Transition from visual condition rating of cracking, shoving and ravelling to automatic data collection

May 2017

TFP Henning, University of Auckland

GJ Morrow, Opus International Consultants Ltd

NZ Transport Agency research report 617

Contracted research organisation –Auckland Uniservices Ltd

ISBN 978-1-98-851227-3 (electronic)
ISSN 1173-3764 (electronic)

NZ Transport Agency
Private Bag 6995, Wellington 6141, New Zealand
Telephone 64 4 894 5400; facsimile 64 4 894 6100
research@nzta.govt.nz
www.nzta.govt.nz

Henning, TFP and GJ Morrow (2017) Transition from visual condition rating of cracking, shoving and ravelling to automatic data collection. *NZ Transport Agency research report 617*. 61pp.

Auckland Uniservces Ltd was contracted by the NZ Transport Agency in 2015 to carry out this research.

Keywords: automated data collection, chip loss, crack detection, data quality, ravelling, scanning laser, shoving

An important note for the reader

The NZ Transport Agency is a Crown entity established under the Land Transport Management Act 2003. The objective of the Agency is to undertake its functions in a way that contributes to an efficient, effective and safe land transport system in the public interest. Each year, the NZ Transport Agency funds innovative and relevant research that contributes to this objective.

The views expressed in research reports are the outcomes of the independent research, and should not be regarded as being the opinion or responsibility of the NZ Transport Agency. The material contained in the reports should not be construed in any way as policy adopted by the NZ Transport Agency or indeed any agency of the NZ Government. The reports may, however, be used by NZ Government agencies as a reference in the development of policy.

While research reports are believed to be correct at the time of their preparation, the NZ Transport Agency and agents involved in their preparation and publication do not accept any liability for use of the research. People using the research, whether directly or indirectly, should apply and rely on their own skill and judgement. They should not rely on the contents of the research reports in isolation from other sources of advice and information. If necessary, they should seek appropriate legal or other expert advice.

Acknowledgements

The research team wishes to thank both Pavemetrics and Data Collection Limited, New Zealand for their contribution of data and information. The NZ Transport Agency and Whangarei District Council are also acknowledged for the data used in this project. Steering group and peer reviewers assisting with the project were:

Steering group

John MacDonald, NZ Transport Agency project manager

Dave Robertson, NZ Transport Agency

Scott Verevis, Whangarei District Council

Angela Wang, Auckland Transport

Peer reviewers

Dr Fritz Jooste, Juno Services Ltd.

Neil Beckett, NZ Transport Agency

Abbreviations and acronyms

COLTO	Committee of Land Transport Officials (South Africa)
CSIR	Council for Scientific and Industrial Research
FWHA	Federal Highway Agency (US)
GEIPOT	Brazilian Transport Planning Agency
HSD	high-speed data
LCMS	laser crack measurement system
LTPP	long-term pavement performance
MTO	Ministry of Transport Ontario, Canada
QA	quality assurance
OAG	Office of the Auditor General
RAMM	Road Asset Maintenance Management
RIMS	Road Information Management Group
SCI	surface condition index
Transport Agency	New Zealand Transport Agency
TRB	Transportation Research Board

Contents

Executive summary

The need for automated defect detection

The manual road condition survey method used in New Zealand (road asset maintenance management ((RAMM)) surveys) was developed during the early 1980s with the primary purpose of feeding into the treatment selection algorithm. For more than 20 years the rating system was adequate for this purpose but as more sophisticated asset management evolved into deterioration modelling and advanced trend monitoring, the data quality from the manual surveys came under scrutiny. Attempts to improve the robustness of the rating system included increasing the recommended sampling size from 10% to 20% of the treatment length plus increasing the requirements for accreditation during the training of raters. Yet, these steps still fall short in increasing the overall usability and repeatability of rated data for the new demands of asset management processes.

Automated defect data collection has been undertaken since the mid-1990s with early technology relying on photographic imaging and processing of road surface data. The technology was particularly popular for application on busy asphalt and concrete motorways in the northern hemisphere but failed to deliver acceptable robustness on chipseal surfaces. This situation changed with the arrival of laser scanning technology, which has overcome the limitations of photo-imaging technology. The measurements now solely depend on laser scanning at a high resolution, which gives a comprehensive 3D image of the road profile. Any defects such as cracks, potholes or surface defects can be identified on the image. The benefits this technology offers to the sector include:

- surveys of 100% of the road are possible
- all aspects of the condition of the surface are captured simultaneously
- the measurements take place at high speed (60 to 80km/h), providing significant safety and traffic management benefits
- ‘removing’ the human element from the measuring allows for more repeat measurements.

Despite the accuracy of the measurement, the constraining factor for the technology is the algorithms that interpret the digital image to identify and quantify specific defects. This has resulted in the main question posed for this project – is the measurement sufficiently robust and is the sector ready to adopt this technology on a wide scale?

This research project

The project’s aim was to focus mainly on the impact assessment with the assumption that the three defect types (cracking, ravelling and shoving) are being accurately collected using the laser scanners. A number of international research projects have confirmed the accuracy of the measurements. During the research, more work was required to validate the results and introduce a new algorithm for the detection of shoving.

The ultimate focus of the research changed slightly to how ready is the technology in its current state for wider adoption in New Zealand.

Main findings from this research project

The research findings were consistent with most international research projects that have studied the accuracy and repeatability of the laser technology. These findings were:

- The laser technology's measurements are accurate and can be repeated – the benefit of having a 100% road length covered by the surveys is particularly attractive for the intended data use.
- The comparison of the laser technology with existing practices remains a challenge and the results from such comparisons should be analysed with care. The comparison between laser scanning and RAMM surveys will never yield ideal results because:
  - The two assessment methods differ fundamentally in the way they define the extent of the defect – one-to-one comparisons are therefore not possible.
  - RAMM surveys only cover a percentile of inspection lengths and a sample outcome will most likely differ from the full-length survey.
  - Ensuring the two assessment types reference the exact same location is difficult and there is a lack of confidence that comparisons are being made between the same road sections.
- The laser technology has identified defects successfully but has also identified a number of false positives. A more detailed investigation during this research into the shoving measurements has identified a number of road features that appear to trigger shoving according to the defined algorithm, but in reality identify a completely different road feature as a shove. The study has also confirmed a number of instances where the rating simply 'missed the shove' as it was not very apparent for a number of reasons.

The implementation plan

- Using laser scanning for detecting road defects should be adopted by all road agencies. This recommendation is made on the basis of the significant benefits that can be realised from:
- more accurate assessment
- better repeatability between surveys from consecutive years
- greater coverage of the road network, ie more roads are being surveyed for 100% of the length.

The laser technology, despite its accuracy, cannot be applied as a 100% automated process. The computer algorithms that analyse the data still need significant 'learning' that can only be achieved if the technology is supplemented by manual validation of the outcome. Someone needs to work through the digital images to find erroneous identifications and feed this knowledge back to the algorithms. Once this is completed, business as normal survey contracts should include calibration procedures, validation and quality assurance protocols.

Further considerations for industry's adoption of laser scanning are summarised in table ES.1.

Table ES.1 Industry impact and adoption

Item	Impact	Further work	Most suitable party to undertake
Defect definition	A new definition of 'defect' needs to be universally accepted by the industry. Sticking to the current definition and quantification will devalue the enhancements from laser scanning.	Incorporate into national condition assessment guidelines.	RIMS/NZ Transport Agency (Include the RIMS Condition Guideline) (RIMS 2016)
Shoving data	The field survey suggested shoving data could be more accurate.	Additional calibration of the algorithm is needed.	NZ Transport Agency (research to be driven by surfacing group)

Item	Impact	Further work	Most suitable party to undertake
Ravelling measurements	Ravelling is an important surfacing defect that has to be included in the surveys, once it provides sufficiently accurate results.	A study is required to calibrate the ravelling identification algorithm to different chipseals in New Zealand.	NZ Transport Agency (research to be driven by surfacing group)
Data standards	New data standards are recommended and should universally be accepted and incorporated in best practice guidelines.	Gain consensus from the industry and incorporate in guidelines.	RIMS/NZ Transport Agency
Procurements	Quality assurance validation and calibration requirements need to be specified.	Develop processes in collaboration with suppliers. Then incorporate in template specifications.	RIMS/NZ Transport Agency

Abstract

Robust condition data feeding into asset management processes is a key step towards having confidence in long-term strategies for renewals and replacements. The manual condition rating system was originally developed as an input into the treatment selection algorithm; however, in later years the data has been used for pavement deterioration modelling and trend monitoring, which are outside the intended scope of the rating system. It was therefore not unexpected that both field inspectors and researchers highlighted shortcomings in the quality and repeatability of manually recorded data. Automated scanning technologies promise to overcome many of the issues associated with manual condition data collection. However, before a wide-spread adoption of the scanning technology is possible, research had to prove the accuracy of the measurements and determine the impact of new data items in the asset management processes. This research addressed both these items and has concluded the technology is ready for adoption in New Zealand. However, fully automated surveys yield less than desirable accuracy with a high portion of false negatives identified. All scanning surveys must be supplemented by appropriate manual quality assurance processes.

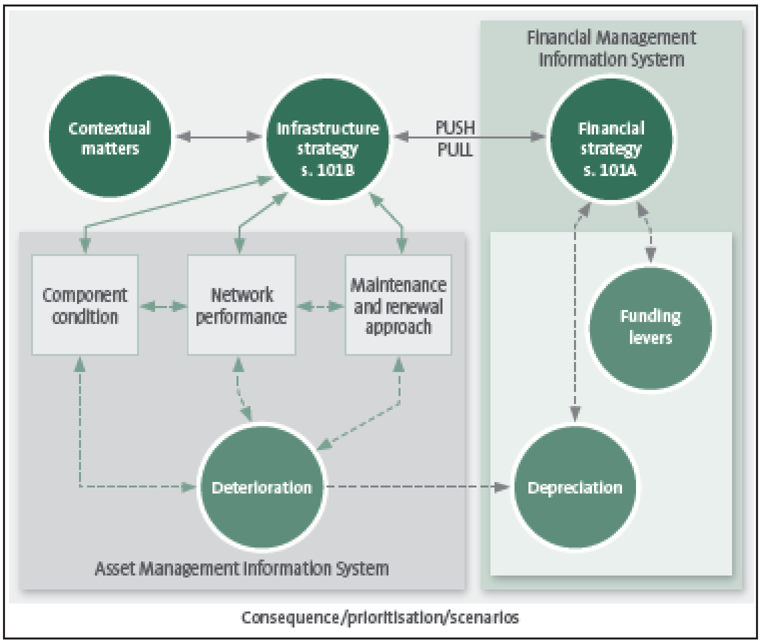
1 Introduction

1.1 Background

A recent report by the Office of the Auditor General (OAG 2014) reviewed the status of infrastructure management practices. While the roading industry was commended for its official and consistent data collection processes, the document also signalled the need for an overall improvement in infrastructure management in New Zealand. The review highlighted the lack of good practices around strategic planning as noted in the asset management plans. Robust strategic asset management is characterised by the following (refer to figure 1.1):

- a thorough understanding of the current and historical performance of the network, underpinned by robust data collection techniques and practices
- a strong understanding of the cost of maintenance on the network and the long-term impact of each maintenance option, underpinned by robust historical maintenance cost information
- an understanding of future performance, in particular potential risk areas needing attention.

Figure 1.1 Aspects of asset management that contribute to good infrastructure strategy (OAG 2014)



The above points emphasise the importance of robust data collection techniques and processes. This research investigated the merits of transitioning from visual condition rating of cracking, shoving and ravelling to automatic data collection.

1.2 Problem statement

Automated data collection is not a new technology. During the late 1970s and early 1980s, the World Bank documented road condition automated surveys for assessing the roughness and road profile during the HDM-III study completed in Brazil (GEIPOT 1981). The sophistication and robustness of automated data collection has improved significantly since the 1980s. In New Zealand, automated data collection technologies form an integral part of the road asset management work of the NZ Transport Agency (‘the

Transport Agency'), which has been collecting roughness, rutting, texture and sideway-force coefficient routine investigation machine (SCRIM) on state highways for almost 20 years. The one aspect that has remained a manual process is the identification of surface defects. The technology for identifying these defects has been in development for some time but given its limitations was only applied to concrete and asphalt pavements. The rough texture of chipseal surfaces resulted in the technology not being applicable initially for this type of surfacing. With New Zealand's surfaced roads consisting of approximately 80% chipseal, it is important to identify early signs of a surface defect as it could quickly progress into a more serious defect given that most of these pavements consist of bound and/or unbound granular layers. It is only recently the automated technology has developed to such a level it can be considered for full adoption in New Zealand. There are, however, a number of questions that need to be answered before adopting the automated defect surveys, including:

1. Is the technology completely ready for full adoption in New Zealand? This is a question that not only refers to the technical capabilities, but also the impact such a transition would have on the industry.
2. Would a transition be affordable for the country and would it truly be a cost-effective change that would result in substantial benefits from the shift?
3. What would be the wider impact on planning processes that use surface condition data? This question not only refers to the change one would expect from the surveys but also to a potentially different sampling regime compared with that of the current manual surveys.

1.3 Objectives and scope

The over-arching purpose of this research was to assist New Zealand road authorities make the transition from a visual condition rating of cracking, shoving and ravelling to an automated data collection regime. This purpose will be achieved through addressing the following key objectives:

- Confirm the automated technology is sufficiently robust to replace the current visual rating. This objective can be seen as a 'proof of concept' stage of the project. A large part of this objective is to understand the technology, its limitations and practical survey aspects that may impact on the implementation of a full-scale roll-out.
- Understand the impact of changing to a fully automated condition assessment. There will be a two-fold impact as a result of adopting a fully automated data collection system. First, there is an expectation of more comprehensive reporting of defects such as cracking; plus, the technology will allow for a 100% sampling regime where currently a 10% to 20% rating regime is used in most cases. Switching to a fully automated survey regime will have an impact on:
  - all reporting processes and trend monitoring
  - all decision processes such as algorithms (eg RAMM treatment selection algorithm) and decision support processes (eg dTIMS).
- Lastly, a switch to automated data collection will have a significant practical impact on the data items, the method of quantification of the defects, the summarisation interval (an average value for 20m intervals) of the databases and reporting lengths. This research aimed to define the most practical data specifications and the most robust manner of storing and analysing the required information.

The scope of this project included:

- proof of concept – confirm the technology will work satisfactorily for New Zealand roads
- know how the change will impact on the industry in relation to reporting and decision making

- lastly and perhaps most importantly for the Transport Agency, to have a road map for adopting this technology throughout the country.

The most likely automated data collection method considered for this study was the laser crack measurement system (LCMS). Earlier research showed this to be the best technology, capable of measuring cracks in chipseal pavements with satisfactory robustness. It is also worth noting this technology is currently available in New Zealand with two operators capable of conducting the LCMS surveys.

2 Literature review

2.1 Purpose and scope of the literature review

The main purpose of the literature review was to investigate previous research into the capability of the technology. For the purpose of this research, the 'how effective' question was of greater importance than the 'how does it work' question. As an outcome of the literature review, further testing/analysis of local-based data needed to be specified in order to establish local acceptance of the technology. A further objective for the literature review was to understand the limitations of each application.

The scope of the literature review included all international and local research that has looked at the robustness and accuracy of automated defect identification. Emphasis was placed on third party research that did not have a direct interest in a particular technology.

2.2 A summary of defect identification techniques

2.2.1 Manual and visual assessment techniques

The earliest evidence of visual assessments goes back as far as the original American Association of State Highway Officials trials conducted since 1920 with the development of the present serviceability index. This index was mainly a function of roughness and rutting, with a panel of motorists giving these properties a qualitative score pending their ride experience (TRB 2007). Since then a number of variations have been developed from this approach mostly involving manual assessments of road pavements, recording all defects on assessment forms and post-processing into composite indices and decision algorithms for maintenance planning.

The visual and manual assessments can be classified into three broad categories:

- Windshield surveys – assessors undertake a 100% assessment of road condition travelling at slow speeds. Assessment forms are completed for individual treatment lengths. An example of this approach is described in TMH-9 (CSIR 1992).
- Walk-over surveys – assessors walk over an assessment length recording all defects on an assessment form. The assessment length differs from agency to agency, some surveying 100% of the road network and others only sampling a portion according to a given percentage, say 10% to 20%.
- Detailed manual surveys on a small scale are also undertaken for the purposes of rehabilitation and/or research. Some examples of these are documented in the South African rehabilitation design guide TRH-12 (COLTO 1997), and the Strategic Highway Research Program (SHRP) assessment manual for the long-term pavement performance (LTPP) sections (FHWA 2003). The New Zealand LTPP programme data collection is another example (see section 3.3 for more details).

2.2.2 Photographic based methods

There are a number of photographic-based defect identification methods. The main driver in developing these techniques stemmed from the limitation of visual assessment on busy motorway sections where any visual method was simply impractical given traffic safety considerations. Broadly the photographic methods include:

- Still cameras that record an area of pavements. The photo is then analysed using:
  - road raters picking up defects from the photos

  - automated pixel analysis of the photographs translated into specific defect identification and quantification.

- Video surveys of road sections using the same interpretation techniques as the still cameras.
- Line scanning cameras, which in principle work the same as earlier versions but are able to yield better resolution on the photos, thus enabling more robust automated pixel analysis.

Figures 2.1 and 2.2 show photographs of digital data collection techniques. One of the known limitations of these techniques is the strong sensitivity to the lighting of the photographed area. This is a particular issue for chipseals where the texture creates shadows that are interpreted as defects when an automated pixel processing is used.

Figure 2.1 Digital data collection vehicle (Wang 2004)

Source: Wang (2004)

Figure 2.2 Line scanning camera – ROADCRACK (Source ARRB)

2.2.3 Laser scanning techniques

LCMS consists of high-power scanning lasers combined with a camera that records the ‘laser line’ on the pavement surface (refer to figure 2.3). The image is then processed to yield the profile of the road as depicted in figure 2.4.

Figure 2.3 3D laser scanning (Laurent et al 2011)

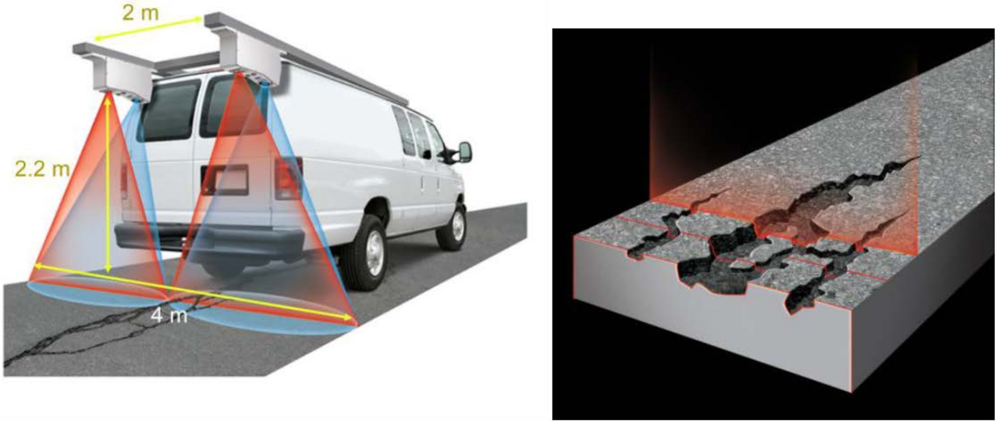
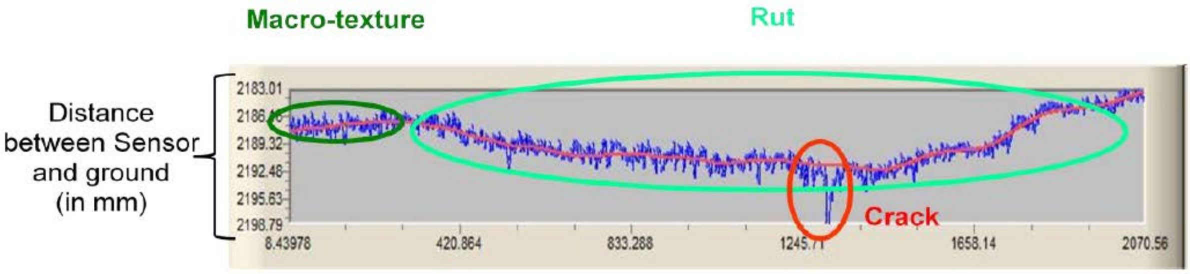
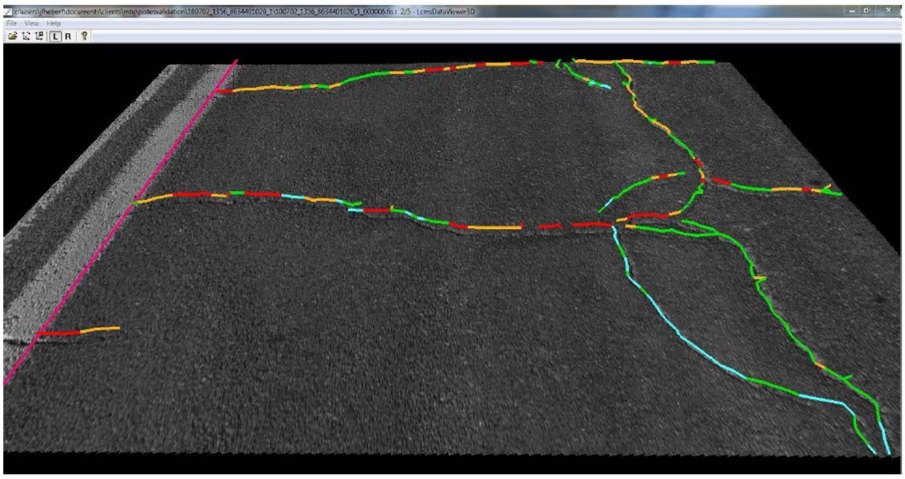


Figure 2.4 Measurements from the LCMS (Laurent et al 2011)



An output from the system yields identified crack pattern images as illustrated in figure 2.5. Note that the crack severity is also recorded as a function of the crack width. Suppliers claim this system is able to detect cracking to a width of 0.5mm.

Figure 2.5 Crack severity analysis (Laurent et al 2011)



Although the system was primarily developed for crack detection the scanned images help identify many other defects such as roughness, rutting, potholes and ravelling, as shown in figure 2.6. For example, figure 2.7 illustrates how the LCMS data could be used for identifying bleeding.

Figure 2.6 Defects identified with LCMS (Source Pavemetrics)

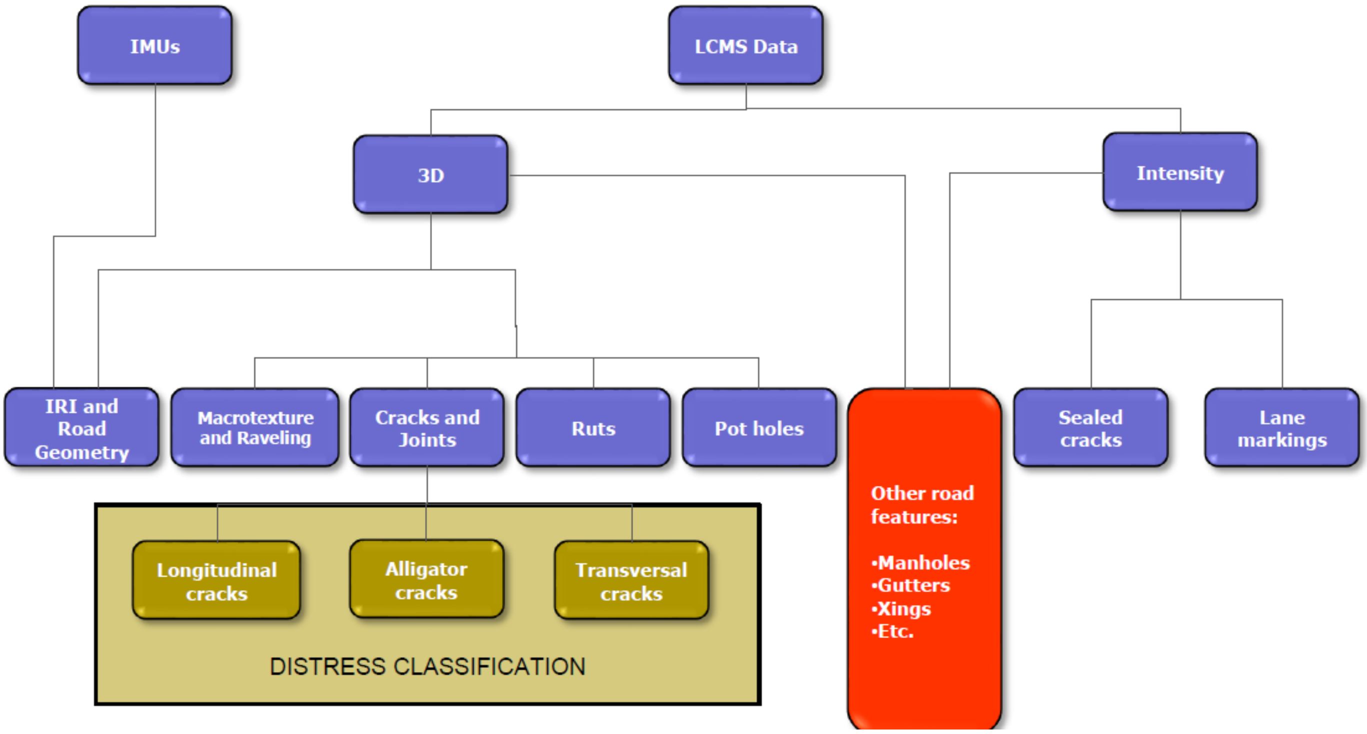
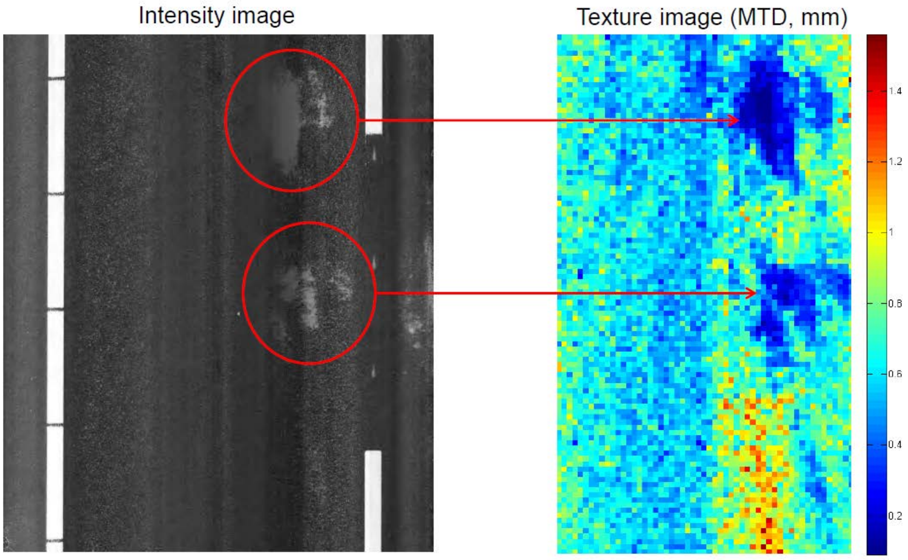


Figure 2.7 Using the LCMS to identify bleeding (Source Pavemetrics)



2.3 Research and testing of automated defect identification

2.3.1 Crack detection

Cracking is one of the most important defects engineers monitor on bituminous surface roads. For asphalt roads, cracking is an indication of layer failure since cracking is one of the design parameters of asphalt surfaces (Austroads 2012). It is also important to monitor cracking on chipsealed roads as, even though it may not indicate a failure per se, it does compromise the water proofing of the seal, letting in water that could cause significant pavement damage. Authorities in New Zealand resurface between 7% and 11% of their road network, and it is estimated that at least 40% of the resurfacing has cracking as a primary driver for the intervention. Knowing that cracking is important to monitor naturally leads to authorities spending considerable amounts of money to undertake road condition measurements and visual surveys. Historically, condition data was solely used for maintenance planning using a decision algorithm. Lately, the data has been used for performance monitoring, benchmarking and performance modelling using systems such as dTIMS (Tapper et al 2013). This increase in the use of data has also resulted in increased demand for more robust, complete and more frequent data.

Given that automated crack detection was the stimulus for the development of the LCMS, it was also expected there would be research and a number of tests confirming the robustness of the technology. Testing of auto-detection methods for other defects such as ravelling and shoving is lacking behind in terms of research into their effectiveness of these measurements. Table 2.1 summarises some of the testing and research into the robustness of the LCMS for crack detection. Note that the effectiveness of the digital technologies was not investigated further as they were proven to be less effective on chipseals, thus questioning their suitability for New Zealand (Wix and Leschinski 2012).

Table 2.1 Research into the robustness of LCMS crack identification and quantification

Research	Reference	Testing	Main findings
FHWA (USA) – Field evaluation of automated distress measuring equipment	Serigos et al (2014)	• A field experiment consisting of 20 sections was developed. • Static manual distress statistics, texture, cross slopes and digital crack maps were collected. • Four vendors were invited to collect automated distress, texture and cross slope measurements at highway speeds. • The results were analysed and compared to assess the difference between automated and manual measurements and to evaluate the change in accuracy between fully and semi-automated results.	From the comparative analyses among the distress statistics reported by each participant and the manual raters, no clear, obvious patterns emerged for all types of distress and time frames. Thus, the researchers could not identify one automated system that was clearly superior to the other. This lack of clear patterns was in part due to the use of different distress classification criteria. It is recommended that an objective and programmable standard or protocol be developed for classifying distresses from automated data to increase the consistency of results. Dynatest and Fugro-Roadware showed a significant improvement in the accuracy of their distress measurements after applying manual post-processing consisting of visual interpretation and correction of the results produced by their systems' algorithms. Additionally, the results reported within four weeks included more types of distresses. These observations show the current need for applying manual interpretation to the automated results produced by state-of-the-art equipment. Dynatest and Fugro produced texture results close to the reference in magnitude with minor error. It is suggested that WayLink-OSU and TxDOT consider updating or calibrating their systems since all measurements presented were greater than the reference values. Note that TxDOT texture results were reported as an average value for each 550ft section, which is equivalent to the 0.10-mile subsection length used to store and calculate PMIS rating sections values. Revising the TxDOT algorithm to report values on a 50ft interval could have resulted in a different conclusion.
ARRB – Cracking – a tale of four systems	Wix and Leschinski (2012)	Comparative trials between different automated crack identification technologies. Testing on asphalt and chipseal pavements.	Both automated systems showed a high degree of repeatability on asphalt surfaces as well as good agreement between their respective cracking intensity results. However, on sprayed seal surfaces, there was a notable difference. While the repeatability of the LCMS on sprayed seal surfaces was good, the reported crack intensity suggested a significant number of false positives and a very poor overall correlation with RoadCrack.

Research	Reference	Testing	Main findings
New Zealand – Did we get what we wanted? Getting rid of manual condition surveys	Henning and Mia (2013)	The main objective of this research was to establish whether laser scanning crack detection methods could effectively identify cracking on chipseal surfaces. The further objective was to determine the effectiveness of crack detection on a larger scale compared with a visual rating that typically looked at either a 10% or 20% sample size.	The outcome of the research suggested the following: • There was a strong correlation between the LCMS and the LTPP cracking data. • The comparison with RAMM network survey data suggested more than 60% of crack lengths were missed according to the 10% sampling length used for the RAMM surveys.
Canada, Pavemetrics – Using 3D laser profiling sensors for the automated measurement of road surface conditions (ruts, macro-texture, ravelling, cracks)	Laurent et al (2011)	General introduction to a specific LCMS is presented along with some robustness testing undertaken in Québec (MTO), Canada.	The LCMS system was tested at the network level (10,000km) to evaluate its performance at automatic detection and classification of cracks. The system was evaluated to be over 95% correct in the general classification of cracks.
Ohio, DoT – PCR evaluation: considering transition from manual to semi-automated pavement distress collection and analysis	Vavrik et al (2013)	This report documents studies that considered the amount of manual processing required to reach an acceptable robustness of the automated equipment.	Automated data collection should be supplemented with manual verifications and QA processes.
TRL, UK – Use of high-resolution 3D surface data to monitor change over time on pavement surfaces	McRobbie et al (2015)	This report focuses primarily on accuracy as a result of location referencing.	Location referencing issues could be addressed through recognised techniques.

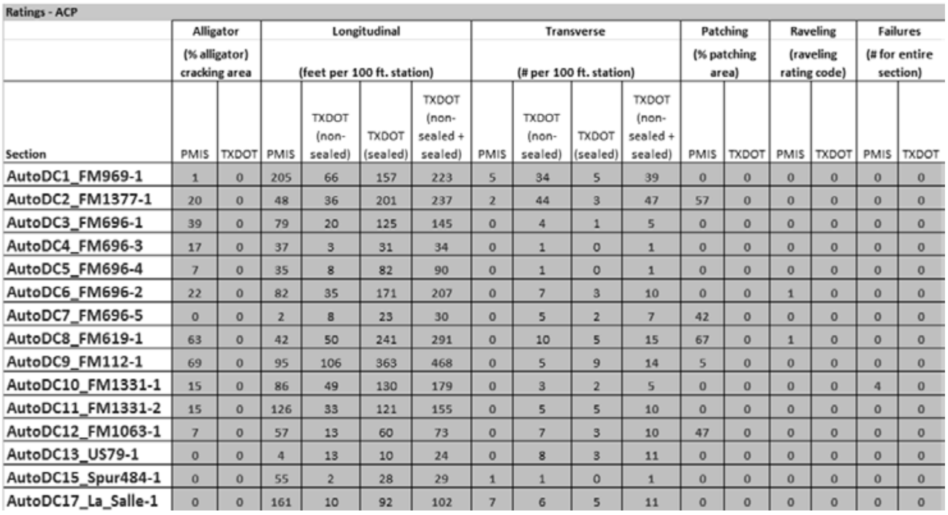
Note: Some examples of the study results are depicted in appendix A.

2.3.2 Other defects

It is fair to comment that the robustness testing of other defects has not been documented to the same extent as for cracking.

The FHWA study conducted by Serigos et al (2014) considered more than just cracking, and included patching, texture ravelling and edge break in their study. Tables 2.2 and 2.3 show some of the comparative results.

Table 2.2 Summary of PMIS manual rating and TxDOT 3D system distress data for hot mix asphalt sections (Serigos et al 2014)



Ratings - ACP

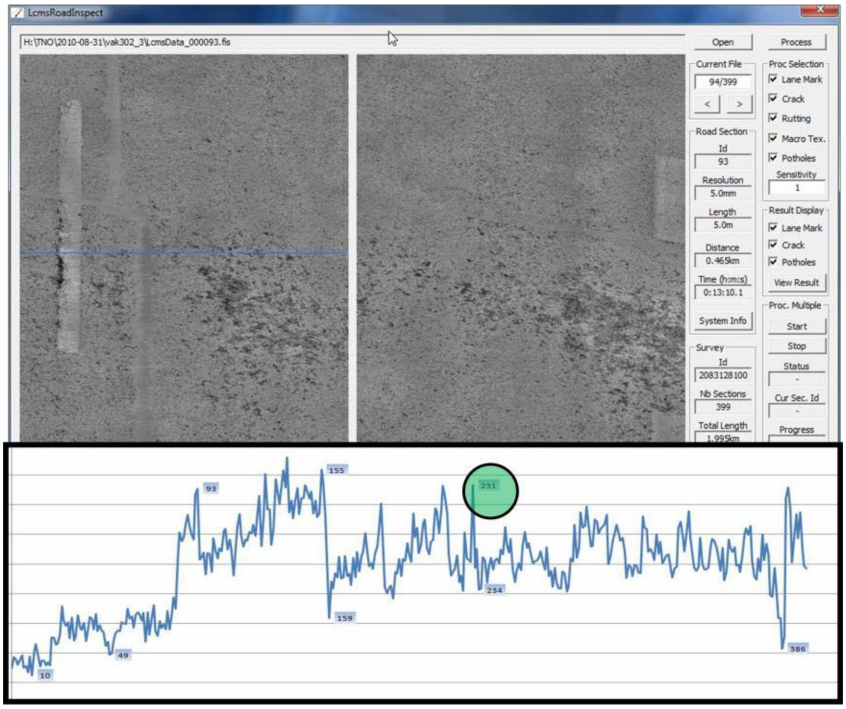
	Alligator (% alligator) cracking area		Longitudinal (feet per 100 ft. station)				Transverse (# per 100 ft. station)				Patching (% patching area)		Raveling (raveling rating code)		Failures (# for entire section)	
Section	PMIS	TXDOT	PMIS	TXDOT (non-sealed)	TXDOT (sealed)	TXDOT (non-sealed + sealed)	PMIS	TXDOT (non-sealed)	TXDOT (sealed)	TXDOT (non-sealed + sealed)	PMIS	TXDOT	PMIS	TXDOT	PMIS	TXDOT
AutoDC1_FM969-1	1	0	205	66	157	223	5	34	5	39	0	0	0	0	0	0
AutoDC2_FM1377-1	20	0	48	36	201	237	2	44	3	47	57	0	0	0	0	0
AutoDC3_FM696-1	39	0	79	20	125	145	0	4	1	5	0	0	0	0	0	0
AutoDC4_FM696-3	17	0	37	3	31	34	0	1	0	1	0	0	0	0	0	0
AutoDC5_FM696-4	7	0	35	8	82	90	0	1	0	1	0	0	0	0	0	0
AutoDC6_FM696-2	22	0	82	35	171	207	0	7	3	10	0	0	1	0	0	0
AutoDC7_FM696-5	0	0	2	8	23	30	0	5	2	7	42	0	0	0	0	0
AutoDC8_FM619-1	63	0	42	50	241	291	0	10	5	15	67	0	1	0	0	0
AutoDC9_FM112-1	69	0	95	106	363	468	0	5	9	14	5	0	0	0	0	0
AutoDC10_FM1331-1	15	0	86	49	130	179	0	3	2	5	0	0	0	0	4	0
AutoDC11_FM1331-2	15	0	126	33	121	155	0	5	5	10	0	0	0	0	0	0
AutoDC12_FM1063-1	7	0	57	13	60	73	0	7	3	10	47	0	0	0	0	0
AutoDC13_US79-1	0	0	4	13	10	24	0	8	3	11	0	0	0	0	0	0
AutoDC15_Spur484-1	0	0	55	2	28	29	1	1	0	1	0	0	0	0	0	0
AutoDC17_La_Salle-1	0	0	161	10	92	102	7	6	5	11	0	0	0	0	0	0

Table 2.3 Summary of texture measurement average errors for each section and vendor (Serigos et al 2014)

Section	*Inner* wheelpath (IWP) - MPD Average Error (mm)				*Outer* wheelpath (IWP) - MPD Average Error (mm)			
	Vendors			TxDOT	Vendors			TxDOT
	Dynatest	Fugro	Waylink-OSU		Dynatest	Fugro	Waylink-OSU	
AutoDC1_FM969-1	-	0.00	-1.55	-	0.02	-0.14	-1.73	-1.26
AutoDC2_FM1377-1	-	0.13	-1.13	-	0.47	0.38	-0.77	-0.79
AutoDC3_FM696-1	-	0.21	-1.45	-	0.45	0.47	-0.98	-3.31
AutoDC4_FM696-3	-	0.01	-1.18	-	0.04	0.06	-0.95	-2.56
AutoDC5_FM696-4	-	-0.01	-1.54	-	0.01	-0.02	-1.12	-1.54
AutoDC6_FM696-2	-	0.03	-1.67	-	0.04	0.09	-1.53	-0.48
AutoDC7_FM696-5	-	0.32	-0.98	-	0.36	0.37	-0.82	-1.76
AutoDC8_FM619-1	-	0.52	-0.98	-	0.73	0.75	-1.14	-2.38
AutoDC9_FM112-1	-	0.35	-0.95	-	0.19	0.37	-0.65	-2.09
AutoDC10_FM1331-1	-	0.33	-2.38	-	0.63	0.54	-9.51	-2.00
AutoDC11_FM1331-2	-	0.55	-1.32	-	0.45	0.56	-1.98	-1.94
AutoDC12_FM1063-1	-	0.62	-1.30	-	0.52	0.53	-1.08	-1.59
AutoDC13_US79-1	-	0.04	-3.16	-	0.50	0.12	-2.07	-2.75
AutoDC14_IH35-3	-	0.06	-0.98	-	-0.07	0.05	-0.61	-1.32
AutoDC15_Spur484-1	-	0.27	-1.86	-	0.22	0.28	-1.09	-2.02
AutoDC16_US77-1	-	0.28	-1.41	-	0.01	0.27	-0.43	-0.81
AutoDC17_La_Salle-1	-	-0.17	-1.32	-	0.10	-0.06	-1.05	0.48
AutoDC18_IH35-1	-	0.05	-1.38	-	0.04	0.23	-0.87	-1.22
AutoDC19_IH35-2	-	0.01	-1.13	-	-0.08	-0.01	-0.93	-1.15
AutoDC20_US84-1	-	0.01	-1.31	-	-0.12	-0.07	-1.14	-1.20
Average	**-**	**0.18**	**-1.45**	**-**	**0.23**	**0.24**	**-1.52**	**-1.59**

The paper authored by Laurent et al (2011) considered an array of defects, for example figure 2.8 illustrates the interpretation of laser image data to identify ravelling on a porous asphalt pavement. It is encouraging to note the seemingly accurate identification of this defect, especially on porous asphalts where ravelling is one of the main maintenance drivers.

Figure 2.8 Example of high ravelling index road section on porous asphalt roads in the Netherlands (Laurent et al 2011)



H:\TNO\2010-08-31\vak302_3\LcmsData_000093.fis

Current File: 94/399

Road Section Id: 93

Resolution: 5.0mm

Length: 5.0m

Distance: 0.465km

Time (h:m:s): 0:13:10.1

Survey Id: 2083128100

Nb Sections: 399

Total Length: 1.995km

Proc Selection: ☑ Lane Mark ☑ Crack ☑ Rutting ☑ Macro Tex. ☑ Potholes

Sensitivity: 1

Result Display: ☑ Lane Mark ☑ Crack ☑ Potholes

Status: -

Cur Sec. Id: -

2.4 The impact of switching to automated defect identification

A significant part of this research was focused on understanding the impact of switching to automated defect identification technology. It is expected this change of technology will have a significant impact on all existing algorithm processes and on the understanding of network performance. Some of the prior sections touched on research indicating the robustness of the respective measurements, which is directly relevant to the expected performance in New Zealand given that the same equipment is available here. However, the impacts such a change will have are unique to New Zealand conditions and have been fully addressed through this research. A few factors that may influence the magnitude of the impact in technology change are discussed in subsequent sections.

2.4.1 ‘Bias’ and errors in measurements

A simple illustration of the relationship between measurement bias and precision is illustrated in figure 2.9. It shows for example that measurements, typically associated with automated methods, are 100% precise, but still have a bias (refer to the top right image in figure 2.9). A good example of this is the Transport Agency’s high-speed data (HSD) collection process that has been in use for more than 17 years on the state highways. During 2008, the configuration on the transverse laser beam was altered to rectify an apparent bias (underestimation) of the rut depth measurements. This correction resulted in the entire state highway network suddenly having a deeper rut depth for the following years than originally expected. Figure 2.10 illustrates the rutting greater than 10mm on the state highways compared with one of the performance-based maintenance contract networks. The sudden increase in rutting following the 2009 survey is evident.

Figure 2.9 Bias and precision (Annis 2011)

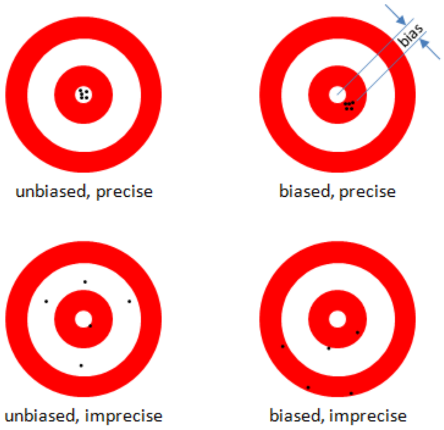
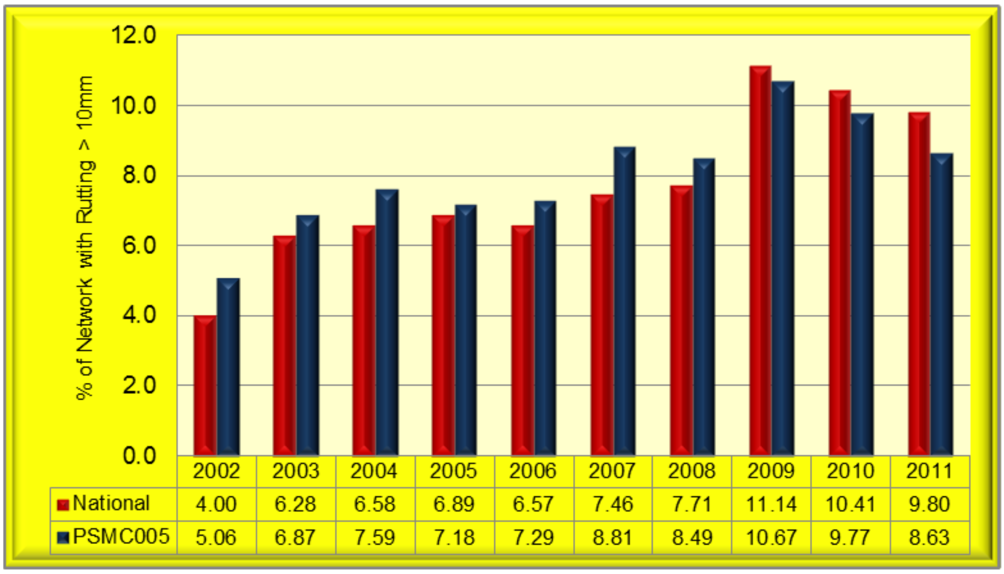


Figure 2.10 Rutting greater than 10mm on the state highways (Source NZ Transport Agency condition report)

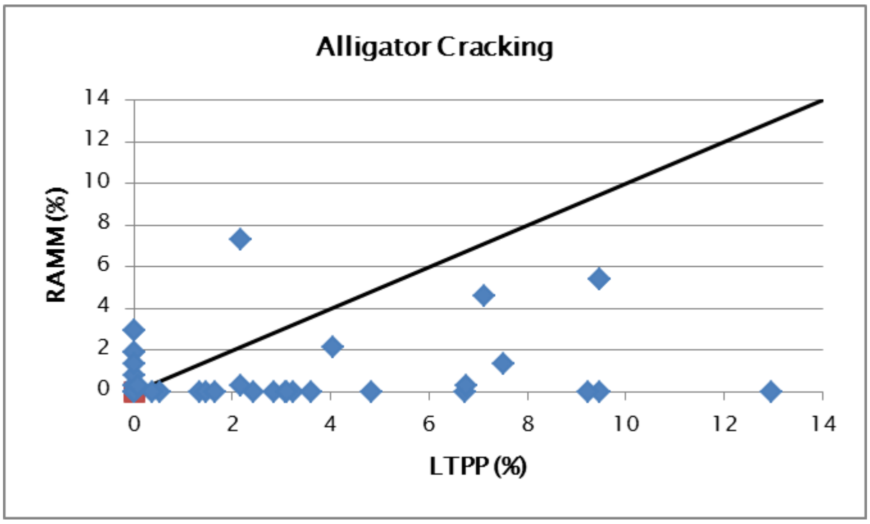


	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
🟥 National	4.00	6.28	6.58	6.89	6.57	7.46	7.71	11.14	10.41	9.80
■ PSMC005	5.06	6.87	7.59	7.18	7.29	8.81	8.49	10.67	9.77	8.63

The exact impact of moving to an automated condition survey is unknown at this stage, although it is expected to be significant. Any manual condition survey probably has some bias associated with a specific rater. Yet, due to the multiple visual surveys used across the country, it is expected the multitude of biased surveys will lead to an unbiased, yet still imprecise result (similar to the bottom right illustration in figure 2.9). Regardless of whether the manual surveys have some bias or not, it is expected there will be a significant bias shift between the manual and the automated defect observations. For example, figure 2.12 illustrates the differences between crack observations on the LTPP sites. From this comparison it is evident the RAMM rater was perhaps underestimating the cracks, when compared with the LTPP surveyor. The bias between two measurement techniques can only be understood if both methods are compared to a third technique (presumably the most accurate measurement or true value).

It is further expected the automated surveys will have an associated bias. For example, given the automated process, certain factors naturally result in bias in the observations. Henning and Mia's (2013) study concluded that the LCMS was perhaps identifying slightly more cracks than existed. As an example, it may have picked up edges of pothole patches as cracks. This is perhaps the reason for the recommendation from the FHWA study (Pedro et al 2014) to have an element of human verification involved in the data processing stages.

Figure 2.11 Comparing RAMM rating surveys with LTPP manual surveys (Tapper et al 2013)

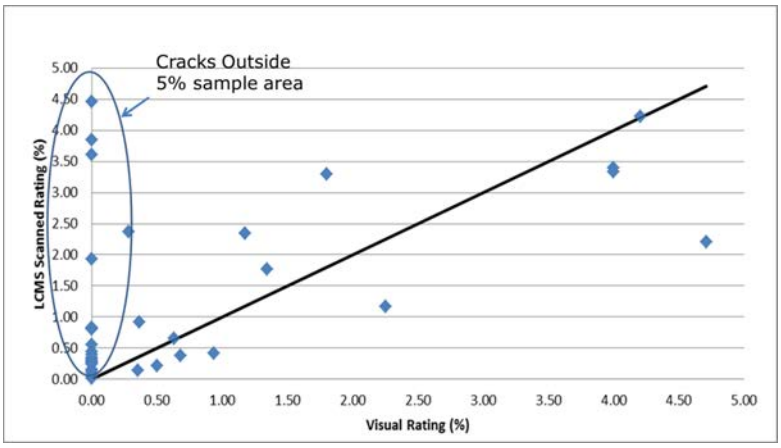


2.4.2 Influence of sampling

One of the main advantages of automated defect identification is that it allows for a 100% survey regime as opposed to the current sampling process adopted for the RAMM surveys. The challenge of any sampling process is that it is only a representation of the true population behaviour. Tapper et al (2013) showed no RAMM sampling size yielded exactly the same outcome as a 100% survey. From a pragmatic perspective, the Tapper et al study (2013) recommended increasing the RAMM surveys to a minimum 20% sample size. Henning and Mia (2013) also considered the impact from different sample sizes, and figure 2.12 illustrates their findings, which suggest the LCMS picks up more cracks than the manual raters. The report named two reasons for the additional observed cracks:

- A certain percentage of additional cracks were observed outside the RAMM-rated assessment lengths that may have showed no cracks.
- In some cases the LCMS identified or confused other defects as cracks, such as mechanical damage to the road surface.

Figure 2.12 Comparing 100% LCMS surveys with 10% sampling RAMM surveys (Henning and Mia 2013)



2.5 Outcome from the literature review

The literature review confirmed most of the assumptions at the onset of the project:

- It confirmed the technology was advanced enough to provide a significant step towards more robust defect identification and quantification. Sufficient research was completed to prove this aspect of crack identification whereas the same confidence does not exist for other defects.
- The research confirmed the shift towards any automated defect detection would result in a shift of the outcome values. Naturally this would also have an impact on all asset management processes that rely on this data. It is also clear this impact assessment could only be based on local studies in order to understand the full impact under local conditions.

The encouraging outcome of the literature research was the strong indication of more consistent and robust data resulting from automated detection methods, which would also assist in more informed decision making.

Recommended further work includes a direct comparison to assess the ravelling and shoving identified. It is further recommended to consider adding flushing to the list of defects for this research.

3 Research methodology

3.1 Understanding accuracy, repeatability and reproducibility of equipment

The first step to understanding the impact of the automated defect identification is to know its capabilities and limitations. It was recognised at the onset of this research that much of this knowledge would come from existing international research. This is described in the literature review.

3.1.1 Direct comparative analysis

The direct comparative analysis completed for cracking (Henning and Mia 2013) needs to be repeated for ravelling, shoving and perhaps flushing, and based on the LTPP data, which is summarised in section 3.3.

The next comparative analysis attempts to quantify the impact of a 100% automated survey versus a 10% and/or 20% sampling regime, typically used by local councils. The objective of this comparison is to gain a full understanding of the impact of changing the assessment technique from manual to automated.

These comparisons could also be vital in putting forward a business case for changing to the automated technique.

Dunedin City Council offered a fully scanned survey across their network for the purposes of this analysis.

3.2 Understanding and managing the impact of changing to an automated assessment

Indications are that a change to automated assessment will lead to an increase in the defects being recorded. There are two ways of dealing with this. The first, and perhaps the easiest, is to adjust the rules according to the newly measured ranges. Although this is effective, it could also be confusing to the industry, which has become used to certain ways of describing road surface conditions. For example most road engineers know that a surface condition index (SCI) of around five would not necessarily be a problem, but a SCI of 10 would suggest a severely affected road. Normalisation techniques will be used to adjust the composite indices such as the surface integrity index and the SCI according to the same principles used during the initial development of these indices (Fawcett et al 2008).

There is also a likelihood of having to change the way in which the defects are described and recorded. Most defects are described according to their seriousness, and the extent to which the road length would be affected. In this regard, the RAMM surveys only record the length of wheel-path affected by the defect. This method was adopted so relatively inexperienced road raters could be employed, without needing a technical understanding of the failure mechanisms. It would be senseless to retain these definitions when automated techniques can provide much richer attributes of the defect. For example, there may be significant ravelling in a small area, or there may be isolated chip loss for the entire road length. Part of this step in the research would be to establish the most effective way of defining the seriousness of the defect and its extent. The LCMS data would be far better utilised with an alternative description of defects than what is currently being used.

Naturally, the change mentioned above would have an impact on current downstream analysis systems including RAMM TSA and dTIMS. With the combined knowledge and experience within the research team who were part of this project, the required changes as a result of different defect definitions, would be quantified and provided to the respective software vendors.

The outcomes from this stage of the research include:

- Make recommended changes to all composite indices for cracking, ravelling and shoving.
- Redefine the descriptions of the degree and extent of defects resulting from the automated measurements.
- Make recommendations to the industry for adjusting rules and triggers defined within software applications and performance measurement frameworks.

3.3 New Zealand LTPP programme data collection

The New Zealand LTPP programme originated in 2000 when 64 LTPP sites were established on the state highways. Two years later the same principles were used for the establishment of an additional 84 LTPP sites on local roads, with approximately half the sections on rural roads and the remaining portion on urban roads.

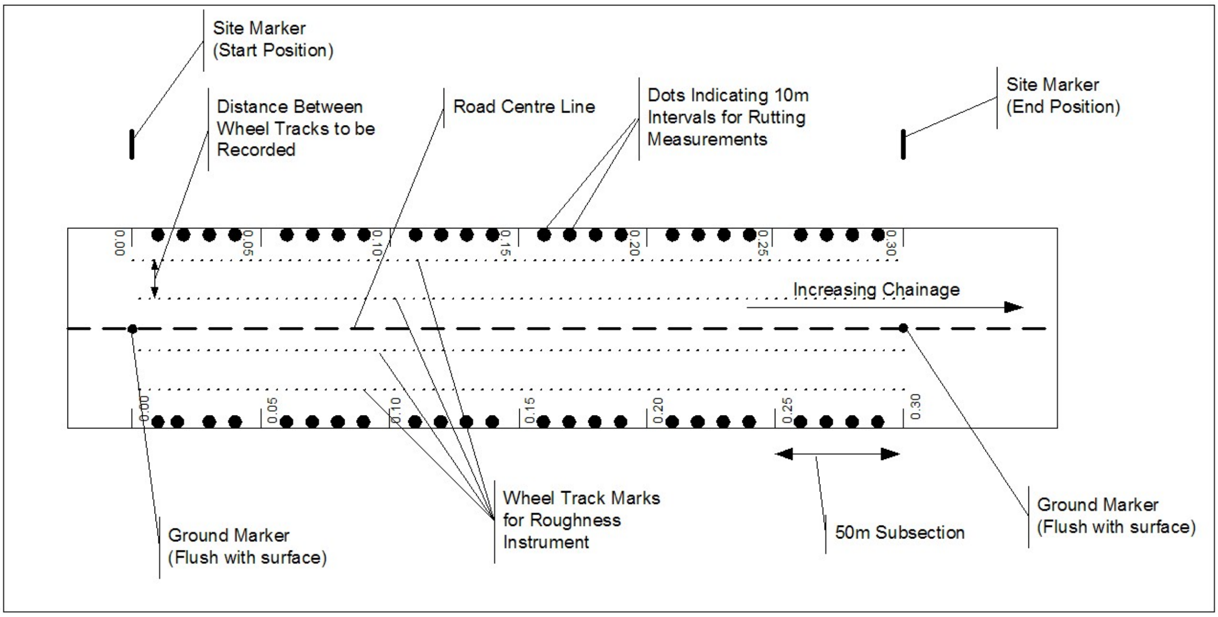
3.3.1 Sites established

Henning et al (2004) describe the rationale of the site establishment in detail. The 140 sites were spread across the country and covered all the factors expected to influence the performance of New Zealand road networks, including:

- climatic and soil condition
- traffic loading
- pavement types/strength
- pavement age/condition
- maintenance regime.

The LTPP sites were 300m in length with each section subdivided into 50m subsections for assessment purposes. The layout of the sites is depicted in figure 3.1.

Figure 3.1 New Zealand LTPP site layout (Henning et al 2004)



The New Zealand LTPP programme adopted a climatic sensitivity rating to classify the combined impact of rainfall and soil moisture sensitivity. For this classification, New Zealand was divided into four environmental sensitivity regions.

3.3.2 Condition data

A private survey contractor undertakes an annual data collection on all LTPP sites. The contract for this survey specifies the required data collection accuracy and repeatability. As only one contractor has been involved since the initiation of the programme, no changes were made to the methodology of data collection. An outstanding outcome from the programme so far has been the quality and subsequent usefulness of the data (Henning et al 2008). The data collection includes:

- a manual assessment of all defects, involving the recording of the exact extent and dimensions of the defects
- manual measurement devices, used for rutting, roughness and texture depth (refer to figure 3.2)
- traffic counts, undertaken using classification loop counters
- the recording of all maintenance
- recording changes to sites through detailed site notes and photographs.

In addition to the above, each site is surveyed annually using the HSD collection survey as part of the Transport Agency state highway network survey processes. Four repeated runs are undertaken in both directions for each site using the HSD equipment. These parallel surveys have resulted in significant research opportunities in the data collection area.

Figure 3.2 New Zealand LTPP programme rutting measurements (Henning et al 2004)

3.4 Outputs expected from this research project

The specific outputs from this research were to provide:

- an understanding of the status of international research on automated defect detection technology and its application to New Zealand pavements and conditions
- quantified measures for the accuracy and repeatability of the measurements, as well as the relationship and confidence levels associated with each
- a list of limitations for the respective measures, eg the algorithms processing the measurements only adhere to the programme's 'interpretation' capability and could analyse the edges of a patch as a block crack. Accepting the measurements will never be perfect, and it is important for the users to understand these idiosyncrasies
- recommended changes to all composite indices that use cracking, ravelling and shoving
- a re-definition of the degree and extent description measures of defects resulting from automated measurements
- recommendations to the industry for adjusting rules and triggers defined within software applications and performance measure frameworks
- the industry with specific guidance related to survey regimes, ie frequency of sampling the network.

After establishing the data structure, this information would be passed onto the software providers in order to adjust their data structures accordingly.

4 Results from comparative analyses

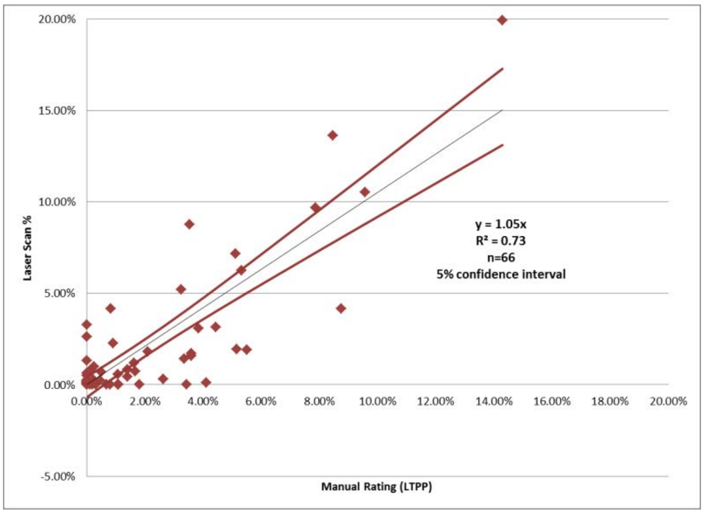
4.1 Accuracy and repeatability

Although not a primary focus of this research, an understanding of the accuracy and repeatability of the scanning technology is important for evaluating the usefulness of the measurements when compared with the manual surveys. The data used for the accuracy and repeatability measures was taken from the LTPP programme and was also used to develop the algorithms for processing the ravelling and shoving data. Overall, the LTPP dataset did not produce sufficient data for an accurate statistical assessment of the ravelling and shoving. Due to the defect size of ravelling and shoving, the accuracy and repeatability measures on other defects would be comparable to the cracking outcome, if not better.

4.1.1 How accurate is the laser scanner?

The results from comparing cracks are presented in figure 4.1, which shows a direct comparison of cracks percentage for the LTPP sub-sections. The manual data was visually recorded and gave an estimated crack length and width. A similar approach was used with the laser to estimate the crack percentage.

Figure 4.1 Comparing scanning laser measured cracks with the LTPP manually assessed cracks



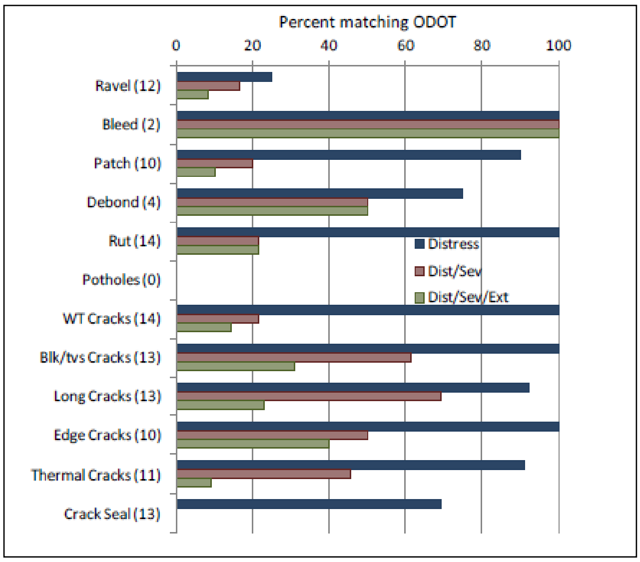
The observations from the figure are:

- The line of equality has a slope of 1.05, suggesting laser measures would have a bias of 0.05% compared with manual measurements.
- The R^2 between the measures is 0.73, suggesting an element of scatter between the two measures. The scatter is on both sides of the line of equality, thus suggesting a lack in bias from the scanner.

The results were significantly better than the RAMM surveys on the same sections (Henning and Mia 2013).

A study from the University of Ohio, compared scanning technology from three different providers with manual surveys (Vavrik et al 2013). Figure 4.2 shows the outcome from one of the devices using the same system as that being tested in New Zealand. The measurements are compared with a visual rating that describes the degree and extent of the defects. According to these tests the scanner identified 85% of the defects and in general the severity of the defects was more accurately measured than their extent. The measurements of ravelling and bleeding were more accurate than those of cracking, thus supporting the statement that if we are satisfied with the accuracy of measuring cracks, we should also have confidence in the measuring processes for ravelling and shoving.

Figure 4.2 Comparing laser scanners to visually assessed information (Vavrik et al 2013)



4.1.2 Is it possible to improve the accuracy of the laser measures?

The accuracy of the scanning laser is similar to the HSD measurement, which uses the stationary laser. The laser measurement itself is highly accurate, but the positioning of the measurements remains a challenge. However, the most challenging part of the technology is the interpretation of the laser measurements by the software and from there identifying the appropriate defects. The complexity of these algorithms is differentiating between the defects. For example, initial algorithms identified line marking as a longitudinal crack. It is also common for the algorithms to pick up a pothole patch and ‘see’ a crack surrounding the patched area. All indications are that most of the laser measurements identify a number of false positives (over-estimate) defects. Researchers from the vendors and stakeholders dedicate significant resources towards addressing the shortcomings in this technology and it is expected that significantly better results will be obtained within the next couple of years.

The Transport Research Laboratory has developed algorithms for aligning the special data from the laser measurement to the actual positioning (McRobbie et al 2015). The research considered a 3D alignment of data to the ‘real’ positioning. Figures 4.3 and 4.4 illustrate what could be achieved through the alignment of measurements. Figure 4.3 shows two longitudinal profiles, one being the base line and the other the test run. The offset of the test measurement is clearly visible. Figure 4.4 illustrates the resulting condition distribution for an aligned and un-aligned scenario. In this example, the alignment resulted in the range of the distribution being half that of the un-aligned measurement.

Figure 4.3 An example of aligning the measurement to its true position (McRobbie et al 2015)

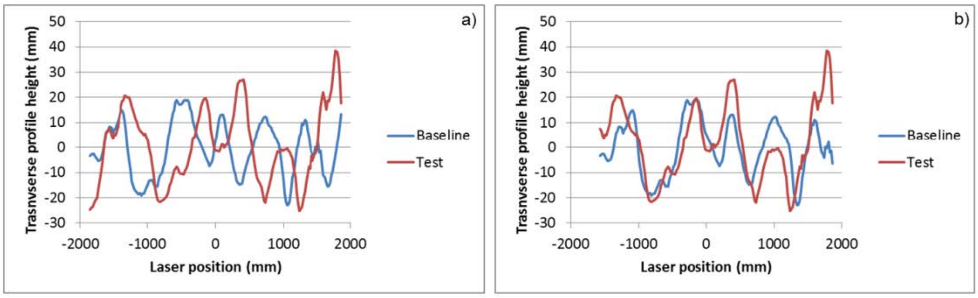
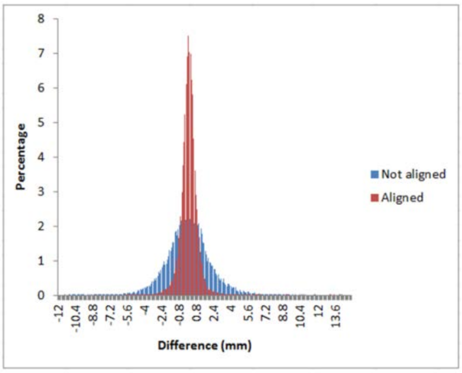
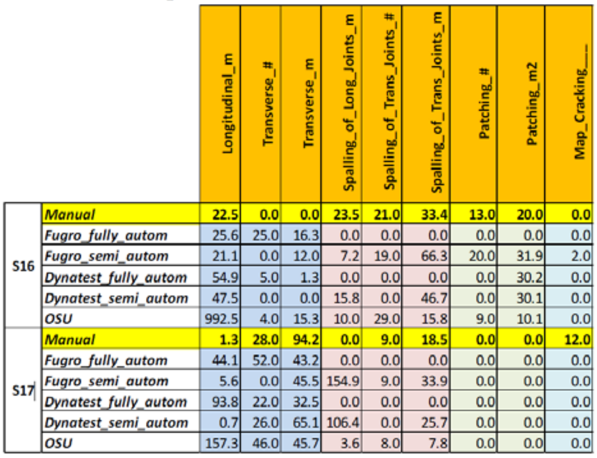


Figure 4.4 The impact of measurement alignment on the accuracy (McRobbie et al 2015)



Perhaps one of the most significant research projects into the accuracy of laser scanning technology was undertaken by the University of Texas at Austin (Serigos et al 2014). A particularly useful element of this research was a comparison between fully automated and semi-automated outputs. The semi-automated processing involved some level of manual correction of algorithms by viewing the profile images. A comparison between these two approaches is presented for two measurement providers in figure 4.5. These results were obtained from scanning data on jointed concrete pavements, yet the principles remain the same. The manual validation specifically highlighted a number of false positive and false negative items.

Figure 4.5 Comparing the fully automated results with manually aided assessments (Serigos et al 2014)



		Longitudinal_m	Transverse_#	Transverse_m	Spalling_of_Long_Joints_m	Spalling_of_Trans_Joints_#	Spalling_of_Trans_Joints_m	Patching_#	Patching_m2	Map_Cracking___
S16	*Manual*	22.5	0.0	0.0	23.5	21.0	33.4	13.0	20.0	0.0
	Fugro_fully_autom	25.6	25.0	16.3	0.0	0.0	0.0	0.0	0.0	0.0
	Fugro_semi_autom	21.1	0.0	12.0	7.2	19.0	66.3	20.0	31.9	2.0
	Dynatest_fully_autom	54.9	5.0	1.3	0.0	0.0	0.0	0.0	30.2	0.0
	Dynatest_semi_autom	47.5	0.0	0.0	15.8	0.0	46.7	0.0	30.1	0.0
	OSU	992.5	4.0	15.3	10.0	29.0	15.8	9.0	10.1	0.0
S17	*Manual*	1.3	28.0	94.2	0.0	9.0	18.5	0.0	0.0	12.0
	Fugro_fully_autom	44.1	52.0	43.2	0.0	0.0	0.0	0.0	0.0	0.0
	Fugro_semi_autom	5.6	0.0	45.5	154.9	9.0	33.9	0.0	0.0	0.0
	Dynatest_fully_autom	93.8	22.0	32.5	0.0	0.0	0.0	0.0	0.0	0.0
	Dynatest_semi_autom	0.7	26.0	65.1	106.4	0.0	25.7	0.0	0.0	0.0
	OSU	157.3	46.0	45.7	3.6	8.0	7.8	0.0	0.0	0.0

4.2 Comparison between RAMM surveys and laser measurements using network data

There will always be a difference in the quantity of defects observed through different measurement techniques. Earlier analysis indicated there is a substantial difference in the defects assessed using the RAMM survey method compared with the scanning laser. Differences in outcomes are not only limited to a direct comparison of the survey techniques, but the greatest differences between techniques stem from statistical aspects. RAMM surveys are normally undertaken on a sampling basis that varies between 10% and 20%. There will always be an expected bias in survey results when comparing a 100% survey with a smaller sample. The assumption is that a sample sufficiently represents the population by both over- and under-estimating the number of defects. Probability theory supports research findings that suggest defects observed for a 100% sample will in most cases be higher than the defects observed on smaller sample sizes.

The comparisons set out in subsequent sections were undertaken for two cases:

1. Direct comparison of RAMM survey results and scanner laser on the same inspection length
2. Comparison on a network basis where the RAMM survey sample was compared with a full scanner survey.

The data was sourced from a survey of the Whangarei network.

4.2.1 Cracking

Figure 4.6 illustrates the direct comparison between cracking data for the inspection length identified using RAMM survey versus the LCMS measurements. The RAMM survey cracking data has been adapted to percentage cracking using the transformation function from the dTIMS analysis:

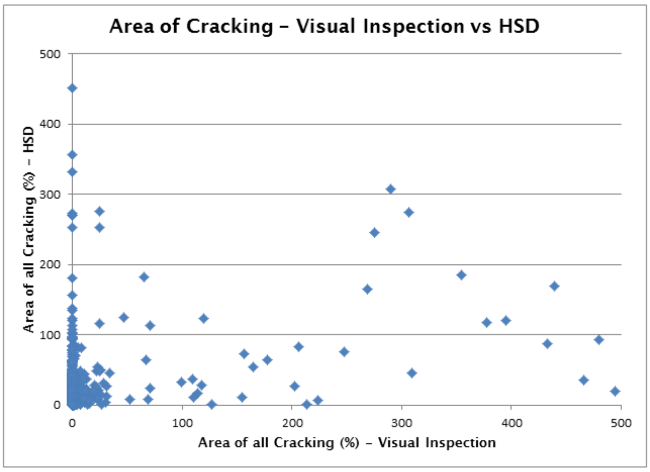
$$Crack\% = 0.0004\left(\text{alligator } x\,\frac{Aveg_{insp\_lengh}}{\text{insp\_length}}\right)^2 + 0.28\,\text{alligator}\,\frac{Aveg_{insp\_lengh}}{\text{insp\_length}} \qquad \text{Equation 4.1}$$

Where: crack% is the percentage of cracking for the inspection length

alligator is the RAMM rated cracking (length of wheel path rated)

insp_length is the inspection length

Figure 4.6 Comparing visual survey and LCMS cracking on inspection length

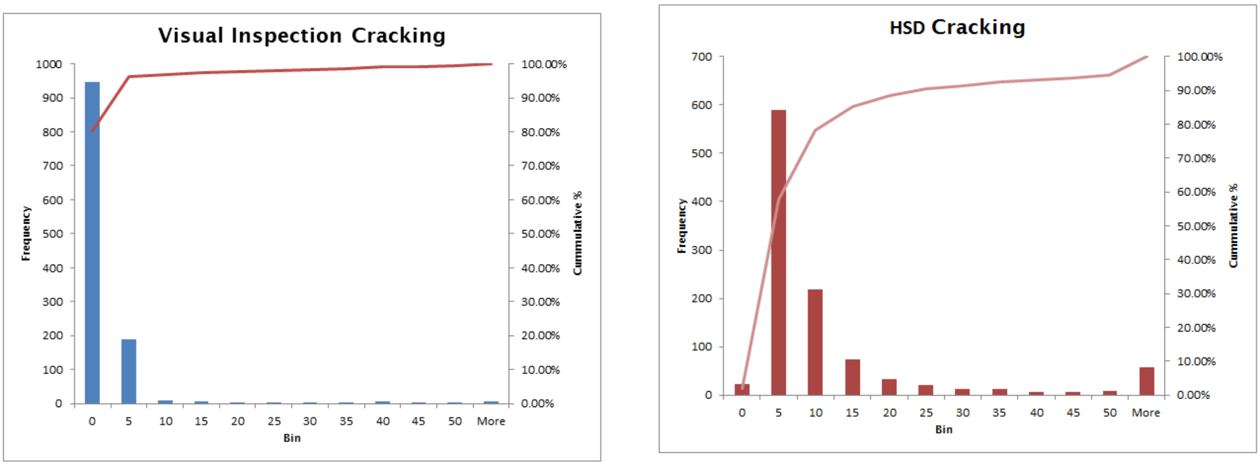


Note: Measurements higher than 100% suggest the length of cracking and its assumed influence area is larger than the inspection length itself. For accurate crack measurements, this value will not exceed 100%.

There is a weak correlation between the cracking data resulting from the two survey methods. Also, there is a high number of inspection lengths with cracking observed through the LCMS, while the raters did not identify cracks on these sections.

The comparison between the distributions of cracking for the network is presented in figure 4.7. Note that the visual rated cracking represents a sample of the network, while the scanning cracking is a continuous measured value.

Figure 4.7 Comparing the cracking distribution for visual survey (sampled) and LCMS 100%



There is a magnitude shift in crack values between the results of the LCMS and those of the visually measured cracking. This difference is attributed to three factors:

- potential false positive cracking identified with the LCMS
- the statistical shift going from a sample to full survey results will result in higher values being observed
- an under reporting from the visual surveys.

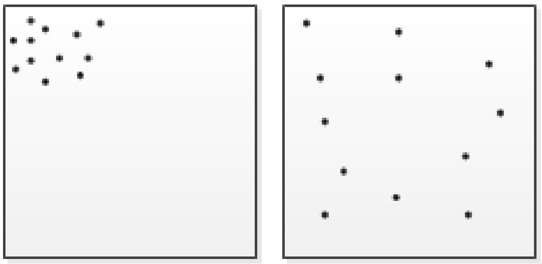
4.2.2 Ravelling

The definition of scabbing/ravelling as per RAMM rating is:

> *Surface with more than 10% of the sealing chip loss. In the case of asphaltic concrete surfaces this will be the area of the pavement showing signs of surface attrition.* (HTC Infrastructure Management Ltd 2000)

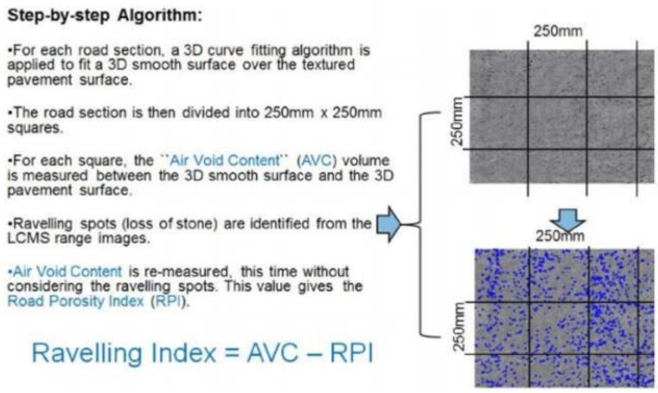
Many other countries use a degree and extent measure to quantify the ravelling. Ravelling can occur in several ways. Figure 4.8 illustrates two road sections with the exact same amount of ravelling (chips being lost). For the road section on the left the ravelling is isolated to a single area, whereas the right-hand figure shows consistent ravelling spread across the entire section. Obviously the maintenance decision for each section would be entirely different.

Figure 4.8 Ravelling on two road sections



The proposed ravelling definition for the LSMC is illustrated in figure 4.9.

Figure 4.9 Ravelling assessment for the LCMS (Source Pavemetrics)



Step-by-step Algorithm:

- For each road section, a 3D curve fitting algorithm is applied to fit a 3D smooth surface over the textured pavement surface.
- The road section is then divided into 250mm x 250mm squares.
- For each square, the ``Air Void Content`` (AVC) volume is measured between the 3D smooth surface and the 3D pavement surface.
- Ravelling spots (loss of stone) are identified from the LCMS range images.
- Air Void Content is re-measured, this time without considering the ravelling spots. This value gives the Road Porosity Index (RPI).

Ravelling Index = AVC – RPI

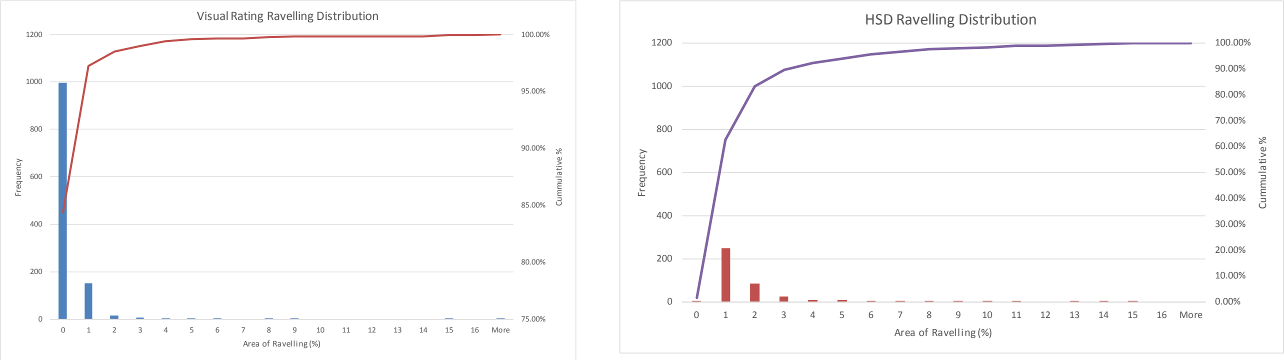
The ravelling table outputs a result for every 1m. The results of all 250mm^2 squares inside the 1m road section are calculated to give:

- average ravelling index (RI) of all squares in section (cm^3/m^2)
- percentage of 250mm^2 squares over RI threshold value (default 100cm^3/m^2). The threshold value can be adjusted
- area affected by ravelling in square millimetres (mm^2).

The suggested multi-description of ravelling makes sense as it overcomes the difficulty of quantifying the ravelling differences illustrated in figure 4.8. The actual ravelling algorithm was developed on asphalt surfaces and it is suspected more work is required to calibrate it to different chipseal surfaces. Figure 4.10

shows the distribution of ravelling on a full network compared with the sampled ravelling for the inspection lengths.

Figure 4.10 Comparing the ravelling distribution for rated and LCMS measurements



The observations from the figures are:

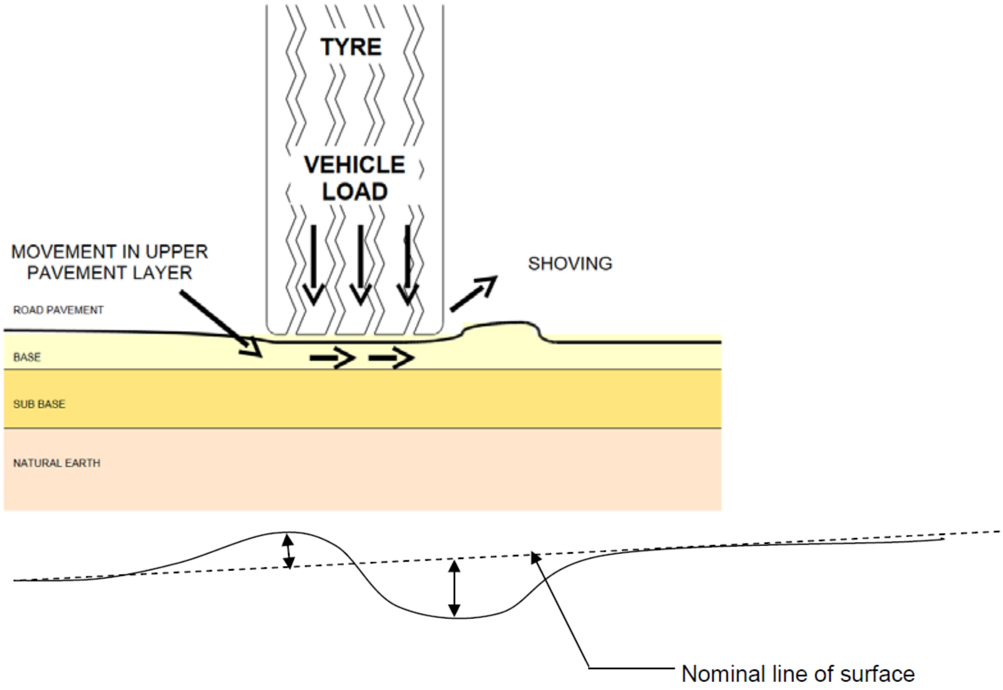
- There are few road sections that have no ravelling according to the LCMS (this is an unlikely outcome).
- There is a relatively good correspondence between the ravelling percentages for sections where ravelling was observed.

The results suggested the ravelling algorithm still requires calibration to the chipseals commonly used in New Zealand.

4.2.3 Shoving

Shoving is the permanent deflection and bulging of the road surface in a transverse generally parallel to the direction of traffic, and/or horizontal displacement of surfacing materials. Shoving is typically caused by braking, accelerating or turning vehicles.

Figure 4.11 Definition of shoving



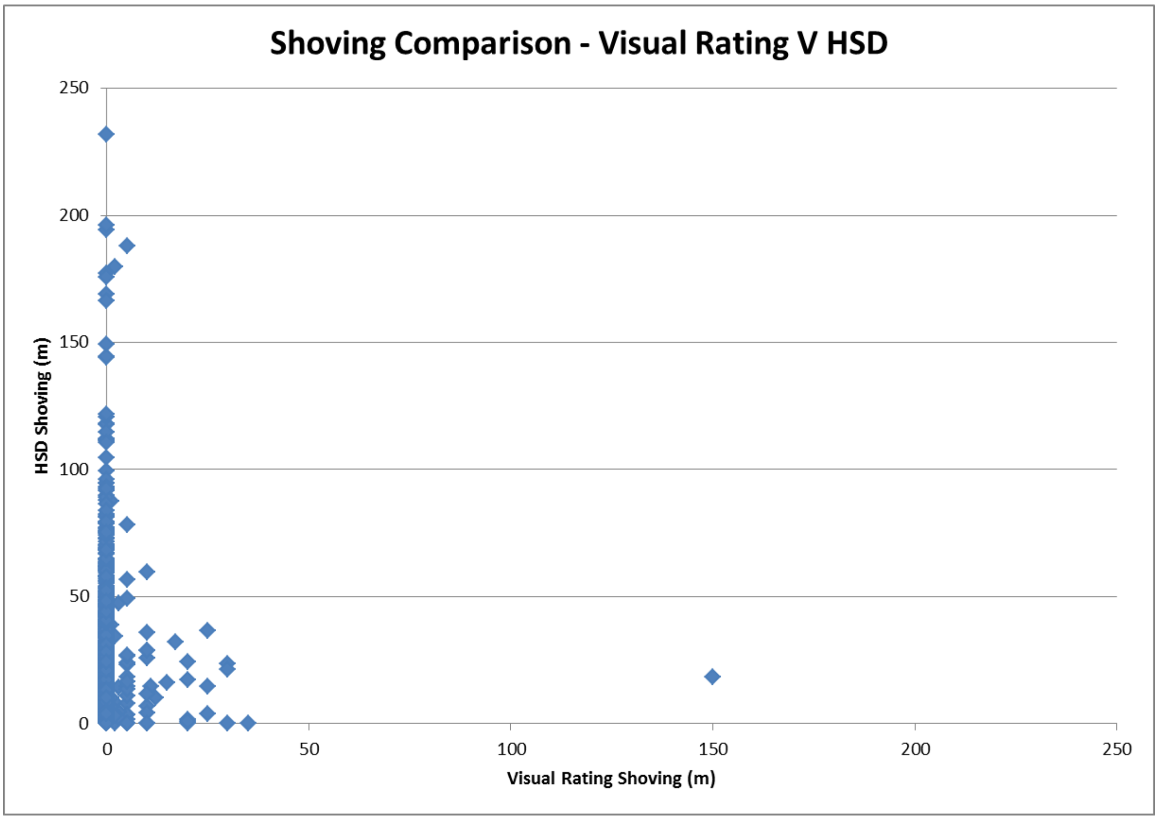
The current depth criteria trigger advised by the supplier, in terms of vertical displacement used by the Pavemetrics algorithm is 5mm. This therefore requires a minimum upward slope of 5mm, at the side of the shove, before it is considered a shove. The algorithm output includes the length of the shove and also the shove count. The shove count is the number of instances the algorithm detects the shoving criteria.

DCL New Zealand advised the algorithm currently checks for the presence of shoving from the centreline to the edgeline or channel/seal interface, for both sides of the road. When in a rural environment, it checks from the centreline to the edge of seal.

The key difference between rutting and shoving is that shoving has displaced material that forms a bulge or heave at the top of a depressed (rutted) area, as shown above.

The direct comparison between the RAMM rated shoving and the automated shove detection is depicted in figure 4.12. The figure shows a comparison between the lengths of shoving recorded by the respective survey methods.

Figure 4.12 Comparing automated shoving detection to manual measurements



It appears from the figure there were a number of false positive measurement picked up through the scanning laser. The following section investigates the actual occurrence of shoving on the network.

4.3 Field validations – shoving measurements

4.3.1 Site selection criteria

Output from the shoving algorithm was categorised into the following groups:

- false negatives (the algorithm reports no shoving, but shoving exists for the same section in the visual condition survey data)
- false positives (the algorithm reports shoving exists, but no shoving exists in the visual condition survey)
- true positives (the algorithm reports shoving exists, and shoving exists in the visual condition data as well)
- true negatives (the algorithm reports no shoving, and no shoving exists).

A sample from each of the following groups was selected to field validate:

- false negatives
- false positives
- limited amount of true positives.

An additional desktop audit check was made on the proposed sites, prior to field validation. This involved checking that shoving data existed in the latest visual condition rating data, completed in 2015. This was to avoid checking sites that did not met the validation criteria. A filter was also used to identify sites with a modelled shove length greater than 20m, to avoid trying to detect very small quantities of shoving. Some of the visual rating sections can be up to 500m in length

4.3.2 Field validation results

The second part of the process was to field validate a selection of sites that met the above criteria. A summary of the results is shown below, with more detail included in appendix B.

Table 4.1 Field validation results

Algorithm outcomes	No. of sites	Shoving exists on site	No shoving on site	Rutting was observed –some may classify it as shoving	Comments
False negative	14	4	10	–	More outcomes with automated measurements correctly reported no shoving
False positives	217	9	208	–	The automated measurement over-identified shoving, where none visually existed
True positives	87	56	10	21	The algorithm picked up field validated rutting.
Total	318	69	228	21	

4.3.3 Field validation observations

While undertaking the site validation, a number of observations were made and recorded via photographs. These are reported by site with commentary on the observations.

Figure 4.13 Note drop off into catch pit, greater than 5mm

Figure 4.14 Scabbed surface or high lip at the channel often greater than 5mm

Figure 4.15 The straight edge shows more clearly the vertical height. The eye can generally pick up >10mm, but finds it hard to pick up <10mm

Figure 4.16 Shoving with a clear vertical height

Figure 4.17 A very shallow rut, hard for the eye to detect, but it is greater than 10mm as the pen fits underneath

Figure 4.18 A closer view of a shallow depression that can be exaggerated when chip loss occurs at the same point. For larger size chip this depth can be greater than 5mm

Figure 4.19 Note the mountable islands in the middle of the road. Shoving was reported through the algorithm at these locations

Figure 4.20 A closer view of the vertical height that exists when scabbing occurs in larger stone chipseals

4.3.4 Summary of lessons learned from field inspection

Some valuable lessons resulted from the filed inspection including:

1. The laser measurements are extremely accurate – but the laser does not always measure the defect it is supposed to identify (refer to the photographs in section 4.3.3).
2. Therefore, manual validation must confirm the algorithm requires the detection of a crest and roll-over shape, not just a vertical height, to avoid rutting being reported as shoving.
3. Consider setting the algorithm vertical height to something that is visually more identifiable. This will help with acceptance of the approach and technology, and align more closely with the current visual

inspection. Once this becomes more acceptable, then the vertical height can be lowered. It is suggested a height of 20mm be used, and then further field validation be undertaken.

4. When undertaking field validation of the modelled shoving output, the results are reported at a 10m interval. This helps to identify more clearly where the shoving has been reported, and narrows down where to look on the road.
5. Undertake a greater sample of outputs and compare these against the location where manual shoving has been reported.
6. Consider the possibility of detecting the presence of shoving using a defined envelope around the wheel path, as opposed to the full width from the centreline to the edgeline or edge of seal.

5 Industry impact

A survey of American state road agencies revealed that only 18% of agencies still undertake manual condition surveys. More than 55% use scanning lasers and the remaining agencies use image analysis for the recording of defects (Vavrik 2013). Much was learned from the different approaches of these agencies to implementing the technology.

The scanning laser has the potential to give far more repeatable survey results than the RAMM visual surveys. However, the comparative analysis from this research has concluded there is still some calibration work required for New Zealand conditions prior to this technology being adopted on a large scale. All the algorithms work, yet, given the difference between chipseal and asphalt surfaces, there is more work required to improve the accuracy of the scanner laser. This process should therefore be seen as part of the overall implementation process. Further aspects to consider during the implementation are discussed in subsequent sections.

5.1 Definition of defects

It was interesting to note some agencies in the US have adopted scanning laser technology but still use historical definition methods. Some, for example, describe defects on a ten or five-point degree and extent scale. The only logical reasoning for converting the data back to its prior definition would be:

- having to be able to show current trends compared with future trends
- not wanting to change data structures in existing software
- decision-making algorithms in pavement management systems to accommodate the new defect definitions.

A different approach is recommended for New Zealand in order to fully utilise the information that could be extracted from the laser surveys. The recommended definitions are summarised in table 5.1.

Table 5.1 Recommended defect definition

Defect	Measurement definition	Notes
Cracking	Record the linear length of cracking expressed as a percentage area crack in relation to the section length. The area of the crack is determined by taking a 0.5m width as the standard influencing area of a crack.	Where more than one parallel crack exists, the area is determined by the length of the cracks times the total width (outside crack plus 250mm)
Ravelling	Ravelling index plus the percentage of effected area affected.	There may be a need to re-define the ravelling index differently for chips seals.
Shoving	Shove length and shove depth	The shoving definition is dependent on a height deferential between 'lip' and depression – it is recommended this should be 15mm

5.2 Data requirements

The data requirements for the scanning laser would be similar to that of the HSD. The requirements are summarised in table 5.2

Table 5.2 Data requirements for the laser scanning data

Item	Requirement
Raw data storage interval	20m separate for each lane
Aggregated data	Summarised to 100m interval
Cracking data items	% wheel path cracking (alligator) % longitudinal cracking
Ravelling	Average ravelling index of all squares in section (cm^3/m^2) Percentage of 250mm^2 over ravelling index threshold value (default 100cm^3/m^2). The threshold value can be adjusted
Shoving	Length of shoving (m) Height of shoving (mm)
100m aggregated data	Minimum, maximum and average value for measurements

5.3 Impact on performance measures

None of the individual defects are currently being used for performance monitoring in New Zealand. Agencies report on surface condition through a composite SCI given by:

$$SCI = min\,(100, (CI + AI)) \qquad \text{(Equation 5.1)}$$

$$CI = min\,(100, max\,(4 * ACA + 0.5 * ARV + 80 * APT + 20 * APH + 1.2 * AFL)) \qquad \text{(Equation 5.2)}$$

$$AI = 3 * min\,(100, max\,(0, ((AGE2 - SLIF) / (SLIF * 12))) \qquad \text{(Equation 5.3)}$$

Where:

CI = condition index

AI = age index

ACA = percentage of alligator cracking

ARV = percentage area of ravelling

APT = percentage area of potholes

APH = percentage area of pothole patches

AFL = percentage area of flushing.

The weighting of the condition indices is determined by the importance of a particular defect in relation to the others. Each defect is also normalised to the expected range for the index.

The research only considered cracking and ravelling but it is safe to assume the other indices will shortly become part of the scanning. All indications from this research are that the particular defect values will rise, thus causing a sudden increase compared with historical values of the SCI.

Unfortunately, such a ‘jump’ in the index value is unavoidable, yet given the additional value offered through a 100% survey and the consistency in measurements it is worthwhile. The laser information offers an additional benefit of making the trending of specific defects possible. For example, the authority will be able to consider the changes in cracking, ravelling and other defects over time. This will enhance understanding of the exact performance of surfaces on a network basis.

5.4 Impact on decision-making tools

As expected, the change in defect values on a network will also have a direct impact on all decision tools such as the RAMM treatment selection algorithm, dTIMS forecasted modelling and others. Fortunately, these tools were developed with the flexibility to adjust intervention criteria and would therefore only require a calibration of intervention criteria to appropriate levels given the increase in defect information.

Note that dTIMS contains a surface integrity index (SII) that is similar to the surface condition index (SCI) and a similar approach has to be followed to re-adjust this index.

5.5 Quality assurance

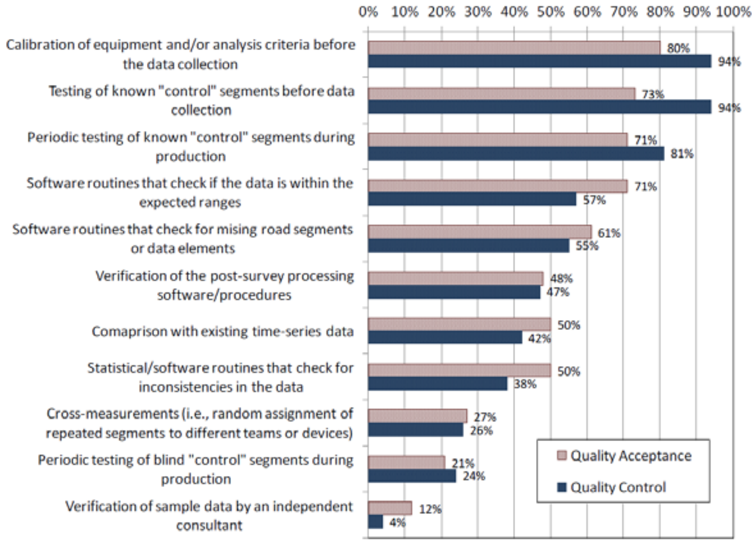
This research has concluded that the outcomes from a survey where QA measures have not been used are not sufficiently robust for wider use in the industry. However, when these measurements were on the Whangarei network taken it was the first time they had been used on chipseals in New Zealand. It is therefore essential to undertake a calibration process on New Zealand roads before the measurements become acceptably robust.

All literature investigated during this research indicated the need for frequent manual confirmation of the results to ensure the robustness of surveys. There are often systematic errors in algorithms that will identify a defect where there is none, but these could be removed with ease manually.

Figure 5.1 lists some QA processes adopted by state road authorities in the USA. The recommended minimum QA test for New Zealand should include (taken from Vavrik (2013):

- *Calibration of equipment and/or analysis criteria before the data collection*
- *Testing of known 'control' segments before data collection*
- *Periodic testing of known 'control' segments during production*
- *Software routines that check if the data is within the expected ranges*
- *Software routines that check for missing road segments or data elements*
- *Verification of the post-survey processing software/procedures*
- *Statistical/software routines that check for inconsistencies in the data.*

Figure 5.1 Percentage of agencies using given QA and quality control measures (N=18) (Vavrik 2013)



5.6 Survey specifications

The survey specification would be fairly similar to that of exiting HSD surveys in New Zealand, with the exception of different QAs to be employed as discussed in section 5.5.

6 Conclusions and recommendations

6.1 Main findings

The research findings, as listed below, were consistent with most international research projects that have studied the accuracy and repeatability of the laser technology:

- The laser technology is accurate and repeatable in its measuring, and the benefit of having a 100% road length covered by the surveys is particularly attractive for the intended data use.
- Comparison of the laser technology with existing practices remains a challenge and the results from such comparisons should be analysed with care. The comparison between laser scanning and a RAMM survey will never yield ideal results because:
  - The two assessment methods differ fundamentally in the way they define the extent of the defect – one-to-one comparison are therefore not possible.
  - RAMM surveys only covers a percentile of inspection lengths and a sample outcome will most likely differ from the full-length survey.
  - Ensuring the two assessment types reference the exact same location is difficult and there is a lack of confidence that comparisons are being made between the same road sections.
- The laser technology has identified defects successfully but has also identified a number of false positives. A more detailed study into the shoving measurements has identified number of road features that appear to trigger shoving according to the defined algorithm, but in reality identify a completely different road feature as a shove. The study has also confirmed a number of instances where the rating simply 'missed the shove' as it was not very apparent for a number of reasons.

6.2 The implementation plan

Using laser scanning for detecting road defects should be adopted by all road agencies. This recommendation is made on the bases of the significant benefits that can be realised from more accurate assessment, more repeatable and greater coverage of the road network.

The laser technology, despite its accuracy, cannot be applied as a 100% automated process. The computer algorithms that analyse the data still need significant 'learning' that can only be achieved if the technology is supplemented by manual validation of the outcome. Someone needs to work through the digital images to find erroneous identifications and feed this knowledge back to the algorithms. Once this is completed, business as normal survey contracts should include calibration procedures, validation and QA protocols.

Further consideration for Industry adoption are summarised in table 6.1.

Table 6.1 Industry impact and adoption

Item	Impact	Further work	Most suitable party to undertake
Defect definition	A new definition of 'defect' needs to be universally accepted by the industry. Sticking to the current definition and quantification will devalue the enhancements form the scanning laser.	Incorporate into national condition assessment guidelines.	RIMS/Transport Agency (include the RIMS Condition Guideline)
Shoving data	The field survey suggested that shoving data could be more accurate.	Additional calibration of the algorithm is needed.	Transport Agency (research to be driven by Surfacing Group)
Ravelling measurements	Ravelling is an important surfacing defect that has to be included in the surveys, once it provides sufficiently accurate results.	A study is required to calibrate the ravelling identification algorithm to different chipseals in New Zealand.	Transport Agency (research to be driven by Surfacing Group)
Data standards	New data standards are recommended, and should be universally accepted and incorporated in best practice guidelines.	Gain consensus from the industry and incorporate in guidelines.	RIMS/Transport Agency
Procurements	QA, validation and calibration requirements need to be specified.	Develop processes in collaboration with suppliers. Then incorporate in template specifications.	RIMS/Transport Agency

7 References

Annis, C (2011) *Statistical engineering*. Accessed August 2011. www.statisticalengineering.com/index.html

Austroads (2012) *Part 2: Pavement structural design. Guide to pavement technology*. Sydney: Austroads.

COLTO (1997) *TRH12 Flexible pavement rehabilitation investigation and design*. Pretoria, SA: Department of Transport, Committee of Land Transport Officials (COLTO).

Council for Scientific and Industrial Research CSIR (1992) *TMH9 Pavement management systems: standard visual assessment manual for flexible pavements*. Pretoria, SA: Department of Transport. Committee of State Road Authorities.

Fawcett, G, T Henning, N Pradhan and M Riley (2001) The use of composite indices as resurfacing triggers. Pp 11–14 in *Fifth International Conference on Managing Pavements, issue 68*, Seattle USA.

FHWA (2003) *Distress identification manual for the long-term pavement performance program*. FHWA-RD-03-031. Washington DC: US Department of Transport, Federal Highway Administration (FHWA).

GEIPOT (1981) *Research on the interrelationships between costs of highway construction maintenance and utilization*. HDM-III final report. BrasIlia.

Henning, TFP, SB Costello, RCM Dunn, CC Parkman and G Hart (2004) The establishment of a long-term pavement performance study on the New Zealand state highway. network. *Road and Transport Research 13*, no.2: 17–32.

Henning, TFP, DC Roux, CC Parkman and D Brown (2008) The utilisation of LTPP data in network management. *New Zealand Institute of Highway Technology (NZIHT) Conference*. Napier.

Henning, TFP and MN Mia (2013) Did we get what we wanted? Getting rid of manual condition surveys. Presented at *14th-Annual-NZ Transport Agency – NZIHT-Conference*, Auckland.

HTC Infrastructure Management Ltd (2000) *Data issue report. New Zealand dTIMS Implementation report.* Auckland: HTC Infrastructure Management Ltd.

Laurent, J, JF Hébert, D Lefebvre and Y Savard (2011) Using 3D laser profiling sensors for the automated measurement of road surface conditions (ruts, macro-texture, raveling, cracks). Accessed 17 December 2016. www.pavemetrics.com/articles-pavemetrics-products/

McRobbie, S, C Wallbank, K Nesnas and A Wright (2015) *Use of high-resolution 3-D surface data to monitor change over time on pavement surfaces. Research into pavement surface disintegration: Phase 3*. 100(4/45/12) HALC. UK: Highways Agency, Network Services.

Office of the Auditor-General (2014) *Water and roads: funding and management challenges*. Presented to the House of Representatives under Section 20 of the Public Audit Act 2001. Wellington: OAG.

Pedro, A, PA Serigos, M Burton, A Smit, JA Prozzi and MR Murphy (2014) Field evaluation of automated distress measuring equipment. *CTR technical report: 0-6663-2*. Austin, Texas: Texas Department of Transport, Federal Highway Administration.

Serigos, PA, M Burton, A Smit, JA Prozzi and MR Murphy (2014) Field evaluation of automated distress measuring equipment. *The University of Texas at Austin report number 0-6663-2.* Texas Department of Transport.

Tapper, M, TFP Henning and KD Dunn (2013) Improvement of visual road condition data. *NZ Transport Agency research report 528*.

Transportation Research Board (2007) Pavement lessons learned from the AASHO road test and performance of the interstate highway system. *Transportation research circular E-C118*. Washington DC: Transportation Research Board Pavement Management Section.

Vavrik, WR, LD Evans, JA Stefanski and S Sargand (2013) *PCR evaluation considering transition from manual to semi-automated pavement distress collection and analysis. State job number 134668.* Columbus, USA: Ohio Department of Transportation Office of Statewide Planning & Research.

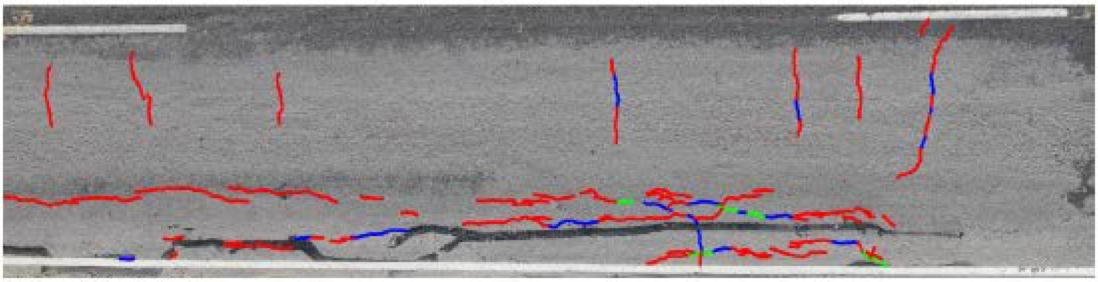
Wang, K (2004) *Tutorial on imaging for pavement distress survey. Automated pavement distress workshop presentation.* TRB Conference, Washington, DC.

Wix, R and R Leschinski (2012) Cracking – a tale of four systems. *25th ARRB Conference – Shaping the Future: Linking policy, research and outcomes*, Perth, Australia.

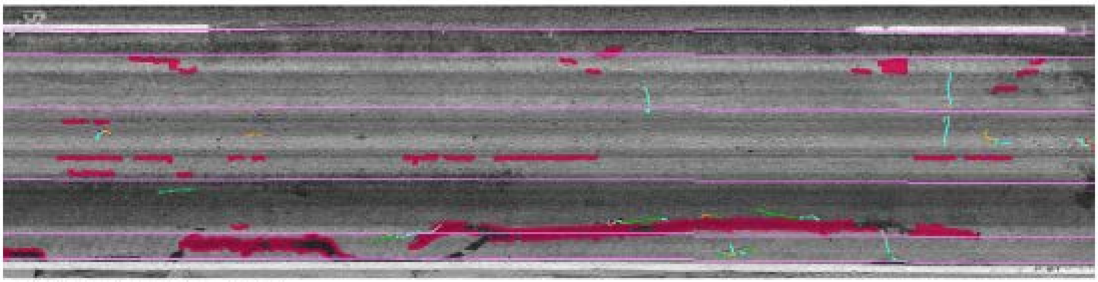
Appendix A: Research findings

A1 FHWA (USA) – Field evaluation of automated distress measuring equipment

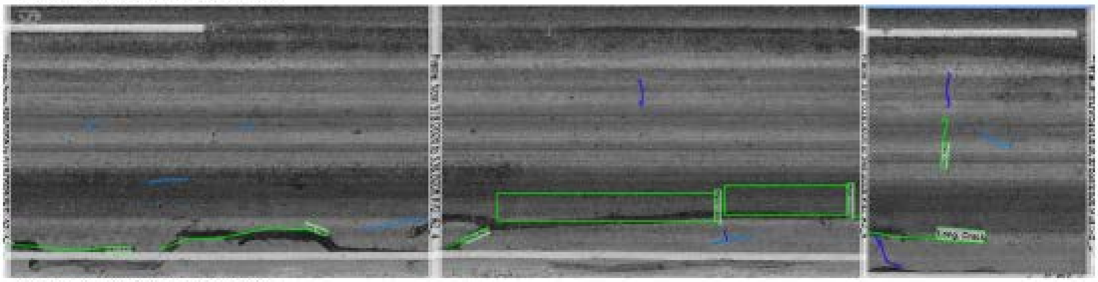
Figure A.1 **Manual and LTPP MDS crack map with Dynatest crack maps before and after manual intervention**



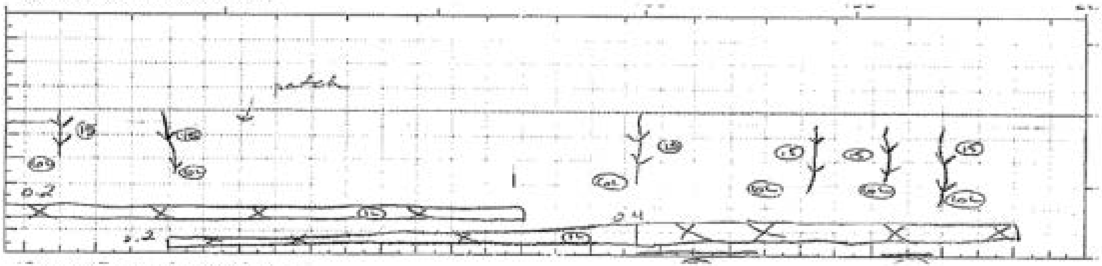
Manual Crack Map



Dynatest (Fully Automated)

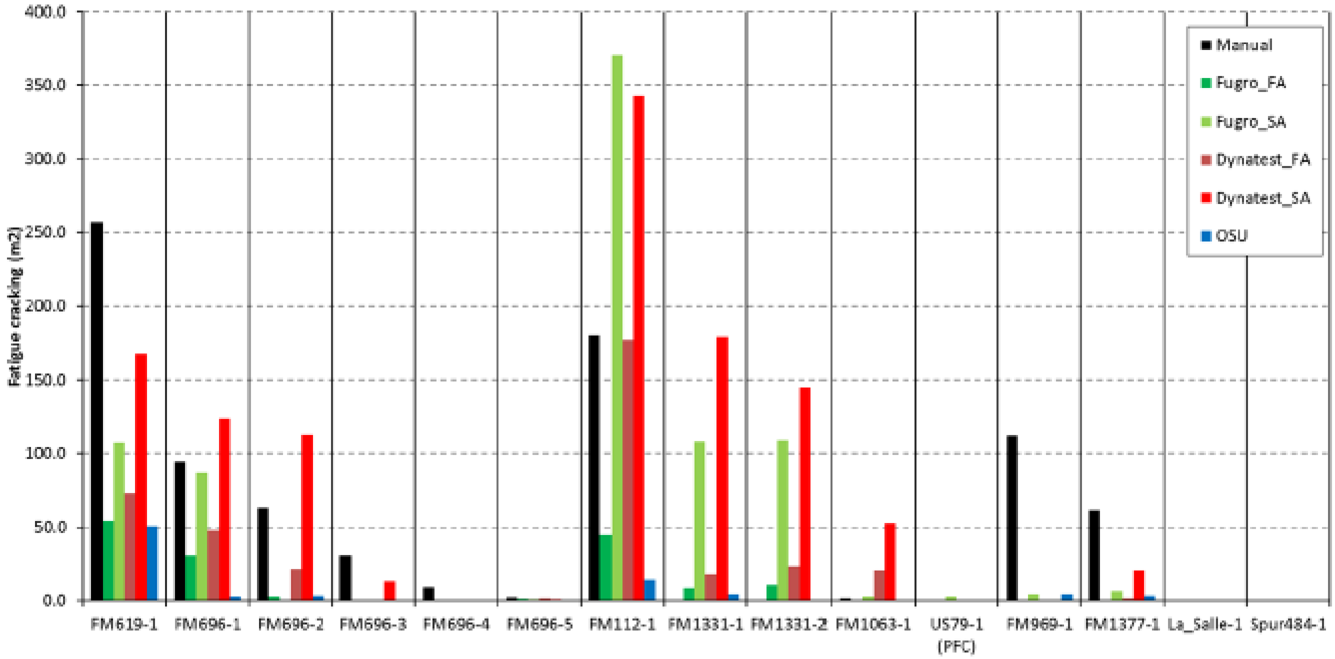


Dynatest (Semi Automated)



LTPP Manual Distress Survey

Figure A.2 Comparison of LTPP fatigue cracking on ACP sections



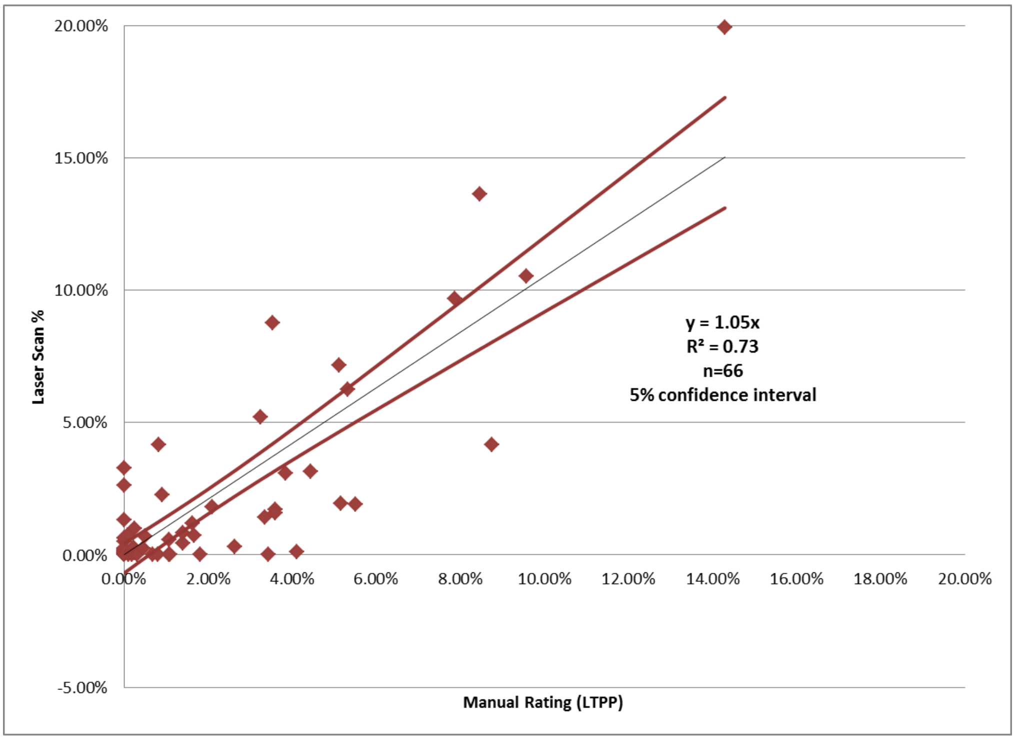
A2 Cracking – a tale of four systems

Figure A.3 RoadCrack and LCMS correlation

System	Statistics	Northbound	Southbound
RoadCrack	r^2	0.70	0.87
	Slope	0.83	0.91
	Intercept	2.8	1.4
LCMS	r^2	0.90	0.96
	Slope	0.99	0.99
	Intercept	0.2	0.5

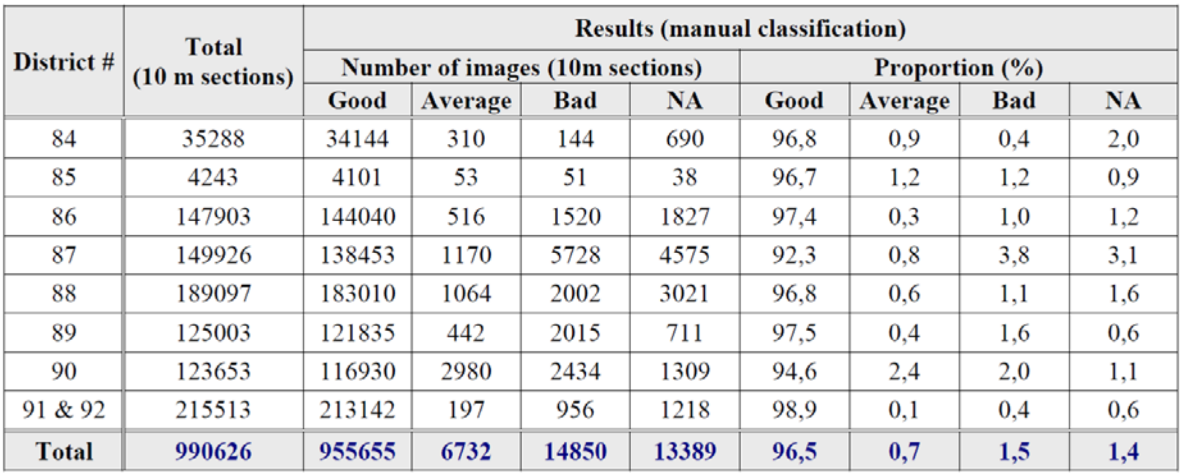
A3 New Zealand. Did we get what we wanted? – getting rid of manual condition surveys

Figure A.4 **Comparing LCMS readings with LTPP survey data**



A3.1 Canada Pavemetrics – using 3D laser profiling sensors for the automated measurement of road surface conditions (ruts, macro-texture, ravelling, cracks)

Table A.1 **10,000km automatic vs manual survey results**



District #	Total (10 m sections)	Results (manual classification)							
		Number of images (10m sections)				Proportion (%)			
		Good	Average	Bad	NA	Good	Average	Bad	NA
84	35288	34144	310	144	690	96,8	0,9	0,4	2,0
85	4243	4101	53	51	38	96,7	1,2	1,2	0,9
86	147903	144040	516	1520	1827	97,4	0,3	1,0	1,2
87	149926	138453	1170	5728	4575	92,3	0,8	3,8	3,1
88	189097	183010	1064	2002	3021	96,8	0,6	1,1	1,6
89	125003	121835	442	2015	711	97,5	0,4	1,6	0,6
90	123653	116930	2980	2434	1309	94,6	2,4	2,0	1,1
91 & 92	215513	213142	197	956	1218	98,9	0,1	0,4	0,6
Total	**990626**	**955655**	**6732**	**14850**	**13389**	**96,5**	**0,7**	**1,5**	**1,4**

Figure A.5 Repeatability results (3 passes) on two MTO road sections

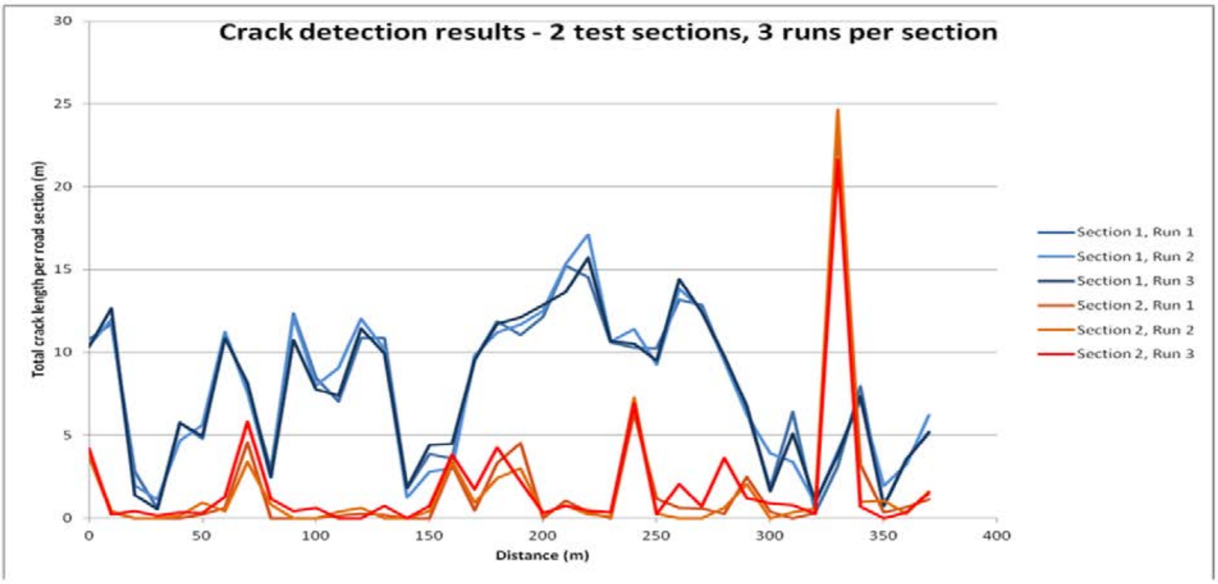


Figure 13. Repeatability results (3 passes) on two MTQ road sections.

Appendix B: Detailed field inspection results

Table B.1 Detail field inspection results Whangarei District Council

road_id	Road name	reading_ date	start_m	end_m	lane	lwp_ shove_ length	lwp_ shove_ count	rwp_ shove_ length	rwp_ shove_ count	sites to visit	Comments from field validation	Category	Outcome
615	Jordan Valley Rd	24/12/2013	3300	3400	L1	2.9	29	1.8	18	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3300	3400	R1	4.5	45	0.1	1	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3400	3500	L1	4.5	45	0.7	7	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3400	3500	R1	1.5	15	1.4	14	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3500	3600	L1	2.2	22			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3500	3600	R1	3.7	37			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3600	3700	L1	1.5	15			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3600	3700	R1	13.3	133			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley	24/12/2013	3700	3800	L1	2.5	25			1	No signs of shoving, lots of stab patches, shoving could	True Positive	POSSIBLE

road_id	Road name	reading_ date	start_m	end_m	lane	lwp_ shove_ length	lwp_ shove_ count	rwp_ shove_ length	rwp_ shove_ count	sites to visit	Comments from field validation	Category	Outcome
	Rd										have been repaired?		
615	Jordan Valley Rd	24/12/2013	3700	3800	R1	0.9	9	0.1	1	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3800	3900	L1	7.8	78			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3800	3900	R1	10	100			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3900	4000	L1	3.6	36	1.7	17	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	3900	4000	R1	2.5	25	0.1	1	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	4000	4100	L1	0.5	5	3	30	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	4000	4100	R1					1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	4100	4200	L1	3.9	39	0.1	1	1	Small amounts of shoving in the LWP, lots of stab patches	True Positive	TRUE
615	Jordan Valley Rd	24/12/2013	4100	4200	R1	0.7	7			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley	24/12/2013	4200	4300	L1	1.4	14			1	No signs of shoving, lots of stab patches, shoving could	True Positive	POSSIBLE

road_id	Road name	reading_ date	start_m	end_m	lane	lwp_ shove_ length	lwp_ shove_ count	rwp_ shove_ length	rwp_ shove_ count	sites to visit	Comments from field validation	Category	Outcome
	Rd										have been repaired?		
615	Jordan Valley Rd	24/12/2013	4200	4300	R1	0.7	7			1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	4300	4400	L1	1.5	15	0.1	1	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
615	Jordan Valley Rd	24/12/2013	4300	4400	R1	10.1	101	4.4	44	1	No signs of shoving, lots of stab patches, shoving could have been repaired?	True Positive	POSSIBLE
632	Matarau Rd	6/01/2014	500	600	L1			1.2	12	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	500	600	L1			1.2	12	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	500	600	L1			1.2	12	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	500	600	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	500	600	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	500	600	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	600	700	L1			0.2	2	1	No signs of showing, big half lane patch running up	True Positive	FALSE

road_id	Road name	reading_ date	start_m	end_m	lane	lwp_ shove_ length	lwp_ shove_ count	rwp_ shove_ length	rwp_ shove_ count	sites to visit	Comments from field validation	Category	Outcome
											the LHS		
632	Matarau Rd	6/01/2014	600	700	L1			0.2	2	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	600	700	L1			0.2	2	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	600	700	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	600	700	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	600	700	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	700	800	L1	0.2	2	0.4	4	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	700	800	L1	0.2	2	0.4	4	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	700	800	L1	0.2	2	0.4	4	1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	700	800	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	700	800	R1					1	No signs of showing, big half lane patch running up	True Positive	FALSE

road_id	Road name	reading_ date	start_m	end_m	lane	lwp_ shove_ length	lwp_ shove_ count	rwp_ shove_ length	rwp_ shove_ count	sites to visit	Comments from field validation	Category	Outcome
											the LHS		
632	Matarau Rd	6/01/2014	700	800	R1					1	No signs of showing, big half lane patch running up the LHS	True Positive	FALSE
632	Matarau Rd	6/01/2014	800	900	L1	53.9	539	4.6	46	1	Deep rutting	True Positive	FALSE
632	Matarau Rd	6/01/2014	800	900	L1	53.9	539	4.6	46	1	Deep rutting	True Positive	FALSE
632	Matarau Rd	6/01/2014	800	900	L1	53.9	539	4.6	46	1	Deep rutting	True Positive	FALSE
632	Matarau Rd	6/01/2014	800	900	R1	0.6	6			1	Shoved now, but might not have been there at the time of survey	True Positive	TRUE
632	Matarau Rd	6/01/2014	800	900	R1	0.6	6			1	Shoved now, but might not have been there at the time of survey	True Positive	TRUE
632	Matarau Rd	6/01/2014	800	900	R1	0.6	6			1	Shoved now, but might not have been there at the time of survey	True Positive	TRUE
632	Matarau Rd	6/01/2014	900	1000	L1					1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	900	1000	L1					1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	900	1000	L1					1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	900	1000	R1	0.1	1			1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	900	1000	R1	0.1	1			1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	900	1000	R1	0.1	1			1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	1000	1100	L1	2.7	27			1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	1000	1100	L1	2.7	27			1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	1000	1100	L1	2.7	27			1	No sign of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	1000	1100	R1	1.3	13			1	Signs of shoving	True Positive	FALSE

road_id	Road name	reading_ date	start_m	end_m	lane	lwp_ shove_ length	lwp_ shove_ count	rwp_ shove_ length	rwp_ shove_ count	sites to visit	Comments from field validation	Category	Outcome
632	Matarau Rd	6/01/2014	1000	1100	R1	1.3	13			1	Signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4000	4100	L1	0.9	9			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4000	4100	L1	0.9	9			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4000	4100	L1	0.9	9			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4000	4100	R1	1.2	12			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4000	4100	R1	1.2	12			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4000	4100	R1	1.2	12			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4100	4200	L1	4.2	42			1	About 10m shoving in the RWP, no signs of shoving in the LWP	True Positive	TRUE
632	Matarau Rd	6/01/2014	4100	4200	L1	4.2	42			1	About 10m shoving in the RWP, no signs of shoving in the LWP	True Positive	TRUE
632	Matarau Rd	6/01/2014	4100	4200	L1	4.2	42			1	About 10m shoving in the RWP, no signs of shoving in the LWP	True Positive	TRUE
632	Matarau Rd	6/01/2014	4100	4200	R1	2.2	22			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4100	4200	R1	2.2	22			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4100	4200	R1	2.2	22			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4200	4300	L1	0.7	7	2.3	23	1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4200	4300	L1	0.7	7	2.3	23	1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4200	4300	L1	0.7	7	2.3	23	1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4200	4300	R1	3.1	31			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4200	4300	R1	3.1	31			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4200	4300	R1	3.1	31			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4300	4400	L1	0.3	3	0.1	1	1	No signs of shoving	True Positive	FALSE

road_id	Road name	reading_ date	start_m	end_m	lane	lwp_ shove_ length	lwp_ shove_ count	rwp_ shove_ length	rwp_ shove_ count	sites to visit	Comments from field validation	Category	Outcome
632	Matarau Rd	6/01/2014	4300	4400	L1	0.3	3	0.1	1	1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4300	4400	L1	0.3	3	0.1	1	1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4300	4400	R1	0.1	1			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4300	4400	R1	0.1	1			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4300	4400	R1	0.1	1			1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4400	4500	L1	8.5	85	0.3	3	1	About 2m shoving in LWP	True Positive	TRUE
632	Matarau Rd	6/01/2014	4400	4500	L1	8.5	85	0.3	3	1	About 2m shoving in LWP	True Positive	TRUE
632	Matarau Rd	6/01/2014	4400	4500	L1	8.5	85	0.3	3	1	About 2m shoving in LWP	True Positive	TRUE
632	Matarau Rd	6/01/2014	4400	4500	R1			1.1	11	1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4400	4500	R1			1.1	11	1	No signs of shoving	True Positive	FALSE
632	Matarau Rd	6/01/2014	4400	4500	R1			1.1	11	1	No signs of shoving	True Positive	FALSE