0	0.0			
			#3	0.4	0.0	0.9	8.4	0.0	15.1			
			#4	0.4	0.5	0.4				0.3	0.0	0.8
			#5	11.9	10.3	3.0				8.9	8.6	1.9
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		High	#1	0.2	0.0	0.6	0.4	0.4	0.7			
			#2	0.0	0.0	0.0	0.0	0.0	0.0			
			#3	0.0	0.0	0.0	0.0	0.0	0.0			
			#4	0.0	0.0	0.0				0.0	0.0	0.0
			#5	0.1	0.1	0.1				2.3	4.0	0.4
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		Total	#1	25.6	12.5	5.3	18.7	9.7	5.5			
			#2	3.4	4.2	2.3	0.9	1.4	0.8			
			#3	61.0	25.8	12.5	39.3	30.1	26.7			
			#4	24.5	8.0	7.4				9.6	7.9	3.6
			#5	32.0	7.6	6.0				24.3	4.8	3.1
			#6	0.1	0.0	0.3				2.1	3.6	1.0
Long cracking WP	Meters	Low	#1	17.8	14.4	4.3	29.7	16.3	9.4			
			#2	16.7	1.4	2.7	20.3	1.8	2.3			
			#3	43.4	39.0	15.5	71.9	37.8	3438			
			#4	41.9	18.8	15.3				82.6	26.4	8.6
			#5	16.5	9.2	6.7				27.1	0.2	7.3
			#6	0.3	0.4	0.6				0.0	0.0	0.0
		Mod.	#1	8.2	5.9	3.1	18.8	10.5	7.6			
			#2	0.0	0.0	0.1	0.0	0.0	0.0			
			#3	8.5	0.0	5.6	29.8	23.0	23.6			
			#4	1.6	1.9	3.6				1.5	2.1	2.8
			#5	5.2	4.1	1.8				29.0	3.7	7.9
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		High	#1	0.6	0.0	1.3	3.2	3.2	5.8			
			#2	0.0	0.0	0.0	0.0	0.0	0.0			
			#3	0.0	0.0	0.1	2.7	0.5	4.1			
			#4	0.0	0.0	0.0				0.0	0.0	0.0
			#5	0.0	0.0	0.0				0.5	0.5	1.1
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		Total	#1	26.6	15.0	4.3	51.7	12.4	9.9			
			#2	16.7	1.3	2.7	20.3	1.8	2.4			

Distress	Units	Sev.	Sec.	Expert			Individual			Team		
				Mean	ϕ_a	ϕ_w	Mean	ϕ_a	ϕ_w	Mean	ϕ_a	ϕ_w
			#3	52.0	37.3	17.0	104.4	43.8	33.2			
			#4	43.4	17.7	15.5				84.2	26.0	9.8
			#5	21.7	8.2	8.4				37.1	5.4	7.5
			#6	0.3	0.4	0.6				0.0	0.0	0.0

TABLE VI #Partial results of component of variance analysis for PCC pavement sections

Distress	Units	Sev.	Sec.	Expert			Individual			Team		
				Mean	ϕ_a	ϕ_w	Mean	ϕ_a	ϕ_w	Mean	ϕ_a	ϕ_w
Corner breaks	Number	Low	#1	2.7	1.7	0.9	3.1	2.4	1.9			
			#3	0.0	0.0	0.0	0.2	0.2	0.3			
			#4	8.8	2.2	1.9				14.1	1.3	1.7
			#5	4.3	3.6	1.0				3.7	0.0	1.8
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		Mod.	#1	3.4	1.3	1.2	4.8	2.8	1.5			
			#3	0.0	0.0	0.0	0.2	0.3	0.4			
			#4	5.3	1.4	1.1				5.6	0.5	1.6
			#5	7.1	4.0	1.4				9.3	1.4	1.7
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		High	#1	0.0	0.0	0.0	0.4	1.1	0.5			
			#3	0.0	0.0	0.0	0.0	0.0	0.0			
			#4	9.2	0.0	2.5				0.9	0.3	0.9
			#5	0.3	0.0	0.6				1.9	2.0	2.4
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		Total	#1	6.1	0.6	1.3	8.3	1.1	1.0			
			#3	0.0	0.0	0.0	0.3	0.5	0.5			
			#4	14.9	3.6	4.3				20.6	1.8	1.4
			#5	11.6	2.1	1.1				14.9	0.0	1.2
			#6	0.0	0.0	0.0				0.0	0.0	0.0
Long cracking	Meters	Low	#1	3.7	1.2	2.0	3.9	0.0	3.4			
			#2	7.4	4.9	9.7	19.7	14.1	28.9			
			#3	0.2	0.0	0.5	8.3	0.0	34.2			
			#4	4.8	1.0	10.3				0.3	0.3	0.2
			#5	13.8	2.7	4.7				16.0	3.9	2.2
			#6	0.7	0.0	1.8				0.9	1.5	0.5
		Mod.	#1	6.3	2.8	2.4	5.0	1.9	17.1			
			#2	5.1	2.6	2.0	2.5	2.5	1.2			
			#3	0.0	0.0	0.0	0.0	0.0	0.1			
			#4	0.6	0.0	1.3				0.0	0.0	0.0
			#5	12.8	5.3	2.5				3.7	0.0	2.5
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		High	#1	0.0	0.0	0.0	9.1	0.0	3.8			
			#2	1.4	0.0	2.2	1.3	1.8	1.8			
			#3	0.0	0.0	0.0	0.0	0.0	0.0			
			#4	0.1	0.0	0.4				0.0	0.0	0.0

Distress	*Units*	*Sev.*	*Sec.*	*Expert*			*Individual*			*Team*		
				Mean	ϕ_a	ϕ_w	*Mean*	ϕ_a	ϕ_w	*Mean*	ϕ_a	ϕ_w
			#5	0.5	0.0	1.2				0.3	0.0	1.0
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		Total	#1	10.0	1.1	2.0	9.8	0.0	23.4			
			#2	14.0	7.6	10.8	23.5	16.9	29.2			
			#3	0.2	0.0	0.5	8.3	0.0	34.2			
			#4	5.5	0.9	10.4				0.3	0.3	0.2
			#5	27.0	0.6	4.1				20.0	4.3	0.8
			#6	0.7	0.0	1.8				0.9	1.5	0.5

In this analysis, CV values were determined by plotting both the within and between rater standard deviations versus mean for each distress type-severity level combination and fitting the best line through these data (y-intercept was forced through 0). The slope of this best fit line (in percentage terms) is a measure of the CV, assuming a linear increasing relationship between standard deviation and mean. Examples of these CV plots for different distress types and severity levels are shown in Figures 1 and 2. Figure 1 shows the plot for fatigue cracking in AC pavements, while Figure 2 shows a similar plot for corner breaks in PCC pavements. Two regression lines are shown on each plot, one for the within rater standard deviation and the other for the between rater standard deviation. The complete set of figures can be found elsewhere ("Study", 1998).

The resulting CV values and related statistics are summarized in Table VII for AC pavements and in Table VIII for PCC pavements. The following observations were made based on information presented in these two tables and the referenced CV plots:

- Both the between rater coefficient of variation (CVa) and the within rater coefficient of variation (CVw) values seem to vary widely, ranging from close to 0% to more than 300%. However, the larger CVs are primarily associated with those distress type-severity level combinations where the magnitude of distress is small and within a narrow range, as is the case for most high severity distresses. As shown in Tables VII and VIII, the high CV values were significantly reduced once data associated with high severity level distresses were removed from the analysis. The ranges were further reduced when only total quantities for a given distress type were considered, which is consistent with findings for manual distress data that indicated a problem in distinguishing distress at different severity levels (Rada et al., 1998).
- For AC pavements, when considering total distress quantities only, the average CVa values for experts, individual raters, and teams was 27, 34 and 22%, and the average CVw values for experts, individual raters, and teams was 16, 28 and 17%, respectively. Variations within this range appear reasonable for field measurements of this kind.
- For AC pavements, individual raters exhibited higher between and within rater variations as compared with those from the experts and teams, both of which had comparable results. Also, individual raters appear to have difficulty distinguishing between fatigue cracking (low severity) and longitudinal cracking in the wheel path; CVa and CVw values were 72 and 61% for total fatigue cracking quantities, respectively.
- For PCC pavements, when only total distress quantities are considered, the average CVa values for experts, individual raters, and teams were 13, 29 and 26%, and the average CVw values for experts, individual raters, and teams were 25, 100 and 20%, respectively. Variations within this range appear reasonable for field measurements of this kind, with the exception of within rater variability (CVw) for individual raters.
- For PCC pavements, individual raters exhibited higher between and within rater variations as compared with those from the experts and teams, both of which had comparable results. Also, unlike the

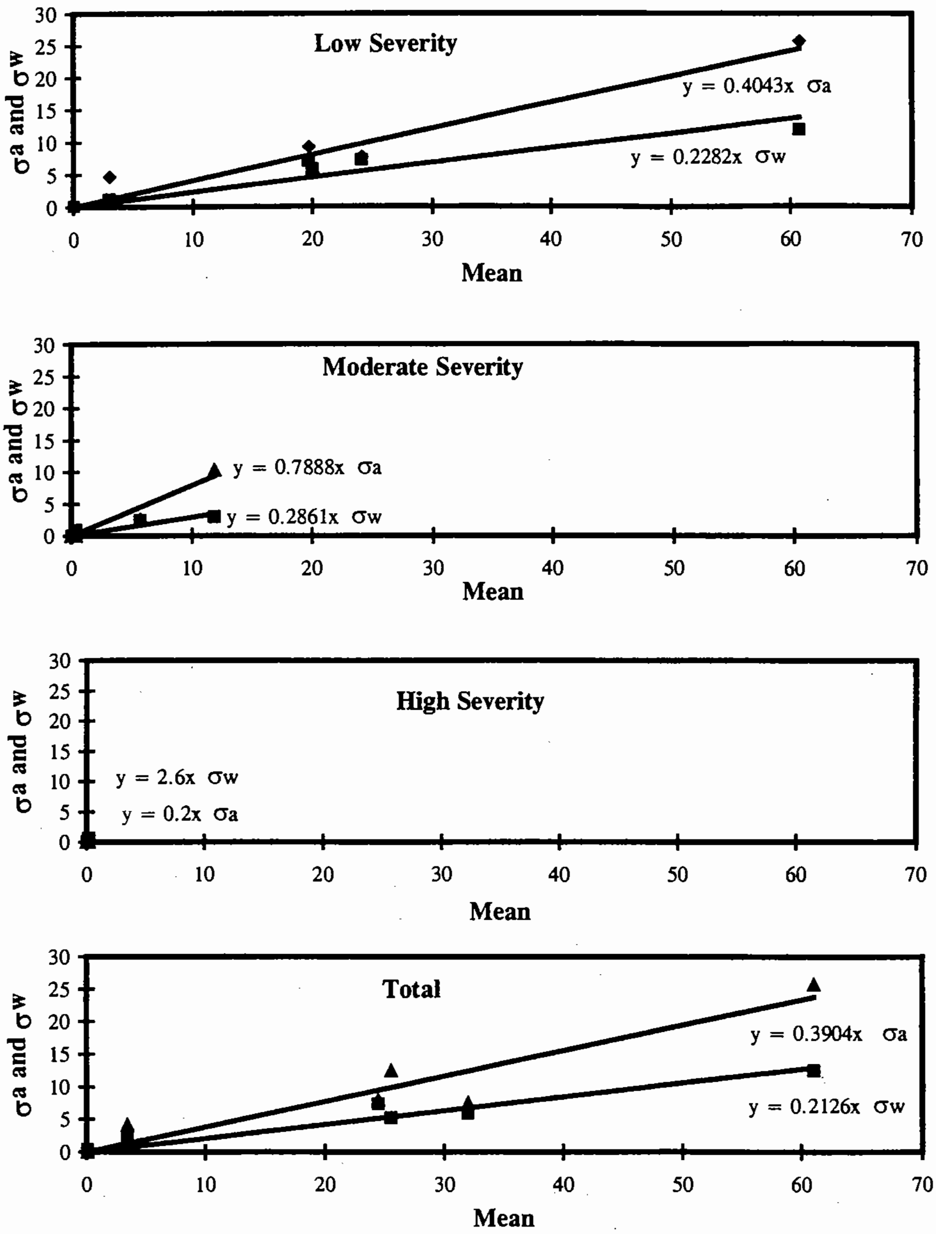


FIGURE 1 Function of σa and σw vs. mean for fatigue cracking of AC pavements, expert surveys

experts and teams, individual raters appear to have difficulty in consistently identifying joint-related distresses. For example, the within rater variation for individual raters was 235% for total quantity of transverse joint spalling, which indicates that additional and/or improved training is required to reduce this variability.

- In general, the results of this analysis appear to indicate that PASCO/PADIAS data variability can be improved (i.e., reduced) through additional

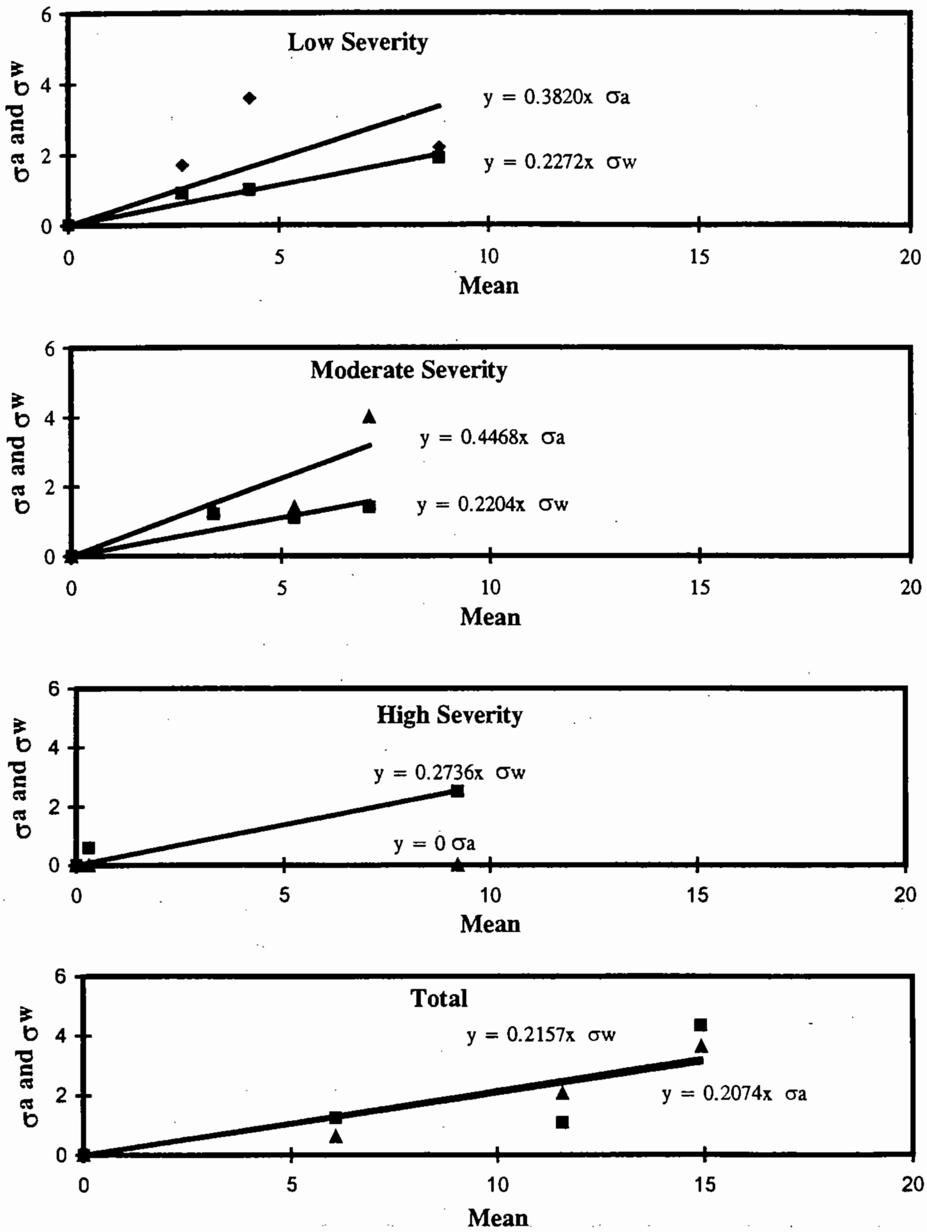


FIGURE 2 Function of σa and σw vs. Mean for corner breaks of PCC pavements, expert surveys

and/or improved rater training as well as through the use of two-person consensus survey teams. The referenced training should emphasize those distress types that were not easily quantified, such as joint spalling of PCC pavement and fatigue cracking of AC pavement. In addition, this training must look at ways of improving rater identification of the different severity levels for applicable distress types.

TABLE VII Coefficients of variation for experts, individuals and teams for AC pavements

Distress	*Units*	*Sev.*	*Expert*		*Individual*		*Team*	
			CVa	CV_w	CV_a	CV_w	CV_a	CV_w
Fatigue cracking	Sq. Meters	Low	40	23	90	56	93	28
		Mod.	79	29	33	133	97	22
		High	20	260	100	175	174	17
		Total	39	21	72	61	29	16
Long. cracking WP	Meters	Low	65	34	50	44	29	12
		Mod.	43	52	71	68	13	28
		High	0	217	66	169	100	220
		Total	55	31	37	29	28	13
Long. cracking NWP	Meters	Low	8	9	41	31	43	34
		Mod.	41	28	47	55	47	34
		High	46	88	116	84	102	118
		Total	7	9	38	29	32	28
Trans. cracking	Number	Low	25	15	19	17	20	17
		Mod.	33	31	30	43	19	24
		High	31	61	58	79	156	38
		Total	16	8	10	11	12	14
Trans. cracking	Meters	Low	24	20	17	18	19	16
		Mod.	20	22	26	49	13	30
		High	31	48	38	74	162	63
		Total	17	12	11	11	10	12
Summary Statistics								
All distress levels		Max.	79	260	116	175	174	220
		Min.	0	8	10	11	10	12
		Avg.	32	51	49	62	60	39
All distress levels except high severity level		Max.	79	52	90	133	97	34
		Min.	7	8	10	11	10	12
		Avg.	34	23	40	44	34	22
Total distress only		Max.	55	31	72	61	32	28
		Min.	7	8	10	11	10	12
		Avg.	27	16	34	28	22	17

Note:
CVa = among rater coefficient of variation
CVw = within rater coefficient of variation

TABLE VIII Coefficients of variation for experts, individuals and teams for PCC pavements

Distress	*Units*	*Sev.*	*Expert*		*Individual*			*Team*
			CVa	*CVw*	*CVa*	*CVw*	*CVa*	*CVw CVa CVw*
Corner Breaks	Sq. Meters	Low	38	23	78	62	9	14
		Mod.	45	22	58	32	13	21
		High	0	27	275	125	92	121
		Total	21	22	14	12	6	7
Long. Cracking	Meters	Low	29	69	59	183	25	14
		Mod.	43	25	50	283	0	68
		High	0	168	3	44	0	333
		Total	13	32	55	167	22	4
Trans. Cracking	Number	Low	4	6	15	18	38	15
		Mod.	64	67	103	151	36	45
		High	74	22	99	74	27	22
		Total	3	5	2	4	28	11
Trans. Cracking	Meters	Low	3	6	6 16	21	30	10
		Mod.	70	65	98	154	37	36
		High	72	24	100	73	28	22
		Total	6	4	7	5	21	4
Spalling of Long. Joints	Meters	Low	6	46	119	116	62	23
		Mod.	76	110	132	209	91	114
		High	160	183	218	192	0	300
		Total	26	20	124	128	61	15
Spalling of Trans Joints	Number	Low	3	63	4	296	17	89
		Mod.	74	76	0	167	0	151
		High	51	44	140	110	67	122
		Total	7	39	2	235	3	70
Spalling of Trans. Joints	Meters	Low	3	59	1	184	0	55
		Mod.	60	104	0	155	9	80
		High	57	76	119	85	75	56
		Total	18	52	0	150	43	28
Summary Statistics								
All distress levels		Max.	160	183	275	296	92	333
		Min.	0	4	0	4	0	4
		Avg.	37	52	68	123	30	66
All distress levels except high severity level		Max.	76	110	132	296	91	151
		Min.	3	4	0	4	0	4
		Avg.	29	44	45	130	26	42
Total distress only		Max.	26	52	124	235	61	70
		Min.	3	4	0	4	3	4
		Avg.	13	25	29	100	26	20

Note:
CVa = among rater coefficient of variation
CVw = within rater coefficient of variation

GLOBAL TRENDS

To gain a general understanding of the variability associated with LTPP distress data interpreted from film using the PADIAS v4.x system, plots of distress quantity at each severity level and total across all severity levels were developed for each of the common distress types identified in the 12 (six AC and six PCC) pavement test sections. Although each expert, individual rater and two-person team interpreted the distress films three times using the PADIAS v4.x system, such repeat measurements are not normal practice used in actual field distress data collection and interpretation. Consequently, only those data from the first interpretation were used to generate the referenced plots.

For a given distress type and severity level combination, the following values are plotted:

- **Reference Value --** Quantity of distress determined by the consensus manual distress survey conducted by four experts. Reference values were used as surrogates of "ground truth" data in this study.
- **Consensus value --** Quantity of distress determined by the consensus expert survey conducted using the PADIAS v4.x system.
- **Minimum, mean and maximum --** Distress quantities derived from first set (repetition) of distress data interpreted by the experts, individual raters and two-person teams using the PADIAS v4.x system.

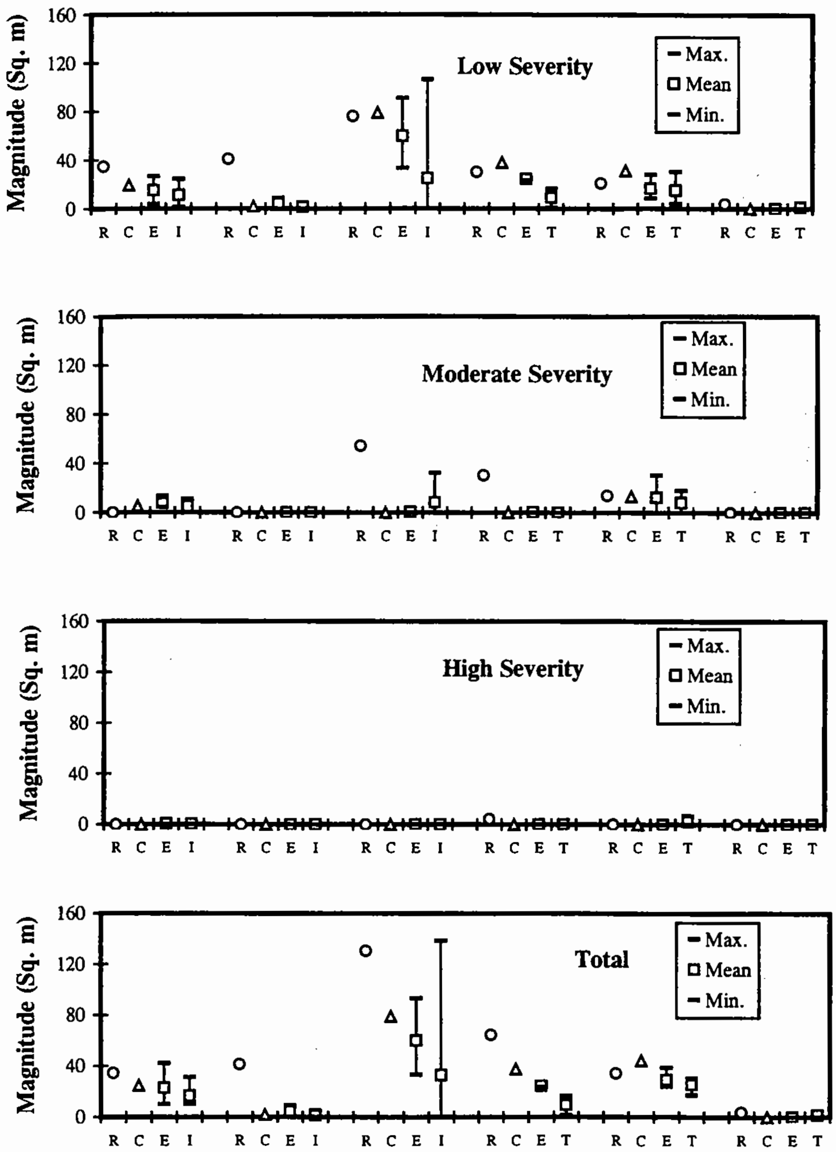
The complete set of figures showing global trends for AC and PCC pavement distress data is contained in a separate report ("Study", 1998). Example plots are given in Figures 3 and 4. Figure 3 shows the global trends for fatigue cracking in AC pavements, while Figure 4 shows a similar plot for corner breaks in PCC pavements. In these plots, the letters "R", "C", "E", "I" and "T" along the X-axis denote the values pertaining to the reference, consensus, expert, individual rater and two-person team surveys, respectively. (Note: for both AC and PCC pavements, test sections 1, 2 and 3 have been separated from 4, 5 and 6 in this comparison since individual raters only looked at the first three test sections and two-person team surveys were only performed on the latter three test sections.)

The following general observations were made from the information presented in these figures:

- There does not appear to be a significant positive or negative bias in the data; i.e., no tendency to consistently rate all distress type and severity level combinations higher or lower.
- For closely related distress types, such as fatigue cracking and longitudinal cracking in the wheel path, compensatory differences between the group ratings and reference values were observed, i.e., group ratings indicated a higher quantity of fatigue cracking and a lower quantity of longitudinal cracking as compared to the reference values.
- Overall, the two-person team surveys show the smallest level of variability, followed by the individual experts and individual raters, respectively.
- For AC pavement test sections, the consensus and group means are relatively close to one another for the various distress types; however, there are significant differences between these means and the reference values. Also, certain surface defects observed during the manual surveys, such as bleeding, could not be reliably identified with the PADIAS v4.x system.
- For PCC pavement sections, the reference, consensus and group means are generally close to one another for cracking related distresses; however, larger discrepancies exist for joint-related distresses and surface defects. Despite this, reference and consensus values generally fall within the range of minimum and maximum values observed for the different distress types.

BIAS AND PRECISION

One of the main objectives of the study was to quantify the bias and precision associated with LTPP distress data derived from film using the PASCO/PADIAS v4.x system. Towards that end, distress data collected by PASCO and interpreted using

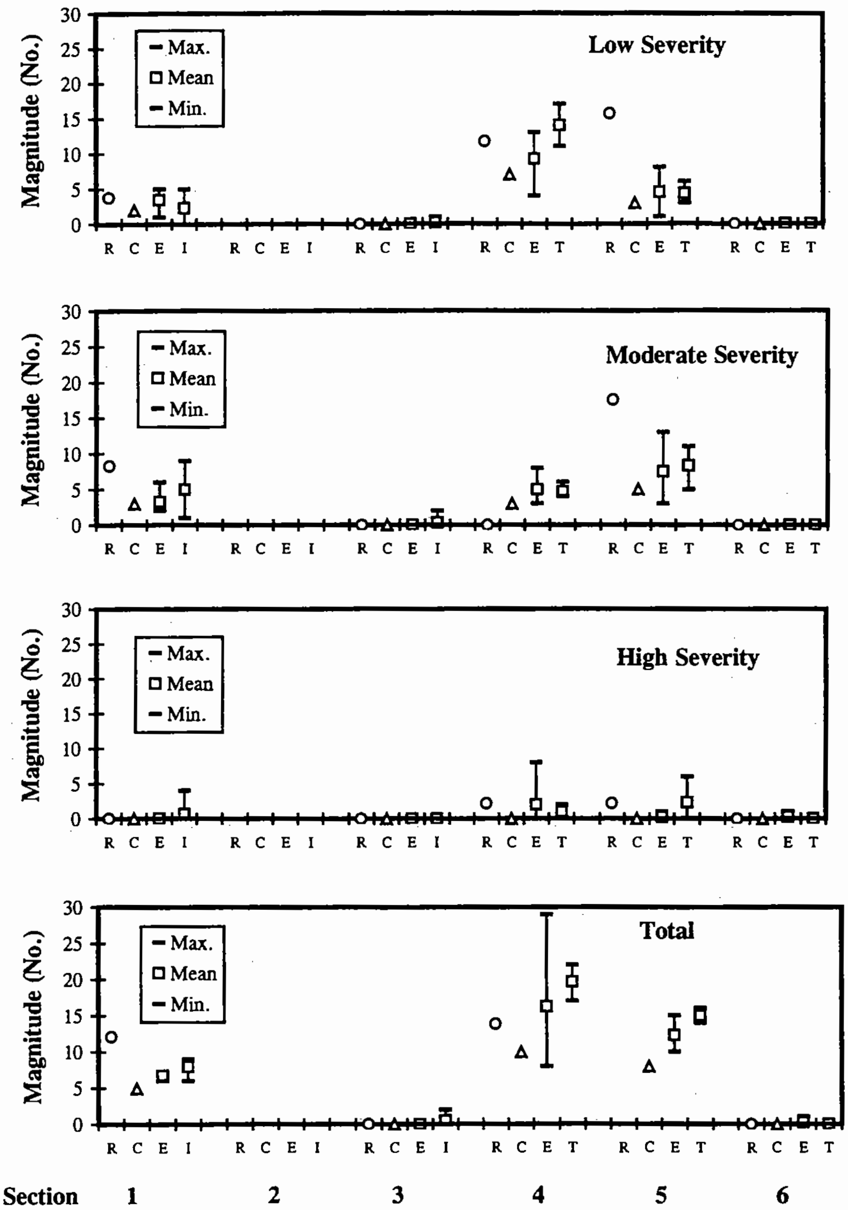


Section 1 2 3 4 5 6

Note: R = reference, C = consensus, E = expert, I = individual, T = team

FIGURE 3 Global trends, fatigue cracking of AC pavements

the PADIAS v4.x system were analyzed quantitatively. Specifically, an analysis of the coefficient of variation (CV) and root mean square of error (RMSE) associated with these data was first undertaken. Indicators of bias and precision were subsequently estimated for these data. Again, to simulate actual field conditions, only the first repeats of the distress data were used in the analyses.



Note: R = reference, C = consensus, E = expert, I = individual, T = team

FIGURE 4 Global trends, corner breaks of PCC pavements

Coefficient of Variation and RMSE Evaluation

As indicated earlier in this paper, the coefficient of variation (CV) is a statistical term normally used for representing the relative variability associated with experimental data. For the data in question, this value was determined by generating CV plots for different distress and severity level combinations as described in an earlier section of the paper. Example CV plots for different distress types and severity levels are shown in Figures 5 and 6. While Figure 5 shows the AC pavement fatigue cracking plot for the expert, Figure 6 shows the PCC pavement corner break plot for the expert. A complete set of CV plots has been presented elsewhere ("Study", 1998).

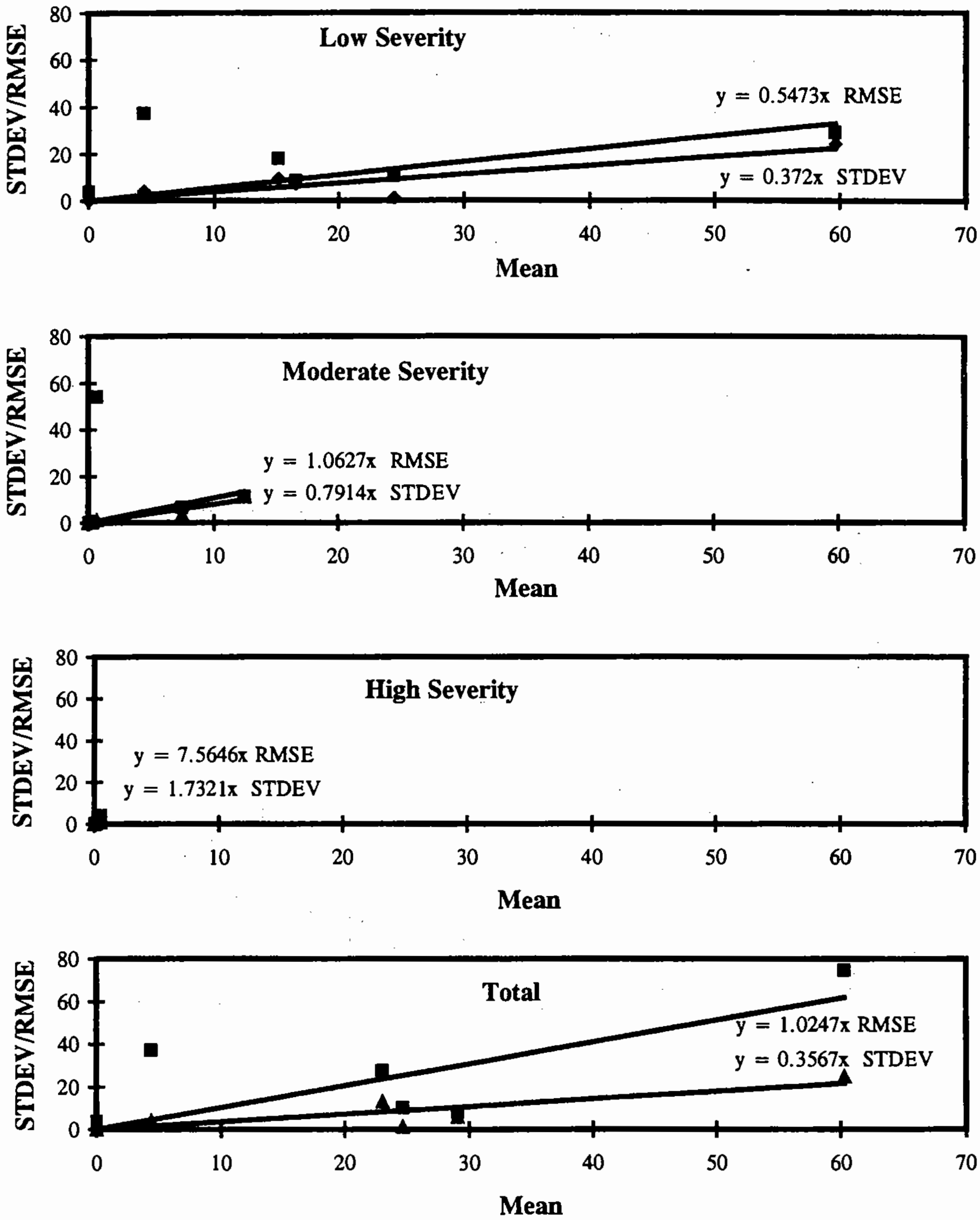


FIGURE 5 STDEV and RMSE vs. mean for fatigue cracking of AC pavements, experts

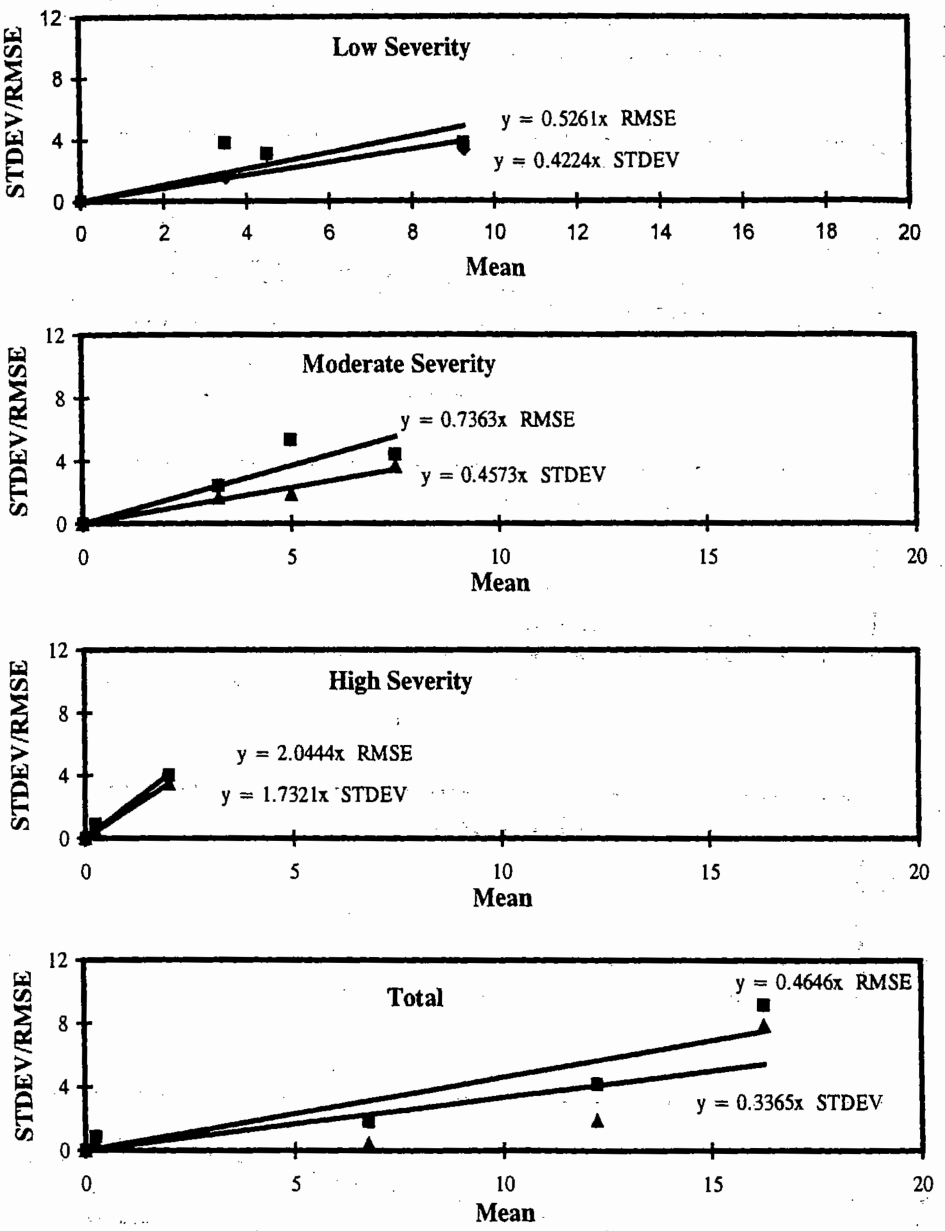


FIGURE 6 STDEV and RMSE vs. mean for corner breaks of PCC pavements, experts

Also included in the referenced figures are regression lines of RMSE (root mean square of error) versus mean, where RMSE is defined as the square root of the sum of the squared differences between reference and individual rater values divided by the number of raters in the workshop. This term combines the variability and bias associated with the different group data relative to the reference values; i.e., distribution of group values relative to the reference.

Based on those figures, the following observations were developed:

- CV values associated with the PASCO distress data range from less than 10% to more than 200%; however, the larger CVs are primarily associated with those distress type-severity level combinations where the magnitude of distress is small, as is the case for most high severity distresses.
- For total distress quantities, the average CV values for AC pavement distress data are 29, 22 and 70% for experts, teams, and individual raters, respectively. Those average values for PCC pavement distress data are 31%, 30% and 90% for experts, teams, and individual raters, respectively. Except for the individual rater CV values, these averages appear reasonable.
- A wide discrepancy was found between the STDEV and the RMSE regression lines, which implied that the difference between the group means and their corresponding reference values might be large. A more detailed analysis of the data will be given in next section.

General Assessment of Bias and Precision

To gain a general understanding of the PASCO distress data, overall bias and precision indicators were calculated for those data across various pavement test sections. A summary of results is presented in Tables IX and X for AC and PCC pavements, respectively. The following terminology was used in these two tables:

- **Pooled Values**: Average of reference and group means computed by pooling all ratings across all six AC or six PCC sections for experts, and across three AC and three PCC sections for individuals and teams.
- **Apparent Bias**: Difference between pooled reference value (from manual distress surveys) and pooled group mean.
- **Pooled Standard Deviation**: Product of pooled group mean and slope of best-fit line from standard deviation versus mean (CV) plots (not directly computed from distress data).
- **Pooled Root Mean Square of Error (RMSE)**: Derived from regression lines introduced in previous section (not directly calculated from the distress data).
- **Coefficient of Variation**: Slope of best straight-line fit (in percentage terms) from standard deviation versus mean plots described in previous section.

In looking at these data, it is important for the reader to understand that bias is being defined relative to the pooled reference values; i.e., pooled group mean minus the pooled reference value. For example, a positive bias indicates that a greater quantity of distress data was identified using the PASCO/PADIAS system as compared to that observed during the manual reference surveys and vice versa. Precision, on the other hand, is being defined in this paper as the variance about the group mean and not the reference value.

From Tables IX and X, the following observations were made:

- For AC pavements, CV values for total distress quantities range between 14 and 55% for experts, 12 and 29% for two-person teams, and 12 and 126% for individual raters. The average CV values for these three groups are 29, 22 and 70%, respectively. These results clearly show that the expert and two-person team surveys provided more consistent data as compared to the individual raters.
- Apparent bias (difference between group mean and reference value) for AC pavement distress data is generally large for all groups, but especially for the individual raters.
- For PCC pavements, CV values for total distress quantities range between 6 and 51% for experts, 9 and 59% for two-person teams, and 4 and 185% for individual raters. The average CV values for these groups are 31, 30 and 90%, respectively. It is also clear that the expert and two-person team surveys provide more consistent data as compared to the by individual raters, which indicates that improvements in distress data variability can be achieved through additional rater training or through the use of consensus team surveys.

TABLE IX Indicators of bias and precision for AC pavement distresses

Distress type	*Unit*	*Distress severity*	*Expert pooled values*						*Team pooled values*						*Individual pooled values*					
			REF	*Mean*	*STDEV*	*RMSE*	*COV*	*Bias*	*REF.*	*Mean*	*STDEV*	*RMSE*	*COV*	*Bias*	*REF.*	*Mean*	*STDEV*	*RMSE*	*COV*	*Bias*
Fatigue cracking	Meters	Low	34.5	20.0	7.4	11.0	37	–14.5	19.8	8.6	6.2	11.8	72	–11.2	49.2	12.4	17.6	31.4	141	–36.7
		Moderate	13.6	3.5	2.7	3.7	79	–10.1	4.6	2.7	2.5	3.1	93	–1.9	22.5	4.5	5.5	20.7	122	–18.0
		High	0.7	0.1	0.2	0.8	173	–0.6	0.0	0.8	1.1	1.3	141	0.8	1.4	0.1	0.2	1.4	224	–1.3
		Total	48.8	23.6	8.4	24.2	36	–25.2	24.4	12.1	3.5	8.6	29	–12.3	73.1	17.0	21.5	51.8	126	–56.1
Longitudinal cracking WP	Meters	Low	6.6	22.9	14.7	24.6	64	16.4	6.8	35.4	8.7	32.5	25	28.6	6.3	39.4	22.9	42.9	58	33. 1
		Moderate	2.6	4.2	2.6	4.3	62	1.7	0.0	6.0	4.7	7.6	79	6.0	5.1	17.6	13.4	19.4	76	12.5
		High	0.6	0.1	0.1	0.6	173	–0.5	0.0	0.1	0.1	0.1	141	0.1	1.1	4.2	4.7	5.6	112	3. 1
		Total	9.7	27.2	14.8	26.7	55	17.5	6.8	41.5	10.9	38.1	26	34.7	12.5	61.2	19.8	57.9	32	48.7
Longitudinal cracking NWP	Meters	Low	81.9	103.4	9.8	26.9	9	21.6	78.1	78.5	35.8	48.4	46	0.5	85.6	87.8	51.1	57.3	58	2.1
		Moderate	39.6	28.9	11.5	23.1	40	–10.6	60.8	31.5	11.9	39.6	38	–29.3	18.3	18.2	8.6	16.2	48	–0.1
		High	22.4	2.4	1.6	13.2	67	–20.0	20.0	4.4	4.5	15.4	103	–15.6	24.8	4.5	4.1	11.9	93	–20.3
		Total	143.8	134.7	19.4	34.5	14	–9.1	158.9	114.4	31.1	63.2	27	–44.4	128.7	110.4	57.2	83.1	52	–18.4
Transverse cracking	Number	Low	33.0	39.5	9.5	13.8	24	6.5	21.3	44.9	12.6	30.6	28	23.6	44.7	42.5	8.6	18.2	20	–2.2
		Moderate	9.3	14.4	5.6	8.4	39	5.1	3.3	6.9	3.0	5.0	43	3.6	15.3	17.2	3.9	5.9	23	1.9
		High	7.5	3.3	1.4	4.7	42	–4.2	2.3	1.1	1.3	1.6	121	–1.2	12.7	10.8	6.9	7.2	64	–1.9
		Total	49.8	57.3	9.2	16.9	16	7.4	27.0	52.9	9.1	32.1	17	25.9	72.7	70.5	8.6	21.8	12	–2.2
Transverse cracking	Meters	Low	37.4	55.3	18.7	24.3	34	17.9	30.7	47.6	12.2	24.0	26	16.9	49.2	12.4	17.6	31.4	141	–36.7
		Moderate	17.5	33.6	6.7	20.8	20	16.1	7.0	9.8	5.0	6.3	51	2.7	22.5	4.5	5.5	20.7	122	–18.0
		High	18.8	9.5	3.7	11.5	39	–9.3	3.2	2.3	3.1	3.1	131	–0.8	1.4	0.1	0.2	1.4	224	–1.3
		Total	73.6	98.3	24.4	34.8	25	24.7	40.9	59.7	7.2	24.9	12	18.8	73.1	17.0	21.5	51.8	126	–56. 1

Note:
Expert surveys were conducted for all the 12 test sections, while team and individual surveys were conducted for six of the 12 test sections.
Furthermore, team and individual surveys were performed on different test sections. The pooled reference valaues for the three surveys are therefore all different.

TABLE X Indicators of bias and precision for PCC pavement distresses

Distress type	Unit	Distress severity	Expert pooled values						Team pooled values						Individual pooled values					
			REF.	*Mean*	*STDEV*	*RMSE*	*COV*	*Bias*	*REF.*	*Mean*	*STDEV*	*RMSE*	*COV*	*Bias*	*REF.*	*Mean*	*STDEV*	*RMSE*	*COV*	*Bias*
Corner Breaks	Number	Low	3.2	3.5	1.5	1.8	42	0.3	5.3	6.1	1.1	1.7	18	0.8	0.0	1.3	0.9	1.5	69	1.3
		Moderate	5.0	3.2	1.4	2.3	46	−1.9	6.7	4.3	1.2	2.4	28	−2.3	2.5	2.7	1.4	1.4	51	0.2
		High	0.4	0.5	0.9	1.0	173	0.1	0.7	1.1	1.2	1.4	108	0.4	0.0	0.3	0.7	0.8	224	0.3
		Total	8.6	7.1	2.4	3.3	34	−1.5	12.7	11.6	1.0	1.3	9	−1.1	2.5	4.3	0.6	1.7	15	1.8
Longitudinal cracking	Meters	Low	8.6	7.5	5.1	5.6	68	−1.1	9.1	5.4	1.4	1.6	27	−3.7	8.1	16.2	32.3	34.5	199	8.2
		Moderate	4.3	3.8	1.4	2.6	37	−0.5	5.8	1.9	0.3	4.0	18	−3.9	2.8	5.3	10.7	11.0	201	2.6
		High	0.7	0.3	0.6	0.8	173	−0.4	1.5	0.0	N/A	N/A	N/A	−1.5	0.0	1.3	2.7	3.0	207	1.3
		Total	13.6	11.7	5.2	6.3	45	−1.9	16.4	7.3	1.2	5.5	16	−9.2	10.8	22.9	42.3	44.6	185	12.0
Transverse cracking	No.	Low	32.3	40.1	1.3	3.1	3	7.8	5.0	8.8	3.6	7.7	41	3.8	59.7	71.2	4.6	6.5	6	11.6
		Moderate	13.0	5.6	3.6	9.6	63	−7.4	5.3	2.1	0.9	3.5	40	−3.2	20.7	6.6	6.8	17.7	103	−14. 1
		High	3.7	3.1	1.8	1.9	58	−0.5	3.0	2.1	0.7	1.0	34	−0.9	4.3	2.8	3.5	3.8	125	−1.6
		Total	49.0	48.8	3.1	6.9	6	−0.2	13.3	13.0	3.6	4.9	28	−0.3	84.7	80.6	3.5	11.6	4	−4. 1
Transverse cracking	Meters	Low	114.6	142.2	14.9	19.8	10	27.6	9.4	21.4	7.5	18.9	35	12.0	219.7	249.0	25.4	28.0	10	29.3
		Moderate	45.4	21.4	13.1	32.3	61	−24.0	18.1	6.4	2.5	10.1	40	−11.7	72.7	23.3	23.3	63.1	100	−49.4
		High	13.0	11.1	6.5	6.9	59	−1.9	10.4	7.8	2.8	3.5	35	−2.6	15.6	9.1	11.2	12.9	123	−6.5
		Total	172.9	174.6	13.9	19.7	8	1.6	37.9	35.6	7.6	9.0	21	−2.3	308.0	281.3	24.5	49.1	9	−26.7
Spalling of longitudinal joints	Meters	Low	5.3	18.4	6.8	18.7	37	13.0	4.5	5.6	3.3	4.9	59	1.1	6.2	8.9	12.4	13.6	139	2.8
		Moderate	1.6	6.5	7.7	10.1	118	4.9	2.4	1.0	1.2	1.5	120	−1.3	0.8	4.0	6.4	7.6	160	3.2
		High	0.3	0.6	1.0	1.1	173	0.2	0.2	0.1	0.1	0.2	141	−0.1	0.4	3.2	7.2	7.9	223	2.8
		Total	7.3	25.4	10.3	26.8	40	18.2	7.1	6.8	4.0	6.0	59	−0.3	7.4	16.2	21.9	25.4	135	8.8
Spalling of transverse joints	No.	Low	0.8	0.3	0.2	0.3	77	−0.6	0.3	0.1	0.2	0.2	141	−0.2	1.5	2.8	4.9	5.4	172	1.3
		Moderate	0.4	0.6	0.5	0.7	92	0.2	0.3	0.1	0.2	0.3	141	−0.2	0.5	0.7	1. 1	1.1	160	0.2
		High	0.6	0.9	0.6	0.7	71	0.3	0.7	0.3	0.3	0.4	82	−0.3	0.5	0.3	0.4	0.6	158	−0.3
		Total	1.8	1.7	0.6	1.1	34	−0.2	1.3	0.6	0.2	0.3	28	−0.8	2.5	3.8	6.2	6.8	165	1.3
Spalling of transverse joints	Meters	Low	0.3	0.3	0.3	0.4	79	0.0	0.1	0.4	0.3	0.4	76	0.3	0.7	1.9	2.4	2.9	121	1.2
		Moderate	0.4	0.4	0.4	0.5	89	0.0	0.3	0.4	0.3	0.4	57	0.1	0.5	0.9	1.2	1.3	134	0.4
		High	0.5	1.7	1.2	1.8	74	1.2	0.6	1.5	0.9	1.6	64	0.9	0.4	0.3	0.5	0.6	145	0.0
		Total	1.2	2.4	1.2	2.2	51	1.2	1.0	2.3	1.1	2.0	51	1.3	1.6	3.2	3.8	4.5	120	1.6

Note:
Expert surveys were conducted for all the 12 test sections, while team and individual surveys were conducted for six of the 12 test sections. Furthermore, team and individual surveys were performed on different test sections. The pooled reference valaues for the three surveys are therefore all different.

- For PCC pavement total distress data, apparent bias is generally small for data interpreted by experts and teams, with the exception of some joint-related distresses; however, apparent bias is large for the individual raters.
- Precision associated with total distress quantities is significantly better than that for the individual distress severity levels; however, apparent bias does not exhibit any observable trends based on distress severity levels.
- Many of the above observations lead to the conclusion that improvements in PASCO distress data variability can be achieved through either additional rater training or the use of consensus team surveys.
- To more accurately quantify the bias and precision associated with PASCO/PADIAS distress data, an expanded experiment that includes more pavement test sections and covers a wider range of distress types and quantities is required.

SUMMARY AND CONCLUSIONS

Statistical Analyses

- From ANOVA, it appears that differences between raters within each group exist, especially for AC pavements. Raters seem to be better at identifying and distinguishing distress types and severity for PCC pavements than for AC pavements.
- Although not a standard practice used in distress data collection and interpretation, the repeatability (within rater variability) of data derived from film for a given test section by the same individual or group appears reasonable. The repeatability of the two-person teams appears to be better than that of individual experts and individual raters.
- Statistical comparison of group means (experts versus individual raters and experts versus two-person teams) indicates that there are no significant differences between them; however, to a large extent, this is due to the high within group variance which tends to mask the results.
- Distress data interpreted by two-person teams and experts show much less variability (both within rater and between rater) than those interpreted by individual raters.

Global Trends

- There does not appear to be a significant positive or negative bias in the data; i.e., no tendency to consistently rate all distress type and severity level combinations higher or lower.
- For closely related distress types, such as fatigue cracking and longitudinal cracking in the wheel path, compensatory differences between the group ratings and reference values were observed, i.e., group ratings indicated a higher quantity of fatigue cracking and a lower quantity of longitudinal cracking as compared to the reference values.

Precision and Bias

- Both the apparent bias and precision for the common distress type-severity level combinations were quantified. The apparent bias seems to be large for most cases and it is not uniform (no clear tendency). However, for PCC pavement, data obtained by experts and teams showed an acceptable bias for cracking related distress. The precision or variability for both AC and PCC distress data obtained by the experts and teams also appeared reasonable, but those associated with the individual raters had very large CV values.
- It is the authors′ opinion that the bias and precision of the PASCO/PADIAS distress data can be improved thorough additional, improved rater training and through the use of two-person consensus surveys. However, to truly quantify bias and precision, an expanded experiment that includes more pavement test sections and covers a wider range of distress types and quantities is required.