NACOE
NATIONAL ASSET CENTRE OF EXCELLENCE

FINAL REPORT

Project Title: A5 Incorporating Uncertainty in PMS Modelling - Phase 1 (Year 1 – 2013/14 and Year 2 – 2014/15)

Project No: 007198

Authors: Dr Peter Kadar and Ranita Sen

Client: Queensland Department of Transport and Main Roads

Date: August 2016

AN INITIATIVE BY:
arrb GROUP
Queensland Government Australia

A5 INCORPORATING UNCERTAINTY IN PMS MODELLING - PHASE 1 (YEAR 1 – 2013/14 AND YEAR 2 – 2014/15)

SUMMARY

Pavement management systems (PMS) require data that faithfully reflect the properties and other operating circumstances of the network. It is a well-known, though frequently ignored, fact that much of the information is uncertain or poorly represented either due to the nature of the data (e.g. environment) or due to the aggregation of the data into disparate segments. Ignoring the uncertain nature of the input transfers the level of uncertainty to the output without acknowledging or quantifying the level of uncertainty.

The need has been identified to take the acceptable risk level (or desirable reliability) into account in the recently developed PMS. Consequently, a project was initiated under the TMR/ARRB research agreement to incorporate uncertain variables in the PMS modelling and budget forecasts.

The adopted approach expands the use of existing deterministic models by utilising the full range (distribution) of the data instead of an aggregated – usually average – representation of the full data set. The proposed approach also utilises a comprehensive set of historical data and forecasts the probability distribution of key dependent variables.

Outcomes from this proof of concept study were as follows:

1. A technique was tested to store data arrays (distributions) in a condensed form in a database and in Microsoft Excel as a text or CSV string, thus it can be stored in a single Excel cell or in a database.
2. Calculations with the condensed data sets were tested and explored.
3. Uncertain data was identified and initial efforts were made to obtain the full data set.
4. The development of a direct link library (DLL) was initiated to link the PMS (dTIMS) to an Excel spreadsheet where the probabilistic calculations will be executed.
5. A composite (Excel and dTIMS) model was developed to demonstrate the feasibility of the method, and illustrate this through a case study.
6. The case study involved running the analysis at several probability levels to allow comparison of outcomes, with the following results:
   (a) The probability that the cost calculated with the average input parameters will be sufficient to achieve the target level of service is 23%, or the risk of not meeting the targets is 77%.
   (b) Even a 50% chance of meeting target LOS can be optimistic, with less than a 40% probability of achieving the target using average inputs.

Once fully implemented, this approach will offer the following benefits:

- Investment savings of the order of 10–15% through appropriately targeted treatments, and improved efficiency resulting from better program justification and less contractual disputes.
- Improved understanding of risk levels and the implicit assumptions that may affect the outcome of any modelling.
- Greater likelihood that maintenance strategies deliver anticipated performance outcomes, having been selected with the clear understanding of the associated risks and reliability of the results.

CONTENTS

TABLES

FIGURES

1 PREAMBLE

Pavement engineers involved in asset management and performance forecasts, in particular, consider performance forecasting as an engineering activity. Whilst the engineering content is inherently part of the process, the process primarily involves forecasting. Whereas, engineering processes focus on clear, unequivocal outcomes, forecasting, by definition, cannot produce similarly unambiguous outcomes for the simple reason that the future cannot be known with 100% certainty. Consequently, there is an apparent conflict between engineering and forecasting procedures. The seemingly conflicting principles, i.e. unambiguous outcome versus qualified forecast, is fortunately only a virtual conflict that can be resolved relatively easily.

The first step is to acknowledge that all engineering design processes include safety factors. Safety factors represent the practical acknowledgement of the uncertainties related to the design process and its input parameters. Safety factors are an effective, albeit rather crude tool to cater for known and unknown uncertainties, thus providing a safety envelope around the design process.

In reality, a design process is also a forecast as it refers to a definite time-frame, e.g. many flexible roads are designed for a 20 year life. This implies that flexible roads will reach their critical condition in 20 years, which is an explicit performance forecast. This forecast is secured by a range of safety factors that multiply the assumed traffic volume substantially. For all practical purposes, the uncertainties around the design are hidden from the practising engineer by the safety factors, creating the impression that the outcome is certain and unambiguous.

Once the uncertain nature of engineering work is acknowledged, the road is open to implement suitable measures and procedures to deal with uncertainties. Acknowledging and accepting the uncertainties in engineering methods requires a mental shift away from deterministic concepts to a seemingly more complex stochastic environment, using the language of statistics. There are deep-seated prejudices against statistics, such as ‘statistics lie’, or can be manipulated. These prejudices mostly stem from a lack of understanding and capacity to interpret information provided in statistical terms. Consequently, leaving the comfort of traditional deterministic engineering approaches requires not only the introduction of the relevant technology and techniques but also assistance in learning and accepting the, for engineers, new terminology and concepts.

2 INTRODUCTION

2.1 Background

TMR is in the process of developing road asset management contracts (RAMC) for South-East Queensland, with a possible transition to outcome oriented, performance–based contracts within the next five years. Underpinning these contracts will be a comprehensive pavement management system (PMS), which employs Deighton's Total Infrastructure Management System (dTIMS) (Deighton Associates Limited 2014), which is currently being developed by ARRB and which will be used to benchmark pavement performance and optimise investment strategies.

Pavement management systems (PMS) require data that accurately reflect the properties and other operating circumstances of the network. It is a well-known, though frequently ignored, fact that much of the information is uncertain or poorly represented (the term 'uncertain' is used to describe both) – either due to the nature of the data (e.g. environment) or due to the aggregation process of representing the network condition. Ignoring the uncertain nature of the input transfers the level of uncertainty to the output without acknowledging let alone quantifying the level of uncertainty.

Deterministic models are widely employed in pavement management to forecast future performance as a trend for each analysis section based on average (or representative) input data. The trend is based on a statistical relationship determined by curve fitting to a scatter of observations. At best, this delivers a result with an approximately equal chance that the true value is either below or above the predicted value. Depending on the level of data aggregation, including chosen analysis lengths, the forecast may be grossly misleading.

The need to be able to quantify, thus manage the level of uncertainty and risk has been identified by TMR. This will become increasingly important as TMR and its contractors gain experience and move towards a second generation RAMC, and will provide the capability to:

- determine the budget required for a given level of service (LOS) with an increased degree of reliability
- incorporate the level of uncertainty (or desired reliability) in their forecast costs/price
- better account for environmental history and future predictions in long-term trends
- evaluate the robustness and risks associated with alternative strategies and funding scenarios.

To address the above issues, a project was initiated in the 2013–2014 financial year under the TMR/ARRB research agreement with the intent of developing probabilistic models to incorporate uncertainty in PMS modelling.

2.2 Objective

The objective of this project is to apply a probabilistic, quantitative, risk-based approach to the modelling of pavement performance. This will allow quantification of the degree of uncertainty of the forecast and will provide TMR and its contractors with the information needed to select a strategy that balances investment cost and risks at the desired level.

2.3 Scope of the Work

The scope of the first phase of the project comprises a proof of concept, and includes:

(a) identifying the uncertain variables

(b) exploring available data condensation (or reduction) technology

(c) investigating the potential loss of accuracy/error due to condensation
(d) trialling the condensation of the forecast distribution
(e) selecting a model for demonstration purposes
(f) implementing the data condensation technology as a DLL and demonstrating its application
(g) reviewing the findings and presenting the results.

2.4 Linkages

The project has the following linkages:

- Austroads work in this area under project *AT1064 Long-term performance monitoring* to develop consistent performance models complements this study and efficiencies will be available by avoiding duplication as this project has focused on both deterministic and probabilistic road deterioration (RD) model development
- the South-East Queensland Road Asset Management Contracts (SEQ RAMC) and associated shadow performance framework initiated by TMR in 2013
- Austroads project *AT1490 Improving the estimation of the cost of accelerated road wear due to increases in axle mass limits* which is also related to AT1064 above as both of these projects produce deterministic road deterioration (RD) models that could be used in probabilistic modelling.

2.5 Anticipated Benefits

Anticipated benefits after the completion of the project are as follows:

- users will be able to forecast performance and thus budget requirements with known, and predetermined, reliability (risk)
- investment savings to TMR of the order of 10–15% through appropriately targeted treatments, and improved efficiency resulting from better program justification and less contractual disputes. A marginal benefit cost ratio (BCR) for this research study of upwards of 30, based on annual program savings in SEQ/cost of research project
- greater likelihood that maintenance strategies deliver anticipated performance outcomes, having been selected on a more rigorous basis taking account of uncertainty in input data
- improved understanding of risk levels and implicit assumptions that may affect the outcome of any modelling
- more accurate understanding of contractors' proposals and the incorporated risks
- reduced risk in managing the network through better appreciation of the reliability (probability) of the forecasts.

3 TERMINOLOGY

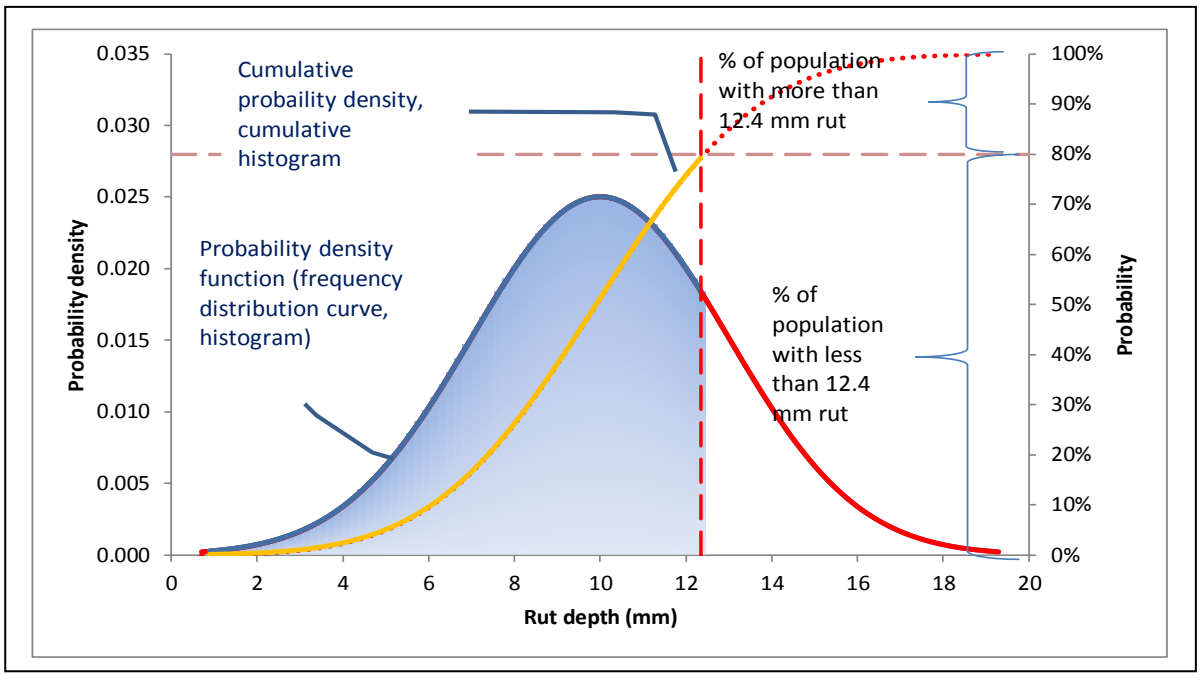
Statistical terms are frequently used interchangeably to describe various issues or events. For example, distribution, scatter and data array may equally describe the population of the data, i.e. data relevant to the same object or road section. To ensure clarity, statistical terms used in this study are defined and explained below.

3.1 Distribution

Data has a shape (Kadar et al. 2015, Savage 2009). Certain shapes can be described by an expression, such as the normal distribution, student distribution, etc. Not all data has a clearly definable shape (dispersion or scatter), in which case, no mathematical formula is available to describe the distribution. The term distribution is used here as a generic term to indicate that there is a number of data points available for a given entity, without referring to any specific pattern or shape.

Distributions are often presented graphically either as a histogram or as a cumulative histogram (Figure 3.1). The usual histogram shows the counts (number) of items in the total data set (population) falling into a specific data range. More commonly, the percentage of the population falling into one of the ranges is shown on the histogram, hence, it becomes a frequency histogram. The connected tips of the bar chart form the frequency function.

Figure 3.1: Probability and risk for predicting rutting



3.2 Percentile

The area under the frequency curve is considered 100%. The cumulative histogram, i.e. the sum of the frequencies therefore is always 100% or 1, depending on the scale selected. The proportion of the total population smaller or larger than a given value is represented either as the area under the frequency curve or as a percentile read from the cumulative histogram. Figure 3.1 shows that 80% of the population has a rut depth less than 12.4 mm, i.e. the 80th percentile of this population is 12.4 mm.

3.3 Probability

Probability is a frequently used and misused term, referring to the likelihood of the occurrence of an event. It is calculated as the ratio of the occurrence of an event and the total number of possible events. For example, the probability of throwing '2' with a six-faced dice is 1/6, as only one side of a dice is marked as '2' out of the six sides. When the distribution is known, the probability of a value larger or smaller than a given value can be ascertained from the distribution or cumulative distribution. For all practical purposes, the cumulative histogram is more suitable to determine a given probability, as it allows reading the relevant percentage directly from the 'y' axis as a percentile. For example, as shown in Figure 3.1 the probability that a member of this population will be less than 12.4 mm is 80%. It can also be stated that the probability of measuring a rut depth larger than 12.4 mm is 20%.

The concept outlined here is similar to that used in describing particle size distribution in soil mechanics, asphalt and concrete mix design. In these cases, the proportion of a given aggregate size in the mixture is defined. For example, a size 14 mm asphalt mix has 90% aggregate less than 14 mm. This can be interpreted as the probability of finding an aggregate in the mix larger than 14 mm is 10%, or the probability of picking up an aggregate less than 14 mm from the dry mix is 90%.

3.4 Risk

Risk can be defined succinctly as the probability and the magnitude of a loss or undesirable event. The definition highlights the two critical components of the term risk, namely probability and loss. Whilst the first item, probability, is a specific quantity usually mathematically definable hence absolute, loss can be rather subjective, i.e. it depends on the nature and 'owner' of the consequences. For example, if a road has a large number of potholes, e.g. a car hitting a pothole has a probability of 90%, the risk to the car owner is high maintenance and repair costs of the car. The driver's risk is spilt coffee (cleaning) and the road agency's risk is expensive road repair or even being sued for car repairs and associated costs. The focus here is on the probability, rather than on the consequences, the latter being the subject of a subsequent study. The probability of an event, e.g. damage to the car, is usually related to a tolerable level of condition say 20 mm rut depth on x per cent of the network. The proportion of the population (road network) exceeding this limit is usually considered 'at risk'.

3.5 Uncertain Variables

Variables that may have more than one value are considered uncertain. The reason for the presence of more than one value can be either an inherently unknown nature of the variable, such as the weather or related to the intent to characterise an entity, such as a road section, displaying a range of values for the same property. For example, rutting may vary along a road; this varying rut depth is usually condensed into a single value, assuming that this is representative for the whole road length. Regardless which central tendency method (average, median or mode) is used, some information is irreversibly lost. Consequently, the assumed representative values become uncertain. A more detailed discussion on specific uncertainties is provided in Section 5.3.

3.6 Deterministic and Stochastic Models

Deterministic models have singular input parameters and yield a singular, unambiguous outcome. Stochastic models use the population of an input parameter instead of a single representative number. When the input is a population (a scatter of numbers), so is the outcome. In brief, models that have distributions as input and distributions as outputs, are referred to as stochastic models. Stochastic models are also referred to as probabilistic models.

4 METHODOLOGY

Probabilistic models utilising the scatter of several input parameters have most commonly been based on Monte Carlo (MC) simulation utilising deterministic relationships. Monte Carlo simulation has been successfully used to model complex problems. Monte Carlo simulation is, however, not suitable for long-term performance simulations of a road network broken down to many smaller entities.

In a PMS, the deterioration of each parameter is modelled for each road segment for each year. This would require a very large number of Monte Carlo simulations and storing all sources (data) and calculated distributions. For a typical 10 000 km length of network with 1 km long segments using 20 variables (inclusive of KPIs) for a 20-year modelling period would mean 10 000 x 20 x 20 = 4 000 000 Monte Carlo simulations and storing the same number of distributions. Considering that each Monte Carlo simulation consists of about 1000 simulations, commonly available computing capabilities are not sufficient to cope with this demand.

When the input data is a distribution, the output of the calculation will also be a distribution. Consequently, these distributions need to be stored during and after the calculations. Commercial databases are typically not geared to storing distributions related to each element of a network, and in any case, this would make the database very large.

To overcome the identified difficulties, the following problems needed to be resolved:

- replacing Monte Carlo simulation with a procedure yielding similar outcomes but without the heavy overhead burden
- resolving the storage of the input and output distributions
- embedding the calculation into a PMS.

The project addresses each of the above issues.

4.1 Replacing Monte Carlo Simulation

Uncertain variables typically represent a population that may or may not be characterised by a known distribution formula. Instead of attempting to squeeze the data into a mathematically defined probability distribution function (PDF), it was decided to use the data as is, i.e. simply listed as a series of numbers belonging to a parameter related to a given road section. For the sake of clarity, the data set describing an uncertain variable is called a Stochastic Information Packet or SIP (Savage 2009).

SIPs can be characterised as follows:

- A SIP is defined as a list of the relevant data.
- A SIP represents the data with its true ‘as measured’ distribution without the need to find a mathematical formula to describe and represent the distribution.
- SIPs retain each member of the data set that can be recalled and interrogated later, and ensure that no data or information is lost for the current and future analyses.

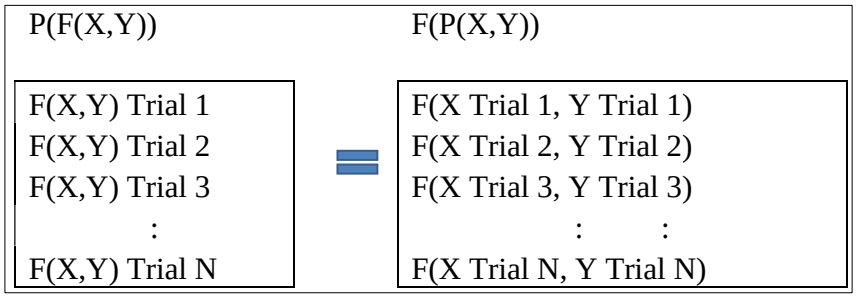
Using SIPs offers a number of advantages, among others that they remove the need to condense, and thus possibly distort the data.

The well-known CSV format used by Excel is the simplest SIP standard.

SIPs may be treated as arrays and can be used in calculations as such. The generic concept of SIP operations is shown in Figure 4.1. Operations with SIPs mean that each and every member of

the population participates in the calculation, i.e. in effect modelling is done with every data point individually. In expressions when two or more SIPs are involved, all SIPs must be of the same length. When SIPs are independent, the sequence of the data in a SIP is not relevant. However, when SIPs represent interrelated data, the sequence of these must be maintained.

Figure 4.1: Generic concept of SIP operation



Note: Where P(F(X, Y)) is the SIP of F(X, Y).

Source: Savage (2009).

Operations with SIPs can effectively replace Monte Carlo simulation, assuming that a SIP contains a large number of constituents.

4.2 Storage of Distributions in Commercial Databases

Calculations with SIPs always yield a new SIP of equal size to the original SIP; hence, the storage of the SIPs needs to be resolved. Large data stacks or arrays (SIPs) can be stored as a txt string, typically in CSV or XML format.

Besides the traditional CSV format, there are currently two concurrent standards for converting SIPs into an XML string, both yielding similar, though not quite compatible strings. The string, also called DIST or DST, depending on the standard used, occupies a single cell in Excel and can be stored in a suitable database. The various methods to convert a series of numbers into a SIP are summarised in the SIP standard (Probability Management 2014).

Both Oracle and SQL have suitable attributes that allow storing XML strings.

The standardised text string format allows storing additional information which identifies and characterises the stored data. The additional information may include the minimum, maximum and average of the data, name, description or precision. The data can be restored from the condensed format with high accuracy.

It is proposed to store all distributions in one of the text string formats in an SQL database as a navchar or XML type attribute.

4.3 Embedding the Procedure in a PMS

Implementing any procedure in industrial-strength software would require substantial programming. To reduce the cost of programming, dTIMS was selected for the purpose of this trial. The selection was based not only on the availability of dTIMS but also on the fact that dTIMS is suitable for using data stored or generated outside dTIMS. This capability has been available for numerical data but not for long text strings. The connection to the textual data has to be created by preparing a DLL to dTIMS.

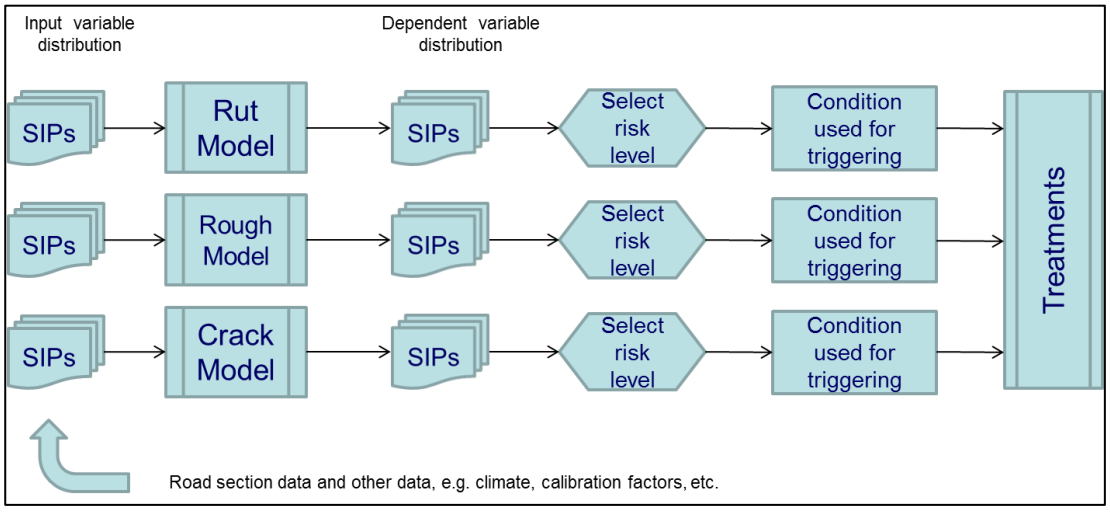
The full probabilistic analysis entails the following steps:

1. select uncertain variable(s)

2. select level of desired reliability (acceptable level of probability of some failure)
3. convert uncertain variables into text strings (SIPs)
4. store SIPs in the database
5. calculate analysis variables (deterioration) for year 1
6. extract the deteriorated parameters (rutting, cracking, etc.) at the selected level of risk
7. use the values selected in the previous step for treatment triggering
8. repeat steps 5 to 7 for subsequent years.

The concept of the process is shown in Figure 4.2

Figure 4.2: Conceptual diagram of the model



The selected deterioration models are proven deterministic models where single (typically average) input values are replaced with distributions. The desired risk level must be decided before the analysis is run and appears as an input parameter. It is feasible, and may even be desirable, to select a different risk level for each variable. For example, the variables affecting safety, such as skid resistance, may be assigned a higher risk level than those assigned to other parameters.

The conceptual design (Figure 4.2) required the following:

- storing XML strings in the dTIMS/SQL database
- dTIMS reading from and writing to an Excel spreadsheet
- the practical realisation of the above encountered issues that were not known beforehand, including:
  - the existing tables in dTIMS could not accommodate large XML strings, due to a size limitation imposed by Microsoft
  - writing to Excel is time consuming and would increase running time significantly.

Consequently, the operation of the system was reorganised in such a manner, that all modelling activities are conducted in Excel, and only the selected percentile results are brought in to dTIMS.

5 DATA PREPARATION

5.1 Data Condensation

Data condensation technology was explored and procedures developed by Thibault (2011) were adopted. The adopted procedures include data condensation and extraction technologies as well as a set of macros to conduct calculations with SIPs in Excel.

After some experimentation, an advanced CSV format was selected that includes some description of the condensed data as well as allowing the user to specify the precision required. The SIP format must follow some standard as retrieving the data and operations with SIPs may depend on the standard.

The condensation technology provides significant collateral benefits, as it allows not only efficient data storage, but also includes efficient data extraction and presentation technologies.

The calculations require that all SIPs have the same length (number of elements). As the source distributions may contain vastly different numbers of elements, resampling of the data must be used to make all SIPs of the same length. Statistical resampling may be based on random resampling or interpolation between the source numbers. This latter method was adopted and used for all source distribution. For phase 1 work, a SIP length of 1270 was used. An example SIP is presented in Appendix B.1.

5.2 Model Selection

For demonstrating the viability of modelling with uncertain numbers, the Austroads pavement deterioration models were selected (Austroads 2010a and Austroads 2010b). These models are well documented and are suitable for implementing in Excel.

5.3 Uncertain Variables in Road Deterioration Modelling

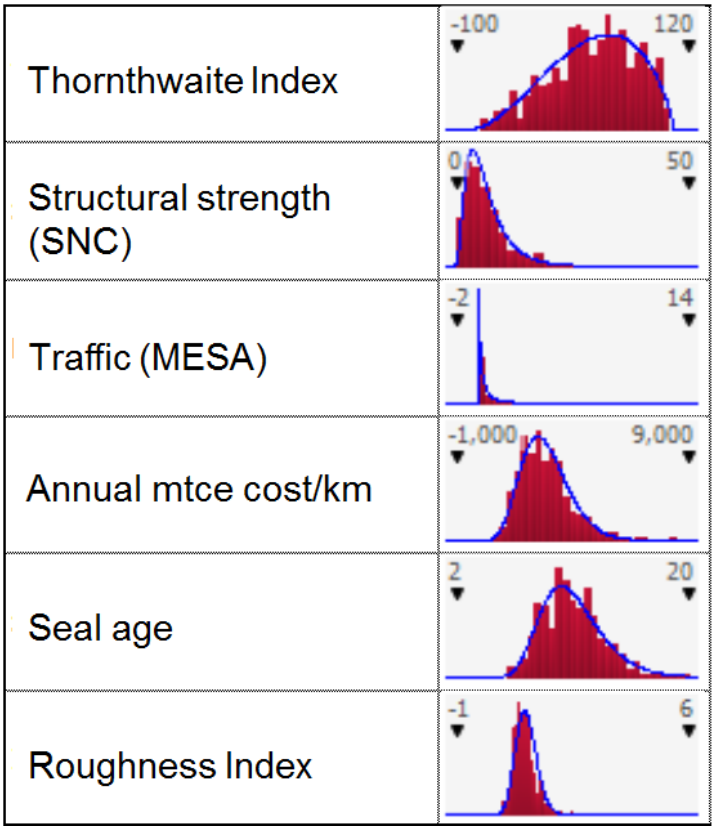
It is generally assumed that the data used for managing assets is a true representation of the asset's properties, condition and the operating environment. In reality, the input parameters are far from being definitive and thus contain a degree of uncertainty. Uncertainty of the input data stems from at least one of the following circumstances:

- Information Quality Level (IQL); the quality of the collected or available data is critical when considering uncertainty. Detailed network level data (IQL-3 and IQL-4) – e.g. collected with high speed electronic devices – will have a lower level of inherent uncertainty than data obtained at IQL-5, e.g. by visual observation or perusing records (Paterson and Scullion 1990).
- Data aggregation; this is required to generate a single representative number from the detailed data for a road section. Most frequently, the average of the data is used for representing the property of a section. Averaging data may misrepresent the nature of the data and in most cases dismisses valuable information on the spread of the data. Actual failure or reduced LOS of an asset is related to the worst and not to the average condition, hence the spread of the data and its extreme values are critical.
- Environmental data is a special case of data aggregation as valuable time series data is compressed into a single parameter such as the weighted annual mean pavement temperature (WAMPT). However, pavement performance is related to the environment at the time of the application of the traffic load, hence even a weighted average of temperature, rainfall etc. can distort the outcome.

- Cost information – unit rates – these tend to vary significantly, depending on several factors ranging from the location and timing of the contracts to the type of contract, and in the adoption in many cases of average treatment costs that do not reflect the variation resulting from designed treatments.
- Insufficient or estimated data is uncertain by definition. Typically, traffic forecast and annual average daily traffic (AADT) data fall into this category as these are calculated from sampling the traffic at various locations and times. Where no data is available, alternatives may be sought to obtain at least approximate – thus uncertain – information (Hubbard 2010).
- Material properties and construction quality are estimated at best but more often than not are assumed only based on local experience.
- Calibration factors are required for most models. Calibration is usually conducted at selected locations and the results are than projected for all similar pavements. This process assumes that the selected locations and thus the calibration is representative of the sub-network. In reality, the derived calibration factors display a scatter (distribution) that should be taken into account.

The analytical work conducted in Austroads project AT1064 (Austroads 2013 & Austroads 2015) indicated a large variety of probability distribution functions (PDFs) for typical pavement deterioration input variables. The variables comprised Thornthwaite Moisture Index (TMI), initial modified structural number for the pavement/subgrade (SNC_0), traffic load (MESA), annual average maintenance expenditure (ME), seal life (seal age) and initial roughness (IRI) as shown in Figure 5.1.

Figure 5.1: Probability distribution functions (PDF) of selected input parameters to pavement deterioration models



Monte Carlo simulation requires representing each set of uncertain data with a closed formula. As the data and its distribution may be different for each road segment, this would be a time consuming process at its best, besides being impractical, as it requires human intervention.

The adopted method does not require this step, it simply uses the data in a SIP format, and therefore it can be easily automated.

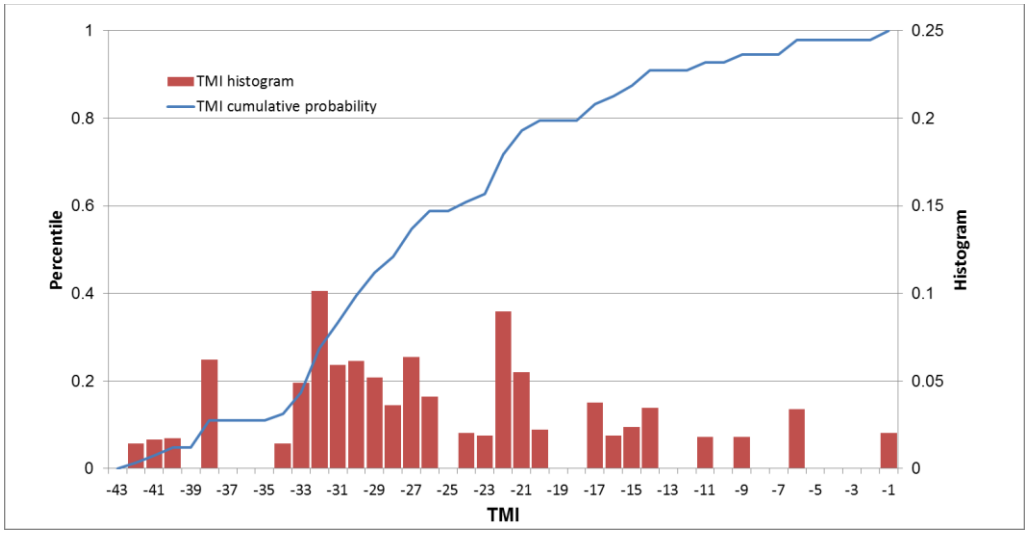
As a part of the current work, the Austroads rutting model has been adopted with the following uncertain variables:

- Thornthwaite Moisture Index (TMI)
- structural strength (SNC_0)
- maintenance expenditure (ME).

5.3.1 The Thornthwaite Moisture Index

Evapotranspiration data for 50 years was used to calculate the annual TMI. The variation over the 50-year period was significant and highlighted the magnitude of simplification and distortion when the average TMI is used (Figure 5.2).

Figure 5.2: Thornthwaite Moisture Index distribution

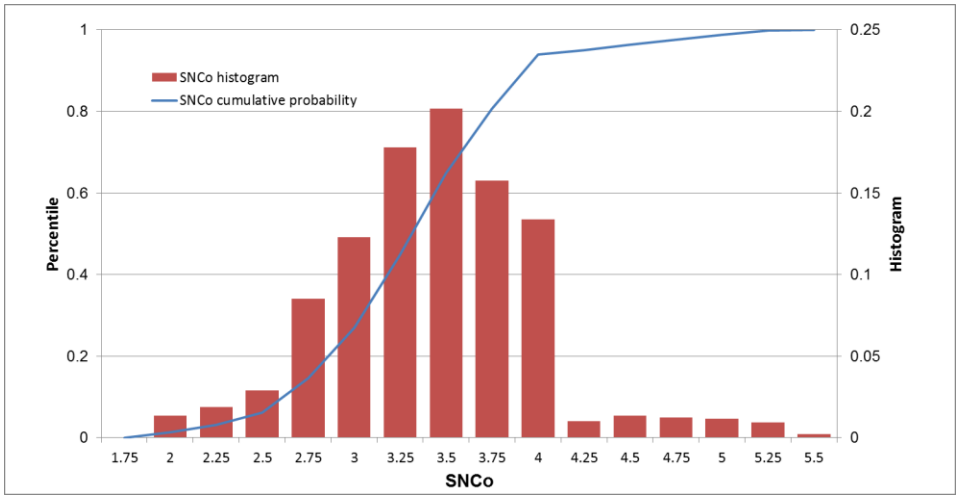


The calculation method is summarised in Appendix B.2. In the subsequent calculations, the full distribution of the TMI was used, resampled to 1270 elements.

5.3.2 Structural Strength (SNC_0)

The distribution of SNC_0 was estimated using empirical information. The assumed distribution is shown in Figure 5.3.

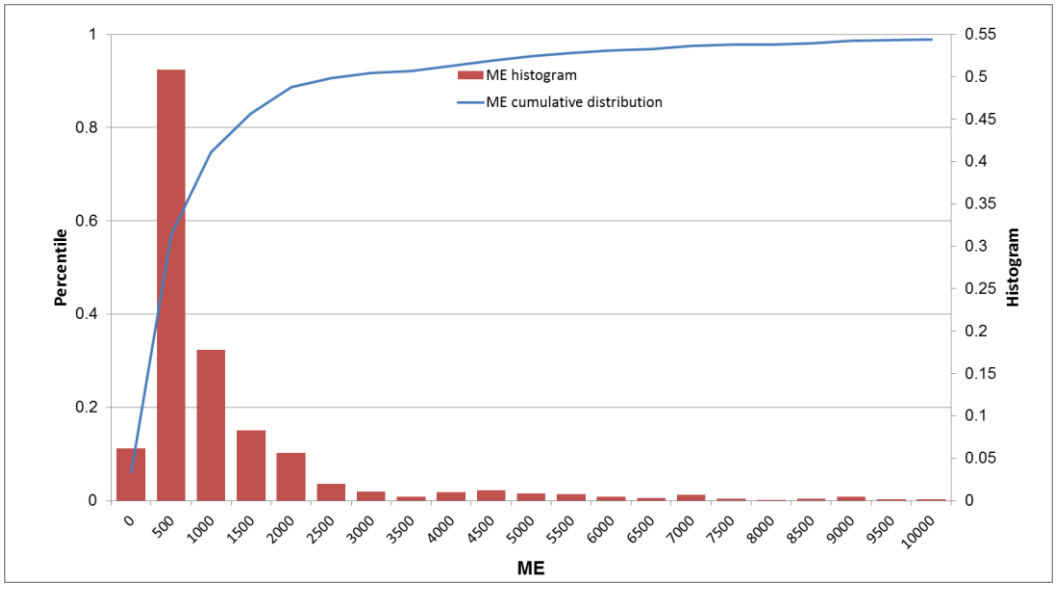
Figure 5.3: SNC_0 distribution



5.3.3 *Maintenance Expenditure (ME)*

Maintenance expenditure data was obtained from TMR. The data span a huge range (from $50 to $10 000 per km per year) (Figure 5.4). As about half of the 170 000 data items were less than $500, all items less than $300 were excluded from the further calculations. The remaining approximately 80 000 data points were resampled to form the SIP containing 1270 items.

Figure 5.4: Distribution of maintenance expenditure (ME)



5.3.4 *Model Calculations*

The conventional spreadsheet model was converted to SIP arithmetic by using bespoke functions (Thibault 2011) replacing the standard Excel operations. These functions are contained in an Excel add-in that enables operations with SIPs. Every operator was replaced with an equivalent SIP operator that executed the same calculation but with arrays (SIPs). The SIP operators and their functions are published in Thibault (2013), so they are not discussed in detail here. The SIP operations will be referred to as SDXL functions.

The general syntax of the SDXL functions is *sdFunction* (*distribution1, distribution2*), where *sd* refers to the add-in, *Function* refers to the operator and the distributions refer to the arrays (SIPs)

to be used by the function. For example, adding two distributions will be conducted with the following function: *sdAdd* (*distribution1*, *distribution2*), where *distribution1* and *distribution2* can be a range of cells containing the data, a named range or reference to a cell containing a condensed SIP.

Functions are available for most operators in Excel. To maintain clarity and transparency, complex expressions of the Austroads models were transcripted in several steps to SDXL, making the Excel spreadsheet look more complex than it is in reality. Examples of the converted formulas are given in Table A 3.

5.3.5 Input Data

A small test network consisting of 20 sections of a flexible road was created for this proof of concept study. The input parameters are shown in Table 5.1.

Table 5.1: Input parameters

Parameter	Type	Scope	Source
Initial structural number (SNC_0)	Distribution	Global	Estimated
Thornthwaite Moisture Index (TMI)	Distribution	Global	Estimated
Maintenance expenditure (*me*)	Distribution	Global	Actual
Annual millions of equivalent standard axle loads (MESA)	Scalar	Global	Assumed
Cracked area (%)	Distribution	Section	Actual
Rut depth (mm)	Distribution	Section	Actual
Roughness (IRI)	Distribution	Section	Actual

The input data is summarised in Table 5.2 and Table 5.3. The assumed data was generated using random numbers according to typical patterns experienced on Australian rural networks as found from the Austroads LTPP and LTPPM databases (Austroads 2016 & Kadar et al. 2015).

Table 5.2: Global input parameters

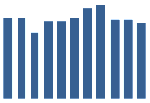
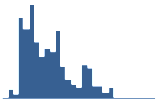
	Minimum	Maximum	Average	Distribution
SNC_0	1.77	5.32	3.3	
TMI	20.0	30	25.1	
me p.a. ($/km)	0	69 594	1139.5	
MESA	0.5	0.5	0.5	n/a

Table 5.3: Sample condition input parameters

	Minimum	Maximum	Average	Distribution
Cracking (% area)	0.5	8.5	3.2	
Rutting (mm)	2.0	16.1	6.0	
Roughness (IRI)	1.26	4.91	2.8	

One of the underlying requirements for SIP operations is that all SIPs must contain the same number of elements. As the source data rarely met this requirement, the length of the SIPs had to be resized. This was achieved by using the *sdResize* (SIP, n) function that adjusts the number of items in the SIP to *n* without altering the shape of the distribution. There are other functions also available, such as conventional bootstrapping, which allows successive, more complex programming to be employed.

5.4 Deterioration Predictions

5.4.1 Rutting

The rut depth progression for selected probabilities (percentiles) is illustrated in Figure 5.5. The average trend is slightly above the 50th percentile, indicating that the average is larger than the median. The cumulative distribution of rutting in years 1 and 10 gives insight into the problems of using just one central tendency indicator, be it the average or median (Figure 5.6). Assuming a trigger limit of 10 mm, the road section has not reached this trigger limit after ten years when the average value is used. Yet in year 10, about 30% of the road is beyond the trigger limit and about 5% has rutting larger than 14 mm, trending towards a dangerous rut depth.

Figure 5.5: Rut depth progression

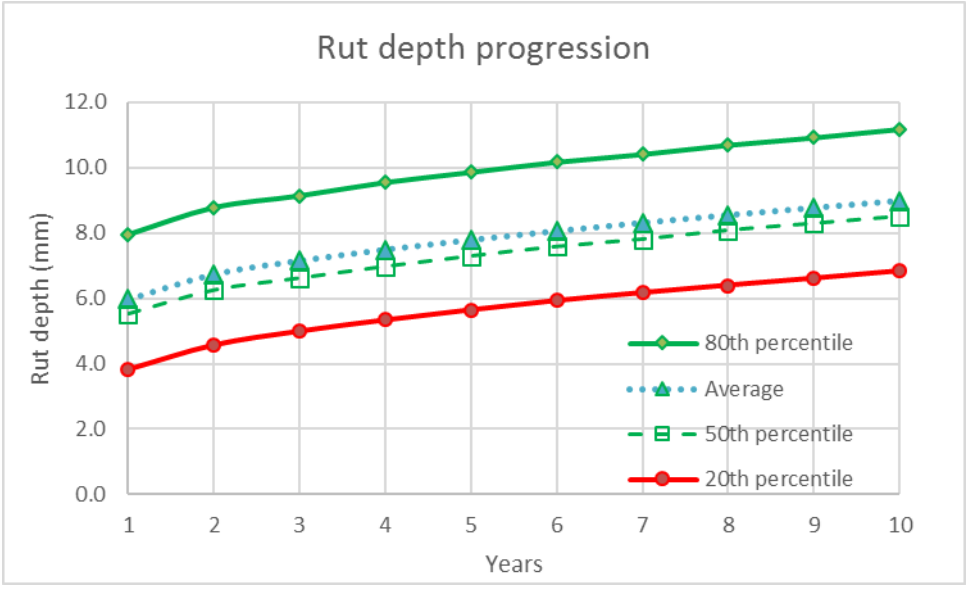
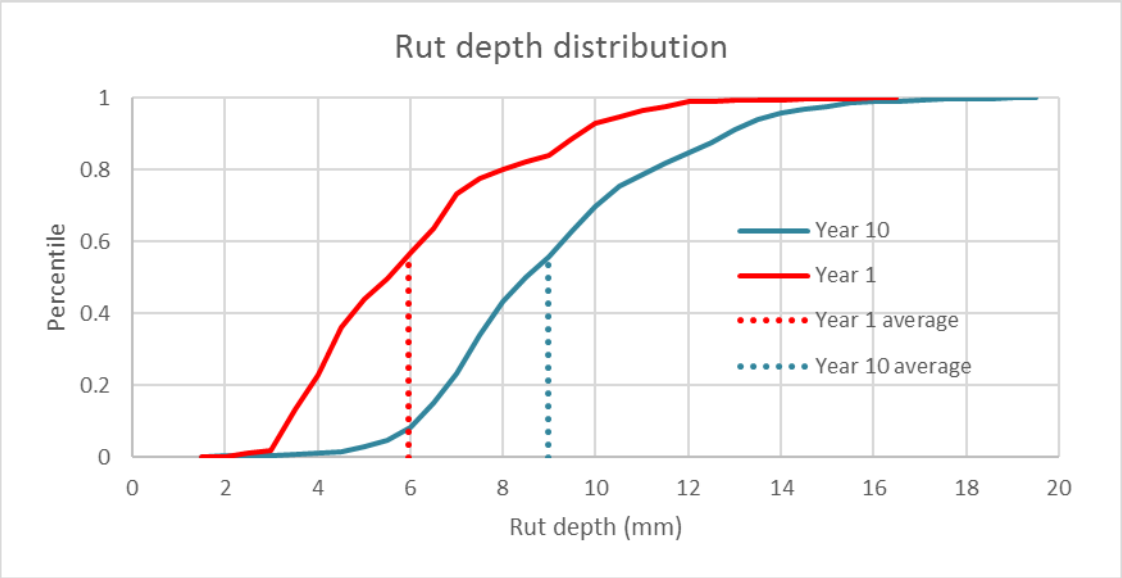


Figure 5.6: Rut depth distribution



5.4.2 Roughness

The roughness progression for selected probabilities (percentiles) is illustrated in Figure 5.7. The average trend is very close to the 50th percentile, though not exactly the same. The cumulative distribution of roughness in years 1 and 10 gives insight into the problems of using just one central tendency indicator, be it the average or median (Figure 5.8). Assuming a trigger limit of 6 IRI, the road section has not reached this trigger limit after ten years when the average is used. However, in year 10, about 20% of the road is beyond the trigger limit, as observed from the distribution.

Figure 5.7: Roughness progression

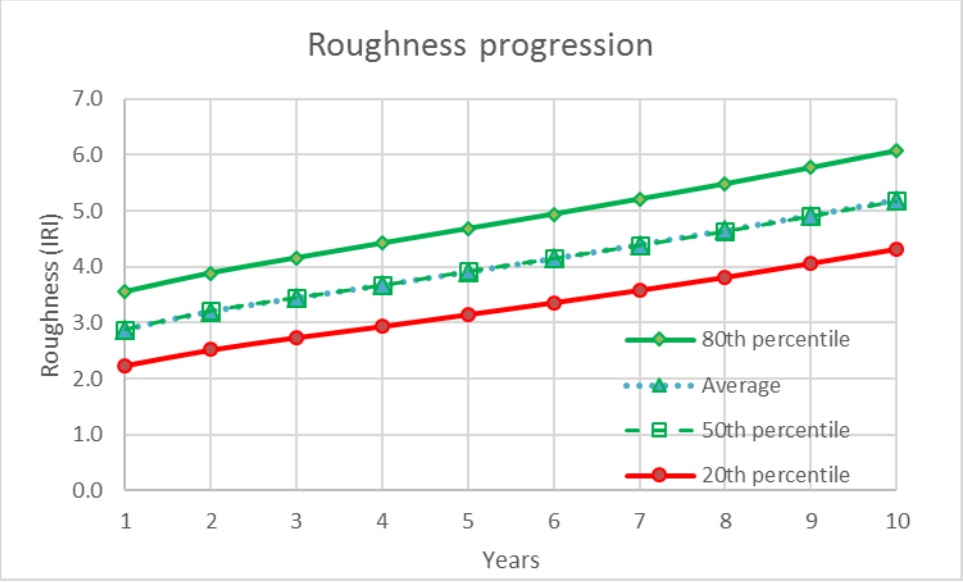
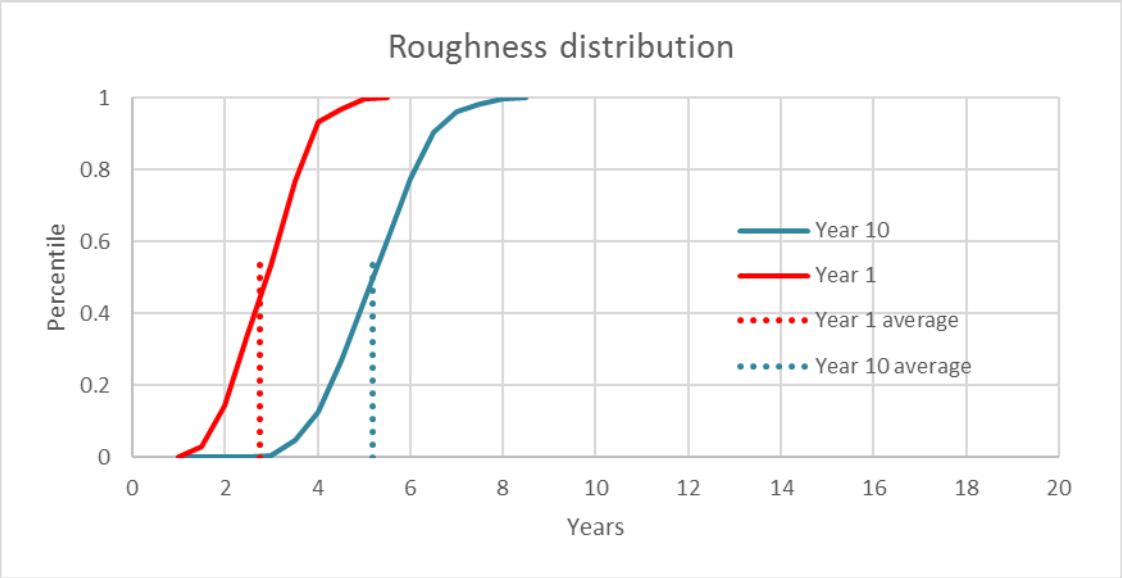


Figure 5.8: Roughness distribution



5.4.3 Transfer to dTIMS

The desired reliability/risk/probability level is selected in the spreadsheet before starting the analysis and the appropriate percentile values are summarised in a table. dTIMS has suitable functions that can read data from an external source, in this case from the Excel table. The designated dTIMS function looks for numerical data; hence, the calculated values must be copied to a selected location, otherwise the Excel functions would be treated as text. dTIMS will read the relevant data for each element (road section) in each year. Consequently, modelling uncertain variables in dTIMS is superfluous, as the predicted results will be transferred from Excel.

5.5 Treatment Selection and Works Effects

Treatments are selected by comparing the annual condition variables imported into dTIMS to the trigger condition. This operation is the same as the usual operation of the dTIMS PMS.

After applying a treatment, the condition variable is changed to reflect the effects of the work. For the calculation of benefits, the works effect (WE) must be scalar, similar to the trigger value. For SIP variables, the works effect can be either scalar or an array. If scalar is selected, no later distribution can be calculated and the WE will be controlled from within dTIMS. Alternatively, the reset (WE) can also be a distribution so that the distribution of the variable(s) can be calculated in future years. For the proof of concept study, this option was selected. The reset or works effect distributions are contained in the spreadsheet. The appropriate SIP is used according to the treatment selected by the dTIMS treatment selection (optimisation) process. In the pilot study, different post-work distributions are assumed after resealing and rehabilitation.

The scalar value for the calculation of the benefit is determined as a percentile or probability level from the WE SIP, similar to that of the roughness and rutting deterioration.

The WE distribution should reflect the quality of workmanship. In a perfect world, the resulting condition would be uniform, i.e. the scatter would be negligible and the cumulative distribution curve would be near or completely vertical. For the pilot study, a modest scatter was assumed.

5.6 Optimisation

The optimised budget allocation requires an objective function, in this case the Pavement Condition Index (PCI). The PCI is calculated as an aggregate of various condition indexes used in

dTIMS. As part of the proof of concept study, the SIP calculations are external to dTIMS together with the storage of SIPs and direct calculation of the PCI is not currently feasible. Calculation of the distribution of derived or dependent variables, such as benefit, cost, PCI, etc., was not contemplated in the proof of concept study. The calculation of the distribution of these will be feasible when SIP operations can be conducted inside dTIMS without using external spreadsheet(s).

For the proof of concept study, the PCI and other dependent variables are calculated at the preselected probability level. To generate the distribution of the PCI, dTIMS modelling has to be repeated at a number of probability levels. Each run yields a budget and associated PCI at the preselected reliability level. The results of these runs can then be combined to present the distribution of the PCI or any other variable.

5.7 Spreadsheet Model

The spreadsheet model and the detailed instructions how to use it are in Appendix A.

6 RESULTS OF THE PROOF OF CONCEPT STUDY

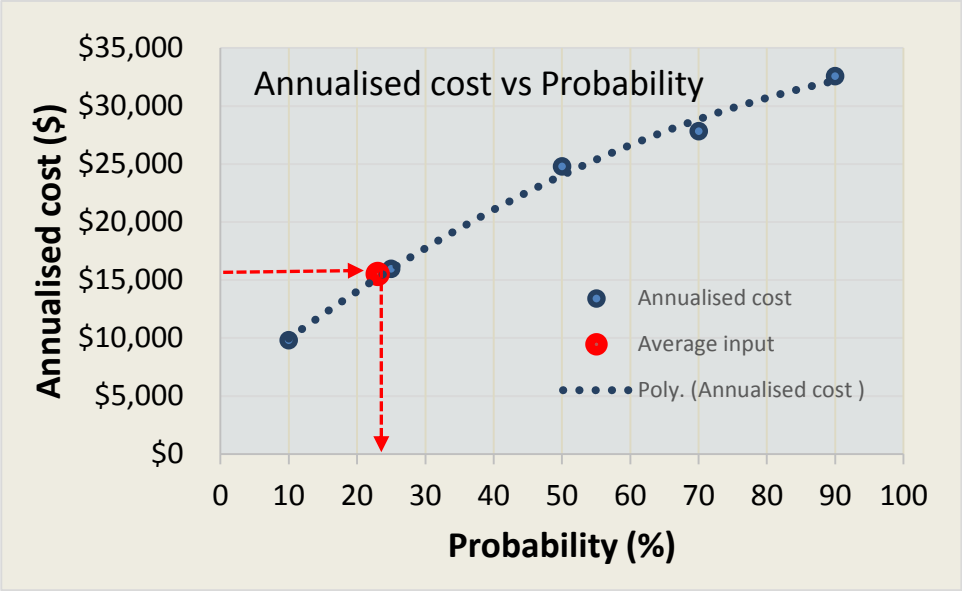
6.1 Proof of Concept

The main objective of this proof of concept study was to illustrate the potential use and benefit of considering the scatter and uncertainty of the input data. To illustrate the use of the process, a cost-probability curve was developed, illustrating that the costs increase with increasing reliability (certainty).

The costs were calculated assuming an unlimited budget, i.e. an ideal situation when no funding restrictions apply. The resulting costs are the maximum possible required to achieve the target condition determined by the triggers. dTIMS was run at several probability levels, that is, the costs to achieve the same condition, but with different reliabilities, were calculated. As the cost requirements may change from year to year, the annualised cost was used for comparison.

Costs were calculated for 10, 25, 50, 75 and 90% probability levels (Figure 6.1). Costs were also calculated using the average values for all input parameters and presented on the same chart.

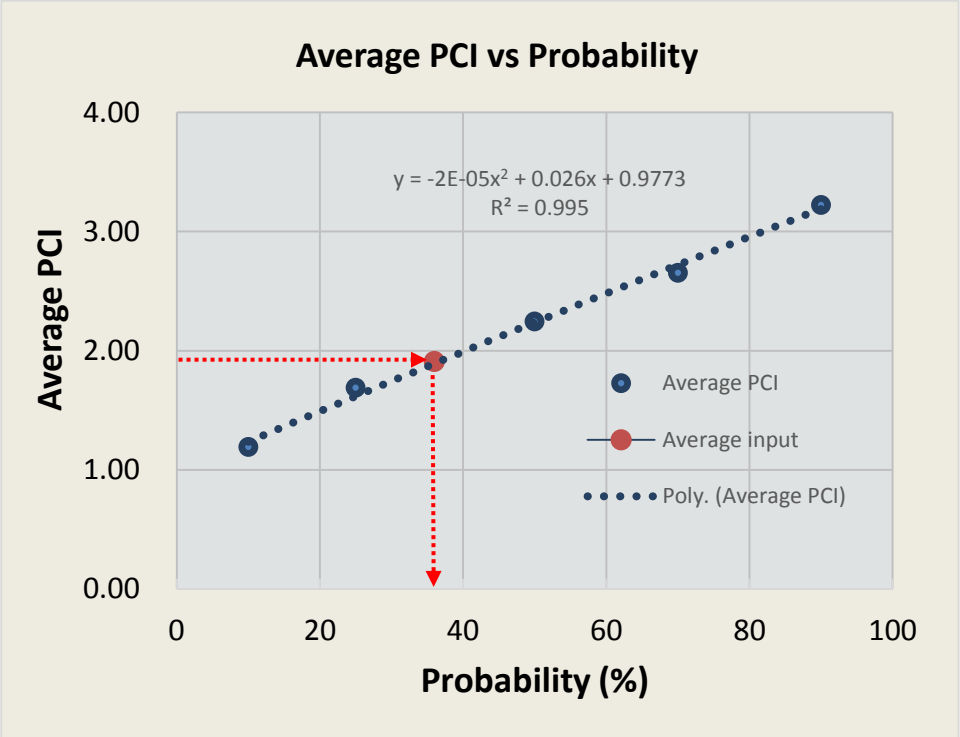
Figure 6.1: Annualised cost versus probability



Based on Figure 6.1, the probability that the cost calculated with the average input parameters will be sufficient to achieve the LoS defined by the trigger levels is 23%. This means that the chance that the funding level determined with the average input parameters will be sufficient is very low (23%) or the risk of not meeting the targets is 77%, that is, quite high.

The trial was repeated with a constrained budget. In this case, the budget was kept constant and the 10-year average pavement condition index (PCI) calculated at various probability levels was plotted against the probabilities (Figure 6.2). The PCI calculated with the average input (PCI = 1.91) fits into the PCI curve at about the 36% probability level, indicating that at the given funding level, the probability of achieving a PCI of 1.91 is about 36%.

Figure 6.2: Average PCI versus probability



The results of both cases prove that it is possible to generate a probability curve that can be used to determine the probability of the outcome. The probability curve is highly dependent on a range of factors related to both the input distributions and the definition of the outcome. As the outcome does vary from year to year, that is, the PCI and the budget demand may vary from year to year, a composite measure had to be calculated to generate the probability curve. Obviously, the definition of the composite outcome will affect the shape of the probability curve. This issue can be eliminated by moving all SIP calculations into dTIMS, as the probability curve would be then generated directly, without post-processing the results into a composite number.

In reviewing the proof of concept method, there is the question of whether it would not be sufficient to use a percentile value of the input parameters, and thus the seemingly complicated SIP calculations would not be necessary. There are at least two reasons for not doing this:

1. When operating with distributions, the resulting distribution will be different; consequently, the probability of the outcome will not be known using the above method.
2. Percentiles are usually calculated from the average and standard deviation; this implicitly assumes a normal distribution that is unlikely to be case.

6.2 Additional Benefits

Predicting the distributions of performance parameters has additional benefits. The distribution in any given year will allow determination of the proportion of the road in an unacceptable condition. The knowledge of the length of roads in unacceptable condition has several advantages, such as:

- The treatment of these identified sections can be planned as opposed to treating these from the reactive or routine maintenance budget.
- The budget can be determined more accurately, i.e. the reliability and accuracy of the budget planning will improve.

- The uniformity of treatment sections will not be particularly important. Uniformity is usually required to ensure the selection of the right treatment at the right time, as uniformity implies relatively limited scatter of the data. This does not matter if the scatter of the data is taken into account in the calculations.
- Relatively large sections or even sub-networks can be analysed without losing accuracy, hence processing time may be materially reduced.

7 RECOMMENDATIONS AND FUTURE WORK

The proof of concept study demonstrated that it is feasible to determine the probability of the outcome of a complex, multi-year works program and budget development process. The proof of concept study used a combination of Microsoft Excel and dTIMS to achieve the desired outcome. This process, however, needs to be developed further to gain robustness suitable for application to large-scale road networks.

The proof of concept study assisted in identifying tasks required to improve the process. The tasks fall into two broad categories, namely software and data preparation. These are briefly discussed below.

7.1 Software Issues

The following issues need to be addressed.

7.1.1 Storage of SIPs

SIPs or distributions are technically XML (text) strings. Both SQL and Oracle have appropriate data types, so storage for large-scale road network databases is readily available. XML strings store distributions and are used and generated during calculations, so storage of these must be available during run time.

The current database structure of dTIMS is not suitable for storing SIPs consisting of more than 256 characters. For all practical purposes, a SIP should contain 100–1000 numbers or up to 2000–8000 characters, so the current limitation in dTIMS is critical. This limitation is related to the fact that the most relevant data tables are already fully utilised and significantly, more items cannot be fitted without further size limitations. This issue can be overcome by storing the SIPs in a new table in the database and dTIMS takes the relevant data from this table. This solution may require additional programming in SQL server and dTIMS.

7.1.2 SIP Operations in dTIMS

dTIMS, like all other PMS, operates with scalars. The SIP operations currently utilise small subroutines based on existing Excel functions. These are embedded into an Excel add-in that can be attached to an Excel spreadsheet. The dynamic library (DLL) can be called from Excel with well-documented functions.

A trial dynamic library (DLL) was developed to explore and evaluate the effort required to implement SIP operations. It was found that the relevant functions could be created with relatively limited effort. The most efficient solution seems to be to develop a dynamic library as an add-in to dTIMS, so SIP operations can be conducted in the future by simply calling an external function. This simplifies coding and at the same time reduces run time. The precondition of this is to store the input and outcome XML strings in the database attached to dTIMS.

Most recently, two functions were created to conduct SIP arithmetic in dTIMS. Whilst the functions operated successfully, due to the limitations discussed earlier only extremely short SIPs could be used. This small test – that went beyond the original scope of the work – proved the feasibility of this approach.

The most critical limitation is the size of existing tables in dTIMS that are close to exhausting the maximum table size in the SQL server. This may be overcome by two different approaches, namely:

- modifying the internal database structure of dTIMS
- performing the relevant calculations in a new piece of software that is linked to dTIMS.

Modifying a complex software is not practical or even possible without the cooperation of the software's owner, in this case Deighton Associates Limited (DAL). Consequently, this avenue may require additional effort besides solving the technical issues. However, cooperation with DAL may assist in solving the technical issues and may bring commercial benefits.

The second option, i.e. creating a universally usable DLL, has the appeal of software independence. Considering the complexity of the task, the DLL will require frequent interaction with the software. Consequently, its use will be limited to software packages providing specific features allowing run time interaction.

7.2 Development of SIP Databases

The source data typically used for PMS analysis is provided in short (5–100 m) road length segments. The data for larger segments can be generated from this by creating SIPs. SIPs can be generated during processing of the data, or can be generated in the source database. Both options are valid, though it is more likely to succeed with the first option initially. This step is not an underlying condition of implementing SIP calculations in a PMS, but it offers advantages on its own.

A SIP database contains the full source data in a condensed format. Current databases, such as ARMIS contain the data aggregated according to a specific segmentation, i.e. typically 100 m for roughness or 20 m for rutting, etc. A SIP database could store all data, without aggregation for a segment. In this case, the segment length would only depend on the database capacity to store a text string, i.e. the maximum allowable size of the text string.

When using an undefined segmentation length, i.e. this is determined after extracting the data, it must be ensured that locations are aligned with the data. This can be ensured by maintaining the sequence of the data in SIPs. As SIP arithmetic is sensitive to the sequence of the data, this issue is critical for future applications.

7.3 Data Preparation

Data preparation is both simpler and more complex when using SIPs. It is simpler as creation of uniform sections is a lesser requirement. However, it is complex due to the large variety of the data.

Some data items are not currently collected. For example, the spread of unit cost rates is usually unknown or based on anecdotal evidence at best. Collecting, organising and storing this data is essential.

Other data items are collected but stored in an aggregated format only. Traffic data is typically provided as an annual aggregate (average). Instead, a traffic spread would be more useful and probably more representative.

A careful evaluation of all data is necessary to weigh the importance and significance of using it in an aggregated or distribution format. This requires a number of trial runs and sensitivity studies.

7.4 Training and Education

It must be acknowledged that engineers are not accustomed to think in terms of distributions, risk and probability. Consequently, the availability of suitable easy-to-use tools is not enough. Extensive training and broadening the attitude towards thinking in this direction is required.

This task is perceived the most difficult, as it is not a technical issue, but dealing with human nature. Education and training is necessary to have the objectives understood and the results broadly accepted.

8 SUMMARY AND ACHIEVEMENTS

Pavement management systems (PMS) require data that faithfully reflect the properties and other operating circumstances of the network. It is a well known, though frequently ignored, fact that much of the information is uncertain or poorly represented either due to the nature of the data (e.g. environment) or due to the aggregation of the data into disparate segments. Ignoring the uncertain nature of the input transfers the level of uncertainty to the output without acknowledging let alone quantifying the level of uncertainty.

The need has been identified to take the acceptable risk level (or desirable reliability) into account in the recently developed PMS system. Consequently, a project was initiated in the 2013–2014 financial year under the TMR/ARRB research agreement to incorporate uncertain variables in the PMS modelling and budget forecasts.

The adopted approach expands existing models by utilising the full range (distribution) of the data instead of an aggregated – usually average – representation of the full data set.

The proposed approach utilises the full set of historical data and forecasts the probability distribution of key variables.

Outcomes from the first phase of the project are as follows:

1. A technique was tested to store data arrays (distributions) in a condensed form in a database and Excel. A data array, i.e. the full data set can now be condensed into a text or CSV string, and thus can be stored in a single Excel cell or in a database.
2. Calculations with the condensed data sets were tested and explored.
3. Uncertain data was identified and initial efforts were made to obtain the full data set.
4. The development of a DLL was initiated to link dTIMS to an Excel spreadsheet where the probabilistic calculations will be executed.
5. A composite (Excel and dTIMS) model was developed to demonstrate the feasibility of the method.

Once fully implemented, this approach will offer the following benefits to TMR:

- investment savings of the order of 10-15% through appropriately targeted treatments, and improved efficiency resulting from better program justification and less contractual disputes
- improved understanding of risk levels and the implicit assumptions that may affect the outcome of any modelling
- greater likelihood that maintenance strategies deliver anticipated performance outcomes, having been selected with the clear understanding of the associated risks and reliability of the results.

The following achievements have been made during the first phase of the project.

Data condensation technology

The overall approach has been clarified and two technologies sourced and preliminary testing undertaken, including:

(a) creation of stochastic information packages (SIPs) as described in Section 4.2 to represent observed data initially within Excel

(b) arithmetic operations with SIPs were implemented in Excel and used to illustrate the difference between deterministic and stochastic models.

Prototype model

A rutting model was adapted to work with SIPs and a ten-year forecast was generated using Excel. The results indicated the impact of asymmetric data distribution, i.e. the most common and typical case. The deterministic models seem to be realistic only when the input data represents symmetrical distributions. As this is rarely the case, as indicated by research project AT 1604 (Austroads 2013, Austroads 2015 & Austroads 2016), deterministic models have limitations in being able to yield reliable results when used alone.

DLL development

A suitable DLL has been developed. The working of the DLL and impact on computing time and efficiency of the software will only be known after the DLL is finalised.

Risk management

One of the main risks is the unavailability of sufficient data at a suitably high resolution to define the distribution of key input parameters. TMR's state-wide asset management team, the regions and the SEQ RAMC project should be encouraged to invest in the required level of data collection to allow a detailed base line to be established. This has largely taken place, and will be improved with time through better feedback from the RAMC and following the deployment of the Traffic Speed Deflectometer (TSD).

Models prove to be insensitive to varying inputs. Calibration of the critical models associated with pavement structural cracking, rutting, roughness and strength deterioration has already taken place and can be updated on a regular basis during the maintenance contracting period as more time series data becomes available. A NACOE study has also commenced with the aim of adapting and calibrating the Austroads AT1064 models to Queensland conditions (Noya & Marzouk 2016).

Model development

The availability of modelling distributions as opposed to a 'representative' value opens up the possibility to develop new models based on the shape of the distribution as opposed to the average. Forecasting the shape of the distribution will yield the so far elusive point of rapid deterioration, i.e. when the rate of increase of deterioration accelerates, indicating the near failure of the pavement.

Forecasting the full distribution also opens up the way to more accurate budgeting, as the proportion of the asset in need of maintenance can be estimated. This allows calculating the extent of pavement damage and subsequent repair needs.

Next steps:

1. On data:
    (a) To explore/investigate which data is the most critical for the outcome. This is a sensitivity analysis that can be conducted with the current prototype model.
    (b) To develop global SIPs for environmental data, e.g. TMI. This has been explored during the current phase but needs to be expanded. Global SIPs are independent from the annual condition data.
    (c) To introduce new data, namely TSD. This may require modifying the prototype model. However, this modification is of a minor nature.

2. Further development of the prototype model
    (a) The current prototype model abandons one of the initial assumptions, namely interactive work with Excel. This should be revisited as new possibilities working with

Excel have become known. These should be explored before contemplating other solutions.

(b) Depending on the outcome of the action in (a), development of an external database and coding of selected SIP functions should be considered.

REFERENCES

Austroads 2010a, *Interim network level functional road deterioration models*, AP-T158-10, Austroads, Sydney, NSW.

Austroads 2010b, *Predicting structural deterioration of pavements at a network level: interim models*, AP-T159-10, Austroads, Sydney, NSW.

Austroads 2013, *Probabilistic road deterioration model development*, AP-T257-13, Austroads, Sydney, NSW.

Austroads 2015, *Further development of probabilistic road deterioration modelling: pilot application*, AP-T294-15, Austroads, Sydney, NSW.

Austroads 2016, *Incorporating uncertainty in pavement management system (PMS) modelling: phase 1*, AP-T304-16, Austroads, Sydney, NSW.

Deighton Associates Limited 2014, *Deighton's Total Infrastructure Management System, (dTIMS),* software, version 9, Deighton Associates Limited.

Hubbard, D W 2010, How to measure anything: finding the value of 'intangibles' in business, Wiley, New York, USA.

Kadar, P, Martin, T, Baran, M & Sen, R 2015, 'Addressing uncertainties of performance modelling with stochastic information packages: incorporating a measure of uncertainty in performance and budget forecast', *International conference on managing pavement assets (ICMPA9), 9th, 2015, Washington, DC, USA*, Virginia Tech Transportation Institute, Blacksburg, Virginia, 11 pp.

Noya, L & Marzouk, S 2016, 'The impact of reduced surface maintenance on pavement condition: NACOE Project A21: annual summary report', contract report 010563-1, ARRB Group, Vermont South, Vic.

Paterson, W & Scullion, T 1990, *Information systems for road management: draft guidelines on system design and data issues: technical paper,* report no. INU77, World Bank, Washington, DC, USA.

Probability Management 2014, *The open SIPmath™ 2.0 standard: the unambiguous communication of uncertainty,* webpage, Probability Management, viewed 1 August 2016, <http://probabilitymanagement.org/standards.html>.

Savage, SL 2009, *The flaw of averages: why we underestimate risk in the face of uncertainty*, Wiley, New York, USA.

Thibault, JM 2011, *On time and under budget: probability management for projects: how to manage the odds of project success*, viewed 1 August 2016, <http://smpro.ca/ProbMan/pm4pm.pdf>.

Thibault, JM 2013, *SDXL reference manual: sample distributions for Excel simulation and modeling*, CreateSpace Independent Publishing Platform, viewed 1 August 2016, <https://www.amazon.com/SDXL-Reference-Manual-Distributions-Simulation/dp/1479199974>.

APPENDIX A PROTOTYPE MODEL

The prototype model is a combination of an Excel file and dTIMS. The Excel file contains the annual deterioration calculations that are automatically accessed by dTIMS. The annual checks for treatment applicability and the triggering of treatments are conducted in dTIMS. The dTIMS operation differs from the usual (deterministic) modelling because the deterioration calculations are conducted in Excel. As some of the model inputs are calculated in dTIMS, e.g. the works effects, these must be input manually into Excel.

The following files are necessary to execute the prototype model:

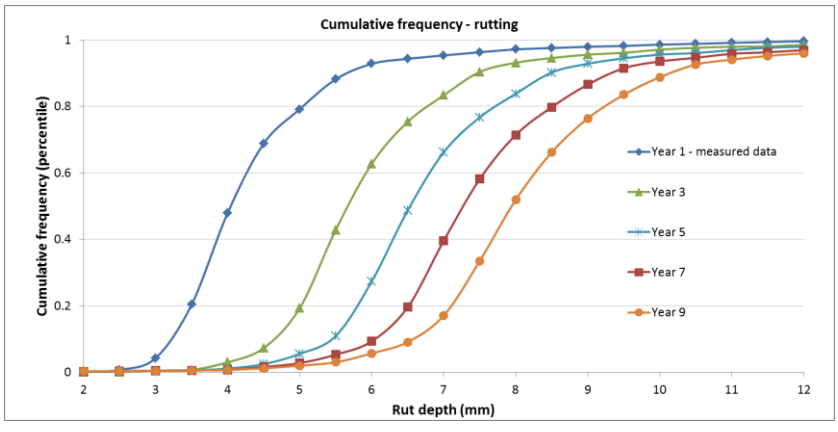
- Excel: **Probabilistic RD models RS**.xls Location: \\nsw-fps\Sydney\1 Projects\3 Completed\2015\007198 007141_A5_TMR_Incorporating Uncertainty in PMS modelling_QLD\4-Working\Probabilistic Deterioration modelling.
- dTIMS setup (SQL 2008) associated with this work can be found on the dTIMS drive: **D:\ARRB\NACOE_A5\NACOE_A5.mdf.**

To execute the model, the Excel spreadsheet must be open and its location must be identified in the relevant expressions in dTIMS. This latter condition is ensured as long as the files are at the above nominated location. If the files are moved, their location must be recorded in the relevant dTIMS expressions linking dTIMS to the spreadsheet model.

A PowerPoint file, titled *"**Probabilistic PMS modelling_setup description**"*, has also been prepared to illustrate the working of the model. The location of the file is \\nsw-fps\Sydney\1 Projects\3 Completed\2015\007198 007141_A5_TMR_Incorporating Uncertainty in PMS modelling_QLD\4-Working\Probabilistic Deterioration modelling\Setup description-presentation.

A prototype rutting model was built on an Excel platform. The probabilistic rutting model is based on the deterministic Austroads model (Austroads 2010a) with the selected input parameters replaced by SIPs. The expressions of the model were modified to accommodate the uncertain parameters. The prototype model can forecast the rutting distribution of a road section for a ten-year period. The forecast cumulative distributions are shown in Figure A 1.

Figure A 1: Resulting rut depth distributions



A.1 Implementation in dTIMS

The implementation of the model in dTIMS is illustrated in Figure A 2. For the sake of clarity and better visibility, parts of Figure A 2 are presented in Figure A 3, Figure A 4, and Figure A 5.

Implementation in dTIMS requires the following steps:

1. Set-up a test road database in Excel.
2. Store input conditions/section specific parameters distributions (SIPs) in the database.
3. Conduct deterioration modelling using condition distributions (SIPs) in Excel and other scalar variables in dTIMS.
4. Extract the stochastic values at the desired probability level, such as the 75th percentile of rutting, roughness etc.
5. Transfer the percentile values to dTIMS for triggering treatments.
6. If treatments are selected, the stochastic parameters are reset in Excel and new distributions are calculated with the reset values.
7. The network level condition and budget demand is calculated.
8. If a distribution of the results is required, the whole process is repeated for a range of probability levels, i.e. percentiles.
9. Check and compare dTIMS outputs (program costs/conditions) for different probability/risk levels.

Figure A 2: Flowchart

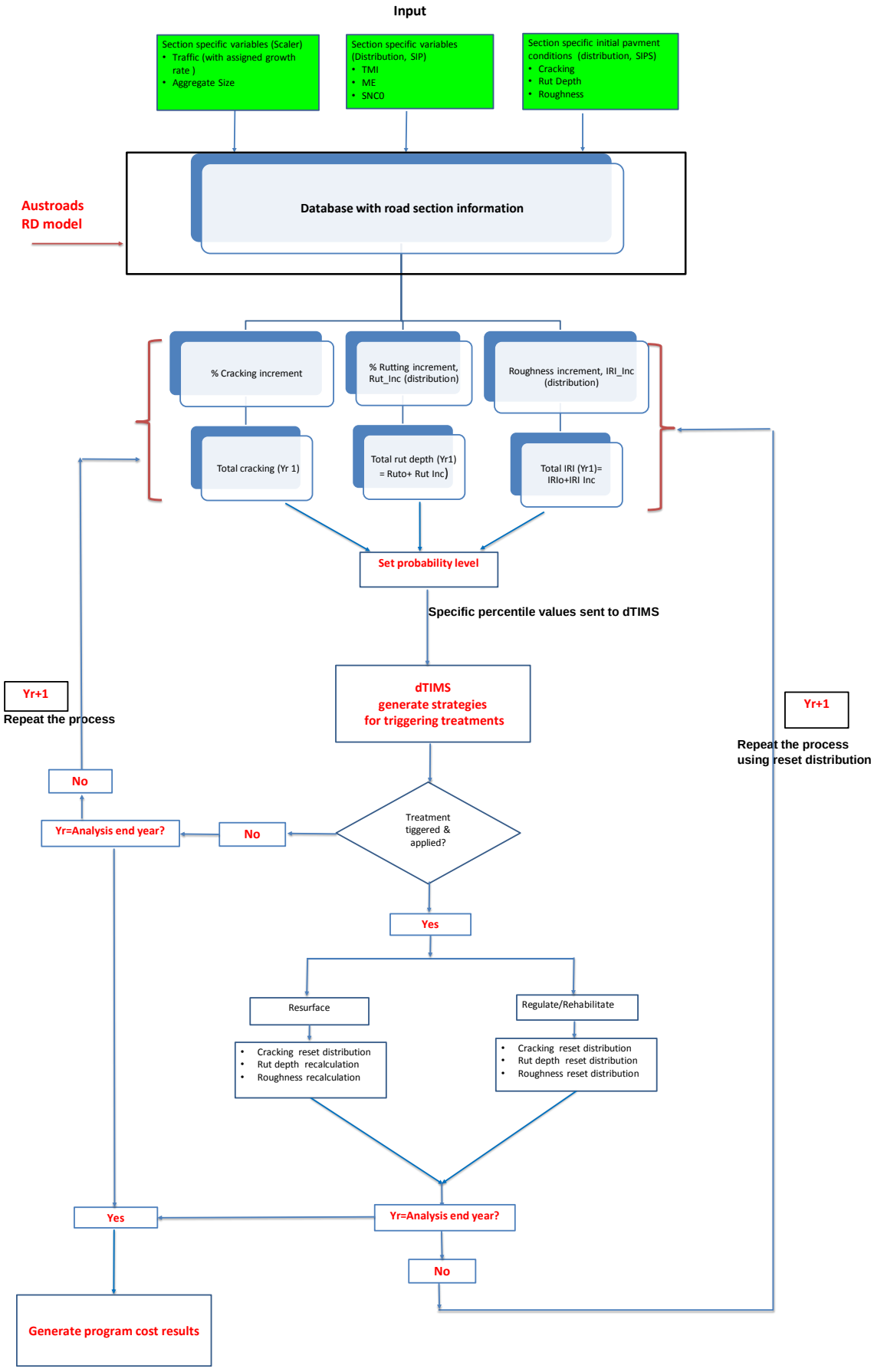


Figure A 3: Flowchart Part 1 – input and Austroads model

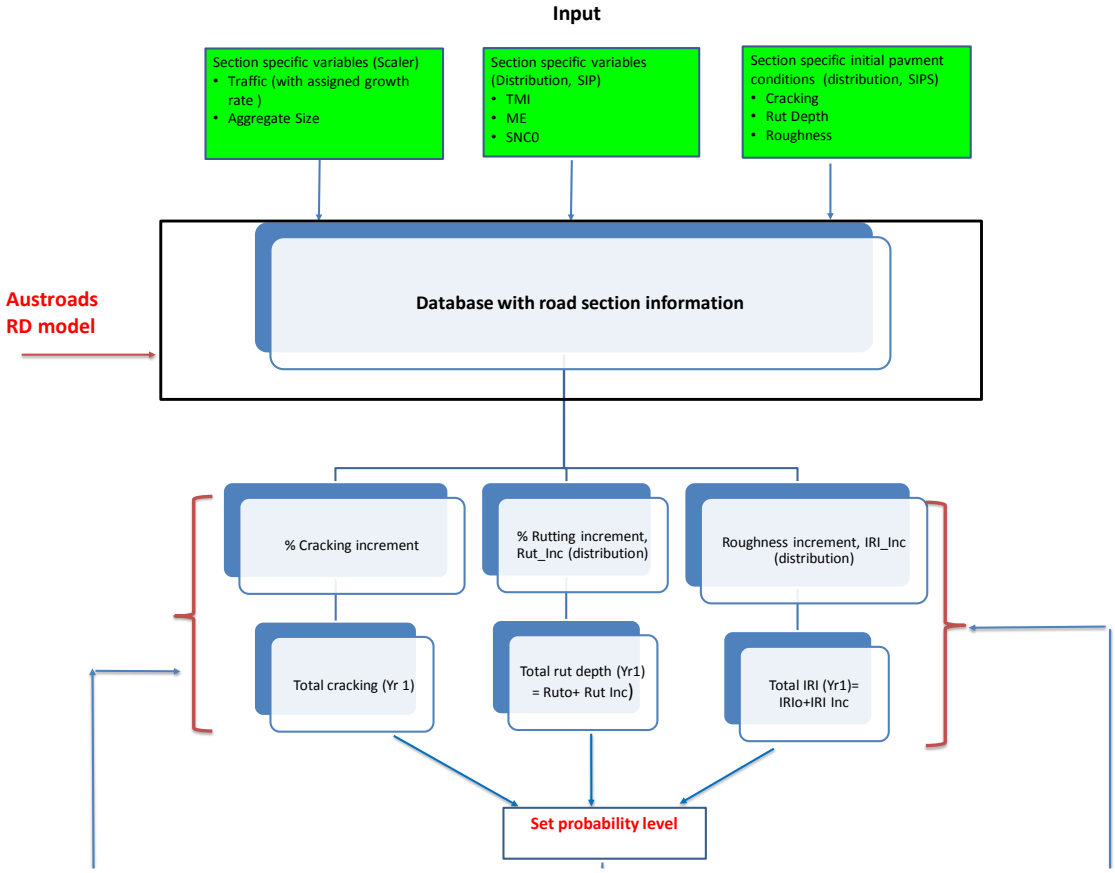


Figure A 4: Flowchart-Part 2 – Austroads model and input to dTIMS

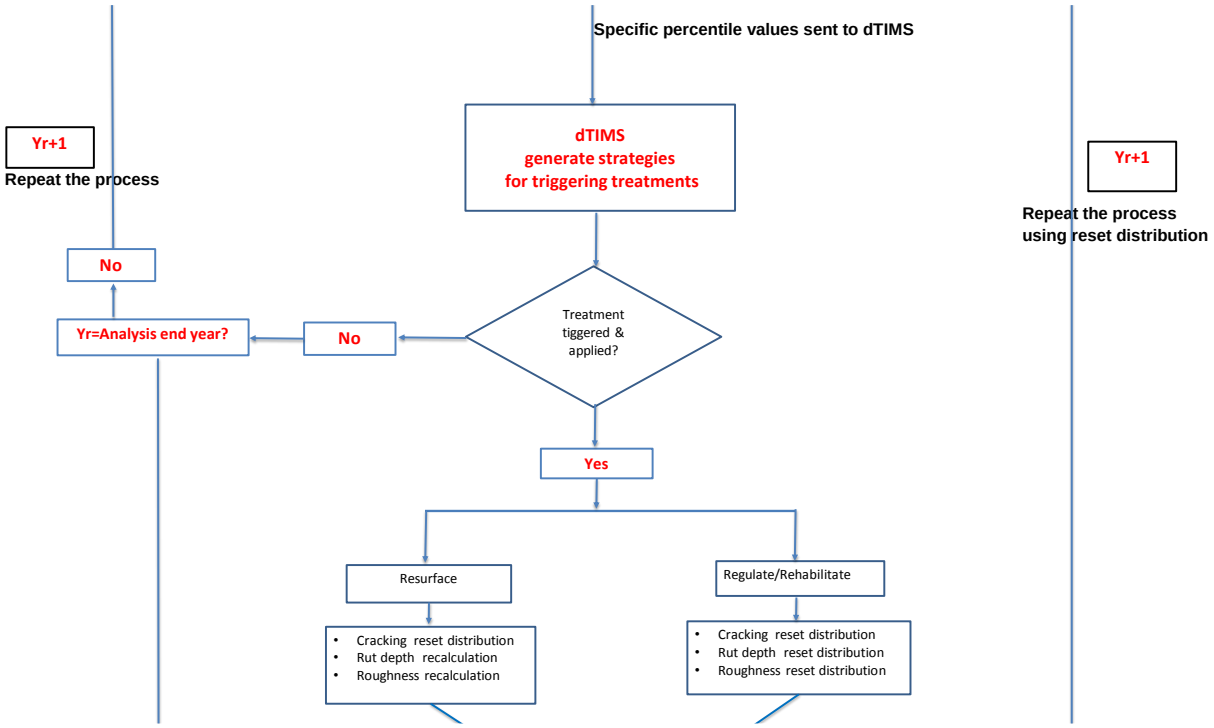
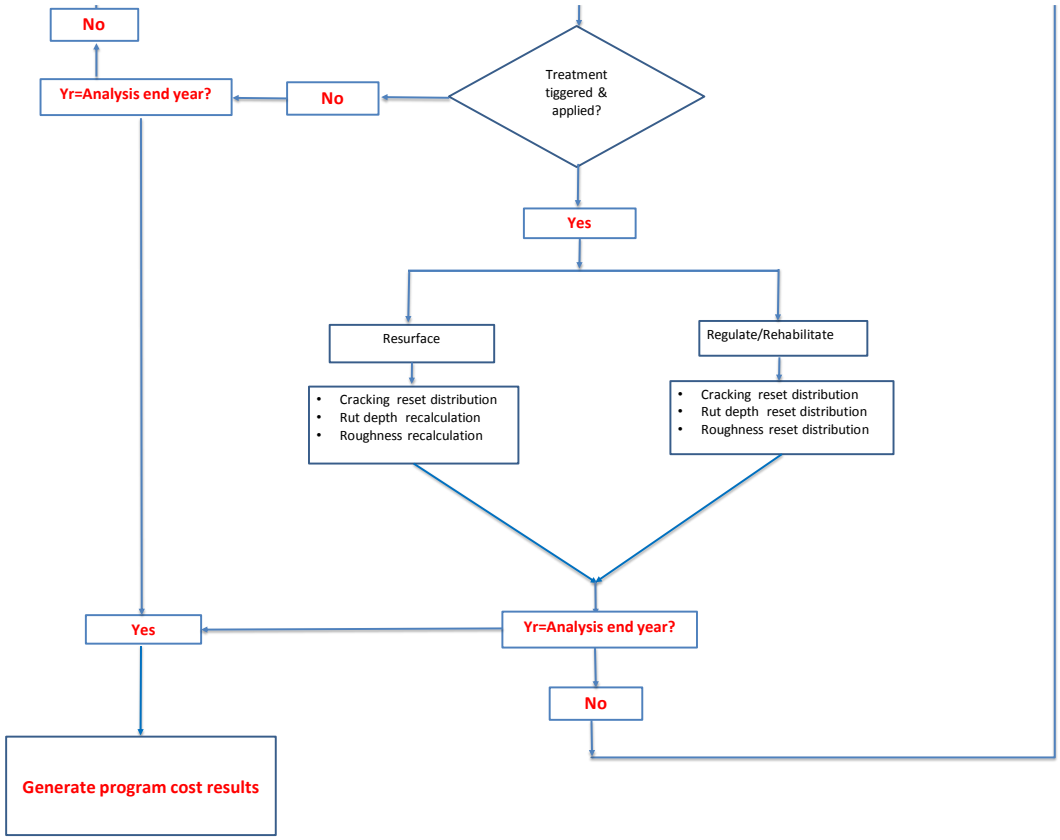


Figure A 5: Flowchart-Part 3 – treatment selection and model calculation in Excel



A.2 Structure of the Excel file

A.2.1 Road sections in the database

The test database in the Excel file currently contains 20 hypothetical road sections each 1 km in length (Table A 1).

Table A 1: Road section list

ID	Road	From	To	Length (m)	ElementID
1	Road01	0	1	1000	Road01_0
2	Road01	1	2	1000	Road01_1
3	Road01	2	3	1000	Road01_2
4	Road01	3	4	1000	Road01_3
5	Road01	4	5	1000	Road01_4
6	Road01	5	6	1000	Road01_5
7	Road01	6	7	1000	Road01_6
8	Road01	7	8	1000	Road01_7
9	Road01	8	9	1000	Road01_8
10	Road01	9	10	1000	Road01_9
11	Road01	10	11	1000	Road01_10
12	Road01	11	12	1000	Road01_11
13	Road01	12	13	1000	Road01_12

ID	Road	From	To	Length (m)	ElementID
14	Road01	13	14	1000	Road01_13
15	Road01	14	15	1000	Road01_14
16	Road01	15	16	1000	Road01_15
17	Road01	16	17	1000	Road01_16
18	Road01	17	18	1000	Road01_17
19	Road01	18	19	1000	Road01_18
20	Road01	19	20	1000	Road01_19

A.2.2 Input Sheets

The Excel file developed for this purpose has three sections separating the inputs (inventory, traffic and condition information), modelling and outputs. However, the sheets are linked internally using Excel formulae and macros so that user inputs are carried to the modelling part and outputs from modelling are linked to the dTIMS input table. Table A 2 summarises the contents of the different sheets within the Excel file. The colours correspond to the colour scheme used in the Excel model.

Table A 2: Excel sheets containing inputs, modelling and outputs

Sheet name	Content	Editable
Introduction	Introduction to the process and operational instructions	No
Operational Flowchart		
Sheet specific notes		
1_Cracking Input	Contains input data / reset data	Yes, specific data input parts only
2_Rut Depth Input		
3_Roughness Input		
4_TMI Input		
5_SNC_0 Input		
6_ME Input		
7_Traffic and Agg Input		
8_Reset variables		
Probabilistic RD model	Contains models, input to dTIMS/Outputs from dTIMS	No, only results can be refreshed/updated
Table		
Results		

A.3 Input variables

A.3.1 Variables used in modelling

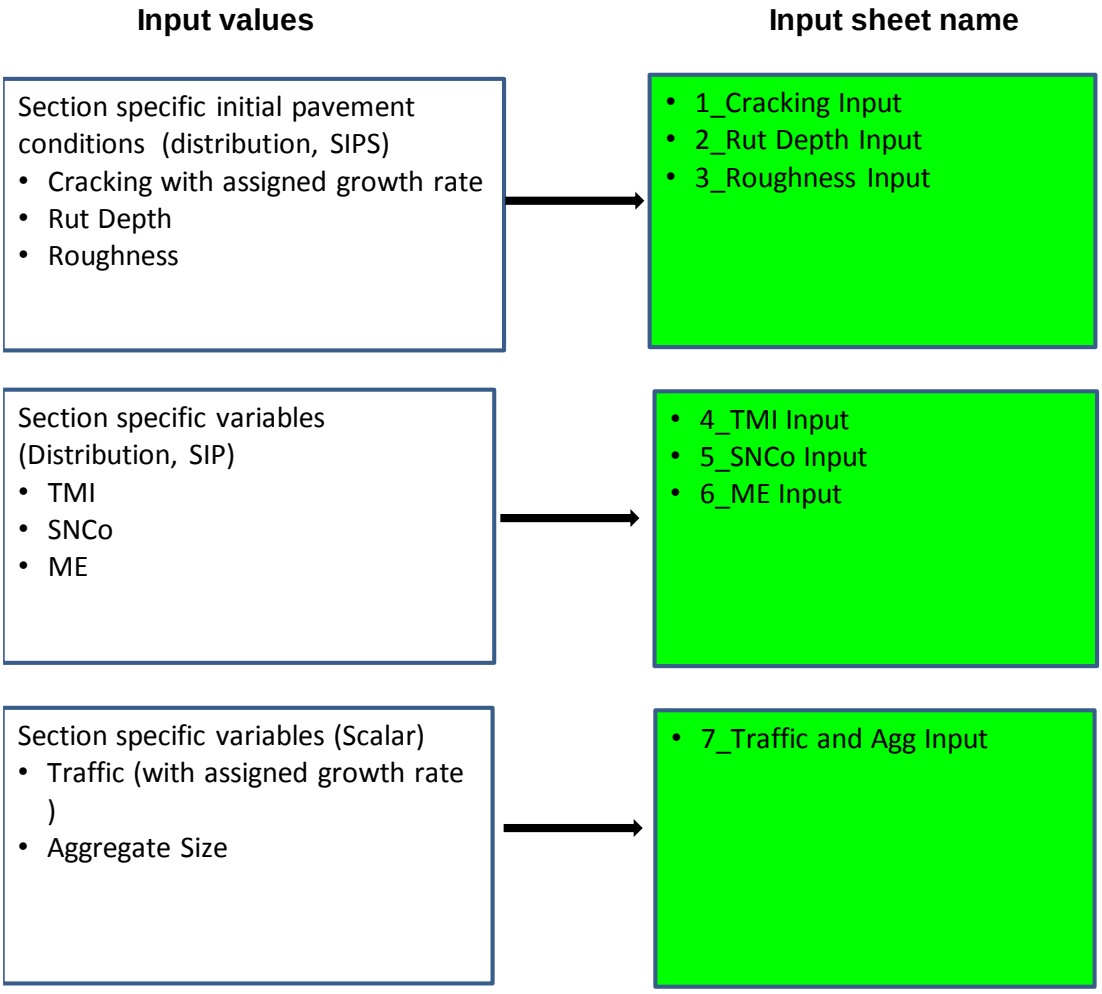
Input variables used in the modelling process include section specific condition variables, traffic and environmental information, strength and aggregate size data. These are as follows:

- Cracking (% area)
- Roughness (IRI)
- Rut depth (mm)

- Thornthwaite Moisture Index, TMI
- Initial structural Number, SNC_0
- Maintenance expenditure, ME
- Traffic (MESA and assigned growth rate)
- Aggregate size.

For all road sections, all input information except traffic and aggregate size are stored as SIPs (currently distributions containing 1000 data items for each section) instead of single numbers in the respective input sheets as summarised in Table A 2 and Figure A 6. Traffic and aggregate size are kept as scalar (single value for each section); however, the traffic growth rate has been taken into account over the length of the analysis period.

Figure A 6: Input variables and respective input sheets



A.3.2 Sample input sheet with distribution

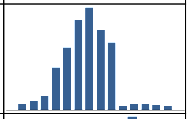
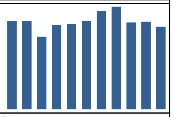
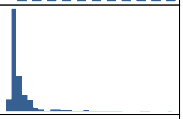
A sample input sheet with initial rut depth (Year 0 input distribution) for 5 road sections is shown in Figure A 7.

Figure A 7: Sample input sheet with distribution

ID	1	2	3	4	5
Road	Road01	Road01	Road01	Road01	Road01
From	0	1	2	3	4
To	1	2	3	4	5
Length	1	1	1	1	1
ElementID	Road01_0	Road01_1	Road01_2	Road01_3	Road01_4
	Rut depth_yr 0	Rut depth_yr 0	Rut depth_yr 0	Rut depth_yr 0	Rut depth_yr 0
SIP	<SIP name="Rut Yr0" cou	<SIP name="Rut Yr0"	<SIP name="Rut Yr0	<SIP name="Rut Yr	<SIP name="Rut Yr0" count=
Starting point of data entry	2.00	2.20	2.00	1.90	2.00
	2.02	2.21	2.00	1.93	2.02
	2.04	2.22	2.00	1.96	2.05
	2.06	2.22	2.00	1.99	2.07
	2.08	2.23	2.00	2.03	2.10
	2.10	2.24	2.00	2.06	2.12
	2.12	2.25	2.00	2.09	2.14
	2.14	2.26	2.00	2.12	2.17
	2.16	2.26	2.00	2.15	2.19
	2.18	2.27	2.00	2.19	2.21
	2.20	2.28	2.00	2.22	2.24
	2.23	2.29	2.00	2.25	2.26

Input data should be entered from the start of the 'green rows', referred to as 'Starting point of data entry'. Currently each SIP string includes 1000 condition data points for each section. The number of elements within a SIP can be changed but within the same model, all SIPs must have the same number of elements. Once the input distribution is entered, it is automatically converted into a SIP. The row named 'SIP' contains input SIPs in XML format.

Figure A 8: Sample input distribution

	Min	Max	Average	Distribution
SNCo	1.77	5.32	3.3	
Thorthwaite index	20.0	30	25.1	
Maintenance expenditure pa. ($/km)	0	69594	1139.5	
MESA	0.5	0.5	0.5	n/a

A similar process is followed for entering data input distributions. Detailed sheet specific instructions for data entry can be found under the sheet named 'Sheet specific notes'.

A.4 Deterioration models

A.4.1 Austroads Models

Selected road deterioration models were adopted from the models developed under Austroads Project AT1064 (Austroads 2010a and 2010b). The database has been set up to cater for all models, but those for which sufficient data was available were implemented. The key models implemented include:

- rut depth progression model
- roughness progression model.

Cracking also uses an input distribution; however, a fixed rate of progression has been assumed for cracking instead of using the Austroads models. A new distribution is calculated each year by adding to the progression of the previous year's cracking.

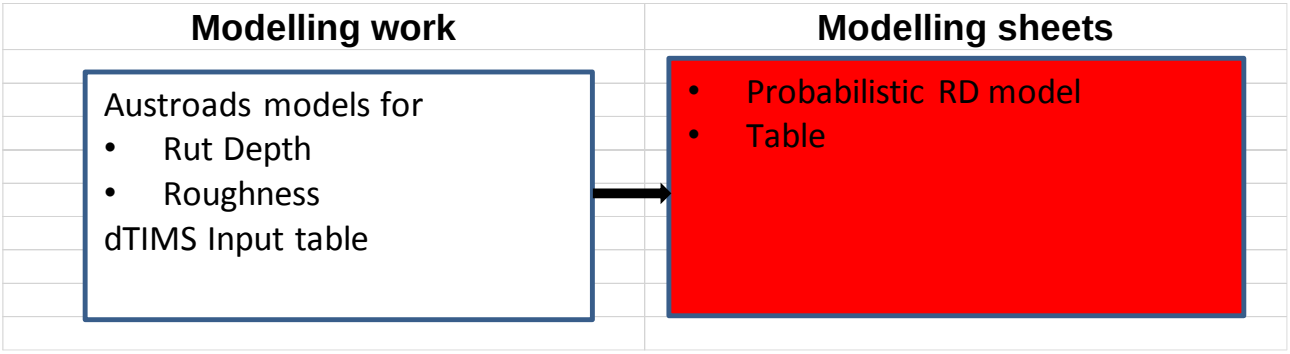
As most of the inputs for the current modelling work are distributions (SIP), the resulting progressions are also distributions (SIPs). Unlike conventional addition, subtraction, multiplication etc., there are specific techniques to calculate results involving two or more SIPs. SDXL add-ins developed by Thebault (2011) have been used for this purpose. Table A 3 shows the difference between the deterministic rut depth deterioration model and the same model when calculated using SIP formulas.

Table A 3: Deterministic rut depth formula and respective SDXL format

Parameter name	Deterministic formula	SDXL format
Annual rut depth increment	(AgeCurrent-1)^0.617*(0.022*(100+TMI)/SNC_0+0.594*MESAcurrent-0.000102*Maintenance Expenditure)-(AgePrevious-1)^0.617*(0.022*(100+TMI)/SNC_0+0.594*MESAprevious-0.000102*Maintenance Expenditure)	toStdSIP(sdsub(sdMul(sdAdd(sdAdd(sdMul(Maintenance Expenditure,-0.000102),(MESACurrent*0.594)),sdMul(sddiv(sdAdd(TMI,100),SNC_0),0.022)),((AgeCurrent-1)^0.617)),sdMul(sdAdd(sdAdd(sdMul(Maintenance Expenditure,-0.000102),(MESAPrevious*0.594)),sdMul(sddiv(sdAdd(TMI,100),SNC_0),0.022)),((AgePrevious-1)^0.617))), "Rut Inc", 6)
Total end-of-year rut depth	Annual rut depth increment+ Rut previous	toStdSIP(sdAdd(Rut Inc SIP,Rut Previous SIP), "Rut=Inc+Rprev", 4)

Currently a 10-year performance modelling period is done automatically within the modelling sheet (Figure A 9) for all 20 road sections. The resulting outputs for all years are also distributions. The user then has the option of setting a desired probability level to be used as inputs into dTIMS.

Figure A 9: Modelling work and respective modelling sheet



A.4.2 Work Effects

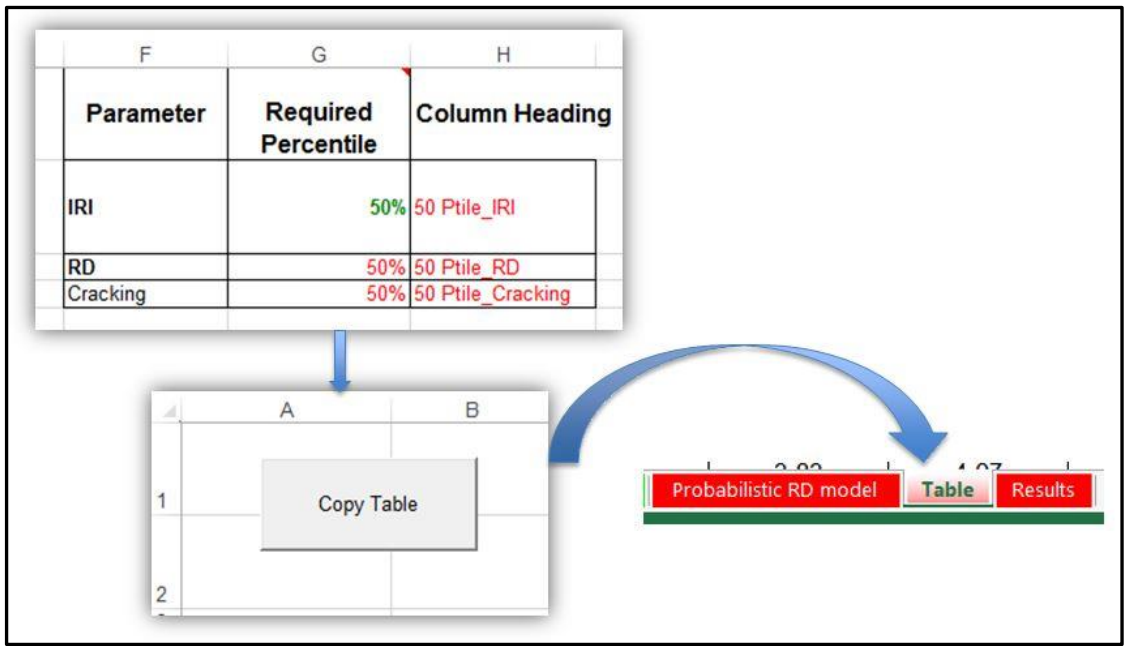
After a treatment is applied, the condition parameters are reset to a certain level. The amount of the reset (and the condition parameters that are reset) depends on the type of the treatment applied. For the current work, three reset distributions have been used for cracking, rutting and roughness and are included in the '**8_Reset variable**' sheet. Resurfacing only resets cracking while regulate/rehabilitation resets all three parameters. For any section, these reset distributions are used after any treatment application and subsequent years are modelled using the same techniques as before using the reset distribution as the initial distribution

A.4.3 dTIMS input

Although most of the cells in the 'Probabilistic RD model' sheet cannot be edited by users, a user has the option of setting a desired probability level for modelled distributions to be used as inputs into dTIMS. Treatments can only be triggered by a single value and not a distribution; hence, the need to select a probability level, which in turn defines the current value for triggering.

Once the probability level is defined in the appropriate cell (see Figure A 10), the 'Copy Table' button should be clicked to refresh the 'Table' sheet with inputs (for dTIMS) for the selected probability. The copy table command copies the relevant values into an area accessed by dTIMS. These values are used by dTIMS for triggering treatments and for resets after treatment application.

Figure A 10: Generation of dTIMS input for a given probability percentage



A sample rut depth input table for triggering treatments and resets in dTIMS is shown in Table A 4 and Table A 5 respectively.

Table A 4: dTIMS input: calculation before treatment with 50Ptile rut depth

ElementID	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_10
Road01_0	4.80	4.80	5.52	5.91	6.23	6.55	6.84	7.10	7.35	7.57	7.80
Road01_1	5.80	5.80	6.52	6.93	7.25	7.54	7.80	8.05	8.28	8.51	8.74
Road01_2	5.10	5.10	5.90	6.32	6.66	6.97	7.25	7.51	7.76	7.99	8.22
Road01_3	5.00	5.00	5.75	6.16	6.50	6.80	7.07	7.33	7.57	7.80	8.00
Road01_4	5.00	5.00	5.67	6.08	6.39	6.69	6.97	7.19	7.42	7.67	7.89
Road01_5	5.40	5.40	6.08	6.47	6.80	7.11	7.38	7.62	7.86	8.09	8.31
Road01_6	5.10	5.10	5.92	6.35	6.69	6.99	7.29	7.56	7.80	8.01	8.23
Road01_7	5.40	5.40	6.17	6.57	6.92	7.22	7.49	7.74	8.00	8.22	8.44
Road01_8	5.20	5.20	5.97	6.37	6.71	7.03	7.31	7.55	7.79	8.03	8.24
Road01_9	5.10	5.10	5.85	6.27	6.61	6.90	7.19	7.44	7.68	7.92	8.13
Road01_10	5.00	5.00	5.75	6.16	6.50	6.79	7.06	7.30	7.57	7.80	8.02

Table A 5: dTIMS input: calculation after treatment with 50Ptile rut depth

ElementID	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_10
Resurf_All	4.80	4.80	5.52	5.91	6.23	6.55	6.84	7.10	7.35	7.57	7.80
Rehab_All	1.50	1.50	2.29	2.71	3.05	3.36	3.63	3.89	4.13	4.35	4.57

A.5 DTIMS output

A.5.1 Treatment triggers using dTIMS

The dTIMS setup used for this work contains the same road database (20 road sections with the same From, To and ElementID) as in Excel. The setup is designed in a way that it reads the values from the 'Table' sheet directly. To ensure the linkage between Excel and dTIMS, the following dTIMS expressions (Table A 6) must be checked before running the setup.

Table A 6: Pre-run checks in dTIMS

dTIMS Expression	Description	Comment
A_Data_Selector	0= reads dTIMS data, 1=reads frm Excel	Should be "1"
Exp_Excel_FileName	Excel file name	Must refer to exact name
Exp_Excel_Location	Excel file location	Must refer to appropriate location

Three treatments are currently used in the dTIMS setup each with its own trigger values for cracking, roughness and rutting. The treatments are:

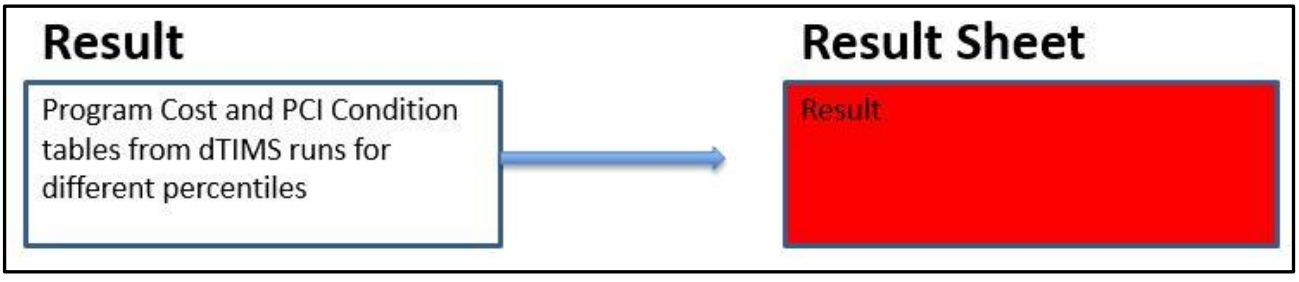
- resurface
- regulate
- rehabilitate.

Each year, dTIMS checks the trigger values with the corresponding parameter values in the 'Table' sheet (e.g. Table A 4) and once the values exceed the trigger values, a treatment is applied (provided there is enough funding). Once a treatment is applied, dTIMS resets the value from the reset table (e.g. Column C_0 of Table A 5) and reads the values for the subsequent year from the same table. If no treatment is applied, dTIMS continues to read the values from the before treatment tables of the 'Table' sheet (e.g. Table A 4).

A.5.2 Results

The dTIMS outputs using different probability levels for the current database are stored in the 'Results' sheet.

Figure A 11: dTIMS output and respective result sheet



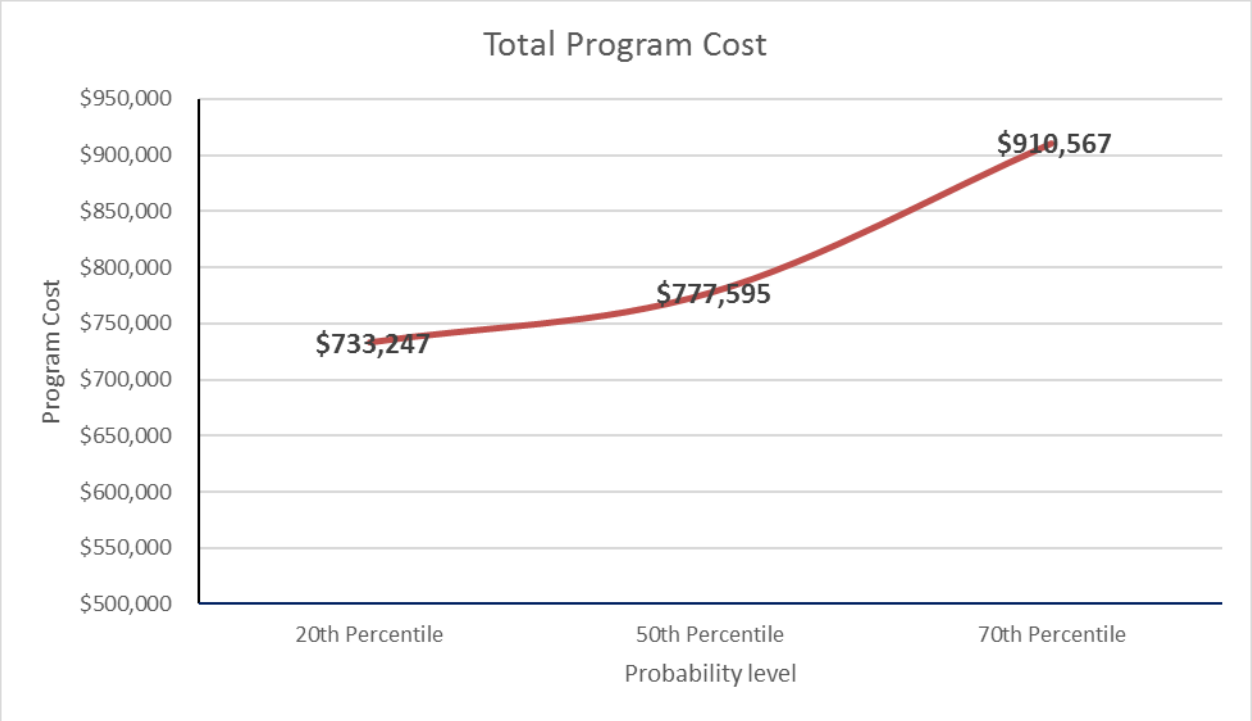
Currently the set probability levels are 20th Percentile, 50th Percentile, 70th Percentile and 95th Percentile. It should be noted that each time while running for a specific probability level, only the table corresponding to that probability level in the 'Result' sheet should be refreshed.

Each probability level requires a separate dTIMS run, i.e. the calculation has to be repeated at each probability level if a full probability curve is required. Otherwise, the calculation is conducted only at the desired probability level.

A.5.3 Sample program cost

A sample program cost has been generated using an unlimited budget and three probability levels for the Excel database described in this report. The results are graphically presented in Figure A 12. It shows an increase in the program cost as the probability/ level of certainty increases.

Figure A 12: Program cost for different probability levels



APPENDIX B DATA CONDENSATION

B.1 Example SIP

An example SIP is shown below. It was formatted with the command '**toStdSIP (TMI,"TMI SIP", 2)**' where the name of the SIP and the number of decimals are defined. The SIP occupies a single cell in Excel:

<SIP name="TMI SIP" count="1270" type="CSV" ver="1.0.0" >-11.00,-42.00,-38.00,-28.00,-32.00,-22.00,-11.00,-17.00,-
40.00,-38.00,-9.00,-26.00,-26.00,-30.00,-16.00,-26.00,-32.00,-38.00,-32.00,-30.00,-22.00,-1.00,-27.00,-6.00,-41.00,-30.00,-26.00,-
26.00,-38.00,-33.00,-33.00,-1.00,-27.00,-15.00,-24.00,-27.00,-32.00,-26.00,-32.00,-32.00,-34.00,-31.00,-31.00,-22.00,-31.00,-21.00,-
31.00,-32.00,-31.00,-30.00,-42.00,-33.00,-22.00,-14.00,-30.00,-28.00,-21.00,-29.00,-6.00,-27.00,-32.00,-17.00,-31.00,-34.00,-14.00,-
38.00,-29.00,-31.00,-28.00,-17.00,-22.00,-38.00,-38.00,-32.00,-27.00,-6.00,-1.00,-15.00,-22.00,-31.00,-27.00,-16.00,-16.00,-22.00,-
29.00,-38.00,-6.00,-27.00,-31.00,-34.00,-30.00,-27.00,-42.00,-33.00,-27.00,-41.00,-14.00,-22.00,-6.00,-23.00,-21.00,-26.00,-42.00,-
22.00,-21.00,-33.00,-22.00,-28.00,-14.00,-23.00,-20.00,-32.00,-17.00,-29.00,-21.00,-27.00,-11.00,-29.00,-27.00,-21.00,-15.00,-31.00,-
30.00,-31.00,-28.00,-21.00,-30.00,-38.00,-32.00,-30.00,-32.00,-33.00,-38.00,-28.00,-21.00,-17.00,-32.00,-6.00,-26.00,-21.00,-38.00,-
38.00,-28.00,-22.00,-32.00,-32.00,-14.00,-22.00,-6.00,-22.00,-6.00,-32.00,-30.00,-21.00,-15.00,-21.00,-24.00,-17.00,-28.00,-33.00,-
11.00,-21.00,-31.00,-27.00,-32.00,-30.00,-34.00,-6.00,-23.00,-17.00,-40.00,-33.00,-32.00,-38.00,-9.00,-41.00,-38.00,-20.00,-28.00,-
27.00,-11.00,-32.00,-30.00,-16.00,-20.00,-11.00,-30.00,-27.00,-30.00,-9.00,-31.00,-29.00,-14.00,-38.00,-23.00,-15.00,-20.00,-30.00,-
17.00,-9.00,-23.00,-26.00,-30.00,-38.00,-42.00,-24.00,-31.00,-32.00,-9.00,-32.00,-31.00,-15.00,-17.00,-40.00,-17.00,-32.00,-23.00,-
29.00,-22.00,-31.00,-28.00,-31.00,-6.00,-22.00,-27.00,-22.00,-15.00,-21.00,-27.00,-30.00,-16.00,-29.00,-16.00,-22.00,-14.00,-14.00,-
24.00,-30.00,-32.00,-14.00,-11.00,-22.00,-32.00,-33.00,-28.00,-31.00,-30.00,-26.00,-21.00,-32.00,-26.00,-31.00,-31.00,-26.00,-28.00,-
6.00,-26.00,-21.00,-31.00,-15.00,-23.00,-31.00,-38.00,-32.00,-31.00,-29.00,-27.00,-16.00,-32.00,-33.00,-30.00,-32.00,-22.00,-16.00,-
22.00,-17.00,-28.00,-42.00,-31.00,-27.00,-30.00,-30.00,-32.00,-28.00,-1.00,-21.00,-26.00,-29.00,-40.00,-14.00,-16.00,-33.00,-26.00,-
41.00,-33.00,-15.00,-32.00,-6.00,-32.00,-30.00,-31.00,-42.00,-6.00,-28.00,-22.00,-27.00,-22.00,-31.00,-1.00,-17.00,-27.00,-29.00,-
33.00,-22.00,-21.00,-27.00,-31.00,-14.00,-16.00,-32.00,-27.00,-27.00,-22.00,-38.00,-16.00,-14.00,-28.00,-22.00,-33.00,-32.00,-15.00,-
21.00,-31.00,-41.00,-33.00,-20.00,-29.00,-22.00,-17.00,-20.00,-22.00,-21.00,-30.00,-21.00,-38.00,-22.00,-27.00,-11.00,-17.00,-38.00,-
21.00,-21.00,-22.00,-6.00,-23.00,-30.00,-11.00,-30.00,-40.00,-16.00,-42.00,-22.00,-38.00,-30.00,-33.00,-26.00,-23.00,-20.00,-26.00,-
28.00,-32.00,-32.00,-30.00,-14.00,-6.00,-41.00,-24.00,-29.00,-41.00,-31.00,-30.00,-38.00,-33.00,-29.00,-38.00,-32.00,-32.00,-22.00,-
30.00,-26.00,-20.00,-31.00,-32.00,-21.00,-22.00,-38.00,-29.00,-20.00,-17.00,-20.00,-27.00,-17.00,-32.00,-38.00,-21.00,-34.00,-26.00,-
41.00,-32.00,-32.00,-17.00,-27.00,-32.00,-21.00,-30.00,-27.00,-1.00,-20.00,-27.00,-11.00,-23.00,-30.00,-6.00,-32.00,-11.00,-21.00,-
30.00,-26.00,-33.00,-38.00,-30.00,-22.00,-14.00,-38.00,-14.00,-6.00,-9.00,-30.00,-21.00,-30.00,-32.00,-21.00,-29.00,-22.00,-26.00,-
14.00,-31.00,-30.00,-32.00,-17.00,-28.00,-33.00,-29.00,-1.00,-31.00,-9.00,-30.00,-38.00,-38.00,-14.00,-31.00,-20.00,-14.00,-38.00,-
28.00,-26.00,-21.00,-22.00,-14.00,-31.00,-33.00,-26.00,-21.00,-17.00,-41.00,-26.00,-9.00,-32.00,-17.00,-9.00,-28.00,-27.00,-17.00,-
22.00,-29.00,-15.00,-38.00,-32.00,-33.00,-26.00,-22.00,-27.00,-26.00,-32.00,-22.00,-6.00,-32.00,-27.00,-38.00,-21.00,-28.00,-14.00,-
40.00,-22.00,-38.00,-40.00,-27.00,-28.00,-16.00,-17.00,-14.00,-20.00,-32.00,-33.00,-23.00,-32.00,-21.00,-31.00,-42.00,-22.00,-29.00,-
38.00,-9.00,-30.00,-30.00,-32.00,-27.00,-42.00,-26.00,-22.00,-26.00,-31.00,-42.00,-27.00,-32.00,-32.00,-15.00,-29.00,-26.00,-24.00,-
40.00,-11.00,-32.00,-33.00,-29.00,-21.00,-30.00,-32.00,-32.00,-29.00,-6.00,-41.00,-15.00,-29.00,-30.00,-30.00,-22.00,-29.00,-32.00,-
15.00,-29.00,-21.00,-22.00,-26.00,-29.00,-16.00,-1.00,-31.00,-22.00,-22.00,-22.00,-20.00,-29.00,-1.00,-22.00,-9.00,-31.00,-28.00,-
29.00,-15.00,-1.00,-31.00,-9.00,-33.00,-1.00,-42.00,-16.00,-22.00,-6.00,-32.00,-22.00,-20.00,-22.00,-38.00,-6.00,-28.00,-17.00,-9.00,-
22.00,-38.00,-26.00,-23.00,-24.00,-15.00,-29.00,-38.00,-9.00,-32.00,-20.00,-22.00,-6.00,-24.00,-34.00,-6.00,-32.00,-14.00,-32.00,-
22.00,-27.00,-32.00,-31.00,-16.00,-31.00,-30.00,-27.00,-21.00,-24.00,-32.00,-28.00,-22.00,-20.00,-6.00,-31.00,-32.00,-31.00,-32.00,-
23.00,-22.00,-32.00,-32.00,-29.00,-32.00,-23.00,-9.00,-30.00,-6.00,-6.00,-22.00,-15.00,-6.00,-27.00,-14.00,-38.00,-6.00,-26.00,-33.00,-
21.00,-22.00,-1.00,-21.00,-40.00,-20.00,-30.00,-28.00,-40.00,-24.00,-28.00,-6.00,-22.00,-22.00,-31.00,-33.00,-24.00,-21.00,-14.00,-
41.00,-30.00,-29.00,-27.00,-32.00,-29.00,-27.00,-38.00,-32.00,-20.00,-24.00,-38.00,-22.00,-30.00,-1.00,-32.00,-22.00,-32.00,-22.00,-
17.00,-16.00,-32.00,-22.00,-38.00,-15.00,-22.00,-27.00,-22.00,-1.00,-16.00,-23.00,-26.00,-21.00,-40.00,-30.00,-33.00,-32.00,-32.00,-
9.00,-30.00,-17.00,-11.00,-28.00,-28.00,-32.00,-26.00,-32.00,-27.00,-38.00,-11.00,-22.00,-32.00,-22.00,-27.00,-28.00,-22.00,-21.00,-
21.00,-27.00,-30.00,-32.00,-29.00,-31.00,-26.00,-6.00,-31.00,-15.00,-22.00,-41.00,-17.00,-24.00,-40.00,-22.00,-40.00,-33.00,-29.00,-
33.00,-16.00,-6.00,-28.00,-22.00,-28.00,-33.00,-33.00,-21.00,-30.00,-30.00,-22.00,-33.00,-32.00,-21.00,-9.00,-11.00,-40.00,-41.00,-
29.00,-29.00,-21.00,-17.00,-21.00,-26.00,-28.00,-22.00,-26.00,-27.00,-22.00,-22.00,-38.00,-17.00,-16.00,-38.00,-22.00,-17.00,-32.00,-
29.00,-24.00,-29.00,-33.00,-27.00,-31.00,-32.00,-34.00,-6.00,-22.00,-6.00,-32.00,-31.00,-14.00,-42.00,-32.00,-9.00,-27.00,-27.00,-
26.00,-14.00,-28.00,-34.00,-9.00,-27.00,-38.00,-21.00,-32.00,-21.00,-28.00,-33.00,-26.00,-31.00,-11.00,-34.00,-30.00,-31.00,-34.00,-
20.00,-42.00,-6.00,-38.00,-26.00,-16.00,-38.00,-21.00,-31.00,-34.00,-29.00,-6.00,-31.00,-32.00,-32.00,-14.00,-30.00,-22.00,-22.00,-
11.00,-31.00,-22.00,-42.00,-31.00,-32.00,-15.00,-38.00,-14.00,-27.00,-17.00,-32.00,-38.00,-11.00,-27.00,-32.00,-27.00,-15.00,-22.00,-
29.00,-30.00,-38.00,-21.00,-38.00,-24.00,-17.00,-21.00,-27.00,-40.00,-24.00,-1.00,-38.00,-33.00,-33.00,-21.00,-29.00,-38.00,-38.00,-
28.00,-32.00,-40.00,-23.00,-33.00,-28.00,-22.00,-33.00,-30.00,-29.00,-22.00,-33.00,-28.00,-33.00,-28.00,-29.00,-30.00,-33.00,-38.00,-
40.00,-42.00,-17.00,-32.00,-42.00,-20.00,-33.00,-31.00,-21.00,-22.00,-17.00,-31.00,-29.00,-17.00,-32.00,-30.00,-28.00,-27.00,-15.00,-
17.00,-27.00,-30.00,-32.00,-29.00,-17.00,-33.00,-11.00,-31.00,-30.00,-22.00,-34.00,-22.00,-38.00,-41.00,-20.00,-24.00,-32.00,-1.00,-
38.00,-1.00,-21.00,-21.00,-29.00,-41.00,-33.00,-27.00,-22.00,-32.00,-32.00,-23.00,-22.00,-33.00,-22.00,-34.00,-6.00,-30.00,-31.00,-
30.00,-27.00,-27.00,-34.00,-32.00,-20.00,-32.00,-38.00,-24.00,-40.00,-27.00,-6.00,-15.00,-38.00,-16.00,-14.00,-33.00,-27.00,-26.00,-
27.00,-29.00,-40.00,-31.00,-21.00,-22.00,-22.00,-15.00,-29.00,-28.00,-34.00,-21.00,-32.00,-31.00,-33.00,-30.00,-32.00,-33.00,-29.00,-
27.00,-1.00,-32.00,-38.00,-30.00,-32.00,-38.00,-32.00,-27.00,-26.00,-31.00,-32.00,-22.00,-21.00,-30.00,-27.00,-32.00,-41.00,-33.00,-
11.00,-20.00,-29.00,-32.00,-31.00,-31.00,-6.00,-33.00,-24.00,-28.00,-20.00,-22.00,-21.00,-27.00,-27.00,-32.00,-21.00,-32.00,-27.00,-
29.00,-14.00,-40.00,-21.00,-33.00,-33.00,-31.00,-31.00,-21.00,-17.00,-33.00,-17.00,-31.00,-14.00,-22.00,-27.00,-15.00,-31.00,-27.00,-
14.00,-29.00,-23.00,-38.00,-26.00,-27.00,-28.00,-17.00,-15.00,-32.00,-29.00,-41.00,-14.00,-20.00,-32.00,-21.00,-38.00,-24.00,-38.00,-
33.00,-21.00,-26.00,-17.00,-21.00,-32.00,-32.00,-38.00,-9.00,-1.00,-30.00,-17.00,-33.00,-14.00,-27.00,-17.00,-31.00,-23.00,-6.00,-
30.00,-22.00,-33.00,-24.00,-38.00,-20.00,-15.00,-22.00,-32.00,-27.00,-33.00,-31.00,-24.00,-26.00,-22.00,-27.00,-27.00,-24.00,-32.00,-
22.00,-29.00,-31.00,-22.00,-32.00,-30.00,-14.00,-14.00,-29.00,-17.00,-14.00,-11.00,-28.00,-32.00,-41.00,-17.00,-27.00,-30.00,-6.00,-
27.00,-17.00,-38.00,-30.00,-15.00,-30.00,-40.00,-23.00,-42.00,-33.00,-30.00,-30.00,-6.00,-1.00,-29.00,-26.00,-1.00,-32.00,-27.00,-
31.00,-15.00,-22.00,-38.00,-17.00,-33.00,-1.00,-38.00,-32.00,-38.00,-30.00,-22.00,-29.00,-26.00,-29.00,-29.00,-21.00,-14.00,-14.00,-
32.00,-6.00,-22.00,-38.00,-40.00,-24.00,-31.00,-30.00,-26.00,-22.00,-31.00,-9.00,-34.00,-22.00,-32.00,-32.00,-22.00,-22.00,-14.00,-
21.00,-32.00,-22.00,-29.00,-26.00,-23.00,-38.00,-31.00,-24.00,-22.00,-1.00,-32.00,-26.00,-29.00,-34.00,-22.00,-38.00,-17.00,-32.00,-

29.00,-33.00,-41.00,-14.00,-22.00,-23.00,-14.00,-28.00,-9.00,-21.00,-29.00,-30.00,-33.00,-33.00,-38.00,-38.00,-16.00,-20.00,-21.00,-27.00,-26.00,-1.00,-11.00,-38.00,-38.00,-29.00,-27.00,-28.00,-22.00,-9.00,-21.00,-14.00,-11.00,-1.00,-17.00,-6.00,-22.00,-27.00,-32.00,-41.00,-34.00,-23.00,-27.00,-26.00,-29.00,-38.00,-21.00,-31.00,-24.00,-16.00,-22.00,-27.00,-29.00,-32.00,-14.00,-38.00,-23.00,-15.00,-41.00,-22.00,-1.00,-32.00,-30.00</SIP>

B.2 Thornthwaite Index

The Thornthwaite Index (TMI) was calculated with the following formula:

$$TMI = \frac{\left(P_{eff} - 48\right)}{0.8} \quad \text{A1}$$

where

P_{eff} = sum of monthly P_{effM}

$$P_{effM} = 1.65 * P_{rM}/((M_M + 12.2)^{1.1111})$$

P_{rM} = monthly precipitation (mm)

M_M = mean monthly temperature (°C)