New Zealand Experience in Comparing Manual and Automatic Pavement Condition Rating Systems

P. D. Cenek and J. E. Patrick, *Works Consultancy Services, New Zealand*
J. F. McGuire and D. A. Robertson, *Transit New Zealand*

The purpose of condition rating surveys is to measure and record defects along a section of road in a standard and objective manner. This provides a measure of the condition of each road section, which in turn can be used to assess routine maintenance and rehabilitation needs. In New Zealand, such a survey is presently accomplished by visual walkover on a 10 percent sample of the road network carried out by trained raters. As part of an effort to continually improve the quality of pavement condition data, a comparative study was undertaken to establish the degree of correlation between walkover survey data and data automatically acquired by the Swedish Road Surface Tester (RST) and Side Force Coefficient Routine Investigation Machine (SCRIM). The principal pavement distress modes of interest were confined to rutting, shoving, scabbing, and flushing. The main conclusion reached was that, although RST and SCRIM have value as survey tools, some form of visual assessment of the actual pavement condition is still required to fully identify surface related defects. In addition, it was demonstrated that for vehicle-acquired condition rating data, particular attention must be paid to selecting appropriate reporting lengths and intervention levels so that they are consistent with both the resolution of the measuring device and the minimum pavement length that justifies resurfacing or shape correction.

The role of experience and judgment in pavement design, construction, and maintenance is so great, given the need to use local materials in natural environments, that good engineering demands that the performance of all roads to be consistently evaluated (*1*). The aims of this evaluation are to

1. Check if the intended pavement function and performance objectives are being achieved;
2. Provide guidance for planning and rehabilitation;
3. Provide feedback for improvements to existing design, construction, and maintenance procedures;
4. Establish a data base on road performance for use by future designers and economic analysts; and
5. Detect condition changes from one year to the next.

The objective assessment of the present condition of a road requires the rating of the individual components that make up that condition. In New Zealand, a pavement management system (PMS) has evolved around manual surveys carried out on a 10 percent sample of the road network. This paper describes the condition rating procedures adopted and the results of a comparative study undertaken to establish the degree of correlation between manual survey data and data obtained by continuous measuring devices such as the Swedish Road Surface Tester (RST) (*2*) and the Side Force Coefficient Routine Investigation Machine (SCRIM) (*3*).

Description of Road Assessment and Maintenance Management System

New Zealand's road network comprises some 10 500 km of state highways and 83 500 km of local roads. Approximately 38 500 km of the network are unsealed. The majority of the sealed roads consist of a granular basecourse with a chipseal surface (bitumen with a one sized aggregate surface), although increasingly asphaltic concrete and porous friction course are being used on urban motorways and arterials.

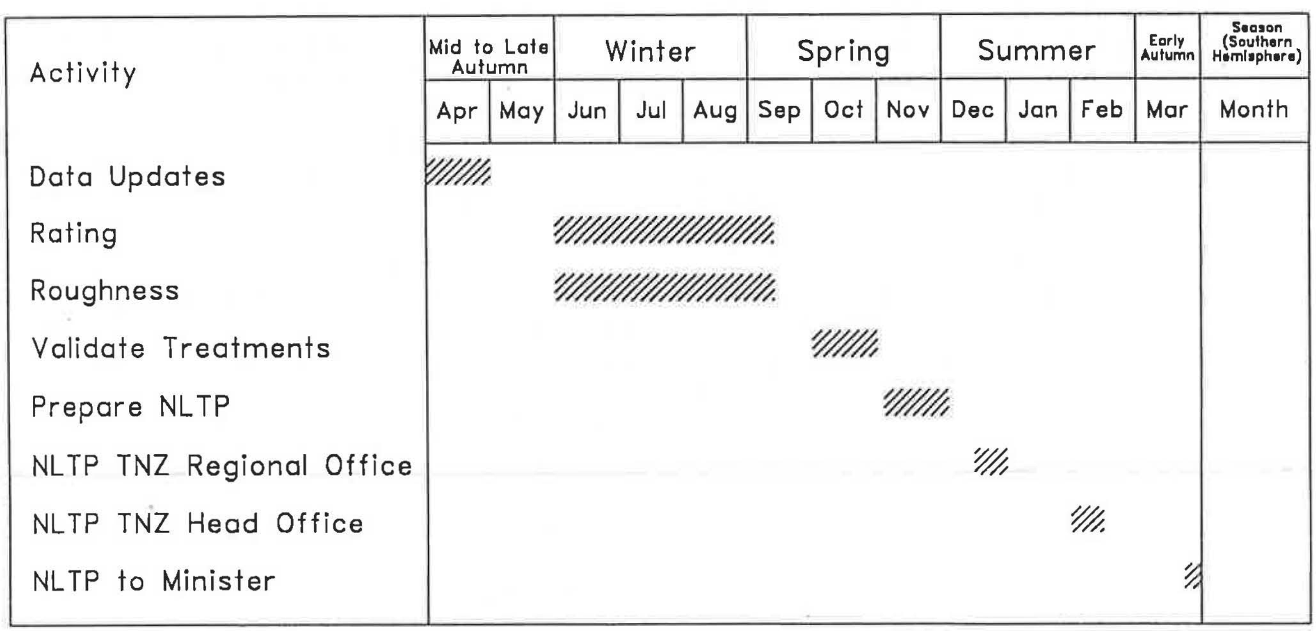
Transit New Zealand (formerly the National Roads Board) is the national road agency. It has the responsibility for managing the state highways and allocating funds (on average 50 percent financial assistance) to the local authorities. Government funding is obtained via road user charges, and gasoline tax, licenses and fees and made available to Transit New Zealand to disburse in accordance with an agreed annual land transport program. Therefore, to effectively manage the road network a PMS is required.

New Zealand's Road Assessment and Maintenance Management (RAMM) system was initiated in the mid 1980s by a local government group and supported by Transit New Zealand. The RAMM system is typical of PMSs in that it contains inventory data of the road asset, condition data including roughness, and an analysis module based on a benefit/cost approach to provide a priority list of treatments.

Central to the RAMM system are annual inspections concerned with pavement surface defects and longitudinal roughness measurement (ride quality). Information regarding surface defects is presently acquired by manual condition rating procedures whereby trained personnel inspect and record defects on a sample per segment basis. Typically this is a 50-m sample per 500-m segment for pavement defects and the full length for surface drainage because of the susceptibility of New Zealand pavements to damage initiated by water ingress. These manual surveys are conducted over the late winter months and are complimented by 100 percent sampling of the road roughness undertaken with the NAASRA roughness meter (*4*).

The schedule of activities must be tightly controlled to formulate the yearly National Land Transport Programme (NLTP). The timing is shown in Figure 1. Several points should be noted.

1. Significant pavement defects are most easily observed and measured at the end of winter before the summer construction season;
2. The data must be acquired by field survey, analyzed, and area treatments for proposed maintenance and reconstruction completed in time to advertise and let contracts for start and completion in the favorable summer season;
3. Typical 1993 unit costs of the various RAMM associated activities are:
   –Initial establishment of the various inventories: US$8.50/lane-km;
   –Subsequent annual inventory condition ratings, including field and office activities: US$5.00/lane-km; and
   –Road roughness survey: US$2.00/lane-km.



Note: NLTP = National Land Transport Programme

FIGURE 1 RAMM schedule of activities.

The pavement surface defects, as defined in the Transit New Zealand Road Condition Rating Manual (*5*), are summarized in Table 1 along with their units of measure. No scoring system is used, instead the actual measurements are recorded and used in the treatment selection analysis. The surveys are conducted by many teams of raters who have attended rating training workshops conducted on behalf of Transit New Zealand.

The RAMM system is now owned by Transit New Zealand, and development is progressed in consultation with local authorities via a RAMM advisory group. Several user groups have been formed to facilitate information exchange and liaison.

As at September 1993, only 6 of 74 local authorities had not undertaken to implement the RAMM system. A directive from the Minister of Transport to Transit New Zealand in June 1993 required all local authorities to have the RAMM system in place no later than June 30, 1994. The New Zealand Government requires all local governments to justify ongoing maintenance funding by means of the RAMM system. The intention is to provide a common benchmarked approach to asset management across all local authorities. The RAMM system is also a cornerstone of Transit New Zealand's assurances to government that budget levels, toward which financial assistance is made available, are set at equitable and efficient levels across the country.

Data Quality Management

To maximize the benefits of the RAMM system and ensure that valid comparisons of the condition of the highways from year to year and location to location can be made, accurate data must be gathered. Steps that may be taken to provide quality data are

- Documentation of measurement procedures;
- Training material and training courses and certification for raters;
- Introduction of quality control, tolerance limits and statistical checks on permissible variability for data;
- Calibration procedures and control procedures for the deployment of measuring equipment;
- Comparison of field data against existing data records to check for anomalies;
- Procedures to adopt when errors are encountered; and
- Quality assurance requirements incorporated in contracts for survey work.

All of the above steps have or are being implemented in New Zealand.

Because most of the data are gathered by the use of straight edges and other measuring equipment (e.g., tapes, wedges, wheels, etc.) uniformity of training is critical. A technical institute, Taranaki Polytechnic, works with Transit New Zealand to provide the course format and resource materials to accredited trainers. On state highways, the raters are employees of the consultants providing road network management services to Transit New Zealand.

The first condition rating survey was conducted on the state highway network in 1989. Training requirements were intensified in 1990, and from data comparisons between the first two years, quality assurance (QA) criteria were developed and tried in 1991. The QA criteria were reviewed and distributed to local authorities in 1992 and recommended for adoption.

TABLE 1 Surveyed Pavement Defects

Defect	Units
Surface	
Rutting	Linear metre > 20 mm from 2 m straight edge
Shoving	Linear metre in wheelpaths
Scabbing	m^2 > 10% chip loss
Flushing	Linear metre in wheelpaths
Alligator cracks	Linear metre in wheelpaths
Longitudinal and transverse cracks	Linear metre of crack
Joint cracks	Linear metre of crack
Potholes	Number > 70 mm diameter
Pothole patch	Number < 0.5 m^2
Edge break	Linear metre
Edge break patch	Linear metre
Drainage	
Blocked	Linear metre
Inadequate	Linear metre < 400 mm deep
Ineffective Shoulder	Linear metre of shoulder that prevents water discharge

From the experience gained in the QA process, limits of variation in the measurement (and recognition) of the defects were developed and issued in 1992. With further experience, these limits are likely to be refined further.

The changes in data brought about by the training procedures is illustrated in Figures 2, 3, and 4 for the length of edge break (reduction of the seal coat by more than 100 mm from the original line of the seal edge), scabbing (areas of chip loss), and alligator cracking. It can be seen that there was a major shift in the distribution of edge break from 1989 to the following years and it appears that the 1989 survey overestimated the rate of this dis-

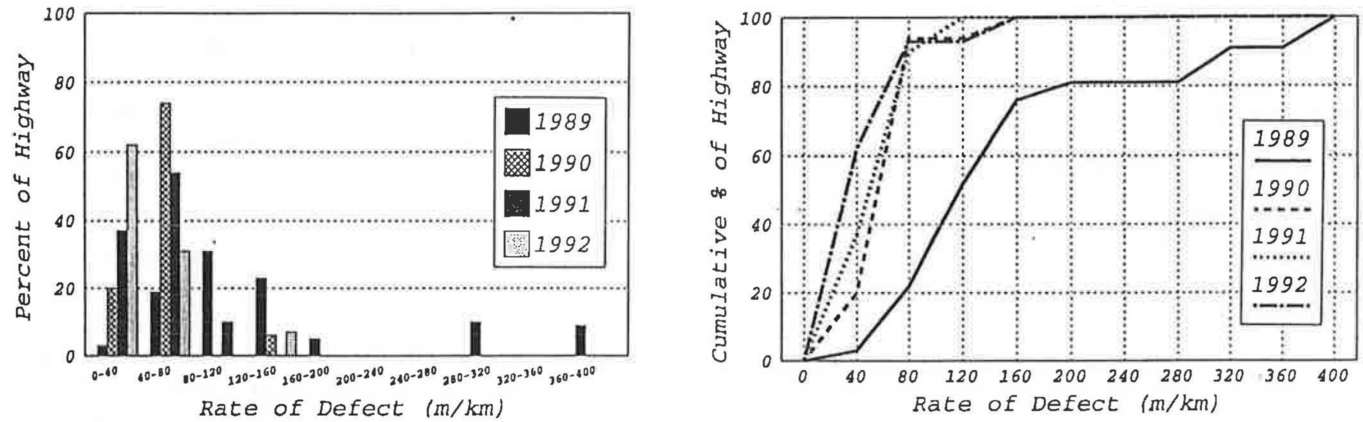


FIGURE 2 Variability of manually obtained edge break data, 1989–1992.

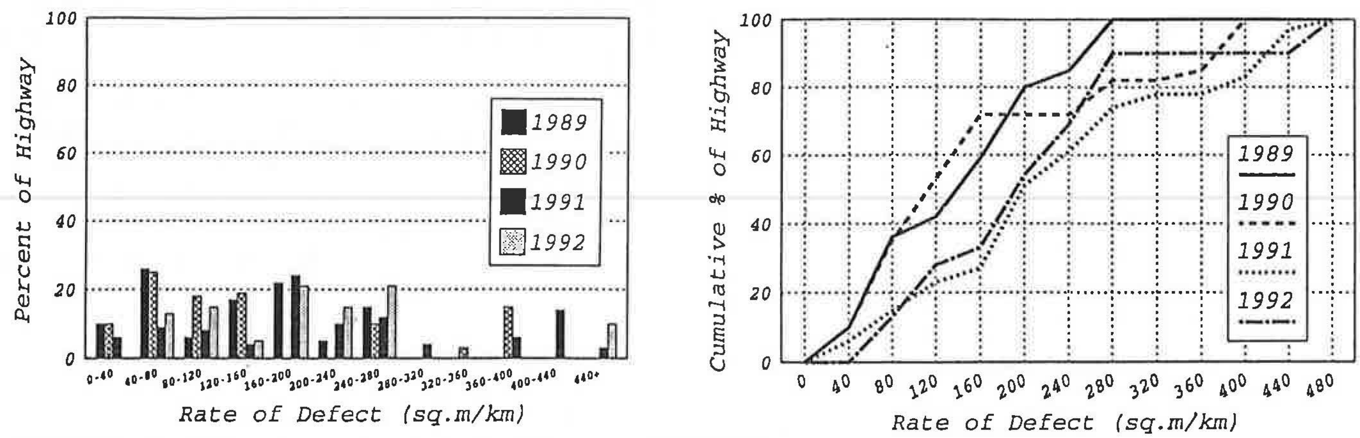


FIGURE 3 Variability of manually obtained scabbing data, 1989–1992.

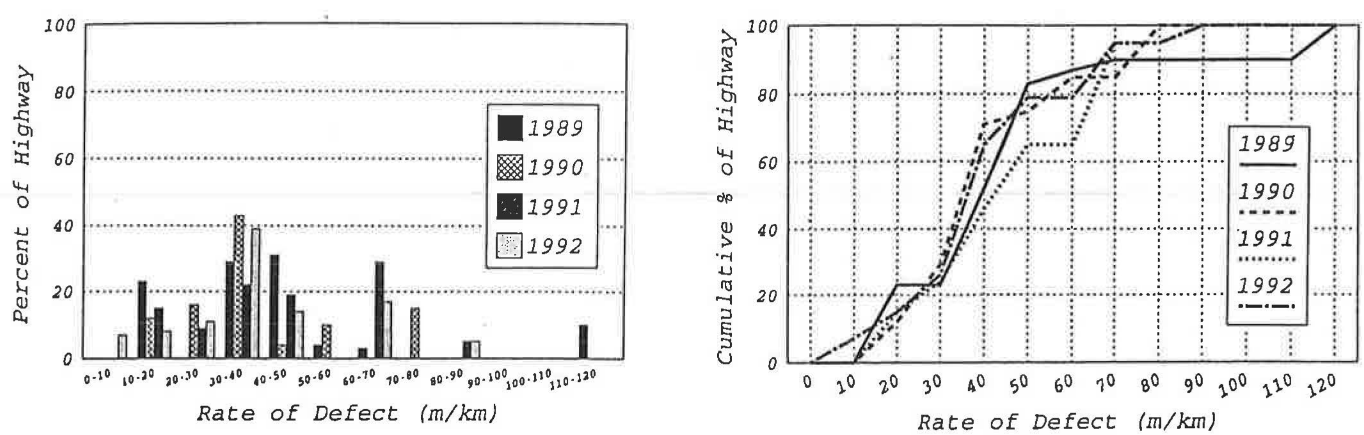


FIGURE 4 Variability of manually obtained alligator cracking data, 1989–1992.

tress. For scabbing, the data suggest that there was a significant change between 1989 and 1990 and 1991 and 1992 surveys. In this case the incidence of this form of distress has increased. Alligator cracking distribution does not appear to have change significantly.

The changes in the frequency of the different types of distress from year to year are greater than can be accounted for by road maintenance activities. It is considered that the closer distributions for all forms of distress that have occurred in the 1991 and 1992 surveys are due to the introduction of formal training and QA procedures.

AUTOMATIC DATA COLLECTION

Although visual surveys are a common and comparatively inexpensive means for collecting pavement defect information, there are a number of recognised problems with this method. These include the subjective nature of the visual surveys, transcription errors that inevitably occur, and consistency of the measured ratings that lead to a low correlation between raters and even among individual raters over time. As a direct consequence of these problems and in response to an increasing need for systematic, objective and safe means of acquiring pavement condition data for input into PMSs, a number of high output continuous measuring devices that can be operated at normal traffic speeds have been developed.

In an effort to determine the most appropriate means for acquiring surface defect information, the use of automatic data collection by vehicle based systems was investigated in 1988 when a RST was imported to New Zealand to demonstrate its suitability for use in gathering input data for RAMM, and again in 1990 when a SCRIM machine performed a survey of the New Zealand state highway network. The availability of 100 percent sampling of roughness, rut depth and surface texture by the RST, and skid resistance, as determined by the 50-km/hr side force coefficient (SFC), provided an opportunity to establish the degree of correlation with the walkover visual survey data stored in the RAMM database.

AGREEMENT BETWEEN AUTOMATIC AND MANUAL CONDITION RATING PROCEDURES

Methodology

Five representative road sections varying in length from 2 to 7 km located on State Highway 1 North (SH1N), New Zealand's main highway, were selected from the RAMM data base. The RAMM data were derived from visual surveys conducted during June and July 1989. These road sections, along with the principal modes of pavement distress identified for each section, have been summarized in Table 2. These modes are briefly described.

- *Rutting:* longitudinal wheeltrack depressions. Only the length of wheeltrack exceeding 20 mm from a 2-m straight edge is recorded in RAMM.
- *Shoving:* horizontal displacement of the surface material, which causes a series of shallow transverse depressions resembling corrugations.
- *Scabbing:* removal of larger surface aggregates leaving craters (i.e., chip loss).
- *Flushing:* road surface has a slick, smooth appearance because binder has flushed (risen) to a level where surface aggregate is only just protruding or where binder has risen to be level with or over the top of the surface aggregate.

Rutting and shoving affect roughness, whereas scabbing and flushing affect macrotexture.

To establish the level of agreement between the walkover visual survey data sampled on a 10 percent basis with the RST and SCRIM data, a graphical approach was adopted. This entailed plotting 100-m averaged results of the RST and SCRIM surveys as a function of distance along the SH1N sections listed in Table 2. Superimposed on the plots, where appropriate, were the intervention/investigation criteria presently specified by Transit New Zealand, along with the locations of the manual inspections. In this manner, the condition of the road section with respect to a certain pavement distress parameter could be readily ascertained along with whether the severity level at the manual inspection location warranted recording by the rater. By way of example, resulting distribution graphs for a site (site 2 in Table 2) are presented in Figures 5 to 8.

Descriptive statistics were also calculated for both the RST and SCRIM data. These statistics were in turn compared with the results of the walkover visual survey data that was normalized with respect to the number of lanes and surveyed road length (i.e., lane-kilometers) to allow the condition of the sections to be ranked with respect to the various recorded distress parameters. The resulting normalized RAMM data are given in Table 3.

Only a summary of the principal findings of the comparative study will be given here; a detailed discussion has been given elsewhere (6).

Rutting

The RST calculates maximum rut depth by a mathematical method, which is analogous to a wire being stretched traversely across the road profile (2). Figure 9 shows the difference between straight edge (as used in manual surveys) and the RST wire surface measurement of rut depth.

A comparison of the RST rut depth distributions with the RAMM rutting ratings identified poor agreement between the measurement methods as to where significant rutting occurred along a road section. Furthermore, at no

TABLE 2 Representative SH1N Condition Survey Sections

Site No.	Section of SH1N	Approximate Location	Length (km)	Characteristic Surface Distress
1	393/9.16-15.97	Whangamarino	6.81	Extensive scabbing and flushing with some localised rutting at the middle and end of section
2	428/0.00-2.37	Huntly	2.37	Extensive scabbing with considerable rutting at middle of section
3	915/6.10-8.71	Manakau	2.61	Extensive flushing with considerable rutting at end of section
4	931/1.50-7.50	Waikanae	6.00	Flushing and shoving along last half of section
5	942/2.09-5.12	Paraparaumu	3.03	Localised scabbing, flushing and rutting

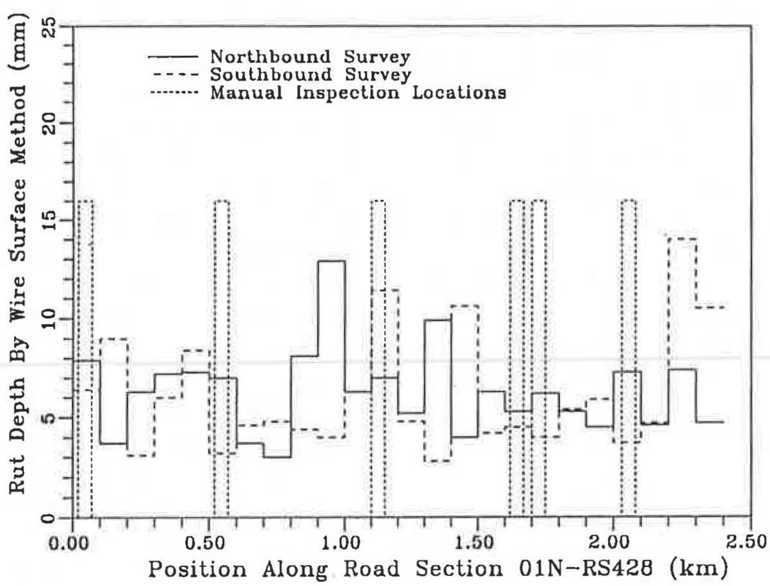


FIGURE 5 100-m rut depth distributions, site 2.

manual inspection location did the RST rut depth exceed the intervention level of 20 mm.

Table 4 shows that sites 1 and 2 have the highest average RST rut depth on the basis of continuous sampling. However, this is inconsistent with the rankings based on normalized RAMM pavement condition ratings given in Table 3 which indicate sites 2 and 3 as having the greatest extent of this type of pavement defect.

The poor agreement between RST rut depth and the manual survey is attributed to

1. Lack of significant correlation between RST and straight edge rut depth described by Jameson et al. (7);
2. The fact that in 1989, some raters gauged rut depth by eye rather than physical measurement—this has subsequently been addressed through training programs implemented since 1990; and
3. The difference in the way the rut depth measurement is presented, for example, an average over a 100-m reporting length (RST) compared with the accumulated length that exceeds a threshold level (manual survey).

Surface Deficiencies

The RST measures both fine and rough macrotexture using the output from laser cameras mounted over each wheelpath. Fine macrotexture covers surface profile

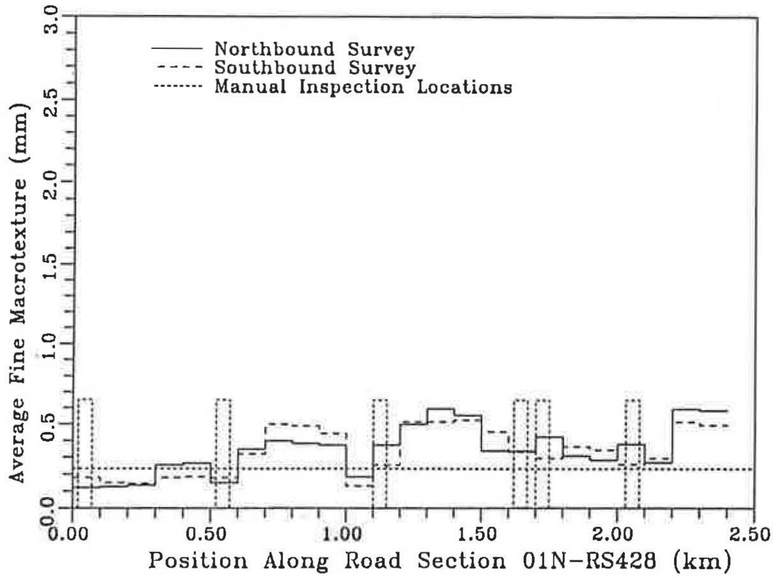


FIGURE 6 100-m RST-RMS fine macrotexture distributions, site 2. (Broken horizontal line represents Transit New Zealand intervention levels.)

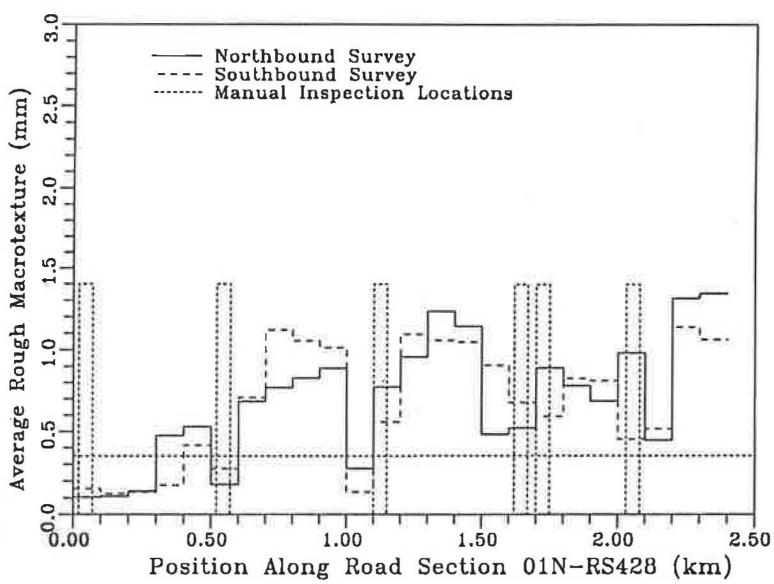


FIGURE 7 100-m RST-RMS rough macrotexture distributions, site 2. (Broken horizontal line represents Transit New Zealand intervention levels.)

wavelengths over the range 1 to 10 mm, constituting macrotexture roughness caused by very small chippings and the sharpness and angularity of chippings. Rough macrotexture covers wavelengths in the range 10 to 80 mm and describes surface roughness caused by large chippings, asperities, and other surface features, which, in the order of size, are less than the tire/road contact zone.

The principal surface deficiencies recorded in visual surveys are associated with scabbing, flushing, potholing, and cracking. In a work by Cenek (*8*) it was demonstrated that the RST could not reliably recognize cracking in chipseal surfaces. However, because both scabbing and flushing involve a loss of surface texture, it was expected that some surface defects could be identified by combining RST fine and rough macrotexture measures with SCRIM data. In particular, RST rough macrotexture should be sensitive to scabbing and potholing, and RST fine macrotexture to flushing. Therefore to establish the degree of agreement between automatic and manual survey procedures in identifying defects in the surface texture, the 100-m averaged RST fine and rough RMS macrotexture and SCRIM skid resistance distributions for the five selected SH1N sections

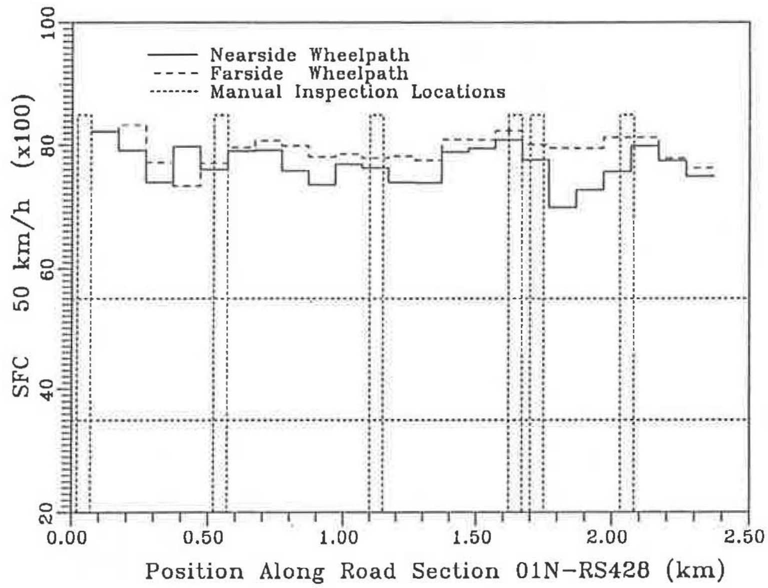


FIGURE 8 100-m SCRIM SFC_{50} distributions, site 2. (Broken horizontal line represents Transit New Zealand intervention levels.)

TABLE 3 Normalized RAMM Pavement Condition Ratings

Site No.	Description	Rutting (m/lane-km)	Shoving (m/lane-km)	Scabbing (m^2/lane-km)	Flushing (m/lane-km)
1	RS393/9.16-15.97	9	0	474	218
2	RS428/0.00-2.37	108	5	400	168
3	RS915/6.10-8.71	110	4	4	286
4	RS931/1.50-7.50	0	61	15	210
5	RS942/2.09-5.12	16	0	16	28

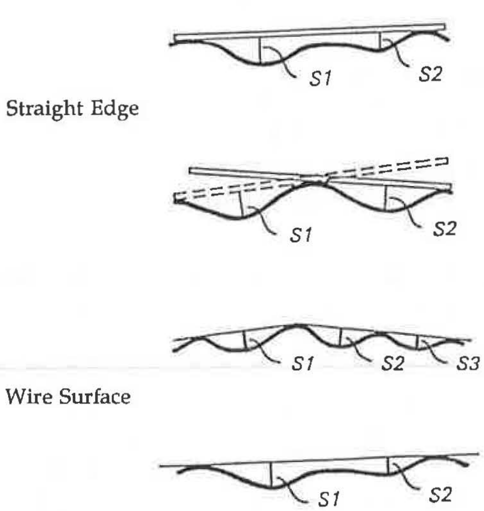


FIGURE 9 Cross-profile measurements; rut depth is the largest value of S1, S2, S3, and so forth (2).

were considered. The associated statistical descriptions for each section are given in Tables 5 to 7.

The distribution graphs showed that only very few 100-m lengths fell below the intervention/investigation criteria for texture depth and skid resistance. Site 2 had the greatest loss of macrotexture, with 1.2 lane-km below the minimum acceptable texture depth, followed by site 3 (0.2 lane-km) and site 4 (0.1 lane-km). Both RST fine and rough macrotexture measures showed the same location and extent of surface texture loss. By comparison, only site 4 had sections where the measured skid resistance fell in the range requiring investigation.

As with the rut depth analysis, there was poor agreement between the RAMM scabbing and flushing ratings and the RST and SCRIM data as to where significant texture defects had occurred along each SH1N section investigated.

Table 5 to 7 show that site 2 has the lowest average texture depth. Of particular interest is the very little difference

TABLE 4 SH1N Descriptive Statistics for RST-Measured Rut Depth

Site No.	Direction	Rut Depth (as Derived from Wire Surface Method), mm						
		Mean	Median	Std Dev	Min	Max	Q1	Q3
1	Northbound	7.14	6.10	3.13	3.20	15.60	4.80	9.50
	Southbound	7.10	6.30	3.06	2.80	16.20	5.10	8.18
2	Northbound	6.35	6.30	2.15	3.00	12.90	4.75	7.30
	Southbound	6.11	4.80	2.98	2.80	14.00	4.05	7.90
3	Northbound	5.42	4.35	3.39	1.60	17.90	3.38	6.35
	Southbound	4.67	4.55	1.61	1.90	8.70	3.50	5.95
4	Northbound	4.27	3.60	2.52	1.00	10.60	2.10	5.75
	Southbound	5.33	4.60	3.31	1.30	15.90	2.50	7.40
5	Northbound	3.26	3.10	1.43	1.20	6.60	2.20	3.93
	Southbound	4.67	2.75	4.91	1.30	23.60	2.08	4.40

TABLE 5 SH1N Descriptive Statistics for RST-RMS Fine Macrotexture

Site No.	Direction	RST-RMS Fine Macrotexture (1-10 mm wavelengths), mm								
		Wheelpath Averaged Statistics							Individual Wheelpath Mean Values	
		Mean	Median	Std Dev	Min	Max	Lower Quartile	Upper Quartile	Outer	Inner
1	Northbound	0.85	0.89	0.10	0.58	0.99	0.83	0.92	0.84	0.85
	Southbound	0.77	0.78	0.09	0.51	0.93	0.73	0.83	0.76	0.77
2	Northbound	0.33	0.34	0.14	0.12	0.60	0.20	0.40	0.33	0.33
	Southbound	0.33	0.33	0.14	0.13	0.53	0.18	0.49	0.33	0.35
3	Northbound	0.52	0.50	0.14	0.24	0.75	0.43	0.65	0.53	0.52
	Southbound	0.49	0.52	0.13	0.19	0.68	0.43	0.58	0.47	0.64
4	Northbound	0.94	0.60	0.67	0.26	2.54	0.47	1.11	0.95	0.94
	Southbound	0.60	0.58	0.20	0.22	1.01	0.44	0.72	0.58	0.61
5	Northbound	0.55	0.56	0.08	0.36	0.75	0.53	0.59	0.54	0.57
	Southbound	0.56	0.57	0.06	0.46	0.70	0.49	0.59	0.55	0.57

TABLE 6 SH1N Descriptive Statistics for RST-RMS Rough Macrotexture

Site No.	Direction	RST-RMS Rough Macrotexture (10-80 mm wavelengths), mm								
		Wheelpath Averaged Statistics							Individual Wheelpath Mean Values	
		Mean	Median	Std Dev	Min	Max	Lower Quartile	Upper Quartile	Outer	Inner
1	Northbound	1.61	1.64	0.15	1.19	1.84	1.55	1.71	1.55	1.66
	Southbound	1.45	1.46	0.16	0.97	1.83	1.38	1.55	1.40	1.51
2	Northbound	0.64	0.69	0.36	0.11	1.32	0.32	0.89	0.63	0.65
	Southbound	0.67	0.70	0.37	0.13	1.14	0.31	1.00	0.65	0.69
3	Northbound	1.18	1.22	0.35	0.46	1.65	0.92	1.47	1.19	1.18
	Southbound	1.11	1.11	0.31	0.35	1.64	0.93	1.31	1.09	1.12
4	Northbound	1.35	1.11	0.52	0.47	2.13	0.91	1.84	1.29	1.41
	Southbound	1.25	1.16	0.43	0.41	2.01	0.90	1.55	1.21	1.28
5	Northbound	1.00	1.00	0.17	0.54	1.28	0.97	1.07	0.95	1.04
	Southbound	1.00	1.03	0.10	0.83	1.20	0.88	1.08	0.97	1.04

TABLE 7 SH1N Descriptive Statistics for 5-m SCRIM SFC_{50} Measurements, Northbound

Site No.	SFC_{50} (x 100) Wheelpath Averaged Statistics							Individual Wheelpath Mean Values	
	Mean	Median	Std Dev	Min	Max	Lower Quartile	Upper Quartile	Outer	Inner
1	70.6	71	3.4	44.0	81.5	69.0	72.5	69.7	71.4
2	78.1	78	3.7	61.5	88.5	76.0	80.5	77.0	79.4
3	69.1	70	4.9	45.5	78.5	66.5	72.5	64.0	74.0
4	70.0	70	6.1	34.5	89.0	67.0	72.5	67.0	73.0
5	70.1	70	3.0	57.5	78.0	68.0	72.0	70.0	70.0

in the mean RST fine macrotexture between outer and inner wheelpaths apart from site 3 in the southbound lane, which shows a significant loss of macrotexture in the outer wheelpath. This result typically indicates the presence of flushing (*8*) and is consistent with the RAMM pavement condition ratings for site 3. Similarly, the RST rough macrotexture data show significant differences in texture depth between outer and inner wheelpaths for sites 1 and 4, and the SCRIM skid resistance data for sites 3 and 4, indicating loss of surface seal either through scabbing, flushing, or patching. Again, these findings are generally consistent with the visual survey ratings. However, when considering mean values given in Tables 5 to 7, site 2 is shown to have the lowest macrotexture depth and site 3 the lowest microtexture. In comparison, the normalized RAMM pavement condition ratings given in Table 3 indicate that the greatest loss of surface seal has occurred on sites 1 and 2.

Many of the above anomalies between automatic and manual survey procedures could have been resolved if measurements of macrotexture and skid resistance were made near the center of the lane in addition to the wheelpaths, thereby enabling the degree of surface deterioration to be more readily identified.

ANALYSIS OF CONTINUOUSLY SAMPLED PAVEMENT DISTRESS DATA

The following additional analyses were performed on the RST- and SCRIM-derived data for each of the road sections listed in Table 2:

1. Descriptive statistics (mean and standard deviation) based on systematic sampling of a 100-m interval at the start of every 0.5 km and continuous sampling were compared to assess the validity of existing sampling procedures used in walkover surveys;
2. SCRIM data based on 5-, 100-, and 500-m reporting lengths were compared to demonstrate the need to relate intervention levels to reporting levels; and
3. Regressions were performed between all the RST measured variables and RST macrotexture and the SCRIM-derived 50-km/hr SFC to establish the degree of correlation among the pavement defect measures.

Influence of Partial Sampling

Both *t* and *F* tests for a 5 percent level of significance were applied to the measures of rutting, RMS fine macrotexture, and 50-km/hr SFC in an attempt to establish whether there were differences between the means and variances derived from 20 percent systematic sampling and those from 100 percent sampling. Unfortunately the minimum reporting length for the RST was limited to 100 m, and so it was not possible to duplicate the 10 percent sampling used in walkover surveys.

With reference to Tables 8 to 10, values of the calculated *t*-statistic all lie well within the critical interval defined by $t_{0.975}$, so the hypothesis that there is essentially no difference between the means derived from 20 and 100 percent sampling is accepted. Similarly, the *F*-test shows that there is no significant difference (i.e., $\sigma_1^2/\sigma_2^2 \leq F_{0.025}$) between the two variance estimates, apart from only two cases. These results, when combined, indicate that we can be 95 percent confident that pavement distress data obtained from 20 percent systematic sampling using automatic means is sufficient to correctly infer the condition of the entire network.

Influence of Reporting Length

With 100 percent sampling, an immense amount of data is collected. This can take a considerable time to summarize and analyze. To overcome this problem, outputs from

TABLE 8 Comparison of Sample and Population Means and Standard Deviations—Rutting

Site No.	Direction	No. of 100 m Length Samples		Mean Rut Depth (mm)		Std Dev (mm)		"t-test"		"F-test"	
		20%	100%	20%	100%	20%	100%	$t_{calculated}$	$\pm t_{0.975}$	$(\sigma_1)^2/(\sigma_2)^2$	$F_{0.025}$
1	Northbound	14	68	6.6	7.1	2.8	3.1	-0.552	1.993	1.27	2.71
	Southbound	14	68	7.2	7.1	3.0	3.0	0.112	1.993	1.00	N/A
2	Northbound	6	24	6.0	6.4	1.0	2.2	-0.420	2.05	4.84	6.29
	Southbound	6	24	5.5	6.1	3.1	3.0	-0.420	2.05	1.07	3.18
3	Northbound	5	26	5.5	5.4	3.9	3.4	0.057	2.04	1.32	3.69
	Southbound	5	26	6.1	4.7	2.2	1.6	1.621	2.04	1.89	3.69
4	Northbound	12	60	4.4	4.2	2.7	2.5	0.246	1.997	1.17	2.20
	Southbound	12	60	6.0	5.3	2.9	3.3	0.674	1.997	1.30	3.00
5	Northbound	6	30	3.2	3.3	1.9	1.4	-0.145	2.03	1.84	3.04
	Southbound	6	30	4.7	4.7	4.9	4.9	0	2.03	1.00	N/A

TABLE 9 Comparison of Sample and Population Means and Standard Deviations—Average RST RMS Fine Macrotexture

Site No.	Direction	No. of 100 m Length Samples		Mean RMS Fine Macrotexture (mm)		Std Dev (mm)		"t-test"		"F-test"	
		20%	100%	20%	100%	20%	100%	$t_{calculated}$	$\pm t_{0.975}$	$(\sigma_1)^2/(\sigma_2)^2$	$F_{0.025}$
1	Northbound	14	68	0.85	0.85	0.11	0.10	0	1.993	1.21	2.13
	Southbound	14	68	0.77	0.77	0.09	0.09	0	1.993	1.00	N/A
2	Northbound	6	24	0.26	0.33	0.08	0.14	-1.14	2.05	3.06	6.29
	Southbound	6	24	0.25	0.34	0.06	0.14	-1.48	2.05	5.44	6.29
3	Northbound	5	26	0.63	0.52	0.09	0.14	1.64	2.04	2.42	8.52
	Southbound	5	26	0.44	0.49	0.15	0.13	-0.742	2.04	1.33	3.69
4	Northbound	12	60	1.03	0.94	0.73	0.67	-0.412	1.997	1.19	2.20
	Southbound	12	60	0.58	0.60	0.22	0.20	-0.306	1.997	1.21	2.20
5	Northbound	6	30	0.55	0.55	0.02	0.08	1.00	2.03	16.00	6.24[a]
	Southbound	6	30	0.58	0.56	0.06	0.06	0.724	2.03	1.00	6.24

[a]Sample and population standard deviations significantly different.

TABLE 10 Comparison of Sample and Population Means and Standard Deviations—$SFC_{50} \times 100$

Site No.	Direction	No. of 100 m Length Samples		Mean SFC_{50} x 100		Std Dev (mm)		"t-test"		"F-test"	
		20%	100%	20%	100%	20%	100%	$t_{calculated}$	$\pm t_{0.975}$	$(\sigma_1)^2/(\sigma_2)^2$	$F_{0.025}$
1	Northbound	14	68	69.9	70.5	2.0	2.1	-0.97	1.993	1.10	2.71
2	Northbound	6	23	79.1	78.1	2.7	2.2	0.91	2.05	1.51	3.25
3	Northbound	5	26	70.2	69.1	1.7	2.9	0.80	2.04	2.91	8.52
4	Northbound	12	60	70.1	70.0	2.1	4.8	0.07	1.997	5.22	3.00[a]
5	Northbound	6	30	69.0	70.1	2.3	2.0	1.16	2.03	1.32	3.04

[a]Sample and population standard deviations significantly different.

continuous devices such as RST and SCRIM are reported over length intervals that are typically several orders of magnitude longer than the sampling length interval. For example, the SCRIM machine measures the SFC in each wheelpath at 5-m intervals. However, in a survey of the New Zealand state highway network performed in 1990, the reporting interval was chosen to be 500 m. Such an averaging procedure can result in a significant smoothing of the raw data. Opportunity was therefore taken to compare descriptive SFC statistics calculated on the basis of 5-, 100-, and 500-m lengths for each of the road sections listed in Table 2, particularly in relation to the identification of road sections with less than a specific SFC level. The results of this analysis are given in Table 11.

As expected, mean values are not affected by the averaging process. However, as a consequence of the reduction in the degree of scatter in the data, the standard deviation reduces and the minimum and maximum values converge as the reporting length is increased from 5 to 500 m. More important, the loss of resolution associated with increasing the number of data points averaged results in an incorrect assessment of the true condition of the pavement in relation to the intervention criteria. For example, Table 11 shows that 25 percent of SHN1 between reference stations 915/6.10 and 915/8.71 (site 4) had a wheelpath-averaged SFC less than 0.55 when derived from the source 5 m measurements, yet if 100- or 500-m reporting lengths are used, this pavement section is shown to have skid resistance characteristics that are neither better nor worse than the others investigated.

Accordingly, it is essential that the reporting length and selected intervention levels be consistent with the minimum pavement length that justifies resurfacing and shape correction. Typically, such a length ranges between 20 and 50 m, the lower value applying to urban roads, whereas the higher to rural roads. Furthermore, the source measurements should never be discarded as they can be useful for ranking road sections with nominally the same mean value of pavement defect parameter that is being considered.

Correlations Between Various RST Measures of Pavement Defect

Tables 12 to 16 show how IRI roughness, rutting, and RMS fine and rough measures of macrotexture relate to each other for the five road sections presented in Table 3.

First, it can be seen that the degree of correlations between the various defect measures varies considerably

TABLE 12 Correlation Between RST Measures of Pavement Distress—Site 1, Northbound and Southbound (southbound in parentheses)

	IRI	Rutting	Fine Macro	Rough Macro
IRI	1.0	0.193 (0.059)	0.299 (0.095)	0.248 (0.084)
Rutting		1.0	0.402 (0.073)	0.361 (0.051)
Fine Macro			1.0	0.851 (0.860)
Rough Macro				1.0

TABLE 11 Averaging Effect on SFC Descriptive Statistics

Site No.	Averaging Length	Average SFC Statistics				Length of Section with SFC <0.55	
		Mean	Std Dev	Minimum	Maximum	Lane-Kilometres	% of Surveyed Length
1	5	0.71	0.03	0.44	0.82	0.25	3.7
	100	0.71	0.02	0.65	0.78	0	0
	500	0.71	0.01	0.68	0.72	0	0
2	5	0.78	0.04	0.62	0.89	0	0
	100	0.78	0.02	0.75	0.82	0	0
	500	0.78	0.01	0.76	0.80	0	0
3	5	0.69	0.05	0.46	0.79	0.65	25
	100	0.69	0.03	0.62	0.74	0	0
	500	0.69	0.007	0.68	0.70	0	0
4	5	0.70	0.06	0.35	0.89	1.15	19
	100	0.70	0.05	0.60	0.83	0	0
	500	0.70	0.04	0.61	0.78	0	0
5	5	0.70	0.03	0.58	0.78	0	0
	100	0.70	0.02	0.66	0.75	0	0
	500	0.70	0.01	0.69	0.73	0	0

TABLE 13 Correlation Between RST Measures of Pavement Distress—Site 2, Northbound and Southbound (southbound in parentheses)

	IRI	Rutting	Fine Macro	Rough Macro
IRI	1.0	0.063 (0.114)	0.050 (0.024)	0.080 (0.038)
Rutting		1.0	0.017 (0.015)	0.038 (0.018)
Fine Macro			1.0	0.931 (0.960)
Rough Macro				1.0

TABLE 14 Correlation Between RST Measures of Pavement Distress—Site 3, Northbound and Southbound (southbound in parentheses)

	IRI	Rutting	Fine Macro	Rough Macro
IRI	1.0	0.189 (0.041)	0.042 (0.006)	0.061 (0.008)
Rutting		1.0	0.019 (0.010)	0.032 (0.032)
Fine Macro			1.0	0.945 (0.782)
Rough Macro				1.0

TABLE 15 Correlation Between RST Measures of Pavement Distress—Site 4, Northbound and Southbound (southbound in parentheses)

	IRI	Rutting	Fine Macro	Rough Macro
IRI	1.0	0.562 (0.497)	0.057 (0.114)	0.135 (0.156)
Rutting		1.0	0.159 (0.050)	0.271 (0.081)
Fine Macro			1.0	0.682 (0.957)
Rough Macro				1.0

TABLE 16 Correlation Between RST Measures of Pavement Distress—Site 5, Northbound and Southbound (southbound in parentheses)

	IRI	Rutting	Fine Macro	Rough Macro
IRI	1.0	0.032 (0.482)	0.000 (0.103)	0.000 (0.025)
Rutting		1.0	0.113 (0.168)	0.096 (0.100)
Fine Macro			1.0	0.957 (0.917)
Rough Macro				1.0

between a northbound and southbound run for a particular section. This result suggests that the lane direction of the pavement condition survey should be recorded in RAMM.

Second, RMS fine and rough macrotexture measures are well correlated and so only one measure of macrotexture appears necessary.

Third, significant correlations sometimes occur between IRI roughness and rutting, and rutting and macrotexture. Such correlations should be investigated further to establish whether they can be used to distinguish between roughness effects caused by rutting and shoving, and macrotexture losses caused by scabbing and flushing.

Further Work

The timing of the RST and SCRIM surveys necessitated the use of the 1989 walkover visual survey data for the comparative study. Unfortunately this was far from ideal. First, the automatically and manually acquired pavement condition data were separated by more than 1 year. Second, and more important, 1989 coincided with the first year that RAMM condition rating surveys were conducted on New Zealand's state highway network, so the data were not as accurate as they should be because of unfamiliarity by some raters with correct measurement procedures. This problem has now been addressed by the introduction of formal training and implementation of QA procedures.

Accordingly, a contract has been let in September 1993 by Transit New Zealand for a more extensive 3-year trial by a vehicle equipped to determine longitudinal roughness, flushing, and rutting. This 3-year trial will enable the year-to-year variability of automatically collected condition rating data to be assessed and a more extensive investigation of the degree of correlation between the automatic and manual data collection methods.

Conclusions

The reported study, using RAMM pavement condition ratings and 100 percent sampling of various pavement condition parameters at highway speeds using continuous measuring devices such as the SCRIM and the RST, has led to the following conclusions:

1. Variability in manually acquired condition rating data can be minimized through appropriate attention to the training of raters and implementation of quality assurance procedures designed to ensure consistency in the measurement of pavement defects.

2. It was found that the RST's pavement condition indexes investigated (rutting and macrotexture) do not generally correlate well with traditional pavement condition ratings obtained using visual or manual survey procedures. Nevertheless, because of the measurement repeatability that can be achieved with the RST, it has value as a survey tool, particularly in regard to periodic monitoring of the network to detect pavement condition changes over 3- to 5-year intervals. The resulting data bases would be useful for assessing the effectiveness of existing design, construction and maintenance procedures, and also for deriving and validating pavement distress prediction models. Survey lengths corresponding to only 10 to 20 percent of the total sealed state highway network should be adequate for such a purpose.

3. The value of making continuous texture depth measurements could be considerably enhanced through relating these measurements to road geometry, in particular horizontal curvature, as scabbing (chip loss) commonly occurs on the outside of the curve because of vehicle cornering forces and by making an additional texture depth measurement in the center of the lane to supplement those in the wheelpaths so that the degree of texture loss can be established.

4. With 100 percent sampling, an immense amount of data are collected. These data can take a considerable time to summarize and analyze. To overcome this problem, outputs from continuous measuring devices are reported over length intervals that are typically several orders of magnitude larger than the sampling length interval. A limited study of the influence of reporting length was performed on SCRIM data. This showed that mean values were not affected by the averaging process. However, as a consequence of the reduction in the degree of scatter in the data, the standard deviation reduces and the minimum and maximum values converge as the reporting length is increased. More important, the loss of resolution associated with increasing the number of data points averaged results in an incorrect assessment of the true condition of the pavement in relation to intervention criteria. Accordingly, it is essential that the reporting length and intervention levels selected be consistent with the resolution of the measuring device and also the minimum pavement length that justifies resurfacing and shape correction.

ACKNOWLEDGMENT

The permission of Transit New Zealand to present this paper is gratefully acknowledged.

REFERENCES

1. Lay, M. G. *Handbook of Road Technology—Volume 1, Planning and Pavements.* Gordon and Breach Science Publishers, New York, 1986.
2. Arnberg, P. W. *The Laser Road Surface Tester (RST): Synopsis of Presentations in Sydney and Melbourne, Australia, March 1986.* VTI Report 114. Linkoping, Sweden, 1986.
3. Salt, G. Research on Skid Resistance of the Transport and Road Research Laboratory. In *Transportation Research Record 622,* TRB, National Research Council, Washington, D.C., 1977, pp. 26–38.
4. Prem, H. NAASRA Roughness Meter Calibrations Via the Road-Profile-Based International Roughness Index (IRI). In *ARRB Research Report ARR No. 164,* Vermont South, Victoria, Australia, 1989.
5. Local Government Training Board. *RAMM Road Condition Rating Manual.* National Roads Board, Wellington, New Zealand.
6. Cenek, P. D. *Appraisal of Automatic and Manual Condition Rating Procedures.* Central Laboratories Report 92-29307, Lower Hutt, New Zealand, 1992.
7. Jameson, G. W., E. Baran, and G. N. Sheldon. Australian Experience with the Swedish Laser Road Surface Tester. *Proc. 14th ARRB Conference, Part 8,* Vermont South, Victoria, Australia, 1989, pp. 244–259.
8. Cenek, P. D. *Evaluation of the Laser RST System.* Central Laboratories Report 88-A9947/1, Lower Hutt, New Zealand, 1988.

The views expressed are those of the authors and should not necessarily be construed as being those of their respective organizations.