Evaluating the network condition changes of transit networks managed under PSMC procurement options

August 2007

Land Transport New Zealand
Research Report 324

Evaluating the network condition changes of transit networks managed under PSMC procurement options

Peter Kadar and Theuns Henning
MWH NZ Ltd

Land Transport New Zealand Research Report 324

ISBN 0-478-28736-4
ISSN 1177-0600

PO Box 2840, Waterloo Quay, Wellington, New Zealand
Telephone 64-4 931 8700; Facsimile 64-4 931 8701
Email: *research@landtransport.govt.nz*
Website: *www.landtransport.govt.nz*

Kadar, P., and Henning, T. 2007. Evaluating the network condition changes of transit networks managed under PSMC procurement options. *Land Transport New Zealand Research Report 324.* 39 pp.

MWH New Zealand Ltd
Level 4, 585 Great South Road
PO Box 12 941, Penrose, Auckland
Tel: 64-9-580 4500
Fax: 64-9-580 4514

Keywords: contract, pavement performance, PSMC, roughness, rutting

An important note for the reader

Land Transport New Zealand is a Crown entity established under the Land Transport New Zealand Amendment Act 2004. The objective of Land Transport New Zealand is to allocate resources in a way that contributes to an integrated, safe, responsive and sustainable land transport system. Each year, Land Transport New Zealand invests a portion of its funds on research that contributes to this objective.

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

While this report is believed to be correct at the time of its preparation, Land Transport 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. They should not rely on its contents in isolation from other sources of advice and information. 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 Land Transport New Zealand but may be used in the formulation of future policy.

Acknowledgements

The authors wish to thank to Transit New Zealand and Land Transport New Zealand for their assistance and support. Special thanks are expressed to our colleagues at MWH for their contribution to this work.

Glossary

Average	An **average**, **mean** or **central tendency** of a data set refers to a measure of the 'middle' or 'expected' value of the data set. It is also called the arithmetic mean
CDF	Cumulative distribution function
dTIMS	Deighton's Total Infrastructure Management System software
HDM4	Highway Development and Management System
KPI	Key performance indicator (used interchangeably with KPM). For the purpose of this report we have used KPM.
KPM	Key performance measure
Median	The middle value that separates the higher half from the lower half of the data set
Mode	The most frequent value in the data set
NAASRA	National Association of Australian State Road Authorities (predecessor of Austroads)
OPRC	Output- and performance-based road contracts
PSMC	Performance-Specified Maintenance Contract
RAMM	Road Assessment and Maintenance Management software
SMA	Stone mastic asphalt

Contents

Executive summary

Performance-specified maintenance contracts (PSMC) have been operational in New Zealand for more than six years. These contracts are driven by key performance measures (KPMs) that are used to define the expectations of the road agencies and monitor the progress and performance of the contractor. As the effectiveness and efficiency of the KPMs are vital for achieving the desired results, it is essential to examine the effectiveness of the current KPMs in controlling and directing the maintenance contracts.

Land Transport New Zealand commissioned MWH New Zealand to investigate the effectiveness of the current KPMs and propose alternative measures if necessary. This report is the summary of the work carried out by MWH New Zealand.

The investigation which took place in 2006 highlighted the deficiencies in representing the data with the average, assuming a symmetrical (normal) distribution. While this simplification is instigated by the broad availability of and familiarity with algorithms for normal distribution, it may cause significant errors when interpreting and distilling network level data. An alternative KPM was explored and was found to represent the network condition better than the currently used average.

True summation of the network condition has significant impact on (a) network management, (b) performance monitoring and (c) forecasting and modelling.

Network management, in fact management in general, relies on information. The more accurate the information, the more likely that management decisions will also be correct. It is, therefore, essential that the network condition is characterised in a manner that reflects it accurately. The proposed method of using the mode instead of the average seems to provide substantially more meaningful network level information and its introduction will improve the effectiveness of network management.

Monitoring the network performance in terms of the average is in conflict with engineering observations. The introduction of the mode has confirmed engineering judgement that, despite the hardly changing averages, some networks seem to be deteriorating. Mode allows insight into the nature of annual fluctuations of averages and reveals when the annual changes reflect measuring error rather than changes of network condition. As mode indicates the property of the majority of the data, large swings in its data are highly unlikely. Consequently, substantial up and down movement of the mode in subsequent years will indicate data error. Using mode, systematic errors, ie errors affecting most of the data, can be detected and the quality of the data easily checked.

Because of the limited representative nature of the average, network level forecasting and planning based on the average is unlikely to be sufficiently reliable. Modelling, or forecasting, must be consistent with the method of network characterisation. In the light

of the findings of this work, models need to be re-evaluated and the validity of the forecasts using mode and averages checked. In changing from average to mode as the optimisation target, it is likely that some deficiencies of long-term modelling and those of the associated work programmes can be eliminated. The key recommendations arising from this study are:

- a review of the current definition of KPMs, and
- the replacement of the currently used data average with the data mode.

Additional parameters, such as the extreme values or percentiles, will increase the ability to accurately convey the road owner’s intentions.

KPMs are currently defined independently from each other. However, they are not truly independent, as maintenance or construction activities tend to have impact on more than one KPM at the same time eg when improving roughness, the seal age and rut depth will also change. The introduction of a synchronised set of KPMs based on more representative definition will potentially increase the contractors’ ability to influence the outcome and at the same time increase the owners’ confidence in the outcome of the contracts. In the next phase of this work, the investigation will include other properties, such as texture depth as well as other networks and sub-networks. It will also refine the methodology to suit the reporting and management requirements of network maintenance contracts. This work is already under way on the Transit NZ network. Before introducing the mode as part of the KPMs, other statistical measures should also be explored to ensure that the mode is the most effective method to characterise the network condition.

Abstract

Performance specified maintenance contracts (PSMC) have been operational in New Zealand for more than six years. These contracts are driven by key performance measures (KPMs) that are used to define the expectations of the road agencies and monitor the progress and performance of the contractor. As the effectiveness and efficiency of the KPMs is vital for achieving the desired results, it is essential to examine the effectiveness of the current KPMs in controlling and directing the maintenance contracts. The report examines the interpretations of the collected data using average and mode. The poor representation of the total network condition by the traditionally used average is illustrated by numerous examples. Alternative representation of the network condition is proposed and illustrated by using the mode of the data set.

1. Introduction and key performance measures

Performance specified maintenance contracts (PSMC) have been operational in New Zealand for more than six years. These contracts are driven by key performance measures (KPMs) that are used to define the expectations of the road agencies and monitor the progress and performance of the contractor. As the effectiveness and efficiency of the KPMs is vital for achieving the desired results, it is essential to examine the effectiveness of the current KPMs in controlling and directing the maintenance contracts.

Land Transport New Zealand commissioned MWH New Zealand to investigate the effectiveness of the current KPMs and propose alternative measures if necessary. This report is the summary of the work carried out by MWH New Zealand.

The outsourcing of road maintenance procurement methods requires an increasingly specific definition of the service to be delivered. The most traditional service delivery method involving in-house resources relies on meeting process or material specifications. It is widely assumed, though not necessarily proven, that meeting specifications will produce the desired long-term results. The in-house workforce is typically governed by a budget and specifications as opposed to network-level performance targets in outcome-oriented contracts.

The introduction of the principles of quality assurance instead of quality control reduced the day-to-day involvement of the owner in the delivery process. Owner-managers moved a step further away from direct involvement, and thus direct control, with the introduction of hybrid and PMSC-style contracts.

Performance-based contracts confronted owners and managers with the need to move away from process specification to define the desired performance. As the contractor assumed stewardship of the road asset, the owner lost the opportunity to influence the outcome through direct involvement. The owner’s desire now needs to be clearly expressed in the performance specification of the maintenance contract.

The concept of performance-based specification highlighted the need for road agencies to clearly define exactly what they wanted to achieve in terms of network performance. Performance-based maintenance contracts forced road agencies to develop their performance-based specifications and through these to quantify their engineering and operational objectives.

An adequate performance-based specification is based on the following essential requirements and/or assumptions:

- The performance requirements are consistent with the policies and objectives of the community and with those of the owner.

- Policies can be expressed with the help of measurable parameters, ie qualitative policies can be translated into quantitative measures or parameters (these parameters are also called key performance measures or KPMs).
- The relationship between quantitative measures and future performance can be modelled reliably. Deterioration models for local conditions are available and are satisfactorily calibrated.
- The input parameters for the performance models can be measured satisfactorily and accurately at a cost commensurate with the asset value.
- The funding level of the asset management activities is consistent with the desired outcome and asset value.

The current work focuses on KPMs, ie the means of quantifying policies. Quantifying policies is essential for setting targets and measuring progress. It might have been satisfactory to formulate wide-ranging statements about wanting good roads and safe transport when these were delivered by an in-house workforce. In a contractual situation the broad objectives must be quantified and be measurable – hence the introduction of the key performance measures.

Besides measuring performance or delivery of service, KPMs are also used to evaluate the effectiveness and efficiency of tenders by using the concept of 'value for money' (Transit New Zealand 2005). Value for money is defined as the functional performance per unit of resources consumed. For a meaningful tender evaluation 'value' must be defined in a manner consistent with the performance indicators. The value delivered by a contractor must reflect policies, ie any service or goods not contributing to declared policies represent none or substantially diminished value.

KPMs thus fulfil a dual role, namely to:

1. express the owner's policies and desires in a measurable format
2. act as the means of measuring progress or the fulfilment of the contract.

In the case of performance-based maintenance contracts, KPMs represent the owner's critical opportunity to convey their policies and the most important means of influencing the engineering outcome of a contract. It is, therefore, logical and essential to probe the extent the currently used KPMs meet the above requirements.

1.1 Key performance measures for road maintenance contracts

According to the United Kingdom local business website (www.businesslink.gov.uk):

*A key performance measure (KPM) helps a business **define** and **measure** progress toward its goals. KPMs are **quantifiable** measurements of the improvement in performing an activity that is critical to the success of a business.*

The generic nature of the above definition reflects its validity and applicability across all industries and business. Specific examples and guidance are given to construction and housing business in the *KPI report for the Minister for Construction* (KPI Working Group 2000). The report contains a brief background and practical examples for key performance measures related to timeliness, quality and user satisfaction. Over 200 web sites can be found under the key words ‘key performance measures’ and ‘road maintenance’, indicating the widespread use of the concept. Other examples, not necessarily related to roads, are abundant. The Ministry of Training (Canada) regularly publishes KPMs for Canadian universities and colleges, see *www.edu.gov.on.ca/eng/general/postsec/uindicator.html* .

While KPMs are embraced by a large number of private and government businesses, the KPMs published seem to be limited to simple expressions. A typical KPM may express the level of achievement by a simple ratio. For example the Los Angeles Department of Transportation and Development’s *2004/2005 Strategic goals and objectives* gives the following key performance measure:

> *Maintain 93% or more of the miles on the National Highway System in fair or good condition*

This KPM is specific about the required performance, but leaves the definition of ‘fair and good’ open for interpretation or definition elsewhere.

In general, KPMs have two key elements, namely

- the specification of the desired performance (in the above example 93%)
- the content definition of the measure of the performance (here ‘fair and good condition’).

The first part of the KPM definition, ie the specification of the performance level is typically given as

$$KPM = f\,(P)$$

where P is a performance measure. The function is usually linear and has only one valid solution – the KPM.

The measure of performance can be any quantifiable parameter, such as income or profit. The use of KPMs has been dramatically expanded by adopting an engineering parameter as performance measure. Engineering or physical parameters introduce a new dimension to measuring performance with KPMs. The issues that need to be considered when establishing a KPM are discussed briefly below.

1. **Measurable**: Can the selected parameter be measured reliably and accurately? Once a performance measure is adopted in a contract it must be measurable in a manner that leaves no doubt about its meaning and value. In a business environment measures are typically derived from accounts or other administrative records. In the

engineering world, other issues, such as repeatability, accuracy, bias etc add a level of uncertainty to the measured value that has to be taken into account.

2. **Representative**: The calculated KPM must be clear and unambiguous leaving no doubts about its value and meaning. It must describe and present the parameter fully and unambiguously.
3. **Controlling power**: The selected performance measure must be in a controlling position, ie it must be capable to affect changes on the business outcome. For example 'the number of sunny days' would be an unsuitable performance measure as it is beyond the control of any business. Similarly, measuring and using measures that cannot be controlled by the business is, at best, futile and, at worst, can lead to serious miscalculations and wrong decisions.
4. **Predictability**: As a consequence from the required controlling power, the performance measure must be predictable by using appropriate models. Modelling performance is an essential requirement for managing medium- to long-term contracts successfully. In practice, modelling capability reflects the understanding of fundamental relationships between cause and causatives. Without clear understanding of these, management decisions are based on guesswork and the KPM will reflect the success of guessing and not the quality of management.

The criteria listed above may be used to evaluate the effectiveness of KPMs.

Before going further in our discussion of KPMs, we need to consider where and how KPMs are used in road maintenance contracts. Road maintenance contracts reflect the intention of the road owners to provide a certain level of service. In the context of maintenance contracts, level of service can be defined for long-term and short-term activities. The latter typically concern a response to unplanned events, therefore, they measure organisational capability and performance rather than network-level pavement performance. Short-term performance measures are typically defined in terms of response time rather than engineering measures. Long-term performance on the other hand is defined by engineering parameters. Our further discussion focuses on long-term performance measures using engineering parameters.

1.1.1 Long-term performance measures

Long-term performance measures became a significant issue with the advent of outcome-based or performance-specified maintenance contracts (PSMC). These contracts typically utilise measurable engineering properties, such as

- roughness
- rutting
- cracking
- deflections
- skid resistance.

The above parameters reflect important aspects of road performance. These parameters are treated in New Zealand and Australia as individual performance measures, reflecting their perceived significance in long-term pavement performance.

1.1.2 Content definition of KPMs

After naming the long-term performance measures, the most critical task is to define how to evaluate the measured parameter. PSMC-style contracts typically use one or a combination of the following definition methods:

- **Average** of the parameter is calculated as the average of the measured parameter for a network or sub-network. This definition is rarely used on its own now as its limitations are well recognised and it is considered to yield only a limited representation of network performance.
- **Definition based on distribution** is used quite widely. This type of definition is based on the recognition that measured parameters follow a certain distribution. This type of definition reflects the intent of controlling the distribution as opposed to controlling only the midpoint or average of the distribution. Typically the average and a percentile are defined, ie at least two points of the cumulative distribution curve are fixed in the contract. The full distribution is defined by 24 points in the performance-based maintenance contracts currently underway in Western Australia.

2. Effectiveness of current KPMs

The effectiveness of KPMs can be measured by their success in achieving the owner's objectives. The road owner's objectives are deemed to be achieved when the KPMs, ie contractual obligations, are met. This definition assumes that the owner's objectives are clearly reflected in the KPMs. For the purpose of the current investigation we separate the two issues, namely the contractor's success in achieving KPMs and the owner's satisfaction with the resulting network performance. By separating the two, we can focus on the underlying question of the capacity of current KPMs to convey the owner's intent.

2.1 Satisfaction with current KPMs

A limited survey was carried out to explore the current level of satisfaction with the outcome of the maintenance contracts. The interviews with selected network managers and stakeholders can be summarised as follows (with the interview reference number in brackets):

- The KPMs are generally achieved, but they do not fully reflect the actual pavement condition. (1, 2, 3)
- The networks are perceived in good condition, but not necessarily by all stakeholders. Subjective opinions are influenced by a number of – sometimes – non-technical factors. (1, 2, 3)
- The KPMs do not drive the contracts – they act more like a weather vane, ie follow rather than lead. (1, 3) If and when KPMs are used as drivers, they tend to instigate reactive strategies. (1)
- Maintenance cost is a significant driver, particularly where traffic control or other work restrictions apply. High maintenance costs promote area-wide treatments, while low ancillary costs may drive the contractor towards patching and other shorter-term measures that yield improved KPMs but still consume the asset. (3)
- The high-speed vehicle data (roughness, rutting, texture depth) seem to be overvalued and at the same time their precision and reliability seems to be questionable. (2)
- A tighter control and definition of the KPMs would be desirable in a manner that better reflects the owner's intent. This should also eliminate the potential possibility of manipulating data. (2)
- A better understanding of the KPMs is essential to develop a more suitable mechanism for controlling contract outcome – until then we get what we deserve. (2)
- KPMs should reflect long-term trends rather than last year's data. (3)
- Asset consumption is a major concern that is not reflected in the KPMs. The current KPMs allow asset consumption – this should be controlled. (3)
- KPMs cannot be major drivers as their current definition and usage do not accommodate engineering considerations. (1, 3)

- KPMs currently imply a high level of forecasting accuracy. This is somewhat deceptive and a +/- one- to two-year anticipated accuracy would be more realistic. (3)
- KPMs cannot replace a good work relationship and trusting alliance between contracting partners. (1, 3)

2.2 Asset owners' intent – what the owners want

The general intent of road asset maintenance is to provide a road network that is in a condition 'fit for purpose' for the duration of its service life. This generic definition may be further refined by introducing key components of usability, durability, and service and comfort measures (World Bank 2005). Other grouping of technical parameters may also be used, such as safety or economic use of roads (see Table 2.1)

Table 2.1 Example grouping of network level technical requirements.

Durability	Preserve the asset for the future, ie limit asset consumption to normal tear and wear
Safety	Ensure safe passage for all road users
Economic use	Ensure that the road can be used at a cost deemed economical for the users (optimise user costs or roughness directly) and limit delays within reason
Service and comfort	Ensure a comfort level commensurate with the intended use of the road

Many other technical, social and environmental aspects can be defined, but most of them can be shoehorned into one of the above categories. The above key categories can then be quantified by a combination of measured parameters. The relationship between these is summarised in Table 2.2.

Table 2.2 Relationship between owner's intent and engineering parameters.

Technical parameter	Durability	Safety	Economic use
Roughness	X	X	X
Rutting	X	X	
Skid resistance		X	
Texture depth		X	
Delay time			X
Strength	X		
Cracking	X	X	X
Surfacing age*	X		

*Age is considered as a substitute for describing bitumen condition, surfacing integrity and flexibility.

The above properties are interdependent to various degrees. There are a number of studies that discuss these relationships, one of the most notable being documented and implemented in the Highway and Development Management System (HDM-4). A comprehensive system, dTIMS, has also been implemented in NZ.

When contemplating the owner's intent, it becomes obvious that no individual engineering performance indicator is capable of expressing this fully. The owner's intent can only be expressed with a combination of technical parameters. This leads to the concept of 'level of service' (LOS) that is gaining wider acceptance. However, before considering LOS, we need to review the currently used KPMs in detail.

2.3 A critical review of currently used KPMs

2.3.1 Definition of content

Current KPMs are typically expressed by using statistical terms such as average or percentile. One of the most commonly used – and abused – terms is the average of a population or data set. Average is probably the simplest way to characterise and summarise data. It is rarely used on its own to define an engineering KPM due to the generally acknowledged inhomogeneous nature of the data and thus road network conditions (a notable exception is the sample specification for OPRC, World Bank 2005) .

The inadequacy of the average in representing the owner's interest on its own is reflected in the many attempts to combine it with additional statistics. These may include:

- minimum and/or maximum values (World Bank 2005)
- percentiles (Southern Tasmania Performance Based Maintenance Contract)
- full distribution of the data (MRWA Term Network Contracts).

All methods, including those listed above, attempt to characterise the network property in a way that reflects its nature correctly and enables it to become an appropriate business driver. The best and most thorough characterisation of the network will be of limited value as a KPM if it steers the contractor in a direction that is not the most beneficial for the asset owner. For this reason, we must explore how the various KPM definitions drive the contractor in the desired direction.

The average of the test results gives a broad picture of the total population, but it does not control the extremes, ie it allows the existence of extreme values as long as they are counterbalanced with extreme values at the other end of the scale. The simplicity of calculating averages makes it relatively easy to use, though its meaningful use should be limited to situations where the data refers to units of the same or at least similar size; in the case of roads of equal length. Otherwise, (length) weighted average should be used.

The limitation of the average to control extreme values is acknowledged by attempts to limit these values. This can be achieved by simply stating the limits (eg 'no values should be less than…') or defining the limit in terms of the proportion of the network that can be outside the limit (eg 'the 90th percentile must be less than…' or '90% of the network must have a property better than…'). The same conditions apply as discussed before, ie the average and percentiles must be weighted by length or the data must refer to sections of equal length. This type of definition is used in Tasmania (Term Network Contract 662). The combination of average and percentile is essentially an effort to embed the desired distribution into the contract. To characterise the distribution, it would be more practical and appropriate to fix two percentiles along the distribution curve, such as the 90th and 50th (ie the median as opposed to the average). By defining the distribution of the property, the owner hopes to gain a reasonable control over the future qualities of the asset.

The owner’s hope can only partially be fulfilled as it is based on the assumption that the distribution is (a) continuous and (b) normal. To overcome the problems associated with the above assumptions, the full distribution curve – defined by 24 percentiles or points – was used as a KPM by Main Roads Western Australia. This clearly reflects and defines the owner’s desire, but also highlights a fundamental deficiency of the method. The prescription of a distribution curve as a contractual requirement assumes that the contractor has the means to control the distribution. In reality, the contractor can only exert influence at best on some points that interact with the total distribution, but has very limited capacity to influence all the points, ie the total distribution. The level of capacity to affect the distribution curve will, among others, depend on the proportion of the network treated by the contractor. The larger the proportion of a network treated, the more control the contractor will have over the outcome. Implicitly the size of the network influences the outcome, as the same treated length in a smaller network will have a greater influence. In other words, breaking down the network into smaller units increases the efficiency of control.

In order to gain better understanding of how the KPM definition impacts on the network, we need to consider the deterioration process and its impact on the distribution of the property. For the sake of clarity, we will use roughness as an example, though the following discussion is relevant to most data covering the full network.

For this particular example, let us assume a simple network that has just been constructed. Assuming good quality workmanship, the network displays the ‘perfect’ distribution (see Figure 2.1 – curve ‘a’), ie when all data have the same value. (For the time being we will ignore variations related to test method or conditions.) Of course when we consider variations due to testing and minor differences inherent to the road, curve ‘b’ – representing minor standard deviation – offers a more realistic representation of the initial overall condition.

Under a uniform traffic load, minor differences in material properties and construction quality appear, causing an increased scatter or variation (see Figure 2.1 – curve ‘c’). Curve ‘c’ however, does not truly reflect the situation, as it is a well known fact that sections in a worse condition will deteriorate faster than sections in a better condition. This is reflected in the incremental deterioration models implemented – among others – in HDM4. According to the incremental model:

$$R_{n+1} = R_n + \Delta R \quad \text{(equation 1)}$$

where roughness is the sum of the current roughness value plus the increment. As the increment is a function of several other properties that also increase from year to year, it is easy to see that the total roughness – or any other modelled property – will show an ever increasing trend. This increasing trend will force the upper end of the distribution to grow faster than the lower end, thus skewing the distribution, ie making it asymmetric, hence not normal. Consequently, the distribution cannot be normal if an incremental model is used (see Figure 2.1 – curve 'd').

Closer inspection of curve 'd' offers another clue about the inappropriateness of normal distribution for modelling deterioration. This can easily be proven by considering the conflicting nature of the normal distribution and pavement deterioration. Normal distribution is by definition symmetrical, ie when the average remains the same but the standard variation increases, the extreme values – the lowest and highest values in the population – will spread accordingly. This, of course, has physical limitations, as roughness may increase indefinitely but cannot improve beyond a certain limit. Consequently the distribution becomes asymmetrical, ie not normal (see curve 'd'). As curve 'd' indicates, increased average and standard deviation may still represent a situation where there is roughness improvement (see curve 'e').

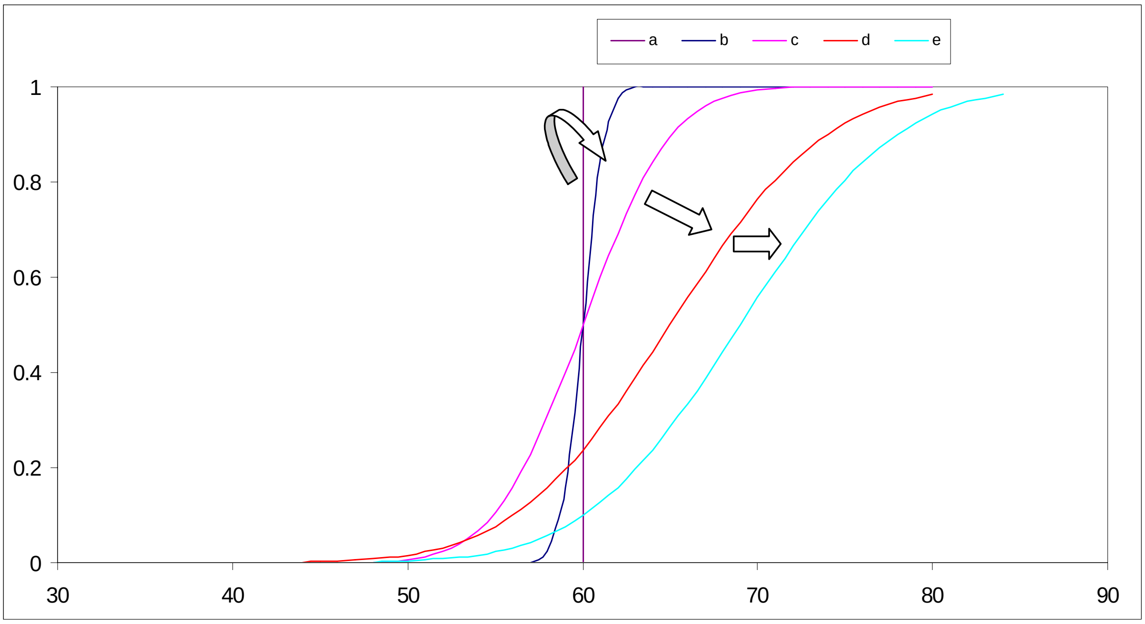

Figure 2.1 Network level roughness deterioration model.

The above considerations assume that (1) no work is done on the network, (2) the traffic pattern and loading does not change and (3) the pavement is of uniform structure and strength.

Following from the above consideration it may be concluded that:

1. Normal distribution cannot truly represent the network condition, therefore, any statistical parameter relying on normal distribution is unsuitable.
2. As long as the KPM is defined with a parameter assuming normal distribution, neither the contractor nor the owner has a true representation of the network property.
3. Without knowing the true nature of the network, it is almost impossible to exercise full control over it.

The above considerations are also pertinent for network level modelling as they imply that deterioration at network level cannot be modelled only by average without taking into account the changing nature of distribution. Current modelling practice is to aggregate data into shorter or longer sections and estimate deterioration based on the average of the analysed (treatment) section. This assumes a level of homogeneity rarely experienced in practice. However, network level performance is an aggregation of the segment level performances. As discussed before, this forms a non-standard distribution.

It should be noted that the above logic is also supported by the general observation that in reality deterioration progresses more slowly than predicted by most models. One of the possible explanations is that the average of the measured performance tends to be stable while the real changes occur around the edges of the distribution. However, this issue falls outside the scope of the current investigation and needs to be addressed elsewhere.

2.3.2 KPMs and long-term planning

A well defined KPM acts as an effective business driver. It is, therefore, necessary to explore the role of the KPM in driving the business to achieve the targeted level of the KPM.

For the following discussion let us consider a KPM that is defined as the average of a parameter and is used exclusively for a contract, ie no other KPMs are considered. Assuming equal treatment lengths, the most effective way to reduce the average value is to reduce the worst number in the population, ie to treat the worst road section. As the most improvement can be achieved by treating the worst section, this must be the most effective use of the available funding.

Selecting the worst section first for treatment drives the average downwards most efficiently regardless of the timing of the treatment. It can be concluded that a KPM defined as an average will always steer towards a ‘worst first’ policy. A KPM defined by average discourages long-term planning and optimisation and encourages selecting a reactive maintenance policy based on selecting assets in the worst condition.

The situation is not significantly different from that where the distribution of the population is defined by two or more points (percentiles) along the distribution curve. By removing the worst item in the population and replacing it with a better one (eg repair the oldest road section and thus replace the seal age with zero) the distribution of the data changes and the median (as well as the average) will move towards the better values (usually closer to the Y axis). The objective is accomplished as the whole distribution has improved. It can be concluded that when using a single KPM for driving maintenance, the ‘worst first’ policy will be the most effective, leaving no room for developing a long-term optimised work programme.

It should be noted that treating the worst section may be statistically the most effective, but not necessarily the most desirable for the road owner/user.

Fortunately all New Zealand and Australian maintenance contracts are driven by several KPMs concurrently which makes the choice less straightforward – and the 'worst first' policy less tempting.

2.3.3 Interaction between KPMs

When several KPMs are used, each KPM may drive the maintenance programme in a different direction by targeting different parameters. However, the road characteristics commonly used in KPMs are not independent variables but rather have a considerable level of interdependency. This interdependency is best seen when the response to trigger conditions (ie the choice of treatments) is considered.

While there is a range of conditions that can be characterised with a number of KPMs, there are only few fundamental treatment options:

Table 2.3 Treatment hierarchy.

Treatment category	Treatment example	KPM
Surface age	Rejuvenation	Surface age
Surface characteristics	Reseal, ultra thin asphalt overlays	Friction, skid resistance, cracking
Surface defects and short wave length unevenness,	Thin asphalt overlay	Rutting
Long wave-length unevenness	Thick asphalt overlay	Roughness
Structural deficiency	Rehabilitation, thick asphalt overlay	Deflection, remaining life

The first two treatment categories may be, and usually are, consolidated into one for most practical purposes, namely surfacing. The above treatment categories also represent a hierarchy as the higher treatments also cover the KPMs addressed by the lower ones.

While the treatments form a hierarchy, the KPMs are determined independently from each other in the maintenance contracts. Independently set KPMs that are interrelated through the treatment hierarchy may cause conflicts when KPMs are used for treatment selection. The resolution of the conflict may be biased; if it is biased towards the owner the contractor spends too much money, but equally it may be biased towards the contractor, in which case the asset is consumed despite meeting the individual KPMs.

As an example let us assume that the treatment area demanded by the roughness KPM exceeds the treatment area required by the seal age KPM. As the roughness treatment results in new surface, the age KPM may be automatically met. In this situation the age KPM is a secondary business driver and likely to be less effective. Unsynchronised interdependent KPMs may cause significant disappointment for one or both parties. It can, therefore, be concluded that synchronising interdependent KPMs is in the best interest of both owners and contractors.

2.3.4 Data reliability and accuracy

The above considerations assume that the physical data aggregated into the KPM is of adequate accuracy.

Roughness is probably the most critical data used as a KPM for network management. Detailed analysis of roughness data highlighted a number of anomalies. Analysis of data collected on an Australian maintenance contract indicated that about 30% of the network displayed improved roughness without applying any treatment (Kadar 2004). A rigorous investigation of 63 calibration sites (Henning et al 2004) demonstrated the importance of ensuring adequate machine calibration and highlighted several influencing factors. Driver and driving conditions, rigorous location referencing – or lack of – are all affecting the reliability of data.

The definition of the KPMs must take into account the repeatability and accuracy of the data. An envelope similar to the confidence interval would allow minor deviation from the KPM without the onus of paying a penalty or bonus.

3. KPMs and major works

3.1 Methodology

The relationship between measured performance and quantity of work, ie the influence of KPM, was investigated. Data was analysed at a network and sub-network level. The network level analysis disregarded the contractual sub-division of the network and treated it as a whole. Network level approach tends to suppress inherent variations. Sub-network level analysis tends to deal with more homogeneous population, so the results may be significantly different as 'balancing' of some trends can be avoided.

For calculating statistical parameters the data was initially represented by assuming normal distribution. As the data is not distributed symmetrically, log-normal distribution was adopted. For the calculation of percentiles exceeding set limits, approximate expressions were developed to avoid the need for integration or the need for statistical tables. Due to the logarithmic nature of the chosen distribution any approximation error may appear magnified. While this is acknowledged, the percentile calculated assuming log-normal distribution is still closer to the true distribution than the same parameter calculated assuming a normal distribution. The difference between the normal and log-normal distributions is illustrated in Figure 3.1.

Visual comparison with the data shows clearly that assuming normal distribution introduces substantial errors. The same chart also illustrates the various statistics used for the analysis.

Where parameters of different magnitudes were compared, the data was normalised, so the trends could be clearly identified.

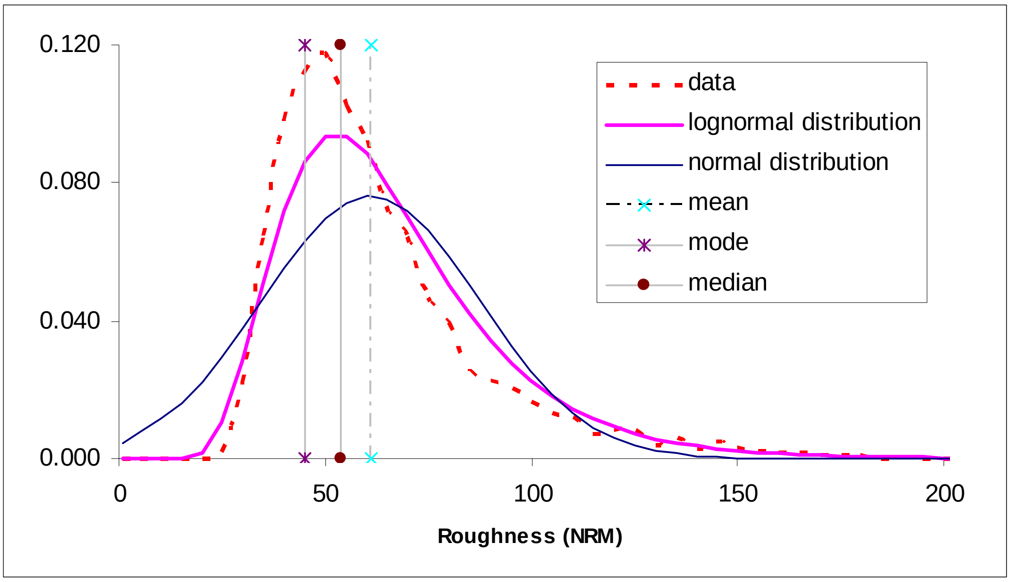

Figure 3.1 Comparison of normal and log-normal distributions.

3.2 Network level analysis

For the investigation of roughness and rutting performance, network data was obtained from the RAMM database. For the current analysis the raw data was used without pre-processing or aggregation. Data items with an event marker were removed to improve the reliability of the data. The annual data sets were separated by using the survey number. For the following discussion, data from the total network was used without breaking it down to sub-network level to illustrate key trends and points.

3.2.1 Roughness

The distribution of the annual roughness data is shown on Figure 3.2. The 100 m NAASRA roughness was used to generate the cumulative histograms with a bin size of five counts. The cumulative histograms are in close proximity, indicating relatively small variation over time. The 1994 data set displays an irregular distribution, which is more a reflection of the selected bin size rather than the data. This issue is discussed further in Section 3.3.

The largest difference between the annual average roughnesses within the contract period is about 2.1 counts (see 2001 and 2002 data in Figure 3.2). The skewness of the data is considered significant, particularly in the earlier data sets. Significant positive skewness indicates asymmetrical distribution distorted to the left, ie with a longer tail towards larger positive numbers. Over a longer period the annual variation of the statistical parameters is quite small, probably approaching the measuring error of high-speed roughness data measurement. Based on this data only, it would be difficult to state whether the limited variation of the roughness level is due to good management or a very low level of deterioration as there is no reference data for any location where no maintenance was carried out. The relationship between treatments and roughness is explored in Section 3.3.2.

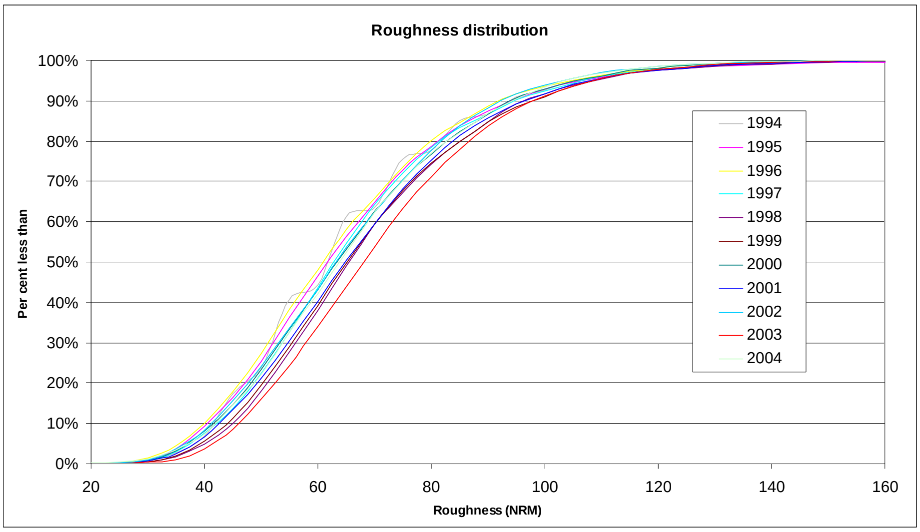

Figure 3.2 Roughness distribution – full network.

Table 3.1 Summary statistics for the network roughness data.

Statistics	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2002 (2)**
Mean	66.6	66.6	65.7	66.0	64.7	66.8	69.4	68.9	66.7	68.2	66.1	66.0	66.7	66.7
Standard Error	0.354	0.354	0.364	0.385	0.363	0.401	0.364	0.358	0.298	0.307	0.295	0.293	0.275	0.303
Median	64	64	63	62	61.5	63	66	66	64	65	64	64	64	64
Mode	54	54	54	52	51	56	67	56	64	65	63	61	61	61
Standard Deviation	22.24	22.24	21.58	23.29	21.65	23.51	21.73	21.54	21.04	21.74	20.69	20.79	20.70	21.45
Sample Variance	494.53	494.53	465.64	542.41	468.60	552.81	472.21	463.79	442.60	472.49	427.99	432.12	428.32	460.28
Kurtosis	1.79	1.79	1.90	8.01	2.23	51.07	2.82	1.82	0.63	1.36	1.25	1.12	0.70	1.78
Skewness	1.11	1.11	1.10	1.78	1.09	3.72	1.14	0.95	0.77	0.92	0.79	0.80	0.71	0.96
Range	184	184	184	253	181	492	200	217	139	169	190	176	157	178
Minimum	21	21	21	20	21	18	18	19	21	24	20	19	18	23
Maximum	205	205	205	273	202	510	218	236	160	193	210	195	175	201
Sum	262669	262669	231408	241624	230239	230133	246877	249296	332749	341940	325028	332756	378635	333544
Count	3943	3943	3524	3661	3558	3443	3559	3619	4987	5011	4914	5042	5674	4998

Based on the presented statistics and chart the recognition of any trend would be quite difficult. However, when the mode and percent above a given limit were selected, trends could be clearly identified. (Mode and average are the same for symmetrical distributions, such as the normal distribution. For asymmetrical distributions the mode provides information on the location of the peak of the distribution, ie where most of the data is (see Figure 3.3).

Based on the averages shown on Figure 3.3, the network remained practically the same over the observation period. However, the mode (ie the position of the majority of the data or the peak of the probability distribution curve) shows that the most frequent data – the peak of the distribution – has moved upwards, indicating the overall aging or deterioration of the road. As the proportion of the data exceeding the set target limit (here 110 counts) decreases, it may be stated that while the road shows normal tear and wear, the proportion of the network beyond the set limit has decreased. The two parameters together indicate a normal tear and wear and at the same time a satisfactory level of renewal.

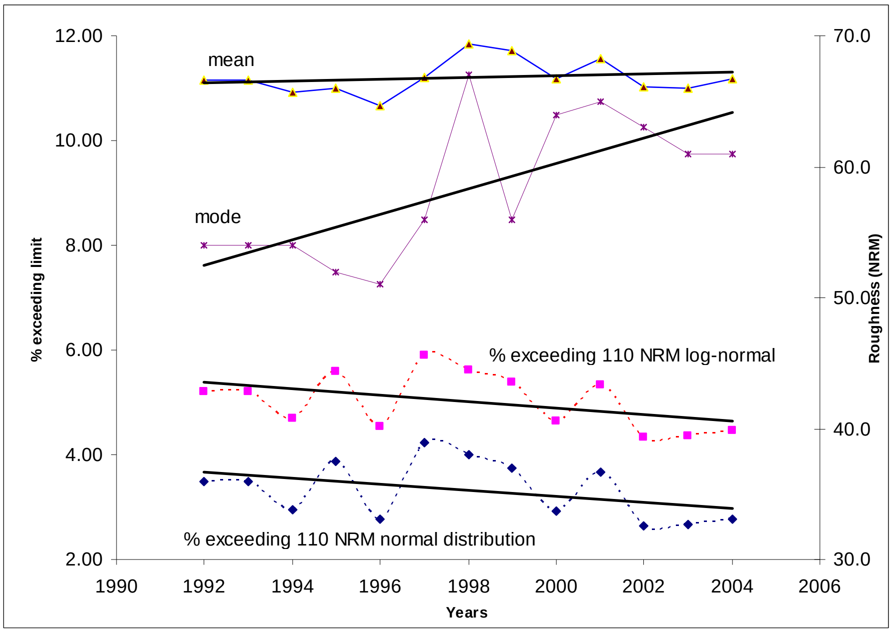

Figure 3.3 Network level roughness trends.

Figure 3.3 also offers a comparison between the normal and log-normal distributions. By definition, the gap between these two varies and increases proportionately with the distance from the intercept of the normal and log-normal probability distribution curves.

3.2.2 Rutting

Rutting data at 20 m intervals was extracted from the RAMM database and was used for the current work. Out of the several rutting parameters included in the database the mean rut depth in the left wheel path was selected. The cumulative histogram shown in Figure 3.4 was generated by using bins of 1 mm. The histograms show close proximity with the most apparent differences between 80th and 90th percentiles. The closeness of the later data sets (2000–2004) is remarkable, while some of the earlier data displays irregular distribution. The summary statistics hardly reflect any change in the network condition (see Table 3.2).

Table 3.2 Summary statistics for the network mean rut depth data.

Statistics	1996	1997	1998	1999	2000	2001	2002	2003	2004
Mean	5.35	3.12	3.88	3.24	3.80	3.90	4.40	4.38	4.28
Standard Error	0.1	0.0	0.018	0.015	0.015	0.017	0.018	0.021	0.020
Median	4.0	2.0	4.00	2.00	3.00	3.00	4.00	3.00	3.00
Mode	4.0	2.0	2.00	2.00	2.00	2.00	2.00	2.00	2.00
Standard Deviation	3.4	2.6	2.58	2.21	2.55	2.80	3.03	3.45	3.39
Sample Variance	11.4	6.8	6.66	4.87	6.49	7.84	9.15	11.89	11.51
Kurtosis	2.5	4.9	5.68	4.67	4.36	4.08	4.13	6.86	6.45
Skewness	1.4	1.7	1.78	1.61	1.59	1.57	1.71	2.05	2.01
Range	24.0	24.0	27	24	26	28	27	37	39
Minimum	0.0	0.0	0	0	0	0	0	0	0
Maximum	24.0	24.0	27	24	26	28	27	37	39
Sum	5475.0	67132.0	83168	70142	110933	108376	119046	120613	124829
Count	1023.0	21506.0	21439	21618	29201	27773	27074	27525	29143

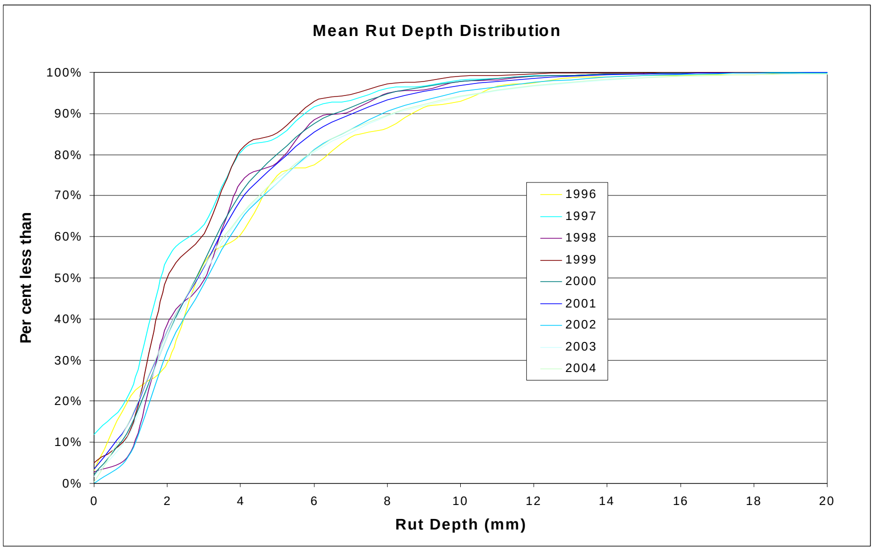

Figure 3.4 Rut depth distribution – network level

Closer investigation of the histograms revealed a significant level of sensitivity to the bin size. The impact of bin size on the cumulative distribution is illustrated in Figure 3.5. To maintain consistency in data processing the 1 mm bin size was used for the current work, though the irregular histograms raised the concern of previous data or data processing issues. Errors related to the bin size can be avoided by using either the full data set or the formula developed for the normal or log-normal distributions, as these are independent from the bin size.

The network distributions also displayed a consistent pattern supporting earlier findings that the actual distribution of the rutting data was asymmetric and would be poorly represented by a normal distribution.

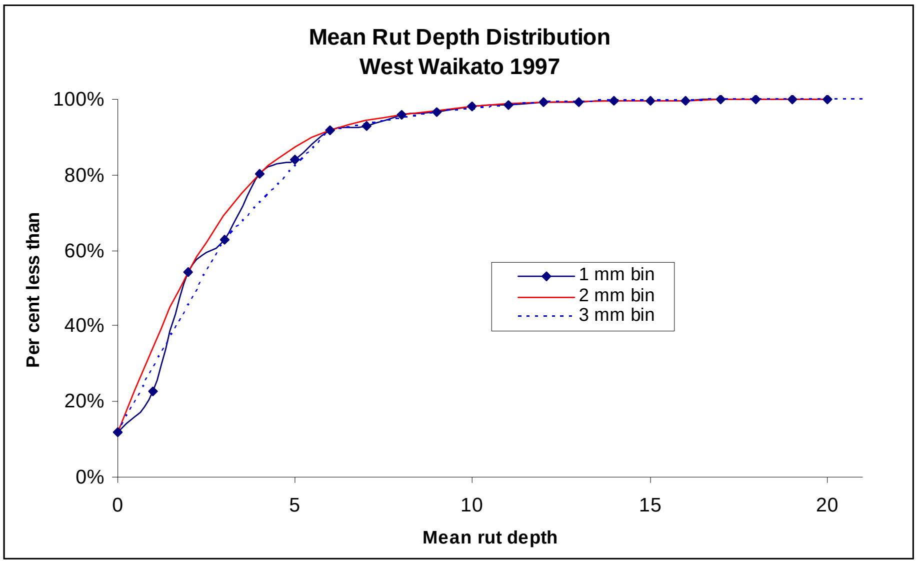

Figure 3.5 Impact of bin size on histograms.

The average rut depth has increased marginally over the observed period. The mode has not changed, indicating that the majority of the data is still centred on the same value after so many years. However, the distribution has become wider, as the percent exceeding the chosen limit (here 10 mm) has steadily increased. The combination of these parameters indicates that the worst parts of the network most likely worsened, while the majority remained unchanged (see Figure 3.6).

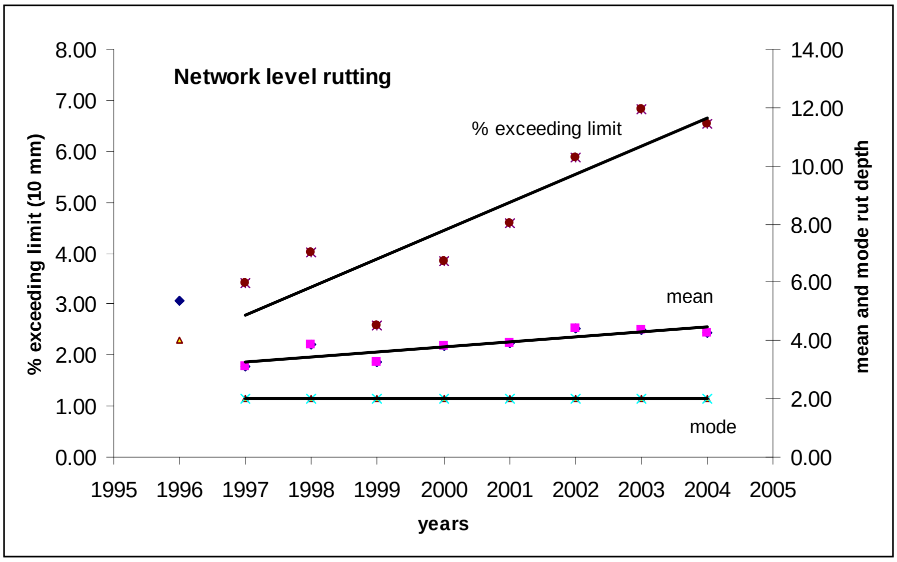

Figure 3.6 Network level rutting.

3.2.3 Work on the network

The intended role of KPMs is to guide the network managers to conduct work in such a manner that set targets are achieved. Consequently, a direct relationship between the KPMs and the nature and quantity of work on the network can be assumed.

In order to keep the current investigation focused, only major works will be considered here. To reflect the intention of the KPMs, pavement work is grouped into three major categories, namely works that:

- affect the surfacing only (seal)
- mainly affect the surfacing, but also improve the shape of the pavement to a lesser degree (shape)
- improve or reinstate strength and shape of the pavement (structural).

These work categories align reasonably well with the major KPMs used (see Table 3.3).

Table 3.3 Work categories and KPMs.

Work category	Work item	KPM
Seal	PSEAL,	Seal life
Shape	Slurry, OGPA, SMA, thin asphalt overlay	Rutting, seal life (and roughness to a limited extent)
Structural	Structural overlay, reinstatement	Roughness, rutting, seal life

The summary of the work conducted on the network covers a period of 50 years. Figure 3.7 shows the work on the network in terms of treated length. The initial construction work was clearly followed by an increasing maintenance or seal/shape work component. When investigating the proportion of the different work categories relative to the total treated length, the increased proportion of the shape/seal treatments after the start of the contract (1998) is apparent (Figure 3.8).

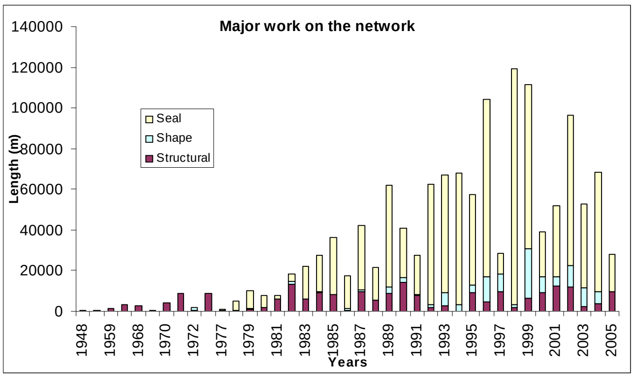

Figure 3.7 Work on the network.

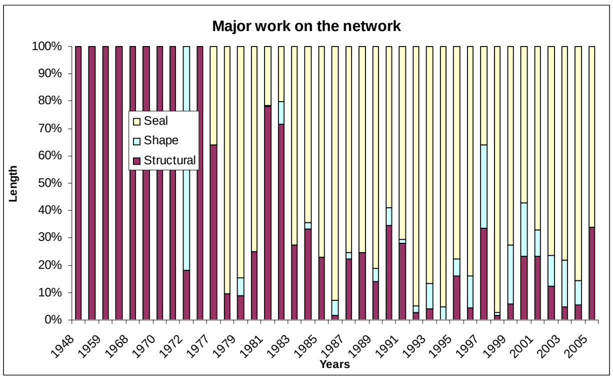

Figure 3.8 Proportional distribution of work on the network.

Earlier data from before approximately 1980, may refer to new construction or first seal, so long-term trends would be difficult to deduce from the early data. Later data (after

1980) reveals an increasing quantity of resurfacing and reshaping work. Some cyclical trend may be observed with significantly reduced total treated length in 1997 and negligible structural work in 1998. The direct correlation between the network level treated length and roughness/rutting was explored and was found to be weak. For the purpose of finding correlations the performance data was shifted by a year forward, ie it was assumed that the results of treatments could be measured.

3.3 Sub-network level analysis

3.3.1 Roughness

Analysis of sub-network level data was conducted on a network sub-divided into four distinct sub-networks. The sub-networks represented a combination of geographical and traffic considerations to create a more homogeneous population. Each sub-network had a different roughness target average and upper tolerable limit. The KPM requirements for roughness are summarised below in Table 3.4.

Table 3.4 Roughness KPM requirements.

Sub-network	Average	Maximum tolerable	% exceeding maximum tolerable
1	92.0	150.0	0.1
2	70.6	120.0	2.0
3	76.4	130.0	1.4
4	64.6	120.0	1.9

The roughness history is presented in Figure 3.9 in terms of average roughness and in Figure 3.10 relative to the KPM. Presenting the average roughness relative to the KPM allows the direct comparison of the four sub-networks on common terms and at the same time highlights their relative performance. The charts clearly indicate that the average roughness was below (ie better than) the target at all times. The relative or normalised presentation of averages emphasises the prevailing trends, ie the substantial reduction of average roughness on three sub-networks and the increase of roughness on one network.

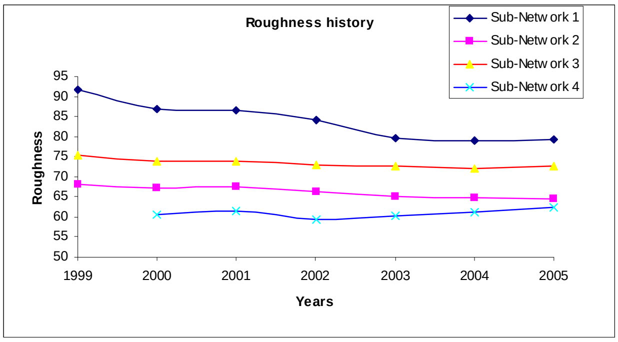

Figure 3.9 Roughness history (NRM).

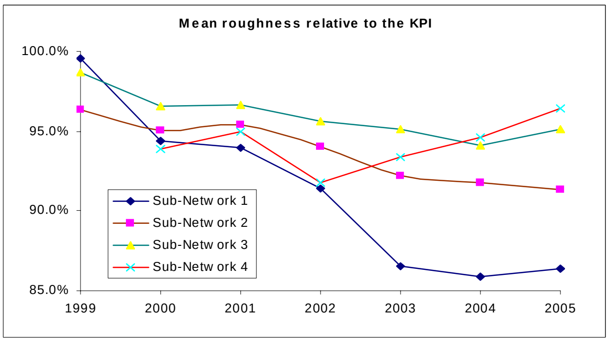

Figure 3.10 Roughness history relative to target roughness.

The mode of the network data indicates a somewhat different trend. The mode represents the roughness value most frequently recorded, ie the peak of the distribution. Figures 3.11 to 3.14 show the mode, mean and percent exceeding the limit for the four sub-networks. In most cases the mean indicates minimal improvement while the mode shows substantial fluctuation and an increasing trend. The charts indicate that while the average is reasonably steady, the bulk of the network has deteriorated. The downward − reducing − trend of the percent exceeding the limit reflects the contractor's effort to control the mean roughness by treating the worst sections. The consistent conflict between mode and mean highlights the high probability of misinterpreting the data by using averages instead of mode. It should be noted that no efforts were made to fit the best curves to the data – the displayed fitted curves illustrate only the major trends.

The substantial movement of the mode from year to year on some of the network may indicate either or both data collection and real roughness problems. However, it is unlikely the peak of the distribution would move about 20 counts in a year, effectively indicating that the most frequent data observations moved up significantly.

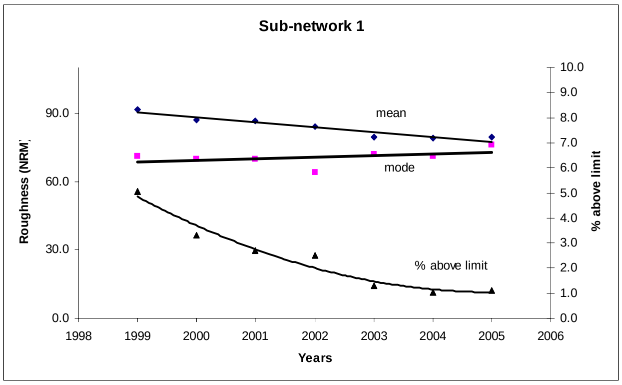

Figure 3.11 Roughness – Sub-network 1

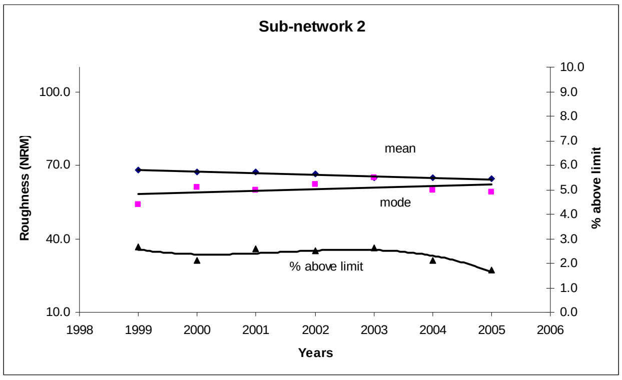

Figure 3.12 Roughness – Sub-network 2

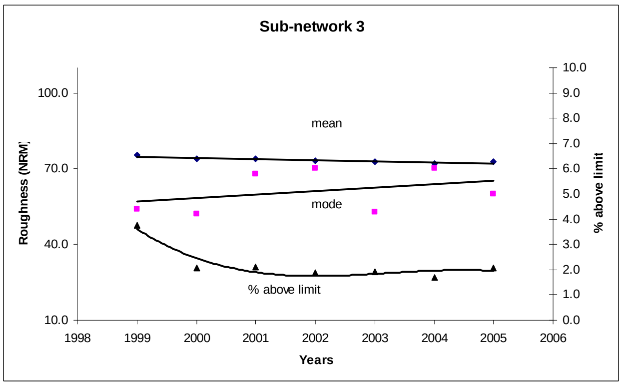

Figure 3.13 Roughness – Sub-network 3

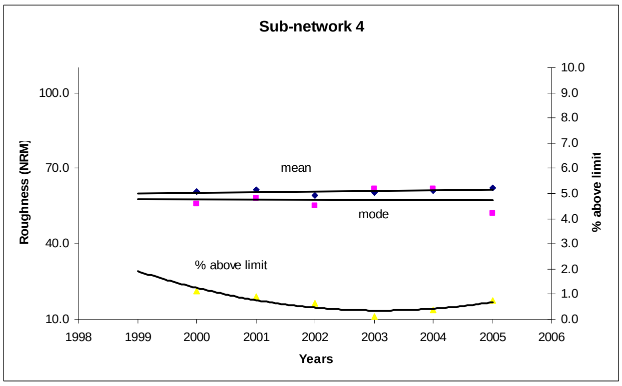

Figure 3.14 Roughness – Sub-network 4

The percent above limit may be derived by fitting or assuming a distribution for the data. This method is less accurate than deriving the result directly from the data, but it is simpler to use and a high level of repeatability can be achieved. The results will depend on the assumed distribution. Typically, the log-normal distribution yielded higher values beyond the target limit than either the normal distribution or the direct calculation. This difference is mainly attributed to the approximate nature of the mathematical representation of the log-normal distribution. It is postulated that the accuracy can be significantly improved by a more rigorous modelling of the distribution of the data.

The actual trends can be more clearly defined when the percent exceeding limit is normalised for the first year of the observation period. While all sub-networks are in a better shape after the first year, sub-network 4 shows some deterioration. The substantial relative changes also indicate that the worst sections were clearly targeted for treatment on all networks.

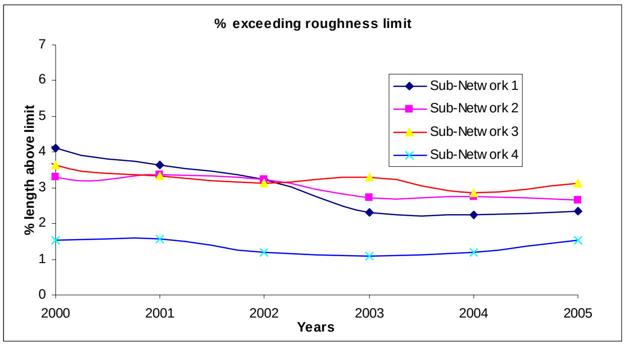

Figure 3.15 Percent exceeding limit.

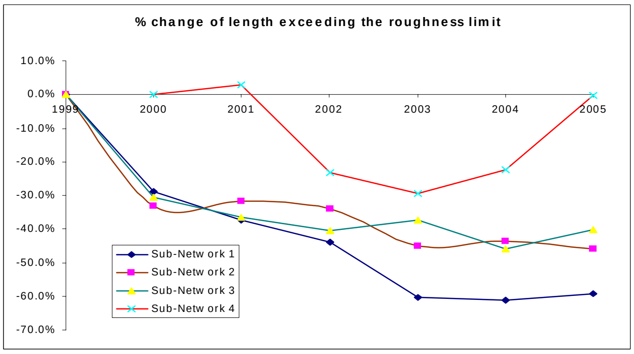

Figure 3.16 Percent exceeding limit relative to first data year.

3.3.2 Roughness-related road works

Roughness-related road works are defined as those that have a direct and immediate impact on roughness. These work types include all structural work, dense- or open-graded asphalt surfacings, slurry and stone mastic asphalt (SMA) wearing courses. Though these affect the achieved roughness improvement to different degrees, all of them reduce roughness. These work items were considered together cumulatively from the first year of the analysis. The work impact is presented here in terms of relative roughness change per unit treated road length (Figure 3.17). The cumulative work, expressed as length is shown on the x-axis and the change of normalised roughness is shown on the y-axis. The normalised change of roughness when plotted against the cumulative work length shows a remarkably close relationship. However, the detailed exploration of this relationship was left for later work. The trends may be used to express the effectiveness of the work done on the network as they reflect the improvement (change) per treated unit length. When considering the change of mean together with the change of percent in excess of the limit, the nature of the work on the network is highlighted. Figure 3.18 depicts similar trends for the percent above limit to that of the normalised means (Figure 3.17) with the exception of sub-network 4. This would indicate that most likely the majority of work was on the worst sections. However, the total volume of work was not sufficient to achieve long-term change. The significant reduction of the mode in the last year contradicts the increased roughness and may indicate data reliability problems.

Network data was explored to find any correlation between the work conducted on the network and the resulting network condition. The explorative work was limited to roughness and rutting on a full network and on four sub-networks. Based on the limited analysis the following issues were raised and explored.

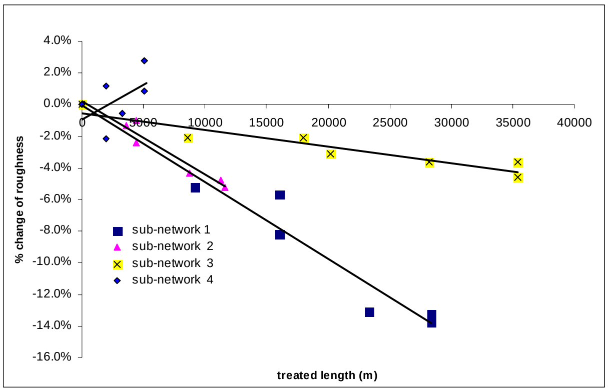

Figure 3.17 Normalised change of mean roughness vs cumulative work.

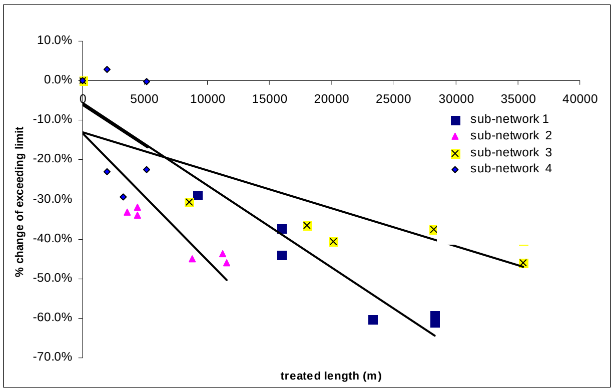

Figure 3.18 Normalised change of percent in excess of limit vs cumulative work.

3.4 Representation of data

The data distribution is asymmetrical and skewed in a manner showing a longer tail towards the higher values (worse condition). Consequently any modelling of the performance that assumes symmetrical (normal) distribution will be inherently inaccurate and may significantly under- or over-estimate the actual condition. Log-normal distribution was introduced to represent the data. The approximate formula that was developed for the purpose of current work proved that this method could produce consistent and more realistic results. By modelling the distribution with any function (such as that of the log-normal distribution) would allow an easy representation of the full distribution instead of using just one or more points of the distribution.

Recommendations:

- Represent data by log-normal distribution in the future.
- Refine the method of calculating log-normal distribution presented here to achieve increased accuracy.
- Calculate the percent exceeding the maximum limit by assuming log normal distribution.

3.5 Characterisation of the data population

So far only average and percentile have been used to define KPMs, ie to characterise network performance. The current work indicated that:

- average is insensitive and can be misleading

- mode may be more indicative of network condition when the data distribution is asymmetrical
- the shape of the cumulative distribution function (CDF) is influenced by the bin size
- the overall condition and particularly long-term trend may be more accurately characterised by the joint use of the mean, mode and percentile when the data distribution is not normal
- when using the mean and mode together, data inconsistencies and bias towards treating the worst sections first can be detected.

Recommendations:

- Explore the use of mode using more data sets and networks.
- Standardise the bin size for each data type for the purpose of calculating cumulative distributions.
- Explore the use of mode to characterise network condition.

3.6 Optimisation and modelling

Optimisation with dTIMS is currently focused on targeting the 'best' mean of selected parameters. Experience showed that in many cases selecting a worst-first strategy will yield similar results to the optimisation of an average. The work presented here supports this observation by highlighting the limitations and inadequacies of using the mean. In reality, the objective of optimisation is to achieve the best overall network condition. As the overall network condition seems to be better represented by the mode, this report proposes investigating the benefits of optimising the mode instead of the average of the network.

Recommendation:

- Investigate using the mode instead of the mean as an optimisation target.

3.7 Work impact

The analysis indicated a strong relationship between the cumulative work and the achieved normalised change in roughness. The methodology developed for the current work magnifies changes and allows direct comparison of various parameters and networks. It is postulated that the methodology is generally applicable and can be used as a common platform for all KPMs. In this case, it may also be utilised in the optimisation of work programmes.

Recommendations:

- Explore further the relationship between cumulative work and normalised change as well as its general applicability.

- Explore the usefulness of this relationship for modelling and network performance evaluation.

3.8 Effectiveness of KPMs

In the current work roughness and rutting KPMs were included. For most of the observed periods the roughness and rutting KPM remained below the target levels without the need for any significant remedial action. Based on the analysed data it is postulated that neither roughness nor rutting KPMs exerted controlling influence on the final work programme. However, efforts to treat the worst sections were detected and thus the benefits of an optimised approach might be compromised.

The analysis demonstrated that the network is poorly characterised when the mean is used without considering the nature of the distribution. The effectiveness of a KPM is a function of its capability to characterise the network condition. If this capability is limited, so will be the effectiveness of the KPM.

Recommendation:

- Analyse key KPMs for more networks in the manner presented here.

4. Summary and conclusions

The intended meaning of the KPMs was reviewed and their mathematical definition was interpreted by using real-life examples from two PSMC networks. The analysis was repeated for network and sub-network level roughness data.

It was found that the explored data was significantly skewed and would far more closely approximate a log-normal distribution than the usual normal distribution (Figure 4.1). A log-normal curve was fitted to the data and was used to estimate the percent exceeding the allowed maximum limits. As expected, the percent of length exceeding the limits was larger when calculated by assuming log-normal distribution than by calculating from a normal distribution. Normal distribution tends to underestimate the percent exceeding the limits.

Following from the analysis of the data distribution, the mean of the data proved to be a poor representation of the population. As the mean coincides with the bulk of the data only when the data is normally distributed, in the investigated cases the mean did not represent the majority of the data. The central tendency of the data (ie the bulk of the data) was better represented by the mode (Figure 4.1).

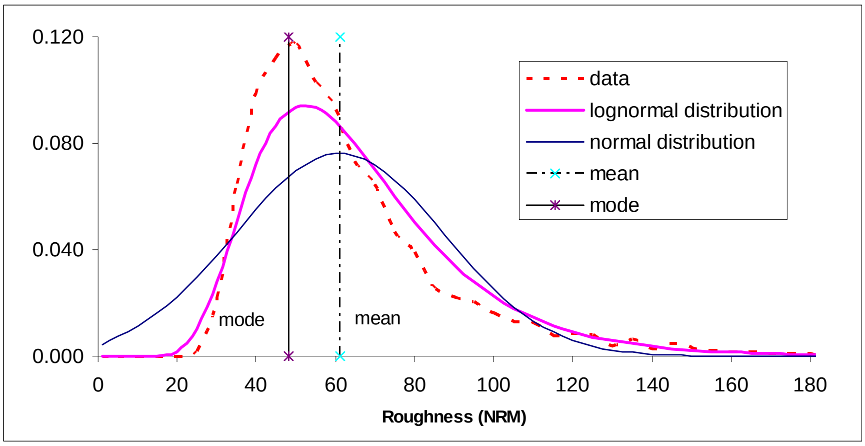

Figure 4.1 Comparison of the data distribution and its representation with normal and log-normal distribution.

When presenting the mean, mode and percent exceeding the limit together, a more realistic interpretation of the network condition became possible. A conflicting long-term trend was observed between mean and mode. In most cases the network or sub-network average remained static or displayed negligible changes while the mode changed significantly. This indicated that most of the network deteriorated, while the average was kept stable by treating the worst sections, as indicated by the decreasing percentage exceeding the limit (Figure 4.2).

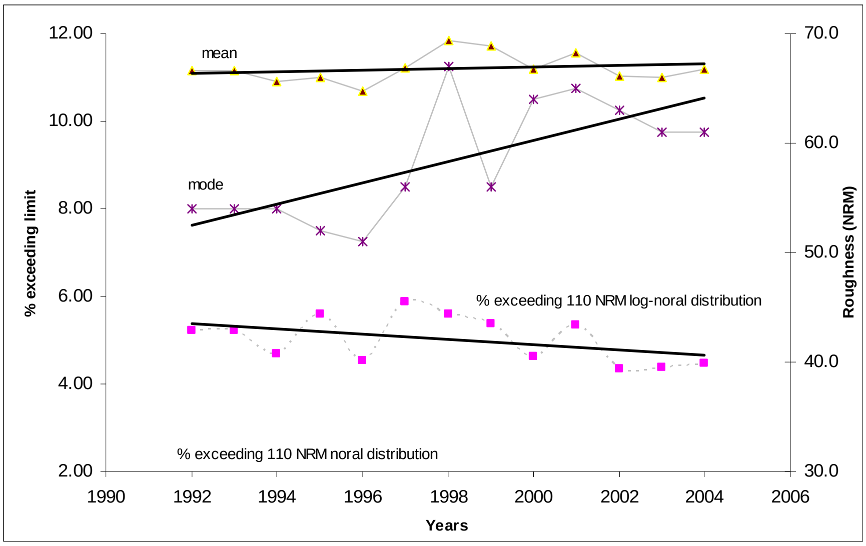

Figure 4.2 Typical trends of mean, mode and percent exceeding the limit.

The work to date yielded a method for better representation of the network condition and also provided a potent tool for interpreting network managers' strategies for network owners. The substantial fluctuation of the mode over time was also indicative of data collection problems. The substantial change of mode while the mean remained more or less the same indicated that the reliability of the data varied significantly. This highlighted the need for tighter specifications for data collection or for an alternative method for data summary and interpretation.

The discrepancy between the mean and mode explains the observed closeness between the optimised and worst first strategies. As optimisation is typically targeted to achieving the best average network level of service, in most cases treating the worst sections is sufficient. However, this does not prevent the deterioration of the greater part of the network. Consequently, if the objective is to achieve the best network condition, and the network condition is best represented by the mode, the optimisation must target the mode as well as the average.

The impact of work on the KPMs was investigated by using roughness data. This analysis utilised the averages as current KPMs are defined in terms of average. The treatments affecting roughness were taken into account by cumulating the treated length from the beginning of the observation period. The total (cumulative) treated length was plotted against the relative (percent) changes of average roughness (Figure 4.3).

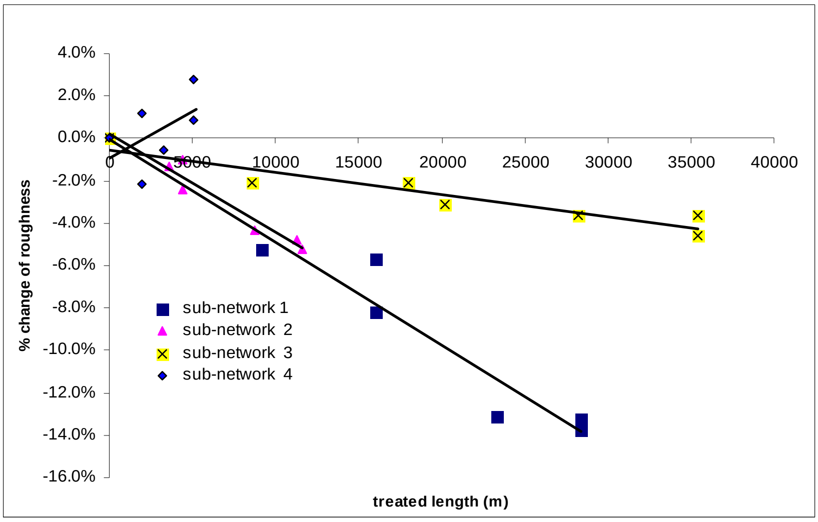

Figure 4.3 The impact of work on the relative change of average roughness.

The analysis was repeated for four sub-networks of the same PSMC contract. The results indicated a close relationship between the cumulative work and the change in roughness. However, the extent of the KPMs' control over the work type and quantity could not be established, as the KPMs never went below the target levels. The linear and close nature of the relationship was unexpected but could be explained in the light of the discussion on the inter-relationship between averages (ie KPMs) and targeting the worst sections first. This analysis may yield different results when using the mode instead of the average.

For the purpose of the above analysis the hierarchy of the treatments was taken into account. The hierarchy of treatments reflects the fact that many treatments affect more than one KPM, eg an asphalt overlay affects the age, roughness, and texture and strength performance indicators. Consequently, these KPMs must be determined by taking these inter-relationships into account. The investigation of this issue fell outside the scope of the current work, but attention must be paid to this in future considerations.

The explored issues may all affect the contracting parties and may cloud the successful completion and achievements of a PSMC contract. In extreme situations the PSMC contracts may become unenforceable as the KPMs may be deemed unreasonable. Consequently, there is every reason to resolve the above issues to the satisfaction of all parties involved.

In summary, based on the work to date, it may be concluded that:

- KPMs defined in terms of average, assuming normal distribution, incorrectly represent the network condition
- log-normal distribution is closer to the true data distribution and should be used instead of normal distribution
- current KPMs, based on averages, discourage true network level optimisation by implicitly favouring 'worst first' policies

- current modelling techniques use the average. By using the mode in modelling, the accuracy of the forecast may be improved
- the precision level of current data collection methods is most likely to be below the level perceived or stipulated by the KPMs implemented in the current maintenance contracts.

4.1 Recommendations

- Document the calculation method of KPMs in detail.
- Document and standardise the treatment of outliers and data flags across all networks and contracts.
- Include the definition of contract areas and sub-networks in the RAMM database.
- Consider the distribution of the data when selecting statistical parameters to describe the data.
- Refine the method of modelling log-normal distribution presented here to achieve increased accuracy.
- Include the edge of the distribution (percent exceeding limit or percentile) in the KPM definition.
- Consider the mode for characterising network condition, subject to further investigation.
- Explore further the relationship between cumulative work and normalised change as well as its general applicability.
- Explore the usefulness of this relationship for modelling and network performance evaluation.
- Analyse key KPMs for more networks in the manner presented here.

The key recommendations arising from this study are:

- a review of the current definition of KPMs, and
- the replacement of the currently used data average with the data mode.

The introduction of a synchronised set of KPMs based on more representative definition will greatly increase the contractors’ ability to influence the outcome and concurrently increase the owners’ confidence in the outcome of the contracts. In the next phase of this work, the investigation will include other properties, such as texture depth as well as other networks and sub-networks besides refining the methodology to suit the reporting and management requirements of network maintenance contracts.

5. References

Henning, T. F. P., Costello, S. B., Dunn, R. C. M., Parkman, C. C., and Hart, G. 2004. The establishment of a long-term pavement performance study on the New Zealand state highway network. *ARRB Journal* 13(2). Australian Road Research Board.

Kadar, P. 2004. The impact of routine maintenance on pavement performance – the missing link. In *Proceedings of the 6th International Conference on Managing Pavement.* Brisbane, Australia.

KPI Working Group. 2000. KPI report for the Minister of Construction. www.berr.gov.uk/files/file16441.pdf

Transit New Zealand. 2005. *Long-term procurement plan*. Transit New Zealand, Wellington.

World Bank. 2005. *Procurement of output- and performance-based road contracts (OPRC) and sample specifications.* Washington DC: World Bank.

Appendix 1

List of interviewed organisations

Code	Company	Name
1	Works Infrastructure	(Michael Haydon)
2	Transit NZ	(Kevin Locke)
3	BECA	(John Hallett)
4	Land Transport NZ	(Gerard Van Blerk)