

# Distress Survey Methodology of the New York State Thruway Authority's Pavement Management System

D. A. GRIVAS, B. C. SCHULTZ, AND C. A. WAITE

A methodology is presented for determining distresses on asphalt and portland cement concrete pavements in an objective and reliable manner. Individual distress magnitudes are established using linguistic scales that consider distress severity and extent along nominal lengths of pavement. Emphasis is placed on generating both the quality and the quantity of data needed for pavement management purposes. The methodology is implemented in the form of a distress survey technique applied to the pavement system of the New York State Thruway. Data recording and handling are automated through the use of laptop computers. The survey technique is taught annually to nontechnical personnel in a comprehensive training program. The results of statistical techniques used to evaluate rater performance and repeatability of survey data are presented and discussed. It is concluded that the implemented distress survey is a reliable procedure that produces distress data at acceptable levels of repeatability.

Distress surveys are widely used techniques to evaluate and monitor pavement condition over time. They provide information needed to characterize pavement surface condition and establish causes of deterioration. Distress data also influence the scope and cost of the work required to restore the integrity of pavement structures.

There are no standard procedures for conducting distress surveys at the present time. Available methods vary widely in both objectivity (data quality) and cost. Such methods include manual field mapping, windshield surveys, visual surveys with automated data logging, rating of previously collected images, and automated real time data collection (for some specific distresses).

In general, there is a significant trade-off to be made between objectivity and cost. Objective survey methods can provide detailed, consistent data, but are time consuming, expensive to perform, and require extensive training. In contrast, subjective methods tend to be fast and inexpensive, but the results are commonly inconsistent and nonspecific. Thus, transportation agencies are faced with the challenge of developing the most objective distress survey technique they can currently afford.

The tasks involved in the development and implementation of a manual yet objective distress survey with automated data entry are described. The survey is part of the Pavement Management System (PMS) of the New York State Thruway Authority (NYSTA). Emphasis is placed on securing the appli-

cability of the survey technique to thruway conditions, and on providing the quality and quantity of data needed for PMS purposes. The survey technique assesses and records pavement distresses in terms of their location, severity, and extent along nominal lengths of the thruway pavement system. Collected distress data are used to achieve a variety of pavement management objectives, namely (a) determine causative factors and identify the scope of needed work, (b) define uniform lengths for project design purposes, (c) establish cost estimates of maintenance and rehabilitation actions, (d) determine pavement deterioration rates, and (e) develop the annual pavement maintenance program.

## DISTRESS SCALE FORMULATION

### NYSTA Pavement Structures

The New York State Thruway consists of 559 mi (2,600 lane-mi) of Interstate type highway. The system was originally constructed entirely of portland cement concrete (PCC) pavement structures (1). As pavements aged, they were overlaid with asphaltic concrete (AC). Several short segments of pavement were reconstructed as full-depth asphalt concrete or portland cement concrete, in accordance with newer standards for slab length and joint treatment. For the purposes of the PMS distress survey, thruway pavements are classified as concrete, overlaid, or shoulders.

Concrete pavements were constructed of 9-in.-thick, wire-mesh-reinforced PCC. Normal slabs are 100 ft long, and have expansion-type load transfer devices. The slabs were placed on 12 in. of granulated subbase course, with no provision for subsurface drainage. Distresses on the surface of concrete pavements characterize damage of individual slabs and damage at the joints between slabs.

Overlaid pavements presently constitute approximately 80 percent of the entire system. They have a composite cross section in the vertical direction, with asphaltic concrete on the top and the original PCC pavement underneath serving as a base. In most cases, underdrains were added when the original PCC pavement was overlaid. Distresses on the surface of such pavements reflect damage of the upper asphaltic layer or the underlying concrete pavement, or both.

Shoulders were originally constructed of chloride-treated granular material or sod. All have since received at least a thin asphalt overlay, while some have been fully reconstructed with asphaltic concrete. Shoulder distresses reflect the con-

D. A. Grivas and B. C. Schultz, Civil Engineering Department, Rensselaer Polytechnic Institute, Troy, N.Y. 12180. C. A. Waite, New York State Thruway Authority, 200 Southern Boulevard, P.O. Box 189, Albany, N.Y. 12209.

dition of the shoulder-lane connection, and the integrity or overall condition of the shoulder surface.

### Distress Components

The AASHTO pavement design guide (2) outlines the minimum pavement distress information needed to make appropriate rehabilitation decisions. Required information for pavement distress description is as follows:

- **Type.** Distress types are distinguished primarily by their causes and location on the pavement surface.
- **Severity.** Distress severities are distinguished by degree of deterioration; they are usually categorized into distinct levels, such as Small, Medium, and Large.
- **Extent.** Distress extent measures the amount of a certain distress type and severity combination per unit area or length of pavement.

Important factors contributing to distress survey objectivity have been identified by Sime and Larsen (3). They include the following:

- Use of descriptive distress definitions (rather than simple subjective numerical rating scales),
- Measurement of distresses (wherever possible), and
- Development of training programs for raters.

The present distress survey has considered both the information and the objectivity requirements stated earlier. Thus, descriptive definitions (linguistic scales) are used to determine severities. Extents are determined in terms of areas and lengths (overlaid pavement), or exact counts (concrete pavement). Actual physical measurements are made at controlled sections to establish the true values of distress magnitudes. Finally, an extensive training program for raters that includes 2 weeks of repeatability testing is conducted annually.

### Distress Types

Because a pavement distress survey represents an intensive data collection activity, a limited number of distress types can be efficiently recorded. It is therefore critical to monitor only those distresses that provide significant input to maintenance and rehabilitation decision making. When evaluating distress types for inclusion in the NYSTA PMS Distress Survey, two primary criteria were used:

- Knowledge of the specific distress and its characteristics (severity, extent) is necessary for maintenance decisions or performance evaluations; and
- Distress characteristics are clearly visible from a vehicle moving slowly (5 mph) on the shoulder.

From an operational standpoint, it was necessary to limit the number of distress types collected so as to render the data collection effort appropriate for a three-person team. The three-person limit (in addition to the driver) was imposed by the number of window seats available in a conventional van.

Extensive consultation with thruway maintenance experts determined that the distress survey should use 14 distress types to assess pavement and shoulder condition. Six of these distresses correspond to overlaid pavements, six to concrete pavements, and two to shoulders. The specific distress types collected are presented in Table 1. Each distress type is assessed using rigorously defined scales and rules of thumb, such as the example presented in Table 2. A detailed description of the factors considered during the development of the distress scales was provided by the AASHTO pavement design guide (2) and by Schultz and Grivas (4).

### DISTRESS SURVEY TECHNIQUE

#### Determination of Distress Ratings

Distress magnitudes are assessed visually; no physical measurements are made during the survey. Ratings are assigned through the use of linguistic scales, which classify distresses on the basis of easily observable characteristics (e.g., relative size, location, orientation, and previous repairs). An example

TABLE 1 NYSTA DISTRESS TYPES

OVERLAID PAVEMENT	CONCRETE PAVEMENT
1. Centerline Cracking	1. Loss of Transverse Joint Sealant
2. Longitudinal Cracking	2. Transverse Joint Spalling
3. Surface Defects	3. Transverse Joint Faulting
4. Rutting	4. Longitudinal Joint Spalling
5. Transverse Cracking	5. Slab Surface Defects
6. Edge Cracking	6. Slab Cracking
7. Shoulder Defects	7. Shoulder Defects
8. Lane/Shoulder Displacement	8. Lane/Shoulder Displacement

TABLE 2 DISTRESS SCALE AND RULES OF THUMB FOR RATING CENTERLINE CRACKING

Severity	Extent	Description	Rating
None	—	No centerline cracking.	N
Small	Local	Tight centerline cracking or sawcut patches on the centerline occur locally in the section.	SL
	General	Tight centerline cracking or sawcut patches on the centerline occur generally throughout the section.	SG
Medium	Local	Open centerline cracking occurs locally in the section.	ML
	General	Open centerline cracking occurs generally throughout the section.	MG
Large	Local	Alligator cracking, generally along the centerline, occurs locally in the section.	LL
	General	Alligator cracking, generally along the centerline, occurs generally throughout the section.	LG
Total	Local	Centerline cracking with fill-type patches and/or significant material loss occurs locally in the section.	TL
	General	Centerline cracking with fill-type patches and/or significant material loss occurs generally throughout the section.	TG

#### Rules of Thumb for Centerline Cracking Assessments

- Open or alligator cracks that have been completely sealed are rated as tight cracks.
- Cracks showing significant material loss that have been sealed are still rated as TOTAL.
- Partially sealed cracks are rated as though they were not sealed at all.
- Cracks in squared-off patches are rated like other pavement cracks.
- Definition of local cracks extend less than 1/3 of the linear distance.
- Definition of general cracks extend more than 1/3 of the linear distance.
- Significant material loss is defined as holes that are at least 6" wide (measured perpendicular to the crack).

of the developed scales is presented in Table 2 for the case of centerline cracking. This technique of describing distress scales was developed specifically to collect data for pavement management purposes (5).

Distress magnitudes are determined by raters through a decision process that considers the following three questions:

- Is the distress type present on the pavement segment?
- What is the severity of the distress type?
- What is the extent of the distress type?

Once the presence of a given distress type has been noted, a distress rating is assigned by using descriptive definitions (scales) to determine its severity and extent. The distress scales serve primarily to distinguish severities, magnitudes of which are assigned the values of None, Small, Medium, Large, or Total. On the basis of a visual estimate or count, the distress is then classified as occurring locally or generally throughout the surveyed section. Thus, each distress is assigned a rating from the possible values of N, SL, SG, ML, MG, LL, LG, TL, or TG along each pavement section surveyed. For the purposes of data manipulation, each of the nine possible severity-extent combinations is assigned an integer value. This process is shown in Figure 1.

Not all distresses require use of the entire scale range. When consistent observations would be difficult, or incremental knowledge of condition would not impact maintenance, the Total category has been abolished. In some cases (e.g., loss of transverse joint sealant), only the None and Large severities are recorded. Table 3 presents the possible distress ratings for the 1989 scales.

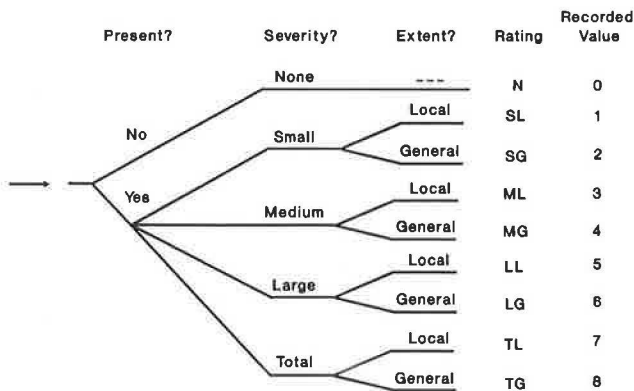


FIGURE 1 Distress rating decision tree.

TABLE 3 POSSIBLE DISTRESS RATINGS (1989 SURVEY)

Distress	Possible Ratings
Centerline Cracking	N, SL, SG, ML, MG, LL, LG, TL, TG
Longitudinal Cracking	N, SL, SG, ML, MG, LL, LG, TL, TG
Surface Defects	N, SL, SG, ML, MG, LL, LG
Rutting	N, LL, LG
Transverse Cracking	N, SL, SG, ML, MG, LL, LG, TL, TG
Edge Cracking	N, SL, SG, ML, MG, LL, LG, TL, TG
Shoulder Defects	N, SL, SG, ML, MG, LL, LG
Lane/Shoulder Displacement	N, SL, SG, ML, MG, LL, LG
Loss of Transverse Jt. Seal	N, LL, LG
Transverse Jt. Spalling	N, SL, SG, ML, MG, LL, LG
Transverse Jt. Faulting	N, LL, LG
Long. Jt. Spalling	N, SL, SG, ML, MG, LL, LG
Slab Surface Defects	N, SL, SG, ML, MG, LL, LG
Slab Cracking	N, SL, SG, ML, MG, LL, LG, TL, TG

## Survey Procedure

The entire 559-mi thruway system is divided into 0.1-mi sections, which are defined by milemarker signposts. The distress survey uses two crews during a 9-week period to rate the entire thruway pavement system. Each crew can survey between 15 and 30 mi per day at a speed that varies from 5 to 10 mph. Pavements in good condition can be surveyed at higher speeds, while more distressed locations require lower speeds. Only the driving lane and right shoulder are rated. The distress condition at these locations is extrapolated onto the other lanes and inside shoulder.

The NYSTA uses two vans, eight rating personnel, and two backup trucks to perform the annual system survey. The survey is applied by a four-person crew using a customized passenger van that travels on the shoulder and is followed by a heavy dump truck with a flashing arrow bar.

Three persons are assigned specific distress types to monitor and rate during the survey. These are designated as Rater 1, Rater 2, and Rater 3. Rater 1 is positioned immediately behind the driver, Rater 2 is seated behind Rater 1, and Rater 3 is seated next to the driver. Raters 1 and 2 have swivel chairs, which allow them to directly face large windows at the side of the van. Table 4 presents the survey crew rating assignments for each pavement type.

The survey begins with the van positioned on the shoulder, next to the starting milemarker. A backup truck is placed at an appropriate distance behind the van, while the driving lane remains open to traffic. When the crew members are ready, the van proceeds along the shoulder at slow speed, followed by the truck. Crew members observe the driving lane and shoulder, and may consult summaries of the distress scales that are available for reference. The van is stopped at the end of each 0.1-mi section. Crew members provide the ratings for their assigned distresses to Rater 1 who records them in a laptop computer. The data logging program provides extensive error and validity checking features, as well as editing capability. A comment feature allows the raters to record the presence of nonstandard distresses and other noteworthy features. After the ratings have been recorded, the van proceeds to the next (usually consecutive) 0.1-mi section. All persons are fully trained in all rating assignments, and they are encouraged to exchange rating assignments several times a day.

## Training

All crew members are qualified to rate in every position, because they all participate in an extensive 3-week annual training program. At the completion of their training, they are rigorously tested and certified on their ability to recognize the standard distress types, and rate them repeatably.

TABLE 4 SURVEY CREW RATING ASSIGNMENTS

	OVERLAID PAVEMENT	CONCRETE PAVEMENT
Driver or Alternate	Edge Cracking Shoulder Defects Lane/Shoulder Displacement	Shoulder Defects Lane/Shoulder Displacement
Rater 1	Centerline Cracking Longitudinal Cracking	Transverse Joint Seal Loss Transverse Joint Spalling Transverse Joint Faulting Longitudinal Joint Spalling
Rater 2	Surface Defects Rutting Transverse Cracking	Slab Surface Defects Slab Cracking

## Costs

The cost for the 1989 thruway system survey was approximately \$143,400. This figure includes overhead, plus all salaries, benefits, travel, equipment, and materials required for 3 weeks of training and 9 weeks of surveying. The cost of backup trucks and drivers is not included. Table 5 presents a summary of the costs. As raters and instructors are "borrowed" from their normal NYSTA maintenance duties for the duration of the annual training and survey, the salary figures given in Table 5 are not truly added costs to the authority.

## REPEATABILITY TESTING

### Conditions of Testing

Repeatability testing determines whether or not differences between multiple ratings at given locations are statistically significant. It is used as a tool for evaluating both individual performance capability and overall objectivity in the survey technique.

Three series of tests were performed to investigate the survey technique. The testing plan was developed to satisfy the following objectives:

- Examine the feasibility of the distress survey technique,
- Evaluate the adequacy of the applied training,
- Establish the repeatability of distress scales and individual raters, and
- Indicate whether the network survey was applied repeatably.

Thus, the applied repeatability testing technique has concentrated on two aspects of the survey, namely (a) individual raters, and (b) distress scales. In an attempt to minimize the effect of the other parameters discussed, all tests were conducted under similar environmental conditions, to the greatest extent possible.

### Data Collection

In order to evaluate the full range of the developed scales, the survey was applied to concrete and overlaid pavements exhibiting various levels of distress. Data collection events are presented in Table 6.

During Test Series 1 and 2, each crew member rated every distress type twice. For each test, this involved six passes of each van, with the raters rotating positions (distress assignments) after each pass.

TABLE 5 APPROXIMATE COSTS FOR THE 1989 THRUWAY PMS DISTRESS SURVEY

Budget Item	3 Weeks Training	7 Weeks Survey	2 Weeks Follow-up	Totals
Salary				
Raters	\$11,890	\$37,210	\$5,315	\$73,542
Instructors	16,600	7,200	2,057	34,946
Travel	4,950	22,575	3,583	31,108
Meeting Room	250	—	—	250
Vehicles	900	2,100	600	3,600
Total	\$34,590	\$69,085	\$11,555	\$143,446

TABLE 6 DATA COLLECTION EVENTS FOR REPEATABILITY TESTING, 1989

Test Series	Number of Events	Personnel	Timing	Number of Data Points
1	1	PMS Engineers	Before Training	30-33
2	2	Rater Trainees	After Training	10
3	15	Raters	During Survey	6 - 27

Data collected during Test Series 1 were used to examine the ability of PMS engineers to determine precisely the same rating of each distress in two different surveys of the same pavement section. This phase of the study provided information on encountered difficulties with distress scales and rating assignments. Comparisons of ratings for each distress helped identify and eliminate any inconsistencies in the applied rating technique. This activity was critical, because the PMS engineers also serve as trainers during the distress rater training course.

Before Test Series 2, the three trainers developed correct ratings for a control pavement section by performing the survey three times and comparing the results to arrive at a consensus in the field. Individual ratings were compared with the correct ones to enable an assessment of each trainee's understanding of the scales, and to identify distresses that were troublesome to all raters. Repeatability of the results was established through comparisons of individual ratings produced during different trips, and applications of statistical testing techniques.

Test Series 3 comprises 15 evaluations performed as spot checks to indicate whether the survey was actually being applied repeatably. Because raters generally held different distress assignments during these evaluations, no repeatability analysis of individual raters was performed. The generated data were used to validate consistency in the application of the distress scales and repeatability of the results.

### Data Analysis

The applied repeatability evaluation involves a statistical hypothesis test about whether the mean difference between paired observations is significantly different from zero. For each distress type, ratings obtained at the same location but on different trips are compared (statistically tested) in a pairwise fashion.

Distress data evaluated in this study were generated through a stratified sampling procedure in which distress samples were taken from the total population of distress values for concrete and overlaid pavements. For each comparison, collected data are ordered in pairs  $[x_{ai}, x_{bi}]$  of each distress rating at the  $i$ th section of pavement, where  $x_{ai}$  and  $x_{bi}$  denote the ratings recorded on Trips A and B, respectively. A paired comparison of these ordered pairs facilitates evaluation of differences between two ratings of a given pavement segment.

The conditions used for repeatability testing are as follows:

- Population variances unknown (but probably equal),
- Population means unknown,
- Generally small sample size ( $n = 6$  to 33), and
- Paired dependent data.



Under these conditions, a two-sided dependent paired-sample  $t$  test is an appropriate technique to assess repeatability. The two-sided technique separates the rejection region into two equal parts in the upper and lower tails of the data distributions. The size of the rejection region is determined by the amount of Type I error that is considered tolerable for the experiment. This value, which is referred to as the significance level  $\alpha$  of the test, is commonly set at 0.05 ( $\alpha = 0.05$ ). In practical terms,  $\alpha$  is the fraction of the number of tests that produces a wrong decision during repeated use. In all three studies, testing was performed at three significance levels, namely  $\alpha = 0.10$ ,  $\alpha = 0.05$ , and  $\alpha = 0.01$ .

Each repeatability test for the distress survey is performed by two-sided, one-sample hypothesis testing of the differences between ratings obtained for a specified test section on subsequent trips. This is achieved by a hypothesis test about the mean differences between ratings for each 0.1 mi, which may be formally stated as follows:

$$H_0: \mu_d = 0$$

$$H_A: \mu_d \neq 0$$

In these expressions,  $H_0$  and  $H_A$  represent the null and alternative hypotheses, respectively, and  $\mu_d$  denotes the mean of the differences between compared ratings. Acceptance or rejection of the null hypothesis is based on the data statistic  $t_{\text{obs}}$ , defined as follows:

$$t_{\text{obs}} = \frac{\bar{d}}{S_{\bar{d}}}$$

in which  $\bar{d}$  is the mean difference, and  $S_{\bar{d}}$  is the estimated standard error of the mean difference. For the given conditions, in which the (independent) differences are considered as random variables, the expression for the estimated standard error  $S_{\bar{d}}$  of the mean difference is

$$S_{\bar{d}} = \frac{S_d}{(n)^{1/2}}$$

where  $S_d$  is the standard deviation of the differences, and  $n$  is the number of observations.

The null hypothesis is rejected ( $\mu_d \neq 0$ ), if  $|t_{\text{obs}}| > t_{\alpha/2, n-1}$ ; and accepted ( $\mu_d = 0$ ), if  $|t_{\text{obs}}| < t_{\alpha/2, n-1}$ , where  $\alpha$  is the significance level, and  $n$  is the number of observations. When a comparison passes the repeatability test (null hypothesis is accepted), this has the interpretation that, on the average, the differences between ratings (for the given distress) on the two trips being compared are not statistically significant. A comparison that fails the repeatability test (null hypothesis is rejected) indicates that, on the average, the difference between ratings is nonzero.

## Results

Data analysis from all three test series resulted in a total of 42 comparisons per distress for concrete pavement distresses and 57 comparisons per distress for overlaid pavement dis-

tresses. Each of these pairwise comparisons was analyzed at the three levels of significance identified earlier.

For each distress and test series, the number of all comparisons that fail the repeatability test (rejection of the null hypothesis) at a given significance level indicates the overall validity of the survey procedure for that distress. Tables 7–9 present the results of testing all comparisons for each distress type. Within each test series, distresses are ranked by difficulty, which is interpreted as the number of comparisons that failed the repeatability test at a given significance level. In order to facilitate comparison between test series, the numbers of such failures are reported as percentages.

The percentage of repeatability failures at  $\alpha = 0.10$ ,  $\alpha = 0.05$ , and  $\alpha = 0.01$  was used as the primary, secondary, and tertiary ranking key, respectively. Thus, a ranking of 1 indicates the least repeatable (most difficult) distress. For example, Table 7 indicates that in Test Series 1, surface defects

TABLE 7 DIFFICULTY RANKING FOR OVERLAID PAVEMENT DISTRESSES (TEST SERIES 1 AND 2)

Test Series	Distress Ranking	Percent of Failed Comparisons		
		$\alpha=0.10$	$\alpha=0.05$	$\alpha=0.01$
1	1. Centerline Cracking	70.0	70.0	50.0
	2. Lane/Shoulder Displacement	60.0	40.0	0.0
	3. Surface Defects	50.0	40.0	0.0
	4. Edge Cracking	40.0	40.0	20.0
	5. Longitudinal Cracking	40.0	30.0	0.0
	6. Shoulder Defects	40.0	20.0	0.0
	7/8. Rutting	0.0	0.0	0.0
	7/8. Transverse Cracking	0.0	0.0	0.0
2	1. Centerline Cracking	44.4	40.7	37.0
	2. Longitudinal Cracking	22.2	14.8	0.0
	3. Transverse Cracking	14.8	11.1	0.0
	4. Shoulder Defects	14.8	3.7	0.0
	5. Edge Cracking	11.1	7.4	7.4
	6. Surface Defects	3.7	3.7	0.0
	7/8. Lane/Shoulder Displacement	0.0	0.0	0.0
	7/8. Rutting	0.0	0.0	0.0

TABLE 8 DIFFICULTY FOR OVERLAID PAVEMENT DISTRESSES (TEST SERIES 3)

Test Series	Distress Ranking	Percent of failed comparisons		
		$\alpha=0.10$	$\alpha=0.05$	$\alpha=0.01$
3	1. Longitudinal Cracking	58.3	33.3	8.3
	2. Transverse Cracking	50.0	50.0	41.7
	3. Centerline Cracking	41.7	33.3	33.3
	4. Shoulder Defects	41.7	25.0	25.0
	5/6. Surface Defects	8.3	8.3	8.3
	5/6. Rutting	8.3	8.3	8.3
	7. Edge Cracking	8.3	8.3	0.0
	8. Lane/Shoulder Displacement	0.0	0.0	0.0

TABLE 9 DIFFICULTY RANKING FOR CONCRETE PAVEMENT DISTRESSES

Test Series	Distress Ranking	Percent of failed comparisons		
		$\alpha=0.10$	$\alpha=0.05$	$\alpha=0.01$
1	1. Transverse Joint Faulting	93.3	93.3	80.0
	2. Lane/Shoulder Displacement	66.7	66.7	66.7
	3. Transverse Joint Spalling	66.7	53.3	46.7
	4. Loss of Transverse Joint Sealant	46.7	46.7	20.0
	5. Slab Cracking	46.7	46.7	13.3
	6. Shoulder Defects	40.0	40.0	33.3
	7. Slab Surface Defects	33.3	26.7	0.0
	8. Longitudinal Joint Spalling	20.0	13.3	6.7
2	1. Longitudinal Joint Spalling	48.1	25.9	22.2
	2. Lane/Shoulder Displacement	33.3	18.5	3.7
	3. Transverse Joint Spalling	22.2	14.8	3.7
	4. Slab Cracking	18.5	14.8	11.1
	5. Loss of Transverse Joint Sealant	14.8	3.7	0.0
	6. Transverse Joint Faulting	11.1	3.7	0.0
	7/8. Shoulder Defects	0.0	0.0	0.0
	7/8. Slab Surface Defects	0.0	0.0	0.0

were the third most difficult distress to rate repeatably. In this case, 50 percent of the comparisons failed the statistical test for repeatability. Similarly, in Test Series 2, longitudinal cracking was the second most difficult distress to rate, but only 22.2 percent of the tested comparisons failed the repeatability test.

## DISCUSSION OF RESULTS

The basic notion in the applied repeatability testing is that, if individuals can duplicate ratings through multiple applications at the same location, then the procedure is precise. Furthermore, if each rating is also correct (i.e., identical to the true value), then the procedure is both precise and objective.

In principle, objectivity evaluation can be also performed using the ANOVA technique. This technique assumes that one can control the main factors influencing the magnitudes of observed distresses; and, furthermore, one can repeat distress observations under identical conditions at specified levels of these (controlled) factors. However, important factors influencing distress observations cannot be easily controlled. Thus, use of the ANOVA technique would tend to obscure information about accuracy.

During all test series, raters were prevented from learning a section through the use of different pavement test sections for different tests. The used test sites did not exhibit a complete range of the magnitudes of every distress. Therefore, in the strictest sense, the conclusions drawn from the study can only be applied to the pavement test sections; i.e., a given distress scale either was or was not repeatable on that particular length of pavement. An ideal test site would have enabled an evaluation of every severity and extent of each distress. However, external conditions contributing to distress development (e.g., climate, materials, and traffic) tend to be uniform over the short distances used in repeatability testing, and, therefore, the desirable wide range of all distress severities does not commonly develop in small areas. When choosing test locations, preference was given to those sites exhibiting a wide range of distresses, but consideration was also given to operational issues, such as travel time to the site. Thus, of necessity, applicability of the repeatability results had to be extrapolated to the entire pavement system from data generated at the chosen sites.

In addition to the variability introduced by individuals and their interpretations of the distress scales, many other factors may affect survey repeatability. These include seat position, rater fatigue, the angle of incidence of sunlight, pavement dampness, light intensity, and others. In principle, it would be possible to design repeatability tests to study each of these factors. However, it is doubtful that such testing would be justified, because most of these factors fluctuate over some natural range throughout actual survey application.

In the applied Test Series 1, comparisons were made between all raters on all trips, whereas in Test Series 2, comparisons were made only between each person's ratings and the true distress magnitudes. In the analysis of data from Test Series 3 (obtained during actual survey application), comparisons were made between values for each distress produced by different raters.

The findings of the first series of repeatability tests indicated that longitudinal joint spalling (concrete pavement) and centerline cracking (overlaid pavement) were by far the most difficult distresses for the raters to evaluate repeatably. These two distresses are observed at the same location on the roadway surface.

Following a 3-week training course and testing program, which culminated in the test results presented as Series 2, Schultz and Grivas (4) concluded that the NYSTA PMS distress scales could be applied repeatably. The rater trainees were then graduated from the training program as certified PMS raters.

Survey speed is a factor influencing data quality. In general, it is expected to increase (up to a certain limit) as raters become more experienced. An average maximum speed of 5 to 10 mph appears to be appropriate in order for the developed survey to be reliable. This was validated during the 15 repeatability spot checks categorized as Test Series 3. Table 8 presents results from 12 checks on overlaid pavement. The results indicate that the raters were somewhat less repeatable during field application than they were during training. Specifically, longitudinal cracking, transverse cracking, and shoulder defects were found to be less repeatable during the system survey than during training. This is attributed to the higher speed used during the system survey, and indicates the need for increased control of field operations.

## SUMMARY AND CONCLUSIONS

A methodology was presented for determining severity and extent of 14 surface distresses on asphalt and portland cement concrete pavements. It included distress scale formulation and operational procedures followed in conducting the survey.

The reliability of the distress survey was evaluated in three series of repeatability tests. Obtained data were analyzed using appropriate statistical techniques, and the results were presented and discussed.

The developed distress survey was applied on the entire thruway pavement network for the first time in 1989. The generated experience and results from additional sample tests of repeatability have led to enhancements in both the reliability and the operation of the survey.

On the basis of the methodology and results presented in this study, the following conclusions can be drawn:

- The implemented distress survey is a viable technique to monitor distresses on thruway pavements.
- The utilized distress scales and survey procedure provide the quantity and quality of data needed for pavement management purposes.
- The applied statistical technique (paired *t* test) can be used to evaluate the repeatability of both distress scales and raters at different stages in the survey development and application.

## ACKNOWLEDGMENTS

The research presented in this paper was sponsored by the New York State Thruway Authority. It is part of a broader

research effort to develop and implement a pavement management system for the authority's pavements. The assistance provided by thruway maintenance personnel during survey development is gratefully acknowledged.

## REFERENCES

1. D. A. Grivas. *Thruway Distress Survey Manual*. Report RPI.NYSTA-1.3, Civil Engineering Department, Rensselaer Polytechnic Institute, Troy, New York, 1990.
2. *AASHTO Guide for Design of Pavement Structures*. AASHTO, Washington, D.C., 1986, p. III-26.
3. J. M. Sime and D. A. Larson. Objective versus Subjective Pavement Distress Evaluation Systems. *Proc., Automated Pavement Distress Data Collection Equipment Seminar*, Vol. 1, Ames, Iowa, 1990, pp. 23-43.
4. B. C. Schultz and D. A. Grivas. *Repeatability of the Thruway PMS Distress Survey Technique*. Report RPI.NYSTA-2.1, Civil Engineering Department, Rensselaer Polytechnic Institute, Troy, New York, 1989.
5. D. A. Grivas. *Development of a Pavement Damage Assessment Methodology and Associated Distress Survey*. Report CE-87-1, Civil Engineering Department, Rensselaer Polytechnic Institute, Troy, New York, 1987.

---

*Publication of this paper sponsored by Committee on Pavement Monitoring, Evaluation, and Data Storage.*