

Technical Report Documentation Page

1. Report No. FHWA/TX-11/0-6664-2		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Maintenance Test Section Survey: Data Collection and Analysis				5. Report Date May 2011, Published August 2012	
				6. Performing Organization Code	
7. Author(s) Jorge Prozzi, Lu Gao, Seokho Chi, Claudia Centurion, Mike Murphy, Yetkin Yildirim, Zhanmin Zhang				8. Performing Organization Report No. 0-6664-2	
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 1616 Guadalupe St, Suite 4.202 Austin, TX 78701				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No. 0-6664	
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, TX 78763-5080				13. Type of Report and Period Covered Technical Report July 2010–May 2011	
				14. Sponsoring Agency Code	
15. Supplementary Notes Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.					
16. Abstract <p>A Maintenance Test Section Survey (MTSS) was conducted as part of a Peer State Review of the Texas Maintenance Program conducted October 5–7, 2010. The purpose of the MTSS was to conduct a field review of 34 highway test sections and obtain participants' opinions about pavement, roadside, and maintenance conditions. The goal was to cross-reference or benchmark TxDOT's maintenance practices based on practices used by selected peer states. Representatives from six peer states (California, Georgia, Kansas, Missouri, North Carolina, and Washington) were invited to Austin to attend a 3-day Peer State Review of TxDOT Maintenance Practices Workshop and to participate in a field survey of a number of pre-selected one-mile roadway sections. It should be emphasized that the objective of the survey was not to evaluate and grade or score TxDOT's road network but rather to determine whether the selected roadway sections met acceptable standards of service as perceived by Directors of Maintenance or senior maintenance managers from the peer states.</p> <p>The pavement sections were selected such that the sample contained a wide range of conditions including Very Good (like new) to Very Poor (extensive cracking, rutting, and rough ride) and in immediate need of maintenance or rehabilitation. In addition to pavement conditions, the roadside and traffic marking maintenance conditions were evaluated by the researchers when making final section selection. It was also important to sample sections within each facility type; therefore, the sample contained sections from the interstate, national, and state systems as well as numerous Farm-to-Market roads. Two county roads were also included in the sample.</p> <p>The MTSS participants traveled in six vans, and rated the one-mile sections traveling at highway speeds over a four-hour period. The results were recorded by the participants on survey sheets handed in at the end of the Survey. The evaluation was based on a simple 1.0 (Well Below Expectations) to 5.0 (Well Above Expectations) scale for each category. These results of the peer state rating were compared to the Texas Maintenance Assessment Program (TxMAP) annual ratings for these sections and the ratings of other participants, grouped according to their background and experience in highway maintenance assessment.</p>					
17. Key Words Maintenance Management; TxDOT Maintenance Peer Review; Maintenance Assessment Program; Pavement Maintenance Survey; Pavement Maintenance; Roadside Maintenance; Traffic Operations Maintenance; Road Rally; Condition Survey; Condition Rating; TxMAP; TxTAP			18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161; <a href="http://www.ntis.gov">www.ntis.gov</a> .		
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 44		22. Price	





## **Maintenance Test Section Survey: Data Collection and Analysis**

Jorge Prozzi  
Lu Gao  
Seokho Chi  
Claudia Centurion  
Mike Murphy  
Yetkin Yildirim  
Zhanmin Zhang

---

CTR Technical Report:	0-6664-2
Report Date:	May 2011
Project:	0-6664
Project Title:	Peer State Review of TxDOT Maintenance Practices
Sponsoring Agency:	Texas Department of Transportation
Performing Agency:	Center for Transportation Research at The University of Texas at Austin

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.

Center for Transportation Research  
The University of Texas at Austin  
1616 Guadalupe St, Suite 4.202  
Austin, TX 78701

[www.utexas.edu/research/ctr](http://www.utexas.edu/research/ctr)

Copyright (c) 2011  
Center for Transportation Research  
The University of Texas at Austin

All rights reserved  
Printed in the United States of America

## **Disclaimers**

**Author's Disclaimer:** The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation.

**Patent Disclaimer:** There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine manufacture, design or composition of matter, or any new useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

## **Engineering Disclaimer**

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES.

Project Engineer: Michael Murphy  
Professional Engineer License State and Number: Texas No. 59874  
P. E. Designation: Research Supervisor

## **Acknowledgments**

The researchers express special appreciation to the Peer State Participants for their comments and insights during the Maintenance Test Section Survey: Steve Takigawa, P.E., Deputy Director—Maintenance and Operations (CALTRANS); Jennifer Brandenburg, P.E., State Road Maintenance Engineer (North Carolina DOT); David Bierschbach, P.E., Asst. Regional Administrator for Maintenance (Washington State DOT); Jim Carney, P.E., State Maintenance Engineer (Missouri DOT); Roy Rissky, P.E., Bureau Chief Construction and Maintenance (Kansas DOT); and Eric Pitts, P.E., State Maintenance Engineer (Georgia DOT).

The researchers would also like to thank PMC Members for their input and suggestions for the workshop and preparations for and participation in the Maintenance Test Section Survey: Tammy Sims, P.E. (PD); Toribio Garza, P.E. (Director—MNT); Mario Jorge, P.E. (DE Pharr); Dennis Cooley, P.E. (DE Lufkin); Lynn Passmore, P.E. (DE Brownwood); Lonnie Gregorcyk, P.E. (DE Yoakum); Dr. German Claros, P.E. (RTI Research Engineer); and Sylvia Medina (RTI Contracts Specialist).

In addition, the researchers would like to thank the following additional Maintenance Test Section Survey participants: Bob Blackwell (MNT); Jeff Seiders, P.E. (CST); Lowell Choate, P.E. (AUS—MNT Director); James Williams, P.E., (AUS); Mike Arellano, P.E. (AUS); Magdy Mikhail, P.E. (CST); Jamie Witten, P.E. (AUS); Gary Lantrip (AUS); Bonnie Lister (AUS); John Sabala (SPP); Gretchen Stoletje (SPP); and Tina Geiselbrecht, P.E. (TTI).

## Table of Contents

<b>1. Objectives of the Maintenance Test Section Survey .....</b>	<b>1</b>
<b>2. Characteristics of the Maintenance Test Section Survey .....</b>	<b>1</b>
Maintenance Test Section Survey Process .....	4
<b>3. Sampling Groups and Evaluation Principles .....</b>	<b>4</b>
<b>4. Ratings per Individual and per Sampling Group .....</b>	<b>5</b>
<b>5. Ratings per Section .....</b>	<b>10</b>
<b>6. Data of Individual Sections .....</b>	<b>16</b>
<b>7. Comparison with TxMAP Scores .....</b>	<b>26</b>
<b>8. Conclusions and Recommendations.....</b>	<b>30</b>
Recommendations.....	31





## **List of Tables**

Table 3.1: Participants of the Maintenance Test Section Survey .....	5
Table 4.1: Test of Hypotheses Comparing the Various Groups .....	6
Table 5.1: Test of Hypotheses Comparing the Peer Group Overall Ratings .....	15
Table 7.1: Test of Hypotheses Comparing TxMAP scores with Group Overall Ratings .....	30



## List of Figures

Figure 2.1: General Location and Route of the Maintenance Test Section Survey.....	2
Figure 2.2: Evaluation Form.....	3
Figure 4.1: Aggregated Evaluation of All Sections.....	7
Figure 4.2: Pavement Score of All Sections (per individual).....	7
Figure 4.3: Traffic Score of All Sections.....	8
Figure 4.4: Roadside Score of All Sections (per individual).....	8
Figure 4.5: Overall Score of All Sections (per individual).....	9
Figure 4.6: All Scores of All Sections (per individual).....	9
Figure 5.1: Aggregated Rating of All Sections (Overall Score).....	10
Figure 5.2: Peers versus all Groups (Pavements).....	11
Figure 5.3: Peers versus all Groups (Traffic).....	11
Figure 5.4: Peers versus all Groups (Roadside).....	12
Figure 5.5: Peers versus all Groups (Overall).....	12
Figure 5.6: Comparison of Overall Scores of Peers vs. Expert Group.....	13
Figure 5.7: Comparison of Overall Scores of Peers vs. PMC Group.....	14
Figure 5.8: Comparison of Overall Scores of Peers vs. Other Group.....	14
Figure 5.9: Comparison of Overall Scores of Peers vs. TxDOT Group.....	15
Figure 6.1(a): Section 16—a two-lane FM road with no shoulders.....	17
Figure 6.1(b): All Ratings for Section 16.....	17
Figure 6.2(a): Section 26—a two-lane FM road with no shoulders.....	18
Figure 6.2(b): All Ratings for Section 26.....	18
Figure 6.3(a): Section 28—a two-lane FM road with no shoulders.....	19
Figure 6.3(b): All Ratings for Section 28.....	19
Figure 6.4(a): Section 10—a two-lane FM road with no shoulders.....	20
Figure 6.4(b): All Ratings for Section 10.....	20
Figure 6.5(a): Urban FM road.....	21
Figure 6.5(b): All Ratings for Section 03.....	21
Figure 6.6(a): Rural US Highway.....	22
Figure 6.6(b): All Ratings for Section 29.....	22
Figure 6.7(a): Urban Interstate Highway.....	23
Figure 6.7(b): All Ratings for Section 34.....	23
Figure 6.8(a): Urban Interstate Highway.....	24
Figure 6.8(b): All Ratings for Section 32.....	24
Figure 6.9(a): Rural State Highway with Shoulders.....	25
Figure 6.9(b): All Ratings for Section 12.....	25
Figure 6.10: Maximum Differences in Ratings.....	26

Figure 7.1: Peers Evaluation versus TxMAP Ratings .....	27
Figure 7.2: TxDOT Evaluation versus TxMAP Ratings .....	28
Figure 7.3: Experts Evaluation versus TxMAP Ratings .....	28
Figure 7.4: PMC Evaluation versus TxMAP Ratings.....	29
Figure 7.5: Others Evaluation versus TxMAP Ratings .....	29

# **1. Objectives of the Maintenance Test Section Survey**

The objective of the Maintenance Test Section Survey (MTSS) was to cross-reference or benchmark TxDOT's maintenance practices against equivalent practices used by selected peer states. Representatives from six peer states were invited to Austin (Texas) to attend a 3-day workshop, Peer State Review of TxDOT Maintenance Practices, and to participate in a field survey of a number of pre-selected one-mile roadway sections. The peer states were California, Georgia, Kansas, Missouri, North Carolina, and Washington. It should be emphasized that the objective of the survey was not to evaluate and grade or score TxDOT's road network but rather to determine whether the selected roadway sections met acceptable standards of service as perceived by Directors of Maintenance or senior maintenance managers from the peer states.

The sections were not selected using a random sampling scheme because of the time limitation of the field survey. The workshop schedule allowed only half a day to conduct the field survey and a minimum of 30 sections were to be evaluated. Statewide, approximately 87 percent of the TxDOT-maintained pavement network is in 'Good' or 'Very Good' condition based on Pavement Management Information System (PMIS) scores. Therefore, a random sample would have resulted in very few (if any) sections with pavements in 'Fair', 'Poor' or 'Very Poor' condition and would have offered very limited value.

The pavement sections were selected such that the sample contained a wide range of conditions including Very Good (like new) to Very Poor (extensive cracking, rutting and rough ride) and in immediate need of maintenance or rehabilitation. In addition to pavement conditions, the roadside and traffic marking maintenance conditions were evaluated by the researchers when making final section selection. It was also important to sample sections within each facility type; therefore, the sample contained sections from the Interstate (IH), national (US), and state (SH) systems as well as numerous Farm-to-Market (FM) roads. Two county roads were also included in the sample.

Therefore, sections in fair, poor, and very poor conditions were over-sampled in order to have several sections in each condition among the 34 sections sampled. With no prior intention, it was determined that exactly 50 percent of the sections sampled had an Overall Score Average above 3.0 ("Meets Expectations") and 50 percent had an Overall Score Average below 3.0 based on a scale that ranged from 1.0 ("Well Below Expectations") to 5.0 ("Well Above Expectations").

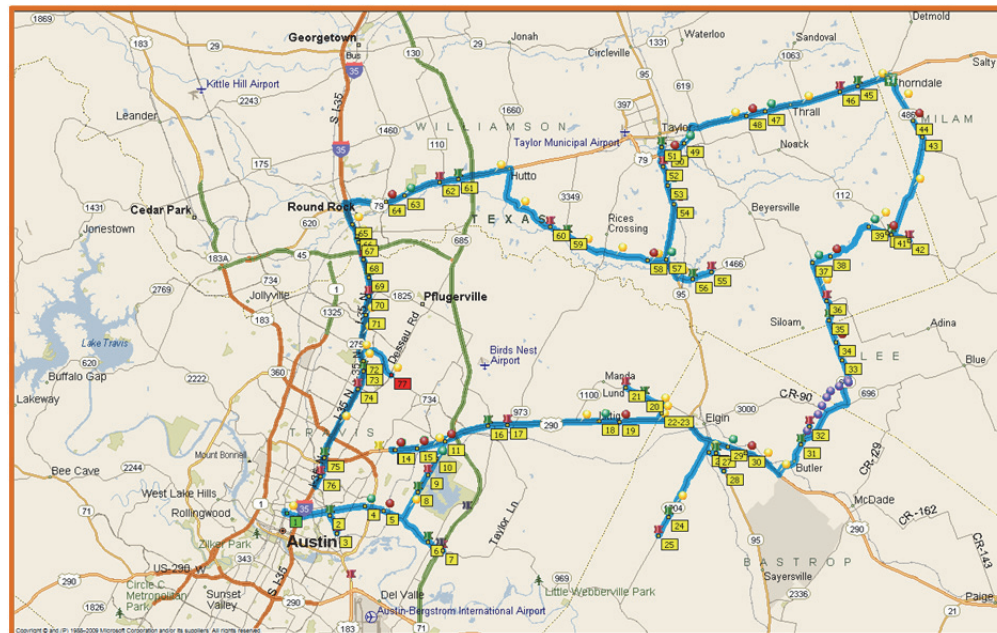
# **2. Characteristics of the Maintenance Test Section Survey**

To address the objective stated above, 34 sections were selected in the Austin District. All selected sections were in the proximity of Austin due to time constraints of the MTSS. The goal was to evaluate at least 30 sections within approximately 4 hours. A total of 34 sections were selected for the survey; of these, 19 sections were part of the Texas Maintenance Assessment Program (TxMAP) statewide evaluation and had been rated earlier in the year. However, 13 sections had not been rated using the TxMAP rating process; two sections were county roads and therefore off-system, and the remainder had ratings that were several months old. Therefore, the TxDOT Maintenance Division was requested to perform TxMAP ratings on all 34 sections just prior to or just after the survey to provide current data for comparison to the MTSS participants' evaluations. (Figure 2.2).

Some of the most interesting areas of the Austin District were selected so as to show some of the challenges that the District has to face to maintain their road network. Some of these challenges include the presence of sulfates in the soil, active clays, and significant amounts of agricultural traffic that can circulate on some of these roads with axle loads above specified limits. Figure 2.1 shows the general location and the route of the survey.



(a) General Location



(b) MTSS Route

**Figure 2.1: General Location and Route of the Maintenance Test Section Survey**

The MTSS took place on October 6, 2010, on the route indicated in Figure 2.1. The route was approximately 160 miles and was conducted from a moving vehicle travelling at approximately 40 miles an hour.

The participants (evaluators) of the survey were purposely given only general instructions about selecting the appropriate rating so as not to bias their opinions about expected maintenance conditions. The evaluators were asked to consider a number of items or attributes that were grouped into four categories: Pavement Score, Traffic Operations Score, Roadside Score, and Overall Score. Figure 2.2 shows the form provided for the evaluation during the rally.

<b>NAME:</b> _____ <b>VEHICLE:</b> _____ <b>SECTION:</b> <b>01</b> _____	<table border="1" style="border-collapse: collapse; width: 40px;"> <tr><td>D</td><td></td></tr> <tr><td></td><td></td></tr> <tr><td></td><td></td></tr> <tr><td></td><td></td></tr> <tr><td></td><td></td></tr> </table>	D										front mid1 mid2 mid3 back	<b>Mark your seat assignment</b>  <b>Highway:</b> <b>SL 111</b> <b>Travel direction:</b> <b>south</b> <b>Beginning TRM:</b> <b>438</b> <b>Ending TRM:</b> <b>439</b>
D													
<b>Pavement Score:</b> (include items such as rutting, cracking, ride, edges, shoulder, failures, etc.)													
5	Well above expectations	Comments:	_____										
4	Above expectations		_____										
3	Meets expectations		_____										
2	Below expectations		_____										
1	Well Below Expectations		_____										
<b>Traffic Operations Score:</b> (include items such as raised pavement markers, striping, attenuators, delineators, shoulder texturing, etc.)													
5	Well above expectations	Comments:	_____										
4	Above expectations		_____										
3	Meets expectations		_____										
2	Below expectations		_____										
1	Well Below Expectations		_____										
<b>Roadside Score:</b> (include items such as vegetation management, trees and brush, drainage, sweeping, encroachment, guardrails, guardrail end treatment, mail boxes, etc.)													
5	Well above expectations	Comments:	_____										
4	Above expectations		_____										
3	Meets expectations		_____										
2	Below expectations		_____										
1	Well Below Expectations		_____										
<b>Overall Score:</b> (rate the section relatively to a similar facility in your state)													
5	Well above expectations	Comments:	_____										
4	Above expectations		_____										
3	Meets expectations		_____										
2	Below expectations		_____										
1	Well Below Expectations		_____										
<b>Additional comments:</b> _____ _____													

**Figure 2.2: Evaluation Form**

## **Maintenance Test Section Survey Process**

A group of six TxDOT vans, ranging in seating capacity from 6 to 15, were used to conduct the MTSS. Each van carried one peer state participant who was located in the front passenger seat. Each van was driven by a TxDOT employee familiar with the roads and route. In addition, a UT/CTR “navigator” was assigned to each van to guide the driver to each survey site using the Microsoft *Streets and Trips 2010* ® software. GPS coordinates were stored to mark the beginning and end of each rating section and used by the navigator to guide the rating team. The navigator would advise the group when the van was approaching a rating section and the moment to begin and end the rating. The navigator also announced the average annual daily traffic (AADT), number of trucks and posted speed limit of the route.

The remaining participants were asked to mark their seating location in the van using the diagram in the upper right hand of the survey sheet. This was done with the intent to consider whether seating location had an effect on ratings. It was expected that the peer state participant would have the best view of the entire roadway while individuals sitting on the left, right, or to the rear of the van might have a somewhat restricted view. Although the analysis of seating position was not carried out, participants from the Strategic Planning and Performance Management Section (SPP) and TTI later used this information when determining how best to plan a Roadway Test Section Evaluation that was to be conducted with a different group of raters by SPP.

At the end of each 1-mile rating section, each participant marked the score for each rating category and provided comments as appropriate.

A five-level scoring system was used; however, these levels are not necessarily consistent with TxDOT’s PMIS five-level scale (i.e., very good, good, fair, poor, and very poor). The main goal of the survey was to assess the threshold level at which maintenance activities are required as perceived by the evaluators from the peer states and by TxDOT personnel. In other words, one of the objectives of the MTSS was to determine whether the pavement condition that triggers maintenance actions in Texas is similar, higher, or lower than that of other states. The evaluators from other states were instructed to evaluate the 34 sections using the criteria they use on similar sections in their state or district.

In establishing the Pavement Score, the evaluators were asked to consider the following items: rutting, cracking, failures, riding quality, pavement edge conditions, and shoulders. For determining the Traffic Operations Score, evaluators were instructed to take into account the condition of the following attributes: raised pavement markers, striping and other pavement graphics, delineators, shoulder texturing, and roadside signs. When evaluating the Roadside score, the following items had to be considered: vegetation management, litter, sweeping, trees, and brush, drainage, right-of-way encroachment, guard rails, guardrail end treatments, and break-away mailbox posts. Finally, the evaluators were also requested to provide an Overall Score for the section that included a combined assessment considering all three categories: Pavement Score, Traffic Operations Score, and Roadside Score.

## **3. Sampling Groups and Evaluation Principles**

A total of 24 people (hereafter referred to as “evaluators”) evaluated the 34 pavement sections. These evaluators have different backgrounds and experiences; therefore, they were grouped into five different sampling groups.



The first group included peer review participants from the six different states mentioned above. This group is referred to, in this report and in the figures, as the **Peer** Group. The second group consisted of TxDOT personnel with significant expertise in pavements from the Austin District, the Maintenance Division (MNT) and the Construction Division (CST) who were not members of the Project Monitoring Committee (PMC). This group is referred to as the **Expert** Group. The third group consisted of the members of the Project Monitoring Committee and is referred to as the **PMC** Group. The PMC is composed of TxDOT subject matter experts, including individuals in senior management positions, who provide guidance and direction to the researchers regarding the research project. The PMC also includes members of the Office of Primary Responsibility (OPR)—in this case the Maintenance Divisions—which is responsible for implementing the research findings. It should be noted that some members of the PMC are also members of the TxDOT Maintenance Peer Review Team. The Maintenance Peer Review Team visits each district on a periodic basis and meets with members of the district staff and administration to review methods used to pick projects, perform maintenance, and conduct maintenance activities. The fourth group consisted of all other participants of the survey who did not belong to any of the previous groups. This group is referred to as the **Other** Group. This group was very diverse and included TxDOT personnel with non-pavement related backgrounds such as biology, geology, traffic engineering, strategic planning and performance measurement, and contract management, among others. This group was also considered to be representative of typical, public road users. Finally, a fifth group was formed by combining all TxDOT personnel. This group is referred to as the **TxDOT** Group. Therefore, the TxDOT Group is the sum of the Expert, PMC and Other groups. The members of the various groups are given in Table 3.1.

**Table 3.1: Participants of the Maintenance Test Section Survey**

<b>Group</b>	<b>Members</b>
Peers	Steve Takigawa, Roy Rissky, Jennifer Brandenburg, Dave Bierschbach, Eric Pitts, and Jim Carney
Experts	James Williams, Neal Munn, Lowell Choate, Mike Arellano, Magdy Mikhail
PMC	Dennis Cooley, Lonnie Gregorcyk, Tammy Sims, Jeff Seiders, Toribio Garza
Others	Bonnie Lister, Epi Gonzalez, Tina Geiselbrecht, Sylvia Medina, Zhanmin Zhang, Robert Blackwell, Gary Lantrip, John Sabala

## 4. Ratings per Individual and per Sampling Group

Before analyzing the ratings of the individual roadway sections, it is prudent to review the evaluators and the sampling groups to determine whether there are any significant differences at the aggregated level.

Figure 4.1 shows the average of all ratings for all sections by each evaluator. The results from the Peer Group are presented first (in purple), followed by the Expert (in green), PMC (in red), and Other Group (in blue), respectively. Some initial observations can be made from the

data displayed in the figure. It is apparent from Figure 4.1 that the Peer and the PMC groups tend to rate the sections lower than the other groups. It is also apparent that the Expert Group tends to rate the sections higher than the rest and that the Other Group is the group showing the largest variability in the results. This was expected because the Other Group consists of TxDOT personnel with very different technical backgrounds.

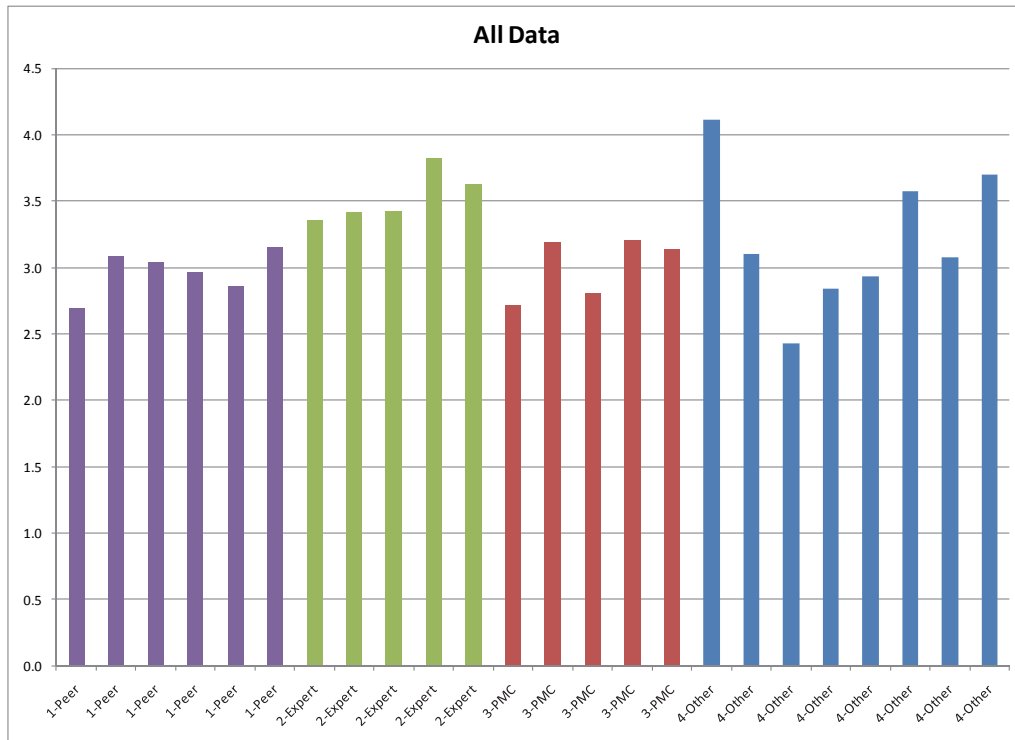
In order to test whether the differences of the aggregated ratings are significantly different amongst the various groups, a series of t-tests were conducted. The results of these tests are presented in Table 4.1. The test were conducted at level  $\alpha$  of 5 percent (i.e., 5 percent probability of Type I error). The null hypothesis ( $H_0$ ) in all cases was that the average ratings of the two groups being considered are equal. If the statistical test rejects  $H_0$ , we can conclude that there is a statistically significant difference between the average ratings of the two groups. The null hypothesis is rejected when the t-statistic from the sample is greater than the critical t-value indicated in Table 4.1.

**Table 4.1: Test of Hypotheses Comparing the Various Groups**

<b>Groups</b>	<b>t-statistic</b>	<b>t-critical</b>	<b>Outcome</b>	<b>Comment</b>
Peer vs. TxDOT	-1.54	-2.07	Cannot Reject $H_0$	No significant difference
Peer vs. Expert	-5.10	-2.26	Reject $H_0$	Significant difference
Peer vs. PMC	-0.35	-2.26	Cannot Reject $H_0$	No significant difference
Peer vs. Other	-1.09	-2.18	Cannot Reject $H_0$	No significant difference

The results show that there is only a significant difference between the Peer and the Expert Groups. It can also be seen in the table that all t-statistics are negative which indicates that the Peer Group rated the sections lower (on average) than any of the other groups.

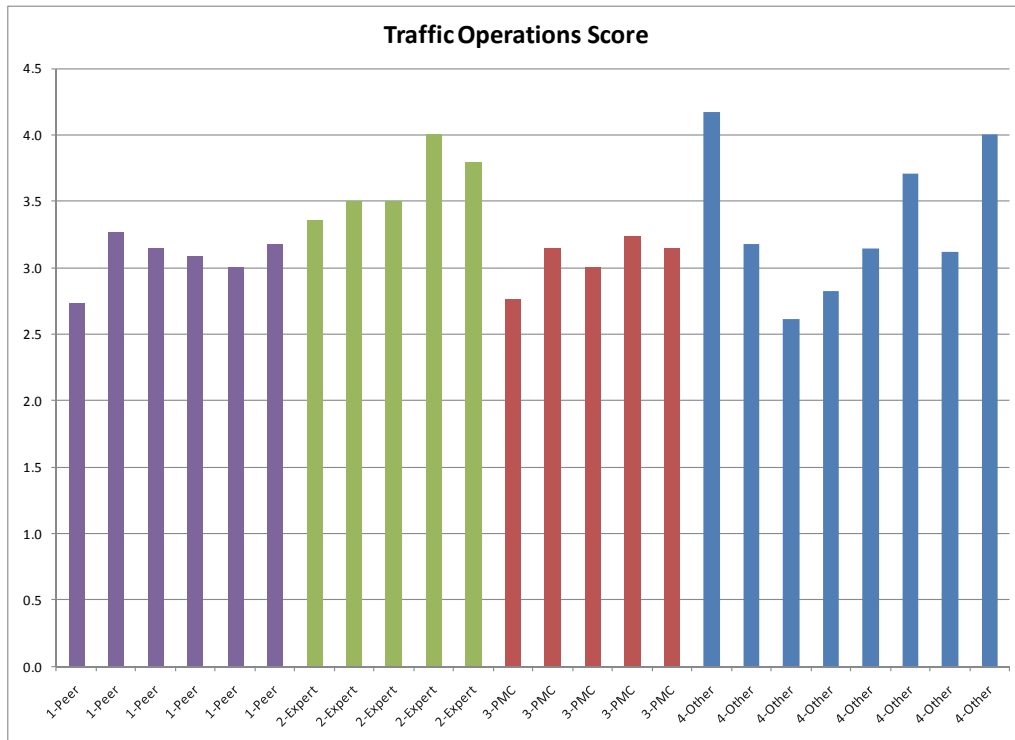
Figures 4.2 to 4.5 show similar results but for each of the attributes rated separately. The results are given for the Pavement, Traffic Operation, Roadside, and Overall rating, in Figures 4.2, 4.3, 4.4 and 4.5, respectively.



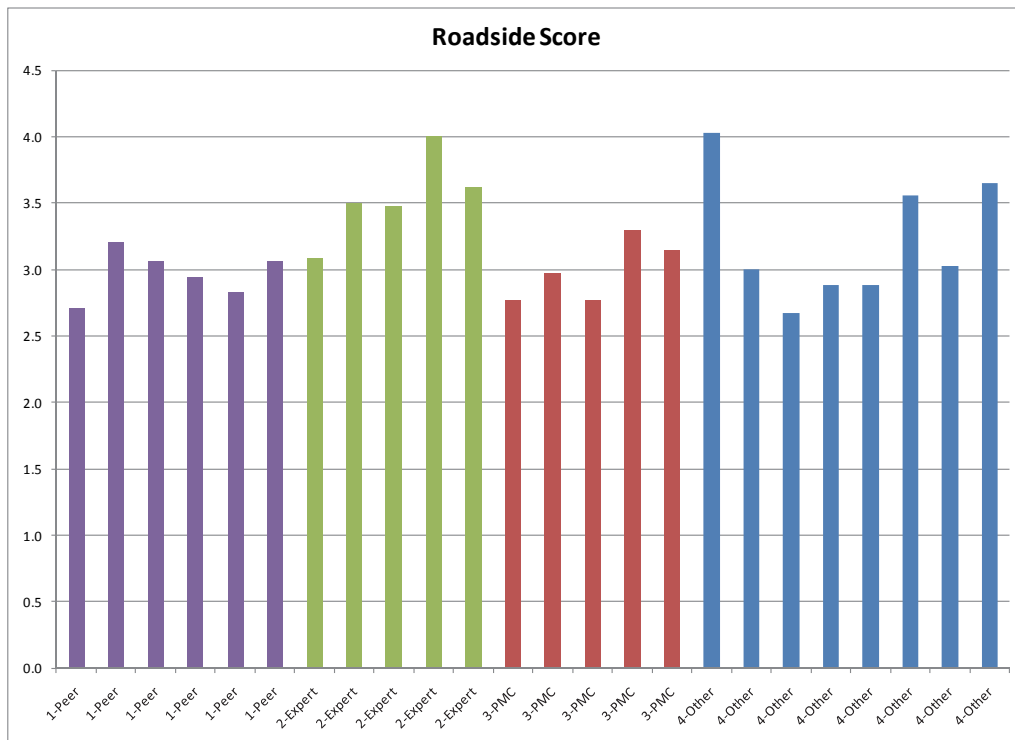
**Figure 4.1: Aggregated Evaluation of All Sections**



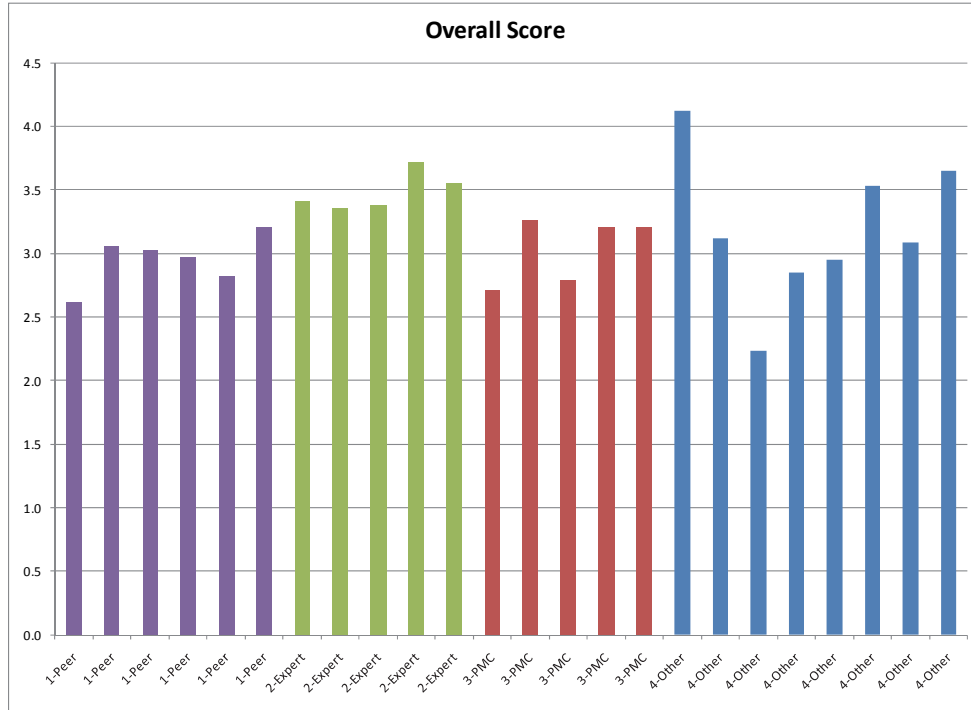
**Figure 4.2: Pavement Score of All Sections (per individual)**



**Figure 4.3: Traffic Score of All Sections**

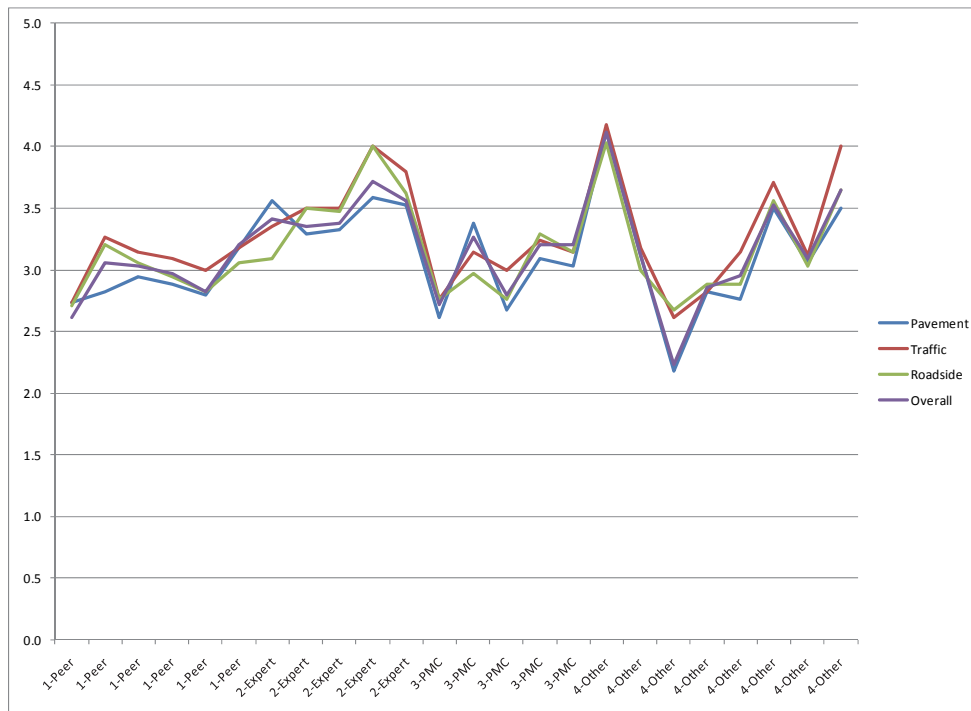


**Figure 4.4: Roadside Score of All Sections (per individual)**



**Figure 4.5: Overall Score of All Sections (per individual)**

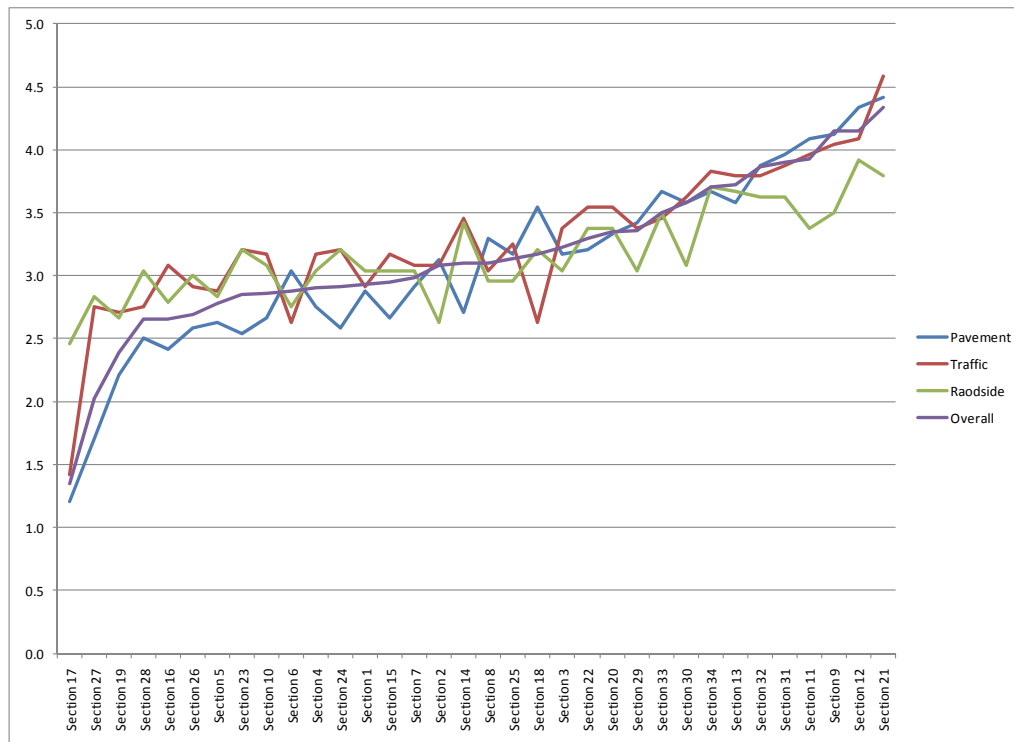
Figure 4.6 shows the ratings for each of the attributes evaluated (i.e., Pavement, Traffic Operation, Roadside, and Overall) for all the sections per individual. The Pavement rating was lower for some sections while the Roadside rating was lower for others. Traffic Operation received higher ratings for most sections.



**Figure 4.6: All Scores of All Sections (per individual)**

## 5. Ratings per Section

In this section of the report, the ratings are presented for each roadway section. Sections are displayed in order of increasing “Overall” rating. This is done to determine whether potential differences in the ratings are systematic and whether these potential systematic differences vary according to the condition of the sections evaluated. Figure 5.1 shows that there is not a systematic difference according to the condition of the pavement section. That is, the variability of the results is almost the same independently of the rating of the sections.

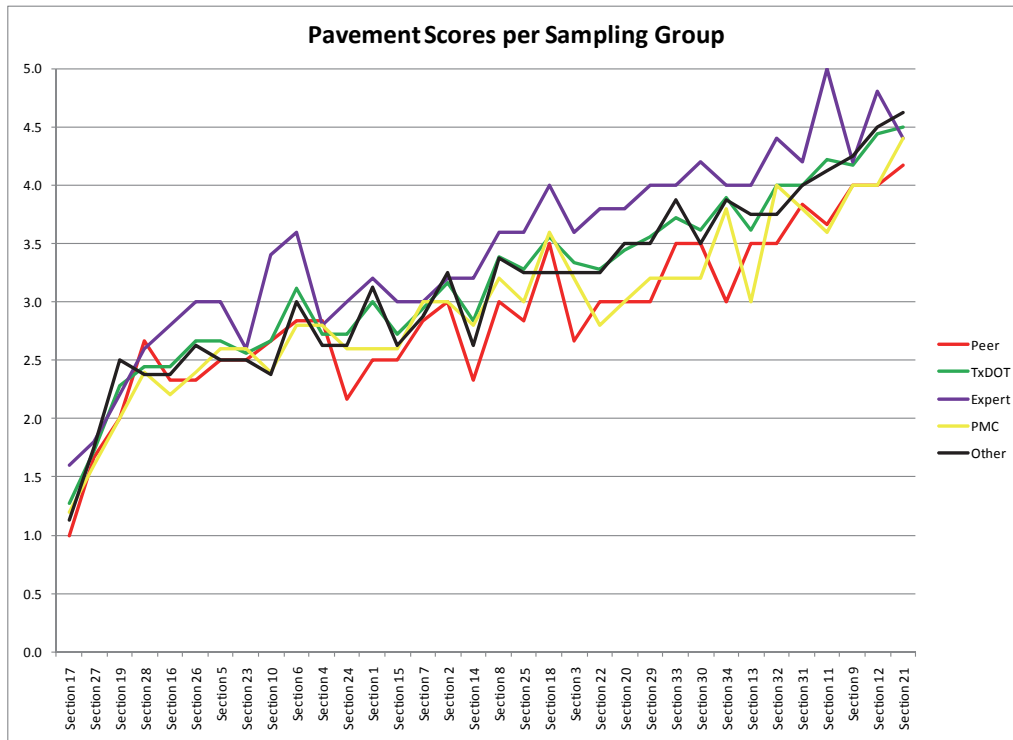


**Figure 5.1: Aggregated Rating of All Sections (Overall Score)**

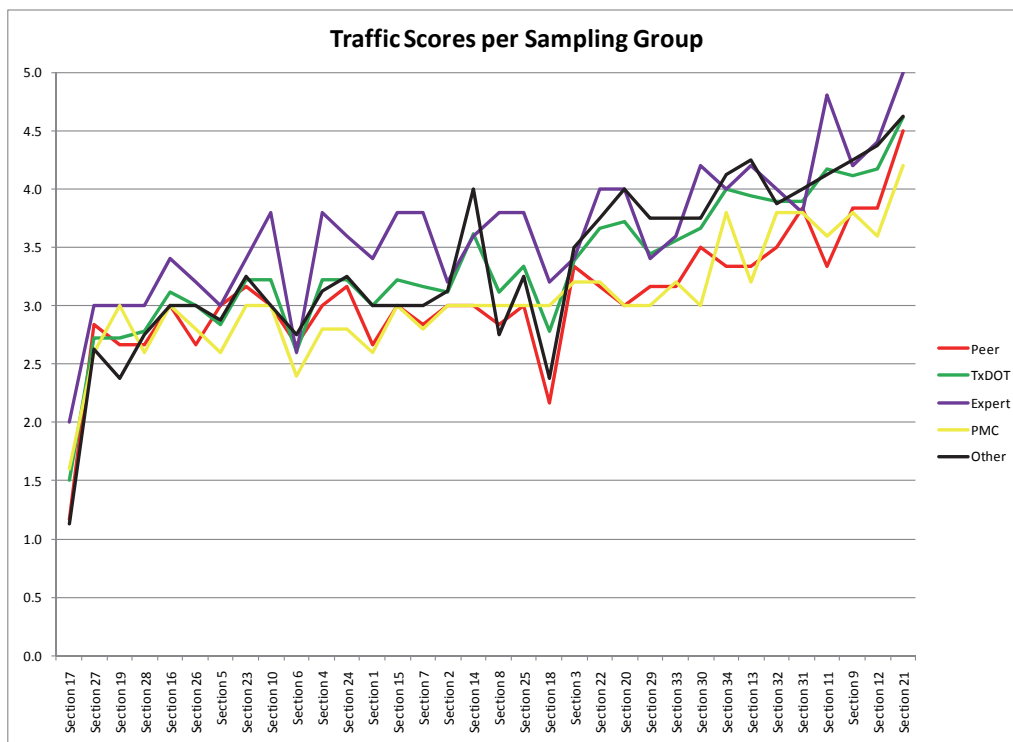
Figures 5.2 to 5.5 show the comparative results between the ratings from the Peer Group and all other sampling groups. The results are presented per category; that is, the comparison of the Pavement, Traffic Operation, Roadside, and Overall ratings are presented in Figure 5.2, 5.3, 5.4, and 5.5, respectively.

It can be observed that for most of the sections evaluated and for all categories rated, as well as for the overall rating, the Peer Group rated the sections consistently lower and the Expert Group consistently higher. This finding supports the results presented in Section 4.

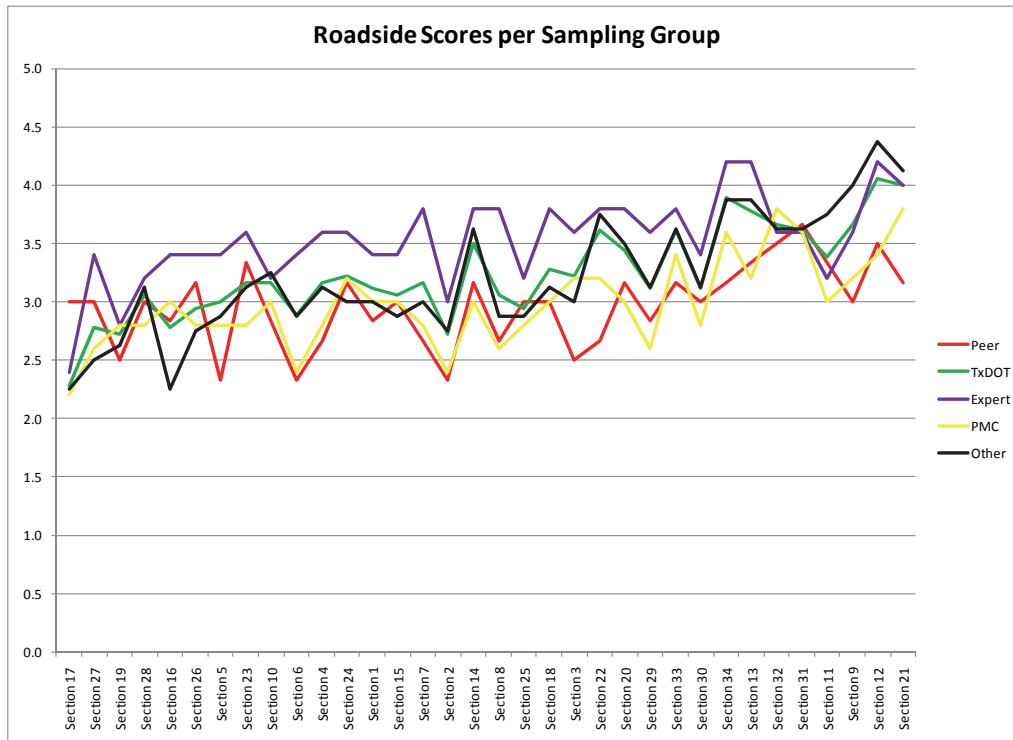
It is also interesting to note that the Pavement rating had the widest range in results (see Figure 5.2). The Roadside rating had the lowest variability in ratings (Figure 5.4).



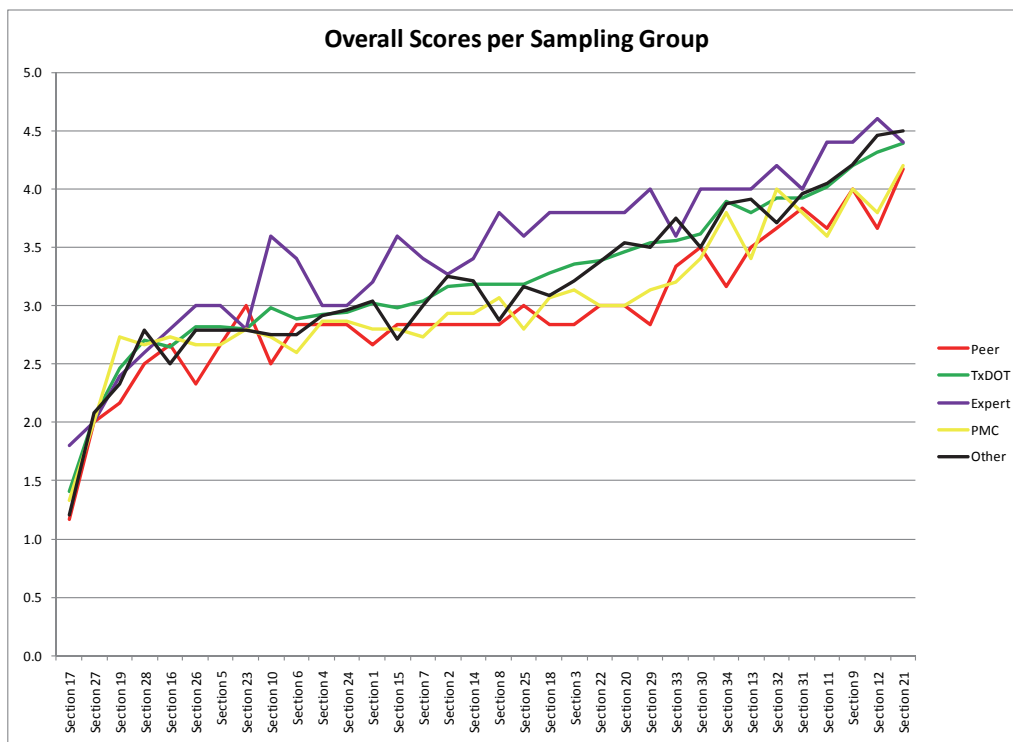
**Figure 5.2: Peers versus all Groups (Pavements)**



**Figure 5.3: Peers versus all Groups (Traffic)**



**Figure 5.4: Peers versus all Groups (Roadside)**

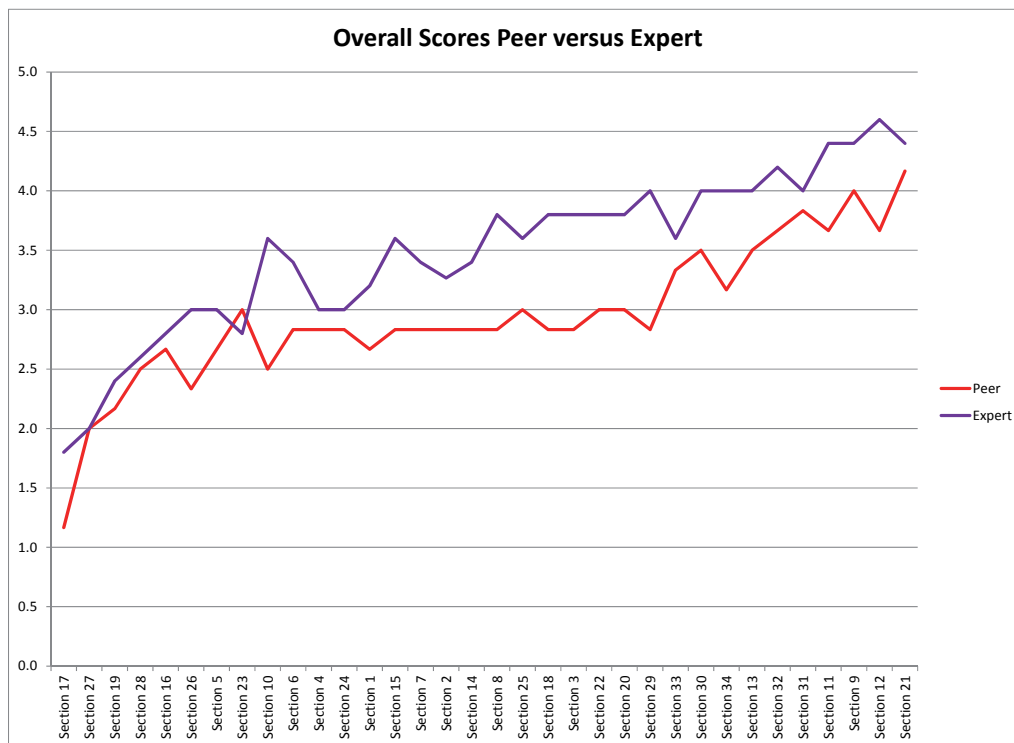


**Figure 5.5: Peers versus all Groups (Overall)**

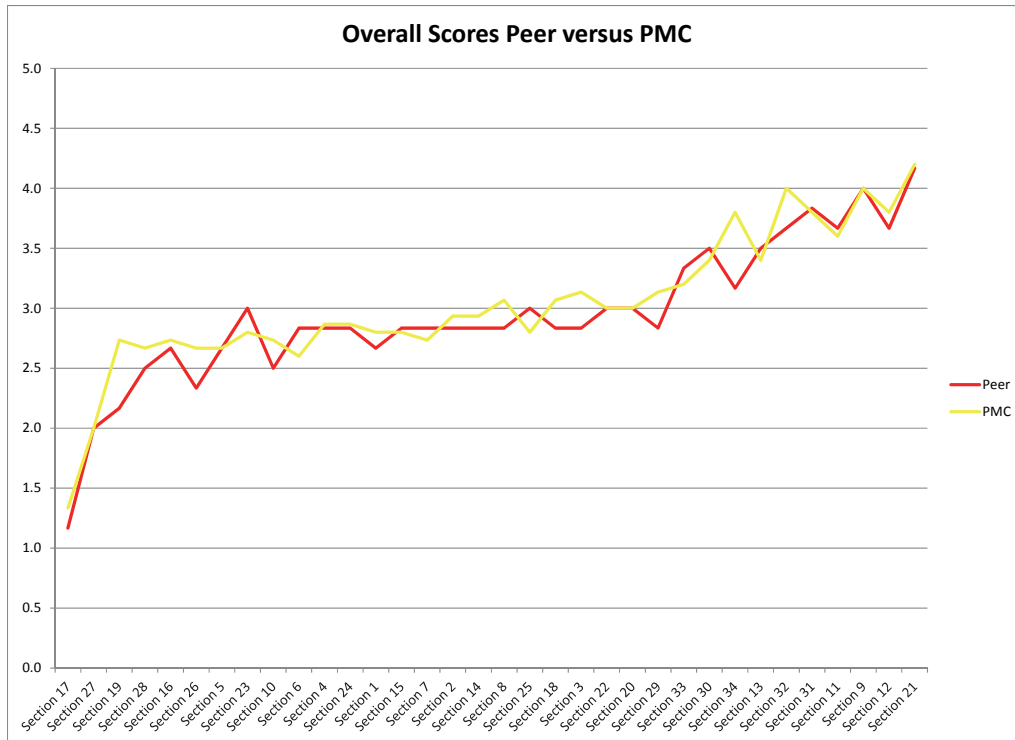


The results presented in Figures 5.2 to 5.5 are quite comprehensive but the figures are somewhat crowded with information. For this reason, the information was disaggregated and presented in Figures 5.6 to 5.9. These figures show the comparison of the ratings of the different sampling groups on a section-by-section basis. The ratings of the Peer Group are compared to those of the Expert, PMC, Other, and TxDOT Groups in Figures 5.6, 5.7, 5.8, and 5.9, respectively.

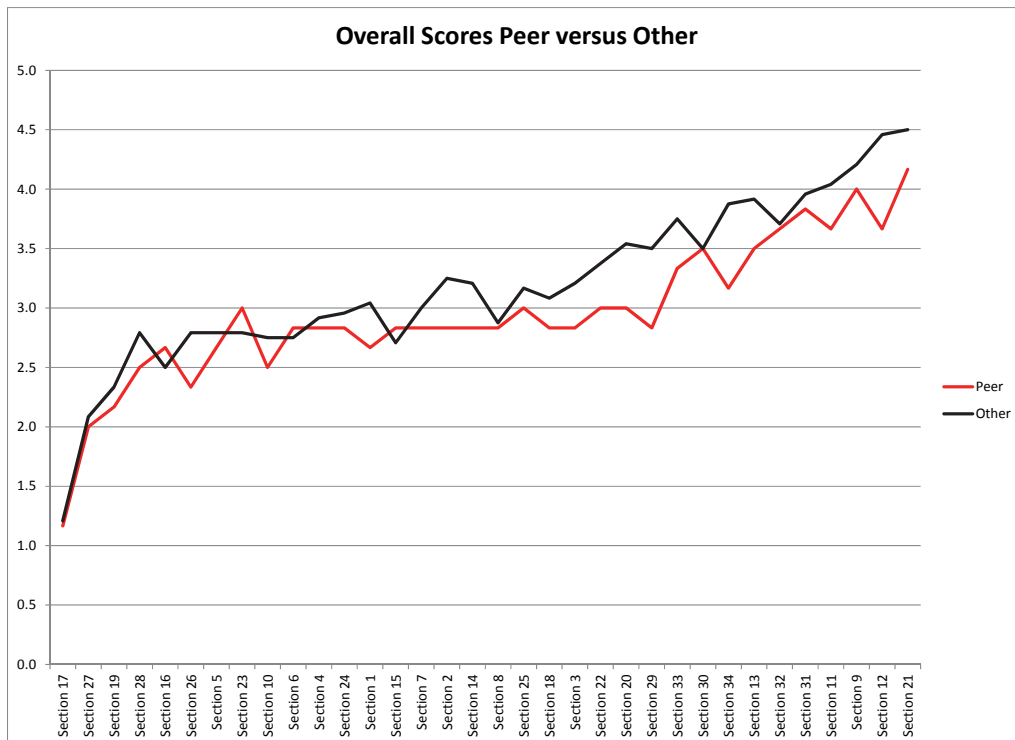
It is apparent from the figures that the ratings of the Peer and PMC Groups are quite similar while the Expert, Other, and TxDOT Groups tends to consistently rate the sections higher than the control group, which is the Peer Group. In order to test these observations, a series of paired t-tests were performed. The results of these statistical tests are presented in Table 5.1.



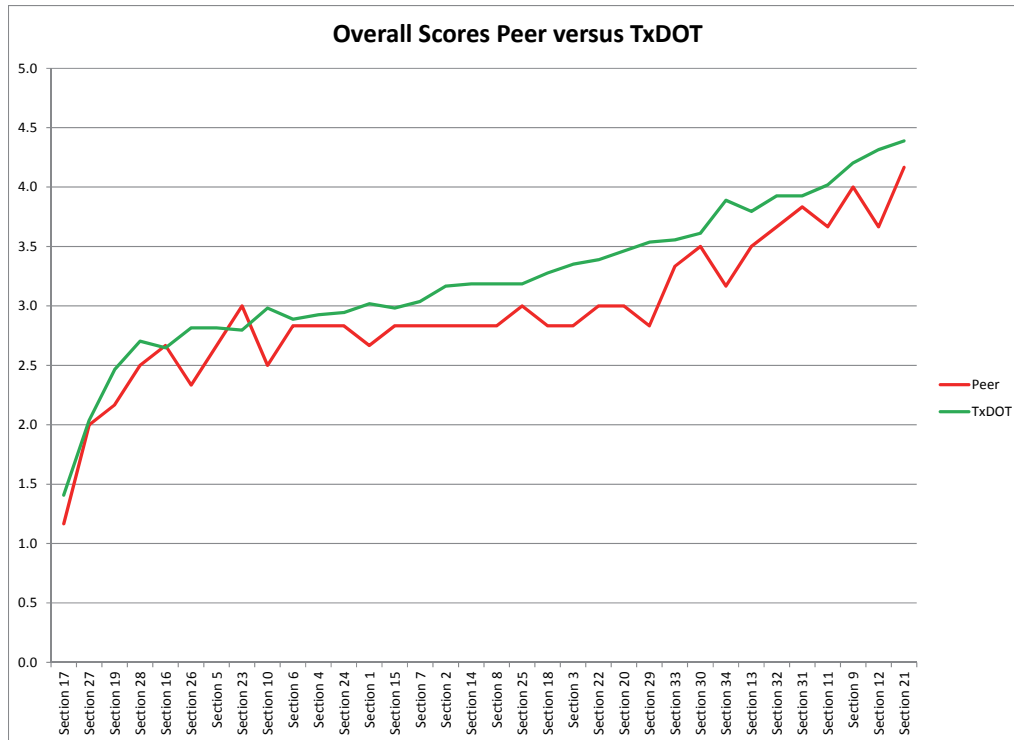
**Figure 5.6: Comparison of Overall Scores of Peers vs. Expert Group**



**Figure 5.7: Comparison of Overall Scores of Peers vs. PMC Group**



**Figure 5.8: Comparison of Overall Scores of Peers vs. Other Group**



**Figure 5.9: Comparison of Overall Scores of Peers vs. TxDOT Group**

The paired t-test is ideal for comparing ratings of the same sections between two different sampling groups. It is a very robust test that does not incorporate any assumption on the distribution of the ratings. It compares the differences in the ratings for each of the sections. A negative difference means that the average rating of the Peer Group is lower than that of the group being compared. The tests were carried out at a significance level  $\alpha = 5\%$ . The sample size was 34 so the critical t-value is -2.03. The null hypothesis proposes that there is no difference between ratings. Rejecting the null hypothesis implies that the difference in ratings is significant. The results of the statistical analyses are presented in Table 5.1.

**Table 5.1: Test of Hypotheses Comparing the Peer Group Overall Ratings**

Group	Difference	t-statistic	Outcome	Comment
Expert	-0.53	-9.28	Reject $H_0$	Significant difference
PMC	-0.09	-2.49	Reject $H_0$	Significant difference
Other	-0.24	-5.83	Reject $H_0$	Significant difference
TxDOT	-0.28	-7.98	Reject $H_0$	Significant difference

Table 5.1 shows that in all cases the ratings by the Peer Groups are statistically lower than those of all other sampling groups. This difference is particularly large when compared with the Expert Group: on average, the Expert Group rated the sections 0.53 points above the Peer Group. This difference is not only statistically significant but it is also quite large.

The differences with the Other and TxDOT Groups are also statistically significant and relatively important. These two groups, on average rated the sections 0.24 and 0.28 points higher

than the Peer Group. The difference in average ratings between the Peer and the PMC Group was also statistically different but very small, less than 0.1 points. This indicates that, on average, the PMC and the Peers rated the sections quite similarly while the Expert and Other Group rated them higher on average. This supports the observations made from Figures 5.6 to 5.9.

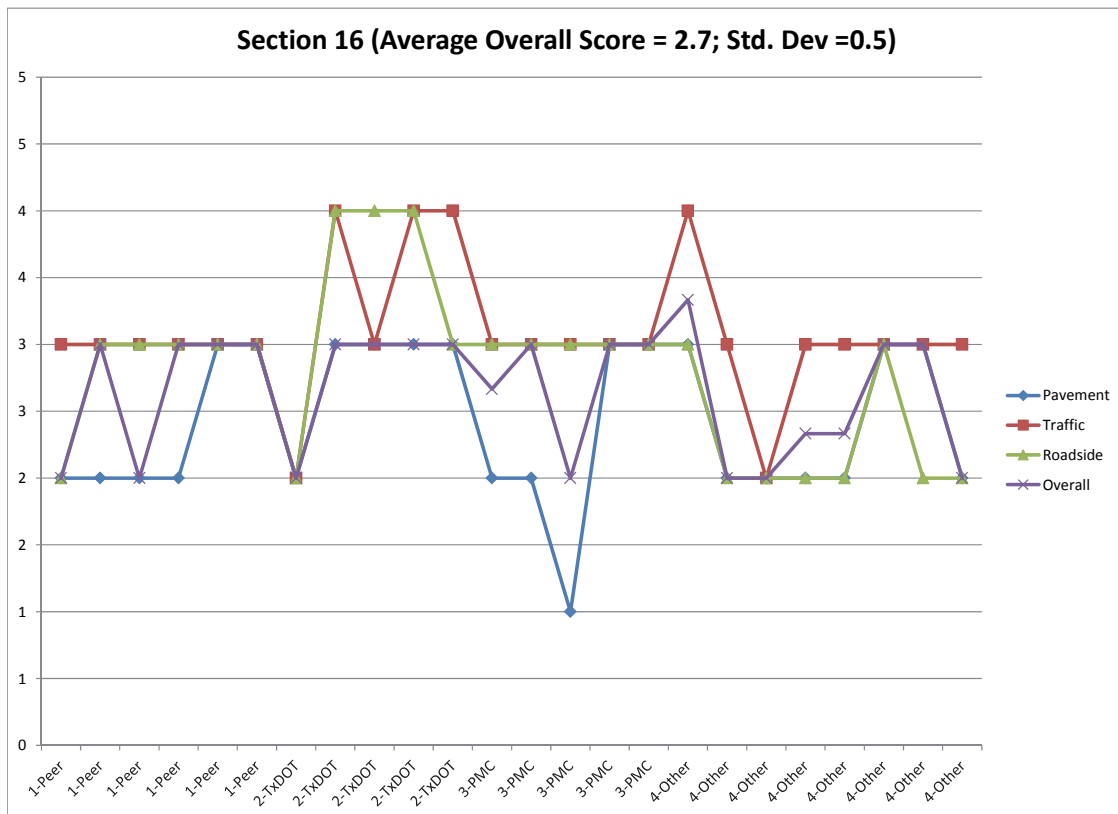
## **6. Data of Individual Sections**

In this section, all the ratings of all the categories for selected sections are presented with a photo taken near the time that the survey was conducted. The objective is to determine whether, for each section, there were any important differences in the ratings. Nine sections were selected with Average Overall Ratings ranging from 2.7 to 4.2. The following sections were selected:

- Sections 16, 26 and 28, which represent some of the lowest maintenance condition sections in the sample. This information is displayed in Figures 6.1, 6.2, and 6.3, respectively.
- Section 10, which represents the midrange of the sections receiving overall ratings below 3.0. This information is given in Figure 6.4.
- Sections 03 and 29, shown in Figures 6.5 and 6.6, represent an average condition with overall ratings of 3.3 and 3.4, respectively.
- Sections 34 and 32, shown in Figures 6.7 and 6.8, represent the midrange of sections receiving overall ratings above 3.0. Their overall ratings are 3.7 and 3.9, respectively.
- Section 12 represents one of the best sections in the sample with one of the highest overall scores.



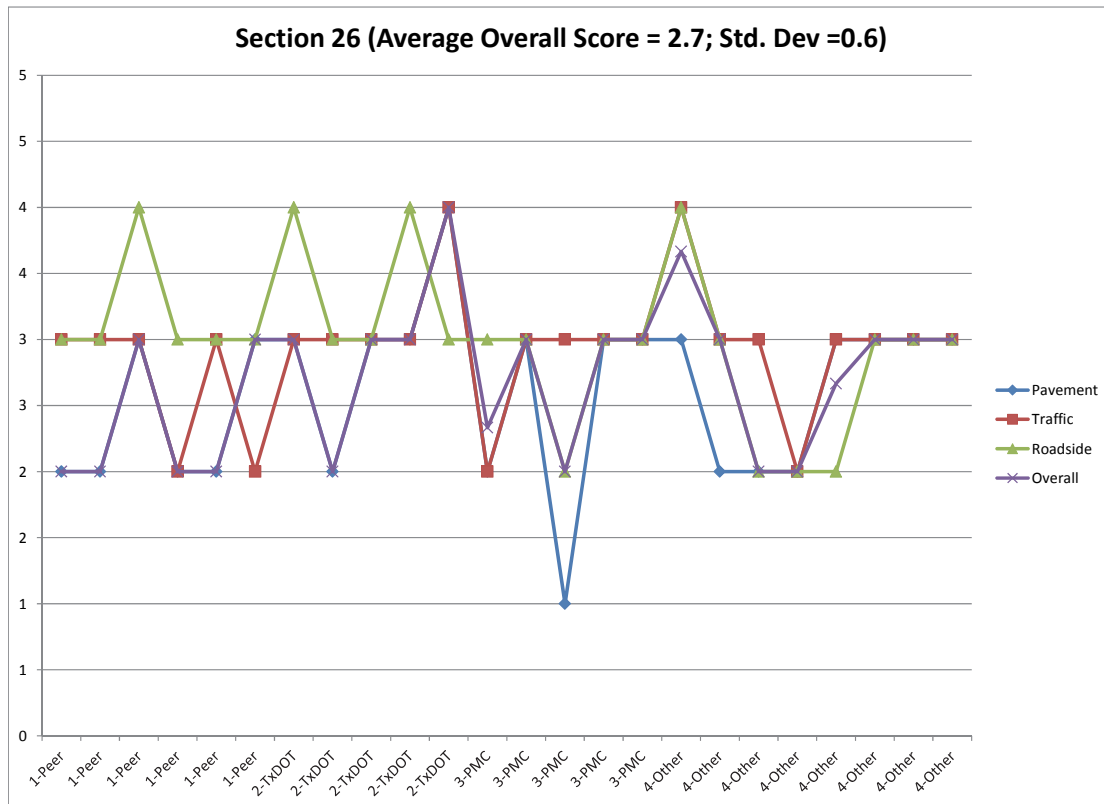
**Figure 6.1(a): Section 16—a two-lane FM road with no shoulders**



**Figure 6.1(b): All Ratings for Section 16**



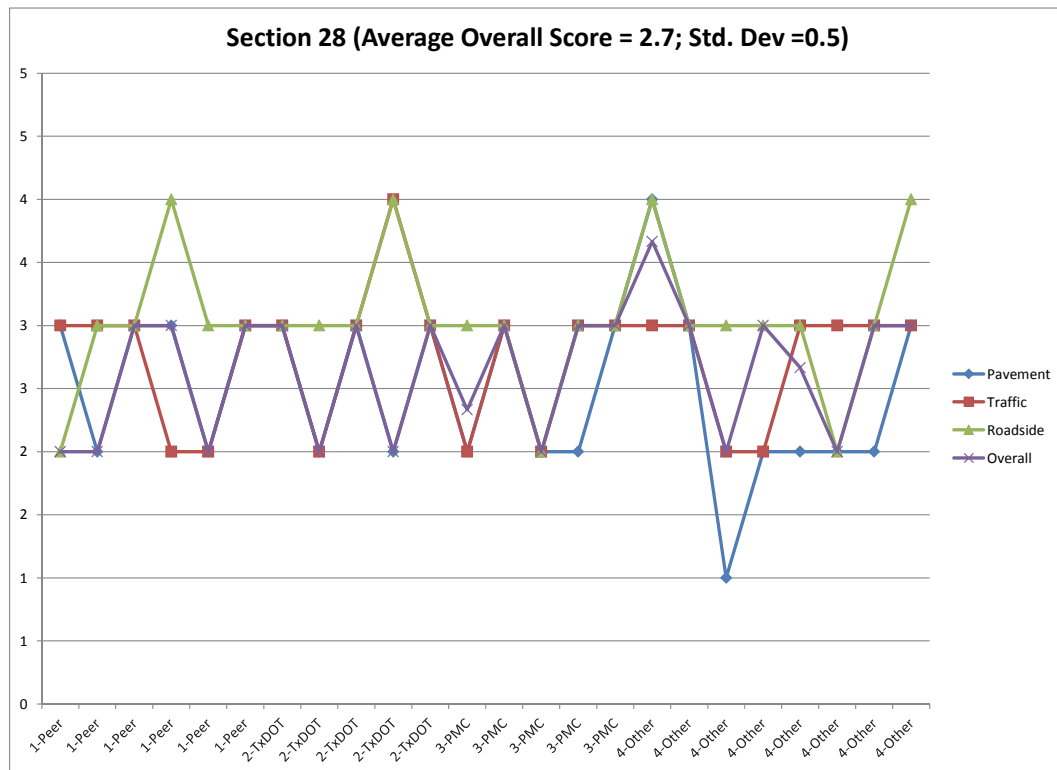
**Figure 6.2(a): Section 26—a two-lane FM road with no shoulders**



**Figure 6.2(b): All Ratings for Section 26**



**Figure 6.3(a): Section 28—a two-lane FM road with no shoulders**

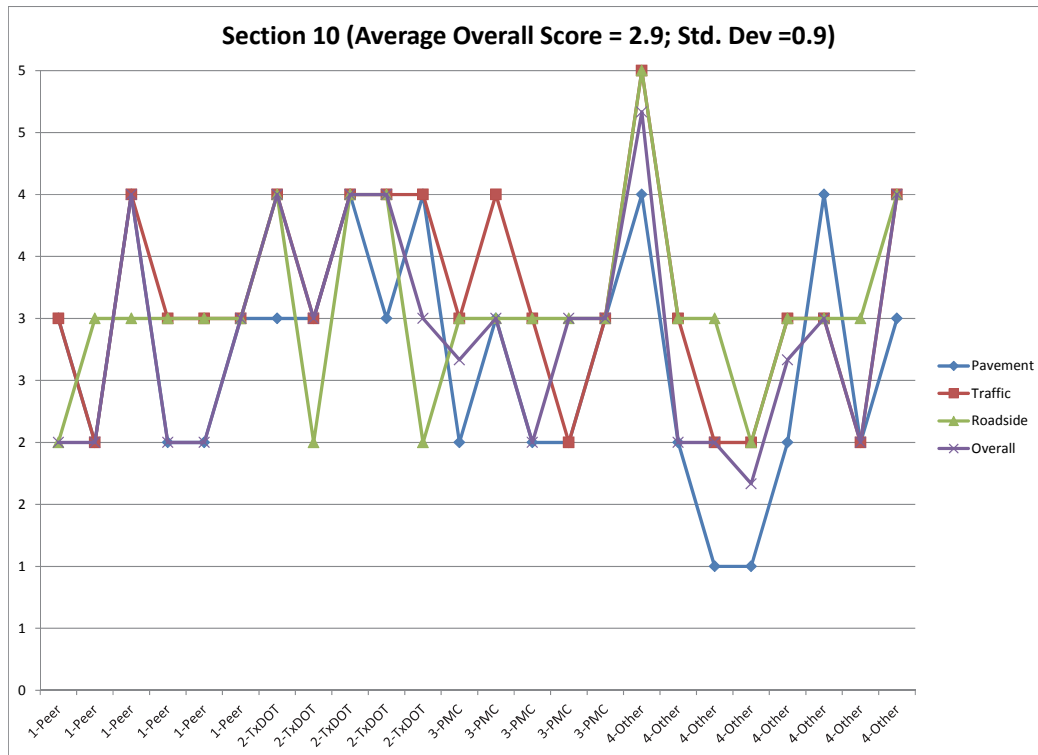


**Figure 6.3(b): All Ratings for Section 28**





**Figure 6.4(a): Section 10—a two-lane FM road with no shoulders**

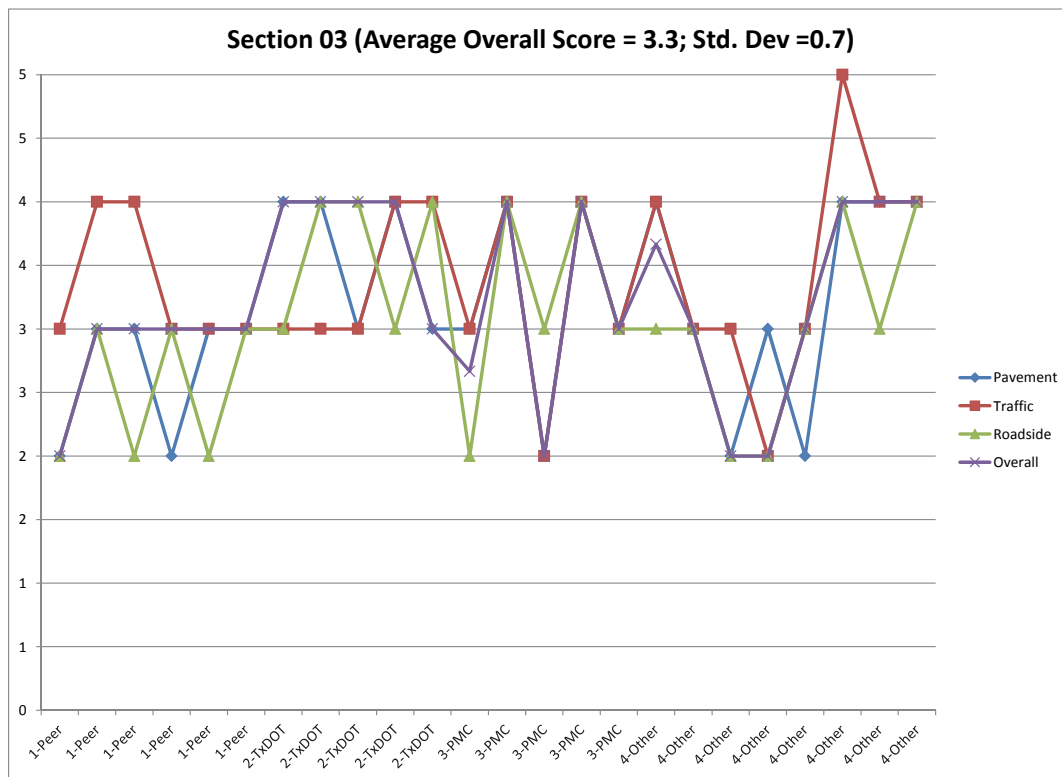


**Figure 6.4(b): All Ratings for Section 10**





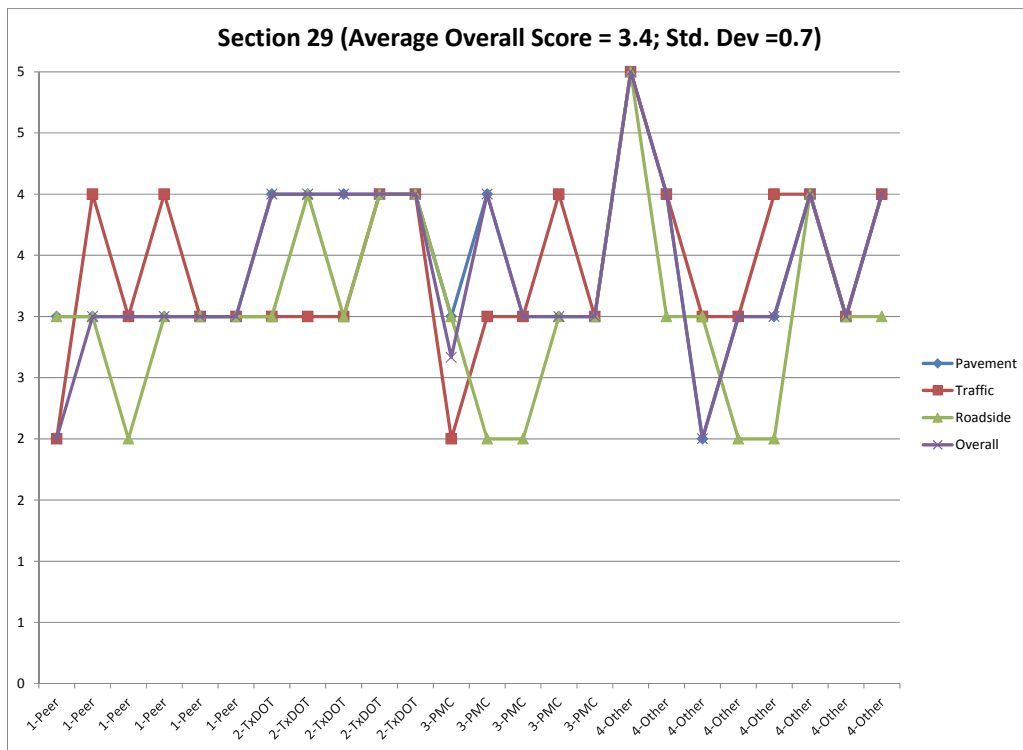
**Figure 6.5(a): Urban FM road**



**Figure 6.5(b): All Ratings for Section 03**



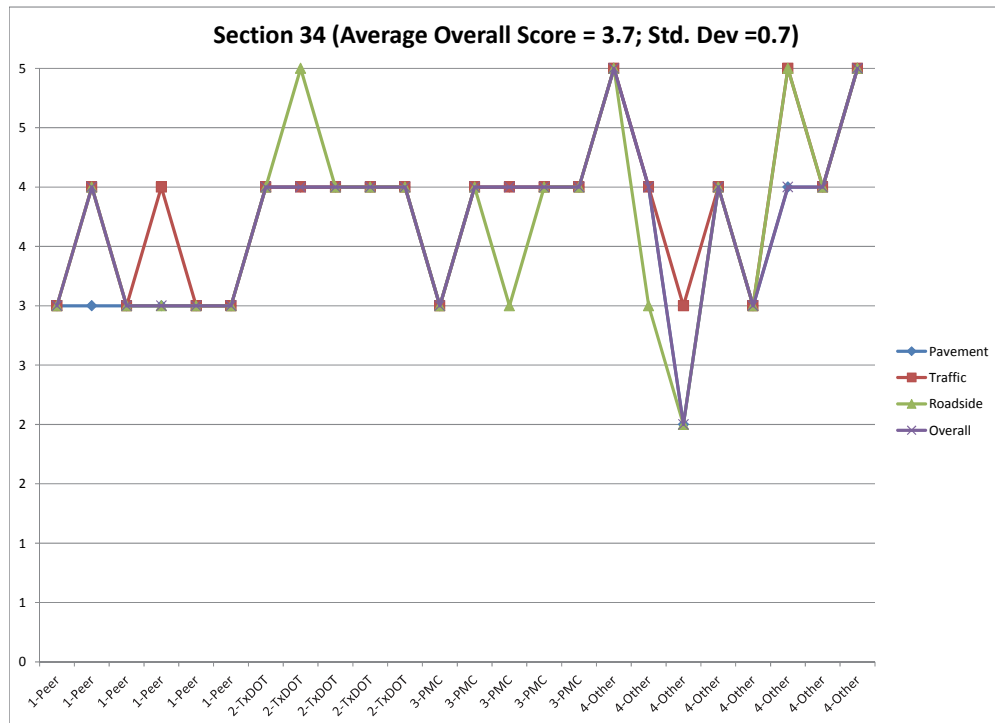
**Figure 6.6(a): Rural US Highway**



**Figure 6.6(b): All Ratings for Section 29**



**Figure 6.7(a): Urban Interstate Highway**

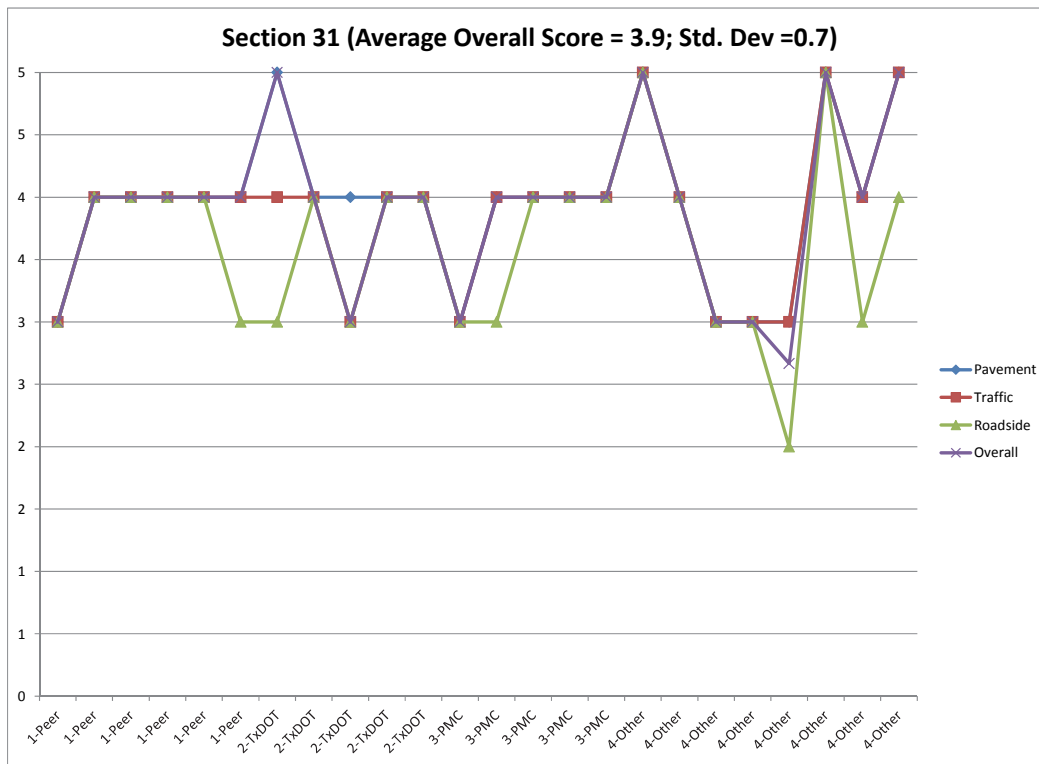


**Figure 6.7(b): All Ratings for Section 34**





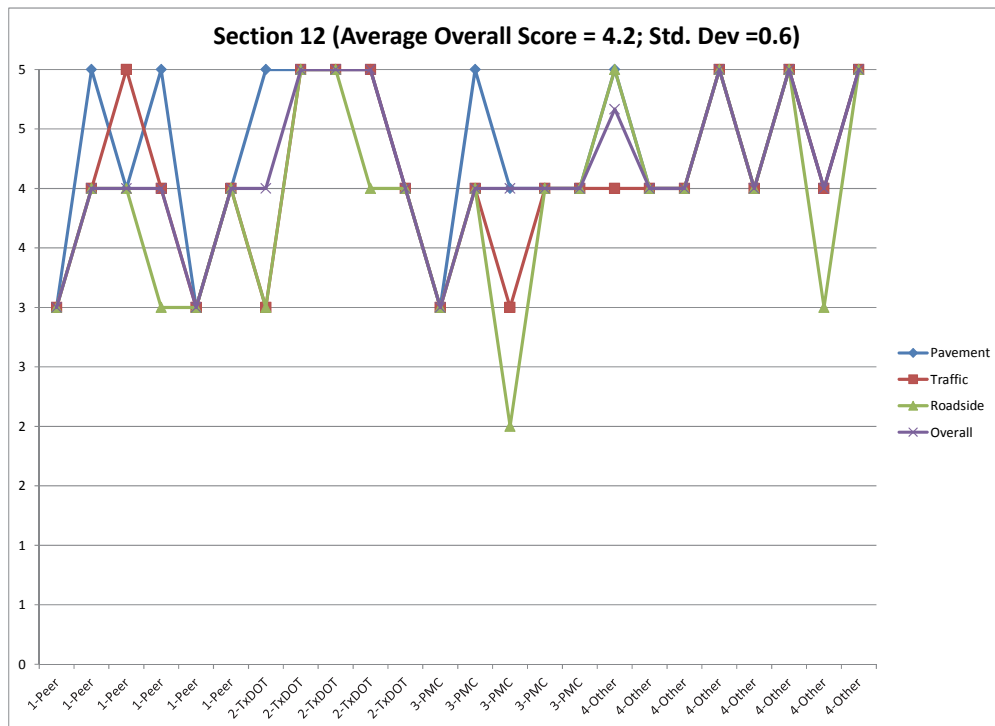
**Figure 6.8(a): Urban Interstate Highway**



**Figure 6.8(b): All Ratings for Section 32**



**Figure 6.9(a): Rural State Highway with Shoulders**

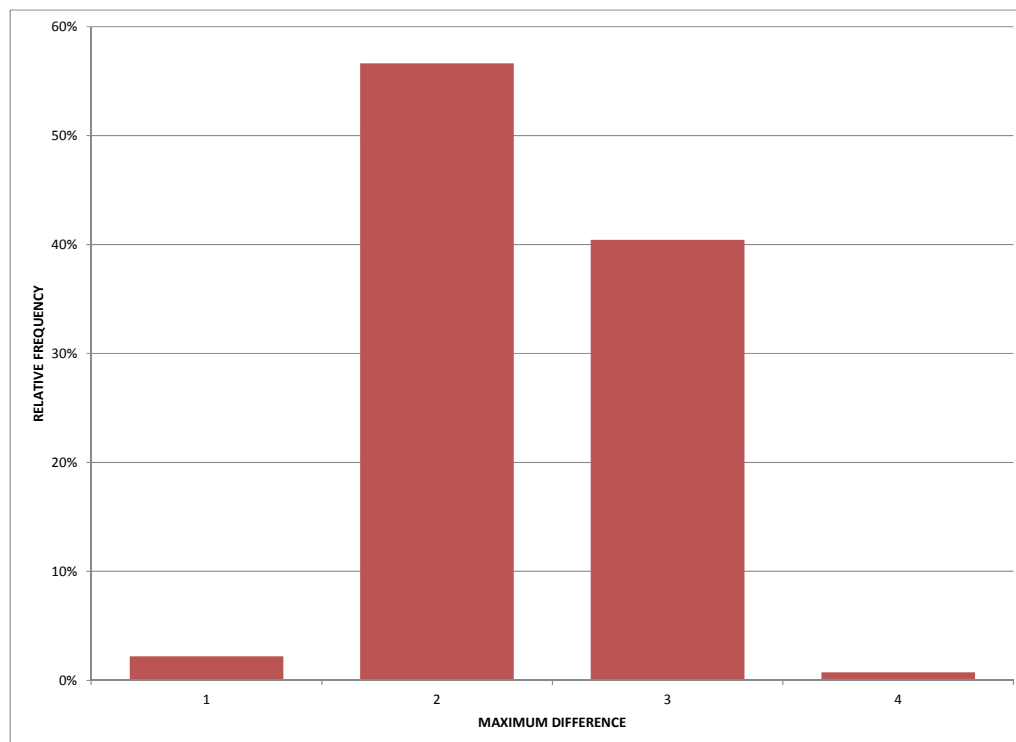


**Figure 6.9(b): All Ratings for Section 12**

Despite the wide diversity of evaluators (in terms of backgrounds and experience) used in this survey and the limited rating guidelines provided to them, the ratings show a surprisingly high degree of consistency in the ratings for all sections, as depicted in Figures 6.1 to 6.9.

For three sections the maximum range in ratings for all evaluations was 1 point. For the majority of sections the maximum difference between the 34 individual evaluators was 2 points. In several cases the maximum difference was 3 points, and in only one case was 4 points.

Figure 6.10 shows the distribution of the maximum difference for the 34 sections and the 4 categories (136 ratings). It can be seen that in almost 60 percent of the cases the maximum difference was only 2 points between maximum and minimum rating given by any of the 24 evaluators. This shows the consistency in the scores. The maximum difference among the 24 evaluators was only 3 points in about 40 percent of the cases.



**Figure 6.10: Maximum Differences in Ratings**

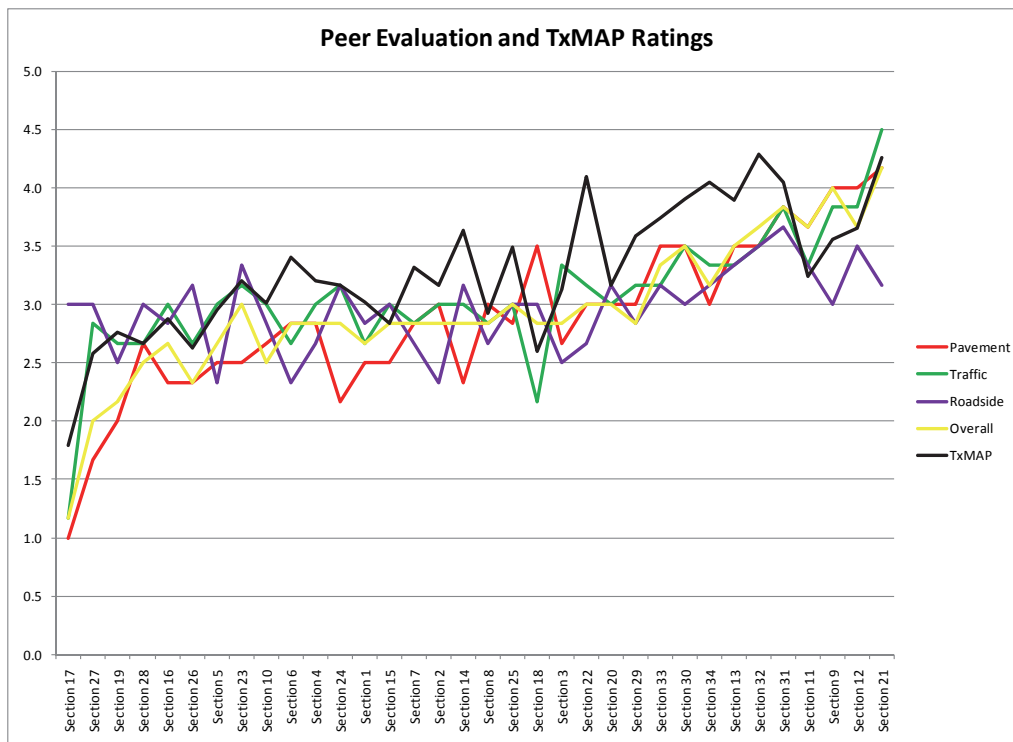
## **7. Comparison with TxMAP Scores**

Although the main objective of the MTSS was to benchmark the maintenance practices in the state, the categories evaluated were consistent with the categories that are evaluated when TxMAP scores are determined. Furthermore, the evaluation of the pavement sections during a TxMAP evaluation is more specific and systematic, and the sections are evaluated at lower travelling speeds allowing the raters to have a better look at the various distress types and the conditions of the various categories evaluated. However, it is still interesting to compare the rating obtained during the survey to TxMAP scores.

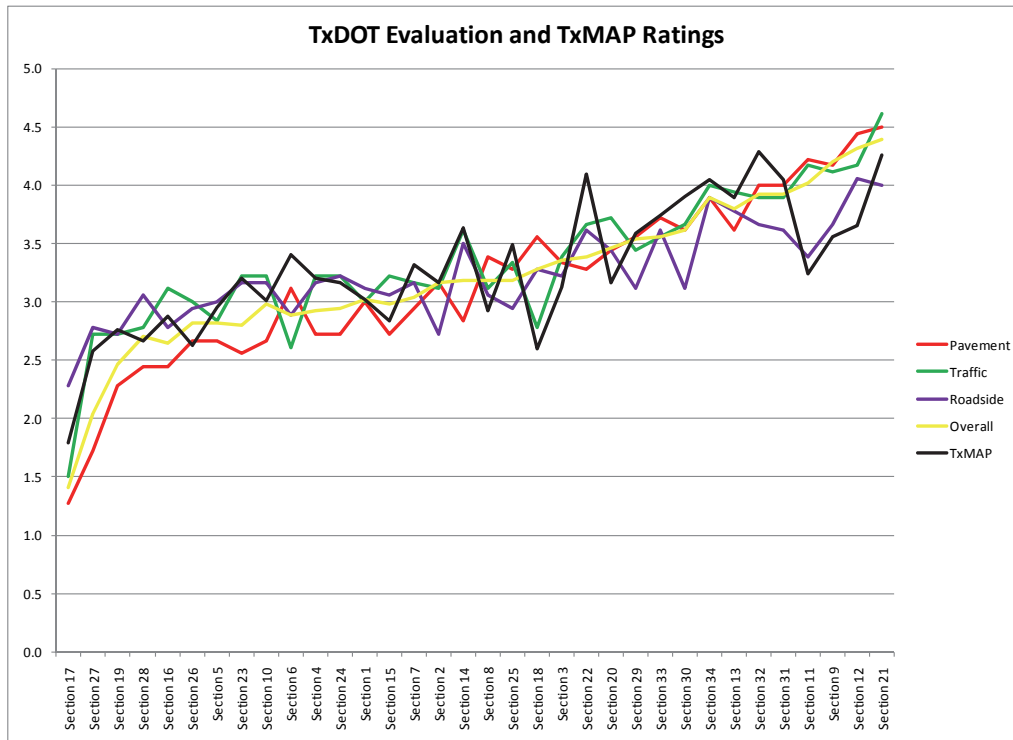
Figures 7.1 to 7.5 display TxMAP scores for all the sections (in black) with the scores of the individual categories (i.e., Pavement Score, Traffic Operation Score, Roadside Score, and

Overall Score) for each sampling group. The rating given by the Peer, TxDOT, Expert, PMC, and Other Group are given in Figures 7.1, 7.2, 7.3, 7.4, and 7.5, respectively.

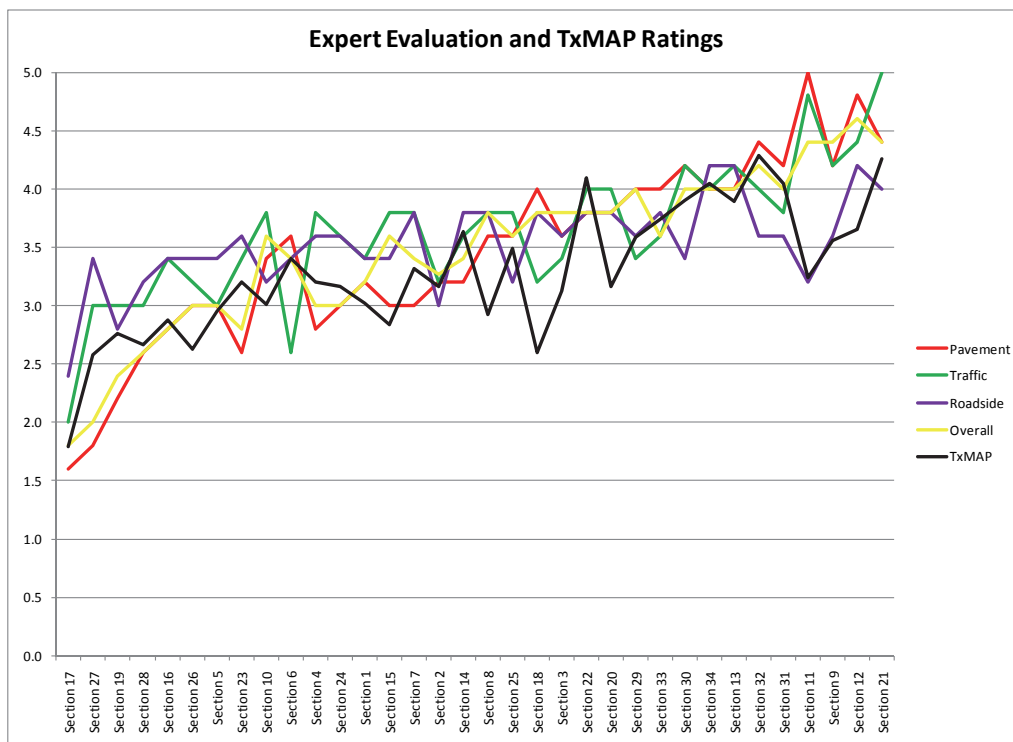
A quick look at the figures reveals that the ratings of the Peer and PMC Groups seem to be below TxMAP scores. However, it is interesting to note that the ratings of the TxDOT, Expert, and Other Groups seem to overlap with TxMAP scores, indicating that the evaluation of these three groups are consistent with TxMAP scores. This may not be a surprise in the case of the Expert Group as this group consists of pavement experts who are familiar with Texas network and TxDOT rating system. It is surprising, however, that the TxDOT and Other (which could also be considered as the Road Users Group) groups seem to have predicted TxMAP scores quite well too. In order to evaluate the validity of these visual impressions, a series of statistical analyses were conducted. The results of these analyses are presented in Table 7.1.



**Figure 7.1: Peers Evaluation versus TxMAP Ratings**

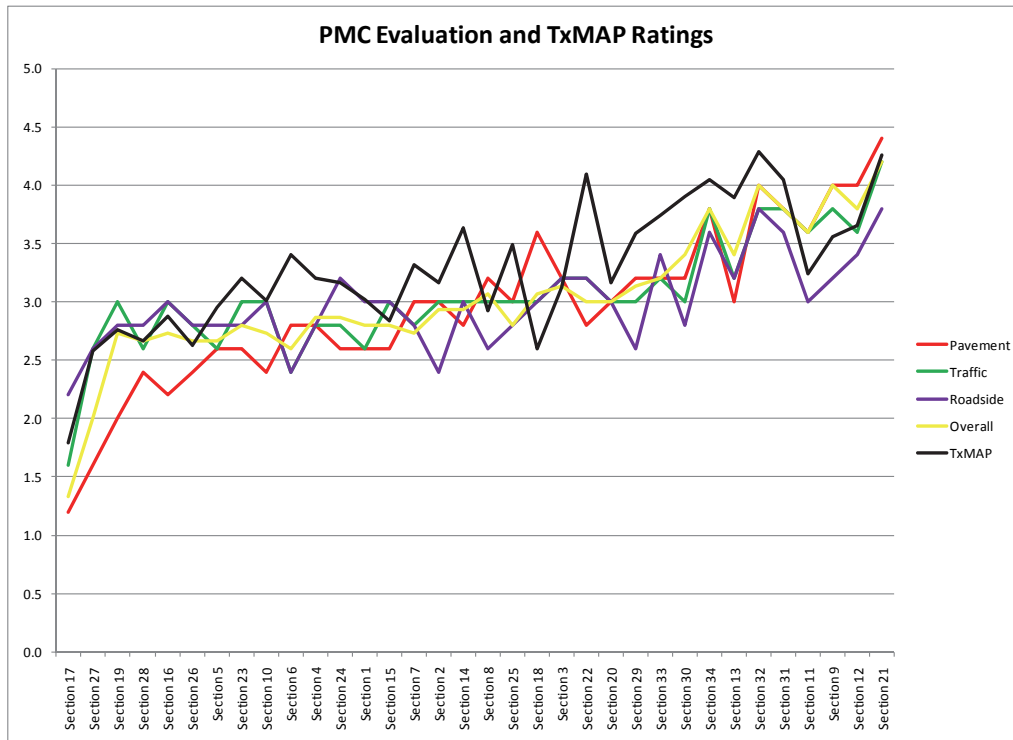


**Figure 7.2: TxDOT Evaluation versus TxMAP Ratings**

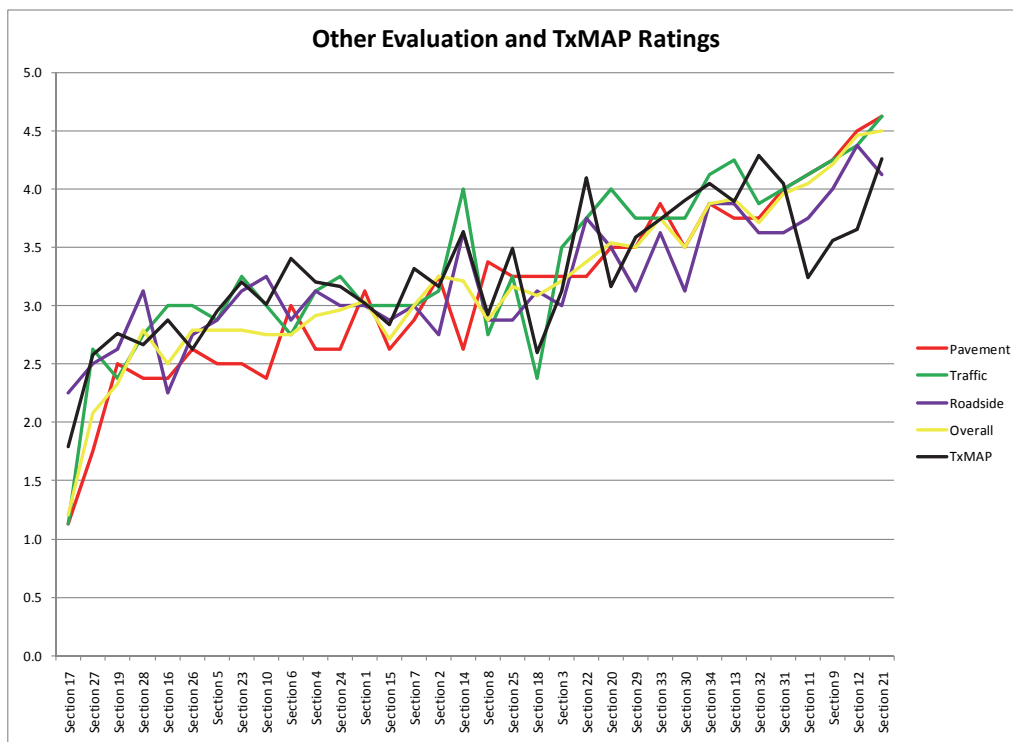


**Figure 7.3: Experts Evaluation versus TxMAP Ratings**





**Figure 7.4: PMC Evaluation versus TxMAP Ratings**



**Figure 7.5: Others Evaluation versus TxMAP Ratings**

To determine whether any of the ratings of the groups are good predictors of the TxMAP scores of the sections, a series of paired t-test were conducted with the data showed in Figures

7.1 to 7.5. In all cases the Overall ratings of the different sampling groups were compared with TxMAP scores. As before, the significance level was 5 percent so the critical t-value was 2.03. A positive difference indicates that the TxMAP scores are higher, on average, than the ratings of the particular group being tested. The detailed results are presented in Table 7.1.

**Table 7.1: Test of Hypotheses Comparing TxMAP scores with Group Overall Ratings**

<b>Group</b>	<b>Difference</b>	<b>t-statistic</b>	<b>Outcome</b>	<b>Comment</b>
Peer	0.34	5.90	Reject $H_0$	Significant difference
TxDOT	0.06	0.92	Cannot Reject $H_0$	No significant difference
Expert	-0.20	-2.50	Reject $H_0$	Significant difference
PMC	0.25	4.21	Reject $H_0$	Significant difference
Other	0.10	1.42	Cannot Reject $H_0$	No significant difference

As hypothesized before the testing, the ratings of the sections by the Peer and the PMC Groups are significantly lower than the TxMAP scores of the same sections. The average differences are 0.34 and 0.25 points below TxMAP scores for the Peer and PMC Groups, respectively. On the other hand, the ratings by the TxDOT and Other Groups are not significantly different from TxMAP scores. That means that the ratings of these two groups are good predictors of TxMAP scores. This is a particularly interesting finding that seems to indicate that TxMAP scores correlate very well with the condition of the pavement sections as perceived by the general road user (i.e., Other Group).

It is also interesting to know that the Expert Group was the only group whose ratings of the sections were, on average, above TxMAP scores. The average difference was 0.20 points above TxMAP scores (Table 7.1). This difference, although statistically significant, is quite small.

It should be emphasized once again that the objective of the MTSS was not to capture or to predict TxMAP scores; there are important differences in the way TxMAP scores are calculated and the way in which the survey was conducted. The objective of the survey was benchmarking only. Nevertheless, it is interesting to see that, for some of the sampling groups, the numerical values are quite close.

## **8. Conclusions and Recommendations**

The MTSS was conducted as part of a Peer Review of TxDOT's maintenance practices on October 6, 2010. During the survey, 24 individuals evaluated 34 pavement sections located east of Austin. This location was specifically selected because subgrade and traffic conditions make this area one of the most challenging areas in the state in terms of maintenance work. Six of the evaluators were invited from six different peer states: California, Georgia, Kansas, Missouri, North Carolina, and Washington. The objective of the rally was to benchmark the condition of the selected sections relatively to that of similar sections in the peer states. The MTSS was a great success and the numerical results of the rally are valid as demonstrated by the low variability within each section and the low variability within each sampling group.

Most of the remaining 18 evaluators were TxDOT employees of very different backgrounds and with different relation to this project. For this reason, they were grouped into three sampling groups: Expert, PMC, and Other. A fourth sampling group was formed by aggregating these three groups into one. This group was called TxDOT Group.

The ratings of the sections were analyzed per sampling group and per section. In all cases, the Peer Group rated the sections lower than the other sampling groups. The differences in the ratings were most significant when compared with the Expert Group and were the smallest when compared with the PMC Group. Although these differences are statistically significant in most cases they are small for practical purposes.

As an additional exercise, the ratings of the sections were also compared with the corresponding TxMAP scores. It was found that the Other and TxDOT Groups ratings were very good predictors of TxMAP scores.

In summary, the MTSS revealed that the Peer Group scored the sections lower as compared with TxDOT personnel. However, this difference was very low in most cases. It should also be noted that, on average, the scores of the Peer and the PMC Groups were very close. These two groups were indeed the most severe evaluators. There seems to be a correlation between the position within the organization and the severity of the rates.

## Recommendations

1. One element not included in the MTSS was evaluation of the sections by a group composed of public highway users (non-TxDOT evaluators). TxDOT may want to consider performing a similar survey that consists of evaluators selected from the general public, legislators, and people related to the transportation sector.
2. As an additional consideration, TxDOT could conduct a survey with the general public in different regions or districts to help gain knowledge about public expectations regarding maintenance conditions on different routes and in urban vs. rural locations. This information could help establish maintenance guidelines or help determine which types of maintenance might be reduced during conditions of limited funding.
3. Although the MTSS was not intended to be a controlled experiment to evaluate the TxMAP rating system, it is interesting that the survey results of the combined TxDOT Group and the Other Group were comparable to the TxMAP ratings. A further study is suggested to determine if a simpler rating process could be implemented to allow evaluation of more miles of pavement.
4. Due to the size of the TxDOT roadway network, a larger sample of rated sections could give a better indication of conditions considering:
  - a. Functional class or route type;
  - b. urban or rural location; and
  - c. district and regional comparisons with statewide conditions.

It is recommended that a further study is conducted to determine the number of miles of roadway that should be evaluated to provide a statistically valid sample.

5. During the course of the Test Section Survey it was apparent, based on evaluator comments, that the absence of paved shoulders, high posted speeds, and narrow lanes on some routes affected the ratings given to the pavement, roadside, and traffic operations ratings. It is recommended that a further study is conducted to evaluate how horizontal and vertical geometry; design vs. posted speed; lane width; and presence of shoulders may affect perceptions of maintenance conditions.