Improved Road Segmentation for Pavement Management Systems in New Zealand

Transfund New Zealand Research Report No. 218

Improved Road Segmentation for Pavement Management Systems in New Zealand

Robert Davies, Statistics Research Associates Ltd
Marian Loader, Opus Central Laboratories

Transfund New Zealand Research Report No. 218

ISBN 0-478-25076-2
ISSN 1174-0574.

Davies, R., Loader, M. 2001. Improved road segmentation for pavement management systems in New Zealand. *Transfund New Zealand Research Report No 218.* 64pp.

Dr Robert Davies, Statistics Research Associates Ltd, PO Box 12 649, Wellington, New Zealand

Marian Loader, Opus International Consultants Ltd, Opus Central Laboratories, PO Box 30 845, Lower Hutt, New Zealand

Keywords: analysis, change points, cusum method, management, New Zealand, normal scores transform, pavement, principal components analysis, roads, statistical analysis, treatment lengths, segmentation

An Important Note For The Reader

The research detailed in this report was commissioned by Transfund New Zealand.

Transfund New Zealand is a Crown entity established under the Transit New Zealand Act 1989. Its principal objective is to allocate resources to achieve a safe and efficient roading system. Each year, Transfund New Zealand invests a portion of its funds on research that contributes to this objective.

While this report is believed to be correct at the time of its publication, Transfund New Zealand, and its employees and agents involved in its preparation and publication, cannot accept any liability for its contents or for any consequences arising from its use. People using the contents of the document, whether directly or indirectly, should apply and rely on their own skill and judgement. If necessary, they should seek appropriate legal or other expert advice in relation to their own circumstances, and to the use of this report.

The material contained in this report is the output of research and should not be construed in any way as policy adopted by Transfund New Zealand but may form the basis of future policy.

Contents

Executive Summary

Introduction

The introduction of more sophisticated pavement management systems, which utilise predictive deterioration modelling (PDM) and 100% sampling of road network condition by automatic means, has resulted in the need to accurately divide the road network into like-performing sections.

Lengths of road that receive the same treatment, and are expected to have uniform condition throughout the pavement life, are commonly referred to as "treatment lengths". PDM relies on the uniformity of these lengths to predict the best treatment and timing for optimal performance from the treatment length. PDM methods use the summary statistics (mean and sometimes the standard deviation) of the condition parameters to model the typical condition of the treatment length. If the treatment length contains a large variation in the condition values, then the summary statistics will not reflect the true condition of any of the treatment length. This could result in the generation of a programme of treatments which does not meet the needs of any of the treatment length. This may result in poor treatment timing, or worse, it may recommend the wrong treatment type.

A robust methodology has been developed, during 2000-2001, to locate change points (i.e. clearly defined breaks) in New Zealand road pavements where significant changes in pavement performance occur. The methodology has considerable potential for generating treatment length subdivisions for new or existing road networks, or for dividing road networks into like-performing (homogeneous) sections for the purposes of deterioration modelling and asset management.

The application of the latest process control and time series statistical techniques, such as the cusum change point method and auto-regressive modelling, was therefore investigated using RAMM[1] data for the entire length of State Highway (SH)1. The result was the development of this robust methodology that involves the use of pavement and surfacing construction data to perform an initial segmentation. After the initial segmentation is made, the cusum change point method is applied to find statistically significant breaks in the pavement where changes occur in roughness, rutting, texture and skid resistance measurements stored in RAMM HSD (High Speed Data) tables.

Methodology

The change point method identifies breaks caused by missing construction data in inventory tables, as well as changes in pavement performance within segments, which have not been previously identified by inspection.

These breaks are worth including in statistical analysis or deterioration modelling as they represent statistically significant values in the pavement condition data. This analysis may result in changes to the timing or choice of treatments assigned to those pavements.

From the perspective of pavement maintenance management and treatment monitoring, the number of additional breaks found from the pavement HSD condition data is quite small so it would be feasible to check these manually. Ideally, adjacent treatment lengths should be amalgamated when their pavement performance measures show little difference.

---

[1] RAMM Road Asset Maintenance and Management

The spacing between Falling Weight Deflectometer (FWD)-derived pavement strength measurements (obtained by structural number (SNP)) is too large to use in the change point methodology. However, SNP measurements are useful in helping to decide which road segments can be amalgamated.

The initial segmentation methodology, using known breaks in construction and the change point methodology, could also be used to create an algorithm for generating treatment lengths for new or existing networks of roads.

Conclusions

Three main applications result from this robust segmentation methodology as follows:

- Creation or validation of treatment lengths for monitoring treatments and maintenance to the pavement sections that comprise road networks.
- Identification of homogeneous stretches of roads for calibrating current pavement deterioration models incorporated in NZ-dTIMS.
- Locating homogeneous roads with known pavement characteristics suitable for long-term pavement performance monitoring, to generate data for developing new deterioration models.

Recommendations

The cusum change point methodology developed and presented in this report can be applied to any pavement condition data held on databases for New Zealand roads to identify statistically significant break points in their construction. Field inspections would confirm the reasons for these break points. Possible reasons include:

- Surface or pavement condition;
- Missing construction data in inventory database;
- Existing treatment lengths.

A field trial could make use of the break points found in this project, or use the PDM method to identify break points in a selected region. The benefits of conducting a field trial to investigate the factors contributing to break points include:

- Calibrating the sensitivity of the cusum method to avoid creating spurious break points.
- Verifying the causes of extreme events in the condition data on the road and thus estimating the influence of the smoothing methods applied to the 20m-aggregated condition data. Many of the causes of extreme events are acknowledged in the condition data collected, for example, for bridge abutments and railway crossings.
- Assessing the accuracy of measurements made in the field. At present the accuracy within RAMM for all maintenance, and pavement and surfacing activities is assumed to be ±50m. A field trial would quantify the level of accuracy of the data stored in RAMM with respect to the position of road features, road maintenance, and construction changes in pavement and surface. Any segments smaller than the limits of measurement accuracy should be excluded from the segmentation process.

Future work, which would benefit from the field trial results, includes:

- Fine tuning of the change-point algorithm. This is an iterative process, and the ideal coefficients cannot be calculated until the majority of the break points have been identified.
- Further investigation of the auto-correlation structure of the data.
- Investigation of criteria for amalgamating adjacent treatment lengths.
- The current RAMM treatment length generator should be reviewed, based on the initial segmentation methodology presented in this report.

Abstract

A robust methodology has been developed, in 2000-2001, to locate change points (i.e. clearly defined breaks) in New Zealand roads where significant changes in pavement performance occur. Both pavement condition and construction data are required to apply the methodology, which utilises standard time-series and process control statistical techniques, i.e. auto-correlations and cusum methods respectively. The methodology has considerable potential for generating treatment length subdivisions for new or existing road networks, or for dividing road networks into like-performing (homogeneous) sections for the purposes of deterioration modelling and asset management, in existing pavement management systems.

1. Introduction

The introduction of more sophisticated pavement management systems, which utilise predictive deterioration modelling (PDM) and 100% sampling of road network condition by automatic means, such as the modelling program NZ-dTIMS, has resulted in the need to accurately divide the road network into like-performing sections.

Lengths of road that receive the same treatment, and are expected to have uniform condition throughout the pavement life, are commonly referred to as "treatment lengths". PDM relies on the uniformity of these lengths to predict the best treatment and timing for optimal performance from the treatment length. PDM methods use the summary statistics (mean and sometimes the standard deviation) of the condition parameters to model the typical condition of the treatment length. If the treatment length contains a large variation in the condition values, then the summary statistics will not reflect the true condition of any of the treatment length. This could result in the generation of a programme of treatments which does not meet the needs of any of the treatment length. This may result in poor treatment timing, or worse, it may recommend the wrong treatment type.

In this paper, road sections with uniform response characteristics are called "homogeneous lengths". A segmentation method that can split roads into homogeneous lengths will aid pavement management and deterioration modelling in three ways:

- Treatment length generation and validation – such a method can split existing treatment lengths where the pavement condition variables are found to be not uniform.
- Calibration of existing models – the method can identify sections of road with uniform condition, and known pavement and surfacing layers that are suitable for calibration of pavement deterioration models.
- Development of new models – the sensitivity of the segmentation statistics can be adjusted to identify sites with uniform condition, low within-site variability, and specific pavement composition, that are suitable to develop new models or test the pavement deterioration model incorporated in NZ-dTIMS.

Any segmentation method will require continuous data that reflects the condition of the pavement and its surface. The high-speed condition data (HSD) which is collected from all state highways managed by Transit New Zealand (Transit NZ), and is readily available, provides this kind of continuous data. Therefore it is logical to use the HSD data to trial the segmentation methods that can be used to improve the standard of the New Zealand road network. The trials are also to ascertain if the methods are suitable for applying to our road network.

Two subsets of Transit NZ roads were selected for the development of the methods:

- State Highway 1 (SH1) from the North and South Islands.
- All the roads in the Transit NZ Northland region where a Falling Weight Deflectometer (FWD) survey had been made in 1999.

Such data was not common elsewhere in New Zealand at the time, though many more New Zealand roads have been surveyed since.

The cusum segmentation method was selected to identify segments where the condition of the pavement or surface was not uniform. It requires variables with close spacing between measurements in order to accurately identify the point where a statistically significant change occurred. The HSD was suitable for this method but the Northland FWD readings were too far apart to be included.

A robust methodology was then developed, during 2000-2001, to locate change points (i.e. at clearly defined breaks) in New Zealand road pavements where significant changes in pavement condition occur. The methodology has considerable potential for generating treatment length subdivisions for new or existing road networks, or for dividing road networks into like-performing (homogeneous) sections for the purposes of deterioration modelling and asset management.

This report outlines the segmentation methodology that was developed using the SH1 data. The Northland data was used as a test case to establish whether or not the FWD values agreed with the splits generated by the HSD.

2. Data Selection

To represent a cross section of regions and traffic loadings, as well as to maximise the number of roads whose structural numbers (SNP), i.e. their pavement strengths, were known through measurement, two databases were created for the analysis. They were of:

- State Highway 1 (SH1), because it runs the length of both North and South Islands, and thus provides a cross section of data throughout New Zealand, and through the full range of environmental conditions. It also carries the highest traffic volumes in each Transit NZ region.
- The roads in the entire Transit NZ Northland region, because an FWD survey had been made there (in 1999) to measure pavement structural strength (SNP). The SNP values were calculated using the adapted New Zealand structural number formula (Salt & Stevens 2001). However, the interval of measurement for this data was 250m. This interval was found to be inadequate when the change point segmentation method was investigated later for this research.

Where possible, the data used in this report was extracted from the Transit NZ State Highway database stored in RAMM (as at October 2000). SNP values were calculated, using Salt's NZ formula, without layer depths (Salt & Stevens 2001), from FWD surveys that had been stored in an Access format at the time of extraction. The final database contained data pertaining to:

- Location data, i.e. State Highway (SH), Reference Station (RS), Route Position (RP).
- Pavement and surface construction data, i.e. age, materials, depth, and width.
- Pavement use data, i.e. hierarchy, Average Annual Daily Traffic (AADT), percentage of heavy vehicles.
- Pavement condition data, i.e. skid resistance, texture depth, rut depth, roughness, and SNP.

The data have been presented as averages over intervals that are nominally 20m in length. Although a small number of observation intervals are shorter than 20m, the number is very small. At present they are treated as normal intervals, and although this is unlikely to make a significant difference to the results, they could also be omitted.

All condition data were summarised over 20m intervals for both increasing and decreasing directions (from Reference Points) for the three financial years of 1997/98, 1998/99 and 1999/2000.

3. Method

3.1 Overall Approach

The SH1 database was used to develop and test a robust segmentation methodology. The methodology was then applied to the Northland region database to determine if the SNP data fitted in with the segments that had been identified by the FWD survey, and to test the homogeneity of the existing treatment lengths.

3.2 Segmentation Method

The Steps used in the segmentation method proceed as in the following summary:

1. Use the road construction and road use data to make a preliminary division into segments.
2. Carry out an exploratory analysis of the surface and condition data.
3. Estimate statistical parameters such as auto-correlation and cross-correlation[1] of the condition data to develop performance tests.
4. Carry out the performance tests to identify segments that are not homogeneous, but excluding trends and outliers.
5. Review the estimates of the statistical parameters in view of the new segmentation, and repeat the earlier analysis if appropriate.
6. Test if any road construction and road use data appear to be introducing breaks which do not have corresponding changes in the condition data. If any particular road construction or road use database field appears to be introducing spurious breaks, repeat the analysis without the breaks related to this field.
7. Compare cusum[1] break points to existing treatment lengths.
8. Compare cusum break points with structural strength measurement data for state highways in Northland region.

[1] Definitions are given in Glossary.

4. Statistical Analysis of SH1 Data

The statistical analysis described in this Section 4 details the steps required to find the significant break points within the pavement use and condition data. The statistical methods employed are auto-correlation, cross-correlation, and cusum techniques. The results are given here in detail so that potential users can apply or adapt them according to their own requirements. The break points obtained using the cusum analysis method are discussed in Sections 5 and 6.

Much of the analysis concentrates on the data from the year 2000 for SH1N and SH1S (North and South Islands respectively) in the increasing direction only. Data for the decreasing direction has been used only as a final check on the analysis (see Section 4.8). Appendix A1 explains the techniques in more detail.

4.1 Initial Segmentation

An initial segmentation was made based on the pavement construction and road use data to subdivide the highways into shorter lengths. Without this initial segmentation, any segmentation method is unlikely to yield reliable results as there are too many spurious changes. It also has the advantages of reducing the computation time, reducing the effect of outliers, and simplifying the calculation of auto-correlation.

Break points found statistically will be subject to random error, whereas the locations of the changes in road construction and road use will be subject only to measurement and recording errors.

Table 4.1 contains the logic sequence used to create the initial segmentation based on the pavement construction and road use data. Gaps in the surface-date variable will create a segment break. For all other variables missing values are treated as the same as the last known value and do not create break points at a change from known to unknown.

Table 4.1 Logic sequence used to create the segmentation algorithm.

Variable	Difference in variable that will trigger a break	Break between known & missing data
Surface Date	> 90 days	Yes
Pavement Date	> 365 days	No
Lanes	Any change	No
Width	> 2 metres	No
Urban / Rural	Any change	No
ADT estimate	> Factor of 1.5 change and increment >100	No
Surface Width	> 2 metres	No
Surface Material	Any change	No
Surface Type	Any change	No
Number of Pavement Layers	>1; but > 10 was treated as missing	No

The variables chosen and the magnitude of the changes should be reviewed in any repeat analysis.

Width and surface width fields may not be appropriate to use where no change in the number of lanes has occurred. However slow and heavy vehicles do use wide verges to allow traffic to pass, and therefore it is not unreasonable to expect that the wear patterns may be different where the road width changes by more than a standard car width. Also, additional changes in road width may indicate changes in construction as well as use. It has been suggested that where there is no change in lane number, the width of the sealed verge would not generally change the driver behaviour, and therefore the wheel track wear would be similar. We recommend that any future work exclude surface width in the initial segmentation step.

The initial segmentation does not require all of chip size, previous surface layer dates, and pavement depth, to be included because:

- Chip size breaks coincide with surface date break points.
- Previous surface layer dates introduced many breaks, but they may not influence the current performance measures unless they coincide with break points for other variables. To avoid creating humps in the road at resurfacing joins, surface layers are overlapped by approximately 10m. Therefore, including previous surface layers in initial segmentation could create artificial break points.
- Typically, pavement depth has a large number of missing or artificial values. Hence it is difficult to estimate total pavement depth reliably.

Table 4.2 summarises the number of breaks made by each of the variables in the North and South Islands for which 2495 and 1879 segments were generated in the initial segmentation process respectively. The average segment length was 20 observation intervals in the North and 25 in the South Island. This seems to be mainly related to the larger number of passing lanes.

Table 4.2 Breaks generated in initial segmentation phase for SH1, North and South Islands.

Variable	Number of additional breaks caused by variable	
	North Island	South Island
Surface Date	1133	1130
Pavement Date	342	305
Lanes	265	84
Width	121	66
Urban/ Rural	53	37
ADT estimate	10	4
Surface Width	49	48
Surface Material	128	77
Surface type	136	78
Number of Pavement Layers	68	50
Total	**2495**	**1879**

Figure 4.1 Cumulative distributions of segment lengths after initial segmentation.

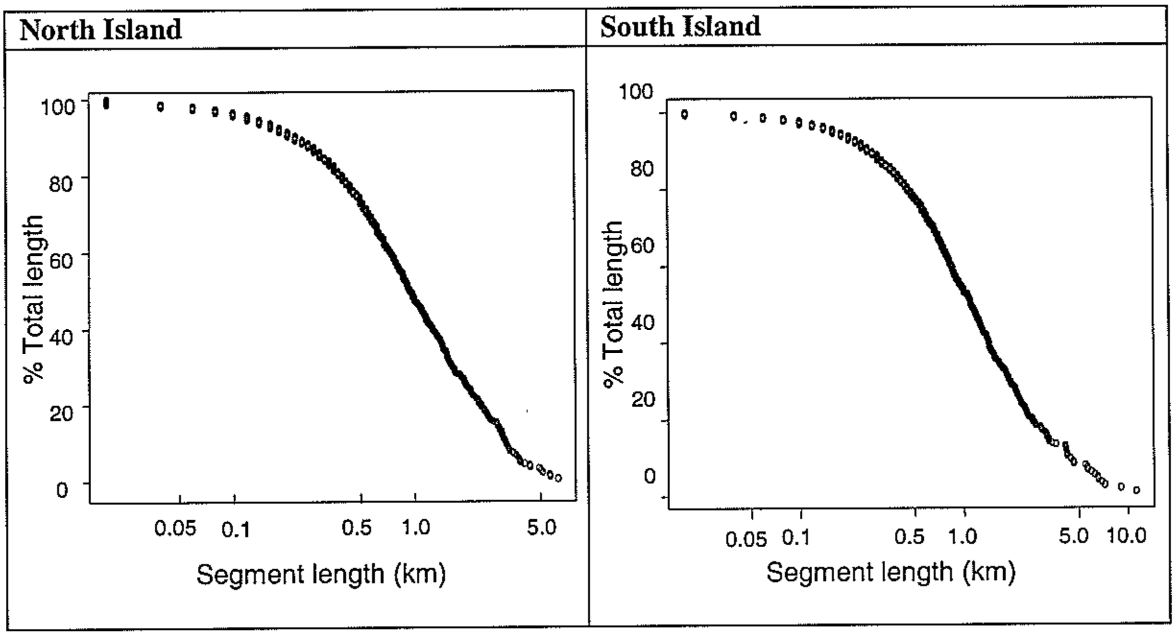
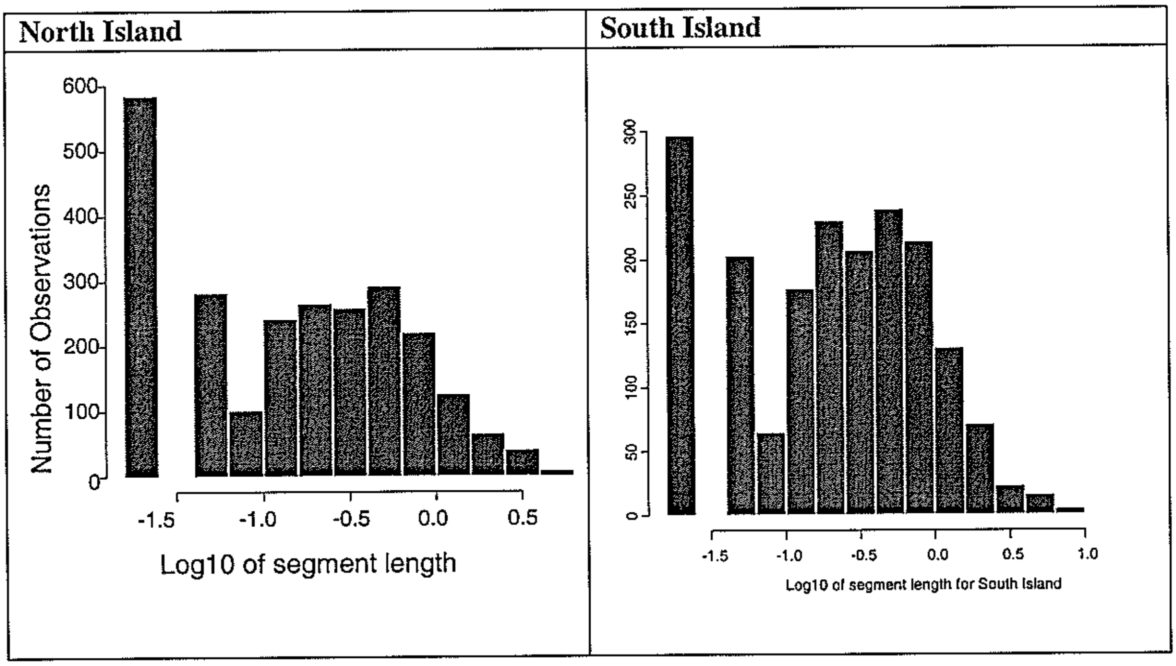


Figure 4.2 Frequency distribution of segment lengths after initial segmentation.



However, many short segments were only one or two observation intervals long (20-40m), some of which will be artificial caused by data entry or measurement errors. At present the accuracy within RAMM for all maintenance, and pavement and surfacing activities, is assumed to be ±50m (G. Hart, pers.comm., July 2001). Possibly segments of up to 100m in length may have been created because of variations in measurement.

Figures 4.1 and 4.2 show the cumulative and frequency distributions of SH1 data in the increasing direction for both North and South Islands.

The segment lengths tend to be shorter in the North Island. The plots in Figure 4.1 show that approximately 70% of SH1N (North) and 80% of SH1S (South) have a segment length of 0.5km (25 observations), or more. Around 45% of SH1N, and 50% of SH1S, have a segment length of at least 1km.

4.2 Reducing the Number of Short Lengths

The peak on the left of each of the two graphs in Figure 4.2 corresponds to a segment of one observation interval (i.e. 20m). Joining this one observation length segment to one of its neighbours would be desirable to reduce the number of lengths requiring processing. If the segments on either side of the short segment meet the criteria listed in Table 4.1, then the three sections could be considered as one. As instances where this occurs is low, being 9 for the South Island, and 15 for the North, combining data does not change the overall picture, and this process was not included in the analysis.

Given the accuracy of the location data (within 50m), another method to eliminate the spurious 20-40m segments could be to join them to one of their neighbours.

This completes Step 1 of the segmentation methodology outlined in Section 3.2 of this report.

4.3 HSD Condition Data Analysis

Mean and variance of texture depth (mm) and rut depth (mm), mean roughness (IRI), and skid resistance (NZ MSSC[2]), were selected from the HSD condition data as candidates for the segmentation process.

Where the distributions were not normal, a transform of the variables was needed to ensure their distributions were approximately normal. Many statistical methods, including the segmentation and time series methods used in this analysis, require the normality assumption to hold. Where the transform would require the logarithm of a zero value, the zero has been replaced by half the minimum of the non-zero values.

---

[2] New Zealand Mean Summer SCRIM Coefficient
SCRIM – Sideways-force Coefficient Routine Investigation Machine

The condition data for texture, IRI, and skid resistance were averaged over the left and right wheel paths before they were transformed. The transforms used to normalise the variables are listed in Table 4.3. Histograms of the transformed variables are included in Figure 4.3. Some outliers are present in the transformed data, but otherwise the data is approximately normal and statistical methods can be applied.

Table 4.3 Transforms to normalise the condition measures of the variables.

Condition Measurement	Transform applied
Mean Texture Depth (mm)	No Transform
Variance of Texture Depth (mm)	Log10
Mean Rut Depth (mm)	Square root (SQRT)
Variance of Rut Depth (mm)	Log10
Roughness (IRI)	Log10
Skid Resistance	No transform

This Step 2 completes the exploration of pavement condition data (Section 3.2).

4.4 Estimate of Cross-Correlation Structure

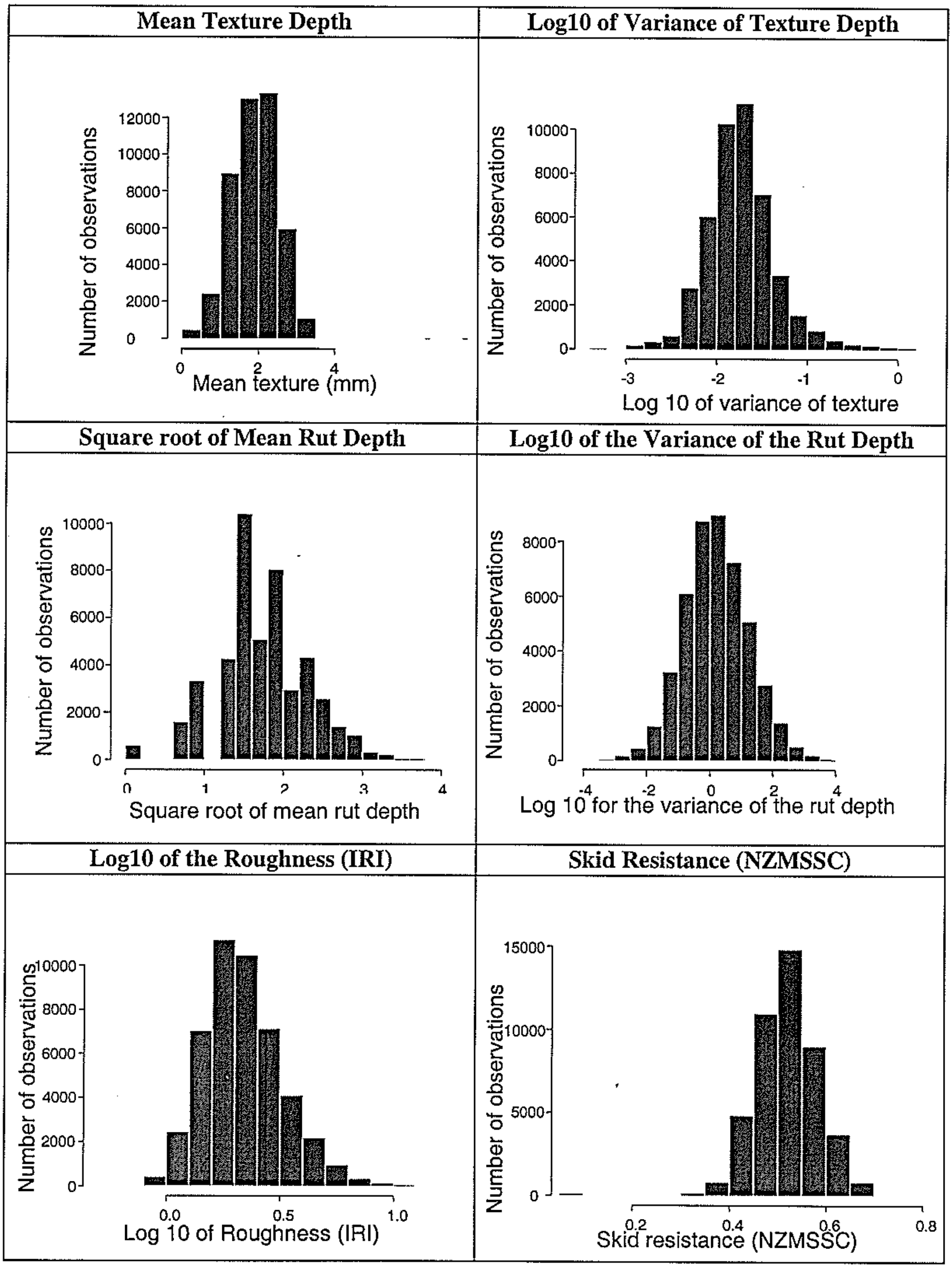
To determine if the performance measures developed from the condition data were independent, a cross-correlation analysis was performed. Table 4.4 contains the correlations between the variables using the S-plus trimmed correlation function (Anon 1999), with the trim parameter set to 0.1. This function trims off the extreme values of the variables, so analysis is less susceptible to effects of outliers than a simple correlation function would be (see Appendix A1.9 of this report).

For Table 4.4, correlations greater than 0.2 are shown in bold. It is not possible to adjust for all the correlations between variables, but it is important to focus on and account for the largest variations. Positive correlations indicate that both variables increase together, negative correlations indicate that one variable increases as the other decreases.

Table 4.4 Correlations between transformed condition variables.

Variable	Mean Texture	Log10 var of Texture	SQRT mean Rut depth	Log10 var of Rut depth	Log10 Roughness (IRI)	Skid Resistance
Mean Texture		**0.33**	0.02	−0.11	−0.15	0.11
Log10 var of Texture	**0.33**		0.16	**0.25**	**0.33**	−0.06
SQRT mean Rut depth	0.02	0.16		**0.64**	**0.27**	−0.03
Log10 var Rut depth	−0.11	**0.25**	**0.64**		**0.5**	−0.02
Log10 Roughness (IRI)	−0.15	**0.33**	**0.27**	**0.5**		−0.08
Skid Resistance	0.11	−0.06	−0.03	−0.02	−0.08	

Figure 4.3 Histograms of the transformed HSD condition variable data.



The variables with high linear correlation are:

- Mean texture depth and variance of texture depth;
- Mean rut depth, variance of rut depth, and IRI.

Skid resistance is not highly correlated to any of the other variables.

A principal component analysis was used to explain the maximum amount of variance with the smallest number of components. This has the advantages of reducing the number of computations in subsequent stages and reducing statistical noise (random error), because the data contained in later components may be mostly attributed to statistical noise. The principal components analysis using the pavement condition data produces three identifiable components (see Table 4.5 in which six components (1' – 6') are generated).

It is better practice to build principal components from a smaller number of variables, using those with larger contributions. In Table 4.5, the variables whose contributions to components were at least 0.3 in magnitude are shown in bold.

Table 4.5 Principal component analysis of condition variables.

Component	**1'**	**2'**	**3'**	**4'**	**5'**	**6'**
Proportion of Variance	0.55	0.21	0.13	0.08	0.02	0.01
Eigen value	2.12	1.30	1.02	0.82	0.44	0.29
Mean Texture	0.02	**–0.78**	0.01	**0.30**	**-0.54**	0.11
Log10 Texture variance	**–0.35**	**–0.56**	**–0.32**	–0.28	**0.62**	–0.06
Sqrt Rut Mean	**–0.51**	0.06	0.27	**0.56**	0.10	**–0.58**
Log10 Rut variance	**–0.60**	0.13	0.19	0.12	0.01	**0.76**
Log10 Roughness (IRI)	**–0.50**	0.10	–0.15	**–0.58**	**–0.55**	–0.28
Skid Resistance	0.08	–0.22	**0.87**	**–0.41**	0.10	–0.06

The first row shows how much of the variance of each the six variables (1' to 6') is explained by each component. The first three components (1'-3') account for 89% of the variance between them, and essentially represent the variation caused by roughness, surface texture and skid resistance. These three components are:

1' Roughness, which is a variable made up from a weighted combination of the average roughness (IRI), the variance in texture depth, and the mean and variance of rut depth.

2' Texture, which is a weighted mean of the mean and variance of texture depth.

3' Skid Resistance, which is the mean of left and right wheel path skid resistance.

The remaining components (4'-6') may be mostly attributed to statistical noise.

Having a smaller number of components makes the data more manageable and should increase the sensitivity of tests, provided that summarising has not deleted relevant information.

Another report concerned with using laser profilers (Fong 1999) found that roughness and rutting measurements were affected by texture depth, particularly on chip surfaces where the lasers appear to detect the gaps between the chips. Table 4.4

shows a positive correlation between the variance in texture and both roughness and rut variance. However the correlation between *mean texture* and *roughness* measures is small and negative.

The most likely explanation for this apparent disagreement is that the cross-correlations have been calculated before allowing for the variation in mean levels between segments. This means that the variation between the sections is dominating the calculation. It is appropriate to calculate the cross-correlations before allowing for the variation between segments because the aim of the analysis is to distinguish between sections. Hence the cross-correlations will be related to differences between segments.

Most likely, older segments tend to have lower texture and more roughness and rutting, and this accounts for the low or negative correlations. This effect may overwhelm the effect of the laser profiler, and thus cause confusion with texture and roughness or rut depth. Variance of texture may be higher on the older roads which gives the positive correlation.

4.5 Reduction of Pavement HSD Condition Data

The change point tests described in the Appendix to this report depend on averages of consecutive observations. So the correlations in the analysis were based on medians of disjoint sets of five consecutive observations. The median is used rather than the average to reduce the effect of outliers. If this aggregating is not done, it is possible that the results will be distorted by the measurement errors in the observations.

It is important that the three components are approximately independent. Hence the texture variance variable (log10.tex.var) needs to be replaced with a variable that has the influence of the mean texture subtracted from it. This can be done using Equation 4.1, where a is chosen so that the resulting variable is uncorrelated with the mean texture (texmean).

$$a.l10.tex.var = log10.tex.var - a \times texmean \qquad \text{Equation 4.1}$$

The appropriate value of a is 0.277 for the North Island and 0.224 for the South Island, so the average value of 0.251 is used.

Tables 4.6 and 4.7 contain the correlations of the four roughness-related variables, for the North and South Island databases respectively. The correlations are similar for both islands.

For Tables 4.6 through to 4.9, contributions greater than 0.3 in magnitude are considered significant, and these numbers are shown in bold.

Table 4.6 Roughness correlated variables for SH1N, North Island.

Variable	a.l10.tex.var	sqrt.rutmean	log10.rut.var	log10.iri
a.l10.tex.var		0.237	**0.408**	**0.477**
sqrt.rutmean	0.237		**0.692**	**0.379**
log10.rut.var	**0.408**	**0.692**		**0.615**
log10.iri	**0.477**	**0.379**	**0.615**	

Table 4.7 Roughness correlated variables for SH1S, South Island.

Variable	a.l10.tex.var	sqrt.rutmean	log10.rut.var	log10.iri
a.l10.tex.var		0.175	**0.391**	**0.452**
sqrt.rutmean	0.175		**0.693**	0.296
log10.rut.var	**0.391**	**0.693**		**0.567**
log10.iri	**0.452**	0.296	**0.567**	

To find the best summary of these four variables, a principal components analysis was performed. Results of the component analysis are given in Tables 4.8 and 4.9.

Table 4.8 Principal component analysis of roughness variables for SH1N, North Island.

Component	1'	2'	3'	4'
Proportion of Variance	0.609	0.208	0.123	0.060
a.l10.tex.var	**0.416**	**0.711**	**0.566**	
sqrt.rutmean	**0.484**	**−0.599**	**0.365**	**0.524**
log10.rut.var	**0.573**	−0.253		**−0.778**
log10.iri	**0.514**	0.270	**-0.737**	**0.345**

Table 4.9 Principal component analysis of roughness variables for SH1S, South Island.

Component	1'	2'	3'	4'
Proportion of Variance	0.581	0.230	0.131	0.058
a.l10.tex.var	**0.413**	**0.658**	**0.624**	
sqrt.rutmean	**0.478**	**−0.625**	0.266	**0.557**
log10.rut.var	**0.589**	−0.245		**−0.769**
log10.iri	**0.503**	**0.341**	**−0.735**	**0.301**

The first row shows how much of the variance of each the four variables are explained by each component. The values of 61% for the North Island and 58% for the South Island show that component 1 dominates but does not completely preclude the possibility of useful information being in the other components. However, only component 1' was used in the analysis. The remaining four rows show how each component is made up of the variables we are analysing. By default, S-plus function leaves a blank when a contribution is small. This first component is close to being an average of the four variables but with a little more weighting for *log10.rut.var*. These weights are based on the assumption that the variables have been normalised by dividing by their standard deviations. Table 4.10 contains the weightings, after dividing the values for the first component by their own standard deviations.

Table 4.10 Variable component weightings for roughness component.

Component	a.l10.tex.var	sqrt.rutmean	log10.rut.var	log10.iri
North Island	0.1261	0.2583	0.4839	0.07933
South Island	0.1171	0.2387	0.4870	0.07297
Average	0.1216	0.2485	0.4854	0.07615

The average value was used in the calculations in this analysis for both the South and North Islands.

The change point analysis is based on the three principal components, which will be referred to as the *roughness*, *texture* and *skid* components respectively from this point. The cross-correlations of these three variables using medians of five consecutive measurements are in Tables 4.11 and 4.12, for the North and South Islands respectively.

Table 4.11 Correlations between the three change point variables, North Island.

Variable	Texture	Roughness	Skid
Texture		0.11	0.12
Roughness	0.11		–0.11
Skid	0.12	–0.11	

Table 4.12 Correlations between three change point variables, South Island.

Variable	Texture	Roughness	Skid
Texture		–0.03	0.16
Roughness	–0.03		–0.03
Skid	0.16	–0.03	

Some correlation still exists but it is probably small enough not to be a problem, and no attempt has been made to reduce the correlation between texture and skid.

4.6 Auto-Correlation Estimate

All the condition series are highly auto-correlated. By this we mean that a measurement on each observation interval is related to those on neighbouring intervals. An estimate of the auto-correlation is needed after allowing for the effect of segments.

Outliers and the differences between mean segment values in different segments complicate the auto-correlation.

To remove the cross-segment variation, the auto-correlation was calculated in each segment and results were averaged over all segments. This does not take out the effect of undetected segments but will be a much better estimate than calculating the auto-correlation over the whole series.

The auto-correlation for the three variables has been calculated with a lag of up to five observations:

- Lag 1 shows the correlation between adjacent observations (X_i and X_{i+1});
- Lag 2 shows the correlation between observation separated by one observation, X_i and X_{i+2}.

To reduce the effect of outliers, which are unusually large or small observation values, the Normal Scores Transform method, described in Appendix A1.8, and the Trimmed Correlation method, described in Appendix A1.9, were used.

The correlation decreases as the lag increases. The S-plus trimmed correlation function was used, with the trimming parameter set to 0.1 and with the normal scores transform. Segments of length less than 20m or with missing observations were omitted. The results are shown in columns 3 and 4 of Tables 4.13 and 4.14. The normal scores version gives slightly lower results than the trimmed version.

Appendix A1.5 describes simple models for adjusting the auto-correlation structure. The AR1 column in Tables 4.13 and 4.14 shows the auto-correlations when a first order auto-regressive model is fitted to the normal scores data, using the lag 1 auto-correlation to fit the auto-regressive parameter. The auto-correlations calculated by the model tend to decay faster than we actually observe.

The AR1, MA1 column shows the auto-correlations when a first order auto-regressive (AR1), and a first order moving average model (MA1) are fitted to the data using the first two auto-correlations to fit the parameters. In general, this fits better, but still decays more quickly than the observed values. However, because the additional change points have not been located, we expect that the observed auto-correlations to be a little higher than those predicted.

Appendix A1.5 introduces a correction factor when calculating the standard deviations of sums of auto-correlated series. Tables 4.15 and 4.16 show the values of the correction factor based on the auto-regressive model and the combined auto-regressive/moving average model. The correction factor, when using the combined auto-regressive/moving average model, generally has slightly larger values.

Table 4.15 Correction values for variables, based on AR1 and combined AR1 and MA1 models, for SH1N, North Island.

Variable	AR 1	AR 1, MA 1
Mean Texture Depth	2.118	2.200
Roughness	1.456	1.628
Skid Resistance	2.154	2.155

Table 4.16 Correction values for variables, based on AR1 and combined AR1 and MA1 models, for SH1S, South Island.

Variable	AR 1	AR 1, MA 1
Mean Texture Depth	2.142	2.212
Roughness	1.433	1.599
Skid Resistance	2.068	2.011

Table 4.13 Auto-correlations within variables for SH1N, North Island.

Variable	Lag	S-plus Trimmed Correlation	Normal Scores Correlation	AR 1Correlation	AR 1, MA 1 Correlation
Mean Texture Depth	1	0.717	0.636	0.636	0.636
	2	0.534	0.425	0.404	0.425
	3	0.407	0.305	0.257	0.284
	4	0.311	0.225	0.163	0.190
	5	0.252	0.174	0.104	0.127
Roughness	1	0.413	0.359	0.359	0.359
	2	0.245	0.203	0.129	0.203
	3	0.159	0.128	0.046	0.114
	4	0.115	0.088	0.017	0.065
	5	0.092	0.068	0.006	0.036
Skid Resistance	1	0.716	0.645	0.645	0.645
	2	0.515	0.417	0.417	0.417
	3	0.383	0.286	0.269	0.269
	4	0.283	0.201	0.174	0.174
	5	0.208	0.138	0.112	0.112

Table 4.14 Auto-correlations within variables for SH1S, South Island.

Variable	Lag	S-plus Trimmed Correlation	Normal Scores Correlation	AR 1* Correlation	AR 1, MA 1* Correlation
Mean Texture Depth	1	0.723	0.642	0.642	0.642
	2	0.532	0.430	0.412	0.430
	3	0.415	0.314	0.265	0.288
	4	0.331	0.243	0.170	0.193
	5	0.277	0.194	0.109	0.129
Roughness	1	0.394	0.345	0.345	0.345
	2	0.230	0.192	0.119	0.192
	3	0.155	0.129	0.041	0.107
	4	0.126	0.096	0.014	0.060
	5	0.105	0.077	0.005	0.033
Skid Resistance	1	0.677	0.621	0.621	0.621
	2	0.456	0.368	0.396	0.368
	3	0.321	0.234	0.239	0.218
	4	0.225	0.151	0.149	0.129
	5	0.167	0.098	0.092	0.076

* AR 1 – first order auto-regressive;
MA 1 – first order moving average model (described in Appendix A1.6).

Based on this analysis, the combined auto-regressive/moving average model and associated correction factors were incorporated in the segmentation methodology.

This completes Step 3 of segmentation methodology outlined in Section 3.2 of this report.

4.7 Recalibrating Auto-Correlation Coefficients

To check if the auto-correlation correction terms calculated in Section 4.6 of this report are realistic, we compared the predicted number of changes for a series of independent numbers and the observed number of significant results from the change point analysis, using the data for the three principal variables. To perform this test, all segments with at least 40 non-missing observations (i.e. segments that are at least 800m long) and with no construction breaks were used. These were divided in half and a normal scores test (Formula A4, Appendix A1.3), was used to test for a difference between the first half and the second half. Ideally there would not be a significant difference. The results are included in Table 4.17, and show that the number of observed significant results is much too high.

Table 4.17 Significance of normal scores test of auto-correlation estimates for SH1N and SH1S.

SH1	South Island		North Island	
Significant results	**Predicted**	**Observed**	**Predicted**	**Observed**
No. of observations	356	356	358	358
Significant at 50%	178	253	179	274
Significant at 5%	18	57	18	74
Significant at 1%	3.6	20	3.6	21
Significant at 0.1%	0.4	3	0.4	1

Part of the reason will be changes in road structure that have not been detected by earlier tests. However, most probably, the correction for the time-series structure in the series has not been sufficiently large. In some ways one cannot distinguish between these two possibilities. The dividing line between classifying a change in the pavement performance data as being related to the time-series structure or being related to a change in the pavement construction is somewhat arbitrary. Quite possibly, at least part of the time-series structure in the data is due to an hierarchy of small changes in construction. However, it is probably best to think of the time-series structure as something we are not trying to classify as breaks.

Therefore the correction factors introduced to allow for the time-series structure are not large enough. This does not invalidate the previous analysis but it does indicate that the number of breaks we have detected will contain some that are more related to the time-series structure than to construction changes.

To obtain agreement for the 50% significance level, an additional factor of 1.3 was found (using trial and error) to be required. Table 4.18 contains the results obtained when the correction factor of 1.3 is included.

Table 4.18 Significance of normal scores, with an additional auto-correlation correction factor of 1.3, for SH1N and SH1S.

SH1	South Island		North Island	
Significant results	Predicted	Observed	Predicted	Observed
No. of observations	356	356	358	358
Significant at 50%	178	167	179	191
Significant at 5%	18	11	18	7
Significant at 1%	3.6	2	3.6	0
Significant at 0.1%	0.4	2	0.4	0

Re-running the analysis of Section 4.6 with the revised auto-correlation correction factors for the three variables from Tables 4.15 and 4.16, we obtain improved estimates of auto-correlation figures as tabulated in Table 4.19.

Table 4.19 Revised auto-correlation factors for SH1N and SH1S.

Texture	Roughness	Skid
2.8677	2.0976	2.7078

The revised significance of construction break points tabulated in Table 4.20 still suggests that most breaks do correspond to changes in the HSD condition data. Nevertheless around 148 in the South Island and 186 in the North Island do not have corresponding changes in the HSD condition data.

Table 4.20 Revised significance of pavement construction breaks for SH1N and SH1S.

SH1	South Island	North Island
No. of observations	707	765
Not significant at 50% level	74	93
Significant at 5% level	394	403
Significant at 1% level	234	242
Significant at 0.1% level	112	107

This completes Step 6 of the segmentation methodology outlined in Section 3.2 of this report.

5. Segmenting SH1

5.1 Identifying Single Change Points on SH1, North & South Islands

This section describes how to identify changes within the segments that were defined in Section 3.2 (Step 4), using the three variables of roughness, texture, and skid resistance, described in Section 4.5.

The segmentation methodology was applied to segments containing 20 observation intervals (i.e. at least 400m) or more, and with no missing values in the three variables.

The method of identifying the breaks is described in Appendix A1.2 and A1.3. The normal scores transform (Appendix A1.8) has been applied to the data, and the correction factors calculated from auto-correlations (see Section 4.6) have been used.

The initial segmentation method described in Section 4.1 identified 641 segments on SH1S that meet the criteria, and they represent 75% of the total length of SH1S. On SH1N North Island, 700 of the 2495 segments identified were long enough to use the cusum segmentation method. This represents 73% of the total length of SH1N.

The results of the first sweep of the data for the South Island are tabulated in Table 5.1. Only segments yielding a statistically significant result at the 0.01 (1%) level, or better, in either the individual variables or the summary analysis of the three variables are reported.

The columns labelled *First Obs.* and *Last Obs.* identify the observation intervals, and the columns *Texture*, *Roughness* and *Skid (resistance)* show the levels of significance of the test for change points detected using the method described in Appendix A1.2.

Observation and segment numbers are related to the route position (RP) on SH1, starting at the northern end for the data sets of both the North and South Islands.

The column labelled *Summary* shows the significance of the overall test described in Appendix A1.4. Remember that the smaller these figures are, the more significant the result is, and the more certain we are that there really is a break point in that segment. Asterisks indicate missing values in the data. In the three columns for *Texture, Roughness, Skid*, 0.000 means less than 0.0005. Because of the method of computation for significance of the summary statistic, a significance for it that is less than 0.0005 is not possible to obtain.

Only segments showing a significance level of 0.005 or less in the *Summary* test were considered for further analysis. These are shown in bold. There are 30 such lines in total.

Table 5.1 Breaks obtained from first change point analysis of South Island data, in increasing direction from Reference Points.

Segment Number	First Obs.	Last Obs.	Significance Level			
			Texture	Roughness	Skid	Summary
113	1626	1681	0.329	0.005	0.859	0.040
123	1782	1837	0.584	0.006	0.786	0.030
207	2949	3037	0.010	0.754	0.622	0.040
258	**4134**	**4223**	**0.034**	**0.188**	**0.020**	**0.002**
339	6908	7118	0.009	0.036	0.301	0.020
386	8322	8441	0.806	0.041	0.010	0.008
409	**9065**	**9168**	**0.223**	**0.256**	**0.010**	**0.005**
466	**10565**	**10787**	**0.033**	**0.003**	**0.249**	**0.002**
507	11846	12296	0.645	0.003	0.981	0.020
548	13836	14163	0.000	0.264	*	*
559	**14339**	**14636**	**0.005**	**0.090**	**0.000**	**0.001**
581	**15177**	**15456**	**0.015**	**0.005**	**0.005**	**0.001**
604	**15928**	**16037**	**0.584**	**0.001**	**0.868**	**0.004**
629	**16617**	**16722**	**0.031**	**0.096**	**0.152**	**0.005**
774	**19505**	**19548**	**0.196**	**0.037**	**0.127**	**0.004**
787	20014	20574	0.000	*	0.278	*
788	**20575**	**20920**	**0.003**	**0.000**	**0.016**	**0.001**
841	**21669**	**21817**	**0.463**	**0.001**	**1.000**	**0.004**
856	**22280**	**22638**	**0.001**	**0.400**	**0.015**	**0.001**
867	**22853**	**23072**	**0.123**	**0.000**	**0.576**	**0.001**
1036	26025	26141	0.252	*	0.005	*
1074	**27173**	**27227**	**0.091**	**0.005**	**0.999**	**0.004**
1094	**27843**	**27926**	**0.017**	**0.911**	**0.006**	**0.001**
1119	**28736**	**28899**	**0.012**	**0.003**	**0.503**	**0.001**
1192	**30017**	**30089**	**0.062**	**0.005**	**0.875**	**0.003**
1261	**31494**	**31604**	**0.064**	**0.002**	**0.903**	**0.002**
1482	35573	35654	0.700	0.314	0.004	0.020
1507	**36479**	**36538**	**0.455**	**0.058**	**0.026**	**0.005**
1527	**36833**	**36922**	**0.081**	**0.199**	**0.068**	**0.005**
1536	**37283**	**37382**	**0.012**	**0.665**	**0.064**	**0.002**
1633	39257	39420	0.066	0.007	0.572	0.020
1650	**39812**	**39884**	**0.048**	**0.493**	**0.039**	**0.004**
1654	**40001**	**40104**	**0.099**	**0.018**	**0.022**	**0.004**
1669	**40831**	**41040**	**0.008**	**0.001**	**0.293**	**0.001**
1673	41048	41108	0.200	0.019	0.108	0.007
1694	**41792**	**41997**	**0.004**	**0.027**	**0.195**	**0.002**
1732	42820	42936	0.005	0.284	0.371	0.009
1745	43054	43174	0.005	0.953	0.254	0.007
1775	**43927**	**44251**	**0.001**	**0.000**	**0.128**	**0.001**
1802	**44917**	**45145**	**0.111**	**0.183**	**0.000**	**0.001**
1848	45778	46050	0.096	0.011	0.803	0.007
1854	**46128**	**46263**	**0.023**	**0.754**	**0.017**	**0.004**
1855	**46264**	**46392**	**0.008**	**0.019**	**0.833**	**0.005**
1857	**46450**	**46512**	**0.198**	**0.015**	**0.364**	**0.005**
1860	**46563**	**46844**	**0.425**	**0.045**	**0.000**	**0.001**

Figures in bold type – significance is 0.005 or less.
Asterisks – missing values.
Obs. – observation.

Examples of the significant break points from Table 5.1 are included in Figures 5.1 to 5.4. The vertical line in these figures shows where the break was identified. The top three graphs in each of these figures show a slightly smoothed version of the three variables as RP increases. The bottom three show the cusum chart for each of the corresponding variables, in which the dotted lines show approximate significance lines for the cusum chart.

The skid resistance graph has tick marks on the lower axis corresponding to the skid site codes, as shown in Table 5.2.

Table 5.2 Investigatory level codes for skid resistance.

Code	Investigatory Level	Tick mark
1	0.55	Long – points up
2	0.50	Medium – points up
3	0.45	Short – points up
4	0.40	None
5	0.35	Short – points down

Some of these graphs show a well-defined break in more than one of the series, others show a less well-defined break. In some others the effect is more of a trend than a break with no sudden change in level.

5.2 Identifying Multiple Change Points on SH1, North & South Islands

The breaks identified by the segmentation program were added to the list of breaks and the data was run through the program again. A few more breaks were identified, most of which occurred in segments that had already been identified as having breaks. However, the correction factors also dropped slightly, and so a few segments that were not classified as significant before, now became significant.

Table 5.3 shows the additional segments found on SH1S. Only segments showing significance at 0.005 or better in the *Summary* column are included. Note that segment numbers have changed from those reported in the previous chart because of the additional segments. Further sweeps produced no additional breaks.

Table 5.3 Significant breaks identified for SH1S, South Island, after second pass.

Segment Number	First Obs.	Last Obs.	Significance Level			
			Texture	Roughness	Skid	Summary
562	14339	14486	0.000	0.533	0.001	0.001
563	14487	14636	0.002	0.788	0.029	0.001
1696	41048	41108	0.177	0.015	0.095	0.005
1769	43054	43174	0.004	0.943	0.234	0.005
1800	44102	44251	0.005	0.293	0.001	0.001
1874	45778	46050	0.081	0.008	0.783	0.005

Figure 5.1 Observations 14339-14636, showing well defined break points in skid resistance and possibly in texture.

Top row: Smoothed variables: Bottom row: Cusum charts for variables

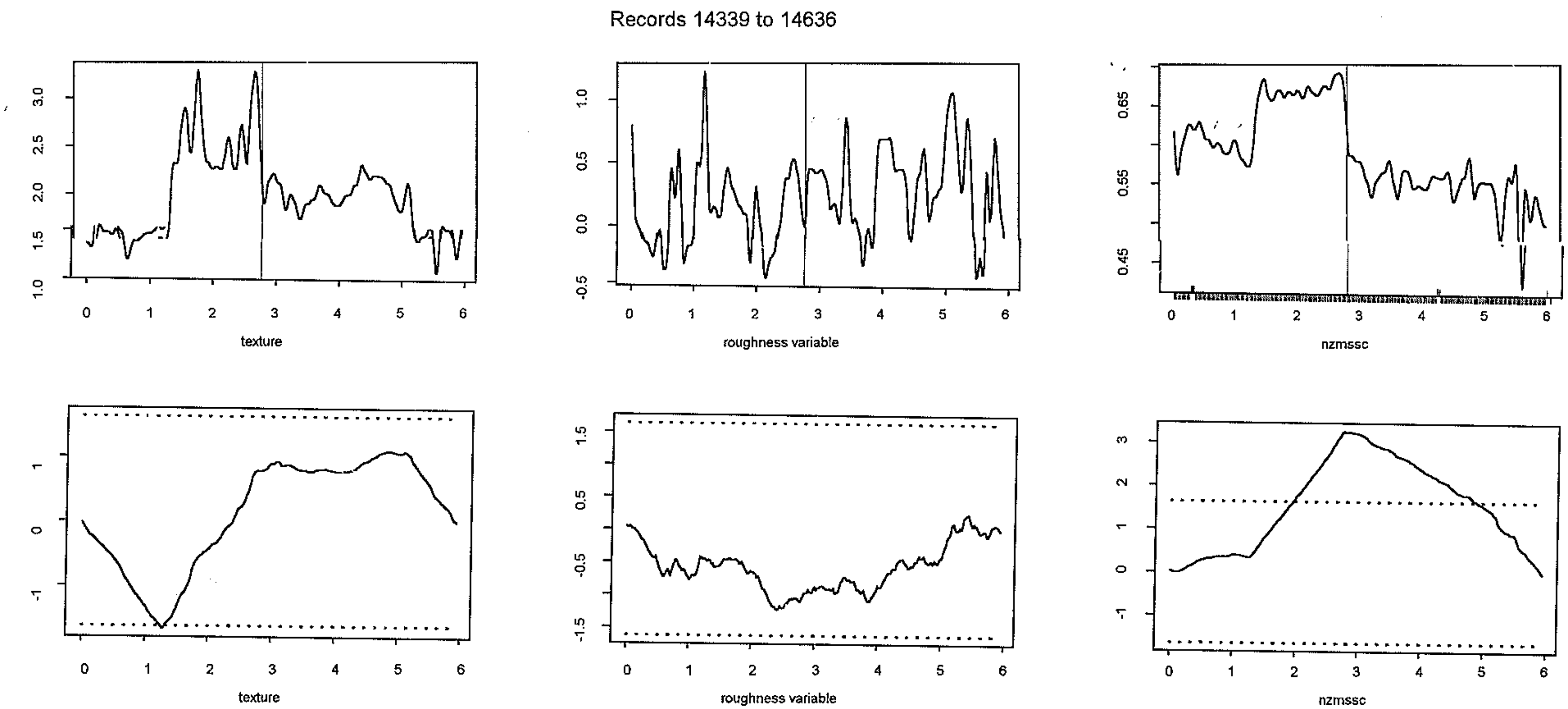


Figure 5.2 Observations 27843-27926, showing well defined breaks in texture and skid resistance.

Top row: Smoothed variables: Bottom row: Cusum charts for variables

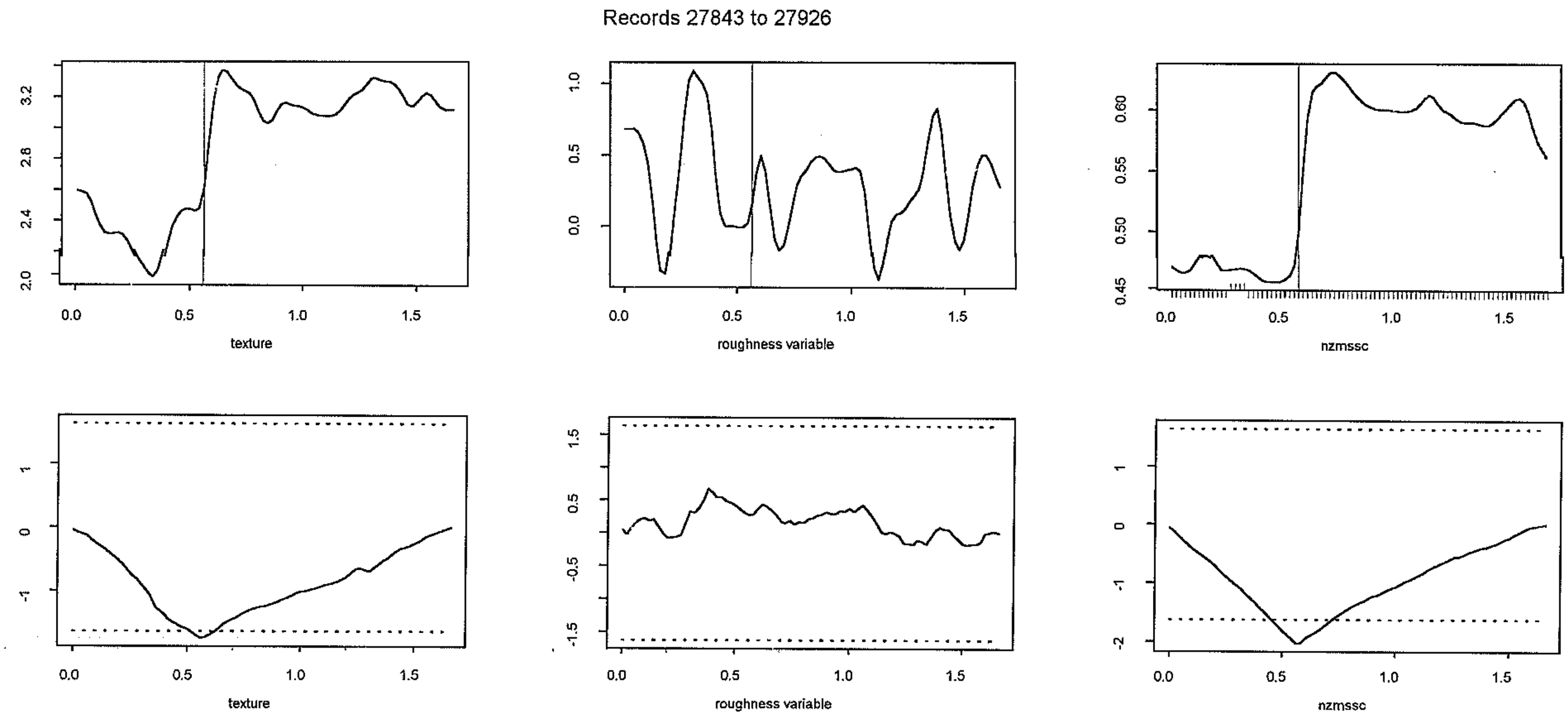


Figure 5.3 Observations 30017-30089, showing well defined breaks in roughness, more of a trend in texture.

Top row: Smoothed variables: Bottom row: Cusum charts for variables

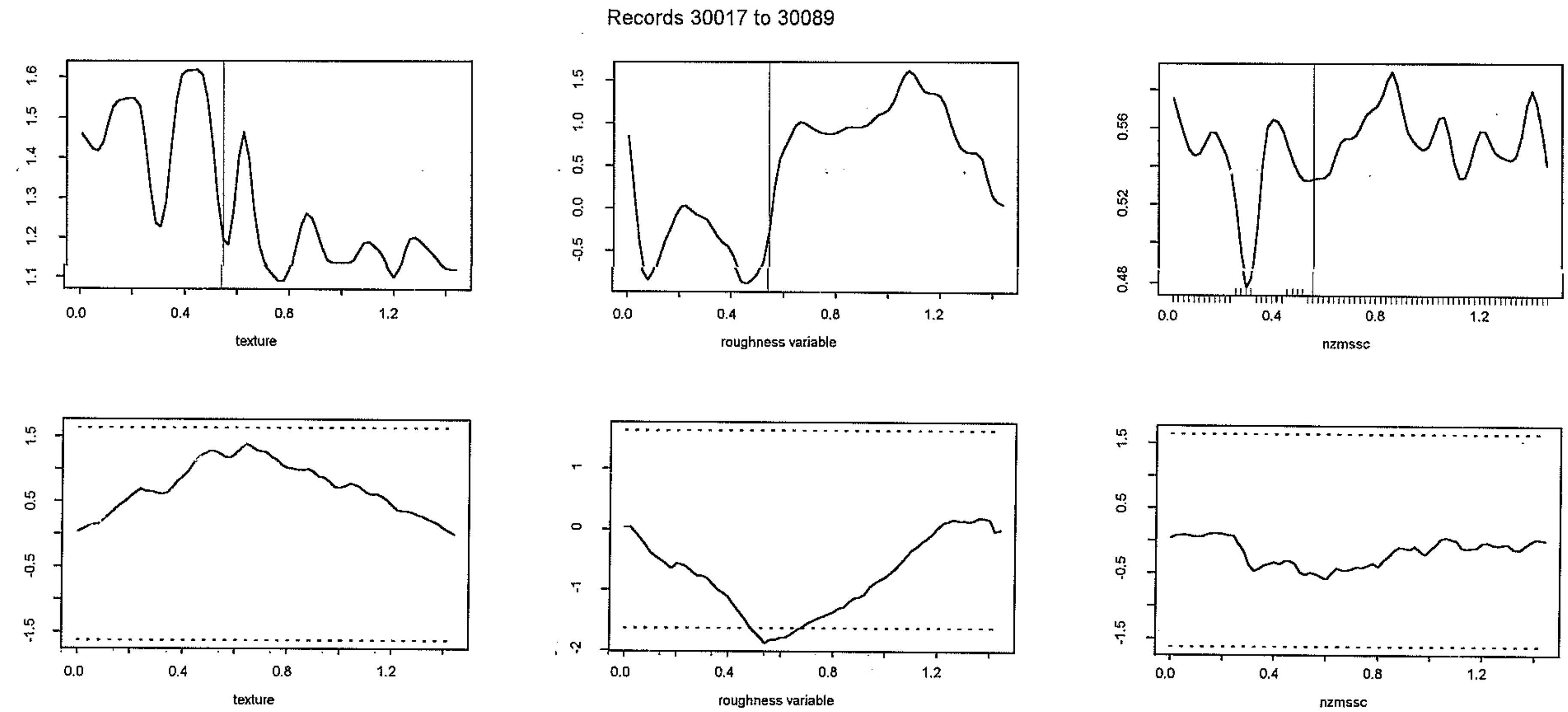
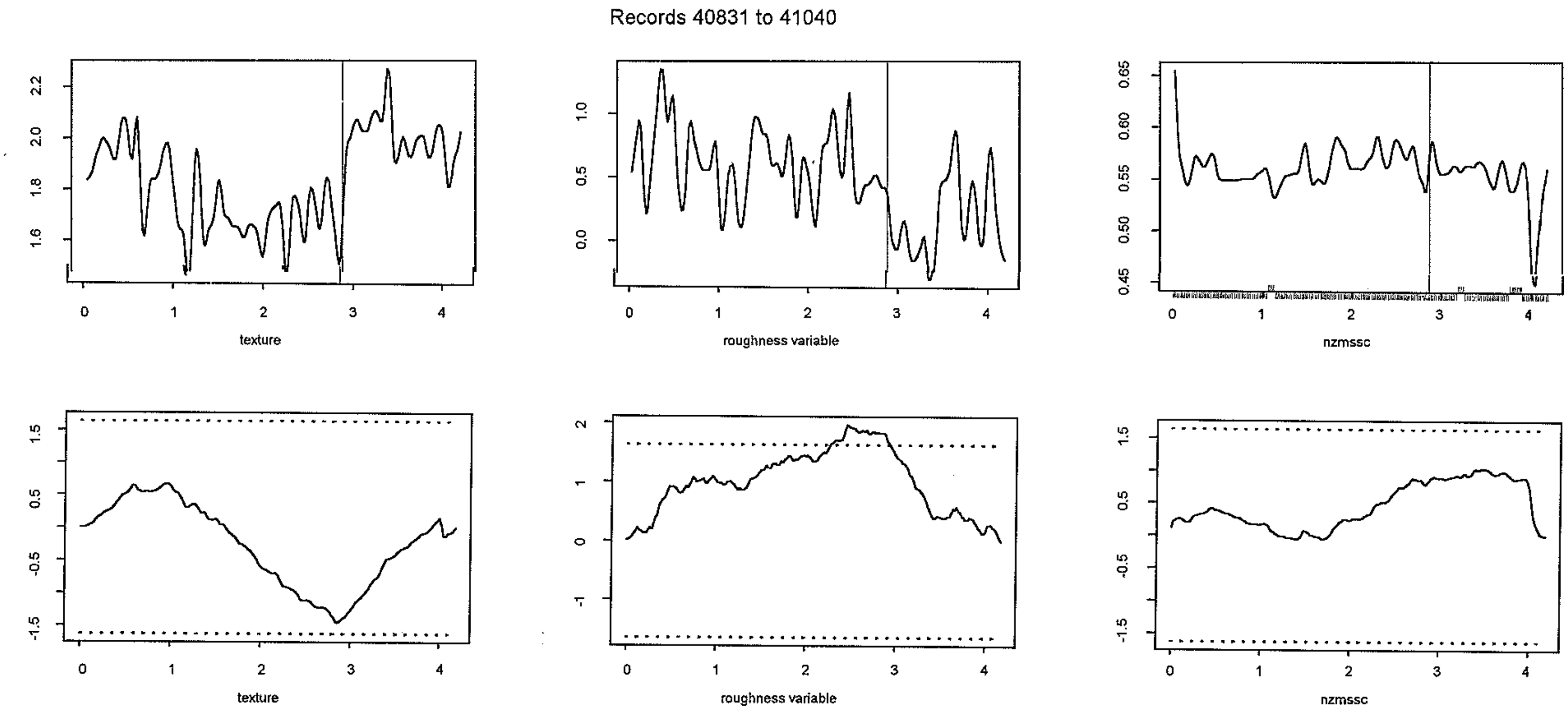


Figure 5.4 Observations 40831-41040, showing less defined breaks in texture and roughness.

Top row: Smoothed variables; Bottom row: Cusum charts for variables



On SH1N three iterations were made. As in Section 5.1, a break was considered to occur if it was significant at the 0.005 level. Twenty-five additional break points were found after the first pass, and an additional four breaks in each of the second and third passes

Finally the data was analysed with the multiple change point method described in Appendix A1.4. This generated six new change points on SH1S, and ten on SH1N. Segments with at least one variable showing significance at the 0.01 level are reported in Table 5.4. Because the method is graphical, identification of breaks has to be carried out manually. In the future, automation based on this research would be useful.

Table 5.4 Significant breaks identified for SH1S, South Island, using multiple change point method.

Segment Number	First Obs.	Last Obs.	Significance Level		
			Texture	Roughness	Skid
340	6908	7118	0.030	0.002	1.000
386	8142	8321	0.008	0.050	0.040
550	13629	13835	1.000	0.006	0.400
1513	35872	36018	0.006	0.200	0.090
1655	39257	39452	0.080	0.005	0.400
1833	44993	45172	0.400	1.000	0.005

5.3 Significance of Breaks Identified from Pavement Construction Data

Step 6 of the segment methodology (Section 3.2 of this report) is to use the change point methodology to check if any of the variables, used for the initial segmentation, produce breaks which do not correspond to statistically significant changes in the three variables derived from the pavement condition data.

Where two successive segments have 20 or more observation intervals, and do not contain missing condition data, a test was carried out on the three variables to find statistical significance of the difference between the segments. The statistical test was based on the normal scores transform of the data so, if outliers occur, the data should be robust. Table 5.5 contains the auto-correlation correction factors used for the three variables. They represent the averages of the North and South Island figures (Tables 4.16 and 4.17).

Table 5.5 Auto-correlation correction factors for combined auto-regressive (AR1) /moving average (MR1) model.

Texture	Roughness	Skid Resistance
2.2059	1.6135	2.0829

Table 5.6 contains the results of the analysis, at various levels of significance. Note that a test significant at the 0.1% level is also significant at the 1% and 5% levels, and a test significant at the 1% level is also significant at the 5% level.

Table 5.6 Test of significance of pavement construction breaks.

State Highway	South Island	North Island
No. of tests	707	765
Not significant at 50% level	23	43
Significant at 5% level	546	555
Significant at 1% level	441	445
Significant at 0.1% level	322	332

Most tests showed significance at the 5% level or better. Tests not showing significance at the 50% level probably represent incorrectly identified breaks, so they are worth looking at more closely, as they probably represent about half the total number of incorrectly identified breaks.

Table 5.7 contains the South Island data for the 23 pairs that were identified as not significant at the 50% level. In no case was a difference caused by a change in lane numbers so this variable is not shown. Cells that contribute to the ability to identify a break have been shaded.

In two cases a short change has been observed in construction data and then the road has reverted to the original. In practice this will not be a real gap but a measurement difference. In a few cases a big change in surface age has been recorded, and it seems likely that there have been data entry errors.

This completes Steps 4 and 5 of segmentation methodology outlined in Section 3.2 of this report.

5.4 Association of Breaks with Pavement Construction Data

The new breaks found from the pavement condition data could be associated with changes in construction data that had not been entered in the inventory database. Another possibility is that they may have been entered with an error in the recording of the surface or pavement construction record, or the placement of the pavement condition data. The new breaks are subject to random error so the approach used was to look at the construction data for five observation intervals before and after the break. In almost every case no change had been recorded, with one exception. In four cases the pavement data switched between known and unknown. This has not been used so far in this project as a reason for a break, and possibly it should be.

5.5 Comparison of Breaks with Current Treatment Lengths

Figures 5.6-5.7 show the locations of the breaks found using the present analysis compared to the current treatment lengths.

Table 5.7 South Island break points coinciding with insignificant pavement construction breaks, at 50% significance level.

First	Last	Surface date	Pavement date	Cway width	Urban / rural	ADT estimate	Surface width	Surface material	Surface type	Number of layers
2671	2692	07/11/90	06/01/89	9.5	2	2800	8	4	4	2
2781	2819	01/11/89	16/02/89	9.9	1	2300	6.2	4	4	2
3913	3977	25/01/95	02/06/94	9.2	1	1900	9.3	9	4	3
3984	4003	25/11/88	25/12/81	9.2	1	1900	6.2	9	4	2
3984	4003	25/11/88	25/12/81	9.2	1	1900	6.2	9	4	2
4004	4029	25/01/95	25/12/81	9.2	1	1900	8.9	9	4	2
4370	4407	01/12/90	25/12/56	7.5	1	1900	7.5	9	4	1
4408	4428	01/12/90	25/12/74	7.5	1	1900	7.5	9	4	2
8464	8493	16/12/97		8.4	1	2000	9.2	9	4	
8494	8513	01/12/87		8.4	1	2000	8.2	11	4	
8947	8973	02/02/93		8.9	1	1700	9.2	4	4	
8980	9009	27/01/93		8.9	1	1700	14	4	4	
15990	16037	03/03/97		10.3	1	8500	9.8	9	4	
16039	16061	03/03/97		6.7	1	4400	9.4	9	4	
16076	16105	31/01/96		10.3	1	4400	9.5	14	4	
16106	16250	31/01/96		10.3	1	8720	9.5	14	4	
16296	16374	01/02/80		10.8	1	10800	10.3	9	4	
16375	16424	25/11/93		10.8	1	10800	10.8	9	4	
16723	16758	16/02/94		11.5	1	13000	13	4	4	
16766	16789	13/02/95	10/03/94	11.5	1	13000	12.5	4	4	1
17220	17248	20/01/95	20/01/95	13.5	2	14500	14.2	5	1	
17273	17301	04/12/96		14.7	2	16000		5	2	
23220	23279	23/11/94		10.1	1	3500	10	9	4	
23337	23356	24/11/94		10	2	4000	10	9	4	
25165	25192	01/03/88		10.5	1	5000	10.5	9	4	
25193	25215	19/03/98		10.5	1	5000	10.5	9	4	
25784	25825	03/03/97	01/03/95	10	1	4420	10	4	4	2
25846	25895	11/03/97	01/04/95	10	1	4420	10	4	4	2
25896	25917	13/02/97		10	1	4420	10	9	4	
25918	25977	01/12/88		10	1	4420	10	9	4	
25918	25977	01/12/88		10	1	4420	10	9	4	
25991	26010	01/12/88		10	1	4420	10	9	4	
31309	31371	01/02/99	25/12/84	8.8	1	3000	10	9	4	3
31383	31409	01/02/99	25/12/85	8.8	1	3000	10	9	4	8
32822	32849	11/04/95	25/12/80	9	1	4000	9	9	4	2
32874	32898	19/12/97	01/04/84	9	1	4000	9	14	4	4
33505	33555	19/11/97	01/04/84	12.6	1	4300	11.5	9	4	5
33574	33630	07/12/98	01/04/84	12.6	1	4300	14	4	4	5
37127	37147	09/10/92	25/12/82	8.5	1	5142	9	9	4	1
37148	37167	08/10/91	25/12/82	8.5	1	5142	8.5	9	4	1
37322	37382	25/02/93	25/12/86	8.5	1	5142	10	9	4	7
37383	37475	25/02/93	25/12/83	8.5	1	5142	10	9	4	1
44252	44309	03/12/96		8.5	1	3915	8.5	9	4	
44310	44351	30/11/98		8.5	1	3915	8.5	9	4	
46179	46263	03/11/93		8.6	1	2198	8.8	9	4	
46264	46296	07/12/92		8.6	1	2198	8.8	9	4	

Blank cells – unknown data

The vertical bars above the dotted line represent break points obtained using the segmentation methodology. Those resulting from an analysis of the condition data are identified as having half the height as those resulting from the initial segmentation process. Break points used in current forward works programmes (FWP) are shown by the vertical bars below the line. A thick horizontal line shows where data is unavailable.

Agreement about breaks is not universal. Occasionally the condition data analysis has found a break corresponding to a treatment length that was not identified by the construction data. Both series, the original FWP treatment lengths and the generated treatment lengths, have a large number of short segments, many of which will be due to variations in recording of locations (remembering that the measurement error can be as large as 50m), but some of them will be caused by short events on the highways such as bridges, corners, intersections, large repairs, etc.

Because of the large number of very short sequences, an attempt was made to join short segments together to reduce the occurrence of multiple break points, and creating segments that are shorter than the measurement error by a single break point. In doing this, a fairer comparison between the treatment lengths and the segments generated in this report will be obtained, since break points within 100m of each other are likely to be caused by the same physical feature or change in pavement condition characteristics.

Using the coarse and fine fit parameters listed in Table 5.8, an attempt has been made to quantify the level of agreement/disagreement between the standard treatment lengths and those generated using the segmentation methodology described in Section 3.2. The results are in Tables 5.9 and 5.10. The coarse fit assumes that the measurement error is ±50m, which is a break point displacement of up to 100m. The fine fit assumes that in most cases the measurement error is closer. to ±20m, generating up to 40m difference in break point placement.

Table 5.8 The two levels of fitting investigated, using treatment length parameters.

Level of Fit	Coarse fit (m)	Fine fit (m)
Minimum segment length	200	100
Maximum break length	100	40
Maximum difference for a match	100	40

All segments that are at least as long as the minimum segment length were retained. Where the two adjacent segments were separated by no more than the maximum break length, a single break was assumed to exist at the centre of the gap between the segments. Because of rounding, there may be an error of up to 10m in such a location. If the gap was more than the maximum break length, the breaks in the gap were ignored for the purpose of the comparison.

Figure 5.6 Comparison of break points on SH1N, North Island.

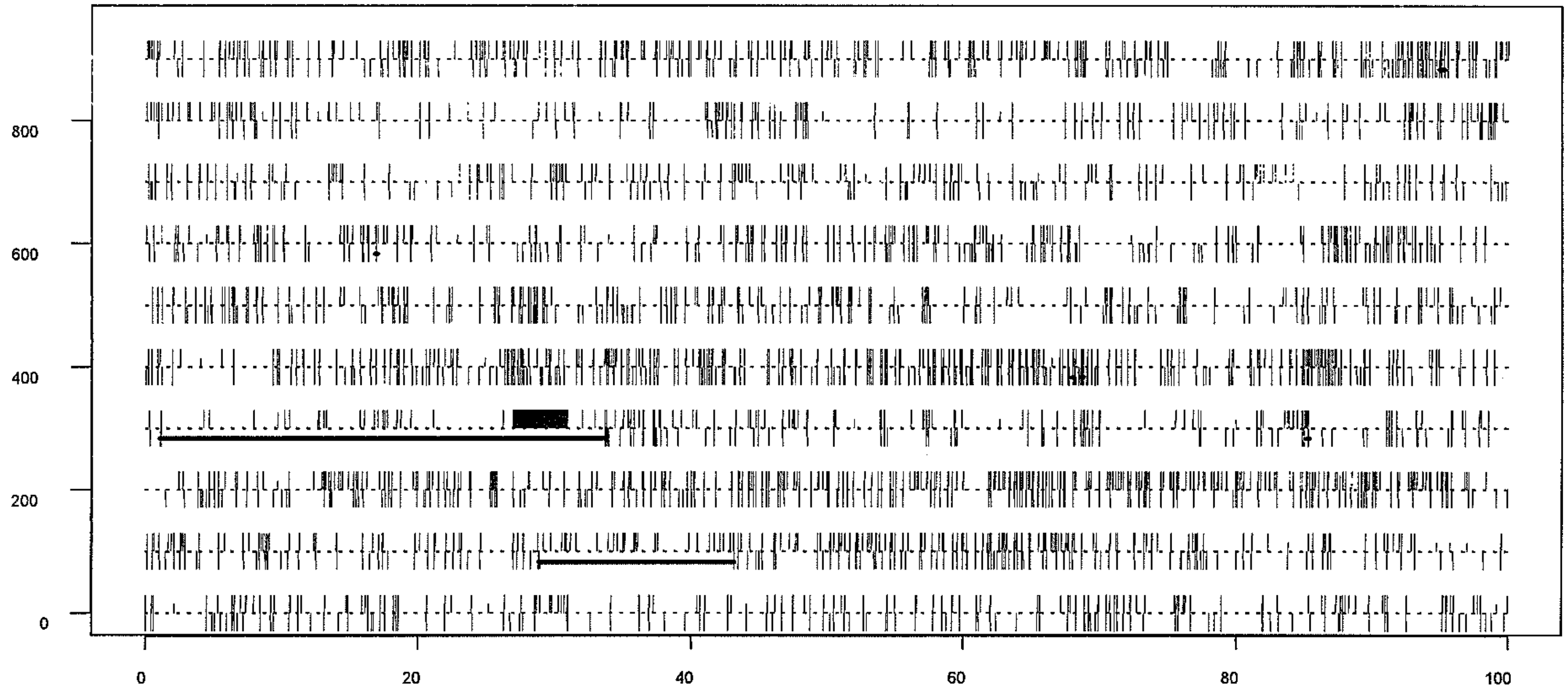
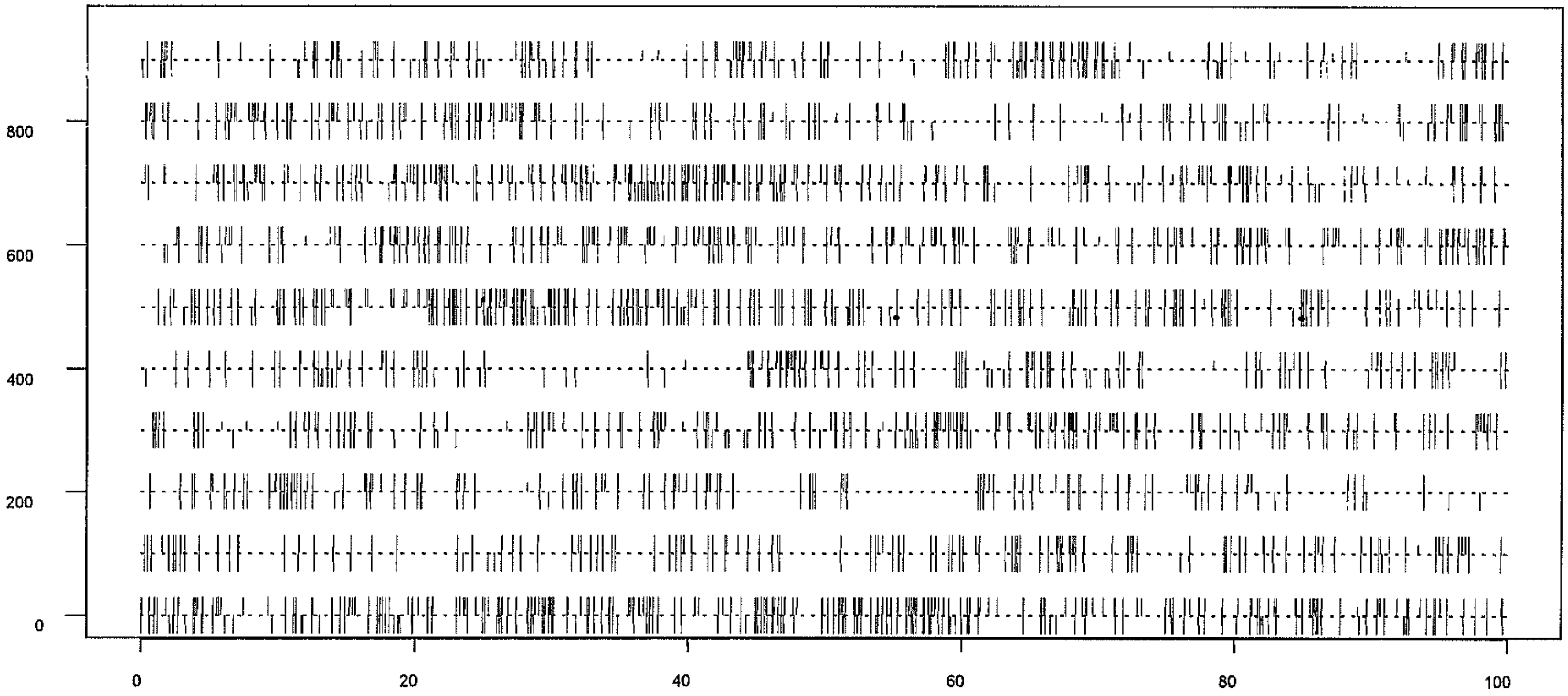


Figure 5.7 Comparison of break points on SH1S, South Island.



This process was applied to both the treatment lengths in the FWP and the segments identified using the segmentation methodology. A break in one of the series was considered to match a break in the next if the distance between them was no more than the maximum difference for a match. The two methods have been compared in the increasing (I) and decreasing (D) directions for both the North and South Island databases respectively.

Table 5.9 Match between FWP and segmentation breaks, North Island.

Match Type	Coarse fit				Fine fit			
	Count		Percent		Count		Percent	
Direction	I	D	I	D	I	D	I	D
Total TL breaks	1536	1536	100%	100%	1536	1536	100%	100%
Initial match	1362	1360	89%	89%	1109	1115	72%	73%
Final match	1380	1375	90%	90%	1116	1126	73%	73%
Breaks not found	156	161	10%	10%	420	410	27%	27%
New breaks (initial)	515	528	34%	34%	1016	1031	66%	67%
New breaks (cusum)	19	25	1%	2%	27	29	2%	2%
Isolated breaks	919	919	100%	100%	1269	1269	100%	100%
Initial match	604	601	66%	65%	711	714	56%	56%
Final match	615	615	67%	67%	715	725	56%	57%
Final - not isolated	786	787	86%	86%	898	911	71%	72%
Breaks not found	133	132	14%	14%	371	358	29%	28%
New breaks (initial)	263	266	29%	29%	577	581	45%	46%
New breaks (cusum)	21	25	2%	3%	28	29	2%	2%

Table 5.10 Match between FWP and segmentation breaks, South Island.

Match Type	Coarse fit				Fine fit			
	Count		Percent (%)		Count		Percent (%)	
Direction	I	D	I	D	I	D	I	D
Total TL breaks	1228	1228	100	100	1228	1228	100	100
Initial match	1140	1140	93	93	1033	1033	84	84
Final match	1146	1146	93	93	1036	1039	84	85
Breaks not found	82	82	7	7	192	189	16	15
New breaks (initial)	169	169	14	14	506	506	41	41
New breaks (cusum)	30	23	2	2	33	23	3	2
Isolated breaks	887	887	100%	100%	1071	1071	100%	100%
Initial match	670	668	76%	75%	772	773	72%	72%
Final match	675	674	76%	76%	775	779	72%	73%
Final - not isolated	814	815	92%	92%	902	905	84%	85%
Breaks not found	73	72	8%	8%	169	166	16%	15%
New breaks (initial)	141	140	16%	16%	381	380	36%	35%
New breaks (cusum)	31	23	3%	3%	33	23	3%	2%

TL Treatment Lengths
I, D Increasing; Decreasing

In each table, the top six rows of the body of the table are for the data where purging of break points has not been applied, and all break points are considered. In the bottom seven rows the purging is carried out. The percentages are with respect to the total number of treatment length breaks in the top six rows and the purged numbers in the bottom seven rows.

The line *Initial match* shows the number of treatment length breaks that were matched by the breaks found from the construction data. In the line *Final match*, breaks found from the cusum process are also included.

The conclusions to be drawn from the comparisons are as follows:

- The data for the increasing and decreasing directions are very similar.
- The comparison with FWP treatment length data for the North and South Islands is roughly the same, bearing in mind that there is a section in the North Island for which treatment length data was not available.
- With the *coarse* match, around 90% of the purged treatment length breaks are matched by break points from our study (70-80% if our data is also purged). The match rate drops a little with the *fine* match, particularly in the North Island.
- In the South Island, if the break points of this study are added to those of the FWP treatment lengths, the result is an increase in break points of 20% for the *coarse* match and 40% for the *fine* match. The figures are higher in the North Island, possibly due to some of the treatment length data being unavailable.

The contributions from the cusum part of the project are relatively small.

This completes Step 7 of the segmentation methodology outlined in Section 3.2 of this report.

5.6 Comparison of Results for Opposing Lane Direction

Figures 5.8 and 5.9 show the segmentation results for the increasing (above the dotted line) and decreasing directions (below the dotted line). The breaks shown by the short lines are those introduced by the performance data.

Most of the breaks introduced by the construction data match in the two directions. However, only some of the breaks from the pavement condition data match. There is always the possibility of marginal situations where the break is real but is detected in only one direction. However, it does appear that the procedure is over-sensitive and is inserting breaks too often.

From the practitioner's point of view, the number of additional breaks produced by this procedure in any given network management area is sufficiently small that plotting the performance data, or visiting the sites to verify the nature of the road condition either side of the change point, is not onerous.

Figure 5.8 Comparing segment breaks in increasing and decreasing directions on SH1N North Island.

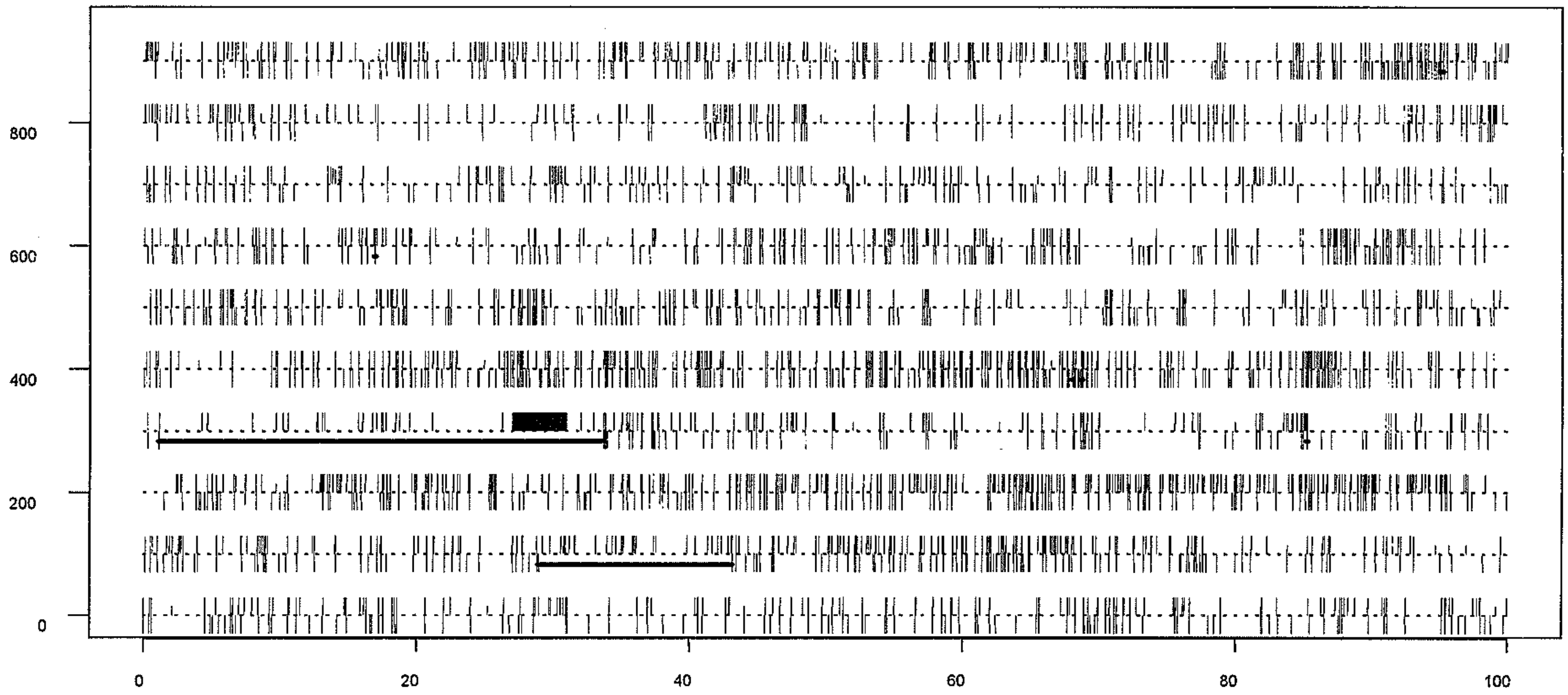
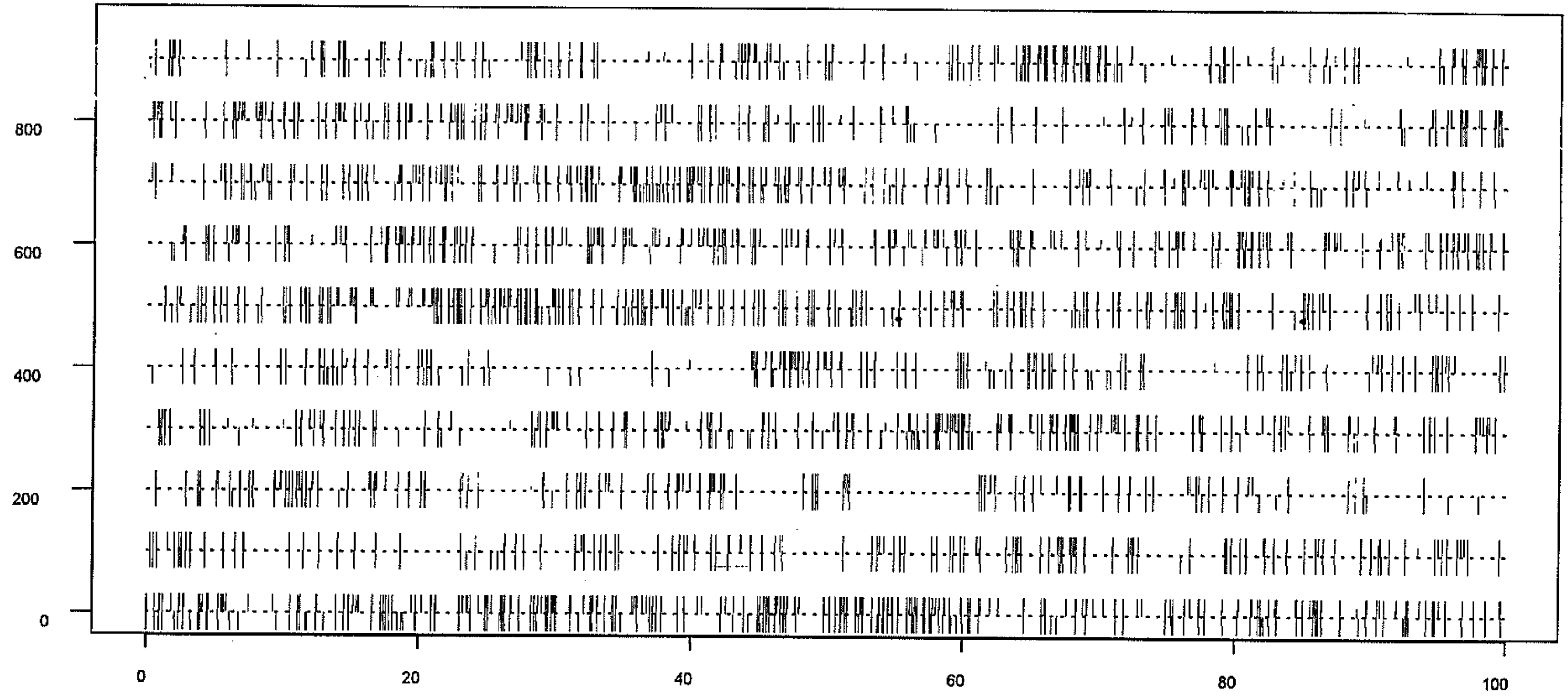


Figure 5.9 Comparing segment breaks in increasing and decreasing directions, on SH1S South Island.



6. Segmenting Northland Region Roads

6.1 Identifying Breaks based on the Variables

The Northland region road network condition data was analysed in the same way as for SH1, detailed in Section 5 of this report, using data in the increasing direction from the Reference Point. The total length of the road analysed was 711km. In the initial segmentation, 1282 breaks were generated. From three iterations of the change-point analysis, 38 breaks were generated in the increasing direction and 29 breaks in the decreasing direction. Figures 6.1 to 6.4 demonstrate the break points in Northland roads for the increasing direction and initial segmentation for the three principal components of texture, roughness and skid resistance. They show that break points are present at all significant changes in the pavement condition data.

6.2 Identifying Breaks based on SNP

SNP values would be expected to indicate changes in pavement strength (and hence changes in construction or treatments). However, FWD surveys are needed to determine SNP, and FWD surveys are costly and have not yet been carried out for all of SH1. As an FWD survey had been carried out in 1991 for the Transit NZ Northland region roads, SNPs for these roads could be calculated, and the usefulness of the procedure was assessed.

In Northland, the average spacing between FWD readings was 123m, if measurements from both sides of the road were used, and the Structural Number (SNP) values ranged between 0.6 and 7.5, with a mean of 4.0. As well, the SNP measurements were plotted with the section breaks to test if the SNP values are uniform over segments (Step 8, Section 3.2 in this report).

Because of the low sampling rate, the SNP values have not been used to identify breaks in Northland region roads, and it is unlikely that they could locate breaks with sufficient precision.

Despite this low sampling rate, SNPs have the potential for providing an additional check on already identified breaks. Figure 6.4 shows that SNPs tend to shift at change points. Therefore we conclude that using only pavement condition data does identify the breaks related to changes in pavement strength, and that SNP is useful for gaining an overall picture of pavement strength.

This completes Step 8 of the segmentation methodology outlined in Section 3.2 of this report.

Figure 6.1 Breaks based on mean texture data, for Transit NZ Northland region roads.

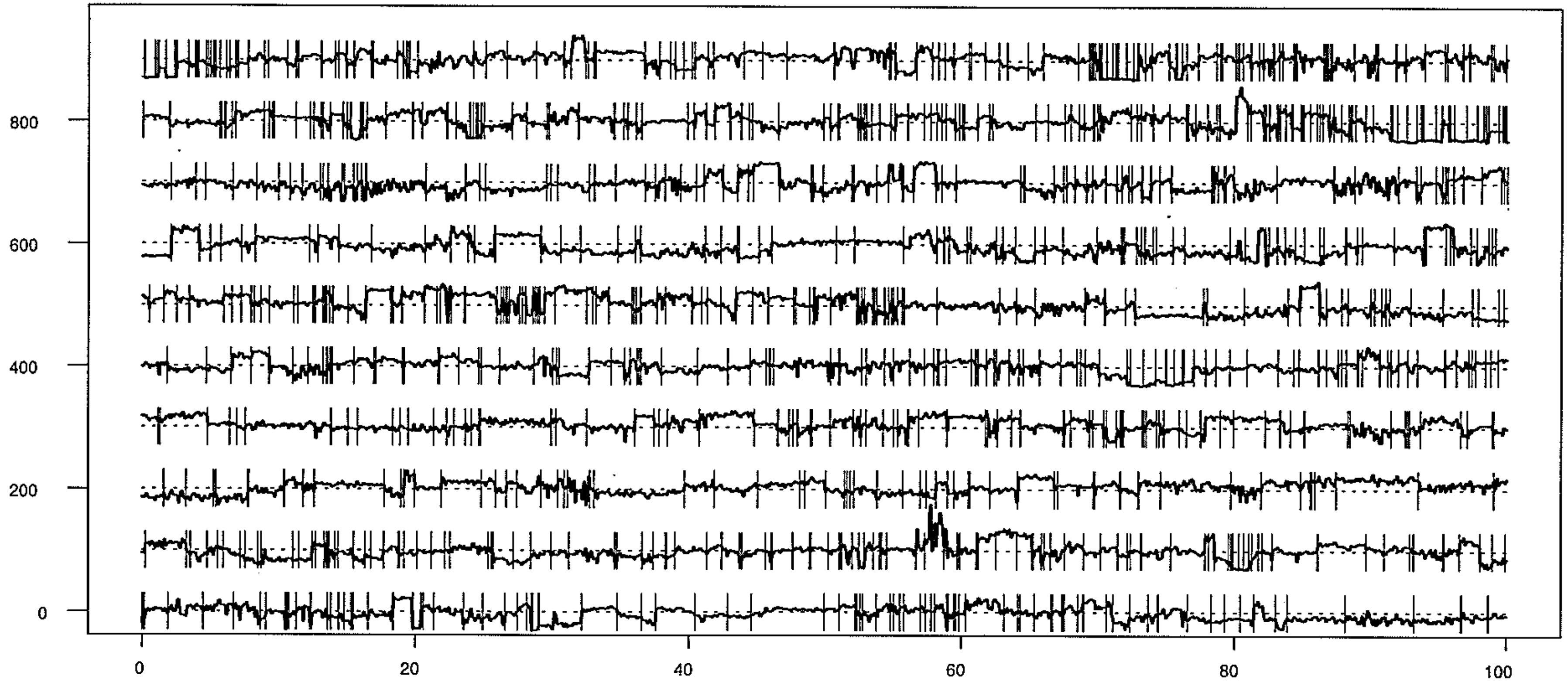


Figure 6.2 Breaks based on combined roughness values, for Transit NZ Northland region roads.

Figure 6.3 Breaks based on skid resistance (NZMSSC) data, for Transit NZ Northland region roads.

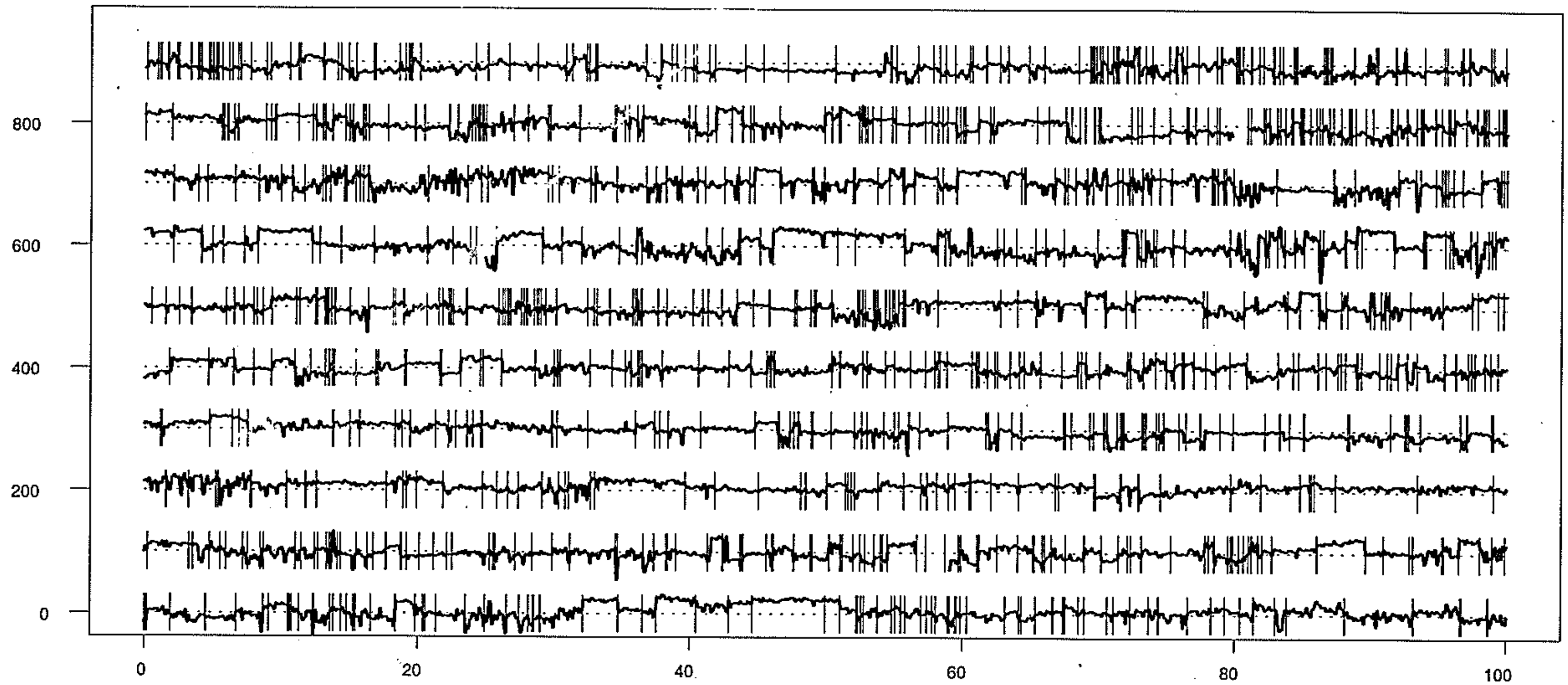
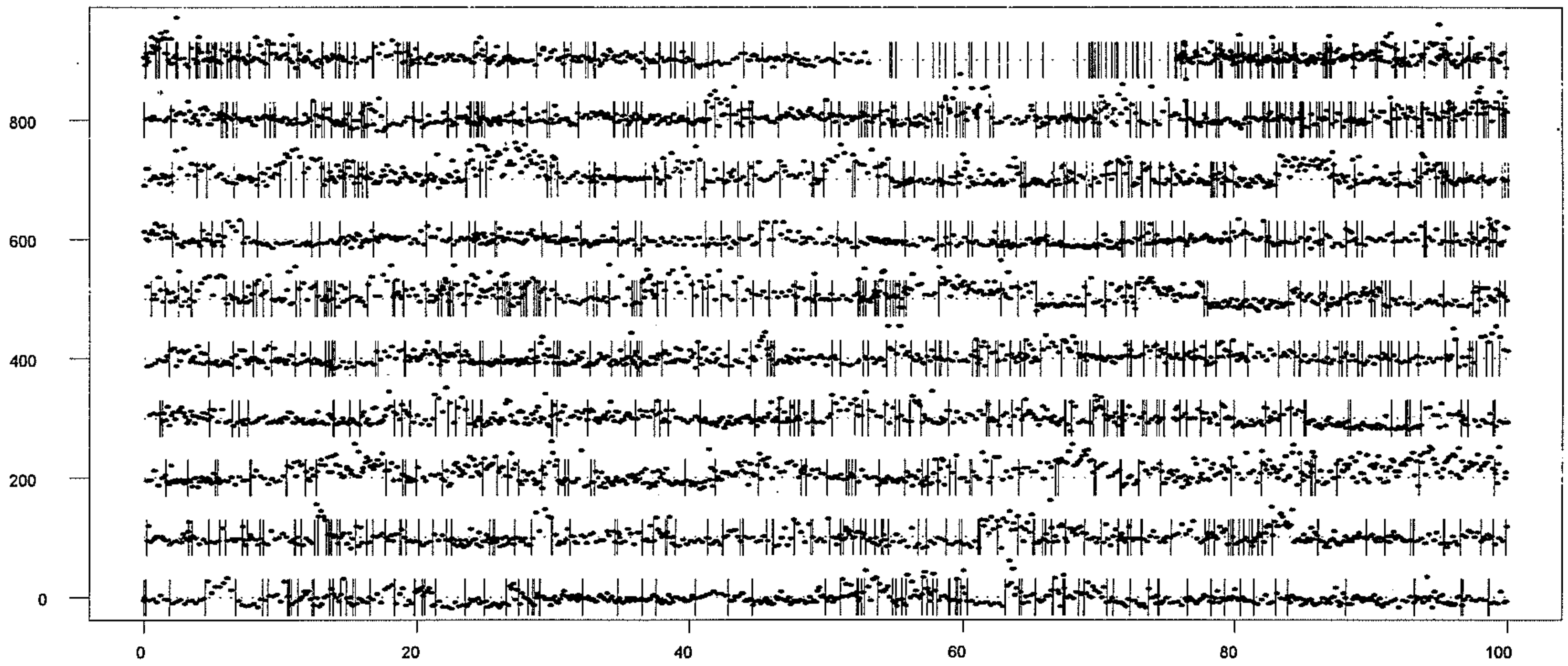


Figure 6.4 Breaks based on structural number values (SNP), for Transit NZ Northland region roads.



7. Discussion & Conclusions

Three main applications result from the development of a robust segmentation methodology as follows:

- Creation or validation of treatment lengths for the on-going management of treatments and maintenance to pavement sections within road networks.
- Identification of homogeneous stretches of roads for calibrating current pavement deterioration models incorporated in NZ-dTIMS.
- Locating homogeneous roads with known pavement characteristics suitable for long-term pavement performance monitoring, in order to generate data for developing new deterioration models.

The cusum change point method that has been investigated can identify breaks caused by construction start and end points. However it is sensible to start with either existing treatment lengths or the initial segmentation methodology, employing pavement layer information, to reduce the computation complexity. Additional breaks caused by variation in the deterioration within treatment lengths, that have not been previously identified by eye, are found through change point tests. These tests are carried out on pavement condition data that has been acquired by multi-function laser-based vehicle measuring systems.

Where the additional break points are clearly defined, they are usually related to omitted construction data. However, some are caused by uneven wear of pavement or surface layers. More often, this shows as a less well-defined break or trend. The less well-defined breaks are worth including as they could cause problems for statistical analyses or predictive modelling if they were ignored.

For the network management application, both the segments generated in this analysis and the FWP treatment length database include a large number of short sections of road, reflecting a high number of short treatment intervals on the highways. Since the match between the segments generated and the FWP treatment lengths is high, the number of additional breaks that can be found using the pavement condition data is relatively small. It would be feasible to check manually to decide which should be included in the FWP treatment lengths. For computer-based pavement management systems, such as NZ-dTIMS, it would be of benefit to include all of them, since each of them represents a significant change in one or more of the predictor variables

Ideally, adjacent treatment lengths should be amalgamated when the pavement condition data measures show little difference. In this case, the analysis needs to concentrate on what changes are of practical significance rather than statistical significance. This has not been attempted in the present study.

The SNP results have not been used as part of the process. However, they are potentially important for the amalgamation process when deciding the adjacent segments that can be amalgamated.

Most of the analysis has been performed on SH1 data. This has provided a long stretch of road with a range of construction and surface characteristics, but it may not be representative of all pavement types because SH1 carries more traffic. Much of it has been designed to carry heavy traffic, and therefore it is reasonable to expect more even pavement performance than other highways where the traffic load is less predictable. However the results of the analysis of the less trafficked Northland database do not give any reason to believe that the components are significantly different on less trafficked highways.

8. Recommendations

The cusum change point methodology developed and presented in this report can be applied to any pavement condition data held for New Zealand roads to identify statistically significant break points in construction. Field inspections would confirm the reasons for these break points. Possible reasons include:

- Surface or pavement condition.
- Missing construction data in inventory database.
- Existing treatment lengths.

A field trial could make use of the break points found in this project, or use the method described to identify break points in a selected region. The benefits of conducting a field trial to investigate the contributors to break points include:

- Calibrating the sensitivity of the cusum method to avoid creating spurious break points.
- Verifying the causes of extreme events in the condition data on the road and thus estimating the influence of the smoothing methods applied to the 20m-aggregated condition data. Many of the causes of extreme events are acknowledged in the condition data collected, for example, bridge abutments and railway crossings.
- Assessing the accuracy of measurements made in the field. At present the accuracy within RAMM for all maintenance, and pavement and surfacing activities is assumed to be ±50m (G. Hart, pers. comm. July 2001). A field trial would quantify the level of accuracy of the data stored in RAMM with respect to the position of road features, road maintenance, and in construction changes in pavement and surface. Any segments smaller than the limits of measurement accuracy should be excluded from the segmentation process.

Future work, which would benefit from the field trial results, includes:

- Fine tuning of the change-point algorithm. This is an iterative process, and the ideal coefficients cannot be calculated until the majority of the change points have been identified.
- Further investigation of the auto-correlation structure of the data.
- Investigation of criteria for amalgamating adjacent treatment lengths.
- The current RAMM treatment length generator should be reviewed, based on the initial segmentation methodology presented in this report.

9. References

Anon. 1999. *S-Plus 2000 User's Guide.* Data Analysis Products Division, MathSoft, Seattle.

Bai, J. 1997. Estimating multiple breaks one at a time. *Econometric Theory 13:* 315-351.

Barksdale, R.D. 1971. Compressive stress pulse times in flexible pavements for use in dynamic testing. *Highway Research Record No. 345:* 32-44.

Beran, J. 1994. *Statistics for Long-memory Processes.* Chapman & Hall, New York.

Chen, J., Gupta, A.K. 2000. *Parametric Statistical Change Point Analysis.* Birkhäuser, Boston.

Divinsky, M., Nesichi, S., Livneh, M. 1997. Development of a road roughness profile delineation procedure. *Journal of Testing and Evaluation 25:* 445-450.

Fong, S. 1999. Effects of texture on profiler measurements. *Central Laboratories Report No. 99-529359.10.* Opus Central Laboratories, Gracefield, Lower Hutt, New Zealand.

Fuller, W.A. 1996. *Introduction to Statistical Time Series.* Wiley, New York.

Hájek, J. 1969. *Nonparametric Statistics.* Holden Day, San Francisco.

HTC Infrastructure Management Ltd. 2000. *Implementation of dTIMS in New Zealand – establishing pavement strength.* Unpublished HTC report (ALGENZ 30/06/2000), Kumeu, Auckland. 22pp.

Salt, G., Stevens, D. 2001. Pavement performance prediction. Determination and calibration of structural capacity (SNP). *Proceedings of 20th ARRB Conference,* Melbourne, Australia.

Stuart, A., Ord, J.K. 1991. *Kendall's Advanced Theory of Statistics,* Volume 2. Oxford University Press, New York.

10. Glossary

10.1 Statistical Terms

ε	error term, independent random variable
k	a number, usual representing the sequence number of an observation in a series of observations
N, n	the number of elements in a series
$O(n)$	a number that is of the same order of magnitude as n
$O(n^2)$	a number that is of the same order of magnitude as n^2
X	a random variable, representing one of the HSD measurements
X_k	the k^{th} element in a series of numbers between 1 and N
$X_1, X_2, \ldots X_k, \ldots X_n$	a series of values for the random variable X
μ	expected value of the first k elements in a series containing n elements
ν	expected value of the last n-k elements in a series
Auto-correlation	measures the strength of the linear relationship between successive members of a series of random variables
Auto-regressive	a particular model for an auto-correlation function
Change point	the point where a significant change is found in a random series
Correlation	measures the strength of the linear relationship between two variables
Cusum	method used to detect small changes in the mean of a process or series using a cumulative sum of the series
First order auto-regressive process	an auto-correlation function made up of correlations between adjacent values within one series (AR1)
First order moving average process	a method of smoothing a time-series by replacing each observation by a weighted mean of that observation and its near neighbours (MA1)
Null hypothesis	a hypothesis about the data that usually corresponds to *no effect,* for example that there is no difference between μ and ν
Outlier	unusual result that cannot be easily accounted for
S-Plus	data analysis software
Step	a change between the expected value before and after point k when $\mu \neq \nu$
Test statistic	a number that can be calculated from the data and used to test a statistical hypothesis; usually large values of the test statistic indicate that the null hypothesis is *not* true
Time-series analysis	technique used to study data collected over a period of time for prediction purposes

10.2 Abbreviations & Acronyms

AADT	Annual Aveıage Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials (after 1974)
ADT	Average Daily Traffic
AR (AR1)	(First order) Auto-Regressive process
COR	Correlation
Cusum	Cumulative sum (see Glossary 10.1 for definition of cusum method)
D	Decreasing
FWD	Falling Weight Deflectometer
HSD	High Speed Data
I	Increasing
IRI	International Roughness Index
LWP	Left Wheel Path
MA (MA1)	(First order) Moving Average process
NZMSSC	NZ Mean Summer SCRIM Coefficient
NZ-dTIMS	NZ HTC Management System modelling program
RAMM	Road Assessment and Maintenance Management system
RWP	Right Wheel Path
SCRIM	Sideways-force Coefficient Routine Investigation Machine
SNP	Structural Number, calculated using Salt's adjustment of AASHTO formula (Salt & Stevens 2001).
SQRT	Square root

Appendix A1. Statistical Techniques

This appendix describes the statistical methods used in this project and applied to state highway data in Section 4.

A1.1 Auto- and Cross-Correlation

These correlation measures indicate how closely two variables are related. For example rut depth, variance of rut depth, and IRI tend to be high in the same places and low in the same places. So they are positively correlated.

If one variable is plotted against another and all the points lie exactly on a straight line with positive slope (i.e. going upwards), the correlation would be 1. If the line has negative slope (i.e. going downwards), the correlation would be –1.

If the point did not quite lie on the line (and the line had positive slope), then the correlation would be less than 1, but it would be close to 1.

If there is no relationship between the points, then the correlation would be zero. For example, if rut depth is plotted against longitude, there would probably be very little relationship. So the correlation would be close to zero.

When driving along a road and measuring texture, for example, each 20m segment would have similar readings to segments that are adjacent to it. That is, adjacent segments will have correlated values. The further apart the segments, the less similar they will be and the correlation will drop. The series of values of correlations for increasing separations (or *lags*) of the segments is known as the *auto-correlation* function.

The correlations between two series, such as texture and rut depth, are known as *cross-correlations*.

Correlation functions can also be calculated for variables that are cross-correlated, i.e. by calculating the correlations between two series, where one of the series is displaced by one or more observations. In the present study such cross-correlations with non-zero lag are not relevant.

When carrying out a statistical analysis of data involving either auto-correlation or cross-correlation, it is important that the analysis takes account of these correlations even if they are not of direct interest in themselves.

A1.2 Change Point Problem, Single Change Point, Single Series

Each of the high speed data (HSD) measurements, i.e. texture, roughness, rut depth and skid resistance, can be thought of as a series of numbers represented by a series of random numbers $X_1, X_2, \ldots, X_k, \ldots, X_n$, where the numbers 1 to n represent their order along the highway. X_k is defined as the k-th element in the series.

Suppose that the first k elements have *expected* value μ and the last $n-k$ elements have *expected* value ν.

An explanation of the expected value is as follows:
Suppose that the series $X_1, \ldots, X_n$ could be produced over and over again by the same random process so that a large number of random series were generated. Then the average value of the repeated values of each of $X_1, .., X_k$ would be μ and the average value of each of $X_{k+1}, \ldots, X_n$ would be ν.

Of course, we have only one series. However the values of $X_1, .., X_k$ would tend to hover around μ and the values of $X_{k+1}, \ldots, X_n$ would tend to hover around ν.

If $\mu \neq \nu$, then there is a step in the process. Typically both μ and ν are unknown and the problem is to decide whether $\mu \neq \nu$, i.e. whether there really is a step.

If the point where the step is suspected is known, then deciding whether, in fact, there is a step is a standard statistical problem. If the point is unknown, then the problem is known as the change point problem.

The cusum method, used for detecting changes in industrial processes, can be used to detect one or multiple steps, and the single step case was considered first.

Formula A1 is the absolute difference between the sum of the first k observations and an estimate of the sum of the first k observations based on the sum of all the observations.

$$\left| \sum_{i=1}^{k} X_i - \frac{k}{n} \sum_{i=1}^{n} X_i \right| \qquad \text{(Formula A1)}$$

If k, the location of the step, was known and several statistical assumptions were satisfied, an optimal test could be based on Formula A1. Since the actual point of the postulated step is unknown, Formula A2 is used, which finds the maximum as k varies. This leads to the *test statistic*[4]:

$$\frac{\max_{0<k<n} \left| \sum_{i=1}^{k} X_i - \frac{k}{n} \sum_{i=1}^{n} X_i \right|}{\sqrt{n \operatorname{var}(X_i)}} \qquad \text{(Formula A2)}$$

Formula A2 includes division by the square root of n times the variance (or an estimate) of the X_i (assumed to be independent of i) to make it independent of scale.

---

4 The test statistic is a number that can be calculated from the data and used to test a statistical hypothesis. Usually *large* values of the test statistic indicate that the null hypothesis is *not* true.

In order to use Formula A2 as a statistical test for determining whether or not there is a step in the sequence $X_1,\dots,X_n$, it must first be known what is the statistical distribution of Formula A2 under the *null hypothesis* (which is that there is no step). The particular advantage of Formula A2 is that the distribution is known[5] when the X_i are statistically independent of each other and have the same distribution, and that this distribution is well behaved (e.g. normal).

Once it has been determined that there is a change point, an obvious estimate of its location is just after the value of *k* where the maximum of Formula A1 occurs.

A possible disadvantage of Formula A2 is that it is less sensitive to a step towards the beginning or end of the sequence than it is to a step near the centre. While such an effect is inevitable, it would be preferable to apply a test that gave a better trade-off between sensitivities towards the ends of the sequence and the sensitivities near the centre.

Formula A3 is a possible candidate test statistic, where Δ is around $n/10$:

$$\frac{\max\limits_{\Delta<k<n-\Delta}\left\{\frac{1}{\sqrt{k(n-k)}}\left|\sum_{i=1}^{k}X_i-\frac{k}{n}\sum_{i=1}^{n}X_i\right|\right\}}{\sqrt{\operatorname{var}(X_i)}} \qquad \text{(Formula A3)}$$

However, in the present report the test statistic in Formula A2 is used as the starting point. Divinsky et al. (1997) also use a version of Formula A2.

A1.3 Change Point Problem, Single Change Point, Multiple Series

In this report, the cusum method has been applied to a number of condition measures, each of which can be thought of as a sequence, say $X_1,\dots,X_n$, $Y_1,\dots,Y_n$ and $Z_1,\dots,Z_n$. Each such sequence may be used to detect the change point. The obvious generalisation of Formula A2 for three series is given in Formula A4.

$$\max_{0<k<n}\sqrt{\frac{\left(\sum_{i=1}^{k}X_i-\frac{k}{n}\sum_{i=1}^{n}X_i\right)^2}{n\operatorname{var}(X_i)}+\frac{\left(\sum_{i=1}^{k}Y_i-\frac{k}{n}\sum_{i=1}^{n}Y_i\right)^2}{n\operatorname{var}(Y_i)}+\frac{\left(\sum_{i=1}^{k}Z_i-\frac{k}{n}\sum_{i=1}^{n}Z_i\right)^2}{n\operatorname{var}(Z_i)}} \qquad \text{(Formula A4)}$$

There does not appear to be a standard theory for the distribution of Formula A4 so a computer simulation was used to find the values. For large *n*, the distribution is approximately independent of *n*, and a sample size of 100 was chosen for the simulation. Independent normal random variables were used for X_i, Y_i and Z_i, and 100,000 simulations were carried out. Table A1.1 shows the estimated critical values for Formula A4 when there are three series (or variables).

[5] The distribution is the same as that for the two-sided Kolmogorov-Smirnov goodness-of-fit test. (Stuart & Ord 1991, p.1188, formula 30.81).

Table A1.1 Estimates of critical levels for cusum method with the three independent normal random series.

Significance Level	Critical Value	Significance Level	Critical Value
0.5	1.16	0.02	1.84
0.4	1.24	0.01	1.94
0.3	1.32	0.009	1.96
0.2	1.42	0.008	1.98
0.1	1.56	0.007	2.00
0.09	1.58	0.006	2.02
0.08	1.60	0.005	2.05
0.07	1.63	0.004	2.08
0.06	1.65	0.003	2.12
0.05	1.68	0.002	2.17
0.04	1.73	0.001	2.25
0.03	1.77		

If, for example, a value of 2.10 was found for Formula A4, then the result is statistically significant at the 0.004 (=0.4%) level, which is quite strong evidence for the existence of a step in the process.

As before, the estimate of the location of the step is just after the value of k for which the maximum is achieved.

A1.4 Change Point Problem, Multiple Change Point, Single Series

It is possible that a length of highway would have more than one break. Then the methods in Sections 4-6 of the main report will still give a reasonable expectation of picking out one of the steps or at least indicating that there is a break. However, the sensitivity may not be high, particularly where one break is followed by another in the opposite direction. Therefore, a test designed for picking out two steps in opposite directions is worth investigating. One possibility is to look at the difference between the maximum and minimum of Formula A1 without the absolute value signs (see Formula A5).

$$\frac{\max\limits_{0<k<n}\left(\sum_{i=1}^{k} X_i - \frac{k}{n}\sum_{i=1}^{n} X_i\right) - \min\limits_{0<k<n}\left(\sum_{i=1}^{k} X_i - \frac{k}{n}\sum_{i=1}^{n} X_i\right)}{\sqrt{n\,\mathrm{var}(X_i)}} \qquad \text{(Formula A5)}$$

This test statistic arises naturally if one starts by looking for a test when the locations of the steps are known, and then adjusts the locations of the steps to maximise the test statistic. The resulting test statistic can be expressed as being composed of three terms, with Formula A5 being the most relevant when one wishes to detect two steps in opposite directions.

Formula A5 is particularly convenient because it requires only $O(n)$ calculations, whereas a reasonable assumption is that it would require $O(n^2)$ calculations when searching for two steps.

As with the multiple series test, critical values need to be found by simulation, and these are given in Table A1.2 (n=100; 100,000 simulations).

Table A1.2 Estimates of critical levels for cusum method to find location of two steps.

Significance Level	Critical Value	Significance Level	Critical Value
0.5	1.11	0.02	1.79
0.4	1.18	0.01	1.90
0.3	1.26	0.009	1.91
0.2	1.36	0.008	1.93
0.1	1.50	0.007	1.94
0.09	1.52	0.006	1.96
0.08	1.55	0.005	1.99
0.07	1.57	0.004	2.01
0.06	1.60	0.003	2.05
0.05	1.63	0.002	2.11
0.04	1.67	0.001	2.20
0.03	1.72		

Unfortunately there does not seem to be any simple generalisation to the multiple time-series situation that does not require $O(n^2)$ calculations.

A1.5 Adjustment for Auto-Correlation

The previous sections assume that the successive observations are statistically independent. In the surface performance measures, this is not the case and some correction must be made to allow for the auto-correlated series.

The approach used was to fit a first order auto-regressive, first order moving average process (see Fuller 1996) to each of the series and then replace the terms such as $\mathrm{var}(X_i)$ in the test statistic with Formula A6:

$$\lim_{n\to\infty}\left\{\frac{1}{n}\mathrm{var}\left(\sum_{1}^{n} X_i\right)\right\} \qquad \text{(Formula A6)}$$

Because Formulae A2, A4, and A5 involve sums of several adjacent terms and because the limit in Formula A6 converges reasonably quickly, this provides a good approximate method.

A1.6 First Order Auto-Regressive, First Order Moving Average Model

A first order auto-regressive model (AR1) is one of the simplest models for expressing dependence between the X_i. It supposes that successive X_i are generated by Formula A7.

$$X_i = \mu + \alpha X_{i-1} + \varepsilon_i \qquad \text{(Formula A7)}$$

In Formula A7, ε_i are independent random variables, and α and μ are constants with $|\alpha| < 1$. Formula A7 implies that the random part of the current observation is equal to a fraction of the previous one plus a random shock. Despite the asymmetric appearance of this formula, the same model is obtained when dealing with normally distributed data and the process is run in the opposite direction.

The correlation dies away exponentially, as in Formula A8.

$$c_k = \text{cor}(X_i, X_{i+k}) = \alpha^{|k|} \qquad \text{(Formula A8)}$$

The limit of Formula A6 is shown in Formula A9.

$$\lim_{n\to\infty}\left\{\frac{1}{n}\text{var}\left(\sum_1^n X_i\right)\right\} = \frac{1+\alpha}{1-\alpha}\text{var}(X_i) \qquad \text{(Formula A9)}$$

Typically the lag 1 auto-correlation $\text{corr}(X_i, X_{i+1})$ is used as an estimator of α.

If the road surface actually follows a first order auto-regressive process, with our measurements being subject to random error, then what follows is a first order auto-regressive (AR1), first order moving average (MA1) process which can be represented by Formula A10.

$$X_i = \mu + \alpha X_{i-1} + \varepsilon_i + \beta\varepsilon_{i-1} \qquad \text{(Formula A10)}$$

The correlation function, therefore, is of the form $c_0 = \text{corr}(X_i, X_i) = 1$, and for $k \neq 1$ Formula A11 applies.

$$c_k = \text{cor}(X_i, X_{i+k}) = A\alpha^{|k|} \qquad \text{(Formula A11)}$$

where A is a function of α and β.

Formula A6 therefore simplifies to Formula A12:

$$\lim_{n\to\infty}\left\{\frac{1}{n}\text{var}\left(\sum_1^n X_i\right)\right\} = \left(1 + \frac{2c_1^2}{c_1 - c_2}\right)\text{var}(X_i) \qquad \text{(Formula A12)}$$

The auto-correlations c_1 and c_2 can be replaced by estimates.

A1.7 Auto-Correlation Structure

In this report the assumption is that correlation can be modelled by a simple auto-regressive (AR)/moving average (MA) model.

Another family of models that are known as long-memory processes (Beran 1994) was originally developed by an engineer for describing river flows. They assume that auto-correlations decay at a power law rate rather than exponentially. It is possible that they would provide a more appropriate description of the auto-correlations and so could result in a better correction factor.

A1.8 Normal Scores Transform

The preceeding theory works best if the data is approximately normally distributed, although it does not have to be exactly normally distributed. Even when transformed, the HSD condition data tends to have occasional very large or very small values. These will tend to disrupt much of the analysis. One of the ways of reducing this effect is through a *normal scores* transformation.

Given a set of observations $X_1,\ldots,X_n$, replace each of the observations, X_i, by its *rank*, R_i so the smallest observation has $R_i = 1$, and the largest has $R_i = n$, where n is the number of observations. The analysis is then performed using the distribution in Formula A13.

$$Y_i = \Phi^{-1}\{R_i/(n+1)\} \quad \text{(Formula A13)}$$

where Φ^{-1} is the inverse of the normal distribution function.

If the data started out as being normally distributed, very little information is lost (apart from location and scale information), provided n is not too small. However, the influence of outliers will be greatly reduced.

To be strictly correct, the transform Formula A13 should be referred to as the *Van der Waaerden* transform or an approximate normal scores transform. The exact normal scores transform is usually referred to as the Fisher-Yates-Terry-Hoeffding test (Hájek 1969: p.148).

A1.9 Trimmed Correlation

An alternative, possibly less drastic, approach for calculating correlations when outliers are present is the trimmed correlation. This discards, say, 10% of the data that do not seem to belong with the bulk of the data. The user sets the actual fraction discarded. This means that unusually large or small observations will have less effect on the calculated correlation than they would have with the standard correlation formula. Usually, this results in a more meaningful estimate of correlation. The S-plus data-analysis documentation (Anon 1999) supplies for details of the calculation method.

A1.10 Change Point Methods

For a general reference for change point methods for independent observations, see Chen & Gupta (2000), who consider tests based on Formula A3 with $\Delta = 0$.

Bai (1997), also using Formula A3, shows that searching for change points one at a time is effective for multiple change points. This result should carry over to tests based on Formula A2. Nevertheless, sensitivity will be quite low when a change is followed by a second change in the opposite direction.

The choice between Formulae A2 and A3 as the change-point formula to use needs to be investigated further. Simulations suggest that a better trade-off in sensitivities in different parts of the series can be found than that provided by Formula A2.

A further question is the type of change that the test is optimised for. Both Formulae A2 and A3 will be sensitive to gradual changes as well as to sudden changes. If one is interested primarily in sudden changes, then one could develop a test that is better optimised for this, and at the same time is less affected by long-range correlation.