CHAPTER 7 – EVALUATION AND DISCUSSION OF ALL MONITORING RESULTS

In Chapters 3 through 6, project data and test results were presented and observations were made. The purpose of this chapter is to highlight some of the data and observations and discuss their relevance.

OVERALL PERFORMANCE LEVELS

Table 12 summarizes product scores achieved from the three types of monitoring – Subjective Observations in Chapter 4, Objective Measurements in Chapter 5, and Onsite Physical Testing in Chapter 6. The sixth column of Table 12 is an average of the three scores earned and serves to rank product performance at the Seedskadee NWR. Since these numbers may imply a higher level of precision than actually existed, the products have been simply grouped, and four groups are apparent. The Lignosulfonate product with the highest score is in the first group, and the Mag/Lig and Caliber products are in the second group. The two enzyme products, TerraZyme and PermaZyme formed the third group, and Soil Sement ranking lowest was in the fourth group. Table 12 also shows relative initial cost, relative application rate, and relative in-place cost.

Table 12. Seedskadee monitoring summary.

Test Section	Product	Subjective Overall Average Score (x10)	Objective Measures Overall Rating (x10)	Physical Onsite Overall Normalized Rank	Product Ranking from All Monitoring	Relative Initial Cost ($/gal)	Relative Application Rate (gal/ 1000 CY for 5 inch depth)	Relative In-Place Cost ($/CY)
I	Terra-Zyme	50	76	61	62	High ($145)	Low (0.01)	Low ($1.49)
II	Ligno-sulfonate	62	86	74	74	Low ($1.30)	High (5.62)	Medium ($7.30)
III	Perma-Zyme	51	81	62	64	High ($98)	Low (0.01)	Low ($0.84)
IV	Soil Sement	45	72	51	56	Low ($3.09)	High (4.10)	High ($12.66)
V	Caliber	56	82	65	68	Low ($1.17)	High (7.20)	Medium ($8.42)
VI	Mag/Lig	60	87	63	70	Low ($0.85)	High (7.20)	Medium ($6.11)

The earlier project at Buenos Aires NWR had a somewhat different ranking. There, the Caliber product performed the best, Mag/Lig was in the second group, and all the other products fell into the third group. The surfacing materials used at Buenos Aires and Seedskadee were different but both were non-plastic materials. Table 13 provides other key parameters for comparing the two projects. Likely, no one product works best everywhere, and owners of unpaved roads should select dust abatement and stabilizer products based locale and on the characteristics of their proposed surfacing material rather than on claims made by any one manufacturer. There is a need for selection criteria to help designers chose what would be the most effective class of

Table 13. Comparison of general characteristics between two NWR stabilization studies.

General Characteristic	Buenos Aires NWR Project	Seedskadee NWR Project
Climate	Desert Climate with Monsoon Seasons	High Desert with Climactic Extremes
Traffic Level	Low: 8 to 25 Vehicles per Day	Low: 4 to 15 Vehicles per Day
Surfacing Type	Borrow Material	Surfacing Aggregate
Material Description	Coarse-grained Gravel and Poorly Graded Silty Sand	Coarse-grained Gravel and Poorly Graded Silty Sand
Maximum Size Gravel	1 ½ inch	¾ inch
Percent Fines Range	4% to 19 %	9% to 13%
Plasticity Index	NP	NP
Organic Matter Content	No Test Results	Very Low (0.4%)
Stabilization Depth	150 mm (6 in)	125 mm (5 in)
Product Application Method	Windrow Mixing	Tiller Method
Best Performing Product	Caliber	Lignosulfonate
Second Best	Mag/Lig	Mag/Lig
Third	Lignosulfonate	Caliber
Fourth	TerraZyme	PermaZyme
Fifth	Soil Sement	TerraZyme
Sixth	PermaZyme	Soil Sement

stabilizer products to use in a specific setting. A preliminary process for accomplishing this objective is proposed in Appendix G. Finally, it is important for owners to also find out how to use the selected product to achieve the best possible result.

PLASTICITY

The materials at both the Buenos Aires and Seedskadee projects were NP or very close to NP. Differences between laboratories in the results reported for Seedskadee were well within expected variability for AASHTO T89 and T90. There are two problems with a low PI. First is that fines create dust when there is nothing gluing them together. Plasticity is the glue, so a surfacing material that is NP but has lots of fines will be dusty. The second problem is that many of the stabilization products used are ineffective without adequate PI. In fact, the typical role that some of them play is to lower a high PI, and they tend to work best with a material that has a PI between 10% and 20%. The exception to this statement is the Lignosulfonate product with a PI of 6% that appeared to add plasticity. This product and the combination Mag/Lig product were the top two performers in the NP aggregate material at Seedskadee. At Buenos Aires, Mag/Lig ranked second and Lignosulfonate was third.

These two performance evaluation projects at Buenos Aires and Seedskadee have brought attention to the need for higher PI. One result of this finding is that CFLHD has increased the PI specification on some future projects.

SILT LOADING AND DUST

This section presents a comparison between objective dust observations and Silt Load Test results. The silt load was the amount of silt available, in ounces per square foot, in the loose surface material that would blow away as dust. In Figures 29 through 34, the agreed dust ratings for each product were compared over the two-year monitoring period to the average silt loading results for the same time period. Generally, the figures show that when there was a lot of silt available, more dust was observed. Since the trends generally moved together, the results of the two tests generally validated each other. Figure 33 that graphs the Caliber product appears to deviate from the general trend at the 20-month event. It must be remembered that the silt test result was an average of only four discreet samples where as the agreed dust rating considered the entire section.

It is interesting that, in all the graphs, the 23-month silt loading result was better than the earlier 11-month and 20-month results. The amount of loose dust size particles was less at the end of the project than earlier, and there was also less dusting observed. Since the weather was dry for all of the events, this phenomenon was curious. One possible explanation was that as the surfacing material broke down from traffic and weathering, the finest particles blew away leaving larger raveled material over the road surface. Under the raveled material was the remaining stabilized surfacing material. Over a longer monitoring period it was possible that silt load results could cycle. The silt test results from the Buenos Aires project did not show a similar consistent up-swing at the final monitoring event, nor were there any objective dust ratings available to compare dust and silt load. Any comparison was further hampered by the fact that only two samples were taken for the silt test from each section at Buenos Aires, so the resulting average of only two silt load values was necessarily less precise than Seedskadee’s four values.

Gradation analysis was performed as part of the Silt Load Test and the summary results are included at the end of Appendix E. The material gathered was only the loose material at the surface, but a comparison to the gradation of the full depth material as initially constructed in October 2004 shown in Tables 5 and 6 in Chapter 3 was interesting. A general observation was made that for the loose surface material, more material passed through the larger sieves (19-mm (3/4-in) down to 2-mm (No. 10) sieves) and less material passed through the 0.425-mm (No. 40) sieve down to the 75 µm (No. 200) sieves. This indicated first that the larger aggregate was likely

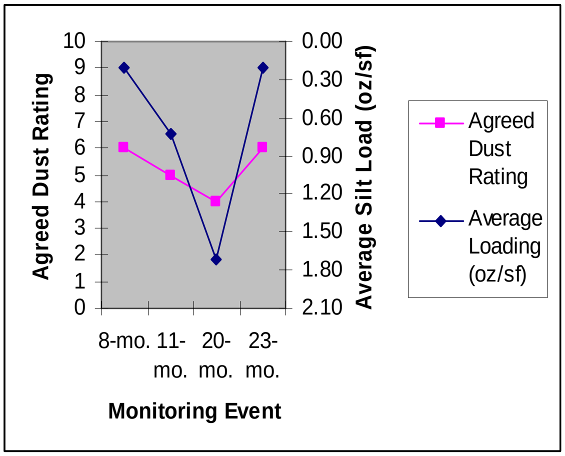

Figure 29. Graph. TerraZyme dust and silt load comparison.

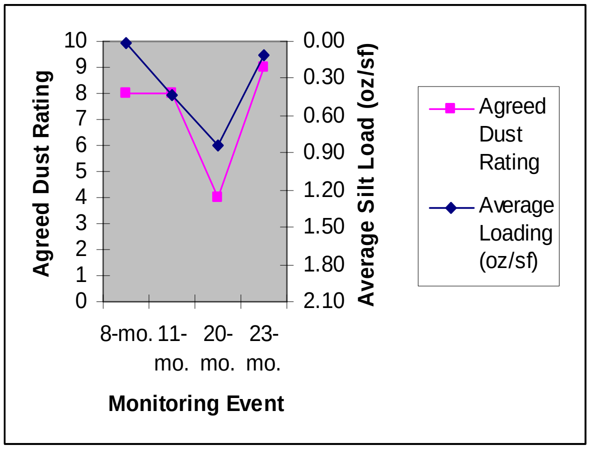

Figure 30. Graph. Lignosulfonate dust and silt load comparison.

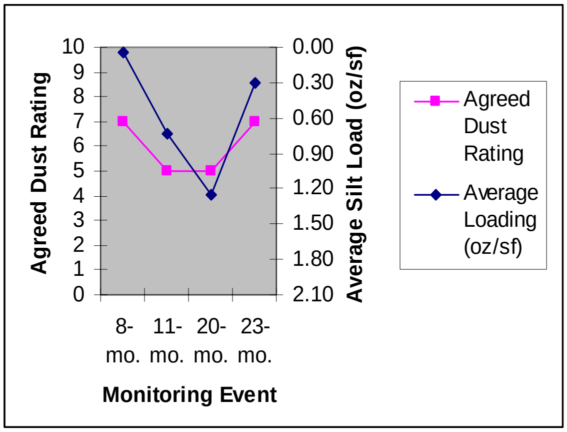

Figure 31. Graph. PermaZyme dust and silt load comparison.

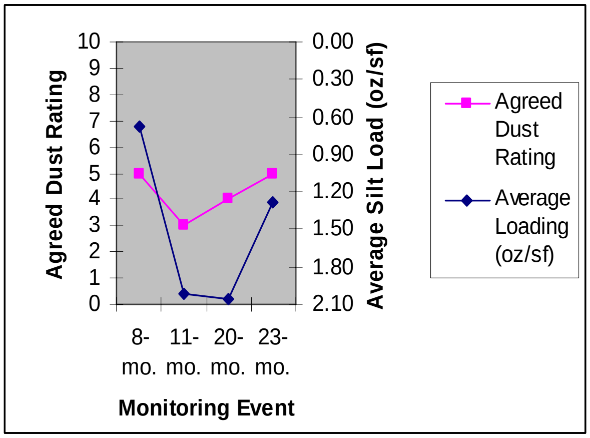

Figure 32. Graph. Soil Sement dust and silt load comparison.

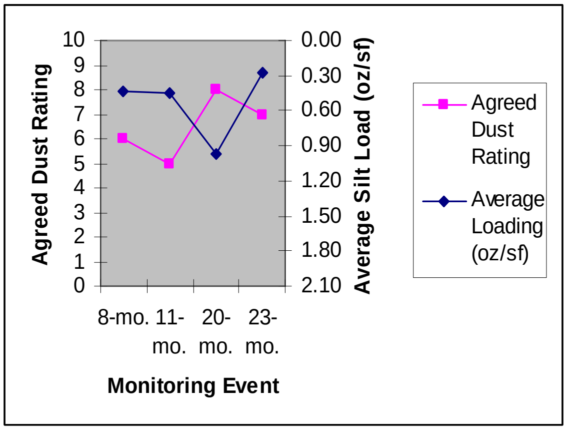

Figure 33. Graph. Caliber dust and silt load comparison.

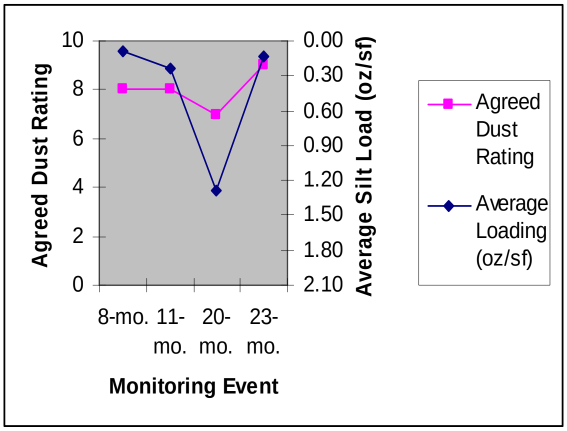

Figure 34. Graph. Mag/Lig dust and silt load comparison.

breaking down over the two years of monitoring. It also indicated that something happened to the smaller particles. In fact, material passing the 75 µm (No. 200) sieve ranged from 9.3% to 12.7% during installation in September 2004, and dropped throughout the monitoring period to a range of 0.8% to 4.2% in the surface layer after two years. This shows that binder material may have been lost to erosive forces such as traffic and climatic conditions. The loss of binder material over time may be one explanation for why the average silt loading for all the products actually decreased (improved) the last month of monitoring.

PRODUCT DISSIPATION

At one point during monitoring of Seedskadee, there was a concern that the stabilization products may have leached out in the first few months. The exception may have been the Lignosulfonate product that is very viscous and was ranked as the best performer at Seedskadee. One way of testing this hypothesis would have been to have a control section with the same traffic but without any stabilization product applied. This was not possible to accomplish on the Seedskadee NWR. Another solution might have been to run before and after tests to determine the amount of stabilization product in material samples.

Since these tests were not done, no conclusive statement can be made about product leaching. However, it should be noted that product leaching was never observed. Additionally, when sweeping up samples for the silt test, in all cases the hard underlying surface appeared to still contain product throughout the two-year monitoring period. A reasonable question for future studies was how can an agency quantify the amount of product present at the beginning of the study and then again quantify it at the end. Currently, the FHWA has only a method specification for application of stabilizer products.

COMPARISON OF CBR VALUES BETWEEN TWO PROJECTS

The average DCP-derived CBR values at Buenos Aires ranged from 57 to 87 for the six products. DCP tests at Buenos Aires were only performed during the last three monitoring events. At Seedskadee, the average in-situ CBRs for its last three events ranged from only 42 to 69. It appears from these numbers that the Buenos Aires material in general was more stabilized and would likely prevent wash boarding, rutting, and potholing for a longer period of time than at Seedskadee.

Since the degree of stabilization was somewhat greater at Buenos Aires, one question asked was whether or not it was an effective use of funds to try to stabilize the crushed aggregate surfacing at Seedskadee. A conclusive answer cannot be stated because there was no control section free of stabilizer product available for comparison at either project. It must be noted however that at Seedskadee, product was visible in all the sections underneath the loose raveled surface material. The significance of that observation is that the stabilizer products did not appear to leach out over the two-year monitoring period.

Each section at Seedskadee behaved somewhat differently in regards to loose aggregate. In one section, loose aggregate was spread evenly over the width of the road. In another, there were wheel paths that were clear of loose aggregate. It was suggested that defined wheel paths might indicate that the rocks were thrown aside from traffic but that no new rocks were breaking loose and moving to the surface. Since the traffic was approximately the same through all the sections, it seems plausible that areas where loose rocks were spread uniformly across the roadway were probably in less stable condition than areas where wheel paths had formed. Where rocks were spread uniformly across the road, it was thought that more rocks were continually coming to the surface as binder broke down. Thus, a hypothesis to be tested was that in sections with loose aggregate spread across the road, the CBR values derived from DCP tests would be lower than in sections where wheel paths have formed. And since each DCP test site was very near to the sampling locations for the Silt Test, and since the Silt Test provides a total mass in grams for the sample collected, it was thought that the total mass of the sample might correlate with the CBR derived from the DCP test performed in the wheel path. Even though there was a large amount of data, there was very little correlation between Silt Test sample size and in situ CBRs.

FULL DEPTH STABILIZATION AND WASH BOARDING

Surfacing materials for both of the projects were stabilized to their full depth. This was 150 mm (6 in) at Buenos Aires and 125 mm (5 in) at Seedskadee. This procedure may have been key at both projects for minimizing wash boarding. However, because there was no true control section that could be constructed at either project, there is no proof that full depth stabilization actually prevents wash boarding. At Seedskadee where recurring wash boarding has typically been the most difficult road maintenance problem, full depth stabilization of the aggregate surfacing worked very well. It is possible too, that the 19-mm (3/4-in) minus specified aggregate surfacing alone may have alleviated washboarding whether or not stabilizers were applied. It should be noted, however, that under the loose raveled surface, stabilizer product was still visible even after two years of monitoring at Seedskadee. Full depth incorporation of the stabilizer products was also successful in largely preventing potholing and rutting at both projects.

Since full depth stabilization was considered, by the evaluation teams on both projects, to be very important in preventing potholing, a discussion follows of the three main methods of incorporating stabilization products into the full depth of surfacing material. Each of the three methods was considered for both Buenos Aires and Seedkadee projects. They are 1) the Windrow Method, 2) the Tiller Method, and 3) the Pug Mill Method. On the Buenos Aires project, forms of the Windrow Method were used, and on the Seedskadee project the Tiller Method was used. The Pug Mill Method was not selected as the preferred method for either of the projects.

Windrow Method

This method involves windrowing the surfacing material to one side, spreading a layer of material, spraying it with the diluted product and water to achieve optimum moisture, blade mixing, and then repeating this process until the specified depth is achieved. The finish bladed roadway is then compacted with a pneumatic roller. This method is easy to do and requires equipment that is generally readily available – a grader, water truck and/or distributor truck, and roller. The layering process assures full depth penetration. Mixing with a grader, however, does not assure uniform distribution across the roadway. The roadway actually needs to be greater than 3.7 m (12 ft) wide to allow room to blade the material back and forth. It is difficult to achieve the correct application rate, and the quality of the job depends on the quality the grader operator can produce.

Tiller Method

Other names for this method include pulverization method and in-place full-depth reclamation. The roadway surfacing material is placed and compacted to the specified depth. Water is applied with a water truck to bring the surfacing material to its optimum moisture content. The stabilizer product is applied through the reclamation machine or, if too viscous, through a distributor truck immediately preceding the reclamation machine. The reclamation machine picks up the surfacing material to the specified depth (125 mm (5 in) on the Seedskadee project) mixes it with the stabilizer product, and lays it back down. The roadway surface is then finish bladed and compacted with a roller. This method uniformly mixes the product with the surfacing material and allows this mixing to occur at the project site unlike the pug mill method. One drawback is that when the product is sprayed on the compacted roadway just in front of the reclamation machine, there is a potential for the product to runoff onto the vegetation at the side of the road before it is picked up and mixed. Additionally, reclamation machines typically cannot make tight turns and therefore cannot be used in tight areas.

Pug Mill Method

This method was not used on either the Buenos Aires project or the Seedskadee project though it was strongly considered. In this method, the stabilization product is introduced into a pug mill at the material production site, and then the treated material is hauled to the project site. This is a controlled process that produces a uniform mixture of product and surfacing material. The equipment needed includes a pug mill, grader, and roller. A water truck is not needed. A spreader box could be used. One limitation of this method is that some products, even after being mixed with water, are too viscous to introduce into the pug mill. Lignosulfonate is one such product. The production rate is slower than with other methods because discrete batches are produced that then need to be hauled to the project site and spread before they set up. A risk of this method is that hauling delays have the potential to cause the material to set up, or react with the stabilization products before reaching the job site.

METHODOLOGY COMPARISON - SUBJECTIVE OBSERVATIONS AND OBJECTIVE MEASUREMENTS

As discussed at the end of Chapters 4 and 5, the subjective comparative system and the objective measurement system each have their strengths.

One recommendation from the Buenos Aires study on dust abatement and stabilizer products was to further refine assessment methods to track performance through time and to strengthen the objectivity and therefore defensibility of the method. The major strength of the subjective comparative inspection system developed for the Buenos Aires project was its ability to recognize subtle differences in performance between the products. What it could not do was track performance trends over time. Thus the evaluation team developed an objective measurement methodology, also based on a zero to ten scale, that attempted to define worst case to best case scenarios for each parameter of dust, washboarding, raveling, rutting, and potholing.

The question was raised early in the monitoring at Seedskadee whether the subjective rating system should be continued considering the 11-point objective measurement system had been developed. It was decided to continue the subjective system through the Seedkadee project, and to discuss both systems in the final report. The remainder of this section evaluates and compares the two monitoring systems.

The relative standings of the products using subjective observations are shown in Figure 35, and the standings using Objective Measurements are shown in Figure 36. At first glance it appears that the objective rating system using field measurements gave much higher scores than the visual comparative system. It must be remembered that the goals and methodologies of the two systems were very different.

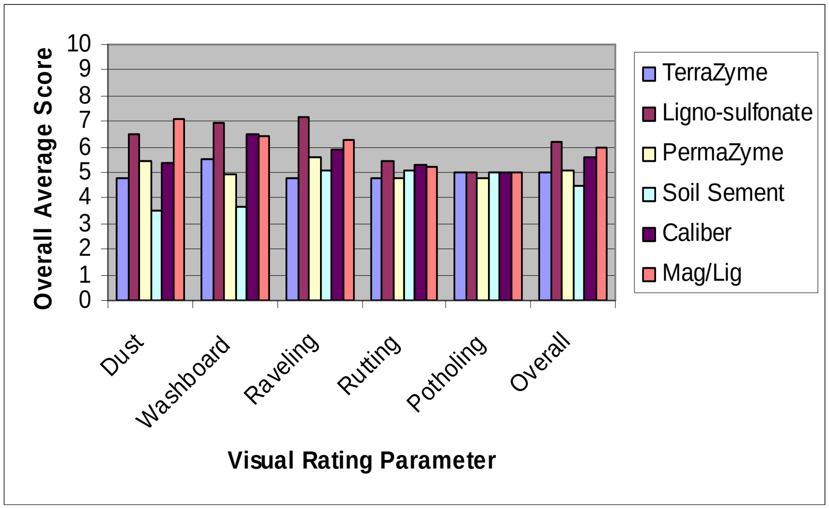

Figure 35. Plot. Relative product standings from subjective observations.

In the subjective comparative system, a different section's product served as the baseline for each monitoring event, and the remaining products were compared to it. Thus, most of the scores hovered around a score of 5 - the score of the baseline product.

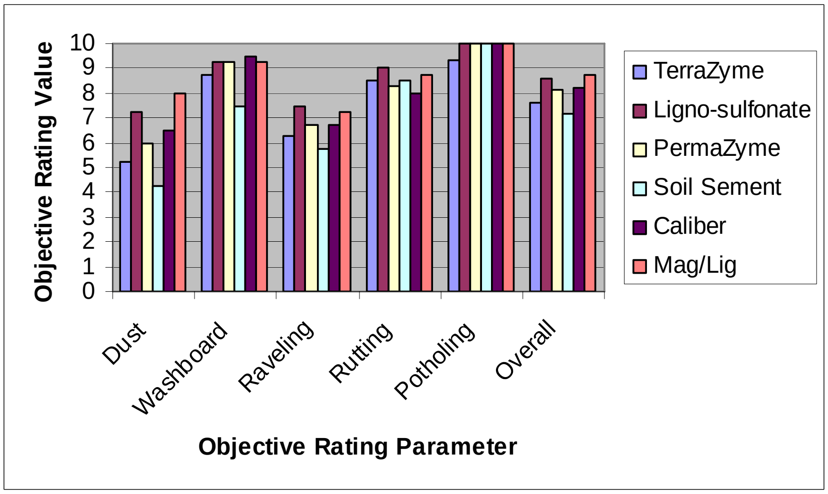

Figure 36. Plot. Relative product standings from objective rating system.

In the objective system, however, averaged field measurements were converted to ratings using descriptive tables from the Appendix B Objective Rating System. Thus the ratings were dependent both on how the descriptive tables were set up and on the specific 7.6-m (25-ft) long areas chosen for measuring. These locations were set up in the Appendix A - Monitoring Order and Mileposts Plan prior to any monitoring to avoid bias in choosing measurement areas.

To compare the results from the two methodologies, a sample set of data consisting of the overall average scores of the products for each parameter using the subjective method and the overall ratings of the products for each parameter using the objective measurement method was selected. If the correlation between the results from the two systems were very good, then in the future either the visual comparative system could be ignored as being much more subjective, or the objective system could be ignored because it required more time and effort. In Figures 37 through 41, objective ratings on the y-axis for each parameter are plotted relative to subjective comparative scores on the x-axis. The six data points on each plot represent the six different stabilization products. A simple regression analysis yielded a best fit line through the data points, and a correlation value, R^2, is also shown. The correlations vary from excellent to poor depending on the parameters, and are discussed below.

Dust Correlation

The Figure 37 correlation for dust results shows excellent correlation, at $R^2 = 0.9696$, between results from the objective measurement monitoring method and the subjective comparative method. This is not surprising because in both methods the evaluation team used visual criteria to estimate the level of dust even though the objective method offered more definitive criteria. Instead, they together agreed on the appropriate objective rating using the Appendix B criteria. This step was done in conjunction with the subjective comparative scoring of dust generation.

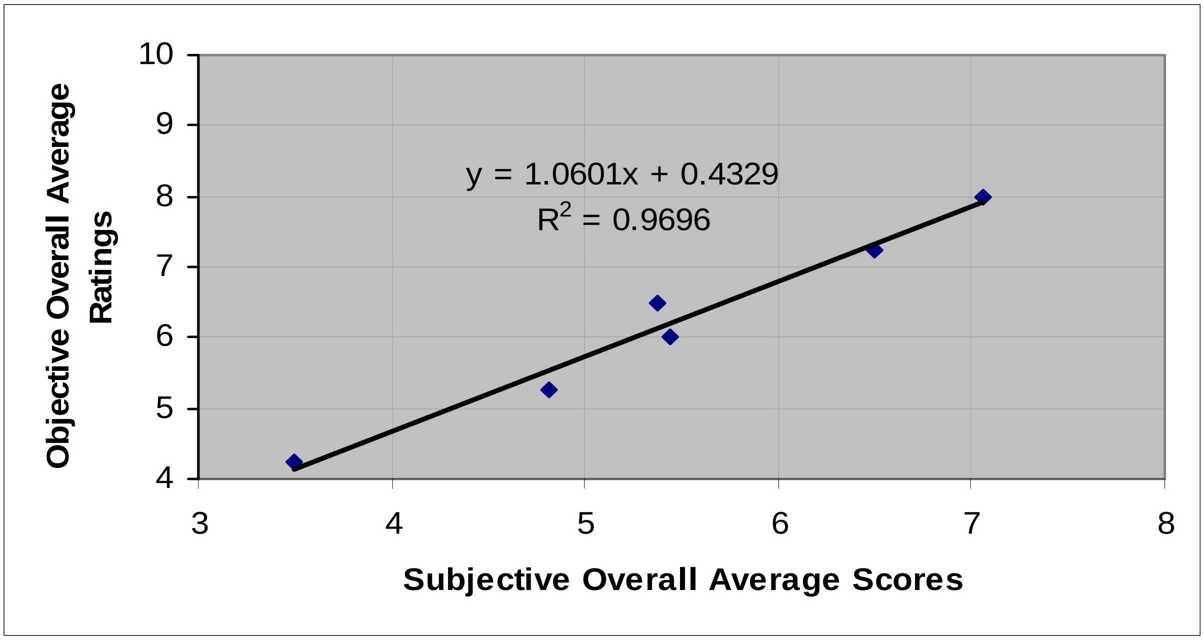

Figure 37. Plot. Correlation of results for dust.

Washboarding and Raveling Correlations

For wash boarding, Figure 38 shows a surprisingly good correlation at $R^2 = 0.7107$. This is also true for raveling shown in Figure 39 with a correlation value of $R^2 = 0.8179$. The reason this degree of correlation is surprising is that the pre-selected 7.6-m (25-foot) long monitoring areas were randomly selected such that information was not recorded from some of the perceived

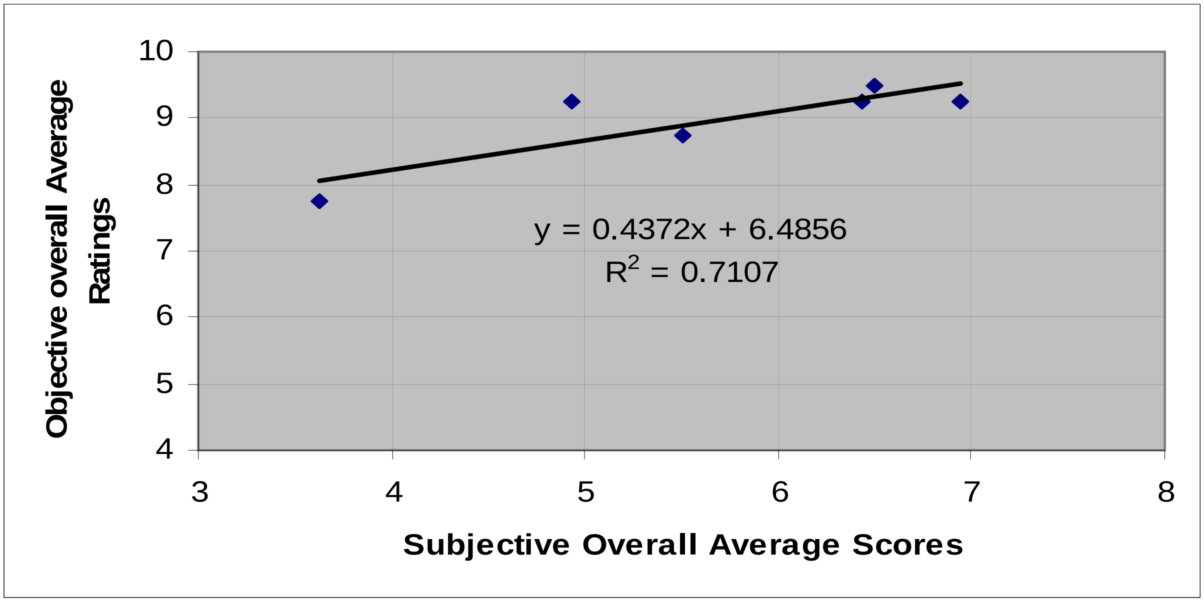

Figure 38. Plot. Correlation of results for washboarding.

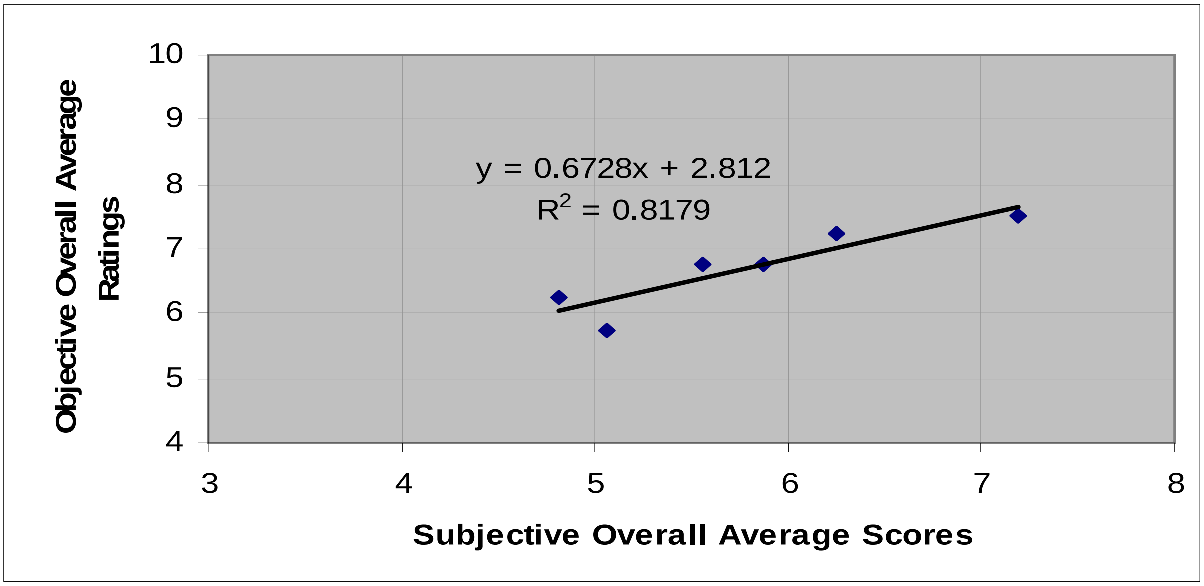

Figure 39. Plot. Correlation of results for raveling.

poorer performing areas within a particular section. This degree of correlation adds confidence that either the subjective method or objective method of evaluating these parameters could be chosen.

Rutting and Potholing Correlations

The correlation between the two methodologies using data from rutting and potholing, however, is quite poor. For the parameter of rutting, Figure 40 shows an $R^2 = 0.1362$, and for potholing Figure 41 shows $R^2 = 0.04$ which is essentially no correlation. These low R^2 values do not

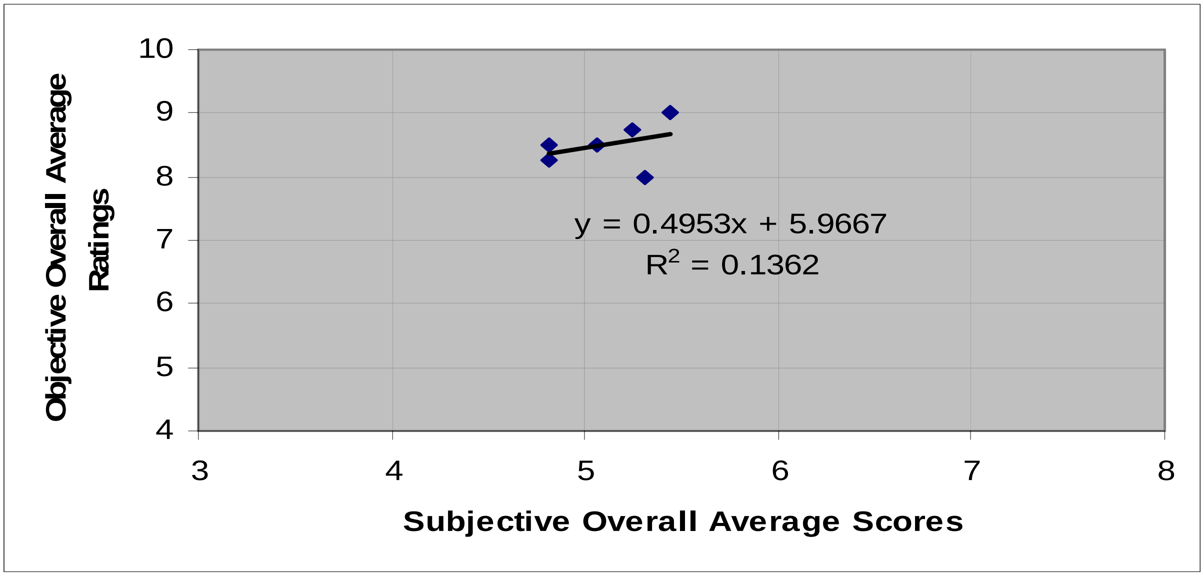

Figure 40. Plot. Correlation of results for rutting.

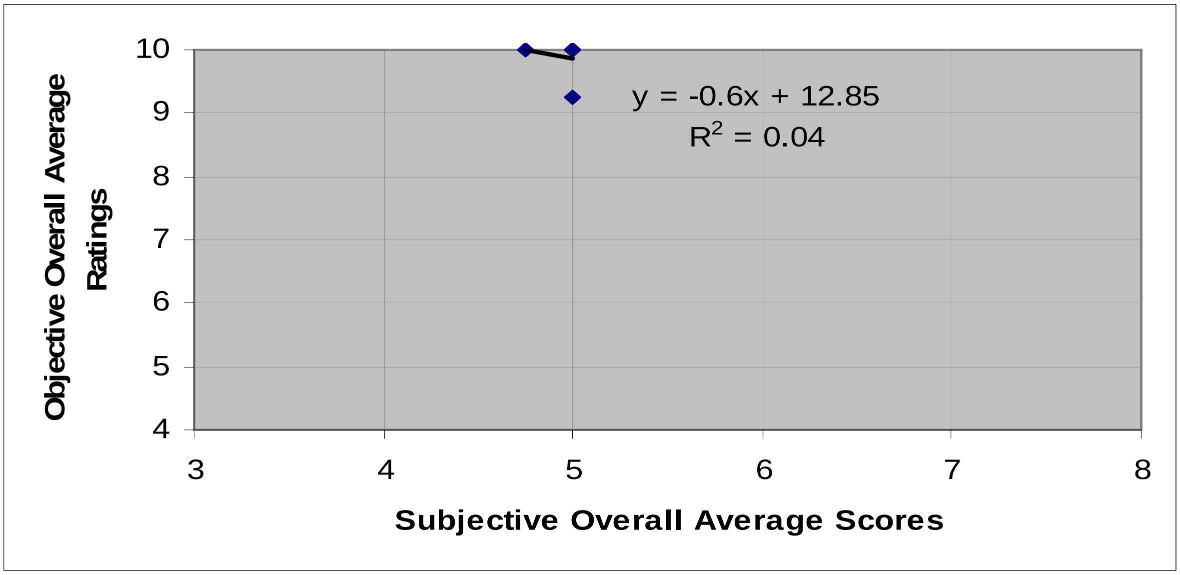

Figure 41. Plot. Correlation of results for potholing.

necessarily mean that the two monitoring methodologies do not give good answers, but rather the resulting values in the data set do not correlate. A data set that reported on a greater spread of rutting values, or one that measured significantly more potholes would have improved the likelihood of a better correlation.

When the evaluation team was subjectively comparing the sections for rutting and potholing, they found very little differences between the sections, and this resulted in average scores that stayed very close to the score of five that was always assigned to the baseline section. That all the scores from the subjective method are close together hampered the use of linear regression for statistical analysis, especially when the sample consisted of only six data points. This is likely the primary reason for the low correlation between monitoring methods for the parameters of rutting and potholing. Other factors also influenced the data. One of the sections had an extremely reduced sampling area because a road section repaired after weather related damage did not contain stabilizer product and therefore were excluded from monitoring. Another factor was that for the parameter of potholing, the planned judging criteria for an anticipated numerous count of potholes did not fit the circumstances encountered of only three potholes in the entire study.

In summary, it appears that the two subjective and objective monitoring systems compare reasonably well. Each has it own strengths and weakness, and future monitoring efforts would benefit by clearly defining the desired goals of the monitoring effort before choosing a monitoring methodology. If time and resources are limited, then the subjective method still could be used to distinguish levels of performance. However, for more justifiable and defensible results, the objective method, even though more time consuming and data intensive, would be the better choice.