age, traffic, climate, and so forth. This is done by comparing each treated section at a site with the control section at the same site. Instead of comparing the means for many sections within different treatment groups, one analyzes the mean difference between each control section and each of the corresponding treated sections. An appropriate statistical test for this purpose is a paired difference test, often referred to as a paired *t*-test.

In an analysis of a long-term effect of a maintenance treatment on IRI, the difference of interest at each site is between the IRI of the treated section, obtained from the most recent profile measurement, and the IRI of the control section, obtained at the same time. Because four comparisons are made (each of the four maintenance treatments versus the control), the significance level, α, of the individual tests should be selected so that $(1 - \alpha)^4$ is equal to the desired overall level of confidence. For four tests to yield a 95% overall level of confidence, the required α is 0.01274. The results of the four paired difference tests are shown in Table 2. The results indicate that of the four maintenance treatments in the core SPS-3 experiment, only the thin overlays had long-term IRIs significantly different from those for the corresponding control sections. These data cover a range of time from 2.0 to 10.4 years, with an average of 6.4 years.

The cumulative frequency distributions of the long-term IRIs are shown for the four treatment groups and the control group in Figure 1. As the plot shows, the median (50th percentile) long-term IRIs are all very close for the slurry seal, crack seal, control, and chip seal groups, whereas the median long-term IRI of the thin overlay is slightly less than those for the remainder of the treatments. However, looking at not just the 50th percentile values but also the complete distributions, it appears that the distributions for not only the thin overlay but also the chip seal are fairly consistently to the left of the distribution for the control, especially at IRIs above about 1.75 m/km. The distributions for the slurry seal and crack seal, on the other hand, follow the distribution for the control very closely. These results reinforce the impression that the chip seal treatment was the second most effective, after the thin overlay treatment, in reducing long-term roughness.

Influences of Other Factors

The second part in the analysis of long-term treatment effects on IRI is to test for significant effects of other factors, that is, time, traffic, climate, pavement strength, and pretreatment IRI. The steps in this analysis are the following:

1. For each site, the change in IRI from the first posttreatment measurement to the most recent measurement is calculated for the control section.
2. For each treated section at the same site, the change in IRI for the treated section at the same site, from the first posttreatment measurement to the most recent measurement, is calculated. The IRI data used should be for dates for which measurements are available for both the control and the treated sections. In a few of the cases in which a linked GPS site serves as the control, there is a difference of a month or two between the posttreatment measurement dates.
3. The difference between the change in IRI for the control section and the change in IRI for each of the treated sections is calculated for each site.
4. For each treatment type, the slope of the difference calculated in Step 3 with respect to each of the factors of interest is analyzed. This may be done by an *F*-test (or, equivalently, by a *t*-test). A factor is concluded (with a 5% chance of error) to have a significant influence on the difference in long-term IRI performance of the control section versus that of the treated section if the calculated *F*-value (the ratio of the regression mean square to the error mean square) exceeds the upper 5% of an *F*-distribution with 1 and $n - 2$ df, where n is the number of sites used in the test.

Although testing for the significance of the effects of different factors applies a statistical test to the linear regression of performance differences with respect to each of the factors, it does not imply a presumption that those relationships are better described by linear regression than by nonlinear regression. For the detection of significant effects of different factors, the question of interest is not whether a linear versus a nonlinear relationship exists but whether any relationship exists, that is, whether the slope of the change in the long-term IRI with respect to the factor of interest is or is not significantly different from zero.

The analysis of long-term treatment effects described here represents the first two steps in building models for determination of the effects of maintenance treatments on pavement performance: (*a*) determining which treatment types significantly affect long-term performance and (*b*) determining which factors (traffic, climate, pretreatment condition, etc.), if any, significantly influence how much effect the treatment types have on long-term performance.

Once the significant independent variables have been identified, the next step in the model-building process—which was beyond the scope of this study—would be to select the model form that best reflects the effects of the independent variables, including nonlinear effects and interaction effects, if any.

TABLE 2 Analysis of Long-Term Effect of SPS-3 Maintenance Treatment on IRI

	IRI (control versus treatment), m/km			
	Thin Overlay	Slurry Seal	Crack Seal	Chip Seal
Mean difference	0.32	0.02	0.02	0.07
n	58	55	57	45
SD	0.65	0.56	0.44	0.63
$T_{\alpha/2, n-1}$	2.57	2.58	2.57	2.60
Confidence interval lower limit	0.10	- 0.17	-0.13	-0.18
Confidence interval upper limit	0.54	0.22	0.17	0.31
Significantly different than control	yes	no	no	no

NOTE: SD = standard deviation (i.e., square root of the variance); $T_{\alpha/2,n-1} = \alpha/2$ percentile of the *t*-distribution with $n - 1$ df.