PB2000-100588

Study of LTPP Distress Data Variability, Volume I

PUBLICATION NO. FHWA-RD-99-074 [illegible]ER 1999

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

LTPP

FOREWORD

This report, which is part of a two-volume set, documents the results of a study to evaluate and quantify the variability of pavement distress data collected in the Long-Term Pavement Performance (LTPP) program. Analyses were performed on both manual and film-derived distress data. General trends of the distress data were first investigated, followed by statistical analyses of repeatability and detection of variability sources. Distress data bias and precision were also quantified. In addition, a comparison of variability between manual and film-derived distress data was conducted. This report will be of interest to engineers involved in pavement design, pavement performance evaluation and prediction, and pavement maintenance and rehabilitation.

Sufficient copies of this report are being distributed to provide two copies to each FHWA resource center and three copies to each FHWA division office and each State highway agency. Direct distribution is being made to the division offices. Additional copies for the public are available from the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, Virginia 22161.

T. Paul Teng, P.E.
Director
Office of Infrastructure
Research and Development

NOTICE

Technical Report Documentation Page

1. Report No. FHWA-RD-99-074	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle STUDY OF LTPP DISTRESS DATA VARIABILITY VOLUME I		5. Report Date September 1999
		6. Performing Organization Code
7. Author(s) G.R. Rada, C.L. Wu, R.K. Bhandari, A.R. Shekharan, G.E. Elkins, J.S. Miller		8. Performing Organization Report No.
9. Performing Organization Name and Address Law PCS A Division of LAW Engineering and Environmental Services, Inc. 12104 Indian Creek Court, Suite A Beltsville, Maryland 20705-1242		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTFH61-97-C-00002
12. Sponsoring Agency Name and Address Office of Infrastructure Research and Development Federal Highway Administration 6300 Georgetown Pike McLean, Virginia 22011-2296		13. Type of Report and Period Covered Final Report January 1997 - December 1998
		14. Sponsoring Agency Code
15. Supplementary Notes Contracting Officer's Technical Representative - Monte Symons, HRDI-13		

16. Abstract

Reliable distress data for pavement performance model development and validation, and other pavement engineering products, are critical to the success of the Long-Term Pavement Performance (LTPP) program. Confidence in distress data requires a measure of error because of the bias and precision components of its variability. No systematic evaluation has been performed to quantify the bias and variability associated with both the manual and PASCO film-based distress data. In view of this, this study was undertaken by the Federal Highway Administration (FHWA) to assess the variability of the LTPP distress data, including those in the Information Management System (IMS) and those currently being collected using either photographic or manual methods.

This main volume of the report contains sources of data used in the analyses, evaluation of manual distress data, evaluation of film-derived distress data, comparisons of data obtained from these two methods, and conclusions and recommendations.

This is Volume I of the two-volume report. Volume II is:

FHWA-RD-99-075 Volume II Appendix A: Tables and Figures for Manual Distress Data Analysis
Appendix B: Figures for PASCO/PADIAS Distress Data Analysis
Appendix C: Figures for Comparison of Manual PASCO/PADIAS Distress Data

17. Key Words Pavement Distress; Pavement Evaluation; Manual Survey; Film-Derived Distress Data; Distress Data Variability; Bias and Precision; LTPP		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 147	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yards	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9 or (F-32)/1.8	Celcius temperature	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

NOTE: Volumes greater than 1000 l shall be shown in m^3.

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.71	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C + 32	Fahrenheit temperature	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

* SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.

(Revised September 1993)

TABLE OF CONTENTS

VOLUME I

Chapter | Page

1. INTRODUCTION 1

  1.1 Background 1
  1.2 Study Objectives 4
  1.3 Data Sources 5
  1.4 Report Overview 8

2. MANUAL DISTRESS DATA VARIABILITY 9

  2.1 Introduction 9
  2.2 Background 9
  2.3 Data Source 10
  2.4 Global Trends 11
  2.5 Impact of Re-Accreditation 12
  2.6 Pavement Condition Index 17
  2.7 Bias and Precision 18
  2.8 Individual Versus Two-Person Team Distress Surveys 38
  2.9 Summary and Conclusions 47

3. PASCO/PADIAS DISTRESS DATA VARIABILITY 49

  3.1 Introduction 49
  3.2 Background 49
  3.3 Data Source 50
  3.4 PADIAS v4.x Distress Data Variability Study 51
  3.5 Global Trends 80
  3.6 Bias and Precision 86
  3.7 Comparison of PADIAS v1.x Versus v4.x Distress Data 97
  3.8 Analysis of Data Interpreted by PASCO's Production Procedure 106
  3.9 Summary and Conclusions 111

4. MANUAL VERSUS PASCO/PADIAS: A COMPARATIVE STUDY 113

  4.1 Introduction 113
  4.2 General Assessment 113
  4.3 Head-to-Head Comparison of Manual and PASCO/PADIAS 113
  4.4 Field Versus Photographic Interpretation by Reference Raters 122
  4.5 Summary and Conclusions 131

TABLE OF CONTENTS (continued)

Chapter | Page

5. CONCLUSIONS AND RECOMMENDATIONS 133

5.1 Conclusions 133
5.2 Recommendations 135

REFERENCES 137

LIST OF TABLES

VOLUME I

Table Page

1. Distribution of Data Sets Used for Comparison of Distress Rating Methods. 6
2. Summary of Accreditation Workshops. 10
3. Group Statistics for Raters Attending Workshop Nos. 1, 2, and 4; AC Pavements. 19
4. Group Statistics for Raters Attending Workshop Nos. 1, 2, and 4; PCC Pavements. 20
5. Pavement Condition Index (PCI) for AC Pavements.. 21
6. Indicators of Bias and Precision for AC Pavement Distresses.. 30
7. Indicators of Bias and Precision for PCC Pavement Distresses. 31
8. Results of Bias Analysis Based on Square Root Transformation for AC Pavements. 39
9. Results of Bias Analysis Based on Square Root Transformation for PCC Pavements. 41
10. Test Sections for PADIAS v1.x vs. PADIAS v4.x Comparison. 52
11. Calculated F-Statistics for AC Sections - Within Group.. 59
12. Calculated F-Statistics for PCC Sections - Within Group.. 60
13. Results of Analysis of Variance (ANOVA) for AC Sections - Within Group. . . . 61
14. Results of Analysis of Variance (ANOVA) for PCC Sections - Within Group. . . 62
15. Results of t-test for AC Pavement Sections.. 64
16. Results of t-test for PCC Pavement Sections.. 65
17. Mean, σa, and σw Values for Experts, Individuals, and Teams for AC Pavements; PASCO Method. 67
18. Mean, σa, and σw Values for Experts, Individuals, and Teams for PCC Pavements; PASCO Method. 70
19. CVa and CVw for Experts, Individuals, and Teams for Asphalt Pavements... . . 78
20. CVa and CVw for Experts, Individuals, and Teams for Concrete Pavements.. . . 79
21. Indicators of Bias and Precision for AC Pavement Distresses - PASCO Method. 94
22. Indicators of Bias and Precision for PCC Pavement Distresses - PASCO Method. 95
23. Comparison of PADIAS v1.x and v4.x, t-test, AC Pavement 98
24. Comparison of PADIAS v1.x and v4.x, t-test, JCP Pavement. 99
25. Comparison of PADIAS v1.x and v4.x, t-test, CRC Pavement 101
26. Relative Indicators of Bias and Precision for AC Pavement Distresses, Individual Raters. 114
27. Relative Indicators of Bias and Precision for PCC Pavement Distresses, Individual Raters. 115

LIST OF TABLES (continued)

Table — Page

28. Apparent Bias: Manual Vs. PASCO/PADIAS, AC Pavements, Individual Raters 119

29. Apparent Bias: Manual Vs. PASCO/PADIAS, PCC Pavements, Individual Raters 120

LIST OF FIGURES

VOLUME I

Figure Page

1. Fatigue Cracking (Sq. Meters) - AC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values. 13
2. Longitudinal Cracking WP (Meters) - AC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values. 14
3. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values. 15
4. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values. 16
5. Pavement Condition Index (PCI) for AC Pavements.. 22
6. Fatigue Cracking (Sq. Meters) - AC Pavements: Standard Deviation/RMSE Vs. Mean. 26
7. Longitudinal Cracking WP (Meters) - AC Pavements: Standard Deviation/RMSE Vs. Mean. 27
8. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Standard Deviation/RMSE Vs. Mean. 28
9. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Standard Deviation/RMSE Vs. Mean. 29
10. Fatigue Cracking (Sq. Meters) - AC Pavements: Standard Deviation Vs. Mean. . 34
11. Longitudinal Cracking WP (Meters) - AC Pavements: Standard Deviation Vs. Mean. 35
12. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Standard Deviation Vs. Mean. 36
13. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Standard Deviation Vs. Mean. 37
14. Fatigue Cracking (Sq. Meters) - AC Pavements, Manual Surveys: Reference, Team, and Individual Values. 43
15. Longitudinal Cracking NWP (Meters) - AC Pavements, Manual Surveys: Reference, Team, and Individual Values. 44
16. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Reference, Team, and Individual Values. 45
17. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Reference, Team, and Individual Values. 46
18. Fatigue Cracking (Sq. Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Individual Raters for Three Repetitions. 54
19. Fatigue Cracking (Sq. Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Teams for Three Repetitions. 55

LIST OF FIGURES (continued)

Figure Page

20. Corner Breaks (No.) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Individual Raters for Three Repetitions. 56
21. Corner Breaks (No.) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Teams for Three Repetitions. 57
22. Fatigue Cracking (Sq. Meters) - AC Pavements, Experts, PASCO Method: σa and σw Vs. Mean. 74
23. Longitudinal Cracking, WP (Meters) - AC Pavements, Experts, PASCO Method: σa and σw Vs. Mean. 75
24. Corner Breaks (No.) - PCC Pavements, Experts, PASCO Method: σa and σw Vs. Mean. 76
25. Longitudinal Cracking (Meters) - PCC Pavements, Experts, PASCO Method: σa and σw Vs. Mean. 77
26. Fatigue Cracking (Sq. Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals, & Team. 82
27. Longitudinal Cracking WP (Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals, & Teams. 83
28. Corner Breaks (No.) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals & Teams. 84
29. Longitudinal Cracking (Meters) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals, & Teams. 85
30. Fatigue Cracking (Sq. Meters) - AC Pavements, Expert Raters, PASCO Method: Standard Deviation/RMSE Vs. Mean. 87
31. Fatigue Cracking (Sq. Meters) - AC Pavements, Individual Raters, ... PASCO Method: Standard Deviation/RMSE Vs. Mean. 88
32. Fatigue Cracking (Sq. Meters) - AC Pavements, Team Surveys, Standard Deviation/RMSE Vs. Mean. 89
33. Corner Breaks (No.) - PCC Pavements, Expert Raters, PASCO Method: Standard Deviation/RMSE Vs. Mean. 90
34. Corner Breaks (No.) - PCC Pavements, Individual Raters, PASCO Method: Standard Deviation/RMSE Vs. Mean. 91
35. Corner Breaks (No.) - PCC Pavements, Team Surveys, PASCO Method: Standard Deviation/RMSE Vs. Mean. 92
36. Comparison of PADIAS v1.x and v4.x - AC Pavements, Transverse Cracking (Meters). 103
37. Comparison of PADIAS v1.x and v4.x - JPC Pavements, Transverse Cracking (Meters). 104
38. Comparison of PADIAS v1.x and v4.x - CRC Pavements, Transverse Cracking (Meters). 105

LIST OF FIGURES (Continued)

Figure | Page

39. Team Survey Vs. Reference and Consensus: Fatigue Cracking of AC Pavement, PASCO's Production Procedure. 107
40. Individual Survey Vs. Reference and Consensus: Fatigue Cracking of AC Pavement, PASCO's Production Procedure. 108
41. Team Survey Vs. Reference and Consensus: Transverse Cracking (Meters) Of PCC Pavement, PASCO's Production Procedure. 109
42. Individual Survey vs. Reference and Consensus: Transverse Cracking (Meters) of PCC Pavement, PASCO's Production Procedure. 110
43. Fatigue Cracking (Sq. Meters) - AC Pavements: Reference, Minimum Mean, and Maximum Values of Individual Surveys for Manual and PASCO/PADIAS., ... 116
44. Fatigue Cracking (Sq. Meters) - AC Pavements: Reference, Minimum, Mean, and Maximum Values of Team Surveys for Manual and PASCO/PADIAS. 116
45. Corner Breaks (No.) - PCC Pavements: Reference, Minimum, Mean, and Maximum Values of Individual Surveys for Manual and PASCO/PADIAS. 118
46. Corner Breaks (No.) - PCC Pavements: Reference, Minimum, Mean, and Maximum Values of Team Surveys for Manual and PASCO/PADIAS. 118
47. Fatigue Cracking (Sq. Meters) - AC Pavements: Field and PASCO/PADIAS Ratings by Reference Group. 123
48. Longitudinal Cracking WP (Meters) - AC Pavements: Field and PASCO/PADIAS Ratings by Reference Group. 124
49. Corner Breaks (No.) - PCC Pavements: Field and PASCO/PADIAS Ratings by Reference Group. 125
50. Longitudinal Cracking (Meters) - PCC Pavements: Field and PASCO/PADIAS Ratings by Reference Group. 126
51. Field Vs. Photographic Ratings by Reference Group, AC Pavements. 127
52. Field Vs. Photographic Ratings by Reference Group, PCC Pavements. 129

1. INTRODUCTION

1.1 Background

One goal of the Long-Term Pavement Performance (LTPP) program is to provide data necessary to improve one of the major problem areas in pavement design and management: performance prediction. A rational system, based on theory validated with field observations, is the eventual objective. The key to this objective is modeling pavement damage by distress type, particularly the development and propagation of critical distresses as a function of traffic, pavement structure, environment, material properties, etc. These type of distress-specific models can only be developed from good quality pavement distress measurements, i.e., good in terms of accuracy, precision, and reliability. Many of the other engineering products to be developed with use of LTPP data are also directly dependent upon the quality of its distress data. In short, **Quality Distress Data are Critical to the Success of the LTPP Program.**

Distress data collection for LTPP began with the decision to use 35-mm black and white photography to obtain frequent images of the surface conditions of the pavement test sections. The PASCO Corporation was selected as the filming contractor. Actual distress data, in terms of distress types, severity levels, and quantities were to be obtained through a film interpretation process conducted after the filming event. As a backup, manual distress surveys were to be performed for test sections where the PASCO vehicle could not travel or where scheduling conflicts required. By policy, PASCO was the primary means to collect distress data for LTPP. In its contract, PASCO was to film LTPP test sections on a schedule intended to provide coverage of each section on a two-year cycle. Some specific exceptions to this schedule were made to obtain before and after photos of the SPS-3 maintenance effectiveness test sections and for GPS-6 and -7 sections where overlays were to be placed.

The distress types and methods to identify and record distress were established in guidelines and procedures issued by the Strategic Highway Research Program (SHRP). The Distress Identification Manual (DIM) was first developed by SHRP contractors and issued as a draft in 1989. This draft DIM was used by PASCO in setting up the processes in its Pavement Distress Analysis System (PADIAS), v1.x software. After review comments were received and implemented by SHRP, the first DIM was published in October 1990. This manual was used by the SHRP Technical Assistance Contractor for the initial film interpretation, which began in the fall of 1990. As the DIM differed slightly from the methodology implemented in the PADIAS, some alterations to the software were made to make it better conform to this version. As film interpretation progressed, it became apparent to the operators and the quality control (QC) reviewers that significant deficiencies existed in the 1990 DIM. Essentially, the imprecise language created difficulties in maintaining consistent interpretations between operators; descriptions and methods of measurement were confusing. As a result, a workshop was held in Arlington, Texas, during April 1991 that was attended by several representatives from each Regional Coordination Office Contractor (RCOC). At this workshop the DIM was reviewed in detail along with field exercises in order to refine and clarify the descriptions, procedures, and intent. Recommendations from this workshop were quite numerous and were adopted by SHRP as corrections to the published manual for internal LTPP use. The SHRP P-001B Technical Assistance Contractor used these corrections as the basis for interpreting or reinterpreting (previously interpreted) film from the following periods:

- SHRP: Round 1 - Fall 1989 through summer 1990 (GPS and SPS)
- SHRP: Round 2 - Fall 1990 through summer 1991 (GPS and SPS)
- SHRP: Round 3 - Fall 1991 through summer 1992 (SPS 3 & 4 and associated GPS sections)

The film interpretation process consisted of the following:

- PASCO filmed a section, developed the film, and performed quality checks for image clarity, contrast, completeness, etc.
- A positive print (one of several produced) was forwarded to the SHRP Technical Assistance Contractor. This film was then reviewed to assess its quality and conformance to SHRP requirements.
- The film was interpreted by the SHRP Technical Assistance Contractor using PASCO's PADIAS v1.x software. Operators who were trained in-house by the Technical Assistance Contractor performed the interpretation.
- Data obtained from this interpretation were forwarded to the RCOC's for review based on their knowledge of the section and to compare against their copy of the film. The RCOC's conducted their review of the film using a large-scale image projected onto a viewing screen.
- Once the RCOC review comments were received and assessed by the SHRP Technical Assistance Contractor against the film, changes were made only where the "missed" distress could be seen in the PADIAS system. Those distresses that were observed by the RCOC's from the magnified image, but could not be seen in the smaller PADIAS image, were **not** recorded.
- When corrections were completed the data were forwarded to the RCOC for upload into the Information Management System (IMS).

The comments from the RCOC reviews demonstrated that more information was available from the film than was contained in the data from the PADIAS interpretation. The RCOC review process used a large projection of the film, which was compared with a map report generated from the PADIAS system. The PADIAS system used an image with 0.30 m representing approximately 3.66 m of pavements, while some RCOC's were using systems with 0.30 m representing approximately 0.90 m of pavements. This magnification difference was obviously a source of error contributing to the variability and completeness in this data. A quantitative assessment of this difference was not performed.

Concerns over the issues of data completeness, interpretation differences, and quality of the film-derived data have persisted since the first RCOC reviews. Some of the technical issues that impact successful film interpretation include:

- **Film quality and limitations.** Film quality varies due to many factors. Primarily, the intensity and consistency of illumination across the image can differ significantly and mask some defects. In other cases the contrast may vary so that cracks are "highlighted" in a test section in one filming event but may not be visible at all in the previous or following event, with no maintenance or rehabilitation having been applied to the section between events. Other issues are less important, such as the swirling effects caused by the camera boom oscillating on a rough pavement or the wander of the vehicle. Film quality does not include the resolution limitation inherent in the film. Crack widths of approximately 2.5 mm down to 1.5 mm cannot be consistently detected because of variable lighting conditions, surface moisture, pavement texture, and film defects. Since low severity cracking is characterized as cracks up to 6 mm [3 mm for portland cement concrete (PCC) pavements], it is obvious that only a portion of the low severity cracking can reliably be detected.

- **Image size**. Aside from the film limitations, image size is the biggest single shortcoming evidenced in data reduction from film using the PASCO supplied system. This system was relatively crude in the resolution of measurements: 0.30-m increments for linear defects and 0.09-sq-m increments for areal distresses. During the reduction process performed for SHRP, the data review process called for the regional contractors to project the film onto a screen and compare their findings to the report (summary and map) generated from the PADIAS reduction. It was found by all regional contractors that many low severity distresses were not being detected using the PADIAS system. The large projected images showed much more detail.

- **Operator variability.** At the time the majority of the data reduction effort was undertaken, there was no LTPP Distress Rater Accreditation program and the 1993 DIM was not available. Each of the operators was given individual training in the use of the equipment and the distress identification. A QC process was implemented to assess the performance of these operators and provide a consistent review of the data.

It is important to note that these sources of variation were not systematically assessed and, therefore, no measurement of variability in the data was developed. However, the qualitative assessment of these data was that variability is high and completeness and repeatability poor.

While the film interpretation work continued, effort was undertaken to revise the DIM and publish a new version incorporating the recommended changes developed from the 1991 workshop. In addition, the concept of calibrating raters across all regions was adopted; calibration meaning training raters to improve consistency in identifying and quantifying distresses. The idea of rater calibration was developed into the Rater Accreditation Workshop, which has become the sole means of implementing quality assurance for manual surveys. This workshop approach began in 1992 with a pilot effort aimed both at completing revisions to the DIM and refining the workshop training methods for general use in accrediting LTPP raters. Implementation of the revised DIM dates to the summer of 1992, as this information and reference material were provided to the raters attending the three "production" workshops conducted that summer. The final version of the DIM was eventually published in May 1993.

Quality control for manual surveys by the RCOC's relied on a mandated policy where:

- Surveys were only to be performed by accredited raters.
- An office review of completed surveys must be performed by an accredited rater.

During the first half of 1992, when LTPP was being transitioned to FHWA, the SHRP contract with PASCO ended. FHWA awarded PASCO a contract to continue the filming process and to provide distress data interpreted from the films. During the review of the analysis system proposed for use by PASCO, it was determined that the data from the small projection provided by the Film Motion Analyzer (FMA) subsystem was not sufficiently repeatable and possibly not sufficiently accurate for the purposes of LTPP. As a result, the interpretation task of that contract was not activated. To provide distress data, the use of manual surveys increased significantly in 1992. The SHRP P-338 report *Distress Identification Manual for the Long-Term Pavement Performance Project* was declared the standard for all LTPP distress data.

Notwithstanding the implemented quality assurance process, assessment of the variability of manually collected data from the accreditation workshops was performed, and the results were not encouraging. One major problem area was surface defects. Detection of these distresses had been underemphasized in workshops because of the need to concentrate on the major distresses (fatigue, longitudinal cracking); therefore, some raters did not rate these types of distress. However, even for some of the major distresses, the variability between single raters was found to be very high, with significant differences from the group mean and reference values. Because of this variability, the use of two-person (or more) teams to perform group consensus surveys was proposed to reduce this variability and to improve consistency. Assessment of two-person consensus surveys was performed and also found to exhibit unacceptably high variability. Review of the distress time series data in the IMS has revealed that, on some sections, illogical year-to-year distress patterns exist that prevent their use in pavement performance analysis.

Because 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, FHWA decided in 1996 to proceed with data reduction from this film using the improved PASCO v4.x software developed by PASCO USA. PADIAS v4.x incorporates the 1993 DIM procedures, with minor exceptions, and is vector based to improve its recording precision. Reduction of these data is currently being performed under the LTPP Data Analysis Technical Assistance Contract (DATS).

1.2 Study Objectives

Reliable distress data for pavement performance model development and validation, and other pavement engineering products, are critical to the success of the LTPP program. Confidence in distress data requires a measure of error because of the bias and precision components of its variability. In turn, measuring these parameters requires evaluating many potential methods and comparing results. As indicated earlier, distress data in the LTPP program consists of both film-derived (PASCO v1.x and v4.x) and manually collected data. Consideration has also been given to enhancing film interpretation with large-screen projection to improve the ability to see smaller distress artifacts. However, in all cases, distress data quality is unknown. The uncertainty as to the correctness of data derived from film (PASCO v1.x) during the early years of the LTPP program and the general lack of knowledge concerning the quality of manually collected data or those presently being derived using PASCO v4.x are serious issues that must be addressed.

In view of this, FHWA undertook a study to assess the variability of the LTPP distress data, including those in the IMS and those currently being collected using either photographic or manual methods. Accomplishing the following objectives would not only go a long way toward achieving the goals of this study, but also would provide a better picture of the issues affecting LTPP distress data to allow for knowledge-based decision making in the future:

1. Assess variability of manual distress data.

2. Assess variability of distress data from film using the PADIAS v4.x system with the current FMA subsystem (small projection).

3. Compare distress data from film using the PADIAS v4.x system to that derived from the PADIAS v1.x system.

4. Assess the agreement of manual survey results to those from film using the PADIAS v4.x system.

The data gathered in support of these objectives are summarized in the next section. The data used for the comparison of the different distress interpretation methods are summarized in table 1 and are described in more detail below as part of the study data sources.

1.3 Data Sources

Manual Distress Data

To achieve the first study objective, assess variability of manual distress data, results from nine LTPP manual distress rater accreditation workshops were used. This data source provided a total of 119 individual manual distress ratings on 18 accreditation pavement test sections [nine on asphaltic concrete (AC) and nine on PCC test sections]; 11 to 16 individual raters per workshop performed the ratings on the same day on the same test sections. Reference surveys of these 18 test sections were also conducted by the workshop instructors, referred to as "expert" raters in this report for convenience, immediately prior to each workshop using a consensus rating method; distress data from these surveys were used as a surrogate for ground truth data in the study. The last two workshops (Reno 1996 and Champaign 1996) also included two-person team consensus surveys to investigate potential improvements in consistency versus individual rater surveys.

Table 1. Distribution of Data Sets Used for Comparison of Distress Rating Methods.

Interpretation Method	Rater	Test Sections											
		AC 1	AC 2	AC 3	AC 4	AC 5	AC 6	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
Manual	Reference	1	1	1	1	1	1	1	1	1	1	1	1
	Individual Raters	16					6	16					6
	Two-Person Team	1					3	1					3
PADIAS	Expert Consensus	1	1	1	1	1	1	1	1	1	1	1	1
	Individual Expert	4	4	4	4	4	4	4	4	4	4	4	4
	Individual Raters	6	6	6				6	6	6			
	Two-Person Team				3	3	3				3	3	3

PADIAS v4.x Distress Data

To obtain data to assess distress data interpreted from film using PADIAS v4.x system, PASCO filmed some of the test sections used in the distress rater accreditation workshops. During the last two manual distress accreditation workshops (Reno 1996 and Champaign 1996), reference surveys were conducted by the expert group on all six test sections used in each workshop (three AC and three PCC test sections, not just the two accreditation test sections), which yielded reference distress data for a total of 12 test sections (6 per workshop). Each of the 12 test sections was also filmed by PASCO approximately one month before the workshop.

Using the PADIAS v4.x software, a group effort was performed to establish consensus values for the film-derived distress data. (Note: these values were assumed to be the best possible set of values for distress data derived from film using PADIAS v4.x; however, manual reference survey data were still considered ground truth). This work was accomplished using film analysis equipment and software located at PASCO USA, Harrisburg PA. The group consisted of expert raters, two of whom were accreditation workshop instructors, who performed the data reduction through a consensus effort by observing the films at the same time, discussing distress types and severity levels, and then mutually deciding on the rating to be assigned.

On completion, distress data reduction from film was accomplished by six other individual raters from PASCO who were responsible for production work. These individuals independently performed distress data interpretation from the film for 6 of the 12 test sections. (Note: the original plan called for interpretation of all 12 sections by each rater, but funding limitations restricted the effort to 6 sections per rater). Three repeat interpretations were performed on each section by each individual rater.

The same individual raters were then paired into three, two-person, survey teams. Interpretations by these two-person teams were performed on the six test sections not previously interpreted by the individual raters. Although the original plan called for interpretation of all 12 sections by the rater teams, funding limitations restricted this effort to six sections per team. This is unfortunate since it does not allow for a direct comparison of individual versus two-person team distress data derived from film using PADIAS v4.x. Three repeat interpretations were also performed on each of these sections by the two-person teams.

Thus, the data produced to quantify the variability of distress data from film using PADIAS v4.x consisted of:

- Consensus surveys for 12 test sections obtained from film (PADIAS v4.x) interpretation methods.

- Data collected independently by individuals from film (PADIAS v4.x) interpretations for six test sections.

- Data collected by two-person teams in consensus surveys from film (PADIAS v4.x) interpretations for six test sections (different from those used in individual ratings).

PADIAS v1.x versus v4.x Distress Data

To undertake the comparison of distress data from film using the PADIAS v4.x system versus that from the PADIAS v1.x system (i.e., third study objective), the following guidelines were used to develop the assessment data set. Using data stored in the IMS, test sections having the following characteristics were selected:

- Three pavement types - AC surfaced, jointed PCC, continuously reinforced PCC (CRC).

- Distresses that challenge identification or quantification - low severity fatigue cracking in AC, transverse cracks in CRC, corner breaks, rigid patches.

- High amounts of distress.

Using the above criteria, 24 test sections were identified for use in the PADIAS v1.x versus PADIAS v4.x comparison.

Distress data reports from the PADIAS v1.x were generated (both maps and summaries) for these 24 test sections from the data available at the LTPP IMS; the actual interpretation of the film had been completed during the SHRP years. PASCO USA also digitized the films for these test sections using PADIAS v4.x and its revised production methodology, including the use of multiple operators, multiple repeats and comparison and correction of discrepancies. The resulting PADIAS v1.x and v4.x data sets served as the basis for comparison of the two methods.

1.4 Report Overview

This report presents the results of the study undertaken by the FHWA to assess the variability of LTPP distress data. Besides accomplishing the specific objectives set forth in the study, this report provides a more complete picture of the various issues associated with the LTPP distress data in order to provide the knowledge base for informed decision making. For example, some of the questions to be addressed include:

- How should we analyze PASCO film for the first five years of the LTPP program to obtain distress data of acceptable quality?
- How do we collect and interpret distress data in the future?
- Do we continue to collect data on all distress types or should emphasis be placed on "important" distresses?
- What are the correlations between methods?
- What is the level of variability for all methods?

Chapter 1 of this report presents an introduction and overview of the LTPP distress data collection efforts to date, the need for and objectives of this distress data study, and the data sources used in support of the data study. Chapter 2 focuses on the assessment of manual distress data, while Chapter 3 is devoted to the assessment of distress data from film using the PADIAS v4.x system as well as the comparison of data derived from this method versus those from the PADIAS v1.x system. The combined results of the manual and film-derived distress data assessments are both compared and discussed in Chapter 4. Finally, the major findings, conclusions, and recommendations from the study are presented in Chapter 5. Much of the supporting data, in the form of figures or tables, are included in the various appendices to this report.

2. MANUAL DISTRESS DATA VARIABILITY

2.1 Introduction

Pavement distress surveys based upon field interpretation and manual mapping and recording of the distress information on paper forms have been used in the Long-Term Pavement Performance (LTPP) program to collect vitally important pavement condition and distress data. Although this "manual" method was used in the past primarily as a backup to the 35-mm black and white photographic-based method, the use of manual distress survey methods has recently increased in intensity and coverage. Many important distress conditions have only been captured with manual-based methods. The study presented in this chapter was undertaken by the Federal Highway Administration (FHWA) to assess the level of variability between individual distress raters and to address the potential precision and bias in these data.

2.2 Background

To promote the uniformity and consistency of distress data collection on the LTPP test sections, one of the early LTPP efforts was to develop standard definitions, measurement procedures, and data collection forms. These guidelines are contained in *Distress Identification Manual for the Long-Term Pavement Performance Project*, Strategic Highway Research Program, National Research Council, Publication No. SHRP-P-338, May 1993. This manual is typically referred to as the DIM (Distress Identification Manual). Although some early manual distress data were collected using an earlier version of this document, those data have been edited to conform to the current version.

The LTPP manual distress survey procedure contained in the DIM is based upon a single rater's interpretation of distresses, preparation of maps indicating the location and nature of distresses within the monitoring portion of the test section, and summarization of the extent and severity of all distresses found. In typical practice, copies of the previous manual distress survey maps are reviewed. The rater then maps distresses identified and notes differences in distress severity or interpretation from the last survey. In addition to the rater's somewhat subjective classification of distresses, the manual survey procedure includes transverse profile measurements on asphalt concrete (AC) surfaced pavements to characterize permanent deformation in the wheel paths (ruts) and fault height measurements on jointed portland cement concrete (PCC) pavements. Only the variability associated with the summary extent and severity rating of individual distresses, which are the most subjective feature of these surveys, are addressed in this study.

The data quality control and quality assurance function for manual distress surveys consists of distress rater accreditation workshops, office review of all manual data by an accredited staff member, and data logic and range checks performed after entry of the data into the LTPP Information Management System (IMS). Accreditation workshops were developed to help promote quality and consistency in manual distress data collection. All surveys entered into the

LTPP IMS must have been conducted by a rater accredited within the last two years. The objectives of the accreditation workshops are to (1) train raters in a consistent and uniform interpretation of the DIM and LTPP standard procedures, and (2) promote consistency between raters' distress interpretations. The second aspect of distress data quality control is an office review of surveys by an accredited staff member. In these reviews, findings from previous surveys are used to detect anomalies, oversights, omissions, and errors prior to entry into the IMS for further processing and public release. The third phase of quality control is range and logic checks on the data entered into the LTPP IMS.

Although extensive efforts were taken in the development of the DIM to define distress types and severity levels in objective terms, the application of these definitions still requires subjective interpretation by the rater. It is this subjectivity that causes the variability between raters, although some variations can be created by arithmetic mistakes made in the summary process. Recognizing that rater variability exists, this study was undertaken to quantify the level of variability so that anyone using manually collected LTPP distress data could incorporate, or at least recognize, this important feature in their analyses.

2.3 Data Source

Results from nine of the LTPP rater accreditation workshops summarized in table 2 were used in this study. This data source provides the distress ratings of 6 to 16 individual raters per workshop who performed ratings on the same day on the same test section. In addition to the individual raters, Workshops No. 8 and 9 also included two-person team consensus surveys to look at possible improvements in distress data variability versus that of individual rater surveys. Reference surveys were also conducted by the workshop instructors immediately prior to each workshop using a consensus rating method; distress data from the surveys are considered to be the closest approximation to ground truth.

Table 2. Summary of Accreditation Workshops.

Workshop No.	Date	Location	No. of Raters
1	June 1992	Reno, Nevada	12
2	July 1992	Reno, Nevada	14
3	August 1993	Minneapolis, Minnesota	15
4	May 1994	Buffalo, New York	11
5	October 1994	Houston, Texas	12
6	April 1995	Reno, Nevada	13
7	October 1995	Minneapolis, Minnesota	12
8	July 1996	Reno, Nevada	16
9	September 1996	Champaign, Illinois	6
		Total:	111

There are two types of workshops: accreditation and re-accreditation. The accreditation workshop provides an extensive introduction to distress definitions and rating procedures contained in the DIM. The accreditation workshop is conducted over a four-day period and offers each rater two practice rating sessions on each pavement type prior to performing a rating on the "exam" accreditation test sections. The accreditation workshops are targeted at individuals who are being accredited for the first time. The re-accreditation workshop, conducted over a three-day period, has less emphasis on introductory material and provides for only one practice rating on each type of pavement prior to the exam rating. Re-accreditation workshops are conducted for previously accredited raters who are familiar with LTPP distress survey methods and require the two-year re-accreditation. The accreditation and re-accreditation process consists of two major parts: a written examination and a two-part field survey examination. The written examination is intended to test the rater's knowledge of LTPP distress definitions and procedures.

The field survey examinations are intended to measure each rater's capabilities in observing, interpreting, and recording distress data. The field examinations are conducted on two 150-m-long pavement sections, one AC surfaced and one jointed PCC pavement (JCP). Each rater is given two hours to complete an **independent** distress rating of each section following all LTPP procedures, including preparation of a detailed scaled map and reduction of mapped quantities onto the summary distress forms. Prior to the workshops, the sections were surveyed in detail using a consensus procedure by the workshop instructors and, in some cases, other knowledgeable personnel, to determine the type, extent, and severity of distresses present. The results of these consensus surveys were used as the reference values against which the individual raters' results were compared for accreditation. These reference values are considered to be a close estimate of ground truth for the distress actually present. To assess the variability in the LTPP manual rating methods, only the summary distress data from the exam accreditation ratings were used. The results of the practice sessions were not used.

2.4 Global Trends

To gain a general understanding of the magnitude of variability associated with manually collected LTPP distress data, plots of distress quantity at each severity level and total across all severity levels for a distress type were developed for each of the common distress types identified at the nine accreditation workshops. For a given distress type and severity level combination, the following values are plotted:

- **Reference Value** - Quantity of distress, for each distress type and severity level, determined by the consensus survey conducted by the workshop instructors immediately prior to the workshop. These reference values are considered a close approximation of the ground truth and are used in this study to estimate the potential bias of the LTPP distress raters.

- **Group Mean** - Average of individual distress quantities, for each distress type and severity level, recorded by each rater on the summary distress forms.

- **Minimum and Maximum Values** - Actual minimum and maximum values from the distress values collected by the group of raters.

In the distress plots, the letters "R" and "I" along the X-axis denote the values pertaining to the reference and the group of individuals, respectively. The complete set of global trend plots for AC and PCC distress types are contained in appendix A. Examples of these plots are given in figures 1 through 4; figures 1 and 2 show the fatigue cracking and longitudinal cracking (wheel path) plots for AC pavements, while figures 3 and 4 present the corner breaks and longitudinal cracking plots for jointed PCC pavements.

The following observations are based on the information presented in these plots:

- Although the magnitude of variability for any given distress type and severity level varies from workshop to workshop, in general the variability is large and the scatter of data tends to increase in magnitude as the quantity of distress increases. Coefficients of variation (standard deviation divided by mean expressed as a percentage) of 30 percent or higher are common.

- For total distress summed across all severity levels for each distress type, the group means are generally closer to the reference value and the between rater scatter is smaller than for the individual severity levels. This is indicative of the greater variability in classification of severity level for a distress type than that associated with distress type identification.

- 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 compared with the reference values.

- 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.

- Because of the relatively high variability between distress rater values, the reference value is almost always within the range of data scatter for all distress type and severity level combinations.

Many of the above observations would be expected and are not considered unusual; however, the relatively high level of between rater variability was surprising and could potentially have a significant impact upon the usefulness of the data.

2.5 Impact of Re-Accreditation

The re-accreditation workshop held in Buffalo, New York, in May 1994 (Workshop No. 4) provided the basis for examination of the impact of re-accreditation on rater variability.

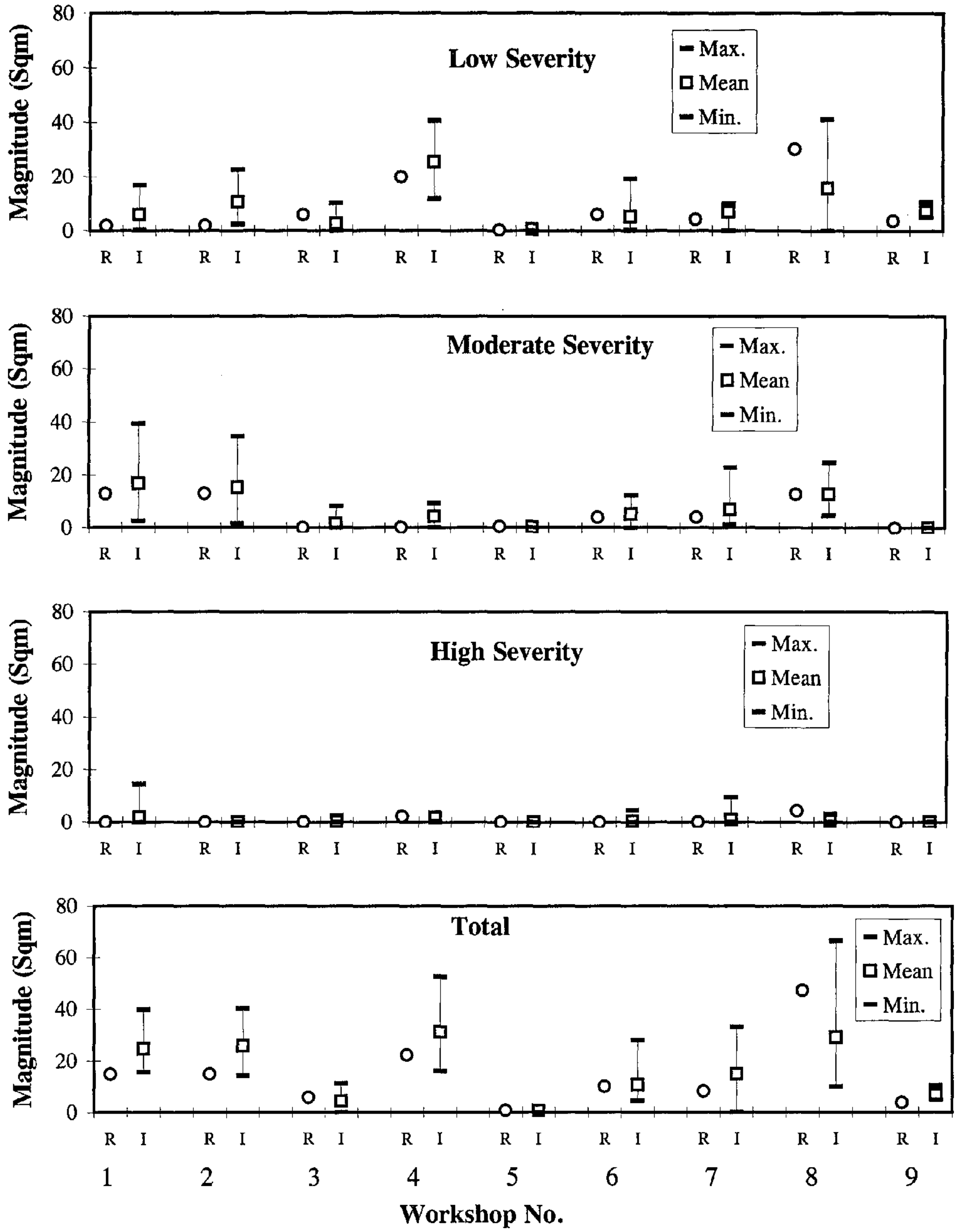


Figure 1. Fatigue Cracking (Sq. Meters) - AC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values.

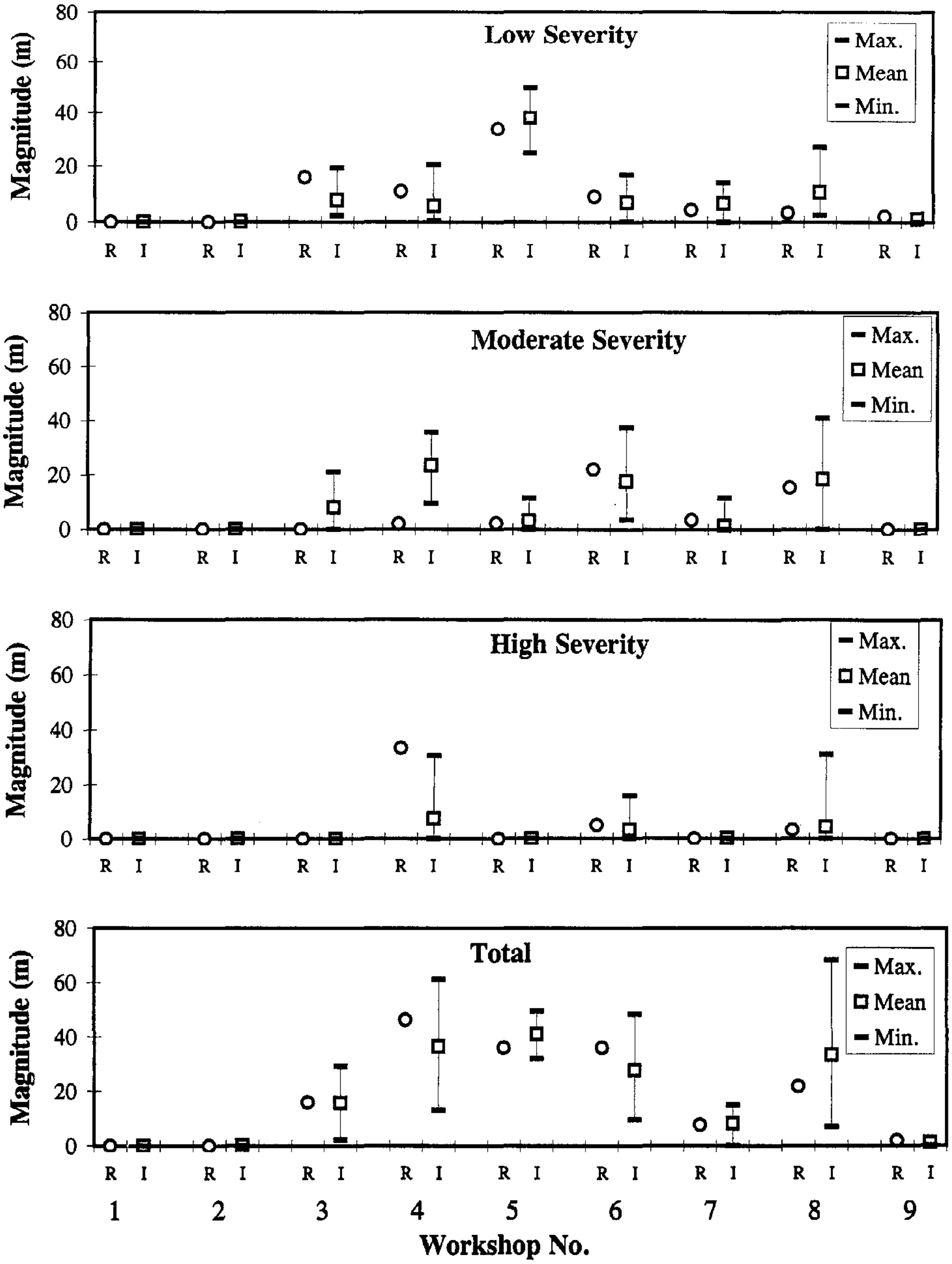


Figure 2. Longitudinal Cracking WP (Meters) - AC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values.

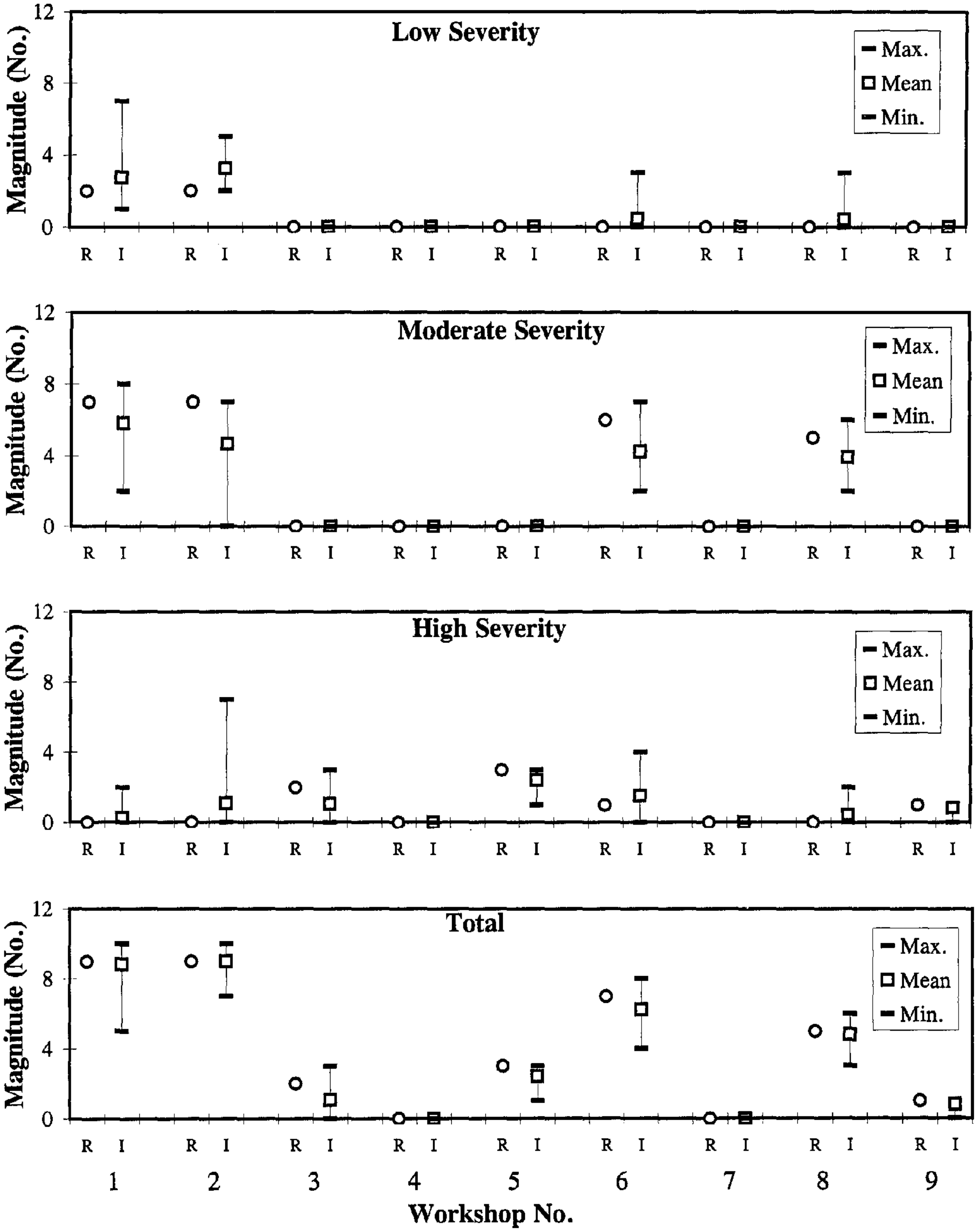


Figure 3. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values.

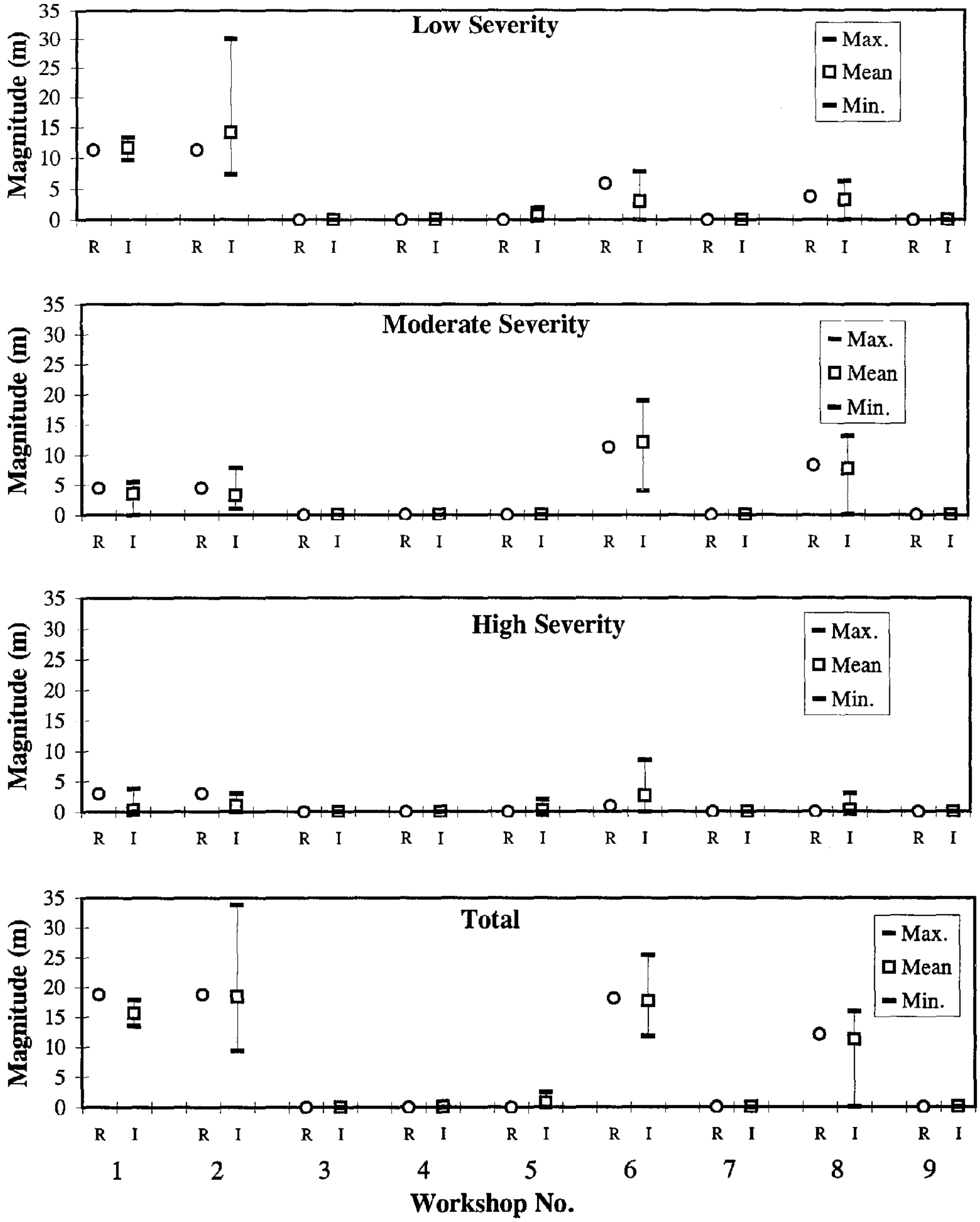


Figure 4. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Reference and Individual Minimum, Mean, and Maximum Values.

This was accomplished by comparing ratings from the 10 raters who had attended either of the two initial accreditation workshops held in Reno, Nevada, in June and July 1992 (Workshop No. 1 and 2) and re-accreditation Workshop No. 4.

Fatigue cracking, longitudinal cracking not in the wheel path, and transverse cracking (number and total length) were common distress types found between the initial and re-accreditation workshops on AC pavements, while transverse cracking (number and length) and joint seal damage of transverse joints were found on JCP pavements. A comparison of the common distress statistics derived for these two workshops is shown in tables 3 and 4 for the AC and JCP pavement test sections, respectively. The percentage of bias shown in these tables is equal to the group mean minus reference value divided by reference value, expressed as a percentage.

On the basis of information contained in tables 3 and 4, it appears there was no significant improvement in the between rater variability and bias relative to the reference values. Both increases and decreases were observed in the coefficient of variation for different distresses/severity combinations between workshops. Although the change in coefficient of variation appears to be quite large for several severity levels, in these cases the larger values are for situations in which the group mean is very small in one workshop relative to the other workshop. The same phenomenon was also observed for the apparent bias of the group mean relative to the reference value.

2.6 Pavement Condition Index

A number of observations and/or conclusions have been made so far in this chapter with regard to distress variability - variability within distress types and severity levels, within and between workshops, and among individual distress raters. Another approach in evaluating distress rating variability is to use a composite distress statistic typical of those used in pavement management systems. For this comparison, distress data from the individual raters at each workshop were used to compute the well-established and widely recognized Pavement Condition Index or PCI value according to the procedure developed by the U.S. Army Corps of Engineers. Because of differences in the definition of distress types and severity levels between the LTPP DIM and PCI methods, this exercise was limited to AC pavements and even then, distress types had to be combined (e.g., longitudinal cracking in the wheel path and not in the wheel path) and a few minor assumptions had to be made (e.g., severity levels assumed to be the same for a few distress types).

To compute the PCI, deduct values are computed for each distress/severity level. In computing these deduct values, weighing curves are used to transform the extent for each distress severity level. These weighing functions act to suppress the sensitivity of the PCI to variability in individual distress ratings, whose extent are less than the trigger levels that indicate the need for application of corrective treatments. Thus, it is expected that the PCI will result in a lower between rater variability and bias from the reference values. The computed PCI value for the individual raters and the statistics for the nine workshops (i.e., AC pavements) are summarized in table 5 and shown graphically in figure 5. The major observations and/or conclusions derived from these data are as follows:

- As expected, there is generally an improvement in the agreement among the individual raters, the group mean, and the reference PCI value. The difference between the group mean and reference value for three of the workshops is 1 PCI, for another workshop the difference is 2 PCI, and for other two others it is 5 and 6 PCI, respectively. The differences for Workshops No. 2, 4, and 8 were 13, 14, and 12 PCI points, respectively, which are considered fair at best. It is interesting to note that the reference PCI is always less than the group mean, except for Workshop No. 7 where the reference PCI is 1 greater than the group mean.

- Variability between individual raters also improved compared with variability of distress data for individual distress/severity levels. The coefficient of variation ranges from a low of 2.0 percent to a high of 18.7 percent.

- PCI values for the two-person teams who participated in Workshops No. 8 and 9 appear to show that the teams are more consistent and that their results are closer to the reference values compared with those from the individual raters.

2.7 Bias and Precision

One of the main objectives of the study presented in this chapter was to quantify the bias and precision associated with LTPP manual distress data. Toward this end, data obtained from the distress rater accreditation workshops were manipulated, and analyses were conducted in two phases. The purpose of the first phase was to evaluate the within group variability (associated with group means) as well as the group bias and variability associated with the reference values. An outlier analysis was also conducted within this analysis phase. Because differences between group means and their corresponding reference values appeared to be small for most distress types, a detailed statistical evaluation was conducted in the second phase to quantify the bias and precision for manual distress data.

Coefficient of Variation and RMSE Evaluation

The coefficient of variation (CV), defined as standard deviation divided by mean value, 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 manual distress data.

In this analysis, the CV was determined by plotting standard deviation versus mean for each distress type-severity level combination and fitting the best line through these data (y-intercept was forced through 0). See appendix A for a complete set of plots for both AC and PCC

Table 3. Group Statistics for Raters Attending Workshop Nos. 1, 2, and 4; AC Pavements.

Distress Type	Units	Distress Severity	Workshop No.	Reference	GROUP Mean	Min.	Max.	STDEV	COV (%)	%Bias
Fatigue Cracking	Sq. Meters	Low	# 1	1.90	7.21	1.00	14.00	5.45	75.60	279.26
			# 4	19.90	25.31	11.60	40.70	10.02	39.57	27.19
		Moderate	# 1	13.00	13.13	1.50	31.70	8.48	64.58	0.98
			# 4	0.20	4.09	0.00	9.30	2.62	63.97	1945.00
		High	# 1	0.00	1.62	0.00	14.40	4.29	265.04	
			# 4	2.20	1.64	0.10	3.40	1.13	69.16	-25.45
		Total	# 1	14.90	21.95	14.20	33.30	6.13	27.94	47.34
			# 4	22.30	31.04	16.00	52.50	10.93	35.22	39.19
Longitudinal Cracking - NWP	Meters	Low	# 1	22.10	32.08	18.80	57.50	10.91	34.00	45.16
			# 4	22.60	26.82	12.10	32.20	5.72	21.31	18.67
		Moderate	# 1	44.00	20.00	9.50	36.75	8.40	41.99	-54.56
			# 4	15.00	4.09	0.00	8.00	2.75	67.18	-72.73
		High	# 1	23.00	32.89	10.00	49.20	14.51	44.12	42.98
			# 4	1.50	0.76	0.00	3.50	1.12	147.72	-49.33
		Total	# 1	89.10	84.96	76.40	106.10	9.07	10.68	-4.65
			# 4	39.10	31.67	20.10	39.50	5.39	17.03	-19.00
Transverse Cracking	No.	Low	# 1	25.00	32.10	27.00	43.00	5.24	16.33	28.40
			# 4	10.00	7.00	4.00	9.00	1.79	25.56	-30.00
		Moderate	# 1	20.00	8.60	5.00	16.00	3.64	42.31	-57.00
			# 4	3.00	5.70	3.00	9.00	1.95	34.24	90.00
		High	# 1	11.00	10.40	2.00	15.00	3.69	35.51	-5.45
			# 4	11.00	4.90	1.00	8.00	2.02	41.27	-55.45
		Total	# 1	56.00	51.10	42.00	63.00	6.22	12.17	-8.75
			# 4	24.00	17.60	15.00	20.00	1.50	8.50	-26.67
Transverse Cracking	Meters	Low	# 1	21.80	35.90	24.70	48.70	7.00	19.51	64.66
			# 4	4.40	4.19	2.10	8.50	1.92	45.93	-4.77
		Moderate	# 1	28.20	16.94	9.50	27.80	6.01	35.48	-39.95
			# 4	1.80	11.21	3.80	17.80	4.15	36.98	522.78
		High	# 1	24.60	23.14	0.00	31.30	8.44	36.45	-5.93
			# 4	20.90	9.49	1.00	18.00	4.36	45.98	-54.59
		Total	# 1	74.60	75.97	64.60	81.90	5.86	7.71	1.84
			# 4	27.10	24.89	19.70	29.10	3.01	12.09	-8.15

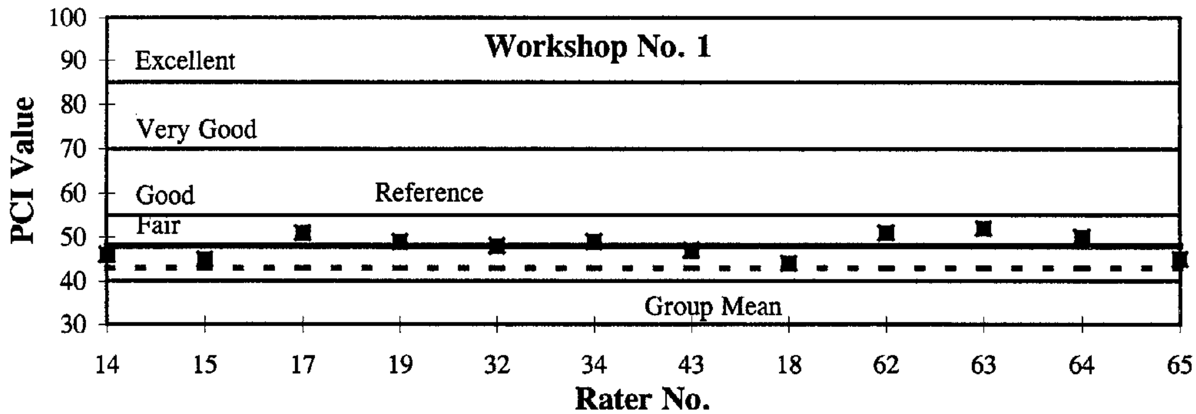
Table 4. Group Statistics for Raters Attending Workshop Nos. 1, 2, and 4; PCC Pavements.

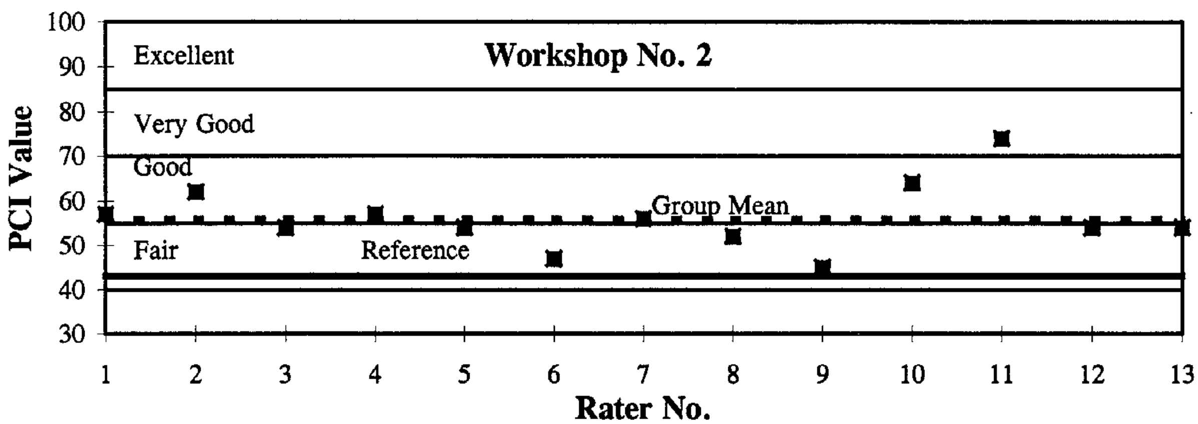
Distress Type	Units	Distress Severity	Workshop No.	Reference	GROUP					
					Mean	Min.	Max.	STDEV	COV (%)	%Bias
Transverse Cracking	No.	Low	# 1	0.00	1.70	0.00	3.00	1.00	59.12	
			# 4	12.00	7.30	4.00	13.00	2.49	34.14	-39.17
		Moderate	# 1	1.00	0.30	0.00	2.00	0.64	213.44	-70.00
			# 4	4.00	11.90	7.00	15.00	2.39	20.05	197.50
		High	# 1	2.00	1.70	0.00	2.00	0.64	37.67	-15.00
			# 4	1.00	0.20	0.00	2.00	0.60	300.00	-80.00
		Total	# 1	3.00	3.70	2.00	5.00	1.00	27.16	23.33
			# 4	17.00	19.40	14.00	25.00	2.94	15.15	14.12
Transverse Cracking	Meters	Low	# 1	0.00	3.93	0.00	5.20	1.45	36.77	
			# 4	30.00	10.57	4.00	22.90	4.90	46.37	-64.77
		Moderate	# 1	3.70	0.74	0.00	7.40	2.22	300.00	-80.00
			# 4	14.00	35.25	24.60	39.40	5.45	15.45	151.79
		High	# 1	7.40	5.57	0.00	7.60	3.00	53.79	-24.73
			# 4	0.60	0.74	0.00	7.40	2.22	300.00	23.33
		Total	# 1	11.10	10.24	5.20	12.60	2.46	23.99	-7.75
			# 4	44.60	46.56	38.80	53.40	3.67	7.89	4.39
Jt. Seal Damage of Transverse Joints	No.	Low	# 1	32.00	32.00	32.00	32.00	0.00	0.00	0.00
			# 4	5.00	5.00	5.00	5.00	0.00	0.00	0.00
		Moderate	# 1	0.00	0.00	0.00	0.00	0.00		
			# 4	0.00	0.00	0.00	0.00	0.00		
		High	# 1	0.00	0.00	0.00	0.00	0.00		
			# 4	0.00	0.00	0.00	0.00	0.00		
		Total	# 1	32.00	32.00	32.00	32.00	0.00	0.00	0.00
			# 4	5.00	5.00	5.00	5.00	0.00	0.00	0.00

Table 5. Pavement Condition Index (PCI) for AC Pavements.

WKSP No.	Location	Date	Group Statistics						Rater No.																
			Ref.	Min.	Max.	Mean	STDEV	COV	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Reno	6/9/92	43	44	52	48	2.56	5.33	46	45	51	49	48	49	47	44	51	52	50	45					
2	Reno	7/14/92	43	45	74	56	7.12	12.69	57	62	54	57	54	47	56	52	45	64	74	54	54				
3	Minnesota	8/10/93	44	41	61	46	5.07	11.03	46	41	44	54	44	46	44	41	61	43	44	48	45	42	47		
4	Buffalo	5/3/94	50	58	71	64	3.89	6.05	58	63	63	66	71	61	67	69	61	61	68						
5	Houston	10/4/94	82	80	86	83	1.68	2.03	83	82	80	86	84	83	83	83	83	82	81	86					
6	Reno	4/4/95	59	60	68	65	2.22	3.41	63	63	66	66	60	65	68	68	65	64	65	68	64				
7	Minnesota	10/3/95	64	47	76	63	7.92	12.50	58	61	69	70	61	72	71	47	58	58	76	59					
8	Reno	7/30/96	30	32	61	42	7.88	18.69	27*	34	41	33	32	37	55	43	41	46	36	44	52	61	42	42	36
9	Champagne	9/17/96	50	43	61	51	6.94	13.71	53*	50*	54*	59	44	48	43	49	61								

* Two-person team.





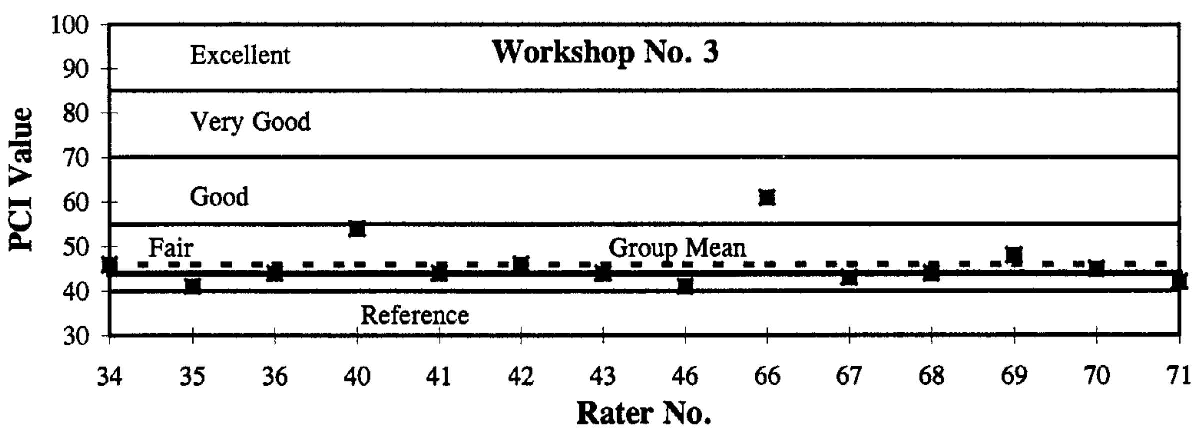


Figure 5. Pavement Condition Index (PCI) for AC Pavements.

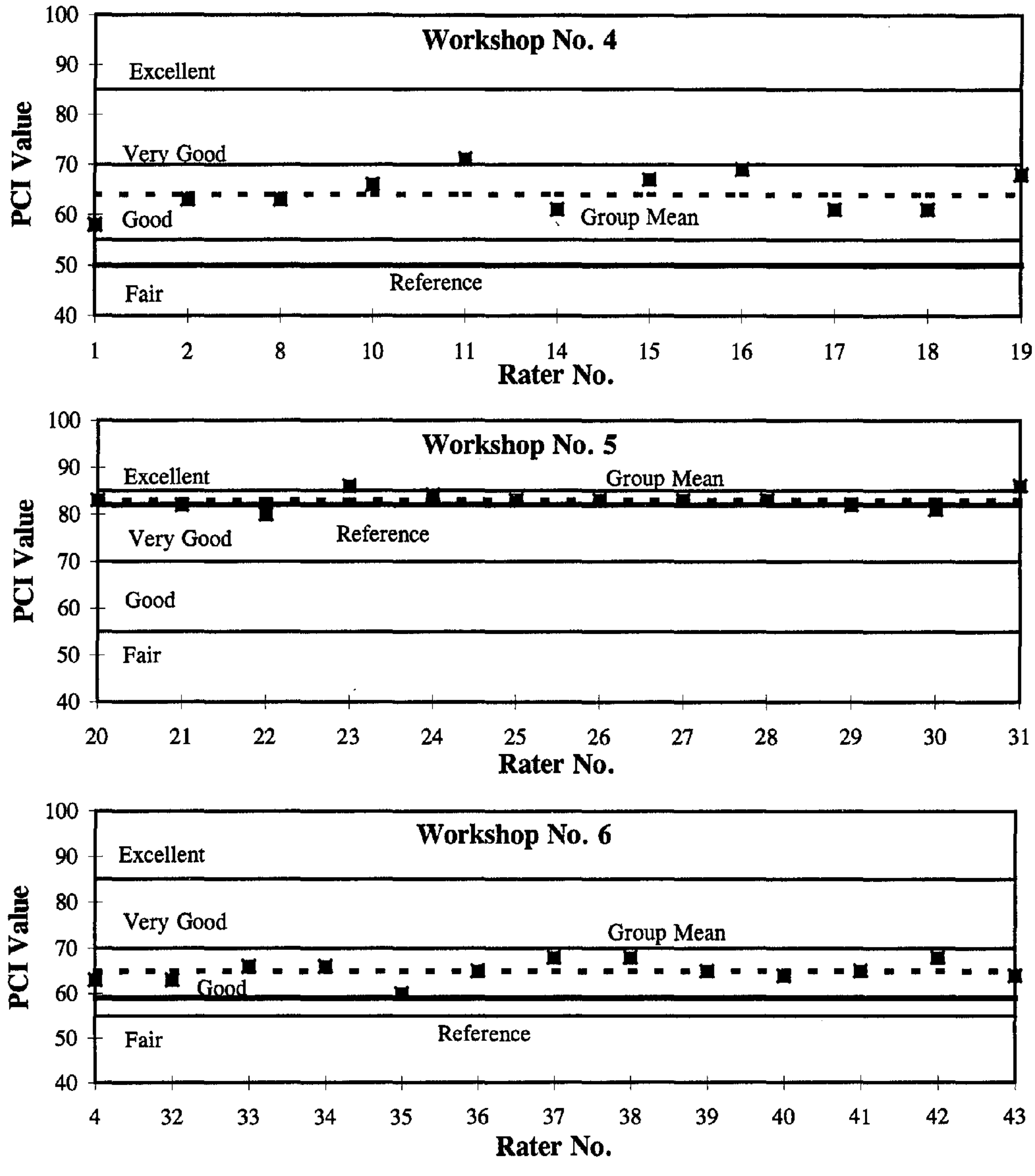
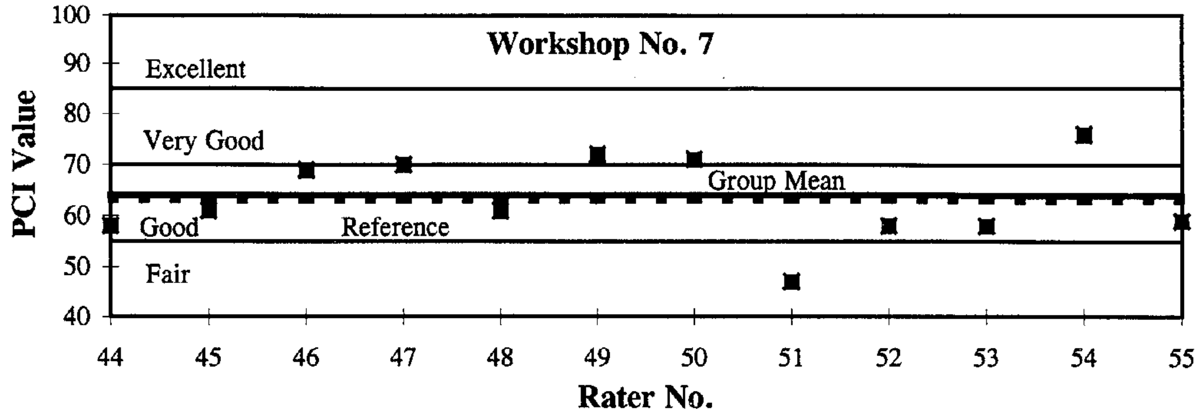
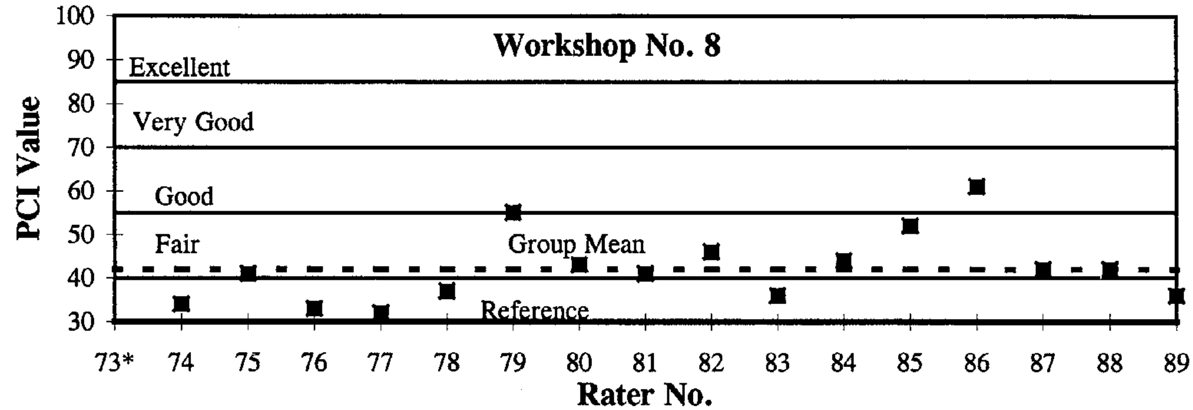


Figure 5. Pavement Condition Index (PCI) for AC Pavements (continued).





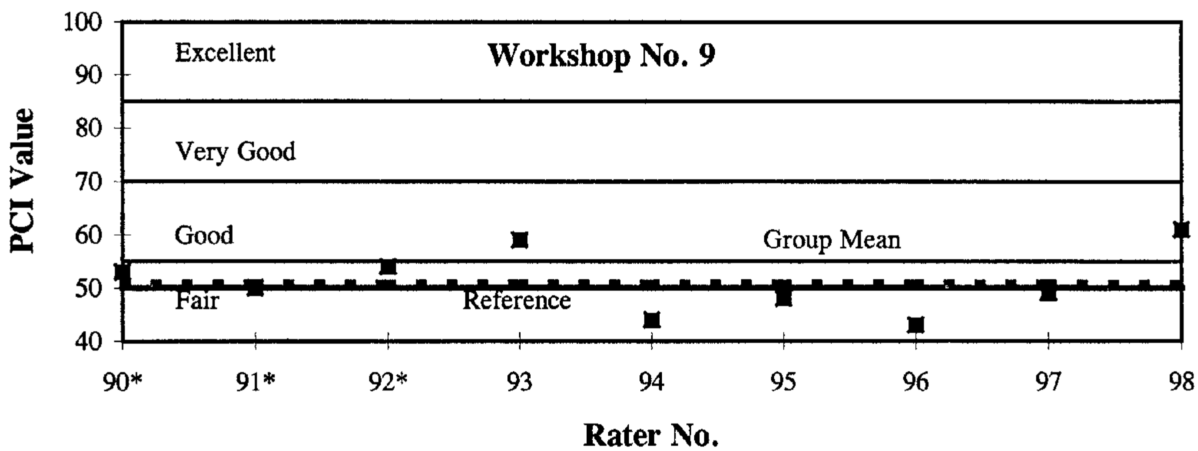


Figure 5. Pavement Condition Index (PCI) for AC Pavements (continued).

pavements. The slope of this best-fit straight line (in percentage terms) forced through 0 is a measure of the ratio between standard deviation and mean over varying ranges of CV, assuming a linear increasing relationship between standard deviation and mean. Examples of the resulting CV plots for different distress types and severity levels are shown in figures 6 through 9; figures 6 and 7 show plots for fatigue cracking and longitudinal cracking (in wheel path) for AC pavements, while figures 8 and 9 show similar plots for corner breaks and longitudinal cracking in PCC pavements, respectively. The regression line marked as STDEV1 in these four figures represents the derived CV for the data in question.

Also shown in these figures are the RMSE and STDEV2 vs. mean regression lines. RMSE (root mean square of error) is defined as the square root of the summation of the squared difference between the reference and the individual rater values divided by the number of raters in the workshop. This term combines the variability and bias associated with the individual rater data relative to the reference values, i.e., distribution of rater values relative to the reference.

The STDEV2 regression line was developed using the standard deviation and mean values from the nine workshops after rejecting observations that were considered statistical outliers. Individual data points were considered outliers if the difference between the individual distress values in each workshop and their group means were greater than three times the standard deviations derived from the STDEV1 regression lines discussed previously. This exercise was only performed on total distress. Once a data point was identified as an outlier, it was also dropped from further analysis at the individual severity level (low, moderate, and high). The objective of this exercise was to examine the potential improvement in variability by excluding raters, i.e., tightening rater certification requirements.

The CV for total distress quantities, without rejection of outliers, ranged from 9 to 38 percent for AC pavements. For PCC pavements, the CV ranged from 8 to 22 percent for cracking-related distresses. The CV were relatively high for joint spalling of PCC pavements, ranging from 25 to 71 percent. Except for fatigue cracking of AC pavements and transverse joint spalling of PCC pavements, no appreciable differences were observed between the STDEV1 and RMSE regression lines, which indicates that the difference between the group means and their corresponding reference values is small. The elimination of outliers did not seem to affect that observation, except for joint spalling of PCC pavements, which appears to be a more difficult distress to quantify compared with others. This result is related to the very conservative outlier exclusion procedure adopted.

In addition to the CV plots, tables 6 and 7 were prepared to assess the apparent bias and precision of distress data across the nine workshops; table 6 shows the results for AC pavement distresses, while table 7 presents those for PCC pavements. The following terminology is used in both of these tables:

- **Pooled Reference Mean** – Average of individual reference mean values from nine accreditation workshops.

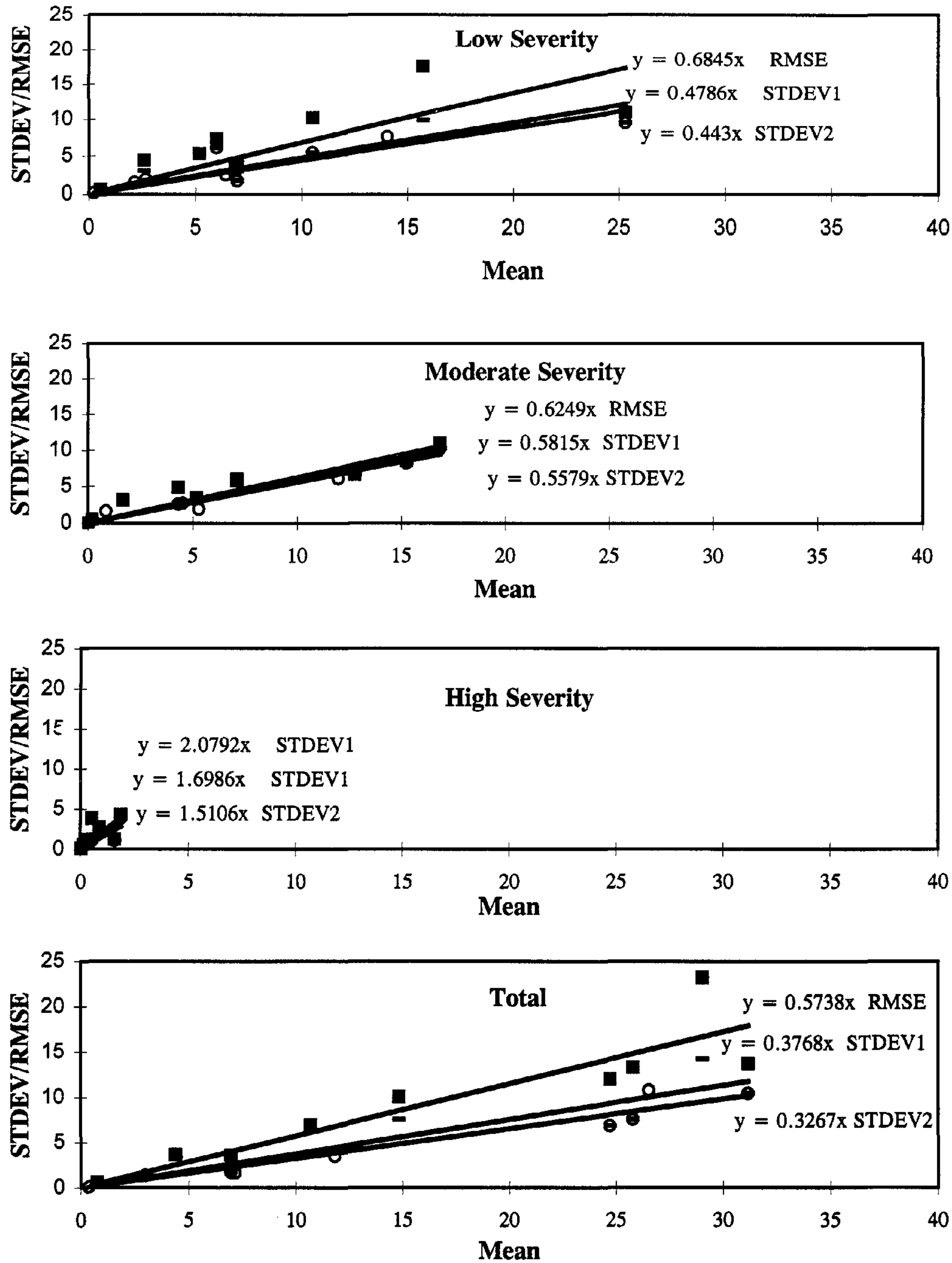


Figure 6. Fatigue Cracking (Sq. Meters) - AC Pavements, Manual Surveys: Standard Deviation/RMSE Vs. Mean.

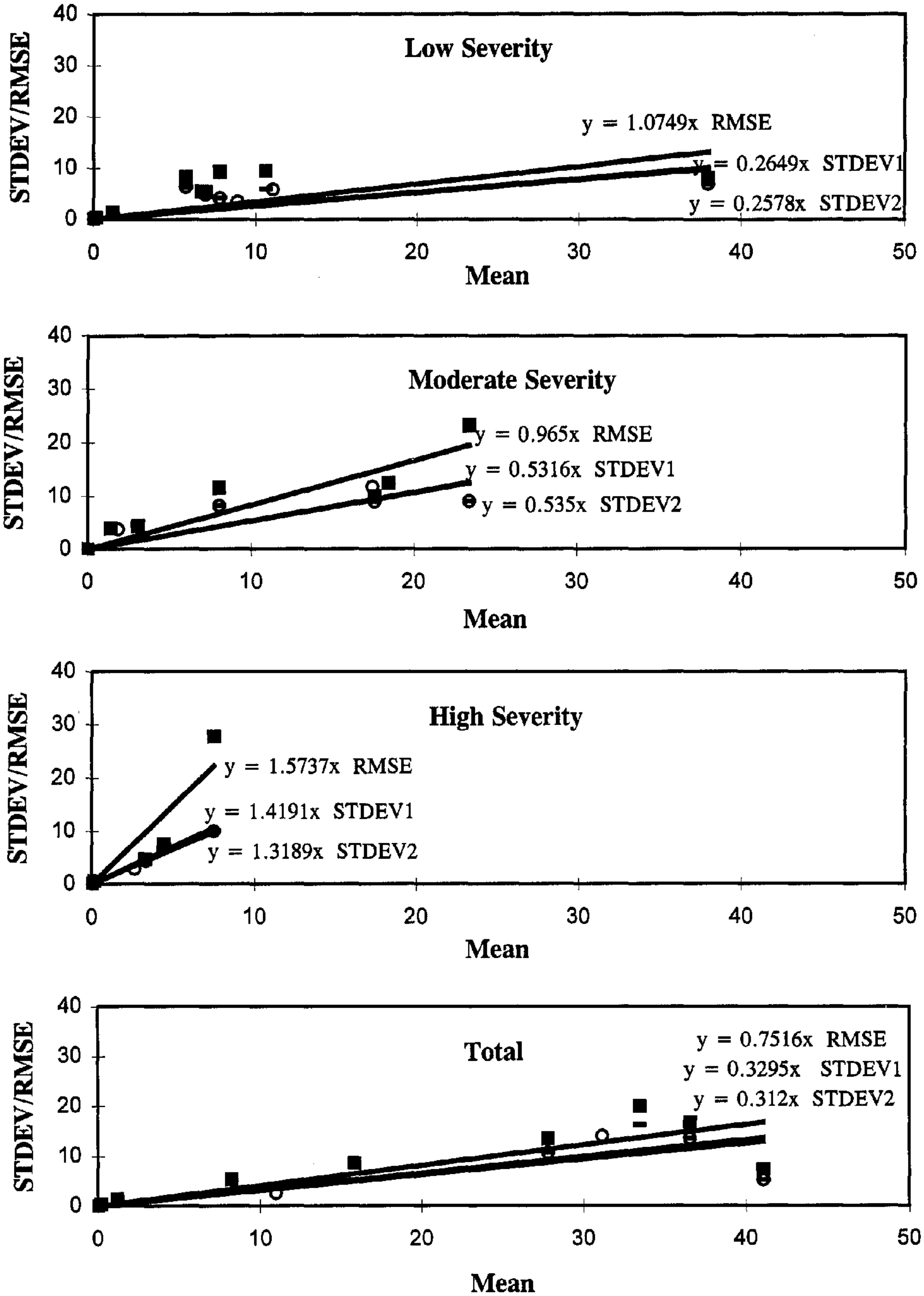


Figure 7. Longitudinal Cracking WP (Meters) - AC Pavements, Manual Surveys: Standard Deviation/RMSE Vs. Mean.

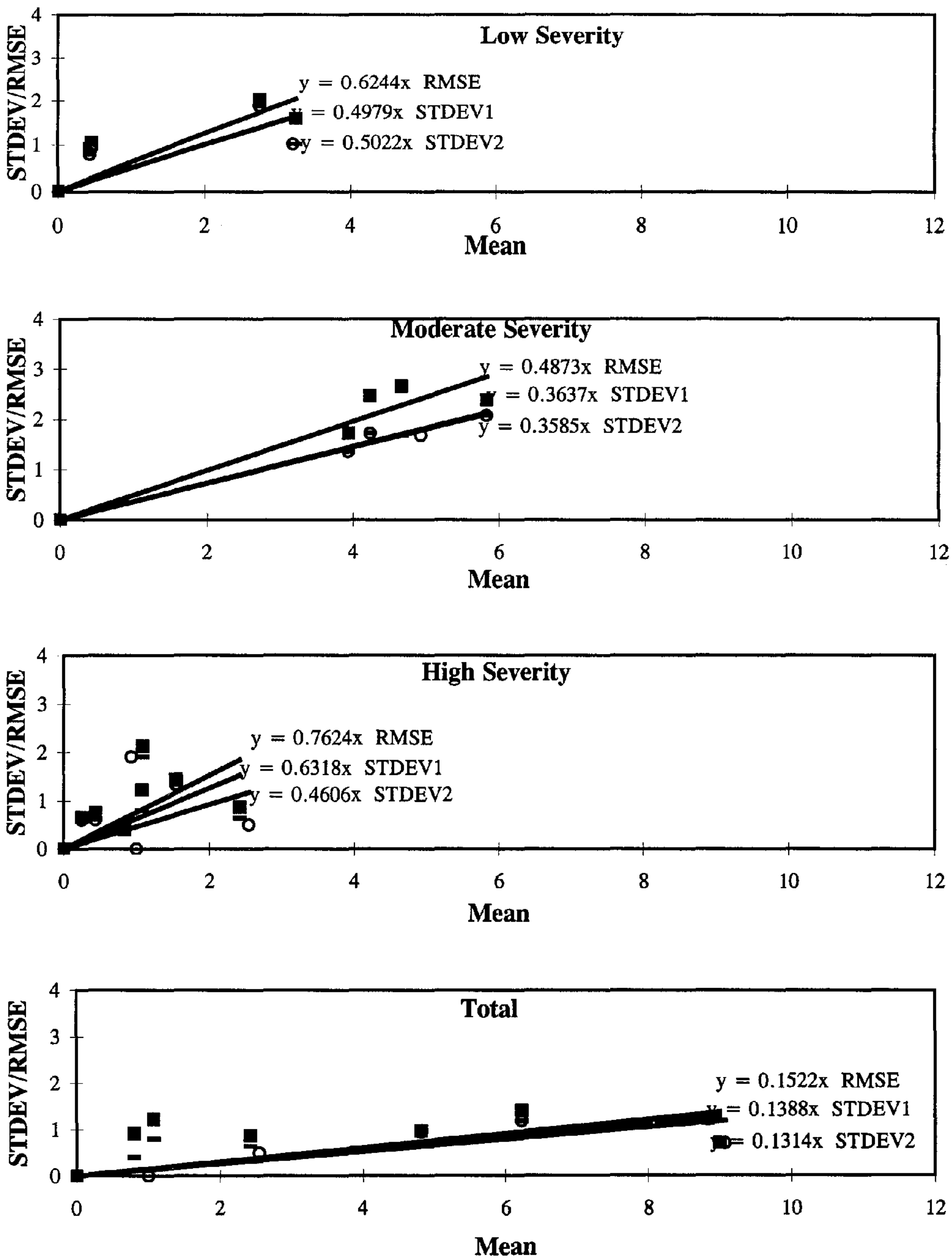


Figure 8. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Standard Deviation/RMSE Vs. Mean.

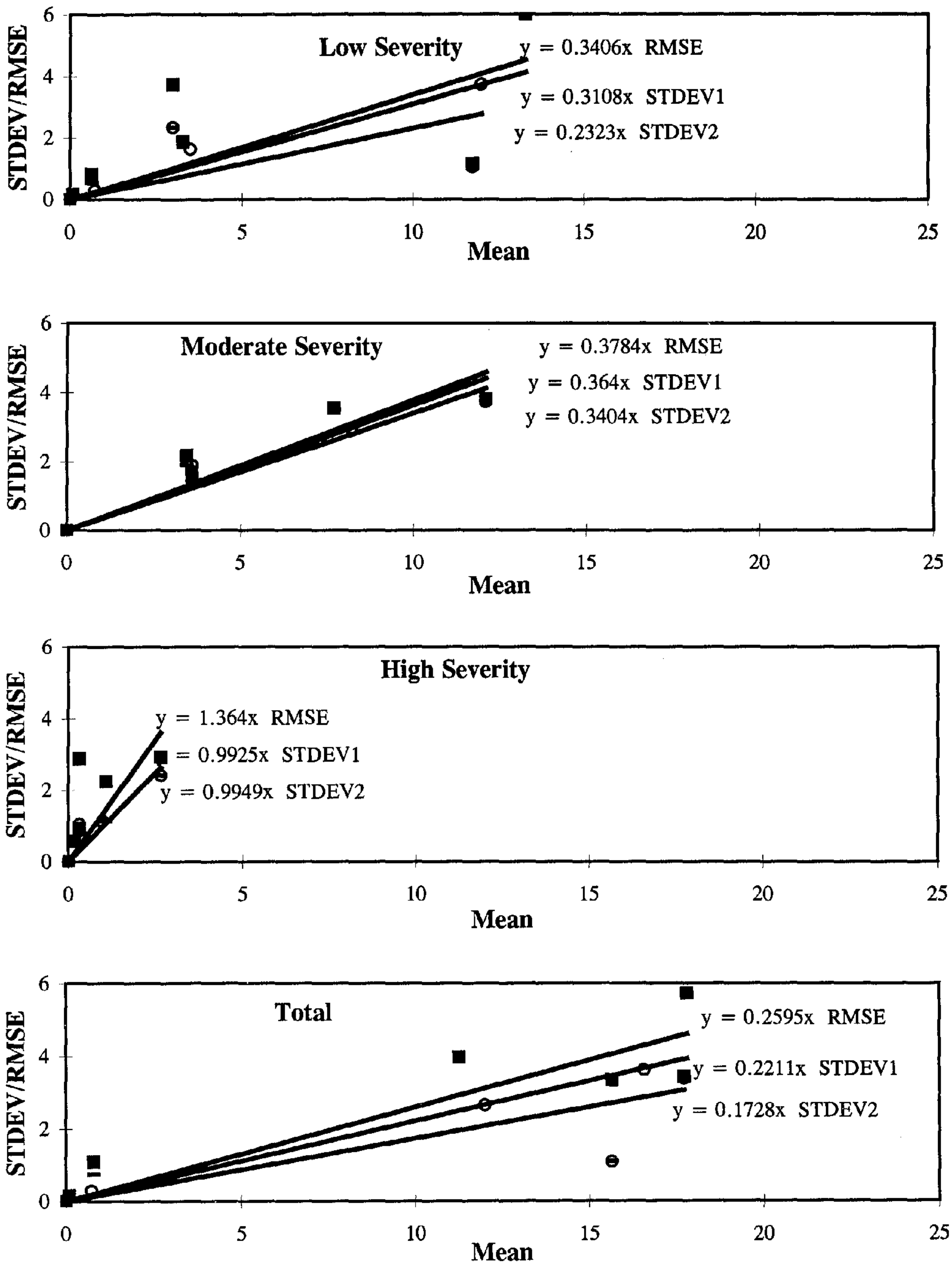


Figure 9. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Standard Deviation/RMSE Vs. Mean.

Table 6. Indicators of Bias and Precision for AC Pavement Distresses.

DISTRESS TYPE	UNIT	DISTRESS SEVERITY	POOLED REF.	POOLED VALUES WITH OUTLIERS GROUP MEAN	STD DEV.	RMSE	COV (%)	BIAS	POOLED VALUES W/O OUTLIERS GROUP MEAN	STD DEV.	COV (%)	BIAS
Fatigue Cracking	Sq. Meters	Low	8.22	8.85	4.24	6.06	47.86	0.63	8.25	3.66	44.30	0.03
		Moderate	5.29	7.04	4.09	4.40	58.15	1.75	6.58	3.67	55.79	1.29
		High	0.72	0.60	1.02	1.25	169.86	-0.12	0.46	0.70	151.06	-0.26
		Total	14.23	16.49	6.21	9.46	37.68	2.26	15.29	5.00	32.67	1.06
Longitudinal Cracking WP	Meters	Low	8.82	8.56	2.27	2.95	26.49	-0.26	8.70	2.24	25.79	-0.12
		Moderate	4.99	7.99	4.25	6.66	53.16	3.01	7.94	4.25	53.50	2.95
		High	4.63	1.71	2.42	5.07	141.91	-2.93	1.52	2.00	131.90	-3.12
		Total	18.44	18.26	6.02	7.43	32.95	-0.18	18.16	5.66	31.17	-0.29
Longitudinal Cracking NWP	Meters	Low	22.88	26.16	18.56	19.50	70.95	3.28	26.29	18.65	70.93	3.41
		Moderate	35.74	26.58	16.79	19.60	63.16	-9.17	26.93	16.90	62.73	-8.81
		High	16.34	17.97	6.69	8.91	37.25	1.63	17.66	6.42	36.37	1.31
		Total	74.97	70.71	14.66	16.30	20.74	-4.26	70.88	14.49	20.45	-4.09
Transverse Cracking	Number	Low	10.56	11.74	2.36	3.46	20.09	1.18	11.96	2.10	17.59	1.40
		Moderate	8.33	7.30	2.57	5.03	35.18	-1.04	7.40	2.55	34.51	-0.94
		High	7.56	5.69	2.13	2.86	37.44	-1.86	5.74	2.15	37.56	-1.82
		Total	26.44	24.73	3.23	3.99	13.06	-1.72	25.09	3.00	11.97	-1.35
Transverse Cracking	Meters	Low	10.92	13.80	2.95	5.12	21.38	2.88	13.80	2.95	21.37	2.88
		Moderate	15.77	16.93	5.96	9.04	35.21	1.17	16.96	5.93	34.97	1.19
		High	17.60	13.89	4.95	6.59	35.65	-3.71	13.83	4.91	35.51	-3.77
		Total	44.29	44.62	4.21	5.32	9.44	0.33	44.59	3.99	8.95	0.30

Table 7. Indicators of Bias and Precision for PCC Pavement Distresses.

DISTRESS TYPE	UNIT	DISTRESS SEVERITY	POOLED REF.	POOLED VALUES WITH OUTLIERS GROUP					POOLED VALUES W/O OUTLIERS GROUP			
				MEAN	STD DEV.	RMSE	COV (%)	BIAS	MEAN	STD DEV.	COV (%)	BIAS
Corner Breaks	Number	Low	0.44	0.77	0.38	0.48	49.79	0.32	0.76	0.38	50.22	0.32
		Moderate	2.78	2.07	0.75	1.01	36.37	-0.70	2.10	0.75	35.85	-0.67
		High	0.78	0.85	0.54	0.65	63.18	0.07	0.86	0.39	46.06	0.08
		Total	3.89	3.68	0.51	0.56	13.88	-0.20	3.72	0.49	13.14	-0.17
Longitudinal Cracking	Meters	Low	3.59	3.55	1.10	1.21	31.08	-0.03	3.45	0.80	23.23	-0.14
		Moderate	3.18	2.98	1.08	1.13	36.40	-0.20	3.05	1.04	34.04	-0.12
		High	0.78	0.50	0.50	0.69	99.25	-0.28	0.47	0.47	99.49	-0.30
		Total	7.54	7.04	1.56	1.83	22.11	-0.51	6.98	1.21	17.28	-0.57
Transverse Cracking	Meters	Low	6.86	6.82	4.12	6.31	60.52	-0.04	6.77	4.02	59.33	-0.08
		Moderate	11.50	11.98	2.68	6.20	22.39	0.48	12.11	2.69	22.20	0.61
		High	6.42	6.21	2.39	2.56	38.49	-0.21	6.15	2.30	37.47	-0.27
		Total	24.78	25.01	2.08	2.39	8.34	0.23	25.04	1.83	7.32	0.27
Transverse Cracking	Number	Low	4.00	4.08	1.52	2.06	37.19	0.08	3.91	1.41	36.19	-0.09
		Moderate	3.44	3.73	0.99	2.20	26.57	0.28	3.71	0.98	26.51	0.26
		High	2.00	1.79	0.73	0.78	40.90	-0.21	1.78	0.73	41.00	-0.22
		Total	9.44	9.60	1.43	1.72	14.87	0.15	9.39	1.33	14.19	-0.05
Spalling of Longitudinal Joints	Meters	Low	5.87	5.88	6.17	6.92	104.92	0.01	4.27	2.25	52.62	-1.59
		Moderate	0.77	0.82	0.58	0.91	70.59	0.05	0.82	0.58	70.89	0.06
		High	0.01	0.50	0.67	0.83	134.20	0.49	0.49	0.66	134.14	0.48
		Total	6.64	7.19	4.91	5.52	68.19	0.55	5.59	1.78	31.92	-1.05
Spalling of Transverse Joints	Meters	Low	1.30	1.09	1.06	1.49	97.04	-0.21	0.75	0.32	42.23	-0.55
		Moderate	0.17	0.39	0.72	0.79	186.57	0.22	0.27	0.41	150.07	0.11
		High	0.22	0.54	0.14	0.35	26.56	0.32	0.48	0.12	24.29	0.26
		Total	1.69	2.02	1.44	2.03	71.20	0.33	1.58	0.53	33.71	-0.11
Spalling of Transverse Joints	Number	Low	2.89	2.50	0.70	1.03	28.18	-0.39	2.32	0.53	22.95	-0.57
		Moderate	0.44	0.55	0.79	0.89	143.25	0.11	0.47	0.60	127.00	0.03
		High	0.33	0.36	0.32	0.42	87.80	0.03	0.31	0.20	65.68	-0.03
		Total	3.67	3.41	0.87	1.12	25.45	-0.26	3.10	0.53	17.15	-0.56

- **Pooled Group Mean** - Average of group mean values from nine accreditation workshops.

- **Apparent Bias** – Difference between pooled group mean and pooled reference value.

- **Coefficient of Variation** – Slope of the straight-line regression between standard deviation and mean values.

- **Pooled Standard Deviation** - Product of pooled group mean and slope of best fit line from standard deviation versus mean (CV) plot.

- **Root Mean Square of Error (RMSE)** - Square root of the sum of squared difference between reference and individual rater values divided by number of raters in workshop.

On the basis of information provided in tables 6 and 7 and the standard deviation versus mean (CV) plots, the following observations and/or conclusions were made relative to apparent bias and precision of distress data:

- Apparent bias for most distress type-severity level combinations is small and it is not uniform, which is consistent with an earlier observation that group means are generally close to reference values. This is especially obvious in the CV plots, where the RMSE and STDEV1/ STDEV2 regression lines are relatively close to each other. Thus, it may be possible to disregard bias and consider the group mean as an unbiased estimate of the true value.

- As illustrated by regression lines in CV plots, standard deviation seems to increase as distress magnitude increases; however, data bias does not seem to be affected by magnitude of distress present on section.

- Precision of distress data relative to group mean appears acceptable, especially for total (all severities) distress quantities. CV values ranged from 9 to 38 percent for AC pavements and from 8 to 22 percent for cracking-related PCC pavement distresses. Quantification of precision for joint spalling in PCC pavements appears to be more difficult; however, the CV was significantly reduced after rejection of data outliers.

- Precision for total distress quantity is generally much better than that for the individual severity levels, which is consistent with an earlier observation that there seems to be a greater variability in distinguishing severity levels.

- Elimination of data outliers from analysis resulted in marginal improvements in apparent bias and precision, which appears to indicate that additional training would not lead to significant improvements for most distress types. Joint spalling in PCC pavements is the only exception, i.e., reduction in CV resulting from elimination of outliers appears to suggest that additional rater training would be beneficial.

Bias and Precision

The analyses conducted in the previous section indicated that the difference between group mean and corresponding reference values is generally small. A logical implication from this observation is that the group mean may be considered as an unbiased estimate of the reference value. In this section, a more vigorous analysis of the data was undertaken to better quantify the bias and precision associated with manual distress data. This analysis involved the following steps:

1. Defining the Relationship Between the Standard Deviation and Mean

It has been shown that, for all distress data, the standard deviation increases as the mean or magnitude of distress increases. A straight-line regression technique was used in the previous section to define the relationship between these two parameters. The resulting lines gave a general indication of data variability; however, to better quantify bias and precision, analyses were conducted to establish regression functions that better define that relationship. Three different functions were evaluated: straight line, logarithm transformation, and square root transformation of the mean values.

It was determined from this effort that regression equations with the square root transformation generally produced a better fit between standard deviation and mean for the three functions considered, and thus was used in the remaining analyses. Figures 10 through 13 illustrate this relationship, according to severity level, for fatigue cracking and longitudinal cracking (in wheel path) for AC pavements and corner breaks and longitudinal cracking in PCC pavements, respectively. The general form of the equation is:

$$StandardDeviation = C\sqrt{(Mean)}$$

where C is the regression constant. The complete set of standard deviation versus mean plots is contained in appendix A.

2. Outlier Analysis

Next, an outlier analysis similar to that in the previous section was conducted. Individual data were considered outliers and eliminated from further analyses if the difference between the individual value and group mean was greater than three times the standard deviation derived from the respective regression line. However, the standard deviation defined by the square root function was used instead of the straight-line function. Thus, a new set of regression equations was developed using the revised data.

3. Bias Analysis

Finally, a bias analysis was conducted for all distress type and severity level combinations in each workshop. A bias was considered to exist if the absolute difference between the workshop

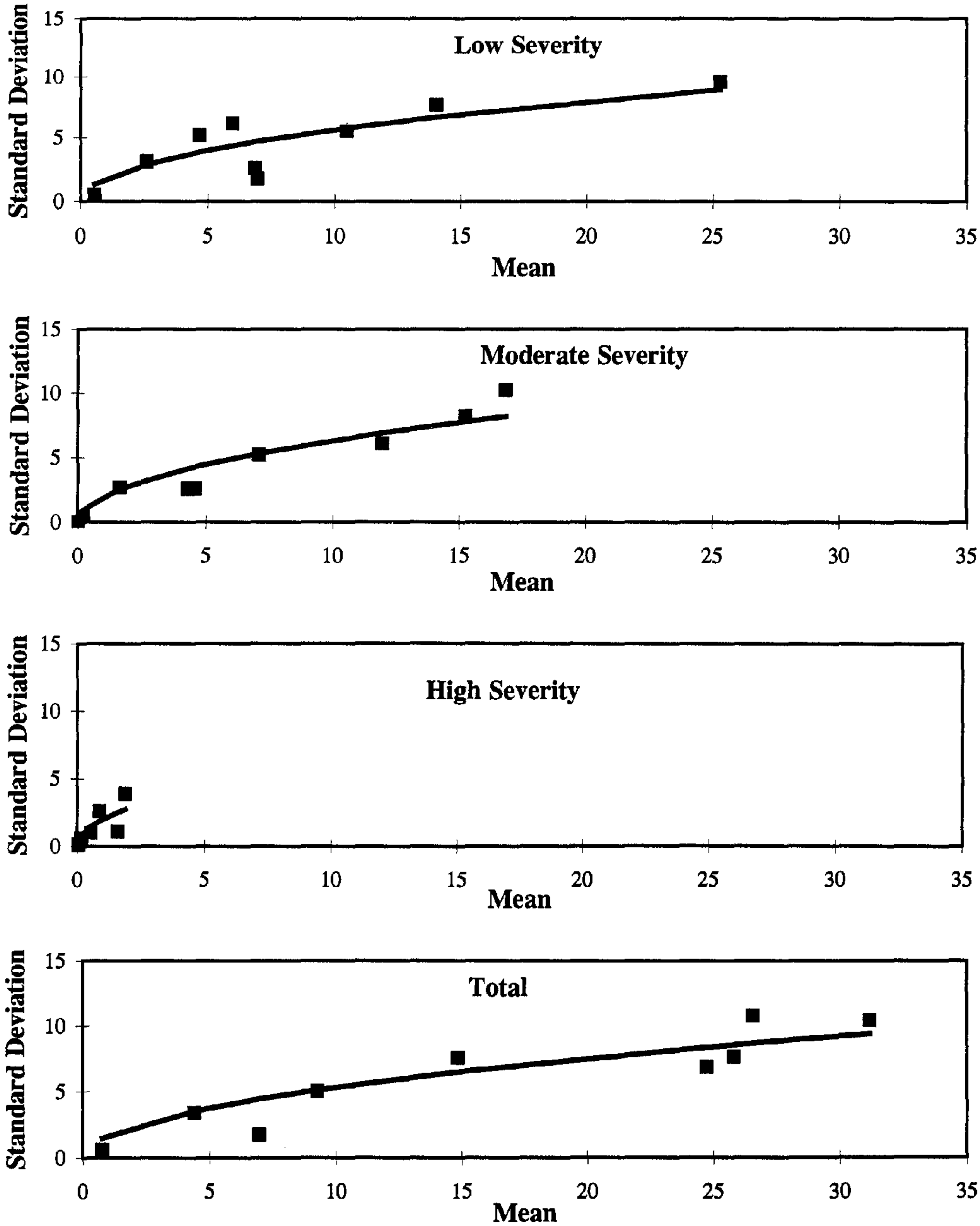


Figure 10. Fatigue Cracking (Sq. Meters) - AC Pavements, Manual Surveys: Standard Deviation Vs. Mean.

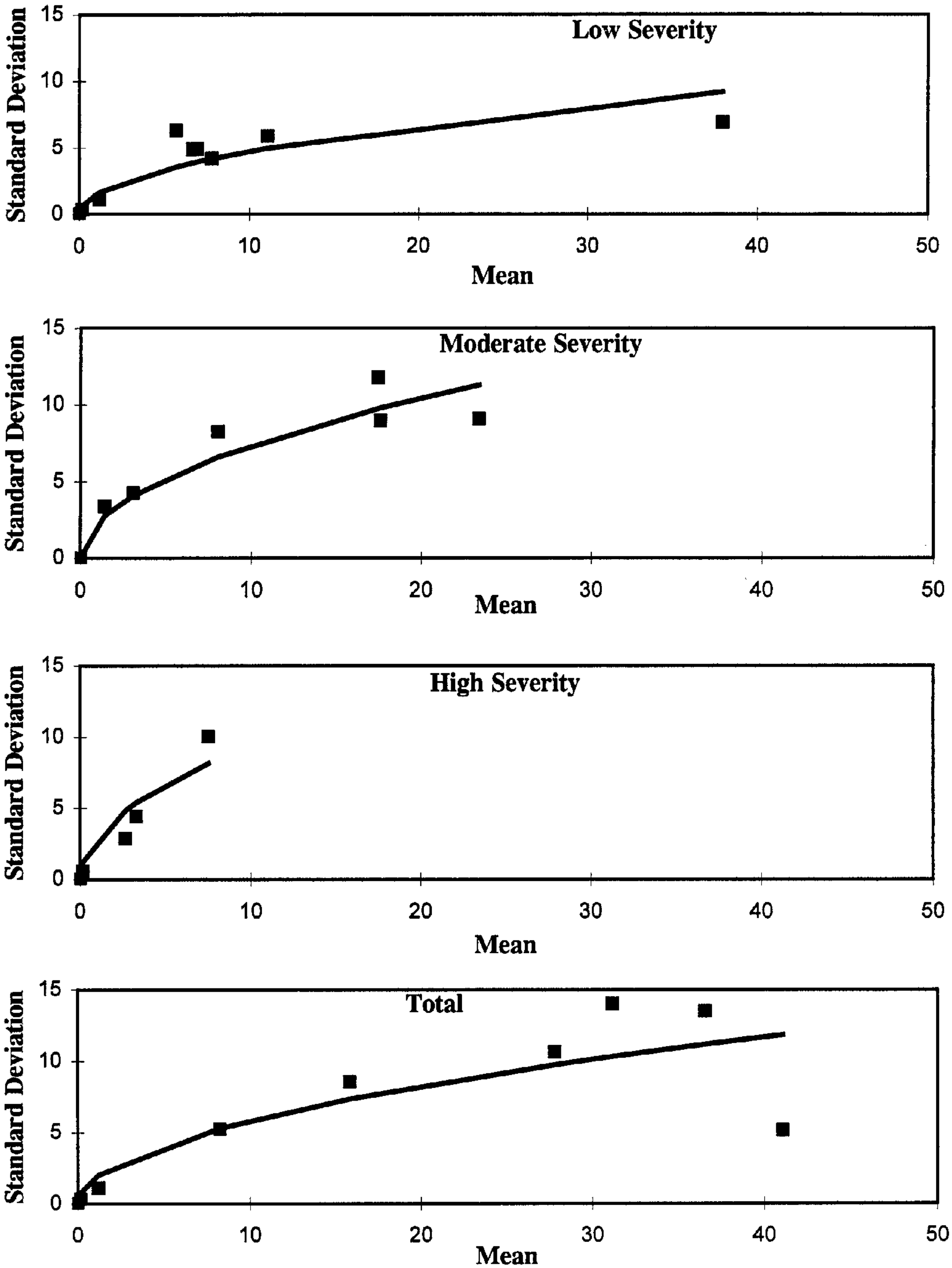


Figure 11. Longitudinal Cracking WP (Meters) - AC Pavements, Manual Surveys: Standard Deviation Vs. Mean.

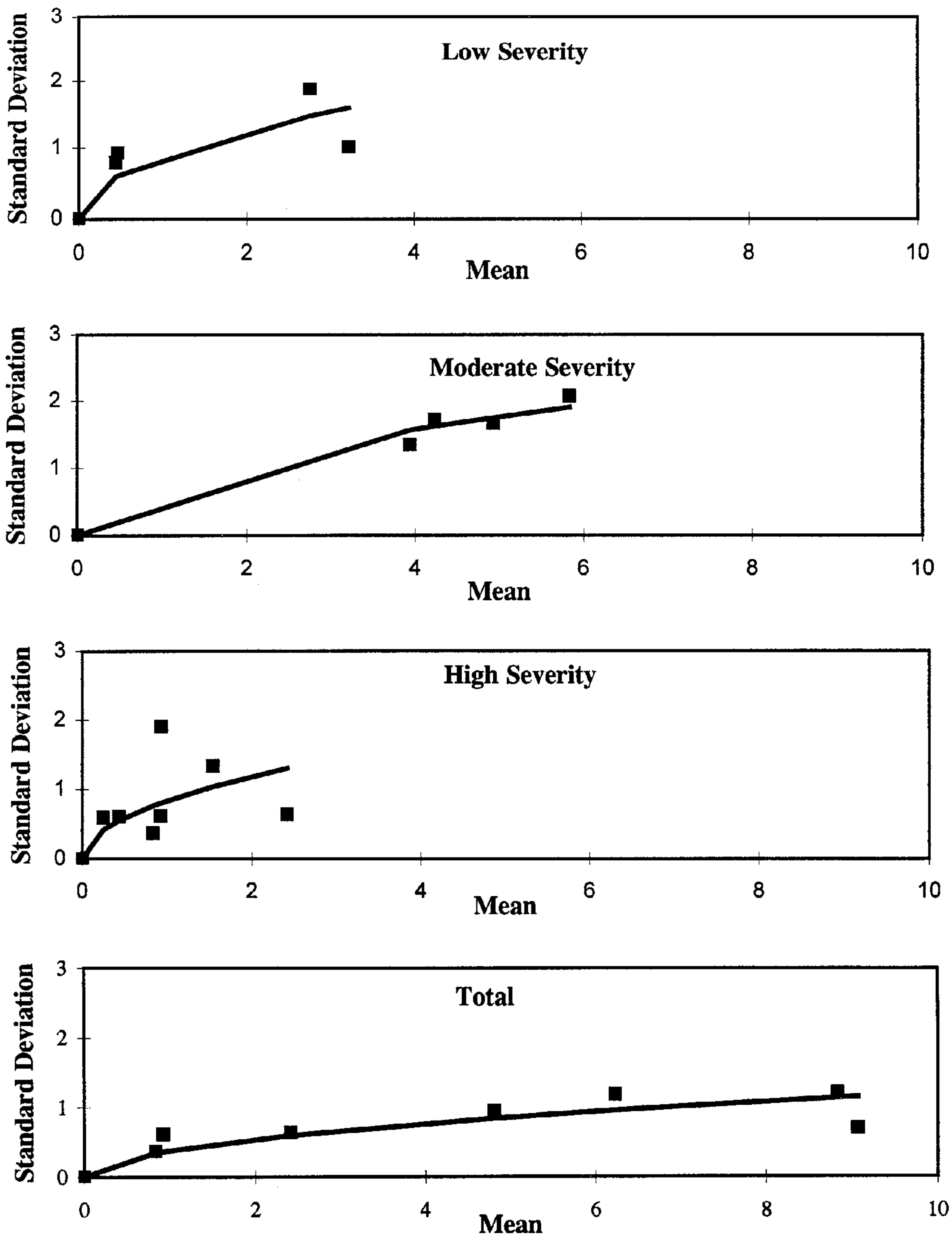


Figure 12. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Standard Deviation Vs. Mean.

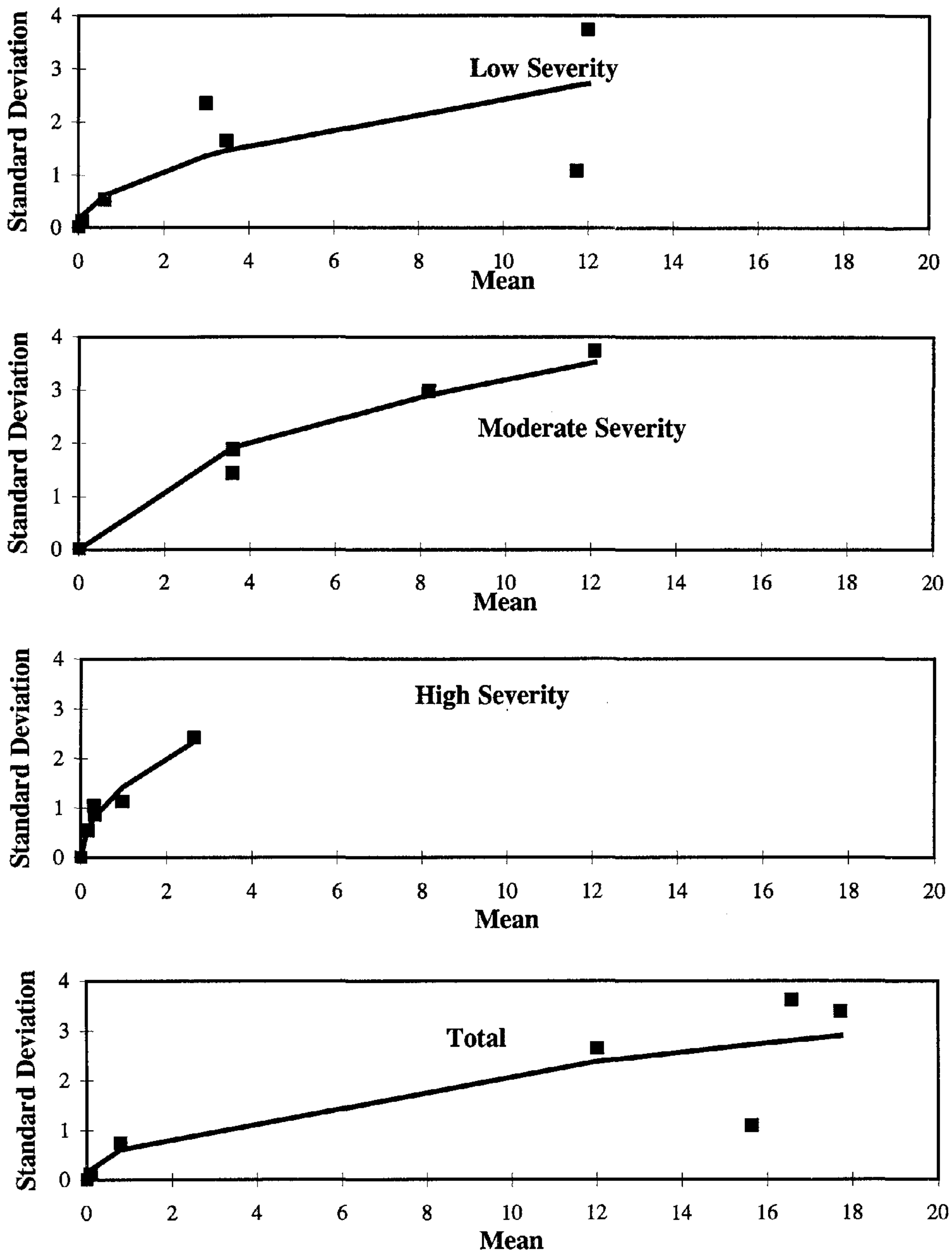


Figure 13. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Standard Deviation Vs. Mean.

mean and reference value was outside the range of three times the derived standard deviation divided by the square root of the number of raters in the workshop. Partial results (total distress only) from this bias analysis are shown in tables 8 and 9 for AC and PCC pavements, respectively. The complete set of analysis results are included in appendix A of this report.

Tables 8 and 9 also show the number of raters before and after the outlier analysis, reference values, workshop means, computed standard deviation from individual data points, derived standard deviation from regression lines, regression equation constants, and their coefficient of determination (R^2); very few data points were identified as outliers. Three symbols are used under the bias column: "N" indicates no bias between group mean and reference value, "Y" indicates that bias does exist, and "NA" indicates that there were no observations.

In general, no bias was observed for AC pavement distresses, except for fatigue cracking where five of the nine workshops indicated some bias. For PCC transverse joint spalling, differences seem to exist between the group mean and reference values; however, these differences often resulted from the small magnitude of mean and reference values, or the small variation (standard deviation) in distress data. Thus, from an engineering viewpoint, one can conclude no bias was observed for PCC pavement distresses.

It can thus be concluded from this three-step analysis that the group mean may be viewed as an unbiased estimate of the reference value. Therefore, assuming a normal distribution and using a 95 percent confidence level, one can state that the true value is bound by the measured value $\pm$ 2*standard deviations, where the standard deviation can be calculated using the regression equations discussed earlier in this section. For example, for AC pavement transverse cracking (meters), the derived standard deviation for a measurement of 30 meters is 3.68 (=0.6716 * sqrt(30)). Or, with a 95 percent confidence level, the true value is within the range of 22.64 to 37.36 meters (=30$\pm$2*3.68).

2.8 Individual Versus Two-Person Team Distress Surveys

It was concluded from analysis of data from the first seven distress workshops that group consensus surveys could lead to an improvement in the bias and precision of manually collected distress data. Accordingly, the last two accreditation workshops (i.e., Workshop Nos. 8 and 9) were designed to incorporate two-person team distress surveys. A complete set of figures comparing the reference, two-person team, and individual distress data from Workshop Nos. 8 and 9, for both AC and PCC pavements, are included in appendix A. Figures 14 through 17 show examples of these plots - fatigue cracking and longitudinal cracking (wheel path) in AC pavements and corner breaks and longitudinal cracking in PCC pavements, respectively. In these plots, the letters "R," "T," and "I" along the X-axis denote the distress values for the reference, two-person team, and individual raters, respectively.

It can be observed from these plots that, although there are improvements in both bias and precision for some distress type-severity level combinations, there are no conclusive trends that support the two-person team surveys as a better method for achieving improved distress data. In many cases, the data scatter is the same, if not worse, when compared with that from the

Table 8. Results of Bias Analysis Based on Square Root Transformation for AC Pavements.

DISTRESS TYPE	UNIT	SECTION ID	DISTRESS SEVERITY	NO. OF RATERS ORIGINAL	NO. OF RATERS OUTLIER	REF.	MEAN	STD DEV	DERIVED STD DEV.	BIAS	SLOPE CONST. C	R^2
Fatigue Cracking	Sq. Meters	AC # 1	TOTAL	12	12	14.90	24.73	6.86	8.37	Y	1.6825	0.827
	Sq. Meters	AC # 2	TOTAL	13	13	14.90	25.80	7.60	8.55	Y		
	Sq. Meters	AC # 3	TOTAL	15	15	5.80	4.39	3.38	3.53	N		
	Sq. Meters	AC # 4	TOTAL	11	11	22.30	31.18	10.43	9.40	Y		
	Sq. Meters	AC # 5	TOTAL	12	12	0.80	0.76	0.58	1.46	N		
	Sq. Meters	AC # 6	TOTAL	13	12	10.00	9.27	5.03	5.12	N		
	Sq. Meters	AC # 7	TOTAL	12	12	8.20	14.84	7.56	6.48	Y		
	Sq. Meters	AC # 8	TOTAL	16	15	47.40	26.55	10.78	8.67	Y		
	Sq. Meters	AC # 9	TOTAL	6	6	3.80	6.97	1.74	4.44	N		
Longitudinal Cracking WP	Meters	AC # 1	TOTAL	12	12	0.00	0.00	0.00	0.00	NA	1.8485	0.720
	Meters	AC # 2	TOTAL	13	13	0.00	0.18	0.34	0.79	N		
	Meters	AC # 3	TOTAL	15	15	16.00	15.83	8.57	7.35	N		
	Meters	AC # 4	TOTAL	11	11	46.40	36.56	13.50	11.18	N		
	Meters	AC # 5	TOTAL	12	12	36.00	41.06	5.16	11.84	N		
	Meters	AC # 6	TOTAL	13	13	36.00	27.81	10.64	9.75	Y		
	Meters	AC # 7	TOTAL	12	12	7.70	8.24	5.25	5.31	N		
	Meters	AC # 8	TOTAL	16	15	21.90	31.17	14.01	10.32	Y		
	Meters	AC # 9	TOTAL	6	6	2.00	1.20	1.10	2.02	N		
Longitudinal Cracking NWP	Meters	AC # 1	TOTAL	12	12	89.10	91.11	10.68	22.85	N	2.3935	0.627
	Meters	AC # 2	TOTAL	13	13	89.10	81.94	7.14	21.67	N		
	Meters	AC # 3	TOTAL	15	14	6.20	6.40	4.09	6.06	N		
	Meters	AC # 4	TOTAL	11	11	39.10	32.12	5.34	13.56	N		
	Meters	AC # 5	TOTAL	12	12	31.40	32.98	5.79	13.74	N		
	Meters	AC # 6	TOTAL	13	13	17.00	20.02	5.04	10.71	N		
	Meters	AC # 7	TOTAL	12	12	23.90	19.97	9.76	10.69	N		
	Meters	AC # 8	TOTAL	16	16	58.20	57.95	14.87	18.22	N		
	Meters	AC # 9	TOTAL	6	6	320.70	289.88	64.24	40.75	N		

Table 8. Results of Bias Analysis Based on Square Root Transformation for AC Pavements (Continued).

DISTRESS TYPE	UNIT	SECTION ID	DISTRESS SEVERITY	NO. OF RATERS ORIGINAL	NO. OF RATERS OUTLIER	REF.	MEAN	STD DEV	DERIVED STD DEV.	BIAS	SLOPE CONST. C	R^2
Transverse Cracking	Meters	AC # 1	TOTAL	12	12	74.60	81.36	5.08	6.06	Y	0.6716	0.655
	Meters	AC # 2	TOTAL	13	13	74.60	76.37	7.53	5.87	N		
	Meters	AC # 3	TOTAL	15	13	16.90	21.91	1.80	3.14	Y		
	Meters	AC # 4	TOTAL	11	11	27.10	24.74	2.91	3.34	N		
	Meters	AC # 5	TOTAL	12	12	18.20	18.72	2.04	2.91	N		
	Meters	AC # 6	TOTAL	13	13	32.00	32.93	2.39	3.85	N		
	Meters	AC # 7	TOTAL	12	12	64.00	56.79	6.39	5.06	Y		
	Meters	AC # 8	TOTAL	16	16	59.90	58.33	5.04	5.13	N		
	Meters	AC # 9	TOTAL	6	6	31.30	30.18	4.70	3.69	N		
Transverse Cracking	No.	AC # 1	TOTAL	12	12	56.00	55.00	5.96	5.81	N	0.7830	0.544
	No.	AC # 2	TOTAL	13	13	56.00	51.23	6.12	5.60	Y		
	No.	AC # 3	TOTAL	15	14	10.00	11.79	2.51	2.69	N		
	No.	AC # 4	TOTAL	11	11	24.00	17.55	1.44	3.28	Y		
	No.	AC # 5	TOTAL	12	12	8.00	9.25	1.83	2.38	N		
	No.	AC # 6	TOTAL	13	13	15.00	13.85	1.70	2.91	N		
	No.	AC # 7	TOTAL	12	12	27.00	23.33	2.78	3.78	Y		
	No.	AC # 8	TOTAL	16	16	26.00	27.38	5.05	4.10	N		
	No.	AC # 9	TOTAL	6	6	16.00	13.83	5.49	2.91	N		

Form of the regression equation:

$$\text{Standard Deviation} = C * (\text{Mean})^{0.5}$$

Table 9. Results of Bias Analysis Based on Square Root Transformation for PCC Pavements.

DISTRESS TYPE	UNIT	SECTION ID	DISTRESS SEVERITY	NO. OF RATERS ORIGINAL	NO. OF RATERS OUTLIER	REF.	MEAN	STD DEV	DERIVED STD DEV.	BIAS	SLOPE CONST.	R^2
Corner Breaks	Number	PCC # 1	TOTAL	12	12	9.00	8.83	1.21	1.14	N	0.3834	0.792
	Number	PCC # 2	TOTAL	14	14	9.00	9.07	0.70	1.15	N		
	Number	PCC # 3	TOTAL	14	13	2.00	0.92	0.62	0.37	Y		
	Number	PCC # 4	TOTAL	11	11	0.00	0.00	0.00	0.00	NA		
	Number	PCC # 5	TOTAL	12	12	3.00	2.42	0.64	0.60	Y		
	Number	PCC # 6	TOTAL	13	13	7.00	6.23	1.19	0.96	N		
	Number	PCC # 7	TOTAL	11	11	0.00	0.00	0.00	0.00	NA		
	Number	PCC # 8	TOTAL	16	16	5.00	4.81	0.95	0.84	N		
	Number	PCC # 9	TOTAL	6	6	0.00	0.83	0.37	0.35	Y		
Longitudinal Cracking	Meters	PCC # 1	TOTAL	12	12	18.80	15.65	1.09	2.72	Y	0.6873	0.801
	Meters	PCC # 2	TOTAL	14	13	18.80	16.58	3.62	2.80	N		
	Meters	PCC # 3	TOTAL	14	14	0.00	0.00	0.00	0.00	NA		
	Meters	PCC # 4	TOTAL	11	11	0.00	0.09	0.12	0.21	N		
	Meters	PCC # 5	TOTAL	12	12	0.00	0.79	0.74	0.61	Y		
	Meters	PCC # 6	TOTAL	13	13	18.20	17.72	3.38	2.89	N		
	Meters	PCC # 7	TOTAL	11	11	0.00	0.00	0.00	0.00	NA		
	Meters	PCC # 8	TOTAL	16	15	12.10	12.01	2.65	2.38	N		
	Meters	PCC # 9	TOTAL	6	6	0.00	0.00	0.00	0.00	NA		
Transverse Cracking	Meters	PCC # 1	TOTAL	12	12	11.10	10.95	2.09	1.82	N	0.5500	0.772
	Meters	PCC # 2	TOTAL	14	13	11.10	11.42	1.90	1.86	N		
	Meters	PCC # 3	TOTAL	14	13	8.70	9.48	1.57	1.69	N		
	Meters	PCC # 4	TOTAL	11	11	44.60	46.41	3.53	3.75	N		
	Meters	PCC # 5	TOTAL	12	12	37.70	39.62	4.04	3.46	N		
	Meters	PCC # 6	TOTAL	13	13	16.20	15.30	1.85	2.15	N		
	Meters	PCC # 7	TOTAL	11	11	7.40	7.28	0.22	1.48	N		
	Meters	PCC # 8	TOTAL	16	15	12.10	13.32	3.04	2.01	N		
	Meters	PCC # 9	TOTAL	6	6	74.10	71.10	4.40	4.64	N		
Transverse Cracking	Number	PCC # 1	TOTAL	12	12	3.00	3.00	0.41	1.05	N	0.6034	0.820
	Number	PCC # 2	TOTAL	14	14	3.00	4.29	1.10	1.25	Y		
	Number	PCC # 3	TOTAL	14	14	7.00	4.50	1.40	1.28	Y		
	Number	PCC # 4	TOTAL	11	11	17.00	19.36	2.80	2.66	N		
	Number	PCC # 5	TOTAL	12	12	18.00	19.25	3.17	2.65	N		
	Number	PCC # 6	TOTAL	13	13	7.00	5.77	1.25	1.45	Y		

Table 9. Results of Bias Analysis Based on Square Root Transformation for PCC Pavements (Continued).

DISTRESS TYPE	UNIT	SECTION ID	DISTRESS SEVERITY	NO. OF RATERS ORIGINAL	NO. OF RATERS OUTLIER	REF.	MEAN	STD DEV	DERIVED STD DEV.	BIAS	SLOPE CONST.	R^2
Transverse Cracking	Number	PCC # 7	TOTAL	11	11	2.00	2.00	0.00	0.85	N		
	Number	PCC # 8	TOTAL	16	16	4.00	5.38	1.80	1.40	Y		
	Number	PCC # 9	TOTAL	6	6	24.00	22.83	2.67	2.88	N		
Spalling of Longitudinal Joints	Meters	PCC # 1	TOTAL	12	12	15.00	14.01	5.68	4.45	N	1.1895	0.734
	Meters	PCC # 2	TOTAL	14	14	15.00	9.96	4.43	3.75	Y		
	Meters	PCC # 3	TOTAL	14	14	3.50	2.91	2.70	2.03	N		
	Meters	PCC # 4	TOTAL	11	10	0.00	0.58	0.57	0.91	N		
	Meters	PCC # 5	TOTAL	12	11	1.30	0.96	1.40	1.17	N		
	Meters	PCC # 6	TOTAL	13	13	1.50	1.87	2.25	1.63	N		
	Meters	PCC # 7	TOTAL	11	10	0.60	0.35	0.34	0.70	N		
	Meters	PCC # 8	TOTAL	16	15	9.30	4.35	1.69	2.48	Y		
	Meters	PCC # 9	TOTAL	6	6	13.60	16.50	3.15	4.83	N		
Spalling of Transverse Joints	Meters	PCC # 1	TOTAL	12	12	0.60	0.12	0.21	0.22	Y	0.6448	0.788
	Meters	PCC # 2	TOTAL	14	12	0.60	0.00	0.00	0.00	NA		
	Meters	PCC # 3	TOTAL	14	13	3.50	1.71	0.83	0.84	Y		
	Meters	PCC # 4	TOTAL	11	10	0.30	1.02	0.46	0.65	Y		
	Meters	PCC # 5	TOTAL	12	9	0.30	0.10	0.16	0.20	N		
	Meters	PCC # 6	TOTAL	13	13	1.10	1.18	0.76	0.70	N		
	Meters	PCC # 7	TOTAL	11	10	5.20	2.70	0.97	1.06	Y		
	Meters	PCC # 8	TOTAL	16	16	1.60	1.51	1.39	0.79	N		
	Meters	PCC # 9	TOTAL	6	6	2.00	5.75	1.37	1.55	Y		
Spalling of Transverse Joints	Number	PCC # 1	TOTAL	12	11	2.00	0.18	0.39	0.28	Y	0.6582	0.188
	Number	PCC # 2	TOTAL	14	12	2.00	0.00	0.00	0.00	NA		
	Number	PCC # 3	TOTAL	14	13	5.00	2.38	1.21	1.02	Y		
	Number	PCC # 4	TOTAL	11	11	1.00	3.27	1.91	1.19	Y		
	Number	PCC # 5	TOTAL	12	10	1.00	0.40	0.49	0.42	Y		
	Number	PCC # 6	TOTAL	13	12	6.00	5.50	2.81	1.54	N		
	Number	PCC # 7	TOTAL	11	11	11.00	10.55	0.66	2.14	N		
	Number	PCC # 8	TOTAL	16	14	2.00	2.21	1.70	0.98	N		
	Number	PCC # 9	TOTAL	6	6	3.00	4.50	0.96	1.40	N		

Form of the regression equation:

$$\text{Standard Deviation} = C * (\text{Mean})^{0.5}$$

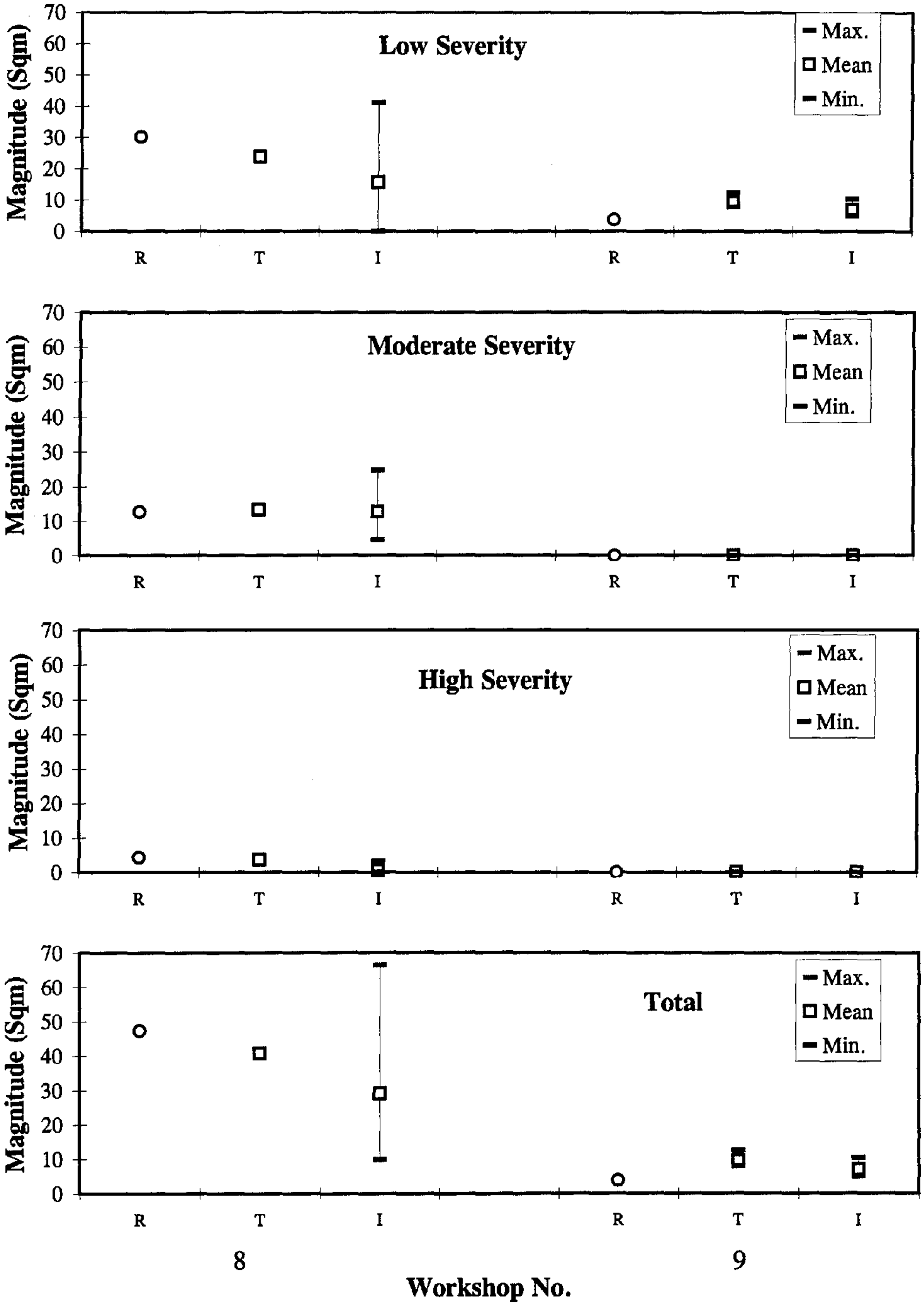


Figure 14. Fatigue Cracking (Sq. Meters) - AC Pavements, Manual Surveys: Reference, Team, and Individual Values.

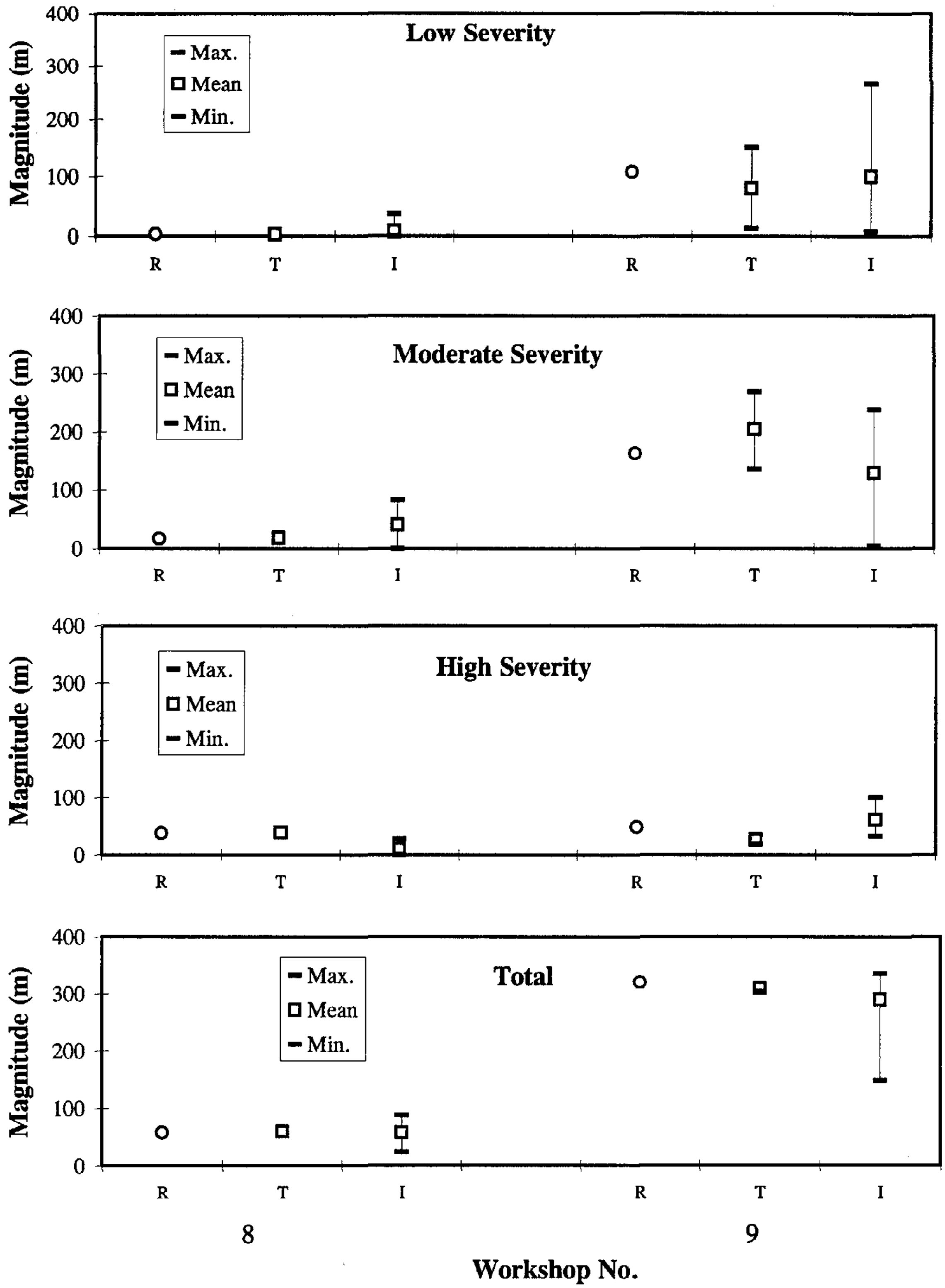


Figure 15. Longitudinal Cracking NWP (Meters) - AC Pavements, Manual Surveys: Reference, Team, and Individual Values.

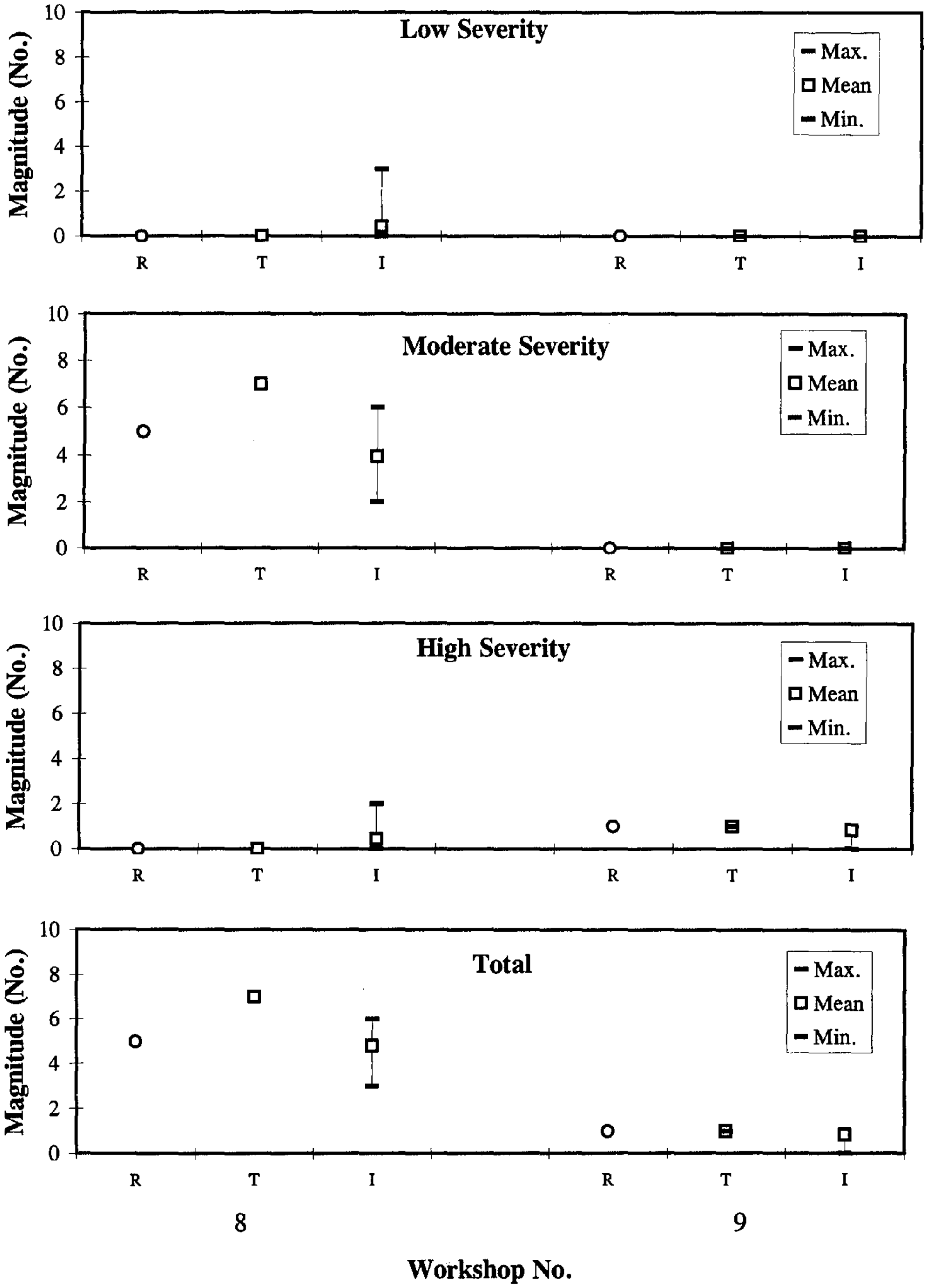


Figure 16. Corner Breaks (No.) - PCC Pavements, Manual Surveys: Reference, Team, and Individual Values.

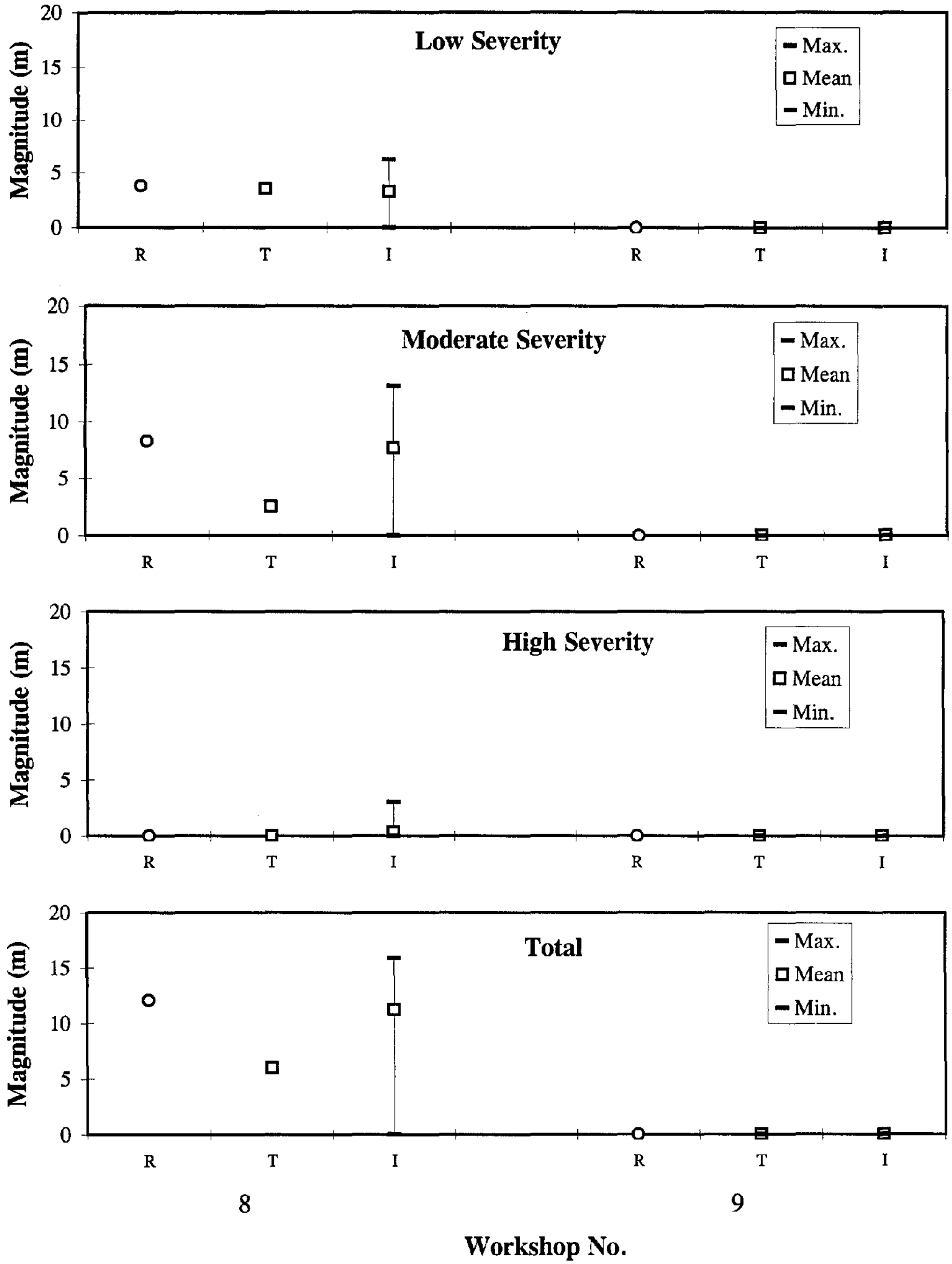


Figure 17. Longitudinal Cracking (Meters) - PCC Pavements, Manual Surveys: Reference, Team, and Individual Values.

individual raters. Thus, the anticipated improvement in bias and precision using two-person consensus surveys does not appear to be supported by the data generated from the last two workshops. In all fairness, however, it must also be recognized that these observations are based on limited data and that additional data are needed to more conclusively arrive at a conclusion relative to the ability of two-person consensus surveys to provide better distress data.

2.9 Summary and Conclusions

One of the primary purposes of the LTPP surface distress monitoring efforts is to provide detailed, distress-specific condition data for use in the development of distress-specific performance prediction models. The LTPP program is relying on both the combination of photographic distress survey technology and the manual distress data to achieve this objective.

Because manual distress surveys are conducted by individual raters whose biases can lead to variability between raters, it was hypothesized that distress data variability existed and that it could potentially be quite large. Thus, the purpose of the study presented in this chapter was to quantify manual distress data variability, with special emphasis on the bias and precision of the data.

Results from nine LTPP rater accreditation workshops conducted during 1992 to 1996 were used as the only source of data in this study. On the basis of analyses of these data, numerous observations and/or conclusions were made. The most important ones are summarized below:

- Individual rater variability for any given distress type-severity level combination is typically large and increases as the distress quantity increases. Also, there is a decrease in variability with an increase in the magnitude of the mean as captured in the coefficient of variation.

- Total distress group means are generally close to the reference value and the scatter of the individual raters is narrower than for the individual distress severity levels, which indicates significant differences in distinguishing severity levels.

- Both apparent bias and precision for the common distress type-severity level combinations were quantified. The apparent bias is small and not uniform, i.e., there is no tendency to consistently rate all distress type and severity level combinations higher or lower. However, the associated precision or variability is very sensitive to the magnitude and range of distress quantities present on a section. The CV ranges from less than 10 percent to well in excess of 100 percent, although it generally decreases with increased distress quantities. Large CV values were observed only when low quantities of a given pavement distress are present. Thus, the large variability indicated by these values may not necessarily indicate poor precision.

- There is a significant improvement in precision for total distress quantities when compared with individual severity levels. The CV in some instances is still high if viewed individually but, as noted before, that is primarily attributable to low distress magnitude and therefore may not necessarily indicate poor precision.

- When all distress type-severity level combinations are viewed in terms of a single composite number such as the PCI value, there is excellent agreement among the individual raters, the group mean, and the reference value. The individual rater variability is also small when viewed in terms of this composite value.

- There does not appear to be a decrease in rater variability resulting from rater attendance at more than one accreditation workshop.

- A limited study was conducted to assess the potential improvement in distress data bias and precision through the use of two-person consensus surveys; however, no clear trends emerged from the data analysis to confirm this anticipated improvement.

The question of acceptable level of distress data variability depends on the intended use of the data. For purposes of pavement management systems, especially at the network level (e.g., PCI threshold values for triggering maintenance and rehabilitation needs or PCI values for describing the overall health of the network), the large distress type-severity level variability identified from the accreditation workshops appears to be acceptable. However, for uses in research directed at developing distress prediction models and similar applications, it is the authors' contention that the level of variability found in this study should be reduced to increase the potential of using these data in the development of such models.

3. PASCO/PADIAS DISTRESS DATA VARIABILITY

3.1 Introduction

Distress data collection for the LTPP began with the decision by SHRP to use PASCO 35-mm black and white photography to obtain frequent, objective surveys of the 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 of surveys 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. Two such processes have been used to date for reduction of distress data from 35-mm film - PADIAS v1.x and PADIAS v4.x systems. The study presented in this chapter was undertaken by the FHWA to assess the level of variability associated with distress data interpreted from film using the PADIAS v4.x system, to define the bias and precision for these data, and to compare differences in observed distresses between data produced by this system and those obtained from the PADIAS v1.x system in the early years of the LTPP program.

3.2 Background

One of the decisions made by the LTPP program under SHRP was to collect distress data using photographic means. Photography was established as the primary distress data collection procedure, producing 35-mm black and white strip photographs of all LTPP test sections, at varying frequencies, over the period from 1989 to 1996. This film was considered as the primary data source; however, the information needed for analyses (the types and amounts of distress recorded by the photographs) had to be determined through tedious, semi-automated processing.

Until 1992, data reduction from film had been accomplished by the SHRP P-001B Technical Assistance Contractor using software developed for SHRP. This system was relatively crude in the resolution of measurements: 0.30-m increments for linear defects and 0.09-square-m increments for areal distresses. In addition, this software was based on early versions of the DIM and was not fully in agreement with the May 1993 version of that manual.

In 1992, FHWA was charged with continuation of the LTPP and awarded a contract for continued collection of distress photography. As part of this contract the vendor, PASCO USA, was tasked with providing both film and reduced distress data. During evaluation of the software system proposed by the vendor, it was determined that the accuracy and repeatability of the data reduction process was poor, and therefore FHWA declined to authorize the vendor to perform data reduction; however, filming continued through the Summer of 1996. In order to continue collecting distress data, manual distress surveys became the preferred means. As a result of this evaluation and other feedback from some data users, questions arose about the quality of the data reduced under SHRP.

Because 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, and much of the film had not yet been interpreted, FHWA decided in 1996 to proceed with data reduction from this film using the improved PASCO v4.x software developed by PASCO USA. PADIAS v4.x incorporates the 1993 DIM procedures, with minor exceptions, and is vector based so that accuracy is significantly better than previous software versions. As a result of this on-going effort, film-derived distress data exist in the LTPP IMS for both PADIAS v1.x and v4.x; however, their quality is unknown. Hence, this study was undertaken to assess the variability of distress data derived from film using PADIAS v4.x, the system currently in use, and to compare data generated by this system with those from the PADIAS v1.x system, which was used to generate the early LTPP distress data.

3.3 Data Source

During the last two manual distress accreditation workshops (Reno 1996 and Champaign 1996), reference surveys were conducted by the instructor group on all six test sections per workshop (3 AC and 3 PCC test sections; not just the two accreditation test sections), which yielded reference distress data for a total of 12 test sections (6 per workshop). Each of the 12 test sections was also filmed by PASCO approximately one month before the workshop.

Using the PADIAS v4.x software, a group effort was performed to establish consensus values for the film-derived distress data. (Note: these values were assumed to be the best possible set of values for distress data derived from film using PADIAS v4.x). This work was accomplished using film analysis equipment and software located at PASCO USA, in Harrisburg, Pennsylvania. The group consisted of expert raters, two of whom were accreditation workshop instructors, who accomplished the data reduction through a true consensus effort by observing the films at the same time and discussing and deciding the type and severity level.

Distress data reduction from film was also done by six individual raters from PASCO USA responsible for production work; they were not a part of the consensus surveys. These individuals independently performed distress data collection from film for 6 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.

Thus, the data available for assessing the variability of distress data derived from film using the PADIAS v4.x system consisted of:

- reference surveys for 12 test sections obtained through manual data collection.

- consensus surveys for 12 test sections obtained through interpretation of film using the PADIAS v4.x system.
- individual rater surveys for six test sections obtained through interpretation of film using the PADIAS v4.x system.
- two-person team consensus surveys for six test sections (different from those used in individual ratings) obtained through interpretation of film using the PADIAS v4.x system.

For comparison of distress data derived from film using the PADIAS v4.x system versus that from the PADIAS v1.x system, the following guidelines were used to develop the assessment data set. Using data stored in the IMS, test sections having the following characteristics were selected:

- All three pavement types - asphalt surfaced, jointed concrete, continuously reinforced concrete
- Distresses that challenge identification or quantification - low fatigue in AC, transverse cracks in CRC, corner breaks, rigid patches
- High amount of distress

Using the above criteria, a total of 24 test sections were identified for use in the PADIAS v1.x versus PADIAS v4.x comparison; these sections are summarized in table 10. For both v1.x and v4.x, distress data were interpreted using PASCO's production procedure, which consisted of film interpretation by two individual raters and then a third rater refereed the interpretation to determine the final severity and amount of each particular distress. The production procedure was used only in the comparison of v1.x and v4.x distress data and in section 3.8, but not in other analyses conducted in this study.

Distress data reports from the PADIAS v1.x were generated (both maps and summaries) for these 24 test sections from the data available at the LTPP IMS; the actual interpretation of the film had been completed during the SHRP years. PASCO USA also digitized the films for these test sections using PADIAS v4.x and its revised production methodology, including the use of multiple operators, multiple repeats and comparison and correction of discrepancies. The resulting PADIAS v1.x and v4.x data sets served as the basis for comparison of the two methods.

3.4 PADIAS v4.x Distress Data Variability Study

To assess the variability of distress data derived from film using the PADIAS v4.x system, a series of analyses were performed using various subjective (plots) and statistical methods. The first of these analyses looked at the repeatability of data generated by the individual experts, individual raters, and two-person teams; three repeat surveys were performed by each of these groups on 12, 6, and 6 test sections, respectively. An analysis of variance (ANOVA) test was then 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 also conducted to assess whether significant differences existed between the groups. In this exercise, distress

Table 10. Test Sections for PADIAS v1.x Vs. PADIAS v4.x Comparison.

Pavement Type (Predominant Distress)	Test Section ID	Survey Date
AC Surfaced		
(Transverse and Longitudinal Cracking)	41034	20-NOV-89
	201005	10-Mar-89
	271016	22-Jun-89
	483689	04-MAR-90
(Low Fatigue)	82008	20-Aug-91
	810506	17-May-90
(Transverse Cracking only)	169032	17-Jul-89
	417018	26-Jul-89
Jointed PCC		
(Transverse and Longitudinal Cracking)	40601	21-Nov-89
	40603	21-Nov-89
	40608	21-Nov-89
	63005	7-Sep-89
	209037	5-Dec-90
(Corner Breaks)	124000	13-Apr-89
	94092	4-Sep-90
	214025	26-Oct-89
Continuously Reinforced PCC		
(Transverse Cracks >35)	105005	21-Mar-91
	175849	24-Jun-89
	415021	26-Jul-89
(Transverse Cracks <35)	395003	3-Oct-90
	485035	25-Jan-90
	245807	11-Oct-89
(Rigid Patches)	265363	5-Sep-89
	195046	18-May-91

data generated by the experts was first compared with that from the individual raters and then to that from the two-person teams. The last statistical analysis, a component of variance analysis, was performed to identify sources of variations within each group. To gain a better understanding of the magnitude of data variability, global trends were investigated through the use of plots. Finally, bias and precision statements were developed for all distress type-severity level combinations where sufficient data were available.

Repeatability Study

Although not a standard practice within the LTPP program, three repeat interpretations were performed on the study test sections by each expert, individual rater, and two-person team involved in the assessment of distress data derived from film using the PADIAS v4.x system. The purpose of the repeat measurements, and hence this analysis, was to assess the consistency of interpretations made by the same person or team at different times, i.e., Do results from the same person or team change if multiple interpretations are performed? To accomplish this, plots of distress quantity at each severity level and total across all severity levels for a distress type were developed for each of the common distress types identified at the 12 (6 AC and 6 PCC) pavement test sections used in this study. For a given distress type and severity level combination, the following values are plotted:

- **Reference value** - Quantity of distress, for each distress type and severity level, determined by the consensus manual field condition survey conducted by three experts. These reference values are considered a surrogate of ground truth and were used in this study to estimate the potential bias and precision of the LTPP distress raters.

- **Consensus value** - Quantity of distress, for each distress type and severity level, determined by the consensus survey conducted by three experts using the PASCO/PADIAS method.

- **Minimum, mean, and maximum** - Distress quantities, for each distress type and severity level, derived from the three repetitions conducted by each of the experts, raters, and teams using the PASCO/PADIAS method.

The letters "R," "C," "E1" to "E4," "I1" to "I6," and "T1" to "T3" along the X-axis of these plots denote those values pertaining to the reference, consensus, individual experts, individual raters, and two-person team surveys, respectively. Examples of these plots are shown in figures 18 through 21. Figures 18 and 19 show the fatigue cracking plots for the six AC test sections, while figures 20 and 21 show the corner break plots for the six PCC test sections. The complete set of plots showing the repeatability of the PADIAS v4.x distress data are contained in appendix B of this report. The following observations were made from the information contained in these plots:

- Although variability of the three repetitions by each expert, individual rater, and two-person team has not been quantified, it appears reasonable. Data consistency varies

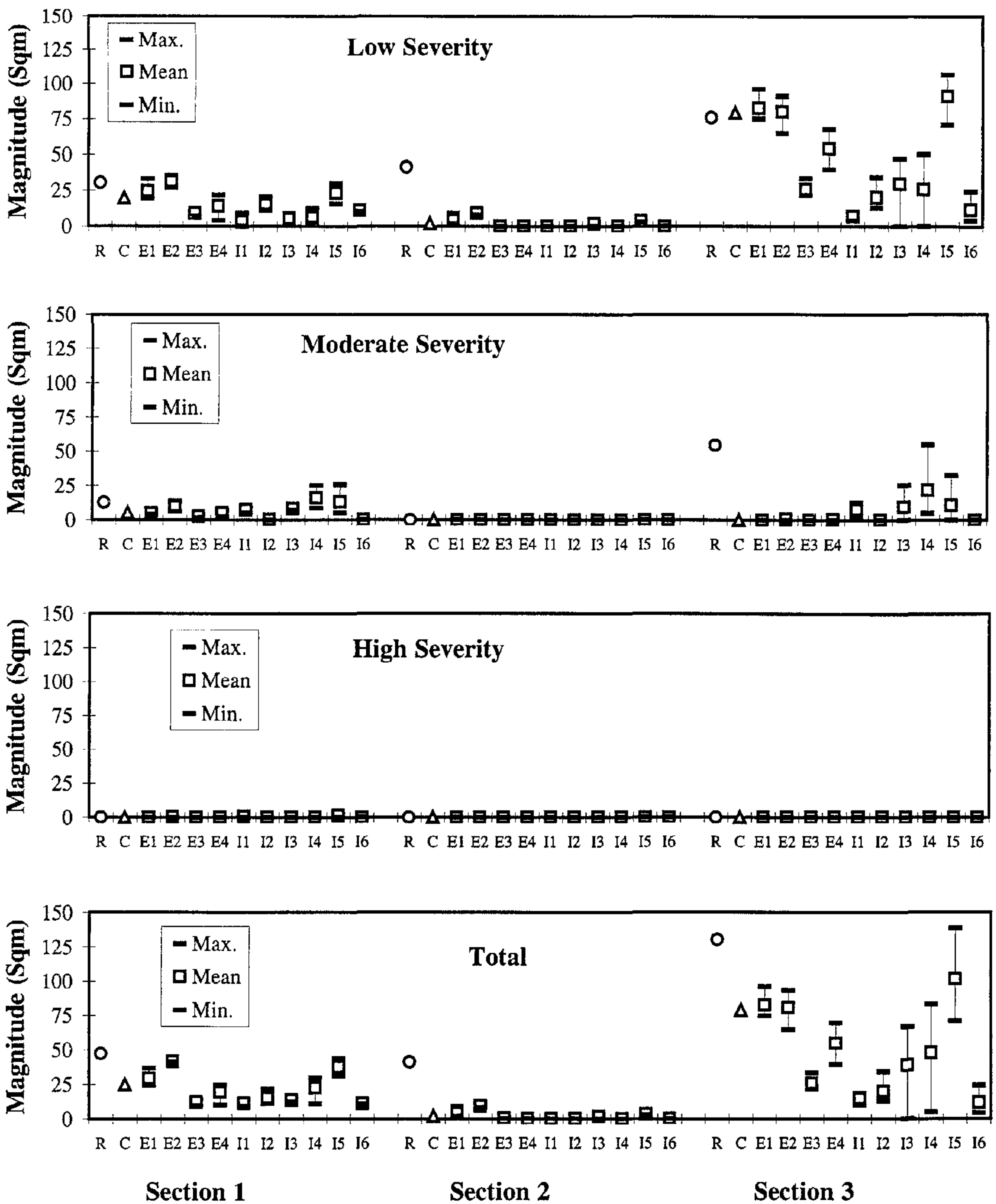


Figure 18. Fatigue Cracking (Sq. Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Individual Raters for Three Repetitions.

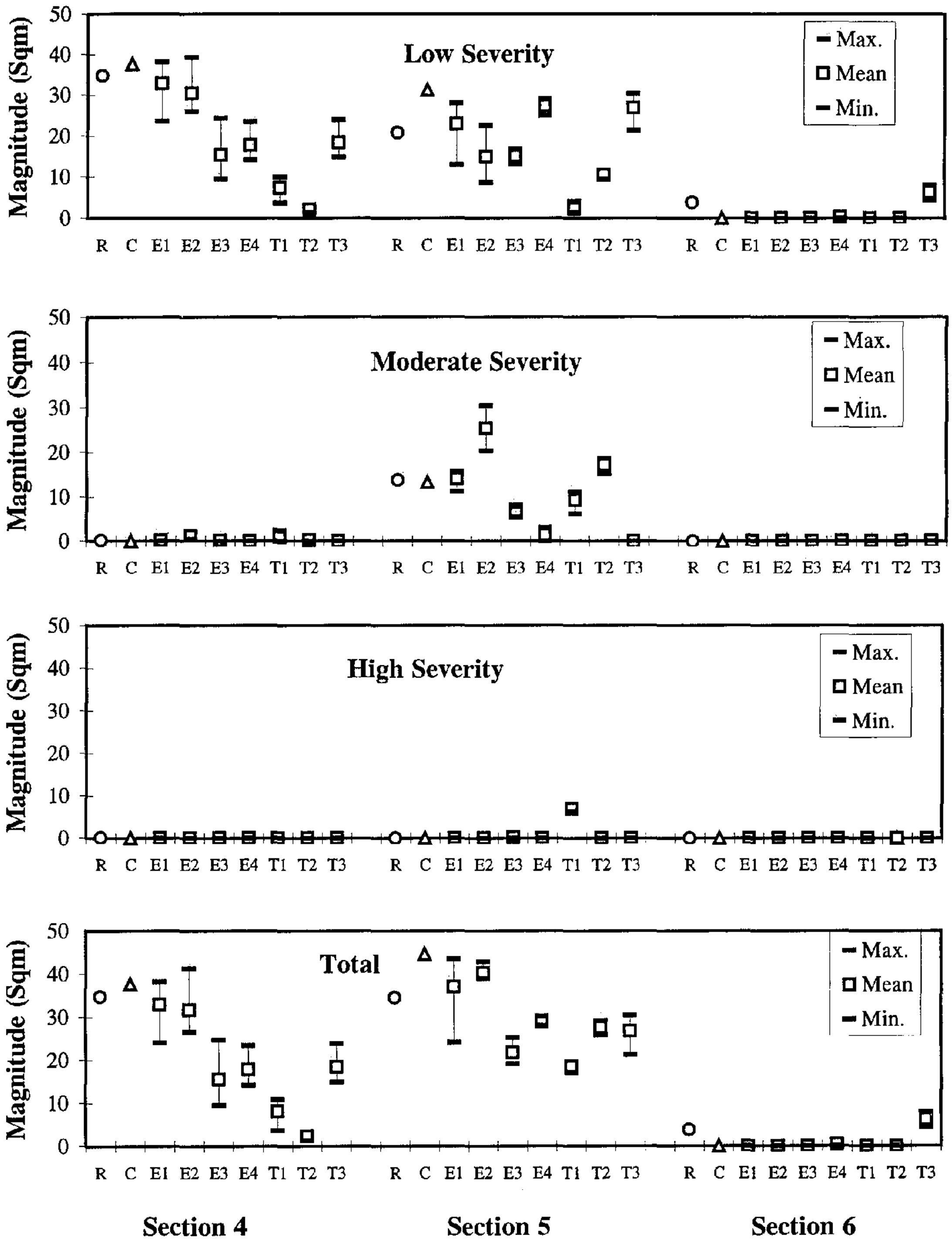


Figure 19. Fatigue Cracking (Sq. Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Teams for Three Repetitions.

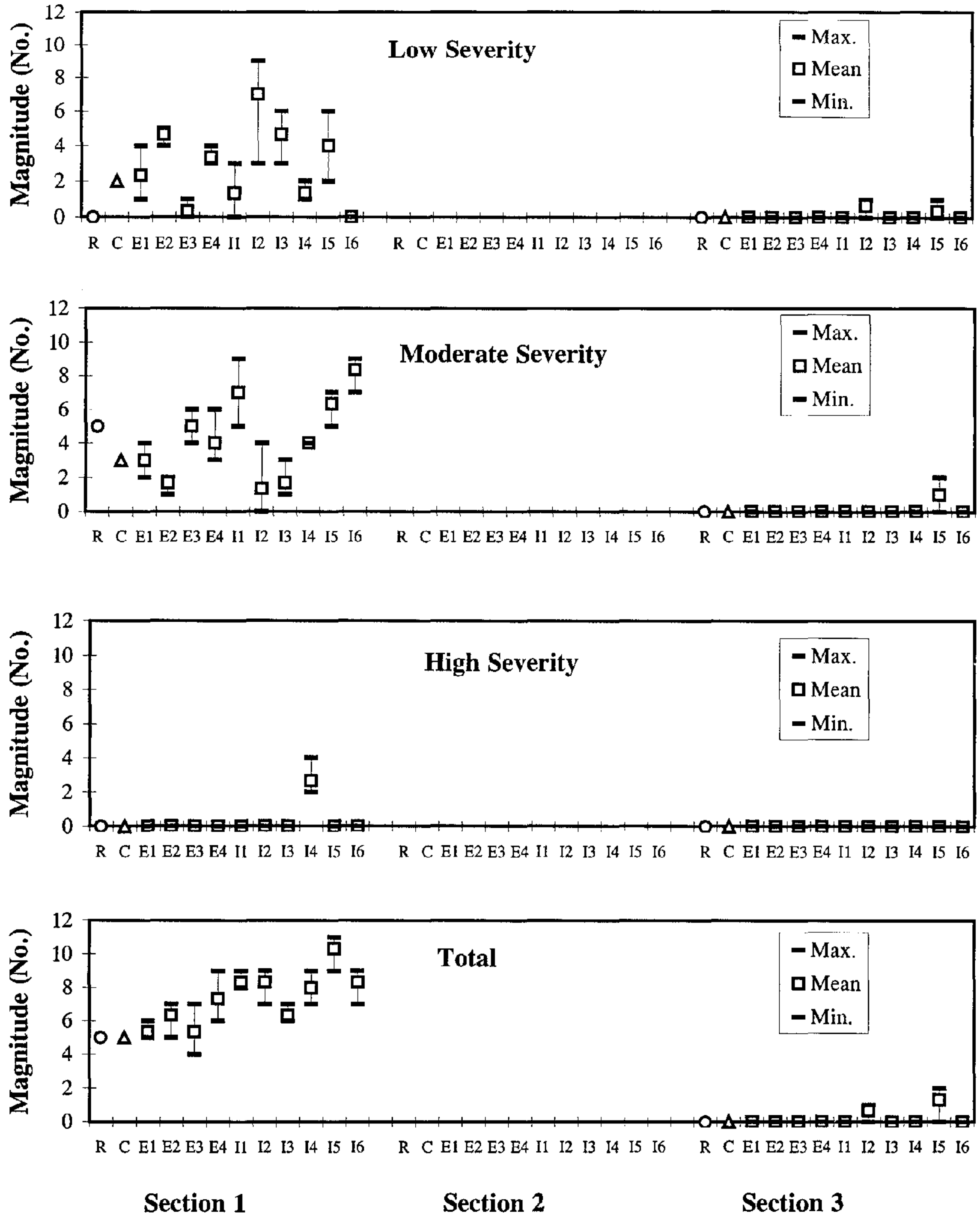


Figure 20. Corner Breaks (No.) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Individual Raters for Three Repetitions.

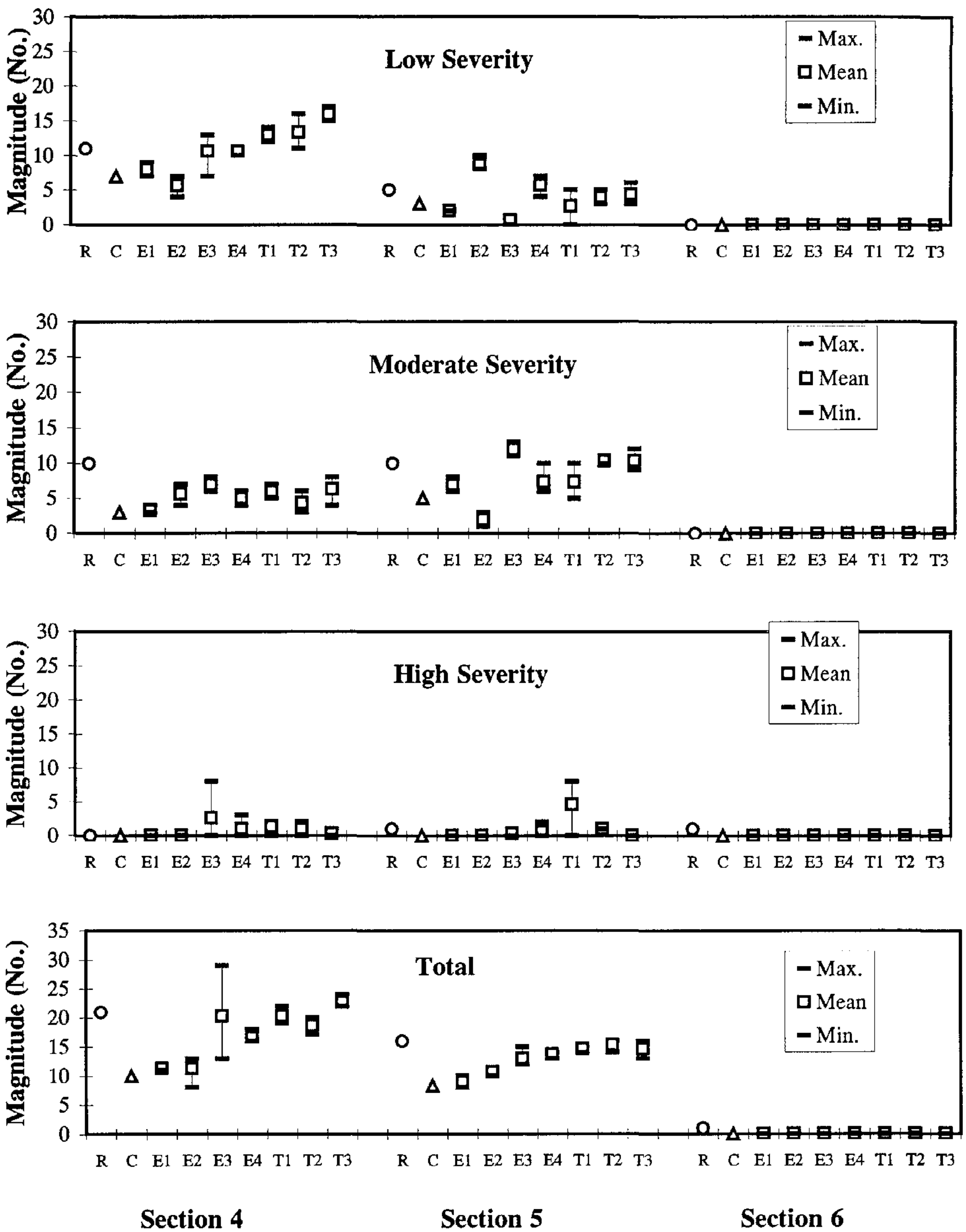


Figure 21. Corner Breaks (No.) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, and Minimum, Mean, and Maximum Values for Experts and Teams for Three Repetitions.

from one individual or team to another and the magnitude of variability tends to increase as the quantity of distress increases.

- Distress data interpreted by the individual experts appear more repeatable than those collected by the individual raters. Some large changes in magnitude from one repeat measurement to another were observed on distress data interpreted by individual raters, i.e., poorer consistency. In contrast, the repeatability of those data interpreted by the two-person teams appears to be slightly better than that of the individual experts.

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 percent, then the hypothesis was rejected - i.e., within group means were not statistically the same.

Tables 11 and 12 show the calculated F-statistics for AC and PCC pavement sections, while the results of the ANOVA test are presented in tables 13 and 14 for these two pavement types, respectively. Three symbols are used in these tables; "N," "Y," and "NA." The letter "N" denotes that the mean values for a given distress type and severity level combination are the same within the group in question; "Y" 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; and "NA" indicates that the particular distress type was not observed by any of the experts, individuals raters, and/or two-person teams. In addition, several cells in table 14 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 variance within each of the groups. For AC test sections 1, 2, and 3, the individual raters agreed with one another only 47 percent of the time, while the experts agreed 43 percent of the time - agreement here refers to equal within group means. For PCC test sections 1, 2 and 3, individual raters agreed 60 percent of the time with one another, while experts agreed 52 percent 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 percent of the time; two-person teams also agreed with each other 45 percent of the time. For PCC pavement sections 4, 5, and 6, the within group agreement was 82 percent and 68 percent 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.)

It was also observed from these data that the level of agreement appears to get worse as the quantity of distress increases, regardless of pavement section or group.

Table 11. Calculated F-Statistics 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	7.30	6.97	10.96	10.50	14.92	8.85	4.43	18.13	3.42	51.12	1.00	37.56
		Mod.	4.36	3.80	NA	NA	0.64	0.88	5.61	0.81	36.04	63.19	NA	NA
		High	1.00	1.69	NA	NA	NA	NA	NA	NA	3.86	281.60	NA	NA
		Total	17.67	10.26	10.96	10.50	13.72	4.81	4.55	15.25	5.91	8.21	1.00	37.56
Long. Cracking - NWP	Meters	Low	13.99	11.62	3.21	6.27	22.14	2.84	3.72	0.52	7.30	6.50	3.49	15.66
		Mod.	2.94	3.84	1.02	2.25	1.03	2.38	4.64	3.48	1.56	97.46	9.71	5.19
		High	0.64	7.49	1.00	NA	NA	1.18	1.75	1.00	2.95	15.67	2.34	2.85
		Total	16.39	2.13	4.68	6.62	16.61	1.81	1.67	0.48	3.50	0.05	2.04	9.05
Long. Cracking - WP	Meters	Low	34.42	10.06	1.84	2.68	19.88	4.53	5.52	29.00	6.58	1.00	2.23	NA
		Mod.	12.00	6.75	1.00	NA	0.22	3.85	1.86	2.59	16.63	5.23	NA	NA
		High	0.64	1.94	NA	NA	1.00	1.04	NA	NA	NA	1.67	NA	NA
		Total	38.05	5.63	1.71	2.68	15.44	6.22	1.72	21.94	3.87	2.58	2.23	NA
Trans. Cracking	Meters	Low	15.86	5.59	4.15	2.84	10.22	4.04	5.10	4.07	3.14	7.45	4.53	5.26
		Mod.	4.68	2.47	1.64	12.96	3.05	1.76	5.96	9.95	3.22	1.66	3.10	0.84
		High	3.26	2.63	NA	NA	1.38	1.38	1.00	NA	7.32	20.08	NA	NA
		Total	12.33	3.03	4.05	2.55	24.80	4.92	10.42	2.63	2.64	6.12	4.13	1.14
Trans. Cracking	No.	Low	20.03	5.56	11.19	5.71	6.11	4.21	15.24	3.94	18.26	20.09	9.17	5.00
		Mod.	5.90	2.86	2.00	12.00	4.03	2.41	6.66	11.08	5.02	2.40	3.23	0.70
		High	2.81	1.60	NA	NA	0.84	2.19	1.00	NA	8.89	52.27	NA	NA
		Total	7.06	1.41	10.79	5.03	17.37	2.87	11.58	2.28	7.44	8.69	14.18	1.97

Table 12. Calculated F-Statistics for PCC Sections - Within Group.

DISTRESS TYPE	UNITS	SEV.	PCC1 EXPERT	PCC1 INDIV.	PCC2 EXPERT	PCC2 INDIV.	PCC3 EXPERT	PCC3 INDIV.	PCC4 EXPERT	PCC4 TEAM	PCC5 EXPERT	PCC5 TEAM	PCC6 EXPERT	PCC6 TEAM
Corner Breaks	No.	Low	12.00	5.88			NA	2.10	4.98	2.92	39.42	0.72	NA	NA
		Mod.	4.56	11.38			NA	3.00	5.98	1.35	24.04	3.00	NA	NA
		High	NA	16.00			NA	NA	0.78	0.88	0.73	3.13	NA	NA
		Total	1.74	5.14			NA	3.36	3.08	6.06	11.88	0.33	NA	NA
Long. Cracking	Meters	Low	2.10	0.56	1.76	1.71	1.00	0.99	1.03	11.21	1.99	10.69	0.71	25.00
		Mod.	5.13	1.04	5.91	13.26	NA	1.00	1.00	NA	14.76	0.27	NA	NA
		High	NA	0.99	0.58	3.99	NA	NA	1.00	NA	0.67	1.00	NA	NA
		Total	1.90	1.00	2.49	2.00	1.00	0.99	1.02	11.21	1.07	90.28	0.71	25.00
Trans. Cracking	No.	Low	4.18	1.51	1.87	2.97	7.08	15.68	0.35	3.00	1.36	3.00	2.01	17.45
		Mod.	3.97	4.42	3.10	2.40	48.46	0.91	3.00	4.00	2.31	4.00	2.01	1.85
		High	1.00	0.69	1.00	NA	62.69	6.33	1.00	1.00	3.60	12.00	2.20	5.78
		Total	9.84	0.78	2.59	2.23	4.95	1.66	0.37	3.00	0.87	1.40	1.14	10.07
Trans. Cracking	Meters	Low	2.95	2.43	5.08	2.79	6.04	21.50	0.17	1.79	1.11	0.85	1.94	30.29
		Mod.	4.21	4.84	3.65	2.22	12.51	0.92	3.90	4.00	2.22	4.08	5.74	3.74
		High	1.00	0.50	1.00	NA	50.31	6.61	1.00	1.00	3.67	13.95	2.12	5.46
		Total	15.67	0.99	6.65	9.06	4.08	0.96	0.29	3.00	1.86	3.19	2.58	83.71
Spalling of Long. Joints	Meters	Low	9.87	4.17	0.67	2.61	18.58	2.04	7.24	3.57	13.44	23.19	2.98	0.96
		Mod.	1.50	1.29	2.43	2.79	5.40	1.25	1.00	7.98	5.97	2.94	1.31	NA
		High	NA	NA	3.30	4.90	1.06	21.00	0.92	NA	1.00	1.00	0.86	1.00
		Total	9.88	3.53	5.63	4.30	18.50	10.01	1.06	13.30	14.91	50.85	4.92	0.95
Spalling of Trans. Joints	No.	Low	0.67	0.84			1.00	7.70	0.14	1.50	4.00	0.88	NA	NA
		Mod.	2.25	0.65			9.00	1.00	12.00	1.00	0.67	0.50	1.07	NA
		High	65535.00	4.80			3.33	4.00	2.80	1.50	1.00	NA	7.56	1.50
		Total	3.33	0.75			1.33	9.20	0.24	0.50	0.61	0.33	1.11	1.50
Spalling of Trans. Joints	Meters	Low	0.41	0.73			0.79	3.14	0.26	0.38	2.22	0.94	1.00	1.00
		Mod.	1.78	0.94			9.27	0.83	2.66	0.10	1.54	2.38	1.06	9.70E+15
		High	768.00	7.43			8.73	2.71	2.46	4.46	1.18	4.00	2.81	7.91
		Total	4.51	0.75			1.33	9.20	0.60	8.49	1.15	1.38	2.68	11.28

Table 13. 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	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y
		Mod.	Y	Y	NA	NA	N	N	Y	N	Y	Y	NA	NA
		High	N	N	NA	NA	NA	NA	NA	NA	N	Y	NA	NA
		Total	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
Long. Cracking - NWP	Meters	Low	Y	Y	N	Y	Y	N	N	N	Y	Y	N	Y
		Mod.	N	Y	N	N	N	N	Y	N	N	Y	Y	Y
		High	N	Y	N	NA	NA	N	N	N	N	Y	N	N
		Total	Y	N	Y	Y	Y	N	N	N	N	N	N	Y
Long. Cracking - WP	Meters	Low	Y	Y	N	N	Y	Y	Y	Y	Y	N	N	NA
		Mod.	Y	Y	N	NA	N	Y	N	N	Y	Y	NA	NA
		High	N	N	NA	NA	N	N	NA	NA	NA	N	NA	NA
		Total	Y	Y	N	N	Y	Y	N	Y	N	N	N	NA
Trans. Cracking	Meters	Low	Y	Y	Y	N	Y	Y	Y	N	N	Y	Y	Y
		Mod.	Y	N	N	Y	N	N	Y	Y	N	N	N	N
		High	N	N	NA	NA	N	N	N	NA	Y	Y	NA	NA
		Total	Y	N	N	N	Y	Y	Y	N	N	Y	Y	N
Trans. Cracking	No.	Low	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N
		Mod.	Y	N	N	Y	N	N	Y	Y	Y	N	N	N
		High	N	N	Y	Y	N	N	N	Y	Y	Y	NA	NA
		Total	Y	N	Y	Y	Y	N	Y	N	Y	Y	Y	N

Note: N = equal group means
Y = at least one different group mean
NA = no distress observed

Table 14. Results of Analysis of Variance (ANOVA) for PCC 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	Y	Y			NA	N	Y	N	Y	N	NA	NA
		Mod.	Y	Y			NA	N	Y	N	Y	N	NA	NA
		High	NA	Y			NA	NA	N	N	N	N	NA	NA
		Total	N	Y			NA	Y	N	Y	Y	N	NA	NA
Long. Cracking	Meters	Low	N	N	N	N	N	N	N	Y	N	Y	N	Y
		Mod.	Y	N	Y	Y	NA	N	N	NA	Y	N	NA	NA
		High	NA	N	N	Y	NA	NA	N	NA	N	N	NA	NA
		Total	N	N	N	N	N	N	N	Y	N	Y	N	Y
Trans. Cracking	No.	Low	Y	N	N	N	Y	Y	N	N	N	N	N	Y
		Mod.	N	Y	N	N	Y	N	N	N	N	N	N	N
		High	N	N	N	NA	Y	Y	N	N	N	Y	N	Y
		Total	Y	N	N	N	Y	N	N	N	N	N	N	Y
Trans. Cracking	Meters	Low	N	N	Y	N	Y	Y	N	N	N	N	N	Y
		Mod.	Y	Y	N	N	Y	N	N	N	N	N	Y	N
		High	N	N	N	NA	Y	Y	N	N	N	Y	N	Y
		Total	Y	N	Y	Y	Y	N	N	N	N	N	N	Y
Spalling of Long. Joints	Meters	Low	Y	Y	N	N	Y	N	Y	N	Y	Y	N	N
		Mod.	N	N	N	N	Y	N	N	Y	Y	N	N	NA
		High	NA	NA	N	Y	N	Y	N	NA	N	N	N	N
		Total	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N
Spalling of Trans. Joints	No.	Low	N	N			N	Y	N	N	N	N	NA	NA
		Mod.	N	N			Y	N	Y	N	N	N	N	NA
		High	Y	Y			N	Y	N	N	N	NA	Y	N
		Total	N	N			N	Y	N	N	N	N	N	N
Spalling of Trans. Joints	Meters	Low	N	N			N	Y	N	N	N	N	N	N
		Mod.	N	N			Y	N	N	N	N	N	N	Y
		High	Y	Y			Y	N	N	N	N	N	N	Y
		Total	Y	N			N	Y	N	Y	N	N	N	Y

Note: N = equal group means
Y = at least one different group mean
NA = no distress observed

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, individual raters, and two-person teams for each distress type. In this exercise, distress data generated by the experts was first compared with that from the individual raters and then with 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 percent, then the hypothesis was rejected - i.e., mean of two groups being compared were not statistically the same.

Comparisons were first made between the means from the individual experts with 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. (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.)

The computed t-statistics and critical t-values are shown in tables 15 and 16 for AC and PCC pavements, respectively. Three symbols are used in these tables; "N," "Y," and "NA." The letter "N" denotes that the between group means being compared for a given distress type and severity level combination are statistically the same, "Y" indicates that the means are not statistically the same, and "NA" indicates that the particular distress type was not observed by any of the experts, individuals raters, and/or two-person teams. In addition, several cells in table 16 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)

Table 15. Results of t-test for AC Pavement Sections.

Distress Type	Units	SEV.	AC1			AC2			AC3			AC4			AC5			AC6		
			Calcu. t-stat.	Critical t-stat.		Calcu. t-stat.	Critical t-stat.		Calcu. t-stat.	Critical t-stat.		Calcu. t-stat.	Critical t-stat.		Calcu. t-stat.	Critical t-stat.		Calcu. t-stat.	Critical t-stat.	
Fatigue Cracking	Sq. M	Low	2.51	2.09	Y	1.94	2.16	N	2.77	2.05	Y	3.81	2.09	Y	1.59	2.16	N	1.84	2.31	N
		Mod.	0.90	2.06	N	NA	NA	NA	2.29	2.11	Y	0.10	2.15	N	0.79	2.09	N	NA	NA	NA
		High	0.91	2.05	N	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.93	2.31	N	NA	NA	NA
		Total	1.57	2.08	N	1.94	2.16	N	1.82	2.05	N	3.76	2.09	Y	2.42	2.10	Y	1.84	2.31	N
Long. Cracking NWP	Meters	Low	0.18	2.07	N	2.24	2.10	Y	2.40	2.05	Y	2.82	2.26	Y	1.09	2.12	N	0.60	2.16	N
		Mod.	0.13	2.05	N	4.65	2.06	Y	2.88	2.09	Y	1.51	2.09	N	2.17	2.23	N	2.28	2.09	Y
		High	1.58	2.06	N	1.00	2.20	N	2.61	2.11	Y	0.82	2.09	N	1.73	2.26	N	0.35	2.20	N
		Total	0.97	2.09	N	2.83	2.10	Y	3.84	2.05	Y	3.06	2.31	Y	3.28	2.09	Y	0.59	2.15	N
Long. Cracking WP	Meters	Low	2.06	2.05	Y	3.28	2.07	Y	1.76	2.05	N	3.89	2.11	Y	2.68	2.09	Y	1.29	2.20	N
		Mod.	3.09	2.06	Y	1.00	2.20	N	2.77	2.10	Y	0.03	2.09	N	1.10	2.23	N	NA	NA	NA
		High	1.67	2.09	N	NA	NA	NA	2.76	2.11	Y	NA	NA	NA	1.20	2.31	N	NA	NA	NA
		Total	4.60	2.06	Y	3.26	2.07	Y	3.17	2.05	Y	3.91	2.12	Y	3.52	2.09	Y	1.29	2.20	N
Trans. Cracking	Meters	Low	0.17	2.06	N	1.03	2.16	N	0.13	2.05	N	3.13	2.09	Y	0.44	2.09	N	2.05	2.18	N
		Mod.	0.17	2.06	N	0.40	2.08	N	1.85	2.05	N	2.33	2.18	Y	1.52	2.12	N	0.07	2.10	N
		High	1.10	2.05	N	NA	NA	NA	2.17	2.08	Y	1.00	2.20	N	0.49	2.15	N	NA	NA	NA
		Total	2.46	2.08	Y	1.03	2.16	N	0.26	2.05	N	2.35	2.09	Y	0.66	2.16	N	1.27	2.13	N
Trans. Cracking	No.	Low	0.48	2.05	N	1.01	2.10	N	0.81	2.06	N	3.99	2.11	Y	1.63	2.16	N	2.08	2.15	N
		Mod.	0.37	2.06	N	0.00	2.05	N	0.37	2.05	N	0.44	2.09	N	1.14	2.12	N	0.57	2.11	N
		High	1.20	2.06	N	NA	NA	NA	2.40	2.09	Y	1.00	2.21	N	0.26	2.16	N	NA	NA	NA
		Total	2.74	2.05	Y	1.00	2.16	N	0.01	2.05	N	4.58	2.11	Y	1.33	2.10	N	1.93	2.15	N

Note: N = equal group means
Y = unequal group means
NA = no distress observed

Table 16. 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	N				1.84	2.11	N	5.20	2.09	Y	0.51	2.11	N	NA	NA	NA
		Mod.	1.60	2.05	N				1.37	2.11	N	0.42	2.11	N	1.69	2.10	N	NA	NA	NA
		High	1.72	2.11	N				NA	NA	NA	0.04	2.13	N	1.63	2.26	N	NA	NA	NA
		Total	4.19	2.06	Y				2.06	2.11	N	3.28	2.13	Y	4.63	2.11	Y	NA	NA	NA
Long. Cracking	Meters	Low	0.23	2.05	N	1.52	2.07	N	1.01	2.11	N	1.52	2.20	N	1.13	2.09	N	0.28	2.09	N
		Mod.	0.31	2.09	N	2.44	2.08	Y	1.00	2.11	N	1.72	2.20	N	5.29	2.12	Y	NA	NA	NA
		High	1.02	2.11	N	0.11	2.06	N	NA	NA	NA	1.00	2.20	N	0.34	2.10	N	NA	NA	NA
		Total	0.03	2.11	N	1.11	2.06	N	1.01	2.11	N	1.74	2.20	N	4.04	2.10	Y	0.28	2.09	N
Trans. Cracking	No.	Low	1.10	2.05	N	1.52	2.07	N	2.32	2.05	Y	5.37	2.18	Y	0.73	2.13	N	0.58	2.16	N
		Mod.	2.99	2.05	Y	1.29	2.09	N	1.04	2.10	N	2.28	2.09	Y	1.18	2.09	N	3.43	2.11	Y
		High	0.90	2.06	N	1.00	2.20	N	0.95	2.06	N	0.28	2.10	N	0.33	2.10	N	0.27	2.11	N
		Total	2.31	2.05	Y	0.68	2.07	N	1.15	2.15	N	6.15	2.16	Y	0.62	2.10	N	0.36	2.16	N
Trans. Cracking	Meters	Low	0.07	2.05	N	2.13	2.06	Y	2.54	2.05	Y	3.35	2.20	Y	0.76	2.15	N	1.58	2.09	N
		Mod.	2.89	2.05	Y	1.19	2.09	N	1.53	2.11	N	2.05	2.09	N	0.47	2.10	N	4.01	2.16	Y
		High	1.53	2.05	N	1.00	2.20	N	1.15	2.06	N	0.48	2.18	N	0.33	2.10	N	0.22	2.09	N
		Total	2.07	2.09	N	2.36	2.05	Y	0.31	2.18	N	3.62	2.18	Y	0.08	2.15	N	0.59	2.11	N
Spalling of Long. Joints	Meters	Low	0.68	2.08	N	7.27	2.20	Y	5.24	2.20	Y	1.17	2.16	N	0.39	2.13	N	4.48	2.18	Y
		Mod.	1.76	2.11	N	2.10	2.16	N	0.12	2.06	N	1.73	2.26	N	2.03	2.31	N	1.18	2.20	N
		High	NA	NA	NA	0.76	2.13	N	2.88	2.07	Y	1.12	2.20	N	0.59	2.23	N	0.89	2.16	N
		Total	1.06	2.08	N	7.73	2.06	Y	5.43	2.18	Y	0.64	2.10	N	0.82	2.15	N	5.54	2.18	Y
Spalling of Trans. Joints	No.	Low	1.33	2.11	N				1.22	2.06	N	1.20	2.13	N	2.19	2.23	N	NA	NA	NA
		Mod.	0.71	2.06	N				2.28	2.15	Y	0.45	2.18	N	0.30	2.12	N	1.77	2.20	N
		High	0.48	2.06	N				2.30	2.11	Y	0.39	2.09	N	1.00	2.20	N	5.12	2.10	Y
		Total	1.27	2.06	N				2.92	2.05	Y	1.42	2.15	N	1.73	2.16	N	6.89	2.09	Y
Spalling of Trans. Joints	Meters	Low	2.22	2.11	Y				0.22	2.11	N	0.17	2.11	N	2.17	2.13	Y	0.73	2.26	N
		Mod.	0.53	2.05	N				1.52	2.09	N	0.32	2.09	N	0.95	2.11	N	0.12	2.09	N
		High	0.19	2.07	N				1.96	2.07	N	0.35	2.10	N	1.32	2.16	N	1.04	2.09	N
		Total	1.78	2.07	N				1.88	2.11	N	0.29	2.09	N	1.67	2.11	N	0.91	2.11	N

Note: N = equal group means
Y = unequal group means
NA = no distress observed

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

The within rater component can be estimated any time the rater repeats 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 that 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.

Unlike the manual distress data, the PASCO/PADIAS distress data available for this study were sufficiently adequate to allow for conduct of the component 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 chapter, i.e., 6 AC and 6 PCC test sections. Tables 17 and 18 present the analysis results for the 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.

As indicated earlier in this report, distress measurement variability can be quantified using the coefficient of variation (CV). Because of concerns over the impact of small distress quantities, 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). See appendix B for a complete set of plots. 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 22 to 25; figures 22 and 23 show plots for fatigue cracking and longitudinal cracking (in wheel path) for AC pavements, while figures 24 and 25 show similar plots for corner breaks and longitudinal cracking in PCC pavements, respectively. 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 resulting CV values and related statistics are summarized in table 19 for AC pavements and in table 20 for PCC pavements. The following observations were made on the basis of information presented in these two tables and the referenced CV plots:

- Both the between rater coefficient of variation (CVa) and within rater coefficient of variation (CVw) values seem to vary widely, ranging from close to 0 percent to more than 300 percent. However, the larger CV's are primarily associated with those distress

Table 17. Mean, σa, and σw Values for Experts, Individuals, and Teams for AC Pavements; PASCO Method.

DISTRESS	UNITS	SEV.	SEC.	Expert Mean	Expert σa	Expert σw	Inidividual Mean	Inidividual σa	Inidividual σw	Team Mean	Team σa	Team σw
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	5.2	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	34.8			
			#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
Long. Cracking WP	Meters	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			
			#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 17. Mean, σa, and σw Values for Experts, Individuals, and Teams for AC Pavements; PASCO Method (Continued).

				Expert			Inidividual			Team		
DISTRESS	UNITS	SEV.	SEC.	Mean	σa	σw	Mean	σa	σw	Mean	σa	σw
Long. Cracking NWP	Meters	Low	#1	30.0	22.9	11.0	31.5	20.2	10.8			
			#2	323.6	13.7	16.0	256.9	103.1	77.7			
			#3	17.5	17.2	6.5	35.8	15.6	19.9			
			#4	116.0	7.6	8.0				74.9	0.0	45.6
			#5	37.3	18.9	13.1				48.8	22.3	16.5
			#6	117.0	34.6	37.9				134.1	75.5	34.2
		Mod.	#1	31.9	7.2	8.9	31.2	13.3	13.7			
			#2	26.7	0.9	10.5	8.4	5.8	9.0			
			#3	3.2	0.3	3.6	13.7	8.5	12.5			
			#4	4.2	2.8	2.5				2.2	1.8	2.0
			#5	7.8	2.6	6.0				19.9	17.8	3.1
			#6	102.9	46.5	27.3				62.0	26.2	22.2
		High	#1	2.7	0.0	3.6	6.2	7.4	5.0			
			#2	0.4	0.0	1.3	0.0	0.0	0.0			
			#3	0.0	0.0	0.0	1.3	0.5	2.1			
			#4	1.0	0.9	1.7				0.4	0.0	1.2
			#5	0.6	0.8	1.0				3.0	4.2	1.9
			#6	6.8	3.5	5.3				8.5	8.3	10.6
		Total	#1	62.4	19.2	8.5	68.9	8.3	13.6			
			#2	350.7	10.6	9.5	265.4	106.5	77.8			
			#3	20.8	16.3	7.1	1.3	0.5	2.1			
			#4	121.2	3.1	6.5				77.5	0.0	45.4
			#5	45.7	14.9	16.3				71.7	0.0	17.3
			#6	224.6	32.0	54.3				204.6	82.4	50.3
Trans. Cracking	Number	Low	#1	10.3	6.4	2.5	11.6	6.6	5.4			
			#2	83.1	11.8	6.4	78.9	7.2	5.8			
			#3	61.1	20.5	15.7	53.2	21.1	20.4			
			#4	58.3	19.9	9.2				93.9	15.2	15.3
			#5	22.8	10.5	4.4				32.3	15.9	6.3
			#6	12.7	5.5	3.3				8.8	1.7	1.5
		Mod	#1	14.1	4.5	3.5	14.8	3.6	4.5			
			#2	0.3	0.3	0.6	0.3	0.8	0.4			
			#3	38.3	11.9	11.8	35.9	11.1	16.2			
			#4	3.9	5.0	3.6				2.9	4.8	2.6
			#5	19.4	6.8	5.9				16.2	2.4	3.5
			#6	0.8	1.4	1.6				0.4	0.0	1.1
		High	#1	5.9	3.1	4.0	8.1	2.0	4.6			
			#2	0.0	0.0	0.0	0.0	0.0	0.0			
			#3	6.6	0.0	4.0	16.8	11.0	14.1			
			#4	0.1	0.0	0.3				0.0	0.0	0.0
			#5	3.9	2.8	1.8				3.4	5.3	1.3
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		Total	#1	30.3	2.8	2.0	34.6	1.8	4.9			
			#2	83.4	12.1	6.7	79.2	6.7	5.8			
			#3	105.8	14.1	6.1	16.8	11.0	14.1			
			#4	62.3	16.2	8.6				96.8	9.7	14.9
			#5	46.1	9.4	6.4				52.0	9.0	5.6
			#6	13.5	7.1	3.4				9.2	1.1	1.9

Table 17. Mean, σa, and σw Values for Experts, Individuals, and Teams for AC Pavements; PASCO Method (Continued).

DISTRESS	UNITS	SEV.	SEC.	Expert			Inidividual			Team		
				Mean	σa	σw	Mean	σa	σw	Mean	σa	σw
Trans. Cracking	Meters	Low	#1	12.8	10.5	4.7	13.5	9.8	7.9			
			#2	126.7	18.2	17.8	119.2	5.9	7.5			
			#3	63.8	33.4	19.1	65.8	36.9	36.7			
			#4	55.0	13.4	11.5				75.4	10.1	10.0
			#5	31.6	17.7	21.0				36.0	18.2	12.4
			#6	33.9	10.1	9.3				25.9	2.5	2.1
		Mod.	#1	33.5	8.7	7.8	34.2	7.5	10.7			
			#2	0.3	0.2	0.5	0.2	0.4	0.2			
			#3	91.6	14.3	17.3	70.2	18.9	37.5			
			#4	7.8	8.0	6.2				1.2	2.0	1.2
			#5	38.3	12.7	14.7				29.0	3.7	7.9
			#6	3.2	5.2	6.2				2.9	0.0	7.6
		High	#1	22.1	7.7	8.9	27.2	4.4	13.5			
			#2	0.0	0.0	0.0	0.0	0.0	0.0			
			#3	20.9	4.2	11.9	43.3	20.3	36.1			
			#4	0.0	0.0	1.0				0.0	0.0	0.0
			#5	8.1	6.0	4.2				6.3	10.2	4.0
			#6	0.0	0.0	0.0				0.0	0.0	0.0
		Total	#1	68.4	7.3	3.8	75.0	4.2	5.1			
			#2	127.0	18.3	18.2	119.4	5.4	7.4			
			#3	176.3	27.8	9.9	179.3	28.0	24.5			
			#4	63.1	13.6	7.7				76.7	7.5	10.1
			#5	78.0	21.0	28.0				71.3	7.5	5.8
			#6	37.0	15.2	14.9				28.8	1.7	7.8

Notes: σa = among rater standard deviation
σw = within rater standard deviation

Table 18. Mean, σa, and σw values for Experts, Individuals, and Teams for PCC Pavements; PASCO Method.

				Expert			Inidividual			Team		
DISTRESS	UNITS	SEV.	SEC.	Mean	σa	σw	Mean	σa	σw	Mean	σa	σw
Corner Breaks	Number	LOW	#1	2.7	1.7	0.9	3.1	2.4	1.9			
			#3	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				3.7	0	1.8
			#6	0	0	0				0	0	0
		MOD	#1	3.4	1.3	1.2	4.8	2.8	1.5			
			#3	0	0	0	0.2	0.3	0.4			
			#4	5.3	1.4	1.1				5.6	0.5	1.6
			#5	7.1	4	1.4				9.3	1.4	1.7
			#6	0	0	0				0	0	0
		HIGH	#1	0	0	0	0.4	1.1	0.5			
			#3	0	0	0	0	0	0			
			#4	9.2	0	2.5				0.9	0.3	0.9
			#5	0.3	0	0.6				1.9	2	2.4
			#6	0	0	0				0	0	0
		TOTAL	#1	6.08	0.62	1.26	8.28	1.14	0.97			
			#3	0	0	0	0.33	0.47	0.53			
			#4	14.92	3.62	4.34				20.56	1.84	1.41
			#5	11.58	2.06	1.08				14.89	0	1.15
			#6	0	0	0				0	0	0
Long. Cracking	Meters	LOW	#1	3.7	1.2	2	3.9	0	3.4			
			#2	7.4	4.9	9.7	19.7	14.1	28.9			
			#3	0.2	0	0.5	8.3	0	34.2			
			#4	4.8	1	10.3				0.3	0.3	0.2
			#5	13.8	2.7	4.7				16	3.9	2.2
			#6	0.7	0	1.8				0.9	1.5	0.5
		MOD	#1	6.3	2.8	2.4	5	1.9	17.1			
			#2	5.1	2.6	2	2.5	2.5	1.2			
			#3	0	0	0	0	0	0.1			
			#4	0.6	0	1.3				0	0	0
			#5	12.8	5.3	2.5				3.7	0	2.5
			#6	0	0	0				0	0	0
		HIGH	#1	0	0	0	9.1	0	3.8			
			#2	1.4	0	2.2	1.3	1.8	1.8			
			#3	0	0	0	0	0	0			
			#4	0.1	0	0.4				0	0	0
			#5	0.5	0	1.2				0.3	0	1
			#6	0	0	0				0	0	0
		TOTAL	#1	9.98	1.1	2.01	9.81	0	23.38			
			#2	13.95	7.59	10.76	23.52	16.86	29.15			
			#3	0.15	0	0.52	8.27	0	34.23			
			#4	5.52	0.88	10.4				0.28	0.33	0.18
			#5	27.04	0.62	4.1				20.04	4.26	0.78
			#6	0.7	0	1.78				0.89	1.51	0.53

Table 18. Mean, σa, and σw values for Experts, Individuals, and Teams for PCC Pavements; PASCO Method (Continued).

				Expert			Inidividual			Team		
DISTRESS	UNITS	SEV.	SEC.	Mean	σa	σw	Mean	σa	σw	Mean	σa	σw
Trans. Cracking	Number	LOW	#1	4.8	3.1	3	6.9	2.6	6.3			
			#2	198.9	6.4	11.9	183.3	26.1	32.2			
			#3	14	7.6	5.4	22.4	10.6	4.8			
			#4	0.2	0	0.6				0.3	0.5	0.6
			#5	3.3	0.6	1.7				3.7	0.5	0.6
			#6	18.2	4	6.8				23.6	9.1	3.5
		MOD	#1	1.8	1	1	3.5	1.4	1.4			
			#2	10	6.6	7.9	23.3	24.3	35.6			
			#3	3.5	6.5	1.6	1.4	0	4.3			
			#4	0.8	0.4	0.5				0.2	0.3	0.3
			#5	3.1	0.7	1				2.6	0.7	0.7
			#6	6.2	1.9	3.2				2.1	1	1.5
		HIGH	#1	0.1	0	0.3	0.4	0	0.6			
			#2	0.1	0	0.3	0	0	0			
			#3	12.5	10.3	2.3	8.9	8.8	6.6			
			#4	0.2	0	0.6				0.2	0	0.5
			#5	0.8	0.8	0.9				0.7	1.1	0.6
			#6	5.2	1.3	2				5.3	1.3	1.1
		TOTAL	#1	6.58	3.93	2.29	10.78	0	5.88			
			#2	209	5.92	8.14	206.6	4.96	7.76			
			#3	30	6.23	5.43	32.78	1.58	3.37			
			#4	6.33	0	3.39				0.78	0	0.67
			#5	7.17	0	1.29				6.89	0.27	0.75
			#6	29.5	1.9	9				31	8.9	3.3
Trans. Cracking	Meters	LOW	#1	10	4.9	6	9.8	3.8	5.6			
			#2	748.6	52.1	44.7	655.7	104.3	135.2			
			#3	34.1	18.7	14.5	60.1	33.3	12.7			
			#4	12.4	0	13.7				0.6	0.7	1.4
			#5	9.7	1	5.4				11	0	2
			#6	43.3	11.6	20.6				56.9	17.6	5.6
		MOD	#1	6.1	3.5	3.4	12.1	5.4	4.8			
			#2	40	28.5	30.3	86.1	85.7	134.5			
			#3	13.8	19	9.7	4	0	13.1			
			#4	1.8	1	1.1				0.6	0.9	0.9
			#5	10.6	2	3.2				9.8	2.8	2.7
			#6	25.6	14.1	11.2				5	3.3	3.4
		HIGH	#1	0.3	0	1.2	0.8	0	1.7			
			#2	0.4	0	1.5	0	0	0			
			#3	42.4	35	8.6	28.2	28.1	20.5			
			#4	0.1	0	0.5				0.3	0	0.9
			#5	3.2	3.3	3.5				2.6	4.3	2.1
			#6	19.5	4.7	7.7				20.2	5.3	4.3
		TOTAL	#1	16.42	8.86	4.01	27.74	0	6.94			
			#2	789.1	40.67	29.64	741.8	55.77	34.01			
			#3	90.35	15.91	15.71	92.28	0	6.54			
			#4	14.32	0	13.45				1.56	1.09	1.4
			#5	23.49	1.66	3.1				23.41	0.77	0.91
			#6	88.3	19	26.2				82.1	18.8	3.6

Table 18. Mean, σa, and σw values for Experts, Individuals, and Teams for PCC Pavements; PASCO Method (Continued).

				Expert			Inidividual			Team		
DISTRESS	UNITS	SEV.	SEC.	Mean	σa	σw	Mean	σa	σw	Mean	σa	σw
Spalling of Long. Joints	Meters	LOW	#1	12.7	7.4	4.3	17.5	20.7	20.1			
			#2	64.5	0	30.8	2.2	3.2	4.4			
			#3	8.2	5.3	2.2	0.2	0.3	0.5			
			#4	0.2	0.3	0.2				0.1	0.1	0.1
			#5	16.7	8.1	4				18.4	11.5	4.2
			#6	8.8	4	4.9				0.7	0	1.2
		MOD	#1	0.1	0	0.1	2.9	2.1	6.6			
			#2	24.6	18.6	26.9	4.6	7.2	9.3			
			#3	0.4	0.6	0.5	0.4	0.2	0.8			
			#4	0	0	0.1				0.2	0.3	0.2
			#5	0.2	0.3	0.2				2.1	1.9	2.4
			#6	1	0.9	2.8				0	0	0
		HIGH	#1	0	0	0	0	0	0			
			#2	13.6	21.8	24.9	6	13.1	11.5			
			#3	0.4	0	0.3	0.1	0.3	0.1			
			#4	0.2	0	0.8				0	0	0
			#5	0	0	0.1				0.1	0	0.3
			#6	0.1	0	0.4				0	0	0.1
		TOTAL	#1	12.78	7.44	4.32	20.41	20.09	21.89			
			#2	102.7	25.06	20.18	12.78	24.22	23.09			
			#3	9.09	5.27	2.18	0.73	1.1	0.63			
			#4	0.46	0.1	0.77				0.29	0.43	0.21
			#5	16.93	8.22	3.82				20.62	12.65	3.1
			#6	9.88	4.38	3.84				0.77	0	1.26
Spalling of Trans. Joints	Number	LOW	#1	0.2	0	0.4	2	0	6			
			#3	0.1	0	0.3	0.3	0.5	0.3			
			#4	1.4	0	0.8				1	0.3	0.8
			#5	0.2	0.3	0.3				0.9	0	0.9
			#6	0	0	0				0	0	0
		MOD	#1	0.6	0.4	0.6	0.9	0	1.5			
			#3	0.4	0.5	0.3	0.1	0	0.2			
			#4	0.8	0.6	0.3				0.7	0	1
			#5	0.2	0	0.4				0.2	0	0.5
			#6	0.3	0.1	0.6				0	0	0
		HIGH	#1	0.3	0.5	0	0.3	0.4	0.3			
			#3	0.5	0.4	0.4	0.1	0.2	0.2			
			#4	0.8	0.7	0.9				0.7	0.3	0.8
			#5	0.1	0	0.3				0	0	0
			#6	1.8	0.7	0.5				0.3	0.6	0.5
		TOTAL	#1	1	0.72	0.82	3.19	0	7.58			
			#3	1	0.14	0.41	0.44	0.55	0.47			
			#4	3.08	0	1.12				2.33	0	1.41
			#5	0.42	0	0.71				1.11	0	1.15
			#6	2.2	0.1	0.7				0.3	0.6	0.5

Table 18. Mean, σa, and σw values for Experts, Individuals, and Teams for PCC Pavements; PASCO Method (Continued).

				Expert			Inidividual			Team		
DISTRESS	UNITS	SEV.	SEC.	Mean	σa	σw	Mean	σa	σw	Mean	σa	σw
Spalling of Trans. Joints	Meters	LOW	#1	0.1	0	0.3	2.6	0	4.8			
			#3	0.2	0	0.5	0.2	0.2	0.2			
			#4	1.7	0	0.9				1.6	0	0.8
			#5	0.3	0.3	0.4				0.8	0	0.6
			#6	0	0	0.1				0.1	0	0.2
		MOD	#1	0.8	0.5	1	1.1	0	1.7			
			#3	0.2	0.2	0.1	0.1	0	0.2			
			#4	1	0.6	0.9				1.2	0	1
			#5	0.3	0.1	0.3				0.4	0.2	0.3
			#6	0.2	0.1	0.4				0.2	0.3	0
		HIGH	#1	0.4	0.8	0.1	0.5	0.6	0.4			
			#3	0.3	0.2	0.1	0.1	0.1	0.2			
			#4	2.4	1.5	2.2				2.1	1.1	1
			#5	0.2	0.1	0.3				0	0.1	0.1
			#6	3.1	1.6	2.1				2	2	1.3
		TOTAL	#1	1.36	1.15	1.06	4.09	0	6.14			
			#3	0.63	0	0.55	0.31	0.21	0.28			
			#4	5.08	0	2.22				4.84	1.47	0.93
			#5	0.73	0.16	0.71				1.27	0.25	0.69
			#6	3.3	1.6	2.1				2.3	2.5	1.4

Notes: σa = among rater standard deviation
σw = within rater standard deviation

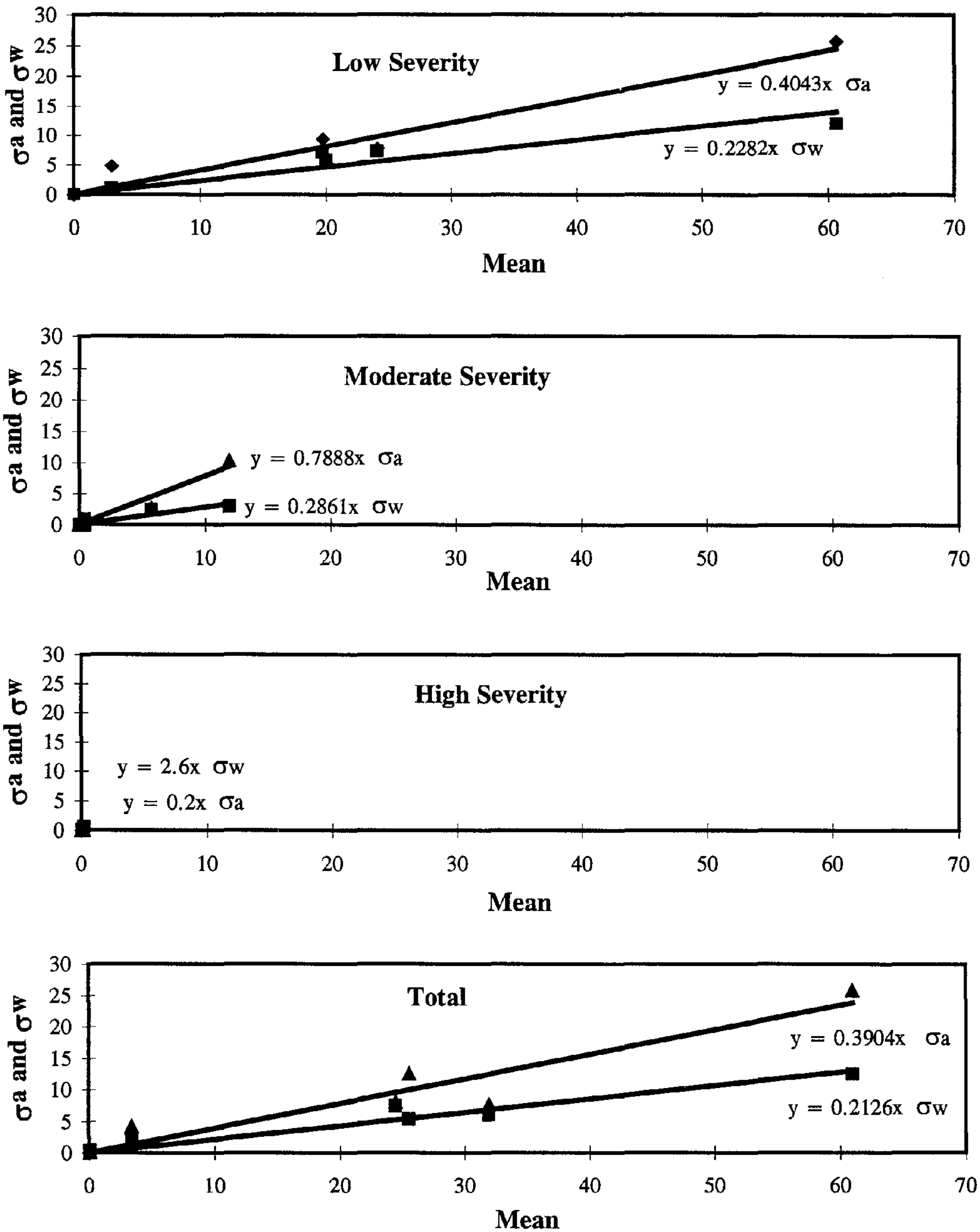


Figure 22. Fatigue Cracking (Sq. Meters) - AC Pavements, Experts, PASCO Method: σa and σw Vs. Mean.

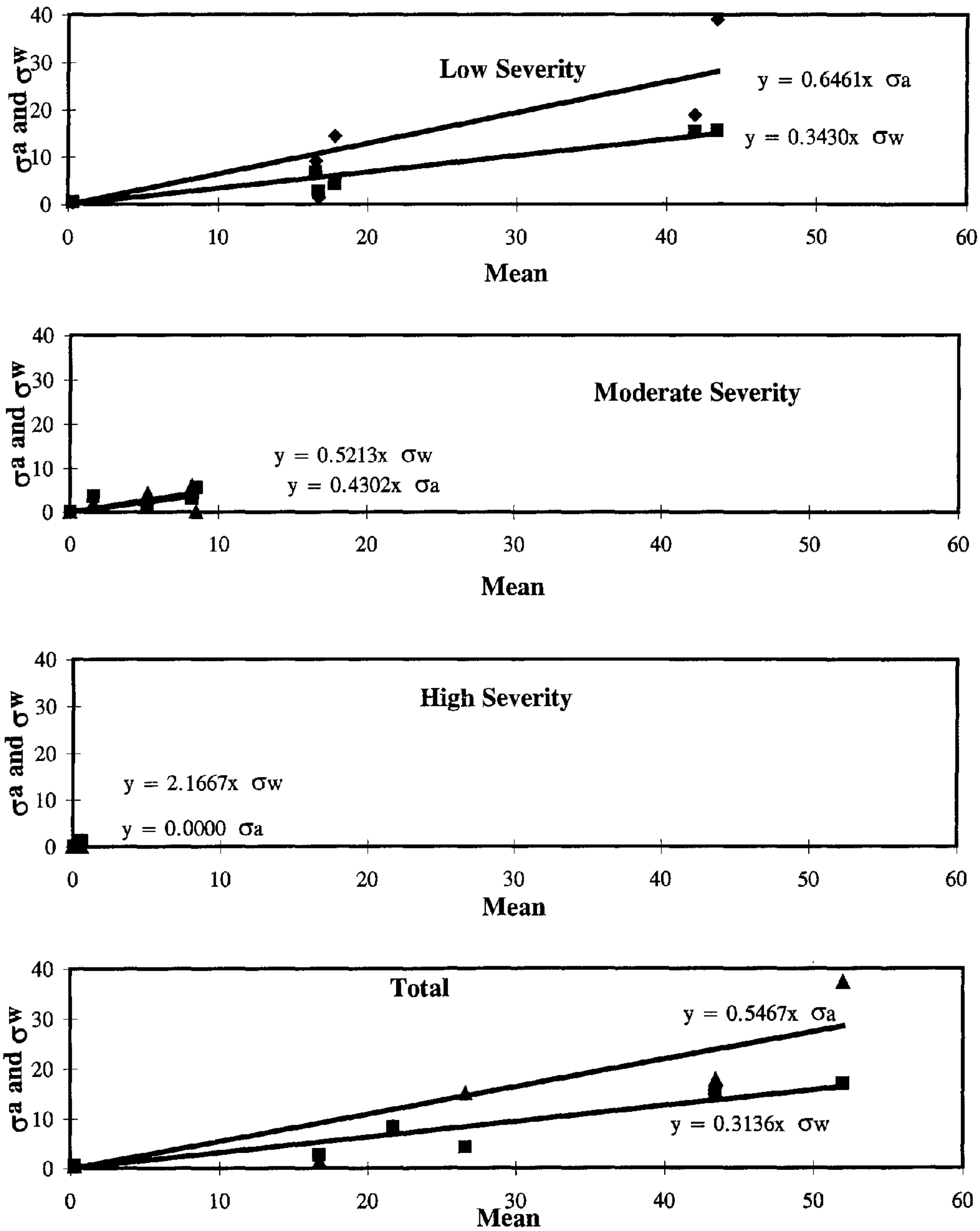


Figure 23. Longitudinal Cracking WP (Meters) - AC Pavements, Experts, PASCO Method: σa and σw Vs. Mean.

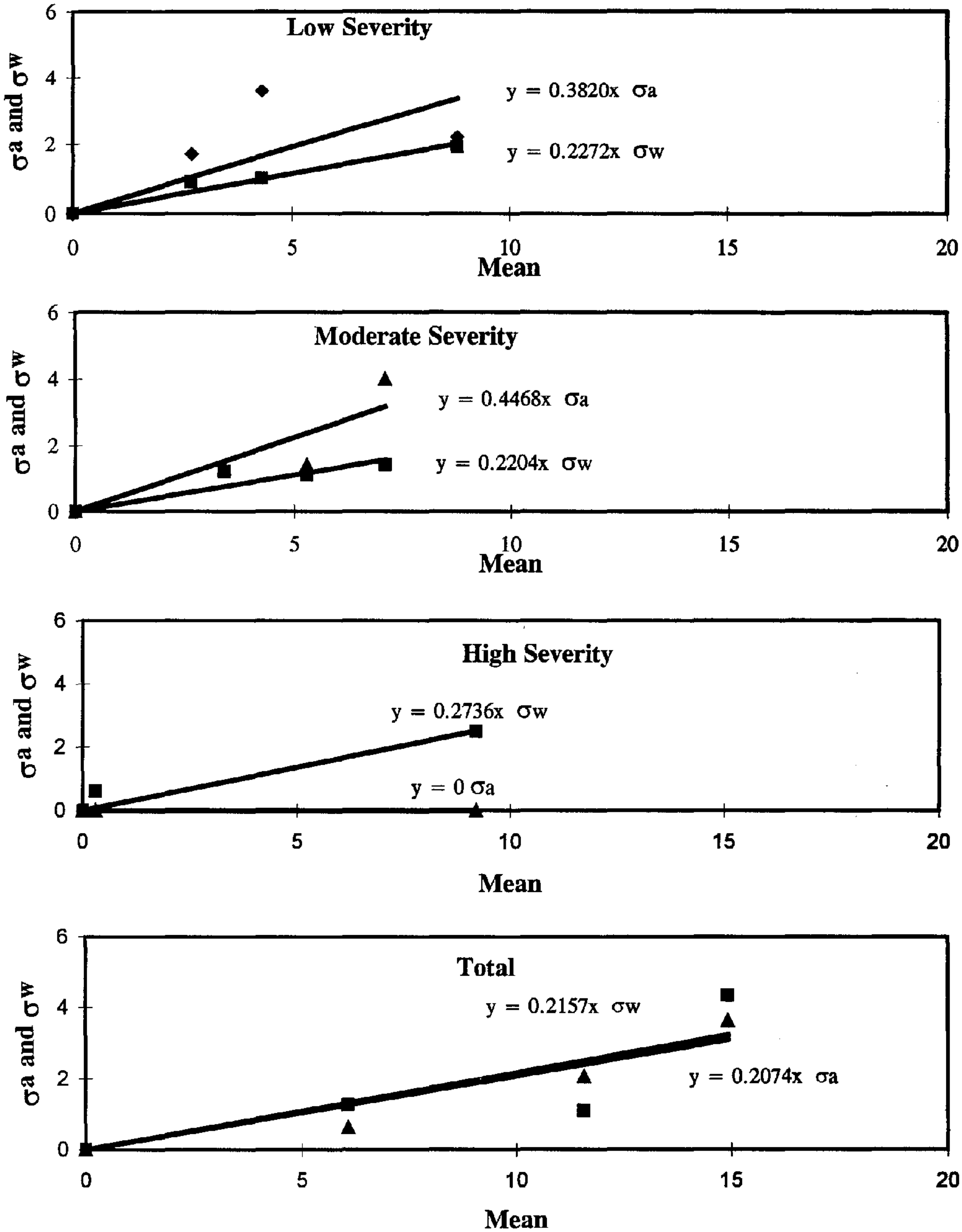


Figure 24. Corner Breaks (No.) - PCC Pavements, Experts, PASCO Method: σa and σw Vs. Mean.

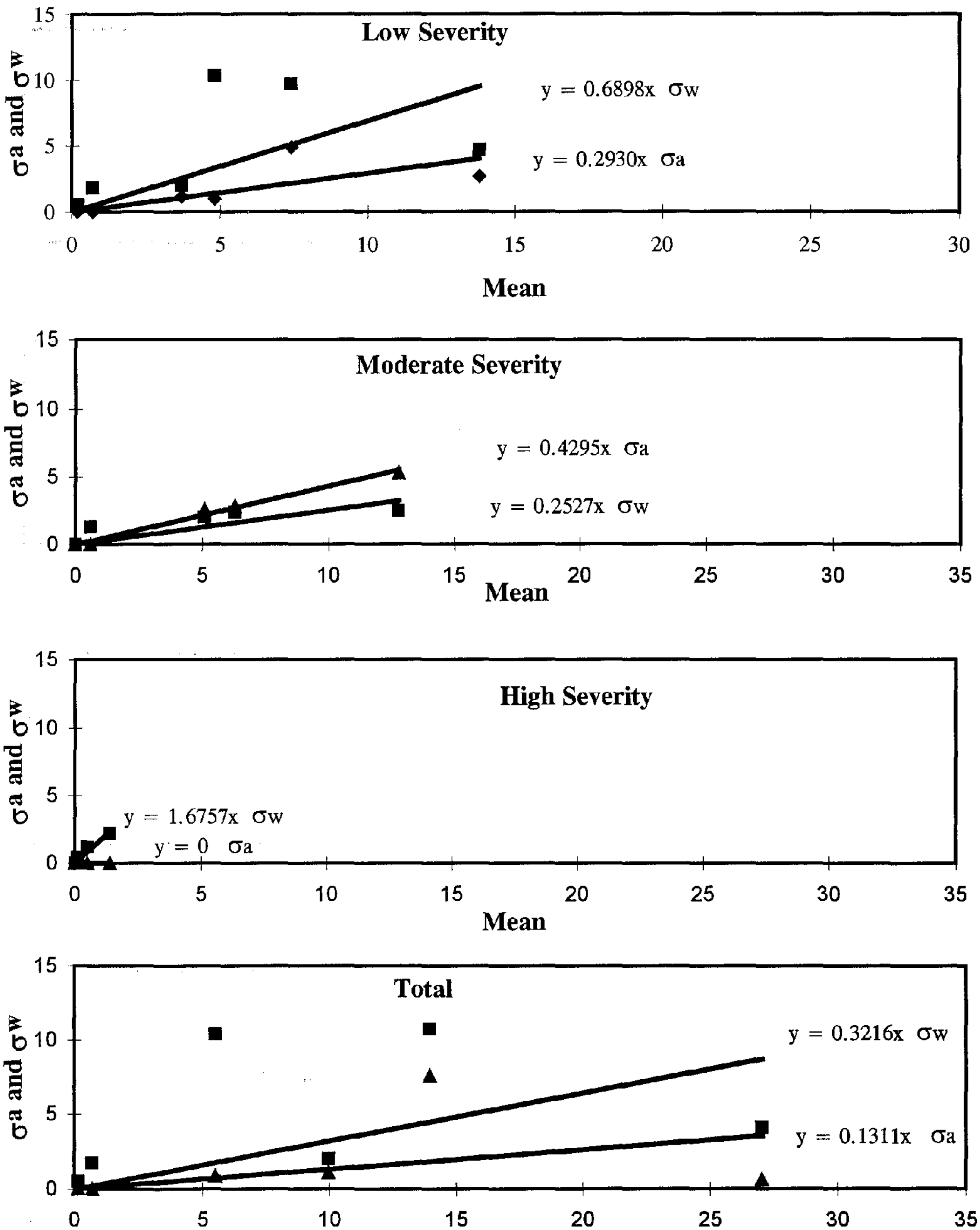


Figure 25. Longitudinal Cracking (Meters) - PCC Pavements, Experts PASCO Method: σa and σw Vs. Mean.

Table 19. CVa and CVw for Experts, Individuals, and Teams for Asphalt Pavements.

DISTRESS	UNITS	SEV.	EXPERT CVa	EXPERT CVw	INDIVIDUAL CVa	INDIVIDUAL CVw	TEAM CVa	TEAM CVw
Fatigue Cracking	Sq. Meters	Low	40.43	22.82	90.16	56.35	93.00	27.65
		Mod.	78.88	28.61	33.17	133.20	96.52	21.63
		High	20.00	260.00	100.00	175.00	173.91	17.39
		Total	39.04	21.26	72.00	60.80	29.05	16.31
Long. Cracking WP	Meters	Low	64.61	34.30	50.10	43.75	28.93	12.02
		Mod.	43.02	52.13	71.11	68.16	13.10	27.67
		High	0.00	216.67	66.12	169.02	100.00	220.00
		Total	54.67	31.36	37.49	28.78	28.25	13.07
Long. Cracking NWP	Meters	Low	8.22	8.53	40.55	30.78	43.17	33.91
		Mod.	40.87	27.74	47.10	54.75	46.71	33.98
		High	45.74	88.25	115.95	84.05	102.14	118.27
		Total	6.93	9.12	38.34	28.69	31.81	28.38
Trans. Cracking	Number	Low	25.42	14.81	19.23	17.47	19.68	16.64
		Mod.	32.65	30.67	29.98	42.99	19.48	23.87
		High	31.21	60.96	57.78	78.81	155.88	38.24
		Total	16.34	8.46	10.04	11.14	11.69	14.42
Trans. Cracking	Meters	Low	24.39	20.21	17.43	18.24	19.36	16.40
		Mod.	19.61	22.42	25.97	49.17	12.89	29.69
		High	30.94	48.39	38.19	73.83	161.90	63.49
		Total	17.04	11.89	11.47	10.87	9.84	11.98
Summary Statistics								
All distress levels		Max.	78.88	260.00	115.95	175.00	173.91	220.00
		Min.	0.00	8.46	10.04	10.87	9.84	11.98
		Avg.	32.00	50.93	48.61	61.79	59.87	39.25
All distress levels except high severity level		Max.	78.88	52.13	90.16	133.20	96.52	33.98
		Min.	6.93	8.46	10.04	10.87	9.84	11.98
		Avg.	34.14	22.96	39.61	43.68	33.56	21.84
Total distress only		Max.	54.67	31.36	72.00	60.80	31.81	28.38
		Min.	6.93	8.46	10.04	10.87	9.84	11.98
		Avg.	26.80	16.42	33.87	28.06	22.13	16.84

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

Table 20. CVa and CVw for Experts, Individuals, and Teams for Concrete Pavements.

DISTRESS	UNITS	SEV.	EXPERT CVa	EXPERT CVw	INDIVIDUAL CVa	INDIVIDUAL CVw	TEAM CVa	TEAM CVw
Corner Breaks	Sq. Meters	Low	38.20	22.72	77.51	61.66	8.63	14.41
		Mod.	44.68	22.04	58.49	31.54	13.42	21.02
		High	0.00	27.36	275.00	125.00	92.08	121.49
		Total	20.74	21.57	13.97	11.95	5.87	7.16
Long. Cracking	Meters	Low	29.30	68.98	58.83	183.50	24.85	13.90
		Mod.	42.95	25.27	50.40	283.20	0.00	67.57
		High	0.00	167.57	2.77	43.69	0.00	333.33
		Total	13.11	32.16	55.24	166.90	21.57	4.01
Trans. Cracking	Number	Low	3.66	6.45	14.76	17.73	37.98	14.89
		Mod.	63.84	66.84	102.51	150.85	35.50	44.87
		High	74.00	21.77	98.68	74.31	26.76	22.19
		Total	3.30	4.77	2.45	4.05	27.53	10.70
Trans. Cracking	Meters	Low	7.13	6.23	16.24	20.63	29.83	10.17
		Mod.	69.56	65.00	98.26	154.32	36.64	36.24
		High	72.48	24.08	99.57	72.81	28.50	22.32
		Total	5.53	4.29	7.39	4.65	21.44	4.38
Spalling of Long. Joints	Meters	Low	6.49	45.71	118.71	116.20	62.41	23.04
		Mod.	75.66	109.63	132.16	209.35	91.01	114.16
		High	160.11	183.05	218.36	191.64	0.00	300.00
		Total	25.94	20.13	124.11	127.89	61.28	15.25
Spalling of Trans. Joints	Number	Low	2.93	62.93	3.67	295.60	16.57	88.95
		Mod.	73.64	75.97	0.00	167.07	0.00	150.94
		High	51.30	43.74	140.00	110.00	67.24	122.41
		Total	6.54	39.50	2.33	235.18	2.67	69.79
Spalling of Trans. Joints	Meters	Low	2.97	58.75	0.59	184.12	0.00	55.45
		Mod.	60.22	104.42	0.00	154.92	8.54	80.49
		High	57.22	76.12	119.23	84.62	75.03	55.89
		Total	17.63	51.97	0.39	149.78	43.47	28.35
Summary Statistics								
All distress levels		Max.	160.11	183.05	275.00	295.60	92.08	333.33
		Min.	0.00	4.29	0.00	4.05	0.00	4.01
		Avg.	36.75	52.11	67.56	122.61	29.96	66.19
All distress levels except high severity level		Max.	75.66	109.63	132.16	295.60	91.01	150.94
		Min.	2.93	4.29	0.00	4.05	0.00	4.01
		Avg.	29.24	43.59	44.67	130.05	26.15	41.70
Total distress only		Max.	25.94	51.97	124.11	235.18	61.28	69.79
		Min.	3.30	4.29	0.39	4.05	2.67	4.01
		Avg.	13.26	24.91	29.41	100.06	26.26	19.95

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

type-severity level combinations where the magnitude of distress is small, as is the case for most high severity distresses. As shown in tables 19 and 20, 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 earlier findings that indicated a problem in distinguishing distress at different severity levels.

- For AC pavements, the between rater variability (CVa) was generally greater than the within rater variability (CVw). When only total distress quantities are considered, the average CVa values for experts, individual raters, and teams was 26.8, 33.9, and 22.1 percent, and the average CVw values for experts, individual raters, and teams was 16.4, 28.1 and 16.8 percent, 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 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 72.0 and 60.8 percent for total fatigue cracking quantities, respectively.

- For PCC pavements, the between rater variability (CVa) was generally lower than the within rater variability (CVw). When only total distress quantities are considered, the average CVa value for experts, individual raters, and teams was 13.3, 29.4 and 26.3 percent, and the average CVw value for experts, individual raters, and teams were 24.9, 100.1 and 20.0 percent, 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 compared with those from the experts and teams, both of which had comparable results. Also, unlike the experts and teams, individual raters appear to have difficulty in consistently identifying joint-related distresses. The within rater variation for individual raters was CVw = 235.2 percent 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 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.

3.5 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 (6 AC and 6 PCC) workshop pavement test sections. Although three sets of PADIAS v4.x distress data were generated by each expert, individual rater, and two-person team involved in this effort, such repeat measurements are not normal practice within the LTPP program. Consequently, only those data from the first repeat 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 the workshop instructors immediately prior to the workshop. 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 appendix B of this report. 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. Example plots are given in figures 26 through 29. Figures 26 and 27 show the global trends for fatigue and longitudinal cracking in the wheel path in AC pavements, while figures 28 and 29 show similar plots for corner breaks and longitudinal cracking in PCC pavements. (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:

- Contrary to the trend observed for manual distress data, total distress quantities did not show reduced variability compared with those for the individual severity levels. However, like the manual data, there does not appear to be a significant positive or negative bias in the data.

- There appear to be compensatory differences between the various groups for closely related distress types such as fatigue cracking and longitudinal cracking in the wheel path for AC pavements.

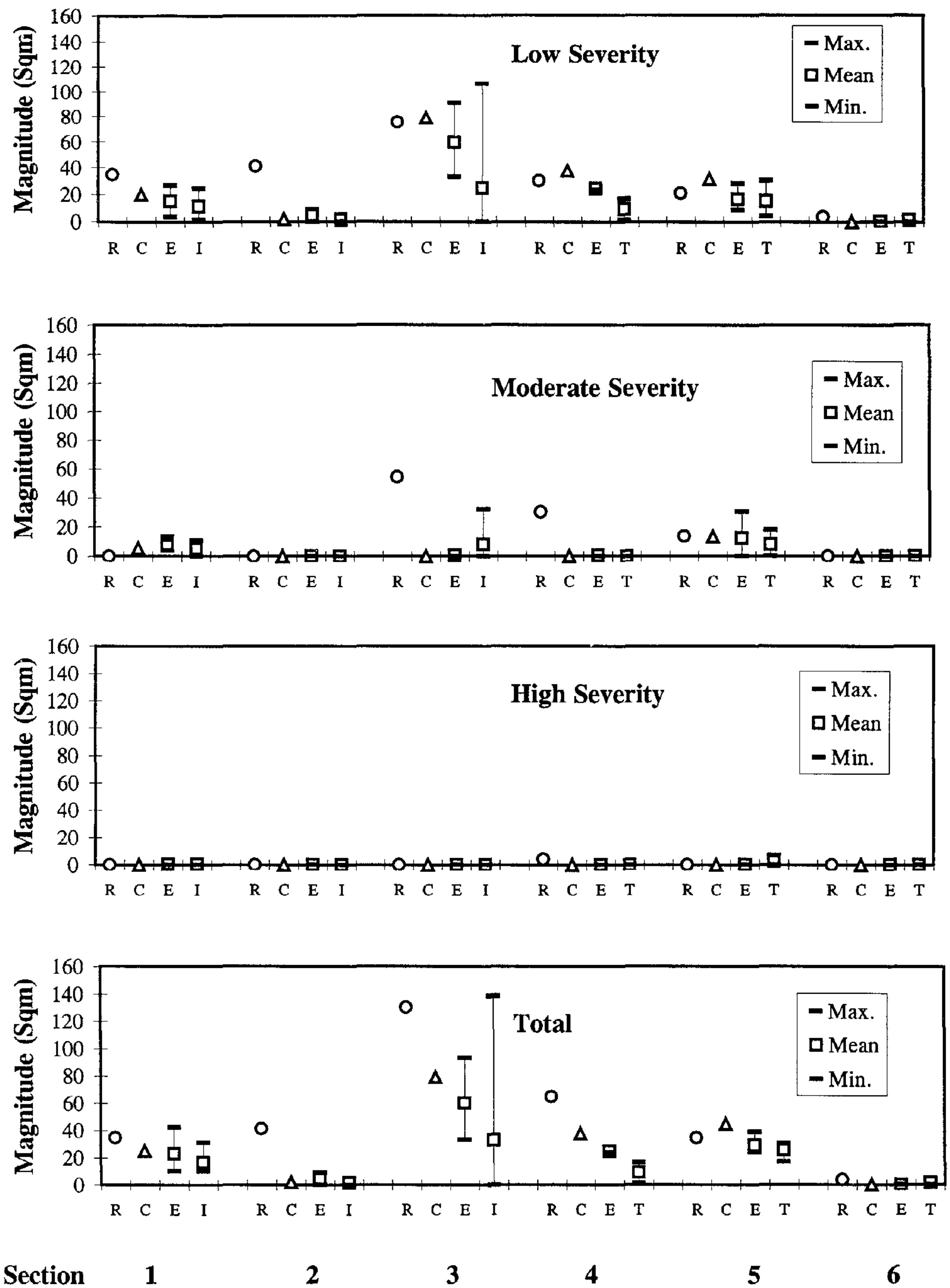


Figure 26. Fatigue Cracking (Sq. Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals, & Teams.

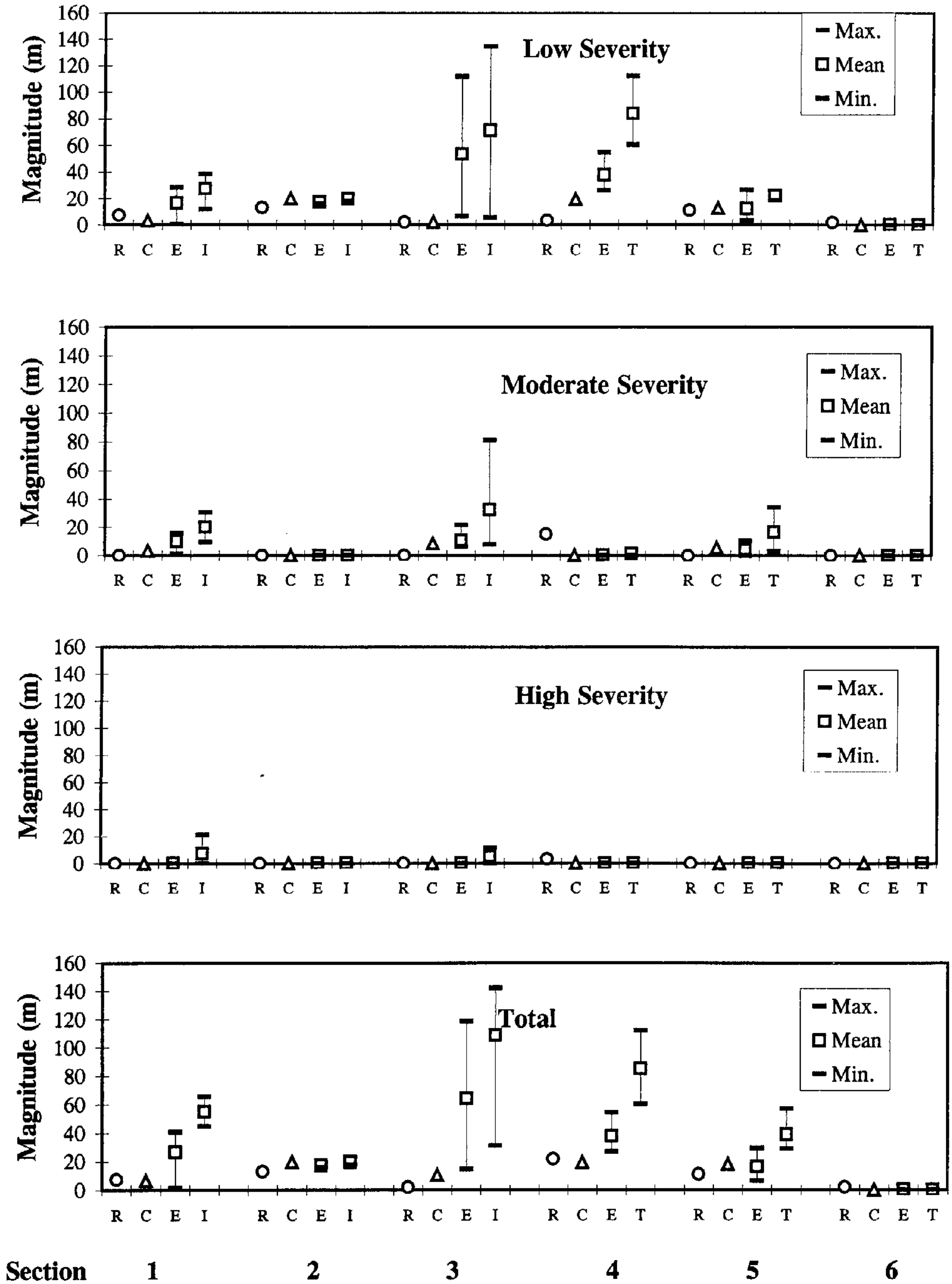


Figure 27. Longitudinal Cracking WP (Meters) - AC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals, & Teams.

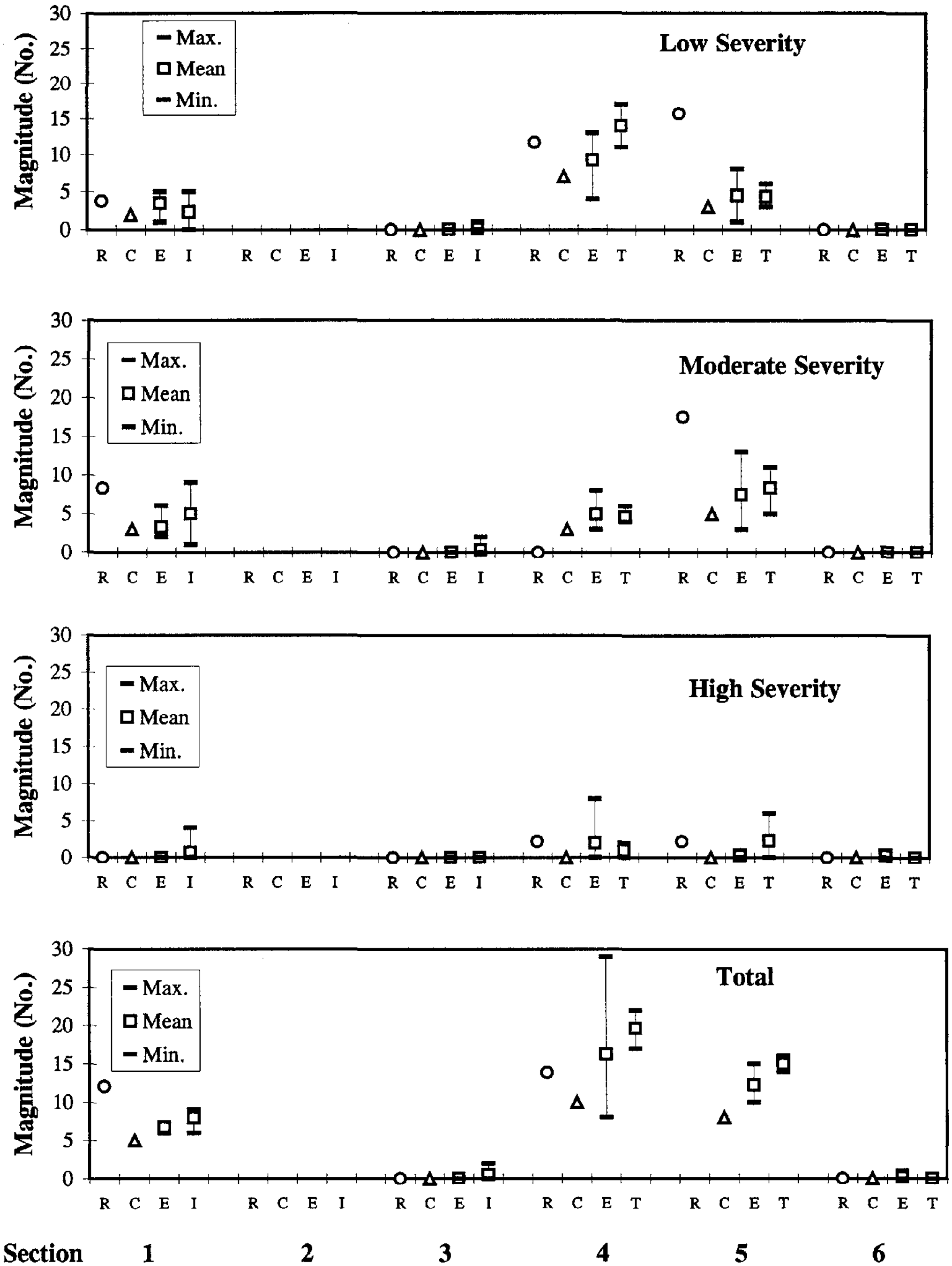


Figure 28. Corner Breaks (No.) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals, & Teams.

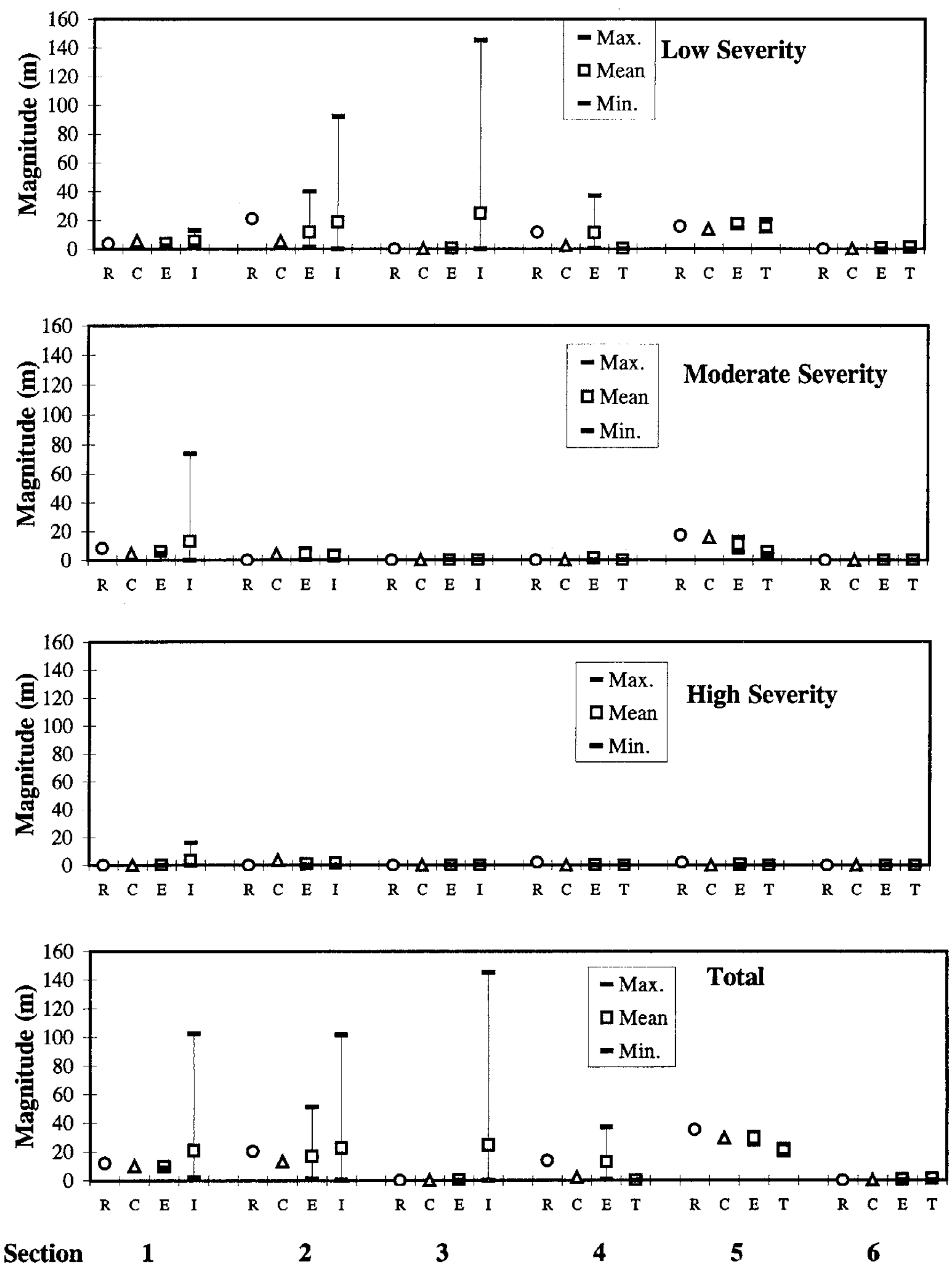


Figure 29. Longitudinal Cracking (Meters) - PCC Pavements, PASCO/PADIAS: Reference, Consensus, Experts, Individuals, & Teams.

- 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.

3.6 Bias and Precision

One of the main objectives of the study presented in this chapter was to quantify the bias and precision associated with LTPP distress data derived from film using the PASCO/PADIAS v4.x system. Toward that end, distress data collected by PASCO and interpreted using the PADIAS v4.x system were analyzed quantitatively. Specifically, an analysis of the coefficient of variation (CV) and root mean square error (RSME) associated with these data was first undertaken. Indicators of bias and precision were subsequently estimated for these data. Both of these analyses and associated results are presented next.

CV and RMSE Evaluation

As indicated earlier in this report, the 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 plotting the within and between rater standard deviation versus means for each distress type-severity level combination and fitting the best line through these data (y-intercept was forced through 0). See appendix B for a complete set of plots for both AC and PCC pavement distresses. The slope of this best straight-line fit (in percentage terms) is a measure of the ratio between standard deviation and mean over varying ranges of CV, assuming a linear increasing relationship between the standard deviation and mean. Example CV plots for different distress types and severity levels are shown in figures 30 through 35. Figures 30 through 32 show the AC pavement fatigue cracking plots for the expert, individual, and two-person teams, respectively. Similarly, figures 33 through 35 show the PCC pavement corner break plots for the expert, individual, and two-person teams, respectively.

Also included in the referenced figures are regression lines of RMSE 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

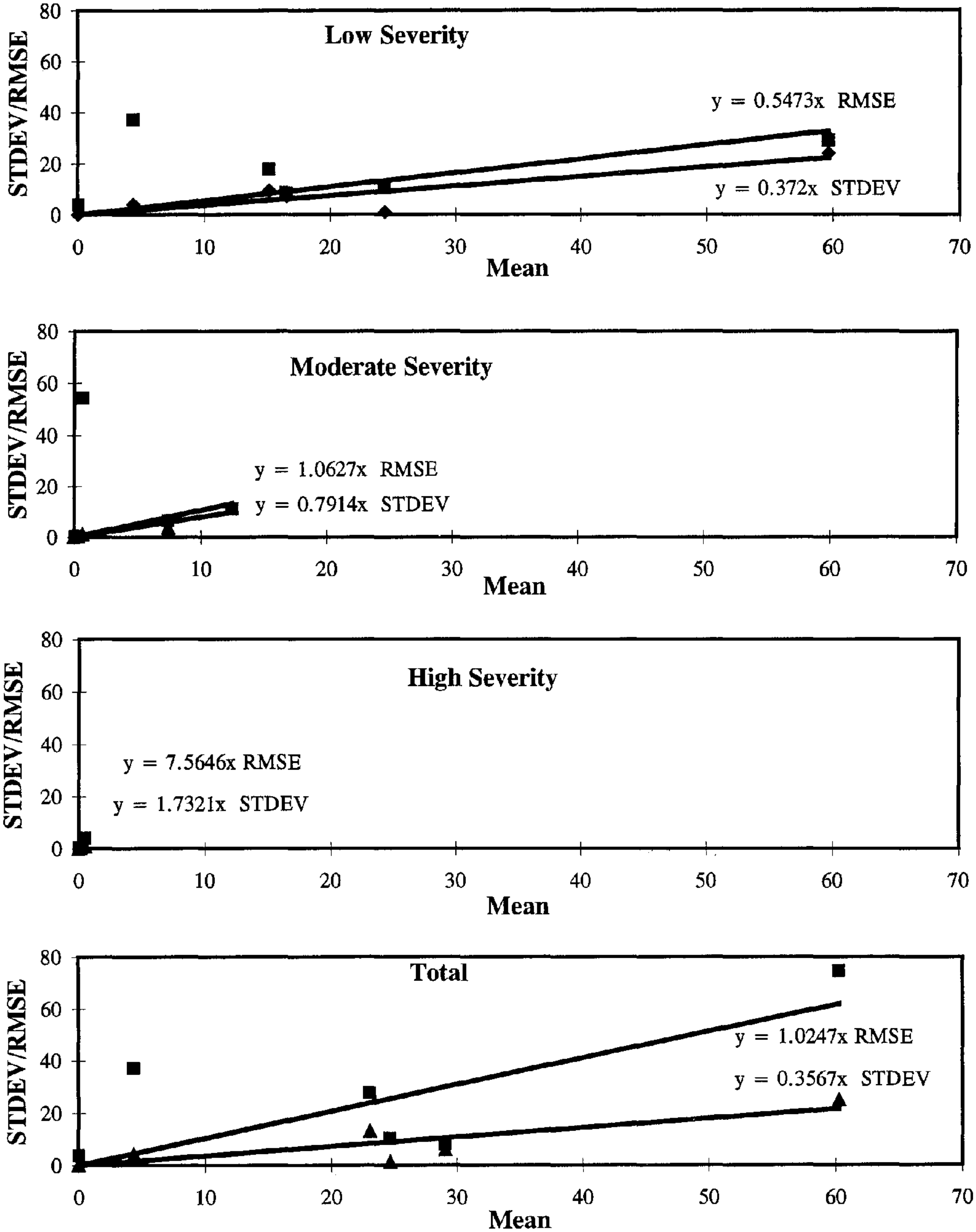


Figure 30. Fatigue Cracking (Sq. Meters) - AC Pavements, Expert Raters, PASCO Method: Standard Deviation/RMSE Vs. Mean.

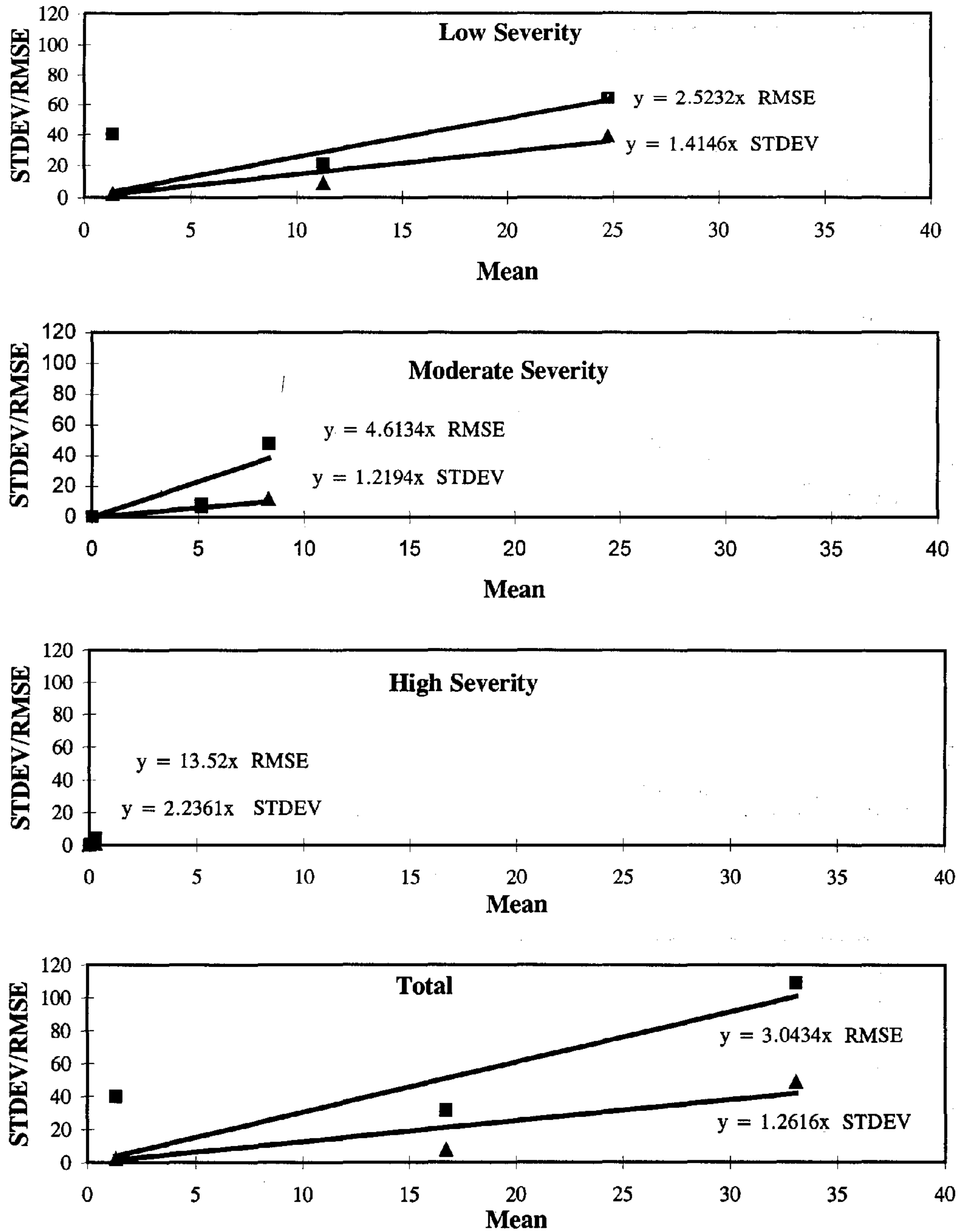


Figure 31. Fatigue Cracking (Sq. Meters) - AC Pavements, Individual Raters, PASCO Method: Standard Deviation/RMSE Vs. Mean.

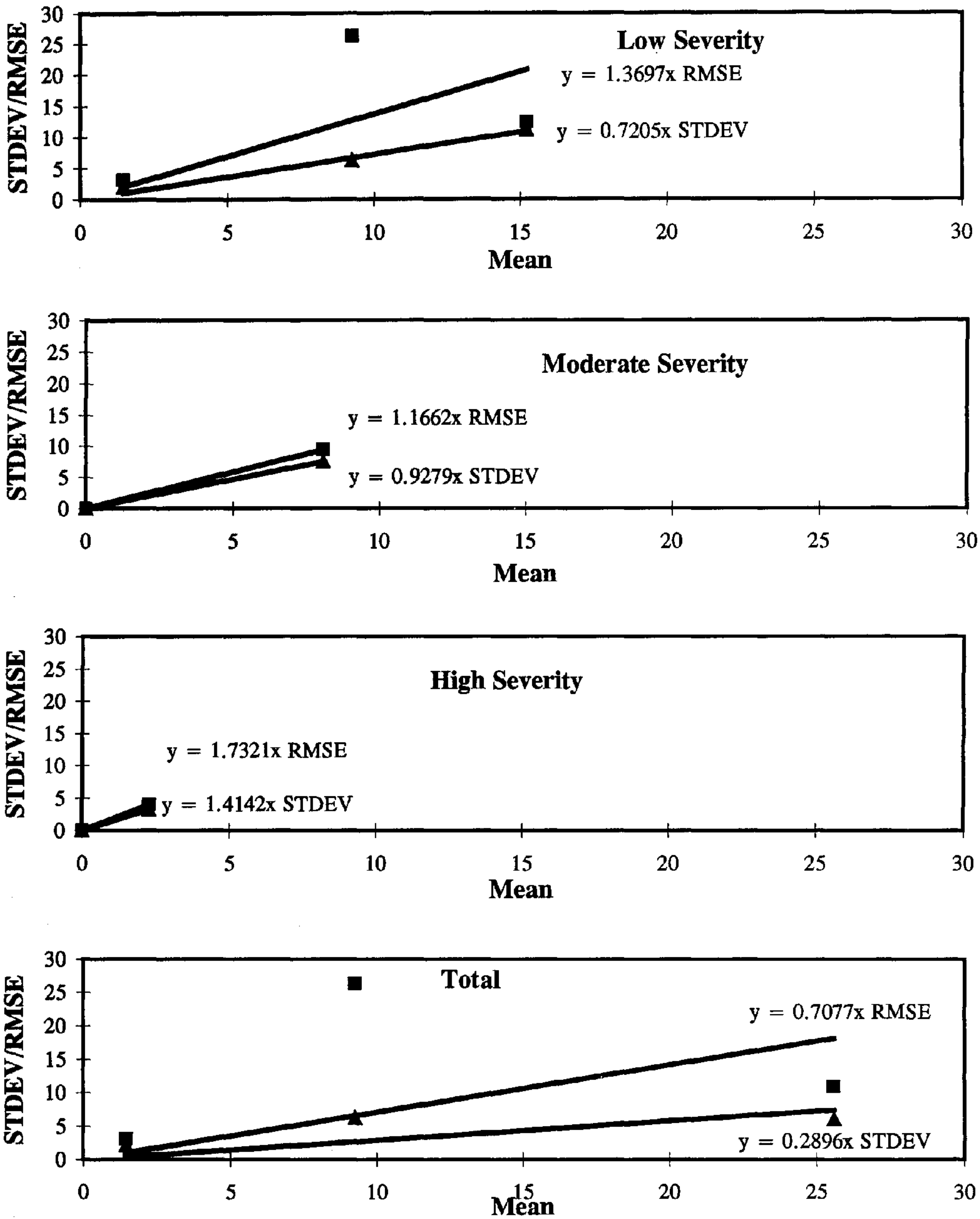


Figure 32. Fatigue Cracking (Sq. Meters) - AC Pavements, Team Surveys, Standard Deviation/RMSE Vs. Mean.

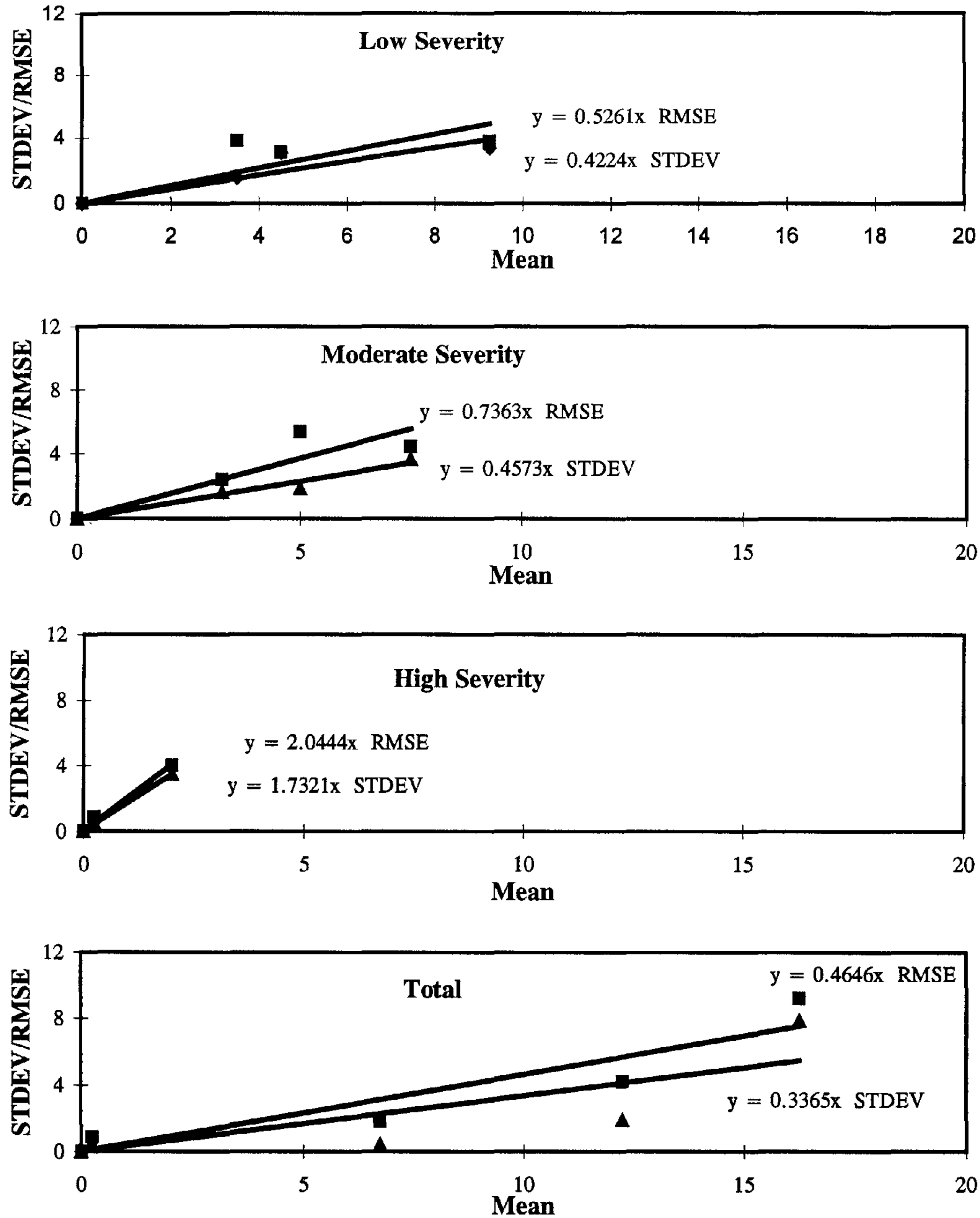


Figure 33. Corner Breaks (No.) - PCC Pavements, Expert Raters, PASCO Method: Standard Deviation/RMSE Vs. Mean.

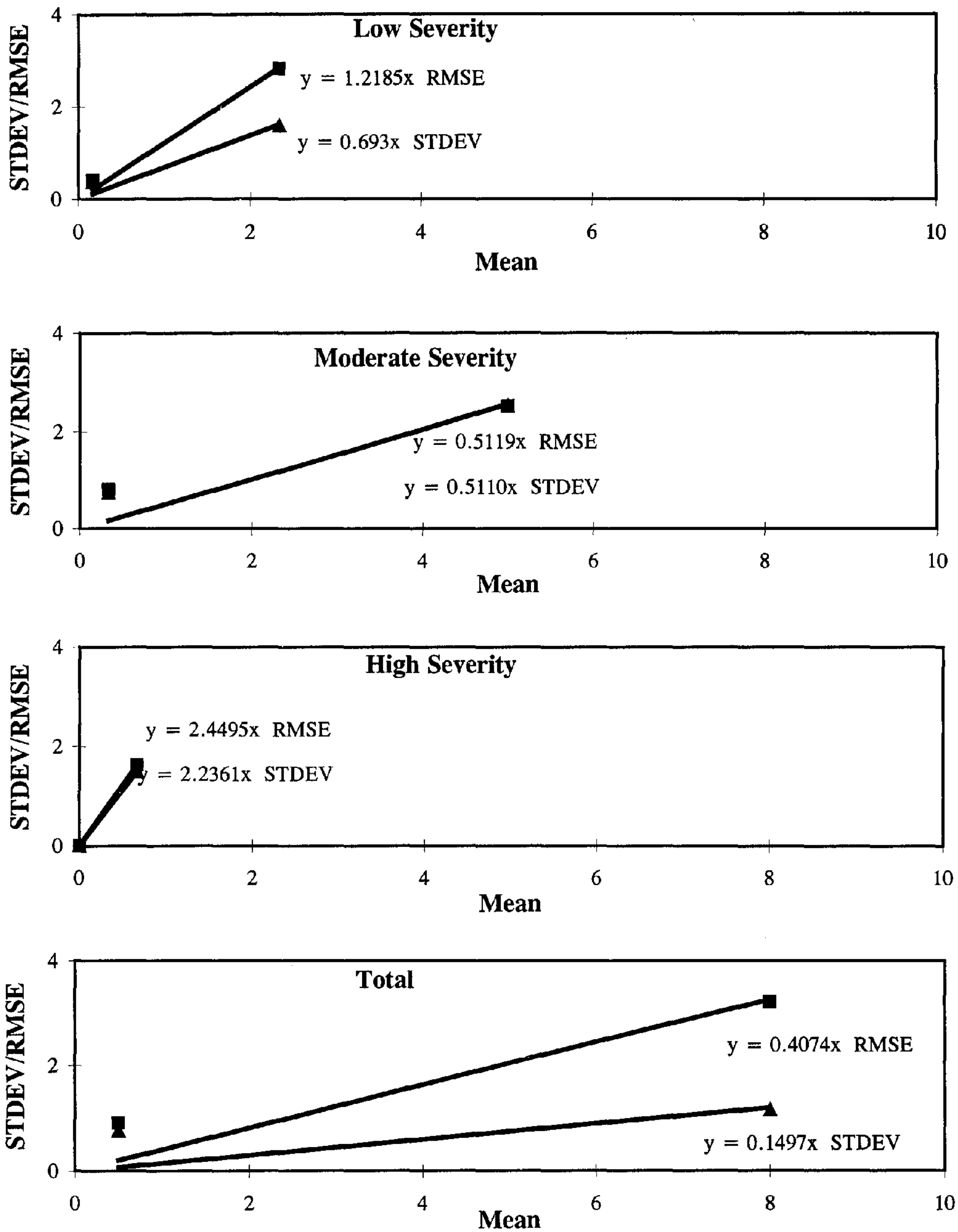


Figure 34. Corner Breaks (No.) - PCC Pavements, Individual Raters, PASCO Method: Standard Deviation/RMSE Vs. Mean.

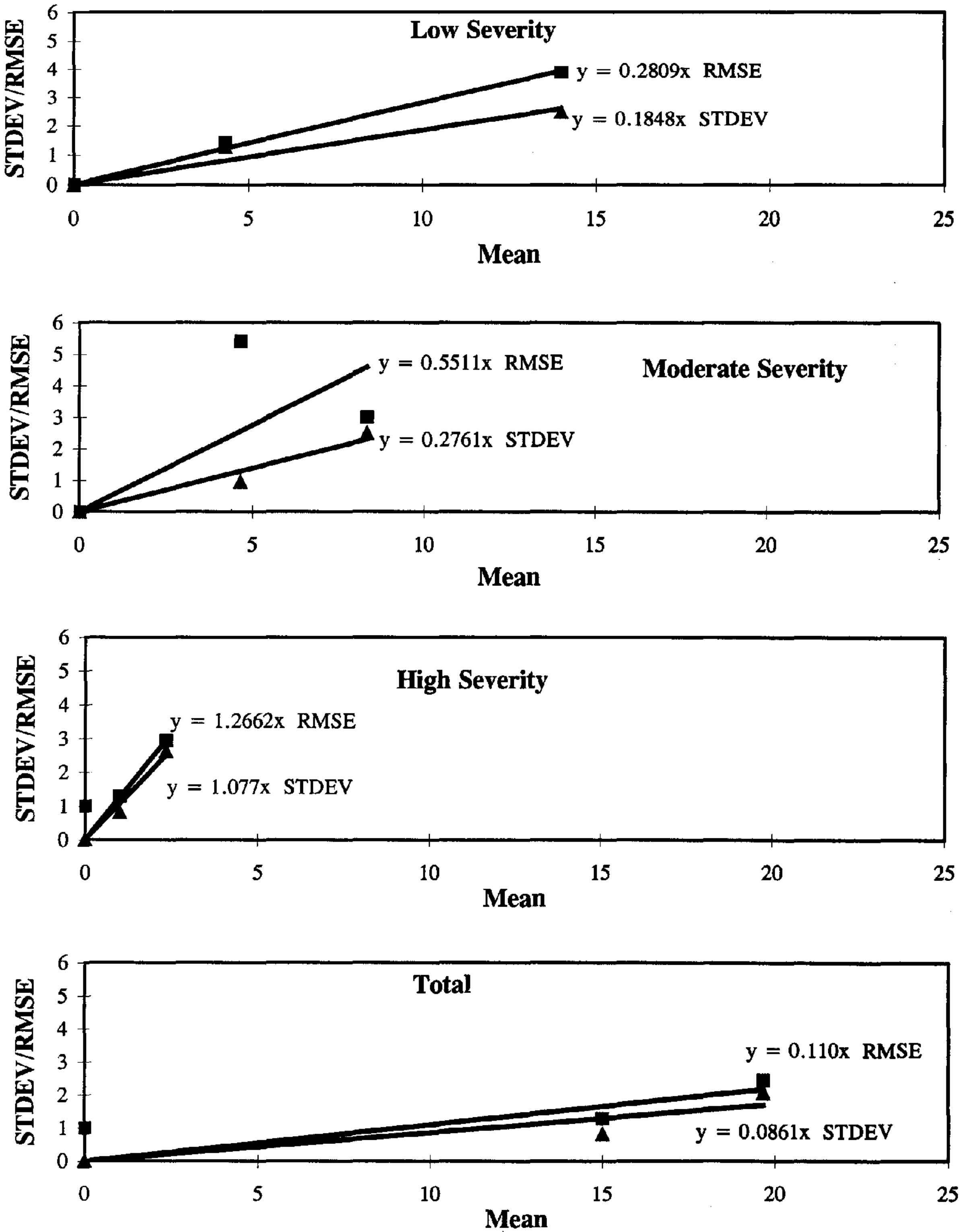


Figure 35. Corner Breaks (No.) - PCC Pavements, Team Surveys, PASCO Method: Standard Deviation/RMSE Vs. Mean.

reference values, i.e., distribution of group values relative to the reference. (Note: unlike the analysis of manual distress data, an outlier analysis was not conducted for PASCO distress data because of the limited amount of data points - 6 sections for expert surveys and 3 sections for individual and team raters.)

On the basis of those figures, the following observations were developed:

- CV values associated with the PASCO distress data range from less than 10 percent to more than 200 percent; however, the larger CV's 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 percent for experts, teams, and individual raters, respectively. Those average values for PCC pavement distress data are 31 percent, 30 percent, and 90 percent for experts, teams, and individual raters, respectively. Except for the individual rater CV values, these averages appear reasonable.

- Unlike the manual distress data, 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 the 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 the various pavement test sections. A partial summary of results is presented in tables 21 and 22 for AC and PCC pavements, respectively. The following terminology was used in these two tables:

- **Pooled Values** - Average of reference and group mean values 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.

- **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 RMSE** - Derived from regression lines introduced in previous section (not directly calculated from the distress data).

Table 21. Indicators of Bias and Precision for AC Pavement Distresses - Pasco Method.

Distress Type	Unit	Distress Sev.	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	34.48	20.01	7.44	10.95	37.20	-14.48	19.80	8.65	6.23	11.84	72.05	-11.15	49.17	12.44	17.60	31.39	141.46	-36.73
		Moderate	13.55	3.47	2.75	3.69	79.14	-10.08	4.60	2.70	2.50	3.14	92.79	-1.90	22.50	4.49	5.48	20.72	121.94	-18.01
		High	0.72	0.10	0.17	0.75	173.21	-0.62	0.00	0.76	1.07	1.32	141.42	0.76	1.43	0.10	0.22	1.35	223.61	-1.33
		Total	48.75	23.58	8.41	24.16	35.67	-25.17	24.40	12.10	3.50	8.56	28.96	-12.30	73.10	17.03	21.49	51.84	126.16	-56.07
Longitudinal Cracking WP	Meters	Low	6.55	22.93	14.73	24.60	64.24	16.38	6.83	35.42	8.74	32.52	24.69	28.59	6.27	39.37	22.91	42.91	58.21	33.10
		Moderate	2.55	4.21	2.60	4.28	61.78	1.56	0.00	5.99	4.71	7.62	78.70	5.99	5.10	17.61	13.44	19.40	76.31	12.51
		High	0.55	0.06	0.11	0.63	173.21	-0.49	0.00	0.08	0.11	0.13	141.42	0.08	1.10	4.21	4.71	5.64	111.81	3.11
		Total	9.65	27.20	14.83	26.72	54.53	17.55	6.83	41.49	10.86	38.09	26.19	34.66	12.47	61.18	19.83	57.91	32.42	48.72
Longitudinal Cracking NWP	Meters	Low	81.85	103.41	9.81	26.92	9.48	21.56	78.07	78.54	35.75	48.38	45.52	0.48	85.63	87.76	51.11	57.34	58.24	2.12
		Moderate	39.55	28.92	11.46	23.06	39.63	-10.63	60.80	31.52	11.95	39.59	37.90	-29.28	18.30	18.16	8.63	16.20	47.52	-0.14
		High	22.40	2.37	1.59	13.19	66.96	-20.03	20.00	4.36	4.49	15.44	103.00	-15.64	24.80	4.47	4.14	11.91	92.66	-20.33
		Total	143.80	134.70	19.44	34.54	14.43	-9.10	158.87	114.42	31.12	63.23	27.19	-44.44	128.73	110.38	57.24	83.08	51.86	-18.35
Transverse Cracking	Number	Low	33.00	39.54	9.47	13.81	23.94	6.54	21.33	44.89	12.59	30.56	28.05	23.56	44.67	42.50	8.59	18.17	20.21	-2.17
		Moderate	9.33	14.42	5.64	8.39	39.09	5.08	3.33	6.89	2.96	4.99	42.98	3.56	15.33	17.22	3.91	5.90	22.71	1.89
		High	7.50	3.29	1.39	4.74	42.11	-4.21	2.33	1.11	1.34	1.61	120.83	-1.22	12.67	10.78	6.88	7.15	63.85	-1.89
		Total	49.83	57.25	9.19	16.93	16.05	7.42	27.00	52.89	9.15	32.07	17.29	25.89	72.67	70.50	8.59	21.82	12.19	-2.17
Transverse Cracking	Meters	Low	37.37	55.27	18.71	24.31	33.85	17.90	30.67	47.56	12.20	23.96	25.66	16.89	49.17	12.44	17.60	31.39	141.46	-36.73
		Moderate	17.48	33.57	6.68	20.77	19.90	16.09	7.03	9.78	4.97	6.26	50.83	2.74	22.50	4.49	5.48	20.72	121.94	-18.01
		High	18.75	9.46	3.71	11.49	39.22	-9.29	3.17	2.33	3.07	3.13	131.45	-0.83	1.43	0.10	0.22	1.35	223.61	-1.33
		Total	73.60	98.30	24.42	34.84	24.85	24.70	40.87	59.67	7.24	24.87	12.14	18.80	73.10	17.03	21.49	51.84	126.16	-56.07

Table 22. Indicators of Bias and Precision for PCC Pavement Distresses - Pasco Method.

Distress		Distress	EXPERT POOLED VALUES						TEAM POOLED VALUES						INDIVIDUAL POOLED VALUES					
Type	Unit	Sev.	REF.	Mean	STDEV	RMSE	COV	Bias	REF.	Mean	STDEV	RMSE	COV	Bias	REF.	Mean	STDEV	RMSE	COV	Bias
Corner Breaks	Number	Low	3.20	3.45	1.46	1.82	42.24	0.25	5.33	6.11	1.13	1.72	18.48	0.78	0.00	1.25	0.87	1.52	69.30	1.25
		Moderate	5.00	3.15	1.44	2.32	45.73	-1.85	6.67	4.33	1.20	2.39	27.61	-2.33	2.50	2.67	1.36	1.37	51.10	0.17
		High	0.40	0.50	0.87	1.02	173.21	0.10	0.67	1.11	1.20	1.41	107.70	0.44	0.00	0.33	0.75	0.82	223.61	0.33
		Total	8.60	7.10	2.39	3.30	33.65	-1.50	12.67	11.56	0.99	1.28	8.61	-1.11	2.50	4.25	0.64	1.73	14.97	1.75
Longitudinal Cracking	Meters	Low	8.60	7.54	5.13	5.63	68.08	-1.06	9.13	5.39	1.44	1.55	26.70	-3.74	8.07	16.23	32.27	34.50	198.77	8.17
		Moderate	4.30	3.81	1.41	2.58	36.93	-0.49	5.83	1.89	0.33	3.96	17.61	-3.94	2.77	5.32	10.71	10.98	201.45	2.55
		High	0.73	0.33	0.58	0.84	173.21	-0.40	1.47	0.00	N/A	N/A	N/A	-1.47	0.00	1.30	2.69	2.99	206.94	1.30
		Total	13.63	11.69	5.24	6.27	44.82	-1.95	16.43	7.28	1.16	5.49	15.96	-9.16	10.83	22.85	42.34	44.65	185.31	12.02
Transverse Cracking	No.	Low	32.33	40.08	1.28	3.14	3.18	7.75	5.00	8.78	3.64	7.73	41.47	3.78	59.67	71.22	4.55	6.48	6.39	11.56
		Moderate	13.00	5.63	3.56	9.62	63.33	-7.38	5.33	2.11	0.85	3.50	40.35	-3.22	20.67	6.61	6.82	17.73	103.14	-14.06
		High	3.67	3.13	1.81	1.93	57.84	-0.54	3.00	2.11	0.72	0.99	33.97	-0.89	4.33	2.78	3.46	3.77	124.64	-1.56
		Total	49.00	48.83	3.13	6.95	6.42	-0.17	13.33	13.00	3.60	4.94	27.67	-0.33	84.67	80.61	3.54	11.58	4.39	-4.06
Transverse Cracking	Meters	Low	114.57	142.16	14.92	19.82	10.50	27.59	9.43	21.41	7.55	18.88	35.25	11.98	219.70	248.97	25.43	27.97	10.21	29.27
		Moderate	45.38	21.36	13.08	32.26	61.26	-24.03	18.07	6.39	2.53	10.07	39.54	-11.68	72.70	23.26	23.28	63.09	100.10	-49.44
		High	12.98	11.05	6.54	6.92	59.16	-1.93	10.37	7.81	2.76	3.49	35.29	-2.56	15.60	9.09	11.17	12.89	122.90	-6.51
		Total	172.93	174.57	13.92	19.66	7.97	1.64	37.87	35.61	7.58	9.00	21.29	-2.26	308.00	281.32	24.53	49.11	8.72	-26.68
Spalling of Longitudinal Joints	Meters	Low	5.33	18.35	6.81	18.66	37.12	13.02	4.50	5.63	3.30	4.87	58.51	1.13	6.17	8.93	12.45	13.63	139.31	2.77
		Moderate	1.60	6.54	7.72	10.13	118.07	4.94	2.37	1.03	1.24	1.52	120.42	-1.33	0.83	4.02	6.43	7.58	160.07	3.18
		High	0.32	0.55	0.96	1.11	172.62	0.24	0.23	0.09	0.13	0.15	141.42	-0.14	0.40	3.24	7.22	7.92	222.68	2.84
		Total	7.25	25.45	10.25	26.78	40.29	18.20	7.10	6.76	3.97	6.03	58.78	-0.34	7.40	16.19	21.91	25.40	135.30	8.79
Spalling of Transverse Joints	No.	Low	0.80	0.25	0.19	0.32	76.74	-0.55	0.33	0.11	0.16	0.19	141.42	-0.22	1.50	2.83	4.88	5.43	172.36	1.33
		Moderate	0.40	0.55	0.51	0.70	92.14	0.15	0.33	0.11	0.16	0.27	141.42	-0.22	0.50	0.67	1.07	1.08	160.08	0.17
		High	0.60	0.85	0.60	0.75	71.06	0.25	0.67	0.33	0.27	0.43	81.65	-0.33	0.50	0.25	0.39	0.62	157.86	-0.25
		Total	1.80	1.65	0.57	1.13	34.50	-0.15	1.33	0.56	0.16	0.27	28.28	-0.78	2.50	3.75	6.19	6.81	164.93	1.25
Spalling of Transverse Joints	Meters	Low	0.34	0.33	0.26	0.45	78.69	-0.01	0.10	0.36	0.27	0.45	75.59	0.26	0.70	1.94	2.35	2.86	121.25	1.24
		Moderate	0.40	0.40	0.36	0.51	88.87	0.00	0.33	0.44	0.25	0.39	56.79	0.11	0.50	0.89	1.20	1.26	134.32	0.39
		High	0.48	1.69	1.25	1.83	74.11	1.21	0.57	1.46	0.93	1.56	63.97	0.89	0.35	0.33	0.47	0.61	144.90	-0.03
		Total	1.22	2.42	1.23	2.16	51.07	1.20	1.00	2.26	1.15	2.01	50.82	1.26	1.55	3.16	3.78	4.45	119.82	1.61

- **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 were identified using the PASCO/PADIAS system compared with that observed during the manual reference surveys and vice versa. Precision, on the other hand, is being defined in this chapter as the variance about the group mean and not the reference value.

The following observations were made from tables 21 and 22:

- For AC pavements, CV values for total distress quantities range between 14 and 55 percent for experts, 12 and 29 percent for two-person teams, and 12 and 126 percent for individual raters. The average CV values for these three groups are 29, 22, and 70 percent, respectively. These results clearly show that the expert and two-person team surveys provided more consistent data compared with 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 percent for experts, 9 and 59 percent for two-person teams, and 4 and 185 percent for individual raters. The average CV values for these groups are 31, 30, and 90 percent, respectively. It is also clear that the expert and two-person team surveys provide more consistent data compared with those 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.

- 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, that 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 through 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 that covers a wider range of distress types and quantities is required.

3.7 Comparison of PADIAS v1.x Versus v4.x Distress Data

As noted earlier, photographic distress surveys were established as the primary data collection procedure at the beginning of the LTPP program. Data reduction from film was accomplished using the PADIAS v1.x software, which was relatively crude in the resolution of measurements and was not fully in agreement with the 1993 LTPP DIM. Although photographic surveys continued to be used through the Summer of 1996, data reduction from film was stopped in 1992 due to concerns over the accuracy and repeatability of data generated by the PADIAS system.

Because PASCO film is the only source of distress information for the first seven years of the LTPP program and most of the film collected between 1992 to 1996 had not yet been interpreted, FHWA decided to proceed with data reduction from film using the improved PASCO v4.x system, which incorporates the 1993 LTPP DIM procedures and is vector based so that precision is excellent. Consequently, film-derived distress data exist in the LTPP IMS for both PADIAS v1.x and v4.x. Hence, the study presented in this section was undertaken to compare the data generated by both of these systems and, depending on the similarities and/or differences between the two methods, to help FHWA decide what to do with the PADIAS v1.x data currently stored in the IMS.

To achieve these objectives, 24 test sections for which PASCO film had been collected and interpreted using the PADIAS v 1.x system were selected; 8 of those sections were AC pavements, 8 were jointed PCC and the remaining 8 were CRC pavements. The criteria used in the selection of these test sections were presented earlier in this chapter, while the final list of test sections was given in table 10. Data reduction was also performed, as part of this study, on the film for these 24 test sections using the PADIAS v4.x system. Thus, complete sets of PADIAS v1.x and v4.x distress data for the 24 test sections were available for this comparison. Also, as pointed out earlier in this chapter, interpretation of both the v1.x and v4.x distress data were performed using PASCO's production procedure.

A paired t-test was first used to compare the two sets of data at a confidence level of 95 percent. The results of this paired t-test comparison are given in tables 23 through 25, which show the computed t-statistic for each distress type-severity level combination. If the computed t-statistic was less than the critical t-value, it is denoted by the letter "N" on the table, which indicates there are no statistical differences between the PADIAS v1.x and v4.x data being compared. A "Y" indicates that the two sets of data were statistically significant, while "NA" means that the distress type-severity level combination in question was not identified.

It can be seen from these tables that, with few exceptions, there are no significant statistical differences between the data interpreted by the two versions of the PADIAS system. For AC pavements, the only exception is the number (not amount) of transverse cracks at low severity and total (see table 23). Similarly, for PCC pavements, the number (not amount) of low severity transverse cracks are also statistically different between the two methods (see tables 24

Table 23. Comparison of PADIAS v1.x and v4.x, t-test, AC Pavement.

DISTRESS TYPE	UNITS	SEV.	AC	
			t-stat.	Is Difference Significant?
Fatigue Cracking	Sq. M	Low	1.13	N
		Mod.	1.51	N
		High	1.00	N
		Total	0.89	N
Long. Cracking	Meters	Low	2.23	N
		Mod.	0.53	N
		High	NA	NA
		Total	2.22	N
Trans. Cracking	Meters	Low	2.33	N
		Mod.	1.84	N
		High	1.00	N
		Total	2.13	N
Trans. Cracking	No.	Low	2.68	Y
		Mod.	1.43	N
		High	0.00	N
		Total	2.46	Y
Patch/Patch Deterioration	No.	Low	0.61	N
		Mod.	1.53	N
		High	1.00	N
		Total	1.11	N
Patch/Patch Deterioration	Sq. M	Low	0.54	N
		Mod.	1.51	N
		High	1.00	N
		Total	1.01	N
Potholes	Number	Low	1.00	N
		Mod.	NA	NA
		High	NA	NA
		Total	1.00	N
Potholes	Sq. M	Low	1.00	N
		Mod.	NA	NA
		High	NA	NA
		Total	1.00	N
Bleeding	Sq. M	Low	1.19	N
		Mod.	NA	NA
		High	NA	NA
		Total	1.19	N
Raveling and Weathering	Sq. M	Low	1.25	N
		Mod.	1.00	N
		High	NA	NA
		Total	1.24	N

Table 24. Comparison of PADIAS v1.x and v4.x, t-test, JCP Pavement.

DISTRESS TYPE	UNITS	SEV.	JPC	
			t-stat.	Is Difference Significant?
Corner Break	No.	Low	1.00	N
		Mod.	0.55	N
		High	1.00	N
		Total	2.05	N
Long. Cracking	Meters	Low	2.11	N
		Mod.	1.10	N
		High	1.00	N
		Total	1.99	N
Trans. Cracking	Meters	Low	0.66	N
		Mod.	0.01	N
		High	1.58	N
		Total	0.54	N
Trans. Cracking	No.	Low	4.15	Y
		Mod.	0.63	N
		High	1.57	N
		Total	1.14	N
Transverse Joint Seal Damage	Meters	Low	2.54	Y
		Mod.	1.99	N
		High	2.09	N
		Total	7.77	N
Spalling of Longitudinal Joint	Meters	Low	1.32	N
		Mod.	1.97	N
		High	1.76	N
		Total	1.70	N
Spalling of Transverse Joint	Meters	Low	2.02	N
		Mod.	0.96	N
		High	0.92	N
		Total	2.29	N

Table 24. Comparison of PADIAS v1.x and v4.x, t-test, JCP Pavement (Continued).

DISTRESS TYPE	UNITS	SEV.	JPC	
			t-stat.	Is Difference Significant?
Spalling of Transverse Joint	No.	Low	1.48	N
		Mod.	0.18	N
		High	1.78	N
		Total	0.05	N
Polished Aggregate	Sq. M.		1.00	N
Popouts	No.		1.06	N
Lane to Shoulder Separation	Sq. M.	Low	3.12	Y
		Mod.	1.16	N
		High	1.00	N
		Total	3.49	Y
AC Patch	Sq. M.	Low	1.66	N
		Mod.	1.59	N
		High	1.75	N
		Total	1.85	N
AC Patch	No.	Low	1.70	N
		Mod.	1.69	N
		High	0.18	N
		Total	1.88	N
PCC Patch	Sq. M.	Low	1.00	N
		Mod.	NA	NA
		High	NA	NA
		Total	1.00	N
PCC Patch	No.	Low	1.00	N
		Mod.	NA	NA
		High	NA	NA
		Total	1.00	N

Table 25. Comparison of PADIAS v1.x and v4.x, t-test, CRC Pavement.

DISTRESS TYPE	UNITS	SEV	CRC	
			t-stat.	Is Difference Significant?
Long. Cracking	Meters	Low	1.75	N
		Mod.	0.93	N
		High	NA	NA
		Total	1.91	N
Trans. Cracking	Meters	Low	1.87	N
		Mod.	1.49	N
		High	1.00	N
		Total	0.65	N
Trans. Cracking	No.	Low	2.51	Y
		Mod.	1.50	N
		High	1.00	N
		Total	1.25	N
Spalling of Longitudinal Joint	Meters	Low	0.85	N
		Mod.	0.96	N
		High	NA	NA
		Total	0.84	N
Scaling	Sq. M.	Low	1.00	N
		Mod.	NA	NA
		High	NA	NA
		Total	1.00	N
Popouts	No.		1.74	N
Lane to Shoulder Separation	Sq. M.	Low	5.88	Y
		Mod.	1.00	N
		High	NA	NA
		Total	5.89	Y
AC Patch	Sq. M.	Low	1.00	N
		Mod.	0.31	N
		High	NA	NA
		Total	1.00	N
AC Patch	No.	Low	1.00	N
		Mod.	0.00	N
		High	NA	NA
		Total	1.55	N
PCC Patch	Sq. M.	Low	1.00	N
		Mod.	1.00	N
		High	NA	NA
		Total	0.36	N
PCC Patch	No.	Low	1.00	N
		Mod.	1.00	N
		High	NA	NA
		Total	1.42	N
Punchouts	No.	Low	1.53	N
		Mod.	NA	NA
		High	NA	NA
		Total	1.53	N

and 25). The only other exceptions for PCC pavements (jointed PCC only) include lane-to-shoulder separation (low severity and total) and low severity transverse joint seal damage.

Next, a subjective comparison of the PADIAS v1.x versus v4.x reduced data was performed using distress magnitude plots generated for each pavement type, distress type, and severity level combination. The complete set of plots is contained in appendix B of this report. Example plots are given in figures 36 through 38, which show the transverse cracking comparison, at all severity levels and total, for AC, jointed PCC, and CRC pavements, respectively.

The following general observations are based on the information presented in these plots:

- Although differences exist, there is excellent overall agreement between the two PADIAS systems for all three pavement types. This is particularly true for the cracking-related distresses, including fatigue (AC only), transverse, and longitudinal cracking (all pavement types). Larger discrepancies between the two systems were observed for surface and joint-related defects, such as joint seal damage, joint spalling, and AC and PCC patches. These discrepancies were not detected by the paired t-test analysis.

- Regardless of pavement type, there is generally better agreement between the two systems for total distress summed across all severity levels than for the individual severity levels.

- Although there is excellent agreement between the two methods and the statistical comparisons show no significant differences, the total quantity of distress identified using the PADIAS v4.x system appears slightly higher than that found with the PADIAS v1.x system.

- Both versions of PADIAS seemed to have trouble identifying low-severity transverse cracking in CRC pavement sections. Typically, a 152-m CRC pavement section is expected to have about 100 to 200 cracks. However, the data from film only showed 5 to 10 cracks over this length of section.

Thus, while significant differences were expected (hence the reason interpretations with PADIAS v4.x began in 1996), both the statistical and subjective-based comparisons performed in this study indicate that there is excellent agreement between those data derived using the PADIAS v1.x and v4.x systems. Although this outcome must be viewed within the context of the overall data study results, it appears to indicate that, although improvements have been made from one version of the system to the other, the amount of information gained by going to PADIAS v4.x is not significant. This, in turn, leads to a number of questions such as: Are distress data derived during the early years of the LTPP program using the PADIAS v1.x system still questionable? If so, does it infer that those data now being derived by PADIAS v4.x are also questionable? Although possible answers to these and others questions are addressed in the final chapter of this report, more definitive ones are beyond the scope of this

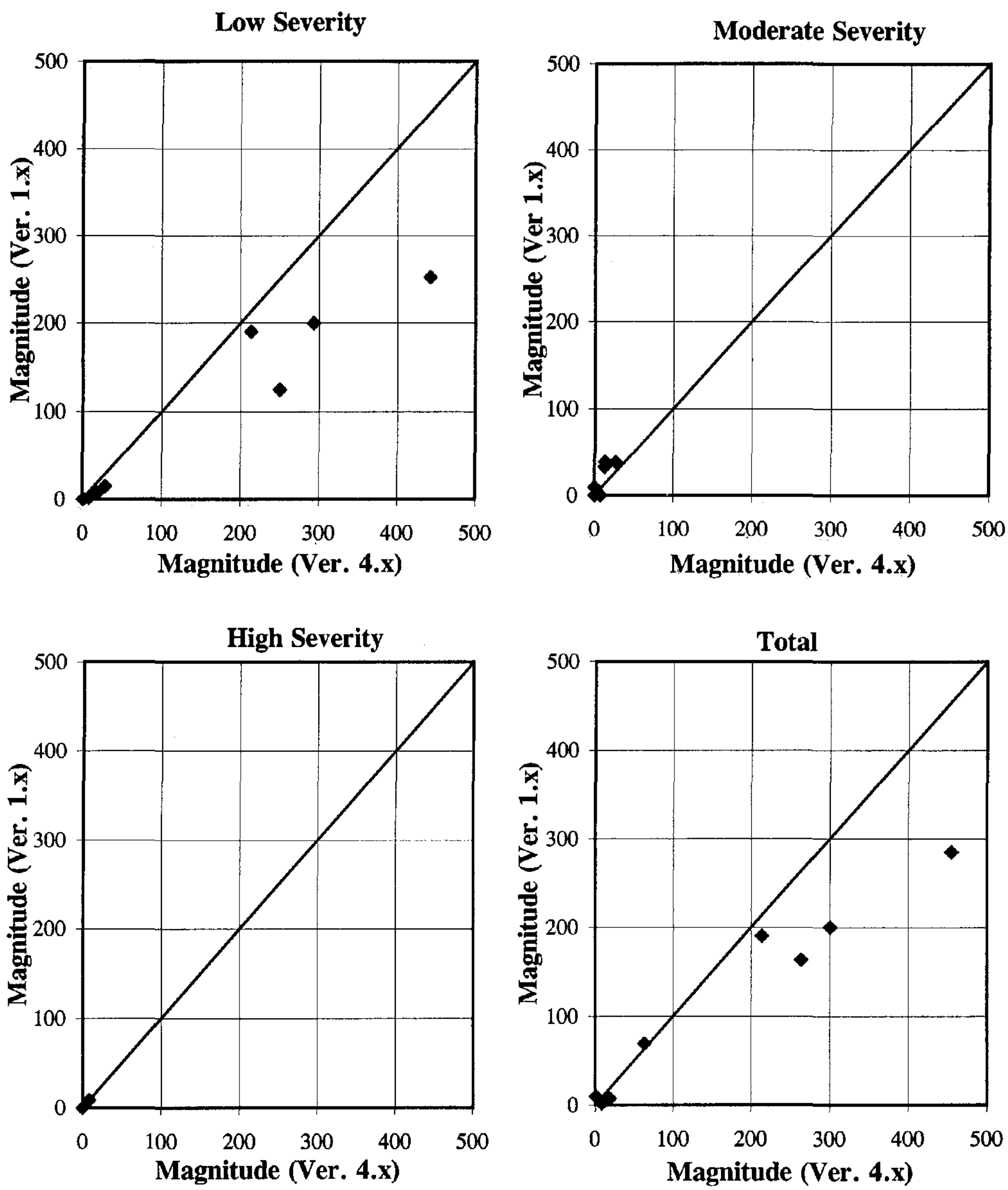


Figure 36. Comparison of PADIAS v1.x and v4.x - AC Pavements, Transverse Cracking (Meters).

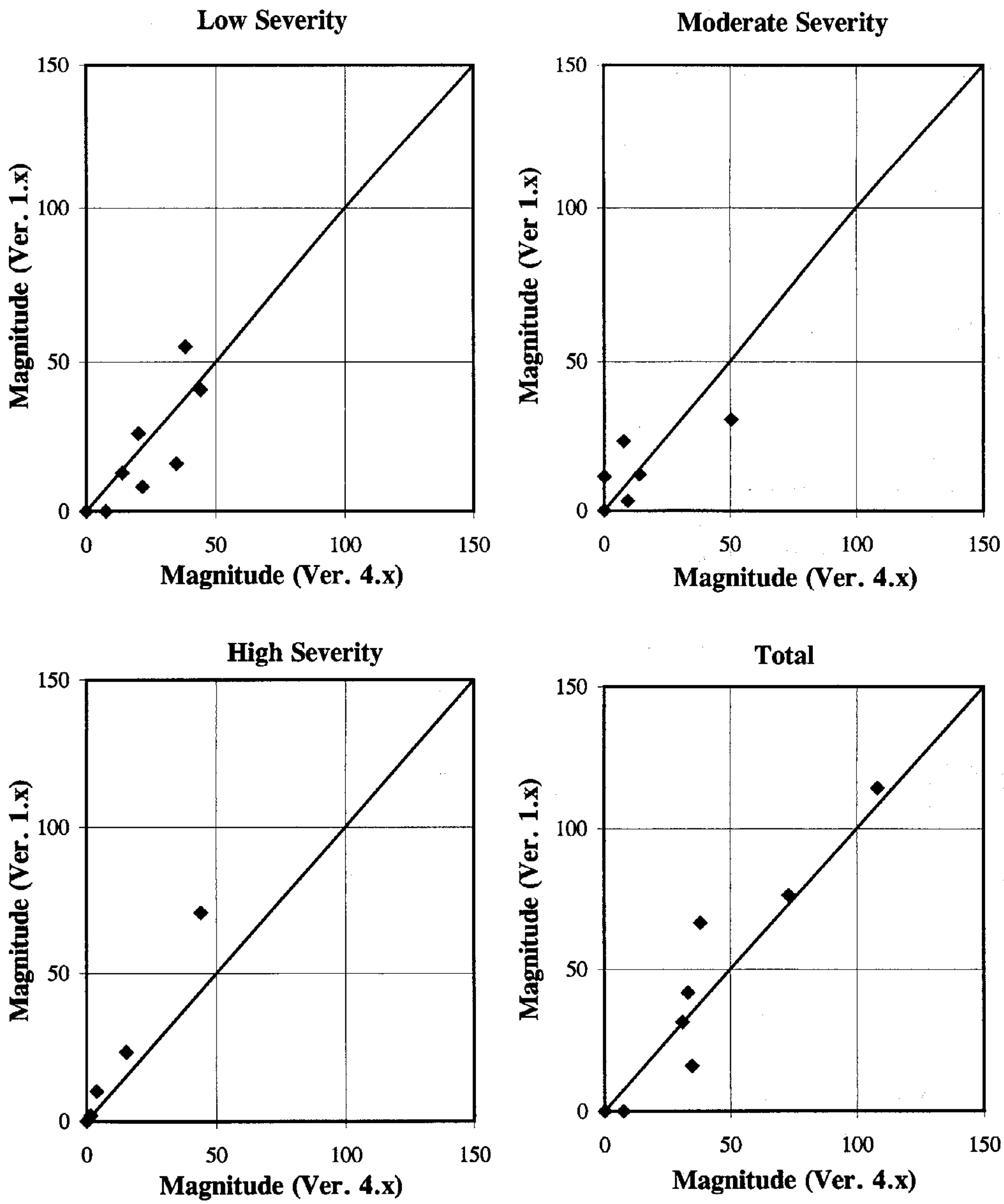


Figure 37. Comparison of PADIAS v1.x and v4.x - JPC Pavements, Transverse Cracking (Meters).

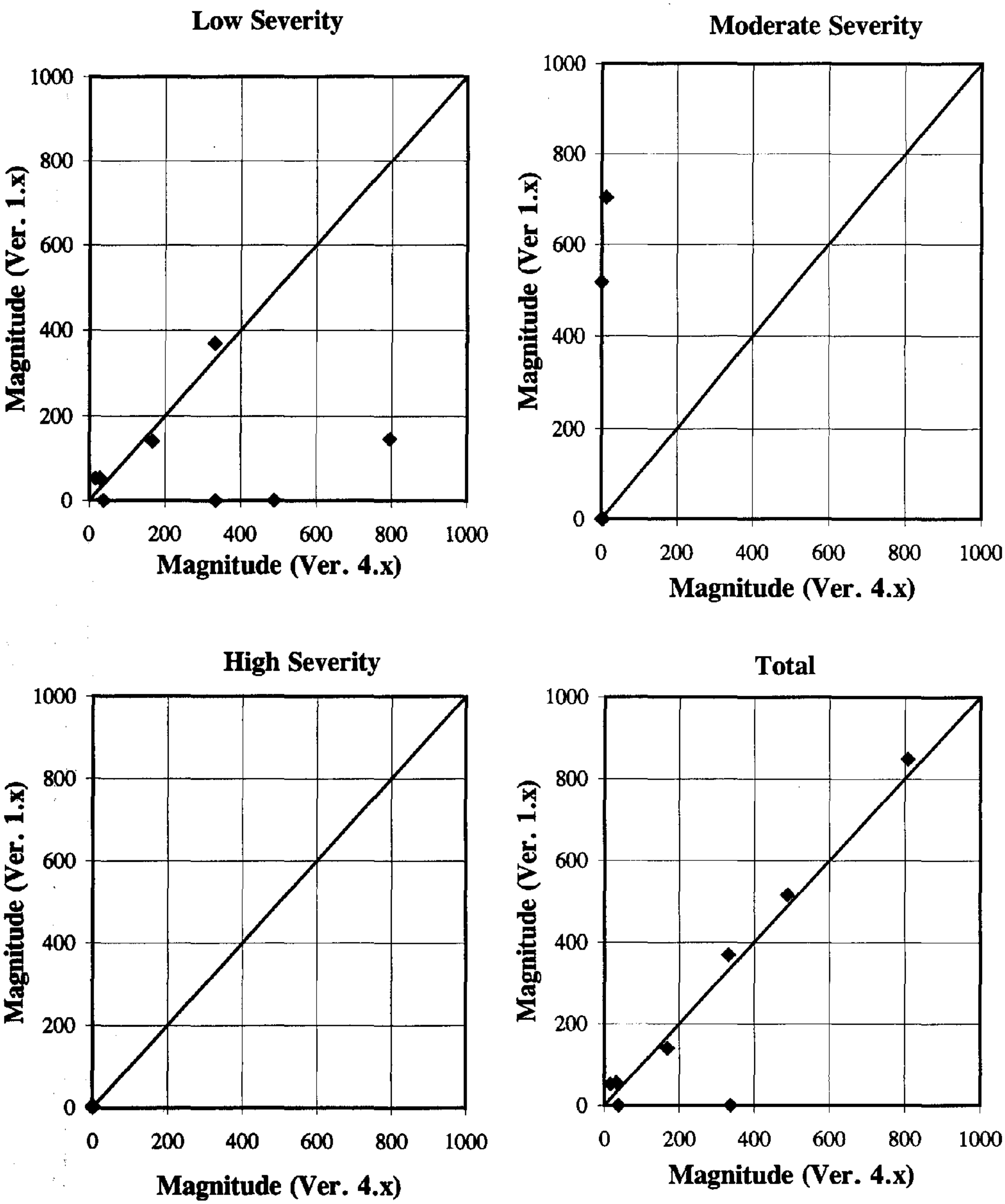


Figure 38. Comparison of PADIAS v1.x and v4.x - CRC Pavements, Transverse Cracking (Meters).

study. Further data studies looking into specific issues such as distress time-series may provide more clear answers, which in turn will better help set future direction.

3.8 Analysis of Data Interpreted by PASCO's Production Procedure

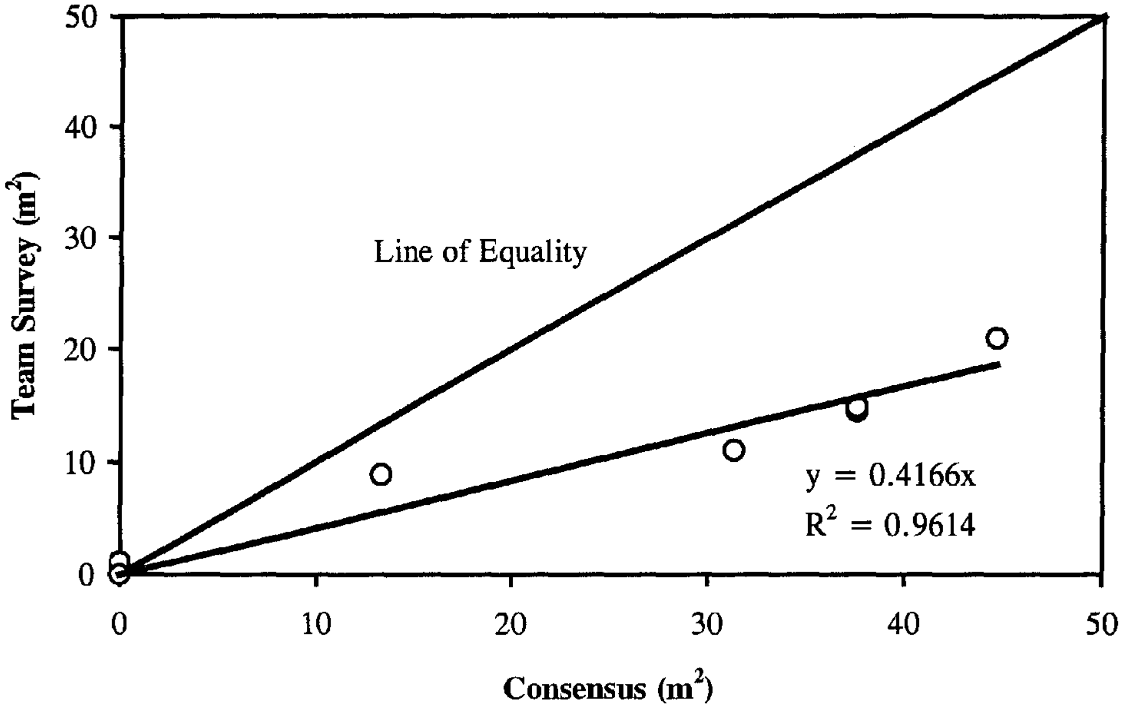
In this section, an analysis was performed to evaluate variability of film-derived distress data interpreted using PASCO's production procedure. This production procedure consisted of film interpretation by two individual raters or two two-person teams and then refereed by a third rater to determine the final severity and amount of each particular distress. For AC pavements, the individual raters interpreted AC sections 2, 3, and 4, while teams surveyed AC sections 4, 5, and 6. For PCC pavements, sections 1, 2, and 6 were surveyed by individual raters and sections 3, 4, and 5 were analyzed by teams.

Since each pavement section was surveyed only once (one distress quantity for each severity level of each distress), no statistical or objective analysis could be performed to quantify variability associated with distress data interpreted by this production procedure. Rather, plots were prepared to compare distress data obtained using the production procedure and their corresponding reference and consensus values. As defined earlier, a reference value is a quantity of distress determined by the consensus manual distress survey conducted by a group of experts. Similarly, a consensus value is a quantity of distress determined by the consensus expert survey conducted using the PADIAS v4.x system.

Example plots are presented in figures 39 to 42, while the complete set of figures can be found in appendix B of the report. Figures 39 and 40 are for fatigue cracking of AC pavements and figures 41 and 42 are for transverse cracking of PCC pavements. In these figures, for each distress type, all severity levels as well as total distress are presented for all three pavement sections evaluated.

For AC pavement sections, both the individual raters and teams consistently underestimated the amount of distresses compared with consensus and reference values for fatigue cracking. For longitudinal cracking in the wheel path, the values were scattered widely and no trends could be observed. These observations were consistent with earlier findings that raters had difficulty distinguishing between these two types of distresses. For longitudinal cracking not in the wheel path and transverse cracking, distress data interpreted by both individual raters and teams were comparable to their corresponding consensus and reference values.

For PCC pavements, distress data obtained by both individual raters and teams using PASCO's production procedure compared favorably with their corresponding consensus and reference values for transverse and longitudinal cracking. For corner cracking, this procedure tended to underestimate the distress quantity. Distress data were scattered for joint spalling, indicating difficulties for raters to identify and distinguish this type of pavement distress. Also, for both AC and PCC pavement sections, teams did not provide better or more consistent distress data than individual raters.



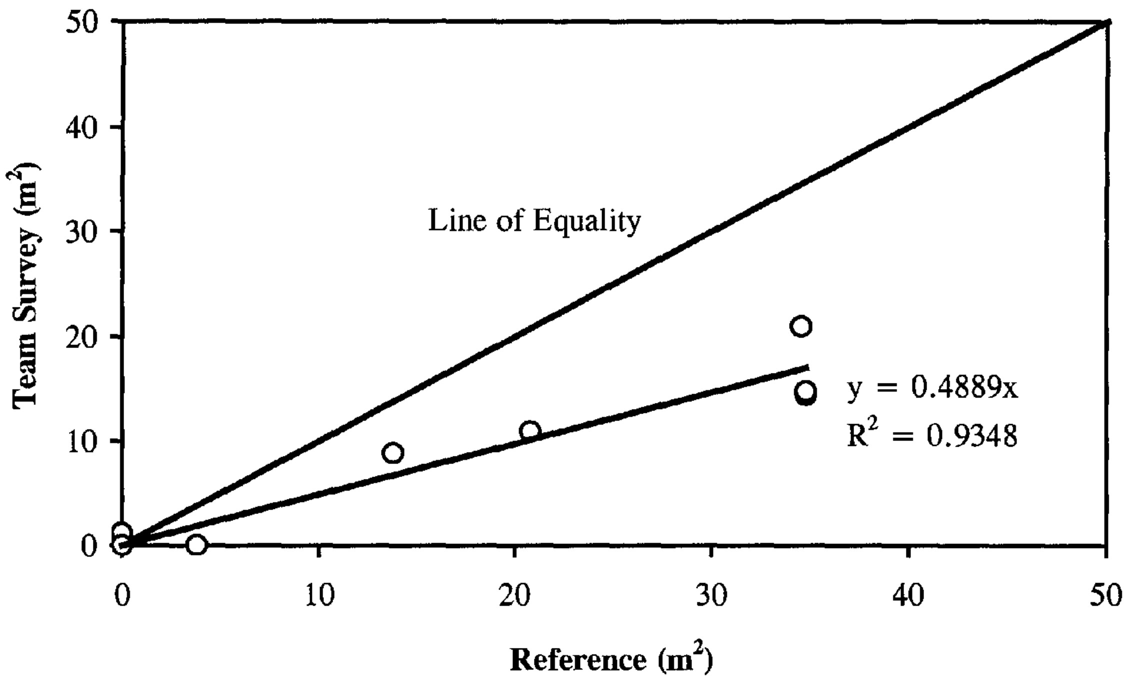
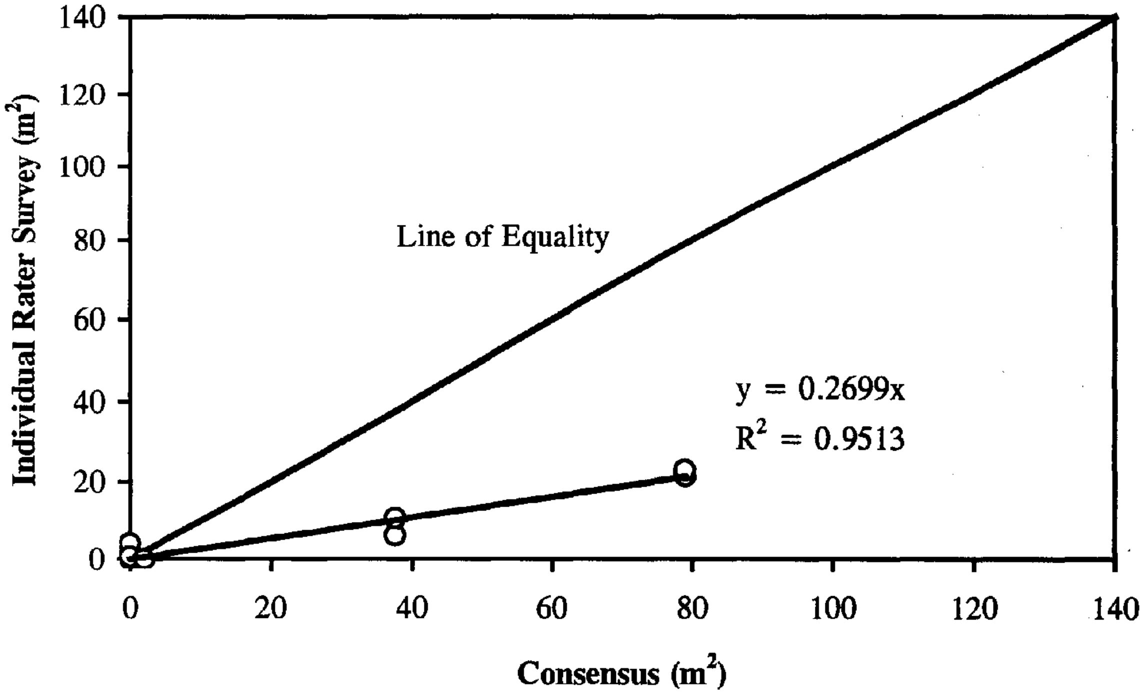


Figure 39. Team Survey Vs. Reference and Consensus: Fatigue Cracking of AC Pavement, PASCO's Production Procedure.



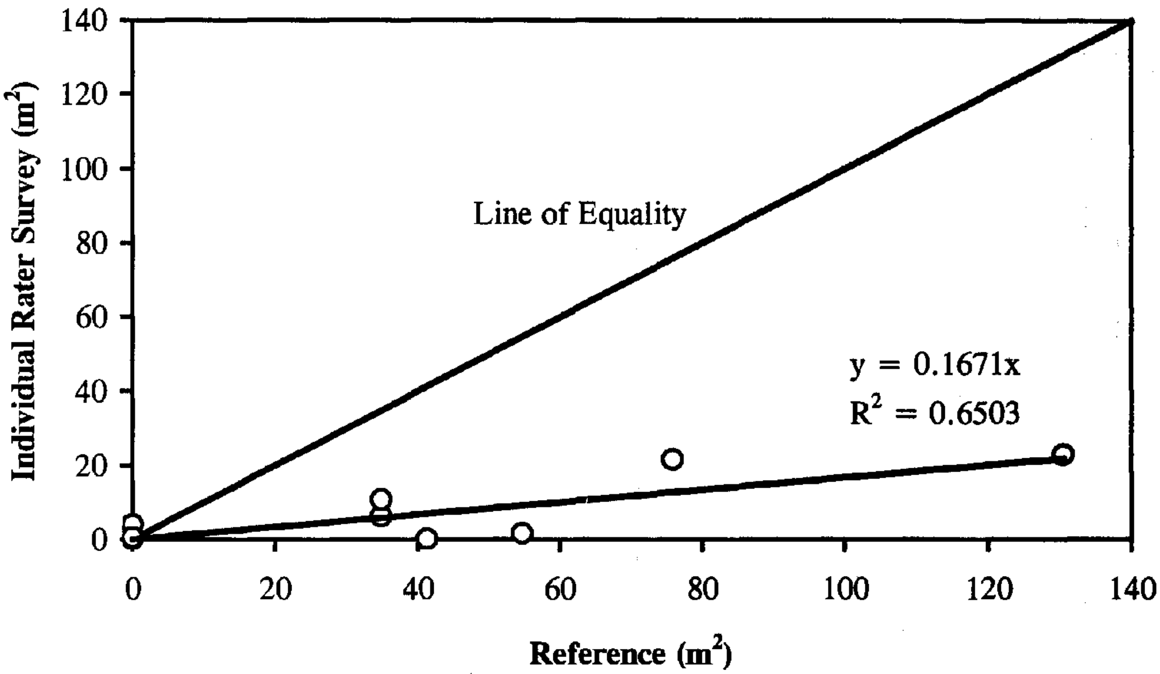


Figure 40. Individual Survey Vs. Reference and Consensus: Fatigue Cracking of AC Pavement, PASCO's Production Procedure.

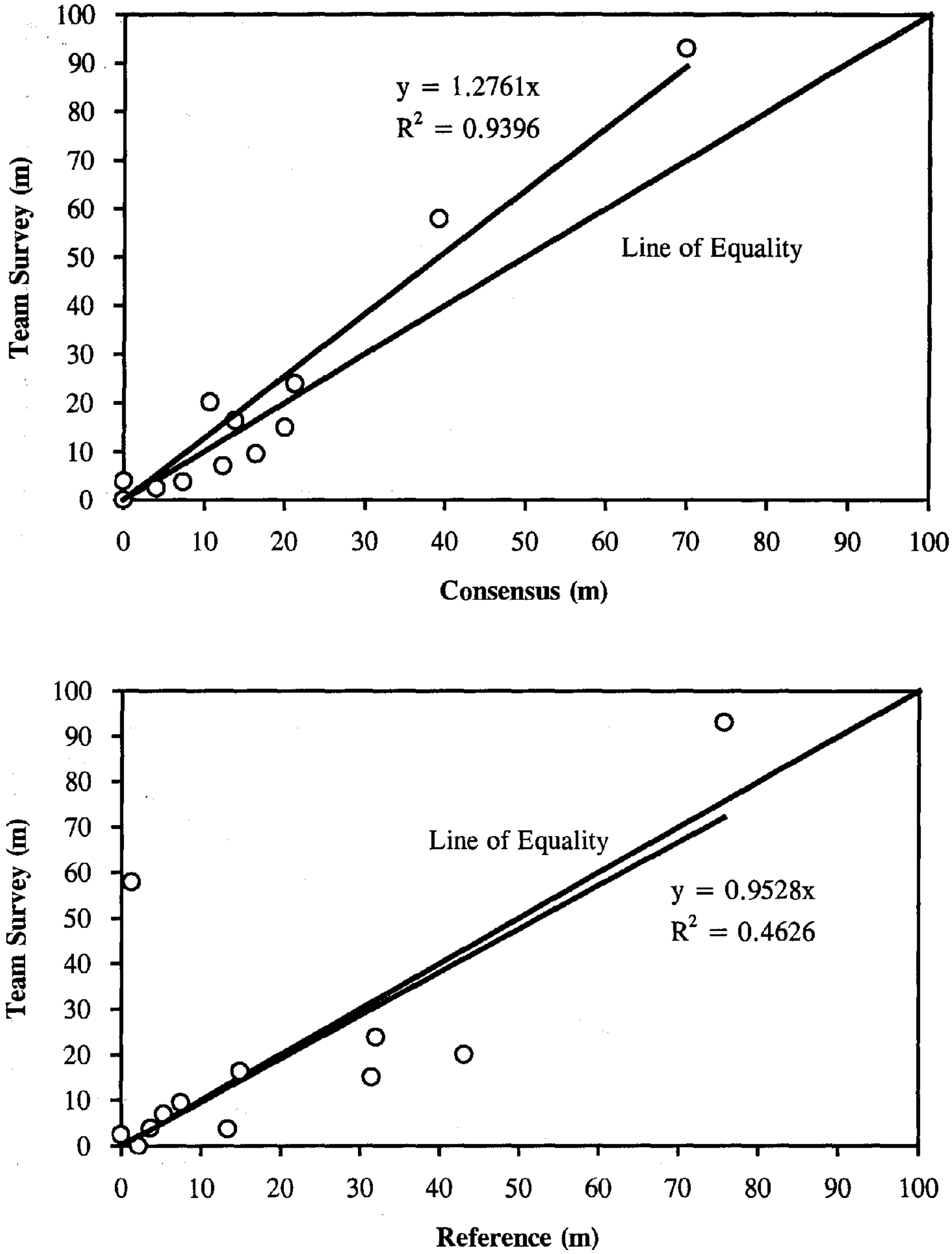
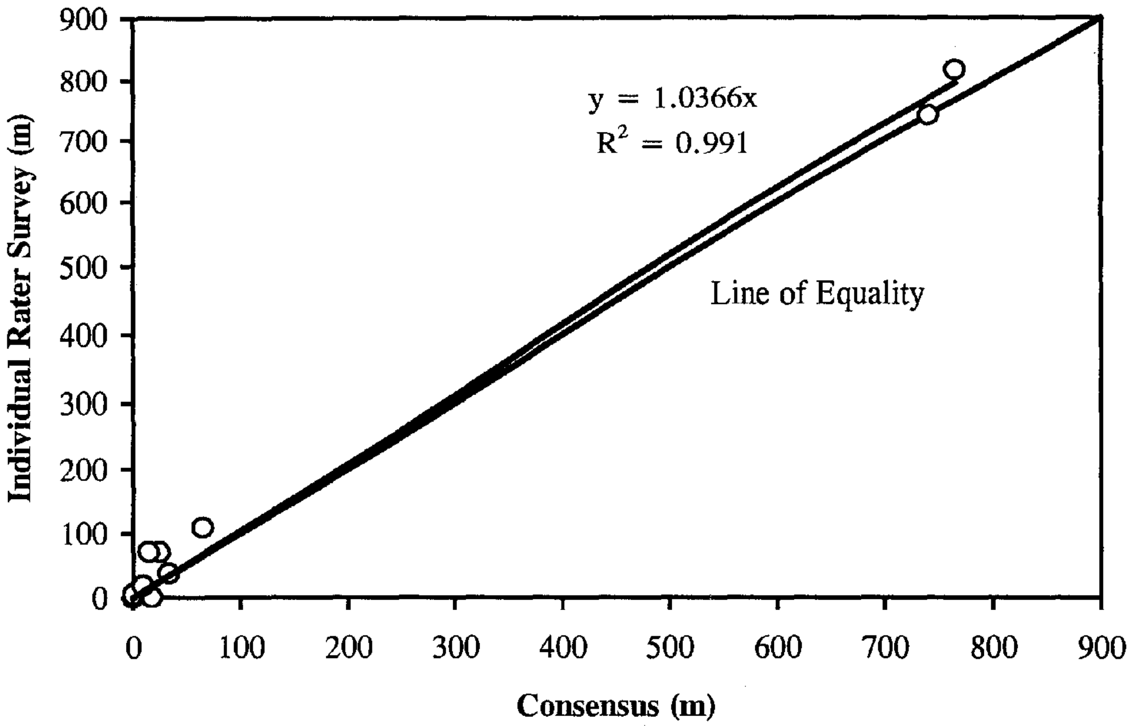


Figure 41. Team Survey Vs. Reference and Consensus: Transverse Cracking (Meters) of PCC Pavement, PASCO's Production Procedure.



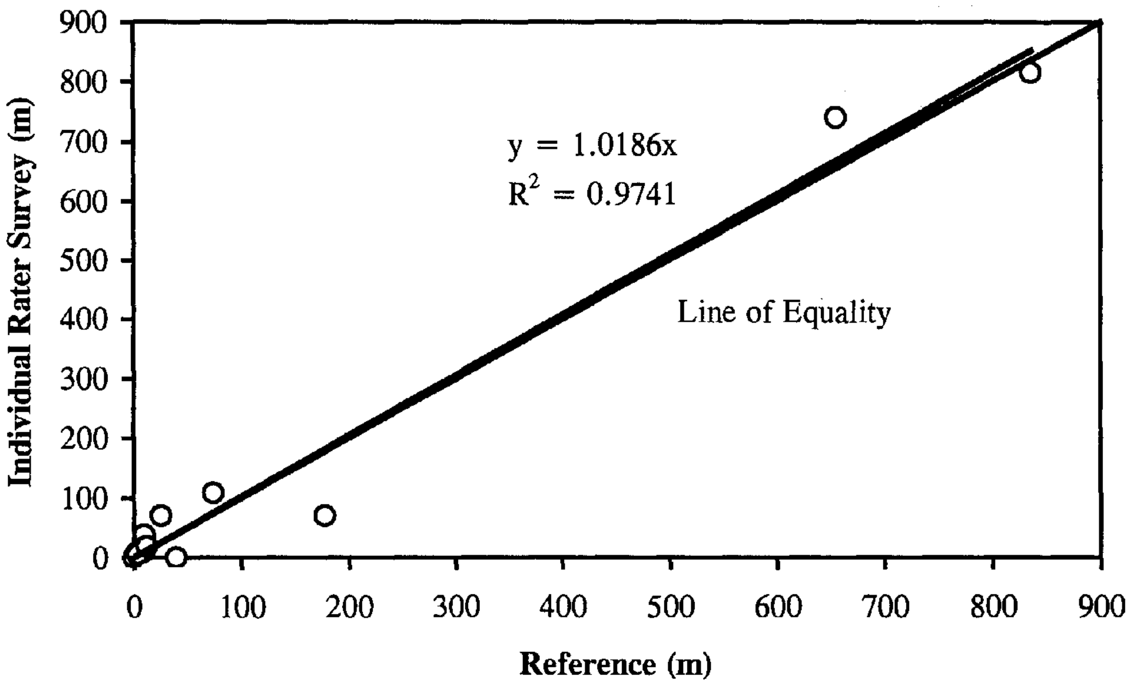


Figure 42. Individual Survey Vs. Reference and Consensus: Transverse Cracking (Meters) of PCC Pavement, PASCO's Production Procedure.

3.9 Summary and Conclusions

PASCO film is the only source of distress data for the first five to seven years of the LTPP program. Film collected through 1992 was interpreted using the PADIAS v1.x system. In 1992, interpretation of PASCO film was discontinued due to concerns over quality of data derived using the PADIAS system. However, in order to support analysis efforts under way at the time, FHWA decided in 1996 to proceed with data reduction from film using the improved PASCO v4.x system. Thus, film-derived distress data exist in the LTPP IMS for both PADIAS v1.x and v4.x, but their quality was unknown. Hence, this study was undertaken to assess the variability of distress data derived from film using PADIAS v4.x, the system currently in use, and to compare data generated by this system with that from the PADIAS v1.x system, which was used to generate the early LTPP distress data.

Film interpretations performed by a consensus team of experts, individual experts, individual raters, and two-person teams on 12 test sections (6 AC and 6 PCC pavements) provided the data used to assess the quality of the PADIAS v4.x distress data. Film for 24 test sections (8 AC, 8 jointed PCC, and 8 CRC pavements) interpreted with the PADIAS v1.x system during the early years of the LTPP program was re-interpreted using the PADIAS v4.x system. These two data sets served as the basis for comparison of the two methods.

On the basis of analyses of these data, a number of observations and/or conclusions were made. The most important ones include:

- Although not a standard practice within the LTPP program, the repeatability 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.

- Although repeatability is good, variance within a given group – experts, individual raters, and two-person teams – is statistically large. Furthermore, this variance appears to get worse as distress quantity increases.

- 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.

- In contrast to the manual surveys, the variability for total distress quantities was not better than that of the individual severity levels. However, there do appear to be compensatory differences for a number of closely related distresses.

- 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 through 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 that covers a wider range of distress types and quantities is required.

- Although differences in data interpreted with the PADIAS v1.x and v4.x systems exist, there is excellent overall agreement between the two systems for all pavement types, especially for total distress quantities and for cracking-related distresses. Larger discrepancies exist for surface and joint-related defects.

- Distress data derived from PASCO's production procedure exhibit good agreement with their corresponding consensus and reference values for longitudinal cracking not in the wheel path and transverse cracking of AC pavement sections, and for cracking-related distresses of PCC pavement sections. However, raters continue to have difficulty identifying fatigue cracking and longitudinal cracking in the wheel path of AC pavements and for joint-related distresses of PCC pavements.

4. MANUAL VERSUS PASCO/PADIAS: A COMPARATIVE STUDY

4.1 Introduction

The results of the analysis of variability associated with pavement distress data from manual field raters and film-based interpretation by PASCO/PADIAS process were presented in two previous chapters. A comparative study of the variability between the two methods is presented in this chapter.

4.2 General Assessment

First, a global comparison of each method was conducted by combining all of the data available for each rating method. Since each rating method was performed on different combinations of test sections, this approach does not provide for a one-to-one comparison; however, it does permit the general trends between each method to be compared. The results of the bias and precision computations are presented in tables 26 and 27. The definitions of the statistics presented in these tables are presented in chapters 2 and 3. For this comparison, two additional relative statistics were provided. The percentage of RMSE and BIAS were computed by dividing the corresponding statistic by the reference value and multiplying by 100. To properly understand this comparison, the differences in the reference values should be compared for each distress type-severity combination. Note that these values are different because the test sections from which these statistics are computed are not the same for each interpretation method.

Overall the bias and precision for the group of manual field distress raters is smaller than for the PASCO/PADIAS interpreters. For example, for total distress of AC pavements, the coefficient of variation (CV) varied from 9 to 38 percent, with an average value of 23 percent, while the CV ranged from 12 to 126 percent, with an mean value of 69 percent for distress data interpreted by PASCO/PADIAS. This observation was generally true for both apparent bias and CV, and for both AC and PCC pavement distress data.

4.3 Head-to-head Comparison of Manual and PASCO/PADIAS

Head-to head comparisons of the two interpretation methods were possible on four test sections, two AC and two PCC, from accreditation workshop Nos. 8 and 9. Both field manual and film-based PASCO/PADIAS distress ratings were performed by a combination of individual raters and rater teams, as shown in table 2 for test sections AC1, AC6, PCC1, and PCC6.

High-low range and average charts were prepared to illustrate the relative difference between the reference values (R), manual (M), and PASCO/PADIAS methods. Figure 43 shows the results obtained for fatigue cracking on AC pavements from individual raters, while figure 44

Table 26. Relative Indicators of Bias and Precision for AC Pavement Distresses, Individual Raters.

Distress Type	Unit	Distress Sev.	MANUAL SURVEY REF.	Mean	STD	RMSE	% RMSE	COV	Bias	%Bias	PASCO/PADIAS REF.	Mean	STD	RMSE	% RMSE	COV	Bias	%Bias
Fatigue Cracking	Meters	Low	8.2	8.8	4.2	6.1	73.7	47.9	0.6	7.6	49.2	12.4	17.6	31.4	63.8	141.5	-36.7	-74.7
		Moderate	5.3	7.0	4.1	4.4	83.2	58.1	1.7	33.1	22.5	4.5	5.5	20.7	92.1	121.9	-18.0	-80.0
		High	0.7	0.6	1.0	1.3	173.6	169.9	-0.1	-16.5	1.4	0.1	0.2	1.4	94.3	223.6	-1.3	-93.0
		Total	14.2	16.5	6.2	9.5	66.5	37.7	2.3	15.9	73.1	17.0	21.5	51.8	70.9	126.2	-56.1	-76.7
Longitudinal Cracking WP	Meters	Low	8.8	8.6	2.3	2.9	33.4	26.5	-0.3	-3.0	6.3	39.4	22.9	42.9	684.7	58.2	33.1	528.2
		Moderate	5.0	8.0	4.2	6.7	133.5	53.2	3.0	60.2	5.1	17.6	13.4	19.4	380.3	76.3	12.5	245.2
		High	4.6	1.7	2.4	5.1	109.5	141.9	-2.9	-63.1	1.1	4.2	4.7	5.6	513.1	111.8	3.1	282.8
		Total	18.4	18.3	6.0	7.4	40.3	33.0	-0.2	-1.0	12.5	61.2	19.8	57.9	464.6	32.4	48.7	390.8
Longitudinal Cracking NWP	Meters	Low	22.9	26.2	18.6	19.5	85.2	71.0	3.3	14.4	85.6	87.8	51.1	57.3	67.0	58.2	2.1	2.5
		Moderate	35.7	26.6	16.8	19.6	54.8	63.2	-9.2	-25.7	18.3	18.2	8.6	16.2	88.5	47.5	-0.1	-0.8
		High	16.3	18.0	6.7	8.9	54.5	37.2	1.6	10.0	24.8	4.5	4.1	11.9	48.0	92.7	-20.3	-82.0
		Total	75.0	70.7	14.7	16.3	21.7	20.7	-4.3	-5.7	128.7	110.4	57.2	83.1	64.5	51.9	-18.4	-14.3
Transverse Cracking	Number	Low	10.6	11.7	2.4	3.5	32.8	20.1	1.2	11.2	44.7	42.5	8.6	18.2	40.7	20.2	-2.2	-4.9
		Moderate	8.3	7.3	2.6	5.0	60.4	35.2	-1.0	-12.5	15.3	17.2	3.9	5.9	38.5	22.7	1.9	12.3
		High	7.6	5.7	2.1	2.9	37.8	37.4	-1.9	-24.6	12.7	10.8	6.9	7.2	56.5	63.9	-1.9	-14.9
		Total	26.4	24.7	3.2	4.0	15.1	13.1	-1.7	-6.5	72.7	70.5	8.6	21.8	30.0	12.2	-2.2	-3.0
Transverse Cracking	Meters	Low	10.9	13.8	3.0	5.1	46.9	21.4	2.9	26.4	49.2	12.4	17.6	31.4	63.8	141.5	-36.7	-74.7
		Moderate	15.8	16.9	6.0	9.0	57.3	35.2	1.2	7.4	22.5	4.5	5.5	20.7	92.1	121.9	-18.0	-80.0
		High	17.6	13.9	5.0	6.6	37.4	35.6	-3.7	-21.1	1.4	0.1	0.2	1.4	94.3	223.6	-1.3	-93.0
		Total	44.3	44.6	4.2	5.3	12.0	9.4	0.3	0.8	73.1	17.0	21.5	51.8	70.9	126.2	-56.1	-76.7

Notes:
REF = reference value
MEAN = average value of individual raters
STD = standard deviaitoin
RSME = root mean square error
% RMSE = RMSE/REF * 100
COV = coefficient of variation, STD/MEAN * 100
BIAS = MEAN - REF
% BIAS = BIAS / REF * 100

Table 27. Relative Indicators of Bias and Precision for PCC Pavement Distresses, Individual Raters.

Distress Type	Unit	Distress Sev.	MANUAL SURVEY REF.	Mean	STD	RMSE	% RMSE	COV	Bias	%Bias	PASCO/PADIAS REF.	Mean	STD	RMSE	% RMSE	COV	Bias	%Bias
Corner Breaks	Number	Low	0.4	0.8	0.4	0.5	107.7	49.8	0.3	72.5	0.0	1.3	0.9	1.5	NA	69.3	1.3	NA
		Moderate	2.8	2.1	0.8	1.0	36.4	36.4	-0.7	-25.3	2.5	2.7	1.4	1.4	54.6	51.1	0.2	6.7
		High	0.8	0.8	0.5	0.6	83.1	63.2	0.1	9.0	0.0	0.3	0.7	0.8	NA	223.6	0.3	NA
		Total	3.9	3.7	0.5	0.6	14.4	13.9	-0.2	-5.2	2.5	4.3	0.6	1.7	69.3	15.0	1.8	70.0
Longitudinal Cracking	Meters	Low	3.6	3.6	1.1	1.2	33.7	31.1	0.0	-1.0	8.1	16.2	32.3	34.5	427.7	198.8	8.2	101.2
		Moderate	3.2	3.0	1.1	1.1	35.5	36.4	-0.2	-6.3	2.8	5.3	10.7	11.0	396.9	201.5	2.6	92.2
		High	0.8	0.5	0.5	0.7	88.1	99.2	-0.3	-35.4	0.0	1.3	2.7	3.0	NA	206.9	1.3	NA
		Total	7.5	7.0	1.6	1.8	24.2	22.1	-0.5	-6.7	10.8	22.9	42.3	44.6	412.1	185.3	12.0	110.9
Transverse Cracking	No.	Low	6.9	6.8	4.1	6.3	92.0	60.5	0.0	-0.6	59.7	71.2	4.6	6.5	10.9	6.4	11.6	19.4
		Moderate	11.5	12.0	2.7	6.2	53.9	22.4	0.5	4.2	20.7	6.6	6.8	17.7	85.8	103.1	-14.1	-68.0
		High	6.4	6.2	2.4	2.6	39.9	38.5	-0.2	-3.3	4.3	2.8	3.5	3.8	86.9	124.6	-1.6	-35.9
		Total	24.8	25.0	2.1	2.4	9.6	8.3	0.2	0.9	84.7	80.6	3.5	11.6	13.7	4.4	-4.1	-4.8
Transverse Cracking	Meters	Low	4.0	4.1	1.5	2.1	51.4	37.2	0.1	2.0	219.7	249.0	25.4	28.0	12.7	10.2	29.3	13.3
		Moderate	3.4	3.7	1.0	2.2	63.9	26.6	0.3	8.2	72.7	23.3	23.3	63.1	86.8	100.1	-49.4	-68.0
		High	2.0	1.8	0.7	0.8	38.8	40.9	-0.2	-10.5	15.6	9.1	11.2	12.9	82.6	122.9	-6.5	-41.7
		Total	9.4	9.6	1.4	1.7	18.2	14.9	0.2	1.6	308.0	281.3	24.5	49.1	15.9	8.7	-26.7	-8.7
Spalling of Longitudinal Joints	Meters	Low	5.9	5.9	6.2	6.9	118.0	104.9	0.0	0.2	6.2	8.9	12.4	13.6	221.1	139.3	2.8	44.9
		Moderate	0.8	0.8	0.6	0.9	118.8	70.6	0.0	6.4	0.8	4.0	6.4	7.6	909.8	160.1	3.2	382.0
		High	0.0	0.5	0.7	0.8	7496.4	134.2	0.5	4379	0.4	3.2	7.2	7.9	1980.6	222.7	2.8	711.1
		Total	6.6	7.2	4.9	5.5	83.1	68.2	0.5	8.3	7.4	16.2	21.9	25.4	343.2	135.3	8.8	118.8
Spalling of Transverse Joints	No.	Low	1.3	1.1	1.1	1.5	114.5	97.0	-0.2	-16.4	1.5	2.8	4.9	5.4	362.3	172.4	1.3	88.9
		Moderate	0.2	0.4	0.7	0.8	471.1	186.6	0.2	132.1	0.5	0.7	1.1	1.1	216.0	160.1	0.2	33.3
		High	0.2	0.5	0.1	0.3	155.3	26.6	0.3	144.8	0.5	0.3	0.4	0.6	124.1	157.9	-0.3	-50.0
		Total	1.7	2.0	1.4	2.0	120.0	71.2	0.3	19.5	2.5	3.8	6.2	6.8	272.5	164.9	1.3	50.0
Spalling of Transverse Joints	Meters	Low	2.9	2.5	0.7	1.0	35.5	28.2	-0.4	-13.4	0.7	1.9	2.4	2.9	409.0	121.3	1.2	177.4
		Moderate	0.4	0.6	0.8	0.9	199.7	143.2	0.1	24.2	0.5	0.9	1.2	1.3	252.0	134.3	0.4	78.3
		High	0.3	0.4	0.3	0.4	126.3	87.8	0.0	8.0	0.4	0.3	0.5	0.6	175.4	144.9	0.0	-7.1
		Total	3.7	3.4	0.9	1.1	30.6	25.5	-0.3	-7.0	1.6	3.2	3.8	4.5	287.4	119.8	1.6	103.8

Notes:

REF = reference value
MEAN = average value of individual raters
STD = standard deviaitoin
RSME = root mean square error

% RMSE = RMSE/REF * 100
COV = coefficient of variation, STD/MEAN * 100
BIAS = MEAN - REF
% BIAS = BIAS / REF * 100

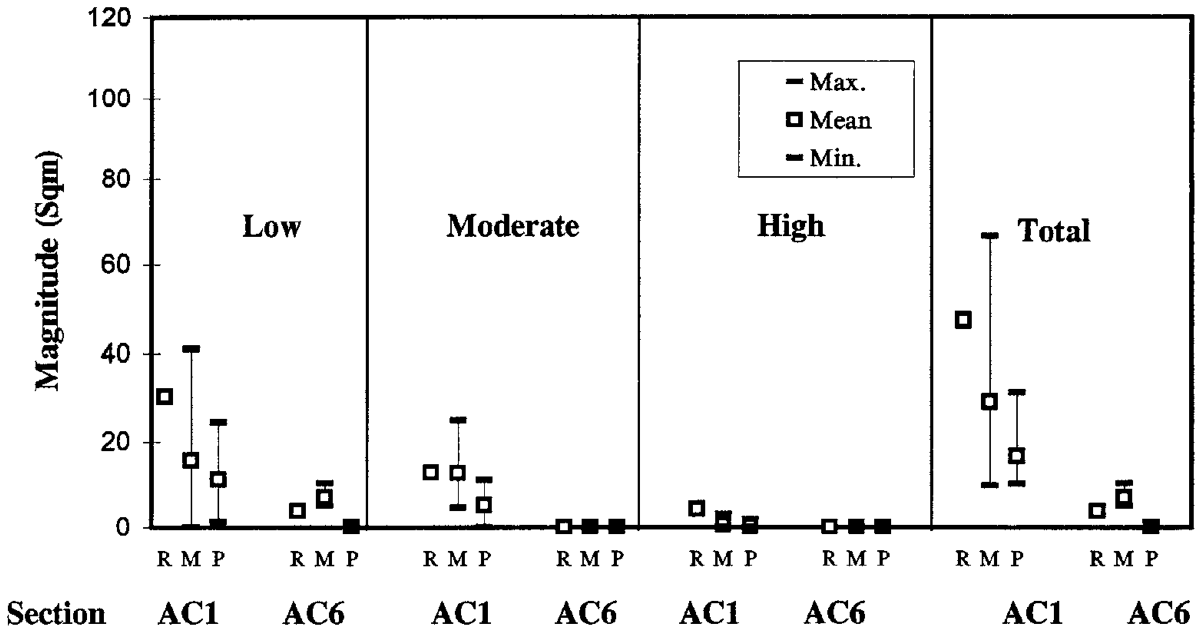


Figure 43. Fatigue Cracking (Sq. Meters) - AC Pavements: Reference, Minimum, Mean, and Maximum Values of Individual Surveys for Manual and PASCO/PADIAS.

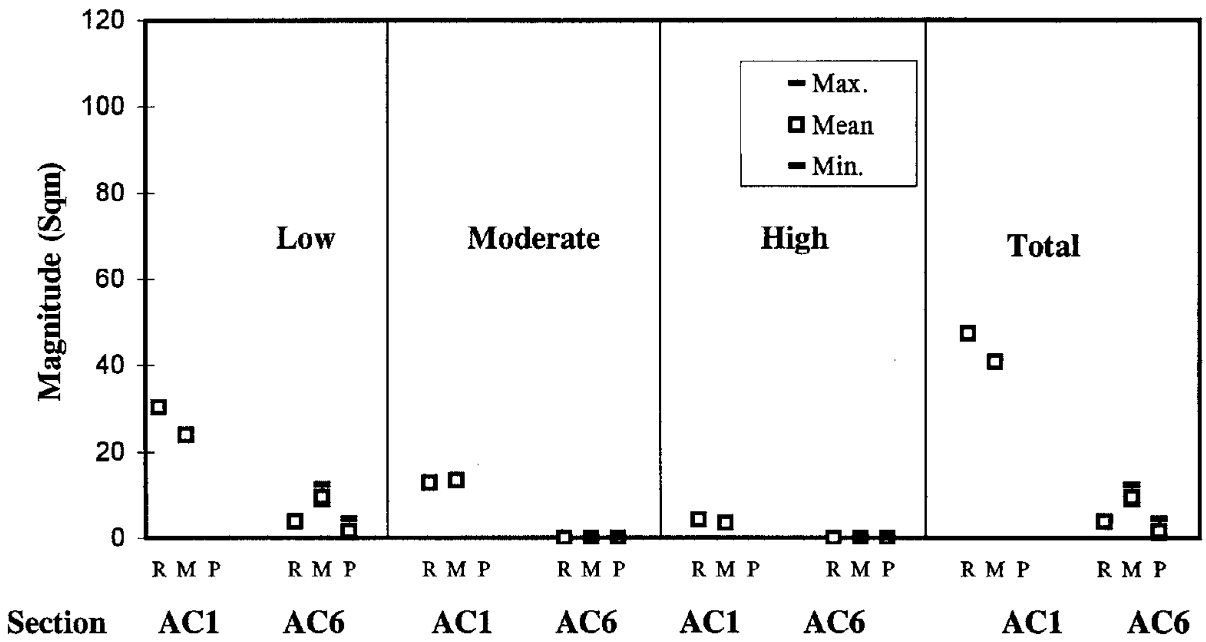


Figure 44. Fatigue Cracking (Sq. Meters) - AC Pavements: Reference, Minimum, Mean, and Maximum Values of Team Surveys for Manual and PASCO/PADIAS.

presents the same results from team ratings. Similar plots for corner breaks on PCC pavements are shown in figures 45 and 46 for individual and team ratings, respectively. A complete set of these plots is presented in appendix C of this report.

The following observations are made for AC pavement distresses:

- Variability associated with field manual distress ratings were found to be both higher and lower than photographic interpretations across distress types, i.e., no definite trends were identified.

- There was a slight trend for mean values from the larger pool of field raters to be closer to the reference values, i.e., lower apparent bias, than the mean from PASCO/PADIAS film interpreters.

- Team variability was lower than individual raters.

The following observations are made for PCC pavement distress:

- There was more variation in the PASCO/PADIAS interpreted ratings than the manual field ratings. Some of this variation appears to be associated with single outlier observations.

- On average, field-determined values are approximately the same as photographic-determined values, with some differences observed for different distresses.

- Overall, team variability is equivalent to individual raters.

The comparative apparent bias statistics for the two rating methods with individual raters are presented in tables 28 and 29. These statistics included manual and PADIAS bias, which is the difference between the mean value and field-determined reference value, and the relative difference between the two biases, which is calculated by subtracting the absolute value of the manual bias from the absolute value of PADIAS bias.

Further, a term "average percent difference" was used to compare the two sets of bias values, also shown in these two tables. For each pavement distress type, the average percent difference is calculated as the average difference of bias divided by the average reference value, expressed as a percentage. The average value for each distress type was computed across the different severity levels for the two pavement sections, except for total distress. The relative difference at total distress level was not used in computing the average percent difference since the total distress is simply a summation of the individual distress levels. For example, for fatigue cracking, the average relative difference of low, moderate, and high severity levels for the two pavement sections is 2.1 m^2, with an average reference value of 8.5 m^2. The average percent difference for fatigue cracking could be computed as 2.1/8.5*100 = 25 percent. The calculated average percent differences varied from -3 percent for transverse cracking (m) to 95

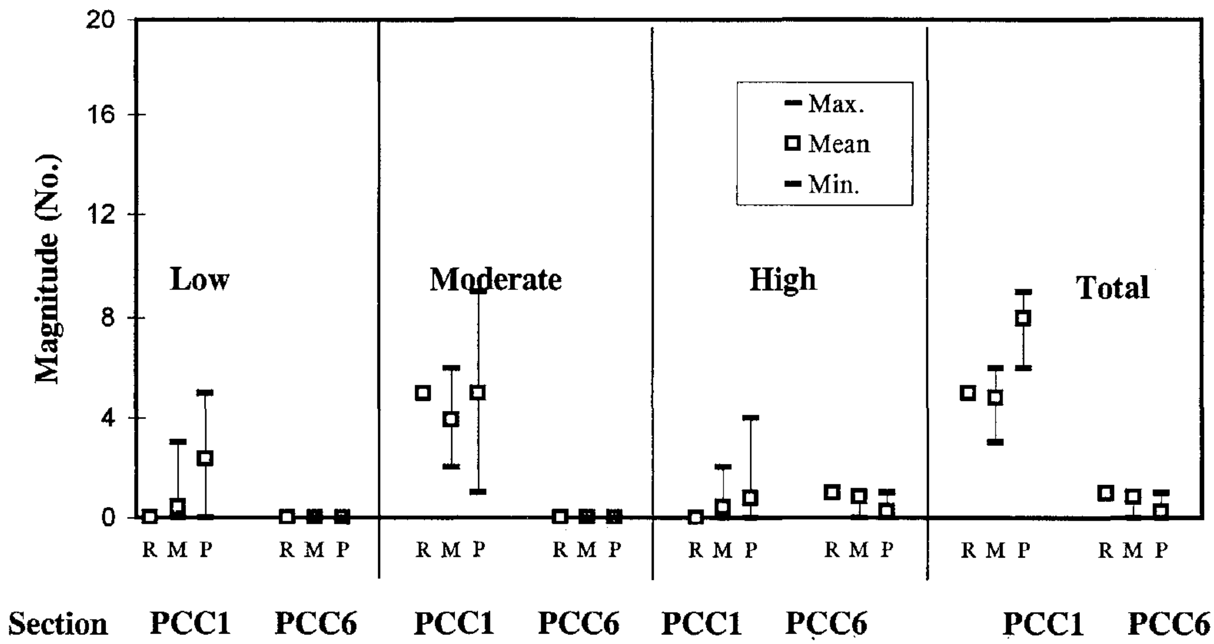


Figure 45. Corner Breaks (No.) - PCC Pavements: Reference, Minimum, Mean, and Maximum Values of Individual Surveys for Manual and PASCO/PADIAS.

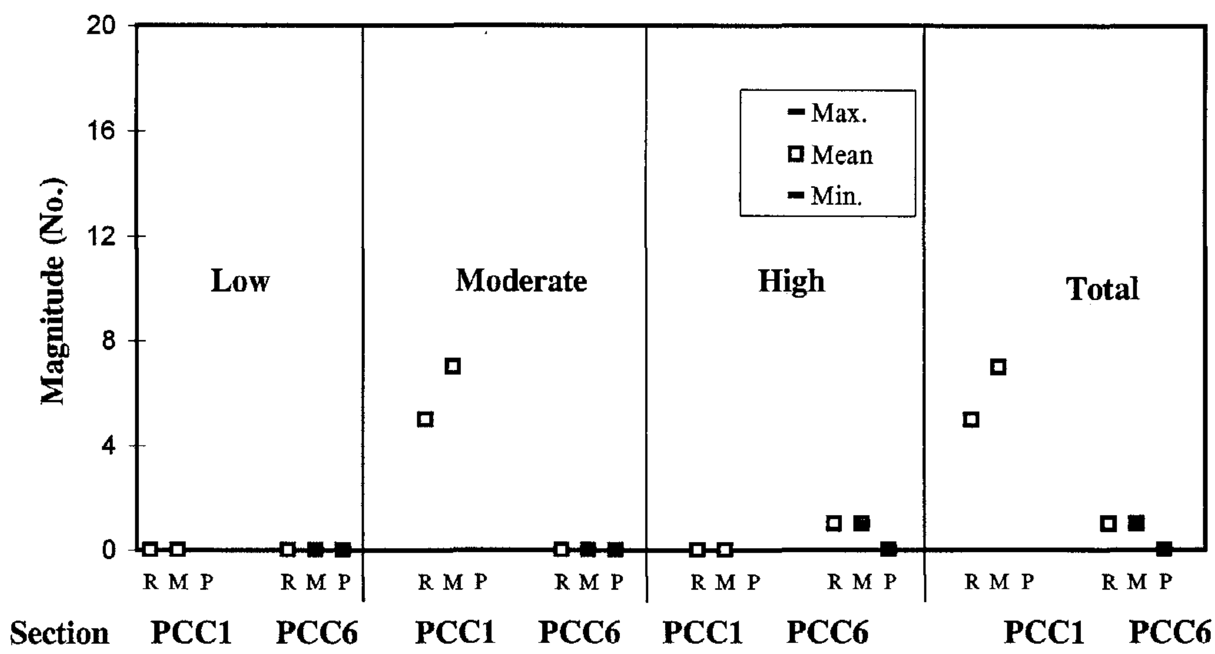


Figure 46. Corner Breaks (No.) - PCC Pavements: Reference, Minimum, Mean, and Maximum Values of Team Surveys for Manual and PASCO/PADIAS

Table 28. Apparent Bias: Manual Vs. PASCO/PADIAS, AC Pavements, Individual Raters.

Distress Type	Section ID	Unit	Severity Level	Ref.	Mean Manual	Bias Manual	Bias PADIAS	Relative Diff.	Avg. %Diff.
Fatigue Cracking	AC1	Sq. Meters	Low	30.3	15.7	-14.6	-19.0	4.4	
			Mod.	12.8	12.8	0.0	-7.6	7.6	
			High	4.3	0.5	-3.8	-4.0	0.2	
			Total	47.4	29.0	-18.4	-30.7	12.3	
	AC6	Sq. Meters	Low	3.8	7.0	3.2	-3.8	0.6	
			Mod.	0.0	0.0	0.0	0.0	0.0	
			High	0.0	0.0	0.0	0.0	0.0	
			Total	3.8	7.0	3.2	-3.8	0.6	25
Longitudinal Cracking - WP	AC1	Meters	Low	3.3	10.6	7.3	24.1	16.8	
			Mod.	15.3	18.4	3.1	4.9	1.8	
			High	3.3	4.4	1.1	4.2	3.1	
			Total	21.9	33.5	11.6	33.2	21.6	
	AC6	Meters	Low	2.0	1.2	-0.8	-1.8	1.0	
			Mod.	0.0	0.0	0.0	0.0	0.0	
			High	0.0	0.0	0.0	0.0	0.0	
			Total	2.0	1.2	-0.8	-1.8	1.0	95
Longitudinal Cracking - NWP	AC1	Meters	Low	3.6	7.9	4.3	25.9	21.6	
			Mod.	17.0	40.4	23.4	11.6	-11.8	
			High	37.6	9.6	-28.0	-27.2	-0.8	
			Total	58.2	58.0	-0.3	10.3	10.1	
	AC6	Meters	Low	108.7	100.2	-8.5	-11.5	3.0	
			Mod.	163.0	129.0	-34.0	-70.3	36.3	
			High	49.0	60.7	11.7	-40.1	28.5	
			Total	320.7	289.9	-30.8	-121.9	91.1	20
Transverse Cracking	AC1	No.	Low	2.0	6.3	4.3	7.2	2.9	
			Mod.	10.0	13.1	3.1	2.0	-1.1	
			High	14.0	7.9	-6.1	-2.2	-3.9	
			Total	26.0	27.4	1.4	7.0	5.6	
	AC6	No.	Low	15.0	12.8	-2.2	-2.5	0.3	
			Mod.	1.0	1.0	0.0	0.8	0.8	
			High	0.0	0.0	0.0	0.0	0.0	
			Total	16.0	13.8	-2.2	-1.7	-0.5	-2
Transverse Cracking	AC1	Meters	Low	1.5	7.0	5.5	9.0	3.5	
			Mod.	22.1	30.7	8.6	4.1	-4.5	
			High	36.3	20.6	-15.7	0.7	-15.0	
			Total	59.9	58.3	-1.6	13.8	12.2	
	AC6	Meters	Low	27.6	26.5	-1.1	11.2	10.1	
			Mod.	3.7	3.7	0.0	3.0	3.0	
			High	0.0	0.0	0.0	0.0	0.0	
			Total	31.3	30.2	-1.1	14.3	13.2	-3

Notes:

relative difference = ABS(PADIAS bias) - ABS(manual bias)

avg. %diff = average (relative difference for low, moderate, and high severity)/
average (reference for low, moderate, and high severity) * 100

Table 29. Apparent Bias: Manual Vs. PASCO/PADIAS, PCC Pavements, Individual Raters.

Distress Type	Section ID	Unit	Severity Level	Ref.	Mean	Bias Manual	Bias PADIAS	Relative Diff.	Avg. %Diff.
Corner Breaks	PCC1	Number	Low	0.0	0.4	0.4	2.3	1.9	
			Mod.	5.0	3.9	-1.1	0.0	-1.1	
			High	0.0	0.4	0.4	0.8	0.4	
			Total	5.0	4.8	-0.2	3.0	2.8	
	PCC6	Number	Low	0.0	0.0	0.0	0.0	0.0	
			Mod.	0.0	0.0	0.0	0.0	0.0	
			High	1.0	0.8	-0.2	-0.8	0.6	
			Total	1.0	0.8	-0.2	-0.8	0.6	30
Longitudinal Cracking	PCC1	Meters	Low	3.8	3.3	-0.5	1.4	0.9	
			Mod.	8.3	7.7	-0.6	4.8	4.2	
			High	0.0	0.3	0.3	3.3	3.0	
			Total	12.1	11.3	-0.8	9.0	8.1	
	PCC6	Meters	Low	0.0	0.0	0.0	0.6	0.6	
			Mod.	0.0	0.0	0.0	0.0	0.0	
			High	0.0	0.0	0.0	0.0	0.0	
			Total	0.0	0.0	0.0	0.6	0.6	72
Transverse Cracking	PCC1	Meters	Low	3.0	6.3	3.3	5.9	2.6	
			Mod.	9.1	7.6	-1.5	2.3	0.8	
			High	0.0	0.0	0.0	1.6	1.6	
			Total	12.1	13.9	1.8	10.2	8.4	
	PCC6	Meters	Low	9.6	17.1	7.5	48.8	41.2	
			Mod.	39.3	27.8	-11.5	-3.7	-7.8	
			High	25.2	26.2	1.0	-9.7	8.7	
			Total	74.1	71.1	-3.0	35.4	32.4	55
Transverse Cracking	PCC1	No.	Low	1.0	3.0	2.0	3.8	1.8	
			Mod.	3.0	2.4	-0.6	0.3	-0.3	
			High	0.0	0.0	0.0	0.4	0.4	
			Total	4.0	5.4	1.4	4.7	3.3	
	PCC6	No.	Low	5.0	6.5	1.5	19.0	17.5	
			Mod.	12.0	8.7	-3.3	-2.8	-0.6	
			High	7.0	7.7	0.7	-2.8	2.1	
			Total	24.0	22.8	-1.2	13.5	12.3	75

Table 29. Apparent Bias: Manual Vs. PASCO/PADIAS, PCC Pavements, Individual Raters (Continued).

Distress Type	Section ID	Unit	Severity Level	Ref.	Mean	Bias Manual	Bias PADIAS	Relative Diff.	Avg. %Diff.
Spalling of Longitudinal Joints	PCC1	Meters	Low	9.0	4.8	-4.2	14.2	9.9	
			Mod.	0.2	1.0	0.8	7.0	6.2	
			High	0.1	0.1	0.0	-0.1	0.1	
			Total	9.3	5.9	-3.4	21.0	17.6	
	PCC6	Meters	Low	7.4	9.0	1.6	0.4	-1.2	
			Mod.	6.2	3.5	-2.8	-3.7	0.9	
			High	0.0	4.0	4.0	0.0	-4.0	
			Total	13.6	16.5	2.9	-3.2	0.3	52
Spalling of Transverse Joints	PCC1	Meters	Low	0.6	0.5	-0.1	3.0	2.9	
			Mod.	1.0	0.8	-0.2	0.8	0.6	
			High	0.0	0.2	0.2	0.7	0.5	
			Total	1.6	1.5	-0.1	4.3	4.2	
	PCC6	Meters	Low	0.3	1.5	1.2	-0.3	-0.9	
			Mod.	0.0	0.6	0.6	0.4	-0.1	
			High	1.7	3.7	2.0	1.9	-0.2	
			Total	2.0	5.8	3.8	2.0	-1.8	79
Spalling of Transverse Joints	PCC1	No.	Low	1.0	1.7	0.7	4.2	3.5	
			Mod.	1.0	0.9	-0.1	0.3	0.3	
			High	0.0	0.3	0.3	0.4	0.1	
			Total	2.0	2.9	0.9	4.8	3.9	
	PCC6	No.	Low	1.0	2.3	1.3	-1.0	-0.3	
			Mod.	0.0	0.8	0.8	0.8	-0.1	
			High	2.0	1.3	-0.7	0.0	-0.7	
			Total	3.0	4.5	1.5	-0.3	-1.3	55

Notes:

relative difference = ABS(PADIAS bias) - ABS(manual bias)
avg. %diff = average (relative difference for low, moderate, and high severity)/
average (reference for low, moderate, and high severity) * 100

percent for longitudinal cracking in the wheel path. The differences in bias were less for transverse cracking than for other types of distresses.

Similarly, apparent bias of manual distress data was generally less than that for distress data obtained using the PASCO/PADIAS method. The average percent differences, in terms of mean reference values, ranged between 30 percent for corner breaks and 79 percent for spalling of longitudinal joints. In general, the average percent differences were greater than those found in AC pavement sections.

4.4 Field Versus Photographic Interpretation by Reference Raters

The ratings from the reference rater group, composed of experienced raters and distress accreditation course trainers, performed in the field and from photographs, were compared. This comparison removes between rater variation from the comparison so that differences between methods can be judged more clearly. The ratings performed by the reference team were done using a team consensus procedure. Figures 47, 48, 49, and 50 present example results of the two rating procedures on six AC and six PCC pavement sections. In these figures, the field ratings are referred to as manual values and PADIAS values are those from the photographic interpretation. A complete set of plots is presented in appendix C of this report.

There was a slight tendency for the field values to be greater than the photographic-interpreted values; however, on some sections the reverse was observed. For AC pavements, field values matched reasonably well with film values for transverse cracking and patching-related distresses. For PCC pavements, the field values were reasonably close to film values for cracking-related distresses, but some significant differences were observed for other distress types.

To further investigate the relationships between reference and consensus values, the reference values were regressed against the consensus values using a straight-line equation. Examples of the resulting relationships are shown in figures 51 and 52. These results confirm that for most distresses the field-determined values tended to be slightly greater than the photographic values, as evidenced by slopes less than one. For many distresses, the intercept was not significantly different from zero, indicating that, on average, good agreement was obtained at low distress levels. The most significant discrepancies between the two methods were observed for PCC spalling and corner breaks. Very little transverse crack spalling and corner breaks were rated from the film when significant amounts were rated in the field. For longitudinal joint spalling, very large amounts of spalling were rated from the film, when none were rated in the field, which caused the regression line to have a negative slope. The R^2 values for AC pavements ranged from a low of 0.36 for longitudinal cracking in the wheel path to a high of 0.92 for non-wheel-path longitudinal cracking. The R^2 values for longitudinal and transverse cracking PCC distresses ranged from 0.78 to 0.96; however, for corner breaks and spalling, R^2 less than 0.1 was obtained.

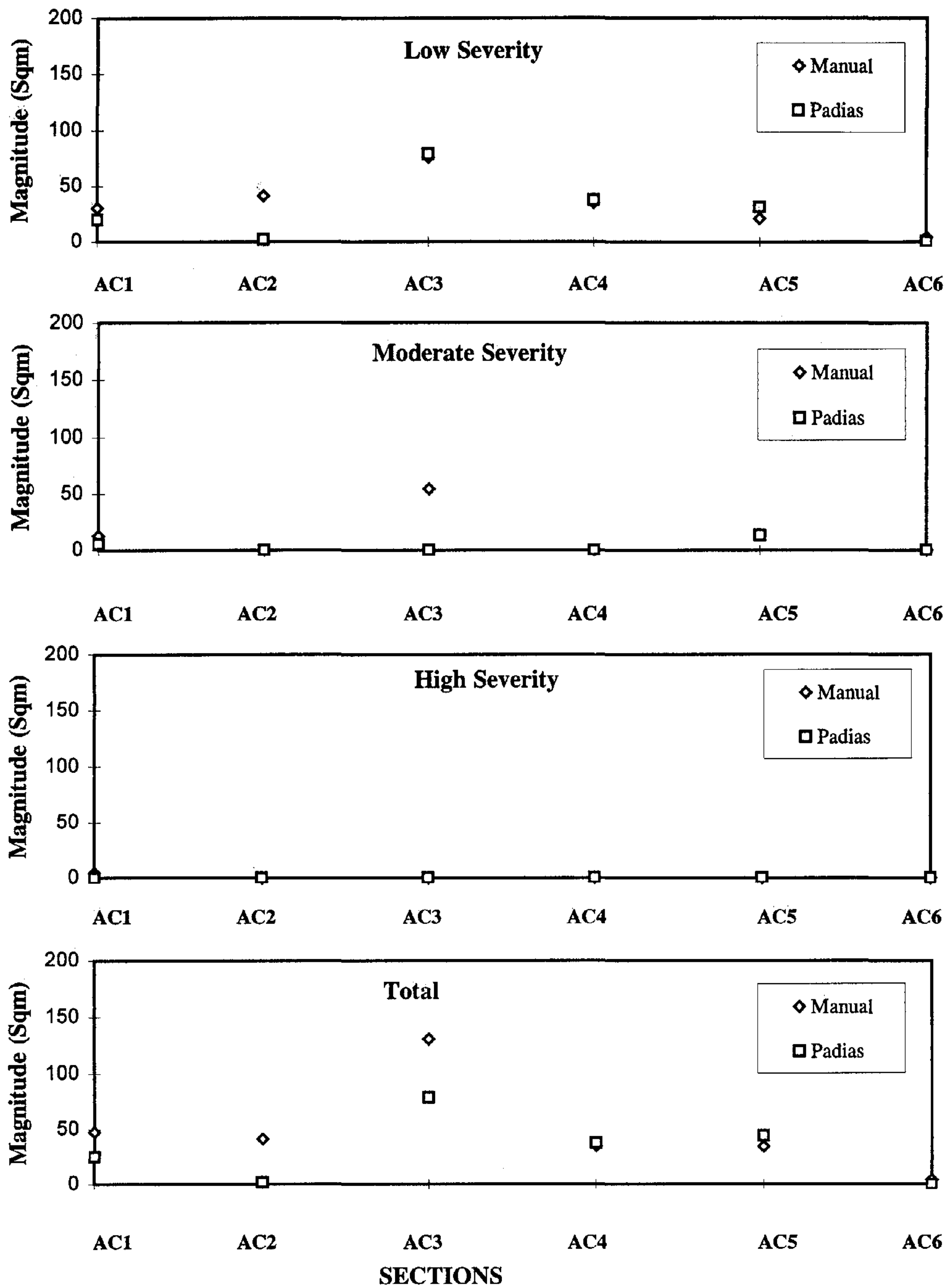


Figure 47. Fatigue Cracking (Sq. Meters) - AC Pavements: Field and PASCO/PADIAS Ratings by Reference Group.

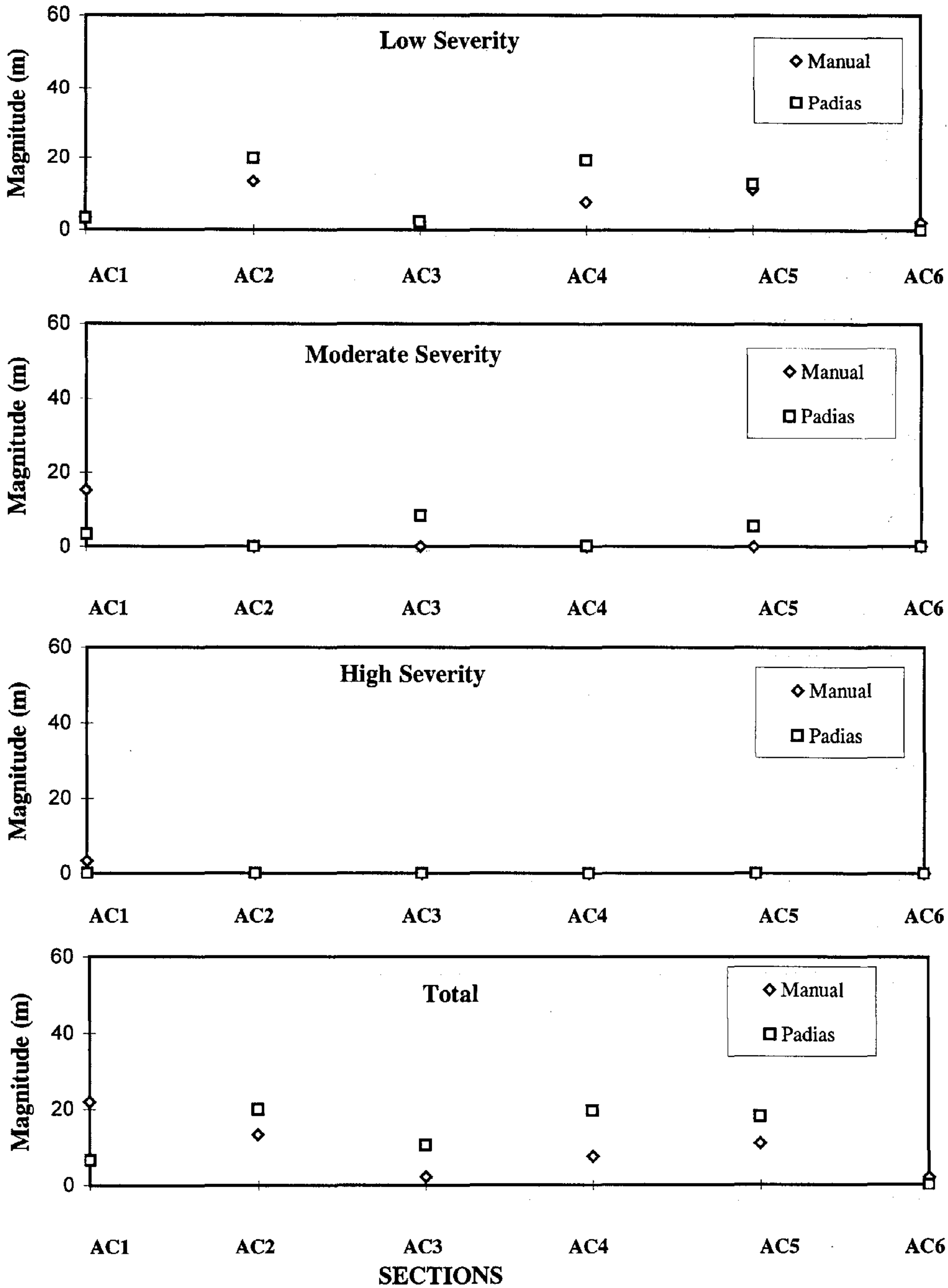


Figure 48. Longitudinal Cracking WP (Meters) - AC Pavements: Field and PASCO/PADIAS Ratings by Reference Group.

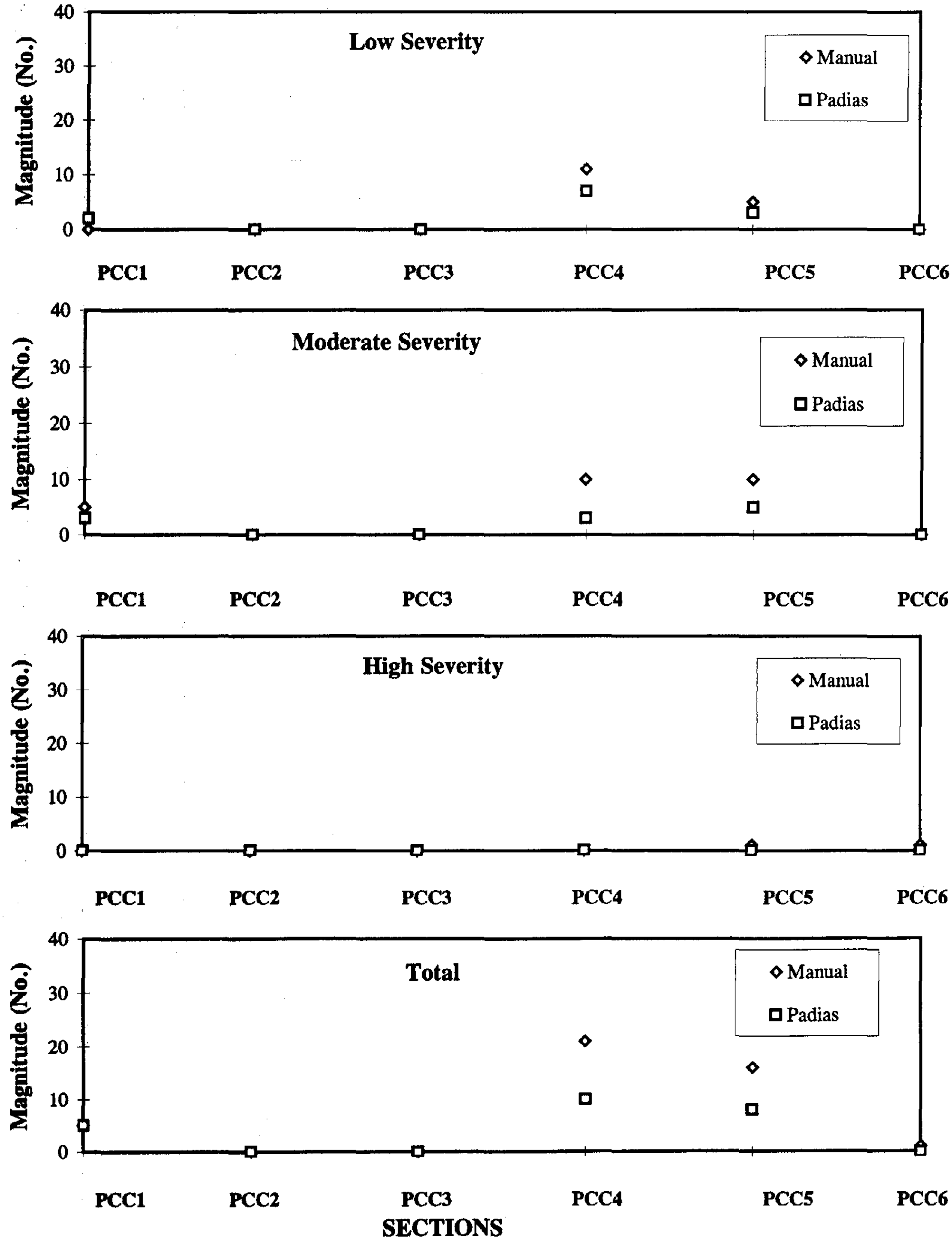


Figure 49. Corner Breaks (No.) - PCC Pavements: Field and PASCO/PADIAS Ratings by Reference Group.

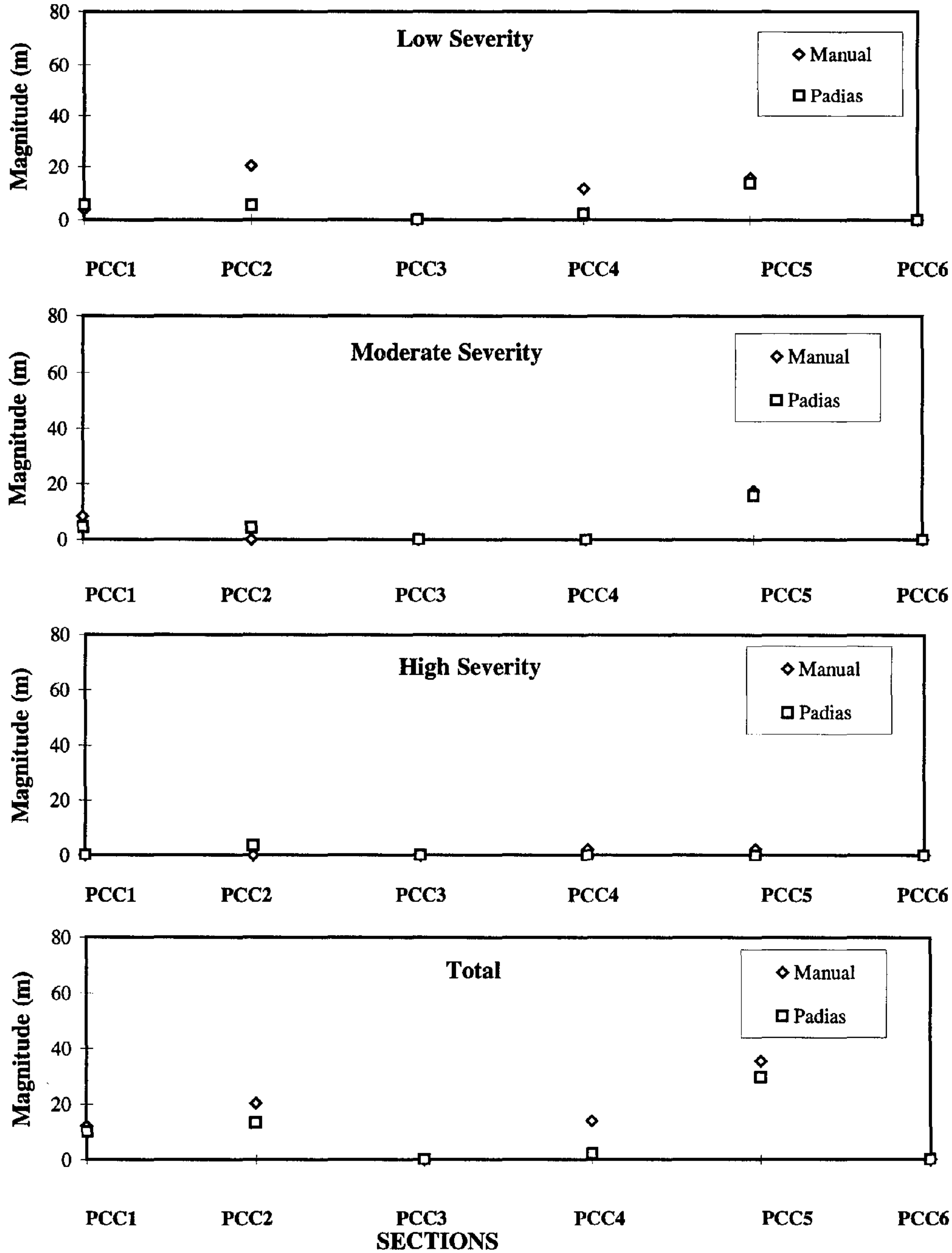


Figure 50. Longitudinal Cracking (Meters) - PCC Pavements: Field and PASCO/PADIAS Ratings by Reference Group.

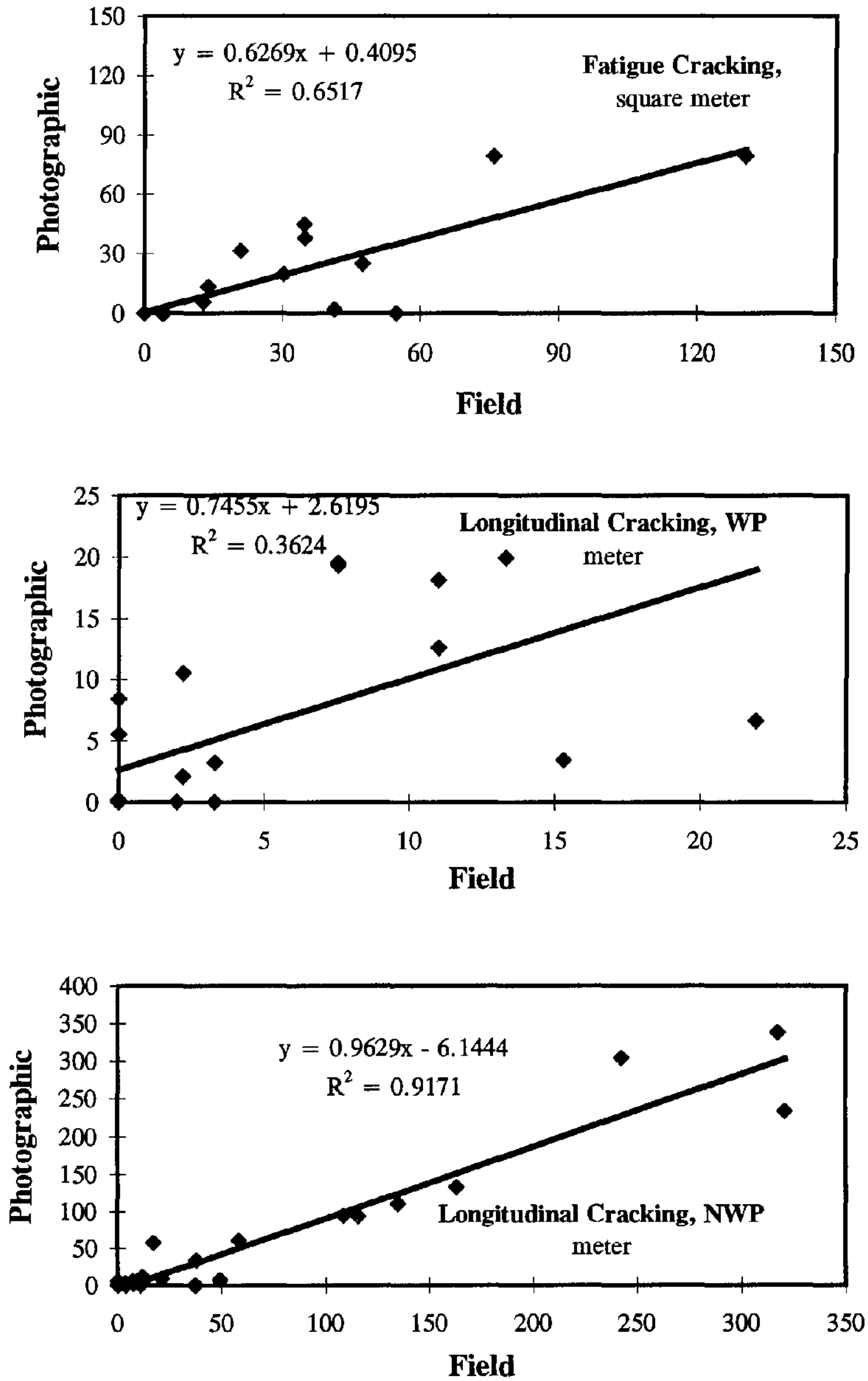
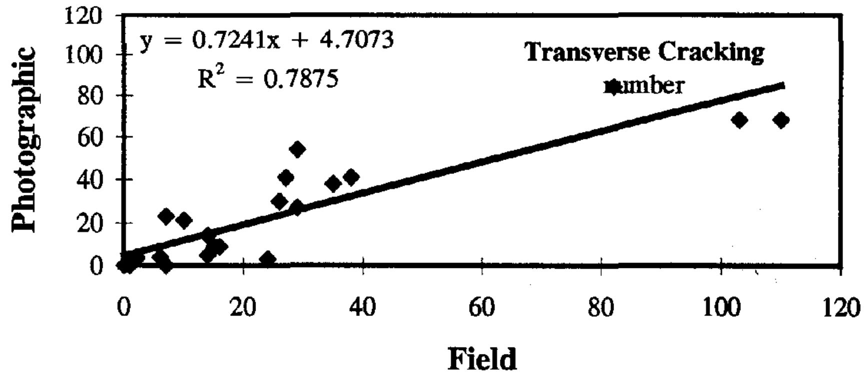


Figure 51. Field Vs. Photographic Ratings by Reference Group, AC Pavements.



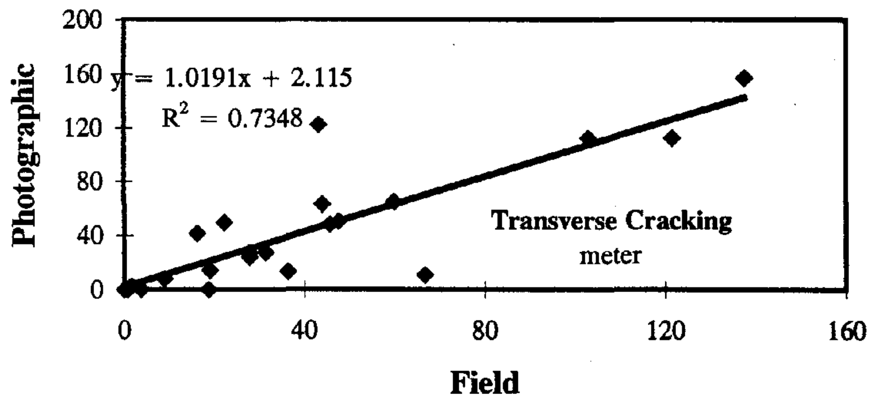


Figure 51. Field Vs. Photographic Ratings by Reference Group, AC Pavements (Continued).

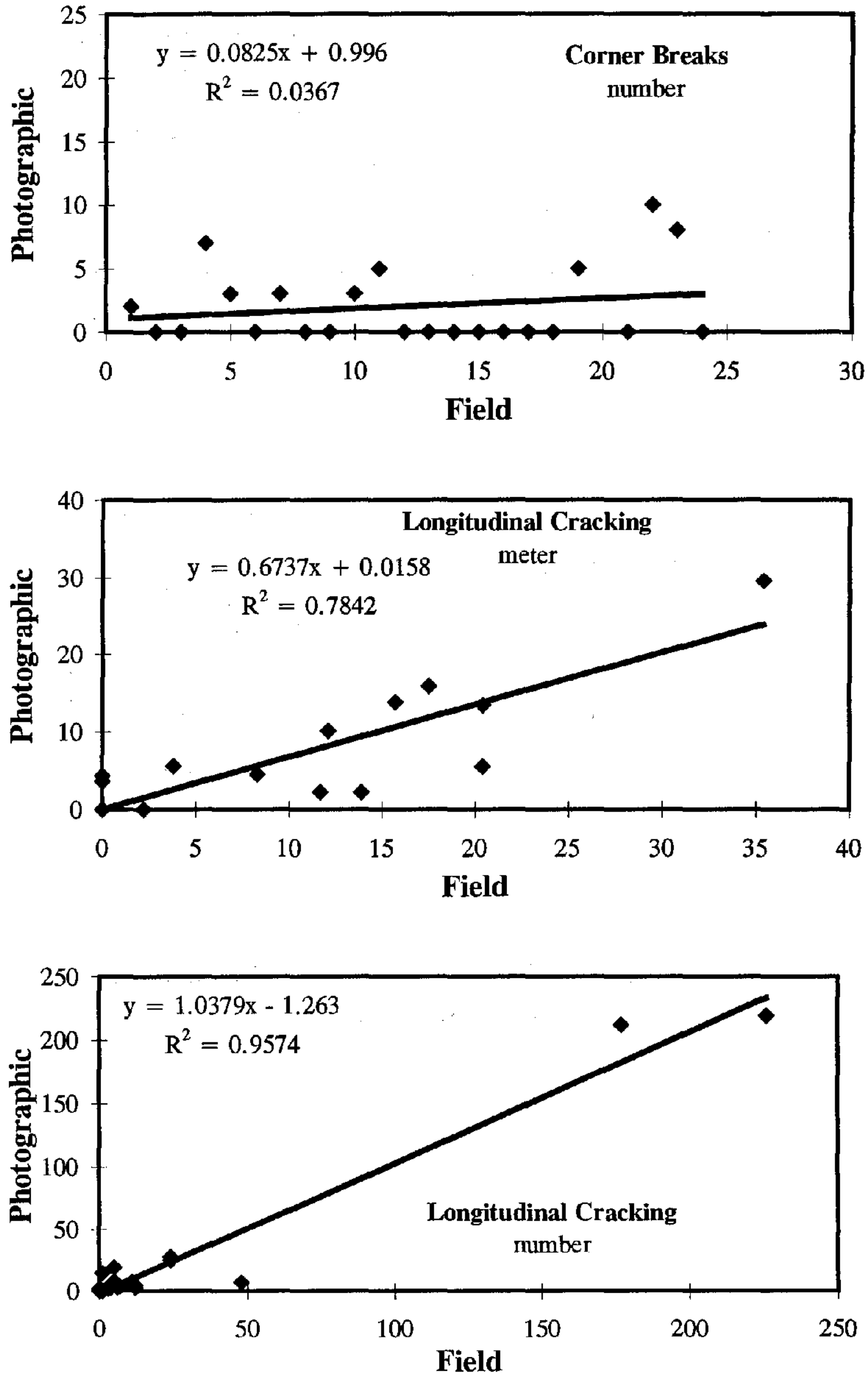


Figure 52. Field Vs. Photographic Ratings by Reference Group, PCC Pavements.

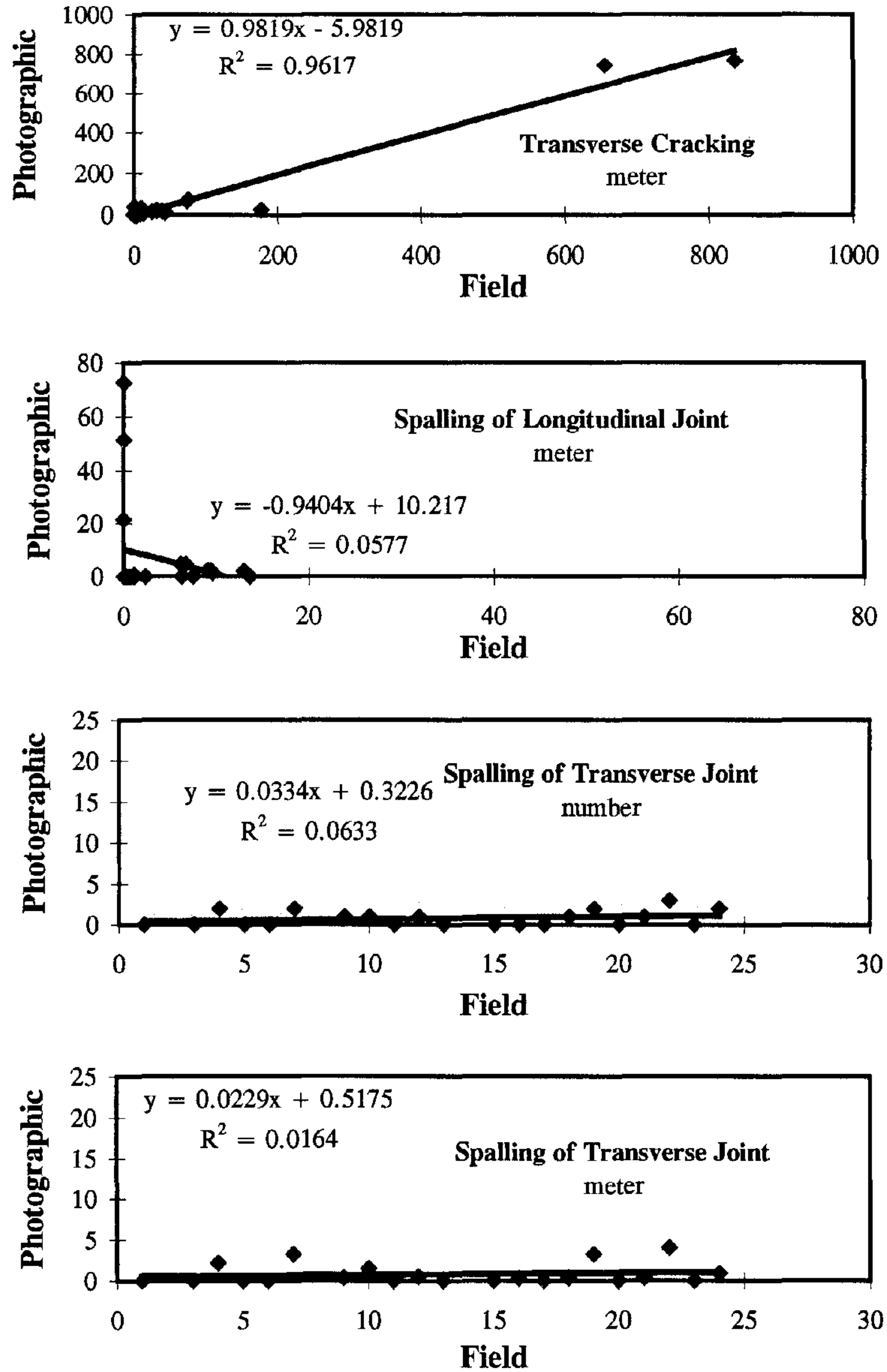


Figure 52. Field Vs. Photographic Ratings by Reference Group, PCC Pavements (Continued).

4.5 Summary and Conclusions

The variability of distress data collected by manual distress surveys and those by PASCO/PADIAS were evaluated in this study. From the analysis, the following conclusions were drawn:

- For both AC and PCC pavements, the overall variability of manual distress data is lower than that for distress data derived using the PASCO/PADIAS method.

- From comparison of the distress data obtained from the four common pavement sections subjected to both manual and PASCO/PADIAS surveys, the apparent bias appeared to be much higher for distress data derived from PASCO/PADIAS than that obtained by manual surveys. The average percent differences were between 5 and 95 percent for AC pavements and from 43 to 74 percent for PCC pavements.

- Reasonable correlation was found between field and film values from the reference rater groups for most of the AC and PCC pavement distresses. The general trend of field-determined values being higher than photographic values was observed. For corner breaks and spalling on PCC pavements, very poor relationships were found.

5. CONCLUSIONS AND RECOMMENDATIONS

Reliable distress data for pavement performance model development and validation, and other pavement engineering products, are critical to the success of the LTPP program. Proper use of distress data in pavement performance analysis requires an understanding of its variability and potential error due to bias and precision effects. In this study, measures of bias and precision were evaluated from measurements on test sections used in the LTPP distress rater accreditation workshops conducted during the period from 1992 to 1996. The test data set included ratings performed in the field and from black and white photographs by individual raters, two-person rater teams, and the reference "expert" group.

5.1 Conclusions

Overall the authors conclude that the current level of variability in the distress ratings from individuals performing field distress surveys on LTPP test sections is unacceptably high and efforts should be pursued to reduce this variability. These conclusions are based on the evidence provided from the data obtained from the LTPP distress accreditation workshops, which may not be representative of variability of the ratings performed on the LTPP test sections. The concern is not the differences between the average of the individual ratings and the reference value, called the apparent bias in this study, but the range of ratings obtained from individual raters, since this reflects the likely variability in the rating performed on LTPP test sections. Further, it is suspected that the discrepancies that have been observed in the distress time histories on LTPP test sections may result from this high variability.[1]

Some of the more specific findings from this work, previously discussed in this report, are summarized below.

Manual Distress Data

- Individual rater variability for any given distress type-severity level combination is typically large, and increases in absolute value as the distress quantity increases, but decreases in proportion to the mean value as measured by the coefficient of variation (CV). The variability or precision is very sensitive to the magnitude and range of distress quantities present on a section. The CV ranges from less than 10 percent to well in excess of 100 percent, although it generally decreases with increased distress quantities. Large CV values were observed only where low quantities of a given pavement distress were present. Thus, the large variability indicated by these values may not necessarily indicate poor precision.

- The variability and bias for total distress quantities (the summation of distress quantities across all severity levels) was much lower than for individual severity levels of the same distress. Thus, even when there was good agreement between raters on the type of distress, there were significant differences between severity levels.

- The apparent bias, the difference between the rating group average and reference value, is relatively small with no uniform tendency to be consistently greater than or less than the reference value. For total distress, the percentage of apparent bias to pooled references varies from -6 to 16 percent for AC pavement sections and from -7 to 19 percent for PCC pavement sections.

- When multiple distress type-severity level combinations are expressed in terms of a single composite index, such as the Pavement Condition Index (PCI) value used in this report, there is excellent agreement (low bias) between individual raters, rater's group mean, and reference values. Individual rater precision (variability) is also very small when expressed in terms of this composite value. This result is as expected because the weighting functions in the PCI calculation greatly reduce variability in the ratings of individual distress types. These indices are most sensitive to very high levels of distress, which approach conditions requiring corrective action.

- There was no strong evidence to suggest that rater variability was improved (decreased) from attendance at more than one accreditation workshop.

- The limited study conducted to assess the potential improvement in distress data bias and precision through the use of two-person consensus surveys resulted in a slight but inconsistent tendency for the team values to be better than those from individual raters.

PASCO/PADIAS Distress Data

- From the repeat film interpretations of the same test sections by the same groups of raters, the repeatability of ratings by the same individual or rating group is small and reasonable. The within rater or rater team variability is also much smaller than the between rater variability. The repeatability of the distress data obtained by two-person teams is smaller than that interpreted by individual experts and individual raters.

- The variance within a given rater group – experts, individual raters, and two-person teams – is statistically large. Furthermore, this variance tends to increase as distress quantity increases.

- Statistical comparison of group means (experts versus individual raters and experts versus two-person teams) indicates no **statistically** significant differences, even when some of the differences are judged to be significant from an engineering perspective. This lack of statistical significance is, to a large extent, due to the high within group variance, which reduces the sensitivity of the statistical test.

- In contrast to the manual surveys, the variability for total distress quantities was not smaller than that for individual severity levels. Compensatory differences for a number of closely related distresses were observed.

- The apparent bias, the difference between the rater's group average and the reference value, was large for most cases, with no uniform tendencies to be greater or less than the reference values. For total distress of AC pavement sections, the percentage of bias to pooled reference ranges between -52 and 182 percent, between -50 and 507 percent, and between -77 and 391 percent for experts, teams, and individual raters. For cracking-related distresses on PCC pavement, expert and team ratings showed acceptably low levels of apparent bias, with percentage of bias to pooled reference generally less than 20 percent. The variability in the ratings by experts and teams for both AC and PCC distress was also relatively small. It appeared that the very large variability between individual raters was due to singular outlier ratings, possibly related to rater experience and training.

- Although some differences in data interpreted with the PADIAS version 1 and version 4 systems exist, overall there is excellent agreement between the two systems for all pavement types, especially for total distress quantities and cracking-related distresses. The largest difference between the two systems was found for surface and joint-related defects.

Manual Vs. PASCO/PADIAS

- For both AC and PCC pavements, the overall variability of manual distress data is lower than that for distress data derived using the PASCO/PADIAS method.

- From comparison of the distress data obtained from the four common pavement sections subjected to both manual and PASCO/PADIAS surveys, the apparent bias appeared to be much higher for distress data derived from PASCO/PADIAS than that obtained by manual survey. The average percent differences were between -3 and 95 percent for AC pavements and from 30 to 79 percent for PCC pavements.

- Reasonable correlation was found between field and film values from the reference rater groups for most of the AC and PCC pavement distresses. The general trend of field-determined values being higher than photographic values was observed. For corner breaks and spalling on PCC pavements, very poor relationships were found.

5.2 Recommendations

Based on the overall conclusions that the variability between distress raters is high, the following steps are recommended to help reduce the variability and hence improve the utility of the distress data:

- **Reduce the number of raters to a smaller pool.** The LTPP regions routinely send individuals to the distress accreditation workshops so they have enough raters to serve as backups and meet operational scheduling needs. Many of these individuals do not perform distress ratings on a routine basis. Because some aspects of distress ratings are still subjective, this larger pool of raters may contribute to large variability in the workshop data.

- **Tighten rater accreditation acceptance criteria.** Of all of the LTPP distress rater accreditation workshops held, only three raters were either not accredited or had some restriction placed on the type of pavements they were allowed to rate. Restricting the acceptance range for accreditation could help improve variability.

- **Add a rating frequency requirement for re-certification.** The old adage of "use it or lose it" certainly applies to distress rating. The LTPP method, being research based, is very detailed, and consistency requires practice. Adding a rating frequency requirement could help limit the pool of accredited raters to a smaller number.

- **Conduct interim consistency checks.** LTPP currently has no uniform method to check on the consistency of the distress ratings on in-service LTPP test sections. Review of regional distress rating practices are performed; however, comparison of the results of distress ratings to reference values on actual test sections is not performed. This type of quality assurance process could help to better define bias and precision components in the LTPP distress data and promote greater uniformity in distress ratings between distress rater accreditation workshops.

- **Use consensus survey teams.** Since consensus surveys from ratings teams are used to establish the reference values for the distress rater accreditation workshops, it is a natural extension of this logic that use of distress raters teams will also help to improve the variability in the distress ratings performed on LTPP test sections. From the limited data set examined in this study, rating teams did result in lower variability on the film-based interpretations, but no definite trend was observed for field interpretations.

In assessing variability in distress ratings, the use of CV is not recommended as a good measure of precision or variability. It appears that relative measures of variability, in which the standard deviation is expressed as a percentage of the maximum amount of distress that can exist on a pavement section, may be more suitable. For those distresses with no easily defined maximum value (such as linear extent of transverse cracking), a relatively high value may be used in place of the maximum value.

The authors believe that for research purposes, target levels of variability in distress ratings of 10 percent are desirable so that 90 percentile confidence levels are less than 30 percent. The problem with this statement is the basis for the percentage calculation. As shown in this work, low distress levels create very large ratios when expressed in terms of the mean value. Sometimes this variability is not significant in absolute terms relative to the high amounts of distress that can occur. More research is needed to assess the impact of indexing these types of variability values to maximum or high distress values or possibly larger "typical" values.

REFERENCES

1. Daleiden, Jerome, Simpson, Amy, and Rauhut, Brent, *Rehabilitation Performance Trends, Early Observations from Long-Term Pavement Performance (LTPP) Specific Pavement Studies (SPS)*, Federal Highway Administration, March 1996.