

Optimising Road Maintenance

12

Discussion Paper 2012 • 12

Mark O. Harvey

Bureau of Infrastructure, Transport
and Regional Economics (BITRE)
Canberra, Australia

Optimising Road Maintenance

Discussion Paper No. 2012-12

Prepared for the Roundtable on
Sustainable Road Funding
25-26 October 2012

Mark O. HARVEY

Bureau of Infrastructure, Transport and Regional Economics (BITRE)
Department of Infrastructure and Transport
Canberra
Australia

December 2012

INTERNATIONAL TRANSPORT FORUM

The International Transport Forum at the OECD is an intergovernmental organisation with 54 member countries. It acts as a strategic think tank with the objective of helping shape the transport policy agenda on a global level and ensuring that it contributes to economic growth, environmental protection, social inclusion and the preservation of human life and well-being. The International Transport Forum organizes an annual summit of Ministers along with leading representatives from industry, civil society and academia.

The International Transport Forum was created under a Declaration issued by the Council of Ministers of the ECMT (European Conference of Ministers of Transport) at its Ministerial Session in May 2006 under the legal authority of the Protocol of the ECMT, signed in Brussels on 17 October 1953, and legal instruments of the OECD.

The members of the Forum are: Albania, Armenia, Australia, Austria, Azerbaijan, Belarus, Belgium, Bosnia-Herzegovina, Bulgaria, Canada, Chile, China, Croatia, the Czech Republic, Denmark, Estonia, Finland, France, FYROM, Georgia, Germany, Greece, Hungary, Iceland, India, Ireland, Italy, Japan, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Moldova, Montenegro, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Kingdom and the United States.

The International Transport Forum's Research Centre gathers statistics and conducts co-operative research programmes addressing all modes of transport. Its findings are widely disseminated and support policymaking in member countries as well as contributing to the annual Summit.

DISCUSSION PAPERS

The International Transport Forum's Discussion Paper Series makes economic research, commissioned or carried out at its Research Centre, available to researchers and practitioners. The aim is to contribute to the understanding of the transport sector and to provide inputs to transport policy design. The Discussion Papers are not edited by the International Transport Forum and they reflect the author's opinions alone.

The Discussion Papers can be downloaded from:

www.internationaltransportforum.org/jtrc/DiscussionPapers/jtrcpapers.html

The International Transport Forum's website is at: www.internationaltransportforum.org or for further information on the Discussion Papers, please email: itf.contact@oecd.org

TABLE OF CONTENTS

1. INTRODUCTION	5
2. COMPONENTS OF THE OPTIMISATION PROBLEM	7
2.1. Deterioration relationships	8
2.2. Maintenance treatments	11
2.3. Road user costs	12
3. THE OPTIMISATION PROBLEM	13
3.1. Welfare maximisation versus cost minimisation	13
3.2. A simplified example of the optimisation problem	13
3.3. The termination problem and residual value	17
3.4. Budget constraints	19
3.5. Multiple treatment types	21
3.6. Other constraints	22
3.7. Alternative objective functions	23
3.8. Other factors affecting maintenance decisions and optimisation	25
4. APPROACHES TO ROAD MAINTENANCE OPTIMISATION	26
4.1. Survey of approaches	26
4.2. Genetic algorithms: the state-of-the-art optimisation method	27
5. MAINTENANCE DEFERRAL	28
5.1. Present value budget constraints versus maintenance deferral	31
5.2. Preventative maintenance	32
5.3. Cost of maintenance deferral	34
6. OPTIMISING THE INVESTMENT–MAINTENANCE TRADE-OFF	28
7. DEFINING AND ESTIMATING THE MAINTENANCE DEFICIT	38
8. OPTIMAL INCENTIVES IN MAINTENANCE CONTRACTS	38
9. IMPLICATIONS FOR ROAD FUNDING	40
APPENDIX: ROAD MAINTENANCE OPTIMISATION LITERATURE	41
BIBLIOGRAPHY	43

Canberra, October 2012

Acknowledgment

The author is indebted to Dr Weiguo Lu of BITRE for assistance with the literature review of approaches to road maintenance optimisation, including preparation of the appendix.

1. INTRODUCTION

Mathematical optimisation models, supported by suitable data, can assist decision making about allocating funds between alternative maintenance tasks and about the size of the maintenance budget over time. The maintenance optimisation problem is, in essence, to find the optimum balance between the costs and benefits of maintenance, while taking into account various constraints (Dekker 1996). For a given road segment, choices have to be made between alternative treatment types and the times to implement those treatments. Where maintenance funds are limited, there is an additional problem of balancing the competing needs of the different segments. Maintenance tends to be underfunded relative to investment because of the smaller, less obvious nature of maintenance works relative to new infrastructure (Semmens 2006, Zeitlow 2006). But deferring maintenance in the short term can be expensive in the long term, a point that can be brought to the attention of decision makers by quantifying the costs of underfunding maintenance.

Maintenance can be defined as 'all the technical and associated administrative functions intended to retain an item or system in, or restore it to, a state in which it can perform its required function' (Dekker 1996). It does not upgrade the asset. In practice, it is common to carry out small upgrades of roads such as widening or shoulder sealing together with rehabilitations. Without maintenance, roads can quickly fall into disrepair leading to increased costs for road users in vehicle operation, time, reliability and safety. If deterioration goes too far, users will be reluctant to use the road with attendant losses of the economic and social benefits the road confers.

The maintenance requirements of gravel, sealed and concrete roads and of bridges differ, however, the same general economic principles apply to all. Gravel roads need to be regraded at intervals of around six months or a year to reduce roughness and resheeted at intervals of some 8 to 10 years. Concrete roads require roughening for safety reasons as usage reduces skid resistance, maintenance and repairs to joints between slabs, crack sealing, and slab replacement. Sealed roads with flexible pavements consist of layers of crushed rock with either a chip seal (a thin layer of bitumen and crushed rock), which keeps out water, or an asphaltic concrete seal (aggregate mixed with bitumen binder), which both keeps out water and adds structural strength. The term 'flexible pavement' refers to the fact that the pavements can deform when loads are applied and then return to their original shape. By contrast, concrete pavements are rigid.

The focus of most literature on optimisation and of this paper is on sealed roads with flexible pavements. They carry most vehicle-kilometres of traffic and command the greater part of maintenance expenditure. Concrete pavements are relatively rare and relatively new, while gravel pavements are only economically warranted for low-trafficked roads. The paper also does not address maintenance of bridges, tunnels, geotechnical structures, and roadside equipment. However, there are similarities between maintenance principles for different types of road infrastructure. For example, Morcous and Lounis (2005) apply the same maintenance optimisation techniques to bridges and Grivas et al. (1993) to concrete pavements as other authors apply to flexible pavements.

Road maintenance can be categorised as:

- Routine: small tasks undertaken frequently — vegetation control, repairing or replacing signs and other roadside furniture, clearing drains and culverts, repainting line markings, patching, crack sealing and pothole repair;
- Periodic: larger tasks undertaken at intervals of several years or more — resealing, resurfacing, overlay, reconstruction; and
- Urgent: unforeseen repairs requiring immediate attention — collapsed culverts, washaways, landslides that block roads (Burningham and Stankevich, 2005).

Optimisation models for sealed roads deal with periodic maintenance and components of routine maintenance that affect roughness or the rate of pavement deterioration, in particular patching, crack sealing and pothole repair. Road providers have considerable scope to vary the types and timing of periodic maintenance interventions. Routine maintenance, on the other hand, comprises tasks that need to be carried out if a road is to remain open to traffic and generally do not vary with traffic volume and composition. For costing purposes, routine maintenance activities not being optimised are usually assumed to be a constant amount per kilometre of road or per square metre of pavement.

The next section of the paper describes the components of the optimisation problem followed by discussion of the problem itself using a simplified numerical example. A large body of published literature exists on techniques for optimising pavement maintenance, mostly from the civil engineering discipline. We provide an overview to indicate the range of techniques applied. The principles for optimising the trade-off between maintenance and construction are then discussed briefly. Ways of defining and measuring the cost of the 'maintenance deficit' and maintenance deferral are proposed. A section is included drawing on the earlier discussion of optimisation principles showing how maintenance contracts can be specified to give contractors the incentive to provide optimal maintenance. The conclusion advances some policy implications from the paper.

2. COMPONENTS OF THE OPTIMISATION PROBLEM

The road maintenance optimisation problem from the point of view of society as a whole involves trading off road agency or maintenance costs against road user costs over time. The three essential components of a road maintenance optimisation model are prediction of future pavement condition, prediction of the effects of maintenance treatments on road condition, and estimation of road user costs as a function of road condition. In the parlance of well-known HDM4 model, these are road deterioration (RD), works effects (WE) and road user effects (RUE), respectively.

Road condition has a number of measureable attributes. The most important ones in the present context are:

- Roughness, measured in metres per kilometre international roughness index (IRI) units¹;
- Rutting, measured by mean rut depth;
- Cracking, measured by area or percent of area cracked; and
- Pavement strength, measured by modified structural number².

Composite indexes combine two or more of the indexes for these attributes into a single measure of pavement quality (Austroads 2007; see for example the 'pavement serviceability index' in Ferreira and Queiroz 2012). Each of these attributes can be objectively measured, but at a cost. Subjective measures based on visual inspection are less expensive and time consuming to implement though less accurate.

¹. Roughness refers to the rideability of the road surface, and indicates the relative comfort offered to road users. It is measured from the movement of a car's rear axle relative to its body as the vehicle travels along the road at a constant speed. Roughness for a lane is reported in terms of the International Roughness Index (IRI) — the average of the results of the application of a computer model of a standard 'quarter-car' to the measured longitudinal road profile of each wheelpath. Two methods can be used to calculate lane roughness from profile data, 'profile averaging' (half-car simulation, giving Lane IRIhc) and 'IRI averaging' (quarter-car simulation, giving Lane IRIqc). The World Bank has adopted IRI averaging (Lane IRIqc) for use in HDM products. The measured profile is filtered with a moving average that has a 250 mm base length to simulate the effect of the tyre-to-road contact area and the way in which a tyre envelops the small sharp unevenness features. The smoothed profile is then further filtered by application of the quarter-car model with specific parameter values that define the Single Wheelpath IRIqc at a simulated speed of 80 km/h. The IRIqc is an accumulation of the simulated motion between the sprung and unsprung masses in the quarter-car model, divided by the length of the profile. The resulting IRI statistic has dimensionless units of metres/kilometre (m/km). Much of the literature from the United States refers to IRI with dimensionless units of in/mi (inches per mile). One IRI (m/km) is equal to 63.36 IRI (in/mi). (Austroads 2007)

². Pavement strength refers to the ability to carry repeated heavy axle loadings before the pavement shows unacceptable signs of structural and surface distress that seriously compromise its function (Austroads 2008a). Structural number (SN) is a measure of the total thickness of the road pavement with each layer given a weight according to its strength, in other words, a linear combination of the layer strength coefficients and thicknesses of the individual layers above the subgrade (Morosiuk et al. 2004). The modified structural number (SNC) takes account of the subgrade contribution to pavement strength. The SNC would equal the SN if the pavement were designed to carry the same traffic on a subgrade with a California Bearing Ratio of 3% (Austroads 2008a).

2.1. Deterioration relationships

Deterministic approaches

Road deterioration models can be deterministic or probabilistic relationships. For the deterministic type, the relationship can be, mechanistic, empirical or a combination of both. The mechanistic approach uses fundamental theories of pavement behaviour to model deterioration trends. This approach produces models that are more easily transferable to different pavements and conditions, but are usually very data-intensive. Empirical models are less structured, relying mostly on statistical analysis of locally observed deterioration trends. Empirical models may not be transferrable to other locations where conditions are different. The combined mechanistic–empirical approach attempts to create models with moderate data requirements and that can be transferred to different pavements and conditions with changed calibration parameters. Since the seminal work by Paterson (1987), a mechanistic–empirical approach has been widely used for modelling purposes. For example, the RD and WE relationships included in the HDM suite of models are mainly structured empirical models (Morosiuk et al. 2004).

Deterministic deterioration relationships can be represented either on a change (incremental) or level (aggregate) basis. Incremental models predict the change in condition from the current situation. They can use any start point and hence are more flexible.

From Paterson (1987), the general form of an incremental model for road deterioration can be summarised as the total change in roughness during a time period equals the sum of changes in roughness due to

- structural deformation: a function of strength, condition, axle loads during the period, the environmental coefficient, and the increase in rut depth during the period,
- surface defects: a function of the increases during the period of cracking, patch repairs that protrude above or below the surrounding surface, and open potholes, and
- age and environmental factors: a function of the environmental coefficient multiplied by the length of the time period.

The structural deformation component in Paterson’s model features a term $134e^{mt}SNCK^{-0.5}\Delta NE$, where m is the environmental coefficient, t is pavement age, $SNCK$ is the modified structural number adjusted for the effect of cracking, and ΔNE is millions of equivalent standard axle (ESA) loads per lane during the period. The environmental coefficient has to be adjusted according the climate and will be higher in wetter areas. A wet environment combined with high axle loadings will lead to a higher deterioration rate. The third term, age and environmental factors, specifies an environment-related roughness increase that occurs regardless of axle loads.

Further relationships are needed to predict the progression of cracking, rutting and potholing. For cracking, Paterson provided an equation to predict the time in years until the first crack appears: $13.2e^{-(\Delta NE/SNC^2)}$. From that time on, $SNCK$ reduces. The implication is that pavement deterioration proceeds at a higher rate after the first crack appears allowing moisture ingress, which is illustrated in Figure 1. The optimum time for a reseal is likely to occur around this time. Paterson notes that potholing tends to cause only a

small contribution to pavement deterioration, with the effect not being realised until later years. Paterson's incremental model forms the basis for modelling the road deterioration (RD) relationship in the HDM-III and HDM4 models.

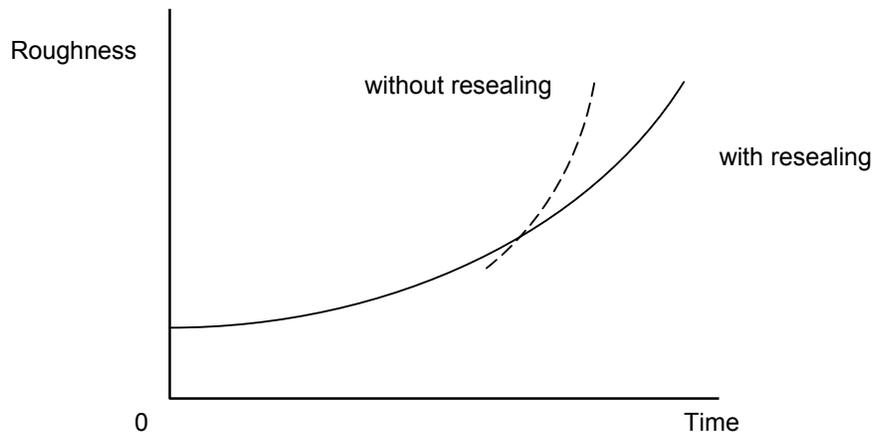


Figure 1. **Roughness as a function of time with and without resealing after cracking sets in**

Source: ATC (2006).

If we write the annual change in roughness as being equal to only the sum of the first term of the structural deformation component and the age-environmental term, we obtain

$$\frac{dR}{dt} = 134e^{mt}SNCK^{-5}\frac{dNE}{dt} + mR(t)$$

Solving this differential equation and setting the initial roughness level to R_0 , gives

$$R(t) = [R_0 + 134SNCK^{-5}NE(t)]e^{mt}$$

which has the same form as Paterson's 'aggregate roughness trend algorithm'

$$R(t) = [R_0 + 725(SNC + 1)^{-4.99}NE(t)]e^{0.0153t}$$

where $NE(t)$ is cumulative ESAs until time t in millions per lane. The aggregate algorithm provides a relationship between roughness and time suitable for strategic analysis assuming pavements are structurally designed for their traffic loadings and well maintained so they suffer only limited deterioration from structural causes.³

³. Paterson and Attoh-Okine (1992) published a modified version of this equation for pavements that do not have extensive distress data. The modified equation should be applied only to pavements that are

Probabilistic approaches

There is a large stochastic element in pavement deterioration caused by unpredictable and unmeasurable factors. Examples are the quality of the materials and workmanship in constructing and maintaining the pavement and drains, the characteristics of the sub-grade, and the weather combined with heavy vehicle loadings. HDM4 algorithms have calibration factors that must be set to match local conditions. 'Calibration sites', where road condition and maintenance works are recorded at regular intervals over long periods of time, are ideal for setting calibration factors. If the deterioration algorithm is well calibrated, it should provide the mean value of a probability distribution within which the actual deterioration path will lie. In common with all forecasts, this distribution will widen as the relationship is projected further into the future. The probability distribution can be derived using Monte-Carlo methods (for example, Jawad and Ozbay 2006).

The Markov chain approach is very common in the literature. Road condition is specified as a discrete variable defined over a number of 'states'. A stochastic process is considered a 'Markov process' if the probability of a future state in the process depends only on the state and on actions taken in the immediate preceding period, not periods further in the past. A Markov chain is a series of transitions between states having the Markov property.

Key to Markovian pavement performance modelling is the specification of Transition Probability Matrixes (TPMs) that indicate the probability that a pavement in each state will change to another state. Transition probabilities are obtained from past data or expert judgement (Li et al. 1997, Morcous and Lounis 2005). In the absence of any treatment, a pavement can only remain in the same state or deteriorate to a lower state, never rise to a higher state. Often, it is assumed that a pavement cannot deteriorate by more than one state in a single time period (Ortiz-Garcia et al. 2006). A different transition matrix is required for each treatment type including the null treatment. As shown below in the appendix, a variety of optimisation techniques have been used in conjunction with Markov chain maintenance models. Model results include total expenditures for each treatment type for the network in each year and lengths or proportions of the network in each condition state at the end of each year.

Traditionally, TPMs are treated as being homogenous (stationary), that is, the road network will always deteriorate according to a fixed TPM. There is an implicit assumption that traffic and environmental conditions stay constant throughout the analysis period, which is not plausible for most real-world pavement situations (Li et al. 1996 and 1997). The problem can be addressed by using a non-homogenous (non-stationary) Markov process where the TPM changes over time.

Morcous and Lounis (2005) discuss the advantages of stochastic models over deterministic models, including better handling of uncertainties, consideration of current pavement conditions in predicting future conditions, and practicality in dealing with large-sized networks due to their computational efficiency and simplicity of use. Hence, they are typically applied for estimating long-term budgets and making needs projections at the network level. Markov chain models cannot be used for planning maintenance on specific roads, such as section X needs treatment Y in year Z.

maintained at low cracking levels (<30% of area). The modified equation is $R(t) = [R_0 + 263(SNC + 1)^{-5}NE(t)]1.04e^{mt}$.

A number of studies have investigated the relationship between deterministic and probabilistic prediction models in pavement management. For example, Li et al. (1997) developed a method to convert a deterministic model into a Markov model. Bekheet et al. (2008) compared the performance of a deterministic pavement prediction model and a Markov-based system. Validation was made of both systems against actual measured pavement condition data. The results showed that both systems performed well.

2.2. Maintenance treatments

There are three main treatment categories for bitumen roads: resealing, rehabilitation and reconstruction.

Morosiuk et al. (2004) lists resealing treatments as cape seal (a chip seal followed within a few days by a slurry seal or microsurfacing) and single or double surface dressing (chip seal or slurry seal). In each case, the treatment may or may not include shape correction. Reseals fill minor cracks, restore skid resistance, and protect the surface from aging. They do not add structural strength to the pavement. Only when combined with shape correction work does resealing improve roughness.

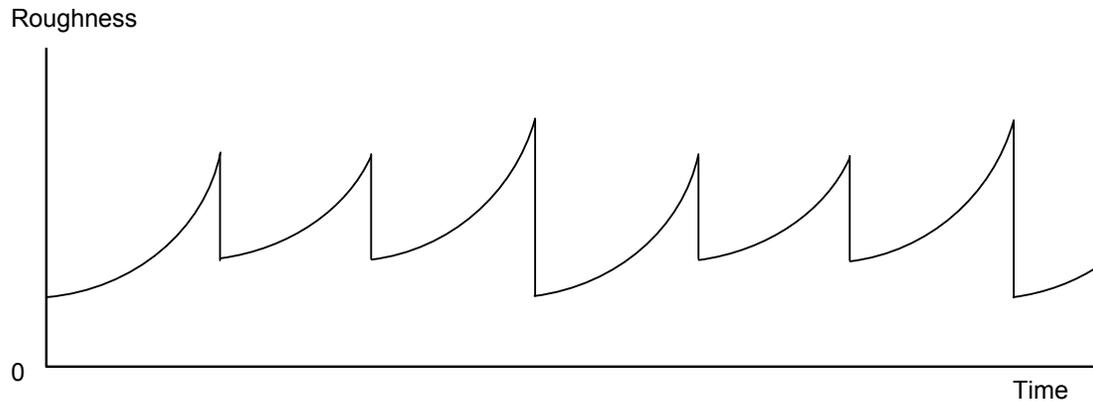
Rehabilitation treatments include mill and replace, overlays of rubberised asphalt, dense-graded asphalt or open-graded asphalt, inlays and thin overlays (Morosiuk et al. 2004). Milling and replacing involves removing the pavement to a certain depth, mixing it with a binder and relaying it. An inlay involves planning off the old surface before the asphalt overlay. Reconstruction involves replacement of one or more pavement layers even down to the subgrade (Morosiuk et al. 2004).

Routine maintenance treatments that affect pavement condition are crack sealing and patching. Crack sealing prevents wide cracks from developing into potholes and inhibits water ingress that will lead to a loss of pavement strength. (Morosiuk et al. 2004). Patching potholes reduces roughness and water ingress.

Each treatment type has a cost per square metre, an amount by which it reduces roughness or to which it restores roughness (reset value), and effects on one or more parameters in the deterioration relationship. A reconstruction or thick overlay will reset roughness to the level of a new pavement, typically 2 m/km IRI or less, as well as raising or restoring the structural number. Thinner overlays and treatments that replace surface layers will reduce roughness and increase strength in varying degrees.

Figure 2 illustrates how roughness progresses over time in cycles starting with a new pavement. Rehabilitation treatments reduce roughness, but not necessarily all the way down the level of a new pavement, and do not restore lost strength in lower pavement layers. Hence, after a number of rehabilitations (two in Figure 2), a reconstruction becomes necessary.

Figure 2. Illustrative pavement life cycles with rehabilitations and reconstructions



Source: ATC (2006).

2.3. Road user costs

Roughness affects road users' costs in a several ways. It reduces vehicle speeds, which increases time taken and alters fuel consumption. Greater rolling resistance increases fuel consumption given speed. Roughness causes wear and tear on vehicles, in particular, on tyres. Road user cost models feature a relationship between speed and volume–capacity ratio, which takes account of 'free speed' and congestion. At low volume–capacity ratios, where there is no congestion, vehicles travel at free speed. Roughness impacts on free speed. In HDM4, free speed is estimated as the 'probabilistic minimum of five constraining speeds based on driving power, braking capacity, road curvature, surface roughness and desired speed' (Odoki and Kerali 2006). Free speed is constrained by the legal speed limit. Time taken, which comprises the greater part of users' costs, is inversely proportional to speed. Research indicates that pavement roughness does not affect speeds until it rises above 4.5 m/km IRI according to Opus (1999) or 5 m/km IRI for cars and 3 m/km IRI for articulated trucks according to research reported by McLean and Foley (1998).

When all the ways in which roughness affects road user costs are combined, relationships between user roughness and user costs tend to be flat up to around 3 m/km IRI (Opus 1999). McLean and Foley concluded that then-current research suggests that over the range from 1.5 to 6.5 m/km IRI, road user costs excluding time rise 4.5% for cars and 5% for articulated trucks per IRI unit. With time costs included, the corresponding rises are 3% and 5.5% respectively. The slope of the curve beyond the flat part is affected critically by the legal speed limit. Opus (1999) stated that the benefit from reducing roughness from 6 to 3 m/km IRI in a 50 km/h zone is only half that achieved in a 100 km/h zone.

While HDM4 makes detailed calculations of road user costs, it has limited optimisation capabilities. Maintenance optimisations on other platforms with greater optimisation capabilities require simple user cost relationships. Such relationships can be obtained by regressing of road user cost estimates per vehicle kilometre from HDM4 against roughness and other variables that affect road user costs such as 'rise and fall', curvature and payload. Typically, for optimisation modelling, the relationship between roughness

and road user cost is assumed to be linear (for example, Li and Madanat 2002) or quadratic (for example, Ferriera and Queiroz 2012).

3. THE OPTIMISATION PROBLEM

3.1. Welfare maximisation versus cost minimisation

For economic analyses of road pricing and investment decisions, the optimisation problem is specified in terms of welfare maximisation. Economic welfare from a road is given by users' willingness-to-pay (the area under the demand curve for the quantity demanded) minus users' costs, external costs and investment and maintenance costs. Willingness-to-pay changes with traffic level only when an improved road generates new traffic or diverts traffic from other roads. For maintenance optimisation, it is usual to assume an absence of any relationship between road condition and traffic level or vehicle mix. Road users base their demand decisions on their generalised costs for an entire trip. Most trips will comprise travel over many road segments at different stages of their life cycles. Unless any individual segment is allowed to deteriorate to the point where it can damage vehicles or cause significant traffic delays, the condition of any individual road segment, other things held equal, will have negligible effect on demand for road usage. Willingness-to-pay is therefore assumed to be constant with respect to the condition of individual road segments. With the sole positive term in the welfare function fixed, maximum welfare occurs when the sum of the costs (the negative terms) in the welfare function are minimised.

The cost minimisation approach avoids the need to specify a base case. For cost-benefit analyses of investment projects, a base case is required in which the investment is not made. There are many alternatives to a particular maintenance treatment at a particular time — the same treatment at another time and a range of other treatments undertaken at a range of possible times. It is possible to specify a do-minimum case against which to compare alternative scenarios of treatments and timings for same road segment. The HDM4 model requires it. However, there is considerable arbitrariness in selecting a do-minimum scenario.

3.2. A simplified example of the optimisation problem

To examine the essential principles of maintenance optimisation, we use a simplified example in which there is just one maintenance treatment type, a major rehabilitation.

Road users' costs are treated as a function of time since the last rehabilitation $c(t)$. This function is a composite of road users' costs as function of roughness $c = c(R)$ and roughness as a function of time, $R = R(t)$. The function $c(t)$ may differ between rehabilitation cycles due to changes in the traffic volume and vehicle mix.

In figure 3, the first rehabilitation is carried out when the pavement is T_0 years old, the second at T_1 and so on. Each rehabilitation restores roughness to its initial level in the

cycle. At time zero, the pavement is δ years old, where $0 \leq \delta \leq T_0$. The road agency incurs rehabilitation costs in year $T_0 - \delta$, then again in year $T_1 + T_0 - \delta$ and so on.

Figure 3. Effect of delaying a rehabilitation

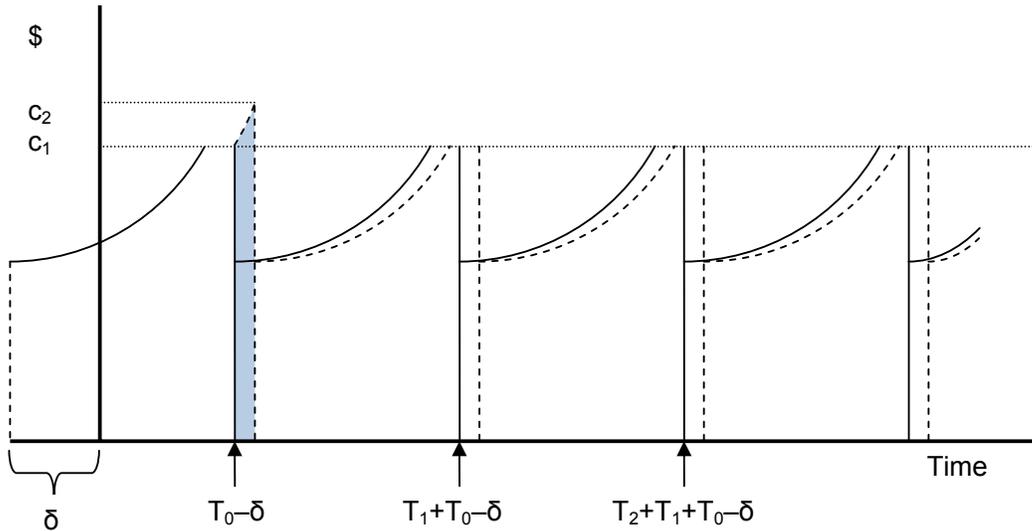


Figure 3 also shows the effect of delaying by one year the time of the first rehabilitation with no changes to the intervals of time between the subsequent rehabilitations. During the year of the delay, immediately after $T_0 - \delta$, users face additional costs equal to approximately $(c_1 + c_2)/2$. This additional cost to society is offset by the cost of rehabilitation to the road agency in year $T_0 - \delta$ being delayed by one year, as well as all future costs, both to the road agency and to users (including user costs with a new pavement in the year after $T_0 - \delta$). The optimum time to undertake the first rehabilitation can be found where the marginal cost of an additional year's delay, which is incurred by users, equals the marginal benefit, which accrues to both users and the road agency. The same rule can be applied to determine the optimum times for all future rehabilitations.

To explore the model further, we make further simplifying assumptions that $c(t)$ and rehabilitation cost, m , are the same for all cycles so the cycles have a uniform time length. The optimisation is done over an infinite time horizon. With continuous compounding, the present value of combined road agency (maintenance) and road users' costs, or Total Transport Costs (TTC) in HDM4 terminology, for a cycle that commences with a rehabilitation is

$$PV_{cycle} = m + \int_0^T c(t)e^{-rt} dt \quad (1)$$

The present value of a monetary amount, a , paid at time zero and then forever afterwards at intervals of T years is $a/(1 - e^{-rT})$ with continuous compounding. The present value over an infinite time horizon of TTCs for a pavement of age δ , in which all cycles are identical is

$$PV_{TTC} = \int_0^{T-\delta} c(t + \delta)e^{-rt} dt + \frac{e^{-r(T-\delta)}}{(1-e^{-rT})} PV_{cycle} \quad (2)$$

To find the optimum cycle time,

$$\frac{dPV_{TTC}}{dT} = c(T)e^{-r(T-\delta)} - \frac{re^{-r(T-\delta)}}{(1-e^{-rT})^2} PV_{cycle} + \frac{e^{-r(T-\delta)}}{(1-e^{-rT})} c(T)e^{-rT} = 0 \quad (3)$$

which reduces to

$$c(T) = \frac{rPV_{cycle}}{(1-e^{-rT})} \quad (4)$$

The optimum occurs where the cost to users of extending cycle time by one year, $c(T)$, equals the benefit from delaying all future cycles by one year, given by the present value of TTCs for future cycles, $PV_{cycle}/(1 - e^{-rT})$, multiplied by the discount rate, the amount of that would be earned by the resources if invested elsewhere for one year. The initial pavement age, δ , is irrelevant. Higher values of m will be associated with higher cost minimising values of T . In other words, the more expensive it is to maintain roads, the lower will be the optimum standard. Since road user costs consist of costs per vehicle times numbers of vehicles, higher traffic levels lead to higher values of $c(T)$ and hence justify higher maintenance standards.

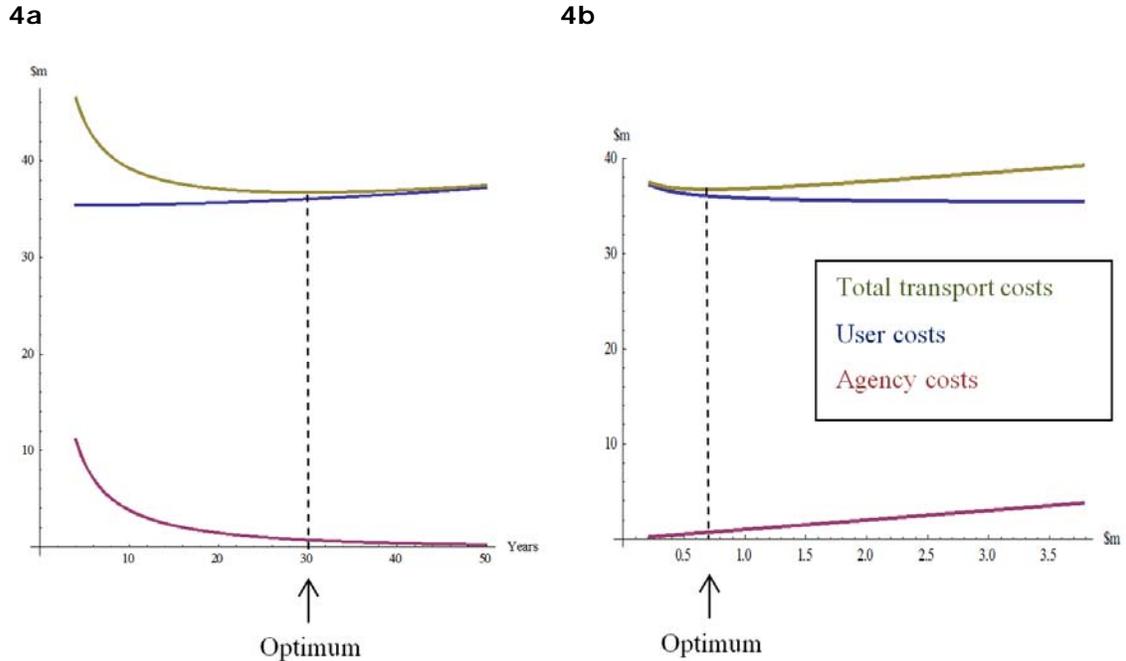
Another way to view the problem is to separate costs to users (PVU) from costs to the road agency or maintenance costs (PVM) as shown in equations (5) and (6).

$$PVU = \int_0^{T-\delta} c(t + \delta)e^{-rt} dt + \frac{e^{-r(T-\delta)}}{(1-e^{-rT})} \int_0^T c(t)e^{-rt} dt \quad (5)$$

$$PVM = \frac{e^{-r(T-\delta)}}{(1-e^{-rT})} m \quad (6)$$

Figure 4 shows these curves and $PV_{TTC} = PVU + PVM$ plotted first, against cycle time (4a), and second, against PVM (4b). The PVM curve in the latter case would be a 45 degree line if the scales were the same on the two axes. The curves have been constructed making realistic assumptions for a one-kilometre length of road with a traffic level of 6000 vehicles per day, a discount rate of 5% and a pavement age at year zero of $\delta = 5$. Pavement deterioration follows Paterson's aggregate algorithm given above (initial roughness 1.5m/km IRI and SNC = 5). A quadratic equation was used for user costs as a function of roughness. Rehabilitation costs of \$1.91 million have been adjusted to make the optimum cycle time 30 years where the roughness level reached is 4.84 m/km IRI. The present value of rehabilitation costs with optimal cycle time is \$0.7 million. Figure 4a shows how agency costs fall and user costs rise as cycle time increases. Moving to the right implies a lower maintenance standard. The reverse occurs in Figure 4b, where moving to right implies a higher maintenance standard. Figure 4b omits cycle times below 10 years to avoid unduly extending the horizontal axis.

Figures 4a and 4b. **Present values of costs as functions of cycle time and the present value of maintenance spending**



At the optimum point, agency costs amount to about 2% of total transport costs. In reality, this would be larger with costs of routine maintenance and resals added. Nevertheless, it is true that in maintenance optimisation problems, road agency costs are small in comparison with user costs. As is usually the case with optimisation problems of this type, the total curve is fairly flat on either side of the optimum. Being out by a few years on either side of the optimum imposes only a small additional cost on society — an additional \$57 000 present value for rehabilitating at 35 year intervals and \$73 000 for 25 year intervals. However, if additional costs of this magnitude were incurred for many kilometres, they could add up to a substantial amount.

Increasing the discount rate from 5% to 10% raises the optimum time interval between rehabilitations from 30 to 34 years. Higher discount rates lead to lower optimal maintenance standards because they increase the gain from delaying maintenance spending.

Expressing the optimisation problem as: find the level of maintenance costs that minimises $PVTTTC = PVU + PVM$, the optimal condition is that

$$\frac{dPVTTTC}{dPVM} = \frac{dPVU + dPVM}{dPVM} = \frac{dPVU}{dPVM} + 1 = 0 \text{ or } -\frac{dPVU}{dPVM} = 1 \quad (7)$$

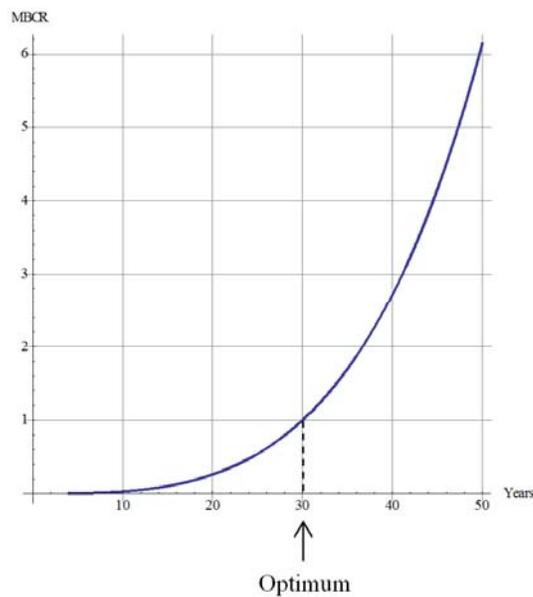
At the optimum, the marginal benefit–cost ratio (MBCR) equals one. The MBCR is the present value of the benefit to users from spending an additional present value of a dollar on maintenance. We can express the optimum condition in terms of the MBCR as

$$MBCR = -\frac{dPVU}{dPVM} = -\frac{dPVU}{dPVM} = \left[c(T) - \frac{r \int_0^T c(t)e^{-rt} dt}{(1-e^{-rT})} \right] / \frac{rm}{(1-e^{-rT})} = 1 \quad (8)$$

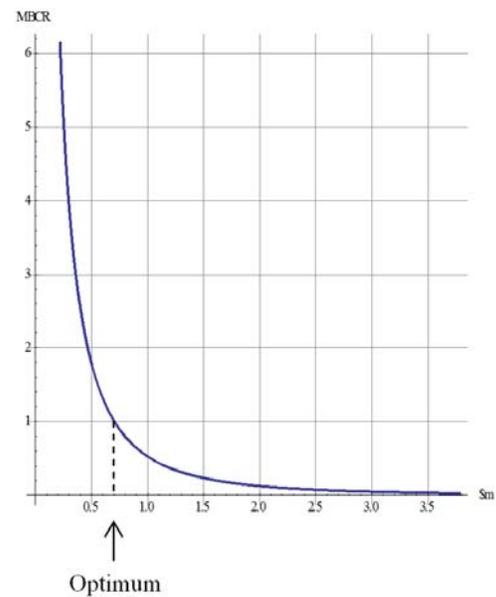
The numerator is the saving in user costs at the end of the cycle as a result of the shortened cycle, $c(T)$, minus the cost of bringing forward future user costs by one year. The denominator is the increase in the present value of maintenance costs with rehabilitation expenditures brought forward by one year. Figures 5a and 5b show the MBCR in our numerical example plotted against cycle time and against the present value of maintenance costs respectively.

Figures 5a and 5b. **Marginal benefit–cost ratio plotted against cycle time and maintenance spending**

5a



5b



The MBCR rises as cycle time increases and falls as maintenance spending increases. The optimum cycle time and present value of maintenance costs can be read off the graphs where the MBCR equals one. The MBCR is quite sensitive to non-optimal maintenance timing — 0.5 for a 25 year cycle time and 1.7 for a 35 year cycle time.

The MBCR equals negative the slope of the TTC curve in figure 4b plus one.

$$-\frac{dPVTTC}{dPVM} + 1 = -\frac{dPVU+dPVM}{dPVM} + 1 = -\frac{dPVU}{dPVM} = MBCR \quad (9)$$

3.3. The termination problem and residual value

To determine the optimal time for any rehabilitation, one needs to know the costs and timings of rehabilitations for many years into the future until such time as discounting

makes changes to future cycle timings of negligible significance. If one terminates the number of years over which maintenance is optimised (the time horizon or analysis period), the model may save costs by extending the last cycle in order to push the final rehabilitation out just beyond the final year. If this occurs far into the future, the effect on rehabilitation times in the near future will be minor, the more so the higher the discount rate. Hence one solution to the termination problem is to extend the analysis period far beyond the period of interest.

A more practical solution is to minimise PVTTTC over a limited number of years minus a 'residual value' or 'salvage value' of the road asset at the end of the analysis period. The absolute magnitude of the residual value is unimportant. What matters is that it varies inversely with pavement age (or directly with pavement condition) at a rate commensurate with the cost of reversing the increase in age (or decline in condition). Then, if the model attempts to reduce PVTTTC by pushing the last rehabilitation just beyond the analysis period saving cost m dollars, the gain is exactly negated by the residual value falling by m dollars.

We can show that, with straight-line depreciation, the residual value approach is approximately equivalent to assuming the last cycle is repeated unchanged into the infinite future. In our simple model, let the pavement age, δ^* , be the age at the end of the analysis period, time t^* . The earlier the last rehabilitation in the analysis period occurs before year t^* , the older will be the pavement in year t^* . The range over which PVTTTC in year t^* can vary with pavement age is m . To demonstrate this, the difference in PVTTTC between an old pavement just prior to rehabilitation, $\delta^* = T$, and a new pavement, $\delta^* = 0$, is

$$PVTTTC|_{\delta^*=T} - PVTTTC|_{\delta^*=0} = \left[\frac{PV_{cycle}}{(1-e^{-rT})} \right] - \left[\int_0^T c(t)e^{-rt} dt + \frac{e^{-rT}PV_{cycle}}{(1-e^{-rT})} \right] = m \quad (10)$$

In our numerical model, with T set equal to the optimal time, $PVTTTC|_{\delta^*=T} - PVTTTC|_{\delta^*=0}$, = \$38.4m - \$36.5m = \$1.9m, the cost of a rehabilitation.

Say the residual value (RV) is set at

$$RV(\delta^*, T) = V + PVTTTC(\delta^*, T)|_{\delta^*=T} - PVTTTC(\delta^*, T) \quad (11)$$

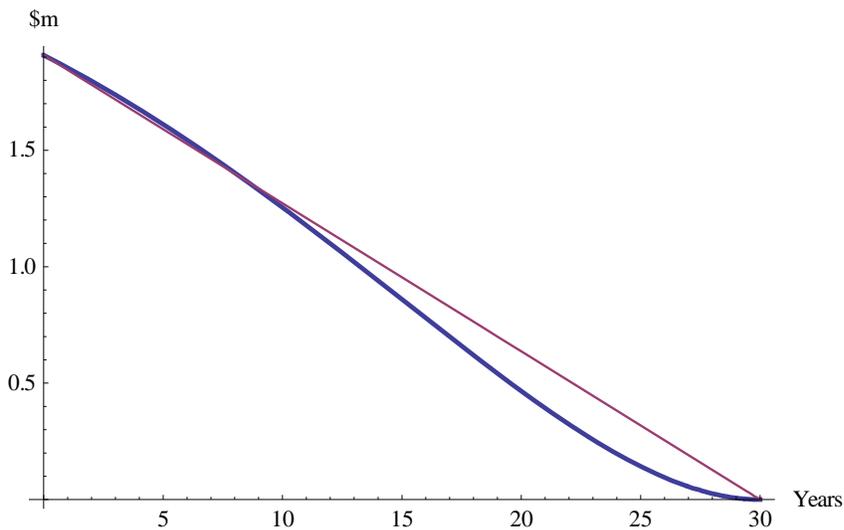
where

- V is a constant equal to the pavement's book value just prior to rehabilitation.
- T is the time between rehabilitations.
- $PVTTTC(\delta^*, T)|_{\delta^*=T}$ is a constant equal to the present value of post-analysis period TTCs discounted over an infinite time horizon to t^* , with cycle time T . In other words, it is the PVTTTC value for a pavement just about to be rehabilitated.
- $PVTTTC(\delta^*, T)$ is the present value of post-analysis period TTCs discounted to t^* . It rises with pavement age from $\delta^* = 0$ to $\delta^* = T$.

Then, minimising PVTTTC from years zero to t^* minus the residual value at t^* discounted to year zero is the same as minimising PVTTTC over an infinite time horizon with identical cycles after the analysis period. When differentiating PVTTTC to find the optimal value of T , the constant terms in the residual value, $V + PVTTTC(\delta^*, T)|_{\delta^*=T}$, drop out and so have no effect on the optimisation.

Figure 6 shows residual value as a function of pavement age, δ^* , given by $PVTTTC(\delta^*, T)|_{\delta^*=T} - PVTTTC(\delta^*, T)$ compared with residual value as estimated by straight line depreciation with T set at the optimal value of 30 years. The residual value for a new pavement is $m = \$1.91$ million in our model, which reduces to zero as the pavement age reaches T years. The graph suggests that straight-line depreciation is a reasonable approximation of how the present value of future TTCs beyond the optimisation time horizon changes between the ages of zero and T .

Figure 6. **Residual value as a function of pavement age: model and straight line depreciation**



Thus, if we let $RV(\delta^*, T) = V + PVTTTC(\delta^*, T)|_{\delta^*=T} + m \left(1 - \frac{\delta^*}{T}\right)$, (12)

the solution to the optimisation problem should not be greatly affected by use of a residual value instead of an extremely long time horizon. To the extent that the actual relationship between residual value and δ^* departs from linearity, there will be some impact on the model's choice of rehabilitation times, but as long as the analysis period is sufficiently long, the impact on rehabilitation times in the near future should be negligible, depending on the discount rate. Other rules for obtaining the residual value, for example, varying it with a measure of pavement condition, will be satisfactory provided they produce a reasonable approximation of the actual residual value curve based on optimisation over an infinite time horizon. Where there is a range of alternative treatment types, the residual value needs to rise by the cost of the particular treatment type and the value added by the treatment should reduce to zero over the life of the treatment.

3.4. Budget constraints

Say we want to minimise PVTTTC, subject to a budget constraint expressed as a present value of rehabilitation costs (PVB), that is,

$$\frac{e^{-r(T-\delta)}}{(1-e^{-rT})} m \leq PVB. \quad (13)$$

Expressing the constraint as a present value is equivalent to assuming funds can be shifted through time by borrowing or lending at an interest rate equal to the discount rate. While not necessarily realistic, it serves as a benchmark because it ensures optimal allocation of limited funds over time and is associated with the MBCR.

For practical optimisation purposes, the difficulty of dealing with a present-value budget constraint over an infinite time horizon can be overcome by undertaking an unconstrained minimisation of $\psi PVU_{0 \text{ to } t^*} + PVM_{0 \text{ to } t^*} - RV$ where ψ is a constant equal to the reciprocal of the MBCR. Assuming the residual value simulates optimisation over an infinite time horizon, the minimum occurs where

$$\psi \frac{dPVU}{dt} + \frac{dPVM}{dt} = 0 \text{ which implies } -\frac{dPVU}{dPVM} = \frac{1}{\psi} = MBCR.$$

Alternatively, one could multiply PVM by a constant equal to the MBCR. The optimisation model would have to be run a number of times to find the MBCR consistent with the budget constraint. With user costs given a lower weighting than agency costs, the shape of actual residual value curve reflecting PVTTC from year t^* to infinity (see Figure 6) would be changed, but numerical modelling by the author suggests the change is insufficient to rule out straight line depreciation as an approximation for the residual value curve.

Where the budget constraint applies to a group of road segments taken together or a network, the optimal allocation of maintenance funds would be found where the MBCR is the same for all segments. If one segment has a higher MBCR than another, shifting maintenance funds from the low-MBCR segment to the high-MBCR segment will generate a net saving in user costs for the two segments. With a present-value budget constraint, optimisation modelling would be simple because one could optimise each segment by itself, applying the same MBCR value to each segment.

Where the relative sizes of the budgets for investment and maintenance are being considered, the optimal split of funds is that which equates the MBCRs for investment and maintenance spending. For investment projects in a budget constrained situation, the economically optimal allocation of funds is found by selecting projects in descending order of BCR until funds are exhausted (or there are no more projects with BCRs above one, in which case the budget constraint is non-binding). The BCR of the last project to be accepted, the 'cut-off BCR', is the MBCR for investment. If the MBCR for maintenance expenditure is above that for investment expenditure, economic welfare could be improved by shifting funds from the investment budget to the maintenance budget, and conversely.

In practice, budgets are not expressed as present values but as amounts that can be spent over a single year or a small number of years. An annual budget constraint would make no sense for a single road segment in isolation. The cost of a rehabilitation for a single segment would far exceed the budget for that segment in the year it occurs. Typically, for modelling purposes, a network budget constraint is set for each year for the first several years over which one is interested, then no constraint thereafter. For example, Archondo-Callao (2008), demonstrating the HDM4 model, imposed uniform budget constraints for the first five years of the analysis only. The tighter the budget constraint for the first five years, the greater the amount of economically warranted

expenditure the model pushes out into the unconstrained period. Maintenance deferral caused by tight annual budgets is considered in detail below.

3.5. Multiple treatment types

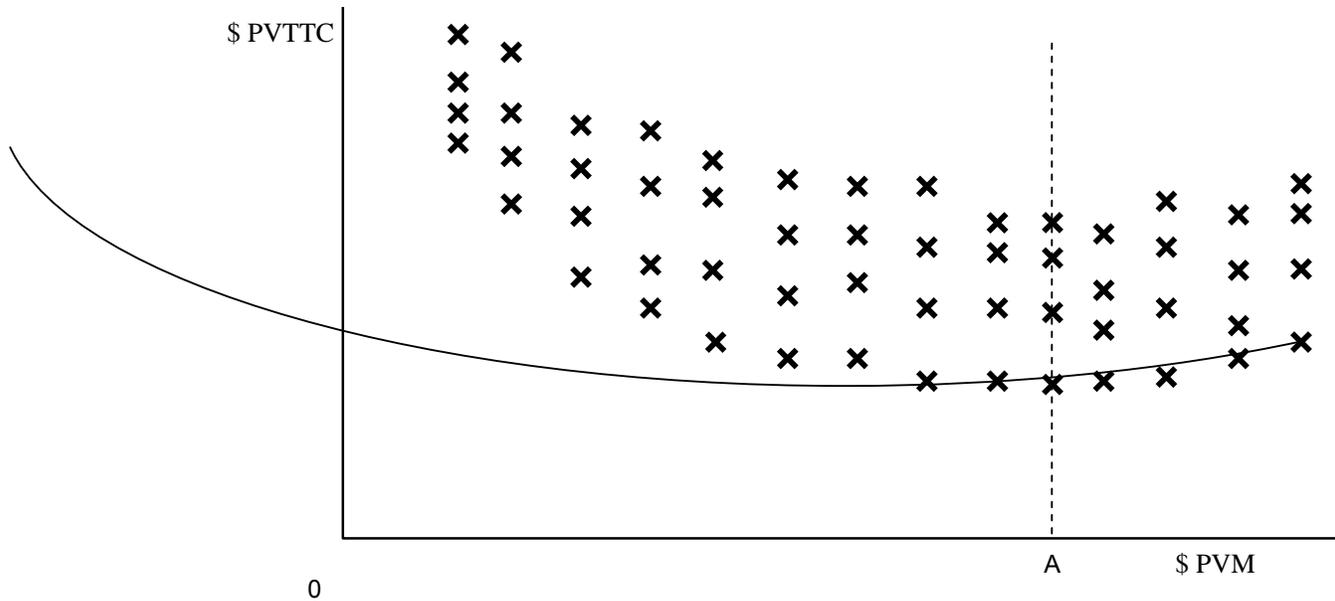
With multiple treatment types to choose between, the optimisation problem becomes much more complex. Instead of a smooth, continuous cost surface with a single minimum point, there are multiple local minimums and discrete choices. Treating time in years as a discrete variable, Golroo and Tighe (2012) note that the number of feasible solutions for N pavement segments with S maintenance actions (treatment types) over a planning horizon of T years is $S^{T \times N}$.

One way to reduce the number of feasible solutions is to schedule maintenance actions over selected years (for example in years 1, 3, 5, instead of 1, 2, 3, 4, 5). Another way is to specify condition-responsive treatment rules instead of years of occurrence, for example, 'rehabilitate as soon as roughness reaches 5 m/km IRI'. The number of possible maintenance actions is increased because the same action can be triggered by multiple condition states. For example, instead of 'rehabilitate' as a single maintenance action, we might have three maintenance actions: 'rehabilitate at 4m/km IRI, rehabilitate at 5m/km IRI, and rehabilitate at 6m/km IRI. The number of possible maintenance actions would increase from S to γS where γ is the average number trigger points per maintenance action ($\gamma = 3$ in our example). But unless the number of trigger points is quite large, the total number of feasible solutions will be less, that is, $(\gamma S)^N < S^{T \times N}$. If the number of trigger points is small, there will be some loss of precision, but as we have shown, small errors in optimum timing have a limited effect on PVTTTC.

A further way to reduce the size of the problem for optimisation modelling is to aggregate segments into bins with similar characteristics in terms of pavement condition, parameters in the deterioration algorithm (structural number, environmental coefficient), traffic level and vehicle mix. Earlier models required high levels of aggregation, due to limited data and computing power. As data availability and computing technology have improved, there has been a trend toward greater disaggregation including dynamic sectioning. The level of aggregation or disaggregation required depends on the question that has to be answered as well as on data and computing power limitations.

Figure 7 shows how the results might appear if PVTTTC values from various solutions under different equality budget constraints were plotted against the present value of maintenance costs. For each spending level, only the solutions with the lowest PVTTTCs are of interest. Joining the minimum values together with a smooth curve would produce a U-shaped curve relationship analogous to the PVTTTC curve in figure 4b. The minimum point of the curve, at spending level A , is the unconstrained optimum. The MBCR for any budget constraint can be obtained from the slope of the curve. Since the slope is evaluated between two points, some distance from each other, the MBCR may be referred to as an incremental BCR (IBCR).

Figure 7. Illustration of present values of total transport costs from maintenance simulations with different budget constraints



Source: Based on diagrams in Tsunokawa and Ul-Islam (2003) and ATC (2006).

If the budget constraint in the model is specified as a series of annual spending constraints, MBCR estimates will differ from MBCRs obtained from present-value budget constraints, depending on the particular years in which spending is increased. The definition of the MBCR for maintenance is considered further below in section 7.4 in the context of maintenance deferral caused by tight short-term annual budget constraints.

3.6. Other constraints

Budgets are not the only constraints imposed in road maintenance optimisation models. Additional constraints may be necessary to prevent corner solutions or the model extrapolating relationships beyond the range over which they apply. For example, Tsunokawa and Ul-Islam (2003), using the HDM4 model, imposed a 5-year minimum overlay interval in the simulations 'to avoid two consecutive condition-responsive overlays from being applied in an impractically short period of time'. The same problem can arise if some treatment types have continuous elements. Overlay thickness can be treated as a continuous variable with thicker overlays costing more to implement. In such cases, there is a possibility that the model will converge towards a corner solution of very thin overlays applied as frequently as the model will allow. The source of the problem is most likely extrapolation of relationships beyond their applicable ranges. It can be addressed either by changing the relationships in the model or by introducing an additional constraint restricting the roughness level at which the treatment is made or the overlay thickness. Changing the relationships is more theoretically correct but adding a constraint is more practical.

At the other extreme, on very low trafficked roads, the model may find it optimal to rehabilitate only at roughness levels that are so high that, in practice, the pavement would be falling apart. It may then be necessary to impose upper limits on the roughness levels at which treatments are undertaken.

Governments may require upper limits to be imposed on roughness levels on some roads to meet community expectations. Sometimes called 'community service obligations' or 'public service obligations', these are cases where a road agency is required to provide services at above economically efficient levels to meet social or equity objectives. Optimal road standards are strongly correlated with traffic levels. In developed countries it is common for rural roads to be provided and maintained at levels that could not be justified by economic criteria given their traffic levels. The cost of over-provision to meet community service obligations is greater where there are budget constraints because such roads divert funds from roads with MBCRs above one. Maintenance optimisation modelling could be used to estimate the cost of community service obligations by comparing the PVTTTC values with and without minimum standard constraints.

Availability of physical resources to undertake certain treatments (manpower, equipment and materials) may impose further constraints (Chan et al. 2001). Davis and Van Dine (1998) included in their maintenance optimisation model minimum and maximum amounts of each treatment that can be deployed in each year. The minimums were 'introduced to avoid a solution that calls for extreme shifts in pavement material production from year to year'.

3.7. Alternative objective functions

Minimising the present value of TTCs without budget or minimum standards constraints yields the most economically efficient solution. The most common alternative approach is to minimise the present value of road agency costs subject to minimum standards constraints. The minimum standards may be determined through community consultation to determine the roughness levels road users consider acceptable (Austroads 2002 and 2009). If road users are not sufficiently aware of the costs of maintaining roads to their desired standards, the result could be uneconomically high standards. However, as ATC (2006) notes, stakeholder consultation can be used to manage expectations as well to obtain information about them.

If community wants cannot be accommodated within available funds, the optimisation problem then becomes maximise standards subject to a budget constraint. This would be straightforward if there were just one standard. However, there will inevitably be different standards for groups of roads with different traffic levels, vehicle mixes and locations patterning the relationship between economically optimal standards and traffic level, vehicle mix and maintenance costs. Road agencies often divide their networks into a hierarchy of sub-networks for this purpose. Multiple iterations may be needed to find the set of standards that, according to subjective judgement, offers the 'right' distribution of standards across sub-networks and regions and fits within the budget. The number of iterations can be reduced by maximising average network condition defined as a weighted sum of roughnesses or other indicators of maintenance standard across the network (Morscous et al. 2005).

As noted previously, HDM4 requires specification of a do-minimum 'base option' to serve as a standard of comparison. Maximising the saving in PVTTTCs compared with the base case ($PVTTTC_{basecase} - PVTTTC$) is the same as minimising PVTTTCs. In the HDM4 Manual,

Odoki and Kerali (2006) define a benefit–cost ratio (BCR) for option m , relative to base option n as $BCR_{m-n} = - (PV TTC_m - PV TTC_n) / PVM_m + 1$. This is only equivalent to the MBCR defined in equations (8) and (9) if both options are on the minimum TTC frontier in figure 7 and are fairly close together. If they are some distance apart on the curve, the slope measured will be for the straight line between the two points, not the instantaneous rate of change implied by the term ‘marginal’. Wrong decisions could ensue if alternative maintenance options, whether for the same road segment or for a network as a whole, were selected to maximise BCR defined thus. As cost–benefit analysis textbooks explain, the decision rule for comparing mutually exclusive options is to maximise net present value, never BCR. Odoki and Kerali’s BCR measure can be used only to compare option m with base option n .

Minimising the present value of road user costs subject to a budget constraint expressed as a present value or as annual amounts over the entire analysis period should lead to the same result as minimising PV TTCs subject to the same budget constraint. This approach could not be used if the budget constraint was expressed as a series of annual amounts for only the first several years and no budget constraints for subsequent years. During the unconstrained period, the model would spend unlimited amounts to keep pavements at the lowest possible roughness levels.

Golroo and Tighe (2012) mention an approach that maximises the present value of ‘effectiveness’ defined as the area under a performance curve multiplied by traffic level and segment length. Similarly, Odoki and Kerali (2006) note that for the IBCR approach to optimisation in HDM4, incremental net present value can be replaced with $\Delta IRI \times \text{length}$, defined as ‘the weighted average change in roughness obtained by comparing the project alternatives using IRI instead of NPV’. As noted previously, road user costs increase little up to a roughness level around 3 to 3.5 m/km IRI units, and then begin to rise at an increasing rate as vehicle speeds are affected. A performance curve related to roughness, could underestimate the negative impact on users of high roughness levels relative to low roughness levels. With the seriousness of high roughness levels downplayed, there would be too little maintenance spending in an unconstrained optimisation, and too much variation in intervention roughness levels between segments permitted in a budget constrained optimisation.

As shown in the literature review below on approaches to optimisation, a number of recent authors have adopted multi-objective approaches. ‘Multi-objective programming’ identifies the Pareto frontier along which no objective can be advanced except at the expense of one or more of the others. Examples of objectives from Fwa et al. (2000) are to minimise maintenance cost, maximise work production (days worked) and maximise network pavement condition. The decision maker then has to select the preferred point on the Pareto frontier. If some objectives are, in fact, constraints, it is easy to eliminate points on the frontier. Indeed, if the constraints are ‘hard’, multi-objective programming approach has no advantage over optimising a single objective subject to constraints except perhaps to address the difficulties genetic algorithms have with identifying optimal solutions close to or on constraints. Multi-objective programming may be advantageous if the constraints are ‘soft’ in the sense of being yet to be determined or open to negotiation. It can help decision makers and negotiators to understand the trade-offs between objectives when selecting a point on the Pareto frontier. It is important to recognise that the Pareto frontier is not an indifference curve or surface. Some points are more economically efficient than others.

3.8. Other factors affecting maintenance decisions and optimisation

The theoretical discussion so far omits many factors that affect maintenance decisions.

Costs of delays to road users while maintenance activities are carried out can be significant. In models, these costs can readily be included with treatment costs. In urban areas with high traffic levels, the need to minimise traffic delay costs affects the type of pavement laid and the times at which the works can be carried out, which adds to treatment costs. HDM4 includes a model for estimating 'work zone effects on traffic and user costs' (Bennett and Greenwood 2004). In urban areas, needs to minimise noise from vehicles on pavements and not to raise pavement levels above gutters and footpaths limit the treatment types available.

The relationships between road safety and maintenance are not well understood and tend not to be included in optimisation modelling. The literature survey in Austroads (2008b) covers relationships between crash occurrence and skid resistance, microtexture, macrotexture, rutting and roughness. The relationship between skid resistance and crashes is well established but studies vary in whether they consider all crashes, wet road crashes or wet road skidding crashes. Rutting becomes a safety concern when water accumulates in the ruts increasing the risk of skidding. Swedish work shows no increase in overall crash rate with increasing rut depth, but shows increasing risk of wet weather loss of control crashes with increasing rut depth and decreasing crossfall (Austroads 2008b). Austroads (2008b) refers to evidence from Sweden of a positive relationship between roughness and crashes. While slower vehicle speeds caused by roughness might be expected to reduce crash numbers and severities, not all drivers reduce speeds sufficiently on rougher roads. Given relationships between road condition measures and crash rates, together with unit crash costs, safety impacts could readily be added to the user cost function in maintenance optimisations. As far as the author is aware, microtexture and macrotexture are not addressed at all in maintenance optimisation literature. Loss of aggregate on the surface, which reduces skid resistance, can, in practice, trigger a maintenance treatment before it becomes due because of cracking or roughness.

Road user's willingness to pay for comfort on smoother roads can be included in the user cost function if suitable estimates are available.

Rehabilitation may be combined with widening or shoulder sealing, which constitutes investment not maintenance because they raise the standard of the existing road above its initial standard. For maintenance optimisation purposes, as long as widening and shoulder sealing have negligible traffic generation or diversion effects and are not being paid for out of the investment budget, the additional costs of the works and user benefits can be treated in optimisation models as if the works were maintenance treatments. Rehabilitation with widening or with shoulder sealing would be considered separate treatments from rehabilitation without these. Widening or shoulder sealing would shift the road user cost function downward.

4. APPROACHES TO ROAD MAINTENANCE OPTIMISATION

4.1. Survey of approaches

Optimisation methods have evolved greatly over the past few decades. The traditional methods of maintenance optimisation were largely based on subjective ranking and prioritisation rules (Morcoux and Lounis 2005). Prioritisation rules can be based either on economic or engineering criteria. Examples of economic criteria are the IBCR and 'marginal cost effectiveness' (improvement in road condition divided by increase in cost). Examples of engineering criteria include road class, traffic volume and quality index. The main weakness of prioritisation methods is that they do not ensure the best possible maintenance strategies when considering long planning time spans (Ferreira and Meneses 2011).

One of the earliest attempts at road maintenance optimisation for a network was made by Abelson and Flowerdew (1975), who used dynamic programming to solve a constrained cost minimisation problem for road maintenance in Jamaica over a 10 year analysis period. Since then, many techniques have been employed to solve pavement optimisation problems, including

- linear programming
- non-linear (including convex) programming
- integer programming
- dynamic programming, and
- genetic algorithms.

Linear programming solves optimisation problems with linear objectives and constraints. Many optimisation problems cannot be satisfactorily represented by linear relationships and require non-linear programming. Convex programming is a special case of constrained non-linear optimisation with the objective function being a concave function and all of the constraints being convex. Dynamic programming deals with large and complex optimisation problems by solving a sequence of smaller problems. Genetic algorithms (GAs) belong to the 'heuristics' family of search methods that provide approximate solutions to optimisation problems. GAs find good solutions, but not necessarily the best, with the benefit of a saving in computing time. They are well suited to solving combinatorial optimisation problems.

The appendix categorises selected studies of pavement maintenance optimisation by method and model type (deterministic or probabilistic, single or multi-objective). Some observations can be made from the table in the appendix.

- Greater reliance is being placed on optimisation instead of prioritisation in multi-year pavement management and planning.
- Up to 2000, most studies were based on single-objective optimisation models. Since then, there has been a growing number of studies using multi-objective

programming approaches with GAs.

- Deterministic performance models are slightly more popular than probabilistic.
- More recent studies tend to rely on GAs.
- HDM4 has extended its optimisation capabilities by relying on convex programming through the use of the steepest descent and conjugate gradient methods.

4.2. Genetic algorithms: the state-of-the-art optimisation method

GAs, first introduced by Holland (1975) and further elaborated by Goldberg (1989), are based on the Darwinian evolutionary principles. Since the early 1990s, various GA methodologies have been developed to solve increasingly complex optimisation problems.

GAs commence by generating a randomly selected parent pool of feasible solutions. The parameters describing each solution are encoded into a genetic representation or chromosome comprised of genes that can be manipulated by 'genetic operators'. Through an iterative process of genetic operations involving copying, exchanging and modifying the genes of chromosomes in the pool, selecting the better solutions and discarding the poorer solutions, the pool evolves towards better solutions. Mutations and new gene pool members may be introduced along the way.

GAs differ from traditional optimisation techniques in a number of ways. GAs retain in memory a pool of feasible solutions rather than one single solution at any one time. GAs use probabilistic transition rules to generate new solutions from the existing pool of solutions, which introduces perturbations to move out of local optimums. The search process is not gradient-based so there is no requirement for differentiability or convexity of the objective function (Fwa *et al.* 2000). GAs can solve multi-objective programming problems. Konak *et al.* (2005) discuss some design issues in multi-objective GA optimisation.

Efforts have been made to find the optimum structure of GAs to handle road maintenance optimisation problems. Chan *et al.* (2001) reviewed methodologies used to handle constraints in GA optimisation and proposed a methodology (the prioritised resource allocation method) that was shown to be computationally more efficient than the conventional techniques such as penalty methods and 'decode and repair'.⁴ Golroo and Tighe (2012) conducted experiments to find the optimal GA settings (in terms of simulation number, mating operator methods and operators' probabilities) for solving complex maintenance optimisation problems. It can be expected that such efforts will continue to improve GA methods leading to better solutions and/or savings in computing time for large optimisation problems.

⁴. If solutions outside constraints are discarded from the gene pool, the GA search algorithm may fail to find an optimum solution close to or on a constraint. One option is to impose a penalty on any solution outside a constraint. The 'decode and repair' method attempts to alter solutions outside constraints to make them feasible. The prioritised resource allocation method in Chan *et al.* (2001) adjusts the solution to ensure it falls within the resource constraints.

5. OPTIMISING THE INVESTMENT–MAINTENANCE TRADE-OFF

Incurring higher investment costs to construct a stronger pavement at the outset saves future maintenance and users' costs. The extreme example is a concrete pavement, which costs much more than a flexible pavement to construct but requires far less future maintenance spending to provide a given level of service to users. For flexible pavements, greater initial pavement strength (higher structural number) leads to a lower deterioration rate, as Paterson's algorithm illustrates.

The optimisation problem has the same form as figure 4a above. With pavement strength on the horizontal axis, as pavement strength increases, construction costs rise and PVTC for maintenance falls. Vertically adding the two curves produces a U-shaped curve. Letting K represent construction costs, the optimum occurs where $\partial K/\partial s = -\partial PVTC/\partial s$. This can be written as $\partial PVTC/\partial K = MBCR_s = 1$, where $MBCR_s$ is the MBCR with respect to pavement strength, defined as the saving in PVTC from an additional dollar spent to increase pavement strength. Higher traffic levels are associated with stronger pavements because the savings in user costs are greater.

The benefits from a stronger pavement could be realised either in the form of lower user costs for the same amount spent on maintenance or lower maintenance costs for the same user costs, or a combination of both. In the models of Small and Winston (1988) and Newbery (1989), the benefit from a stronger pavement is realised entirely in the form of a saving in maintenance costs, with no change to user costs. In their models, the intervention roughness level is exogenous, so a stronger pavement increases the time intervals between rehabilitations with no impact on the present value of user costs. With user costs fixed, their models minimise $K + PVM$. The resulting optimal condition is that pavement strength should be adjusted to set $\partial K/\partial s = -\partial PVM/\partial s$. The marginal investment cost from building a slightly stronger pavement is equated with the marginal benefit of a reduced maintenance cost.

This approach is relevant for the case of a network-wide budget constraint. If pavement strength were increased for one segment, network-wide optimisation would not leave maintenance of that segment unchanged so that only users of that segment would reap the benefit. Rather, the benefit would be realised as a saving in maintenance costs for the stronger segment. The saving would be used to better maintain other segments in the network. The benefit would be transformed into small savings in user costs spread over many segments.

To distinguish between the benefits from additional spending on pavement strength from maintenance, we define the term MBCR with respect to maintenance, $MBCR_m$, as the saving in network-wide user costs from increasing the overall maintenance budget by one dollar. The benefit to society of a one unit increase in pavement strength on segment i is: $-\frac{\partial PVM_i}{\partial s_i} MBCR_m$, where PVM_i is the present value of maintenance costs on segment i . The present value of user costs on segment i is held constant. The benefit

from greater pavement strength is taken in the form of a saving in maintenance costs for segment i , which is added to the overall maintenance budget where each dollar generates a benefit of $MBCR_m$ ⁵. The negative sign is required because $\partial PVM_i/\partial s_i < 0$.

If investment spending is also budget constrained, each additional dollar spent on pavement strength has an opportunity cost above one dollar. For example, if the MBCR for investment were three, an additional dollar from the investment budget spent to build a stronger pavement for a new project would mean forgoing benefits of \$3 elsewhere in the road network. The cost to society of a one unit increase in pavement strength on segment i is: $\frac{\partial K_i}{\partial s_i} MBCR_I$, where $MBCR_I$ is the MBCR for investment.

The optimum pavement strength for segment i occurs where the marginal benefit from a unit increase in pavement strength equals the marginal cost

$$-\frac{\partial PVM_i}{\partial s_i} MBCR_m = \frac{\partial K_i}{\partial s_i} MBCR_I \text{ or } -\frac{\partial PVM_i}{\partial K_i} = \frac{MBCR_I}{MBCR_m} \quad (14)$$

Where both the investment and maintenance budgets are constrained, the optimum pavement strength for segment i is found where the saving in the present value of maintenance costs on segment i from an additional dollar of expenditure on pavement strength, holding the present value of user costs constant, $-\partial PVM_i/\partial K_i$, equals the ratio of the MBCRs for investment over maintenance for the network as a whole.

As already noted, with an optimal split of funds between the investment and maintenance budgets, the MBCRs for the two types of expenditure will be the same making this ratio one. The ratio will be lower, the more constrained maintenance spending is relative to investment spending, which will justify greater spending on pavement strength, which drives down $-\partial PVM_i/\partial K_i$. The effect is to construct fewer new kilometres of road out of the investment budget, using the funds saved to build greater durability into the new pavements. Using HDM4 simulations, Tsunokawa and Ul-Islam (2003) showed that optimal pavement strength is higher with budget constrained maintenance spending. They made an implicit assumption that investment funds are unconstrained ($MBCR_I = 1$). The practice of international aid programs of funding stronger pavements in developing countries where maintenance is accorded a low priority has an economic justification.

It is clearly a 'second-best' outcome to build stronger pavements because there are separate investment and maintenance budgets with different MBCRs. It would be better to ensure similar MBCRs for investment and maintenance.

⁵. $MBCR_m$ here must be calculated from the change in user costs with one dollar added to PVM not from relaxing annual budget constraints over the first several years of the analysis period. Savings in maintenance costs from building stronger pavements accrue over many years, not as a uniform addition to annual budgets over first few years following construction.

6. DEFINING AND ESTIMATING THE MAINTENANCE DEFICIT

An estimate of the 'maintenance deficit' can highlight the extent of a shortfall in maintenance funding. The concept necessarily involves comparison between the existing and a desired road condition or set of policies. The ways of measuring the maintenance deficit suggested below are expressed in terms of comparison with the economic optimum, but there other standards of comparison. Funding to the level of the economic optimum might be considered an unrealistic goal. However, as already noted, there is a strong case for the MBCR for maintenance to be the same as for investment. The amount of maintenance consistent with the MBCR for investment spending could be made the standard of comparison. A potential impediment is that, where considerations other than BCRs play a major role in prioritising investment projects, it may be difficult to identify a cut-off BCR for investment. Predetermined road condition standards are another possible standard of comparison. Specifying standards inevitably involves arbitrariness and subjectivity. They could be set above or below economically optimal levels. A maintenance deficit based on a comparison with 'gold-plated' standards is unlikely to be taken seriously.

The 'maintenance backlog' is the cost of maintenance works that are economically justified at the beginning of the optimisation period. It indicates the funds required to restore network condition to the economically optimal level. Even if the funds were made available to eliminate the backlog in the first year of the analysis period, there may not be sufficient physical resources available to do so. The 'practical maintenance backlog' measure of the maintenance deficit spreads the restoration work over the next several years, typically four or five. The optimisation problem would be to minimise TTCs subject to the constraint that agency costs be equal for the first several years of the optimisation period. The maintenance deficit is then the sum of maintenance costs for the constrained years minus the anticipated budgets for those years.

Tsunokawa and Ul-Islam (2003) refer to the maintenance gap ratio (MGR) defined as $1 - PVM/PVM^*$ where PVM^* is the present value of unconstrained optimal agency (maintenance) costs and PVM is the present value of the maintenance budget. A budget that fully met optimised maintenance needs would have a MGR of zero. The more constrained the budget, the higher the MGR up to a maximum of one when no maintenance is undertaken at all. Estimation of the MGR requires assumptions to be made about the sizes of future budgets over the entire optimisation period.

The MBCR can be a useful measure of the maintenance deficit indicating the value to users of additional maintenance spending and enabling comparisons to be made with the value of additional investment spending.

A related way to express a maintenance deficit is in terms of the net economic value of increasing maintenance funding to the optimal level, either fully unconstrained or practically unconstrained. The result could be expressed as a BCR or a net present value. The BCR would be considerably less than the MBCR because the saving in user costs from the additional dollar spent falls as the maintenance budget is increased. In our numerical example above, if the MBCR was 3.0, which corresponds to a present value of maintenance spending of \$0.36 million compared with the optimal amount of \$0.70 million, the improvement in PVTTTC from increasing spending to the optimal amount would be \$0.25 million = $PVTTTC^* - PVTTTC$ where the $PVTTTC^*$ is the optimum. The BCR would be $1.7 = -(PVU^* - PVU)/(PVM^* - PVM)$. The maintenance gap ratio would

be $1 - \$0.36m / \$0.7m = 0.5$ indicating that the budget is about half the unconstrained optimum.

Maintenance deficit measures could be derived from asset values such as the net depreciation during a year. A net fall in asset value for a network during the year (excluding new assets) could indicate that maintenance spending is failing to keep pace with road deterioration. However, optimal maintenance does not imply that network condition should remain constant over time. For example, optimal maintenance of a network with a bunched pavement age distribution would lead to cyclical expenditure needs.

7. MAINTENANCE DEFERRAL

7.1. Present value budget constraints versus maintenance deferral

Limiting maintenance spending by imposing a present-value budget constraint ensures best use of scarce funds over time, with borrowing or lending as necessary to shift funds across time to when they have the greatest value. Costs to society will be minimised subject to the constraint provided the the interest rate faced by the government is the same as the social discount rate. If they are not, the optimisation problem could be set up to minimise PV/TTC discounted at the social discount rate, subject to a present-value budget constraint discounted at the government's borrowing rate.

Inadequate funding in the short-term can cause departures from the long-term cost minimising path leading to a higher present value of agency costs for the same present value of user costs. As maintenance treatments are deferred, components of the pavement are left vulnerable to damage and so deteriorate more rapidly. The future treatment required to undo the damage can be more considerably expensive than the treatment deferred. If rehabilitation is deferred, damage may occur to lower layers of the pavement and so the next rehabilitation may have to replace pavement layers to a greater depth or involve a thicker overlay, or a reconstruction may be needed. The effect of deferring resealing was illustrated above in figure 1.

In present value terms, the cost saving from deferring maintenance treatments in the short term can be outweighed by the additional cost of more expensive treatments in the future. The distinction between a budget constraint set as a present value and annual budget constraints is therefore important.

Reseals do not by themselves lower roughness but can be combined with surface correction works enabling rehabilitations to be deferred for longer. In theory, such an approach, applied at a high standard, should result in extremely long-lived pavements. In practice, not all reseals and other preventative maintenance tasks will be applied to a high standard and some pavements will not be as resilient as others. A road agency that aims to preserve its pavements as long as possible with reseals and surface correction works will reduce the degree of uniformity in the age distribution of its pavements. At some future time, when the network is comprised mostly of old pavements, there is a risk that widespread pavement failures over a short time period could impose heavy

financial and physical resource demands. This is more likely to occur during a period of wet weather. Probabilistic optimisation modelling could help assess this risk.

7.2. Preventative maintenance

Resealing is not that only maintenance action that preserves a pavement by inhibiting moisture ingress. Roadside vegetation control, clearing of drains and culverts, crack sealing and patching potholes — activities that would be classified as routine maintenance— also protect pavements from moisture. These 'preventative maintenance' activities are low cost compared with rehabilitations and the road agency has only to spend enough to ensure the pavement is protected. For example, resealing may be warranted just before cracks start to appear (every five to fifteen years) but not more often. The optimal amount of preventative maintenance, therefore, in absence of a budget constraint or with a relatively loose budget constraint may be the full amount necessary to protect the pavement without regard to any marginal condition. To explain this point, the present value of preventative maintenance treatments, PVP, has been added to the deterioration function in our numerical model. It is assumed that with $PVP = \$0.25$ million, the pavement is fully protected, and the structural number is 5.0. Higher values of PVP serve only to increase agency costs with no impact on structural number. Below $\$0.25$ million, the structural number reduces linearly to reach 4.0 when $PVP = 0$, at which point the deterioration rate is quite high.

The effect on user costs is illustrated in figure 8, which is a three-dimensional plot of the present value of user costs (PVU) against the present value of rehabilitations (PVM) and PVP. PVU falls as more is spent on either PVM or PVP, except that once PVP reaches the saturation level, $\$0.25$ million, PVU ceases to change as PVP increases. Figure 9 features a contour plot of the surface in figure 8. The iso-PVU curves are vertical for PVP values above $\$0.25$ million, and slope downward where PVP is below $\$0.25$ million. In the latter region, expenditure on rehabilitations and preventative maintenance are substitutable in their effect on user costs.

The downward-sloping straight lines are iso-PVA curves where PVA is the present value of road agency costs defined as $PVA = PVM + PVP$. Since both axes are on the same scale, the iso-PVA lines are at 45 degrees to the axes. For any PVA budget, the optimal combination of rehabilitation and preservation maintenance is found where the iso-PVA line touches the lowest possible iso-PVU contour. The bent line, OAB, is the locus of optimal points. The unconstrained optimum occurs at point B, $\$0.7$ million for rehabilitation, as found previously, plus $\$0.25$ million for preventative maintenance.

Figure 8. Present value of user costs plotted against present values of rehabilitation and preventative maintenance costs

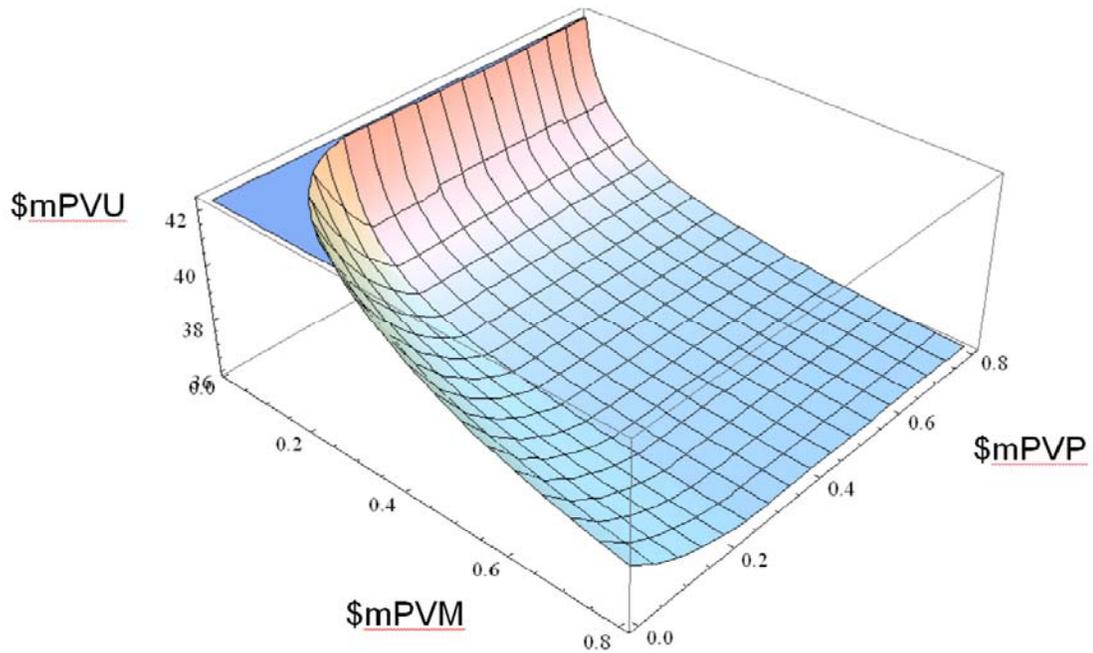
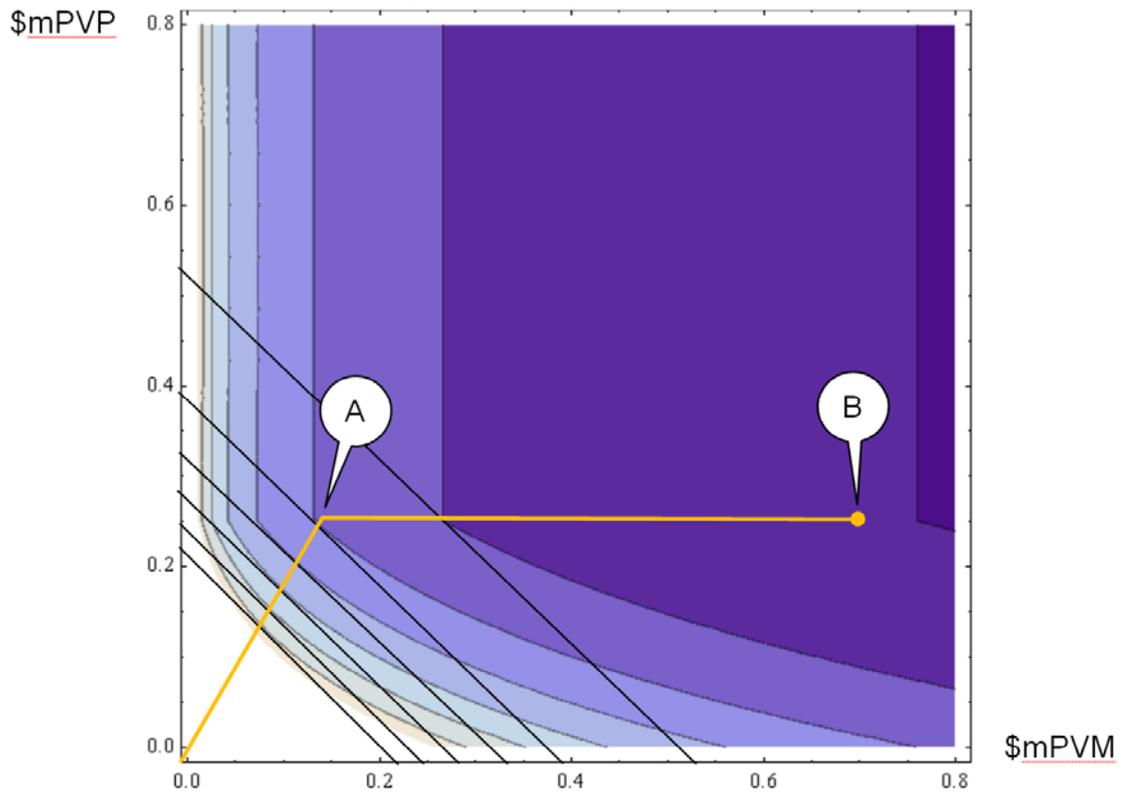


Figure 9. Optimising rehabilitation and preventative maintenance subject to budget constraints



Once a budget constraint is imposed on PVA and is progressively tightened, the optimum point moves leftward along the line AB. The time interval between rehabilitations is increased without any reduction in spending on preventative maintenance. Only when the budget constraint falls below \$0.4 million does it become cost-effective to reduce preventative maintenance. Along the line segment OA, the rule is to balance the two types of maintenance expenditure so that the marginal dollar spent on preventative maintenance and on rehabilitation have the same impact on user costs. If we set up an optimisation problem to minimise $PVU = PVU(PVM, PVP)$ subject to $PVA = PVM + PVP$, the solution would occur where $\partial PVU / \partial PVM = \partial PVU / \partial PVP$.

The model demonstrates that, as the present-value budget constraint is progressively tightened, there is an optimal contraction path in terms of the selection and timing of treatments. Spending on preventative maintenance is not sacrosanct, but it is only optimal to trade off preventative maintenance against rehabilitations for relatively tight present-value budget constraints. Tight short-term annual budget constraints can force the system off the optimal contraction path. For example, say the budget was a present value of \$0.5 million. The optimum lies on the segment AB along which rehabilitations are reduced but full preventative maintenance is retained. A short-term reduction in preventative maintenance for several years in exchange for higher rehabilitation costs later, keeping the present value of agency costs the same, would move the system along the \$0.5 million iso-PVA line to a point below OAB at which user costs are higher.

7.3. Cost of maintenance deferral

Naudé et al. (2008) and (2012) used a case study to demonstrate that savings in the present values of road agency and road user costs can be made by bringing forward maintenance interventions in a situation of tight annual budget constraints. Under the 'Accelerated Road Rehabilitation Program', the Queensland State Government's Treasury provides loans to the road agency, which the road agency repays with interest out of future budget allocations. The case study shows that, in the long term, governments may be better off by borrowing to maintain roads compared with the alternative of deferring maintenance.

Deferring maintenance can be seen as a form of borrowing. Funds are saved in the short-term at the expense of higher outlays in the future. A way to communicate this to decision makers and facilitate comparison with the alternative of borrowing to fund maintenance would be to estimate an 'equivalent interest rate for deferred maintenance' (EIRDM). It is assumed that funds are constrained to the point where only minimum acceptable pavement conditions can be provided. To use higher standards for to estimate the EIRDM might be misleading because, if roads are allowed to deteriorate below the higher minimum standards and never restored to those standards, the equivalent loan is never repaid. The minimum standards could be set using judgement with the aim of ensuring each road is just able to fulfil its economic and social purposes. An alternative, analytical approach would be to use an optimisation model to determine the maintenance program consistent with a given $MBCR > 1$ in the manner suggested above by weighting user costs without any budget constraint. A number of model runs could be undertaken to find the particular MBCR value that balances the needs of users against available funds.

Table 1 Illustrative alternative maintenance and financing strategies

(\$ million road agency cash outflows)

	Strategy 1	Strategy 2	Strategy 3		Strategy 4	
Maintenance strategy	minimise present value of costs with minimum standard constraints	minimise present value of costs with minimum standard and early-year budget constraints	same as for strategy 1		same as for strategy 1	
Finance strategy	no borrowing	no borrowing	borrow at 5%		borrow at 17.2%	
Year	Cash outflow	Cash outflow	Loan	Net cash outflow	Loan	Net cash outflow
	$f(t)$	$g(t)$		$f(t)+loan \leq g(t)$	$f(t)-g(t)$	$f(t)+loan = g(t)$
1	100	50	-50	50	-50	50
2	100	50	-50	50	-50	50
3	100	50	-50	50	-50	50
4	100	75	-25	75	-25	75
5	100	100	0	100	0	100
6	100	200	56	156	100	200
7	100	200	56	156	100	200
8	100	200	56	156	100	200
9	100	200	56	156	100	200
10	100	100	0	100	0	100
Present value (5% discount rate)	772	893	0	772	121	893
Internal rate of return	na	na	5.0%	na	17.2%	na

Table 1 provides a hypothetical numerical example to illustrate the EIRDM concept. Cash outflows are shown for four strategies for maintenance and financing over a 10 year period. Strategy 1, which requires \$100 million to be spent on maintenance in each year, minimises the present value of road agency costs to provide the predetermined minimum standards. Since the objective is to minimise costs to the government rather than to maximise economic efficiency, the interest rate at which the government can borrow should be used as the discount rate in the optimisation model rather than the social discount rate. Assuming the government's borrowing rate is 5%, the present value of agency costs over the 10 year period is \$772 million.

Under strategy 2, a budget constraint of \$50 million per annum is imposed at the start of period and held for long as possible without allowing standards to fall below the set minimums. The \$50 million cap on spending can be continued for only three years. After three years, annual spending must be increased above \$50 million to keep standards from falling below the minimum levels. During the first half of the ten-year period, sub-surface layers of pavements have deteriorated to the point where considerably higher annual amounts than under strategy 1 need to be spent during years 6 to 9 to restore pavements to the minimum standards. By year 10, the necessary restorative work has been completed and annual the expenditure requirement returns to \$100 million. The present value of agency costs is \$893 million, which is \$121 million above the cost under strategy 1.

Under strategy 3, the road agency implements the cost minimising maintenance strategy and borrows sufficient funds to make net cash outflows for the first five years the same as for the first five years of strategy 2. The road agency repays the loan in the second half of period in the way that ensures net cash outflows in all years are less than or equal to those under strategy 2. The present value of net cash outflows is identical to that for strategy 1. In this example, it is more cost effective to implement the cost minimising maintenance strategy and borrow to meet short-term budget constraints than to defer maintenance in the short term and catch up later. Besides saving costs, the repayment schedule can be negotiated to suit fiscal circumstances rather than be determined by the times at which pavements deteriorate below acceptable standards.

Strategy 4 illustrates the EIRDM concept. The road agency implements the cost minimising maintenance strategy but borrows funds and repays them in a way that makes annual net cash outflows in all years identical to those under strategy 2. The streams of cash outflows as functions of time under strategies 1 and 2 are referred to as $f(t)$ and $g(t)$ respectively. The stream of annual cash outflows from the loan is $f(t) - g(t)$. The interest rate for this loan is found by taking the internal rate of return (the interest rate that makes the net present value zero) for the cash outflow stream $f(t) - g(t)$. In the numerical example in Table 1, the EIRDM is 17.2%, well above the rate at which the government can borrow.

Steps to estimate EIRDM are as follows.

1. Determine minimum acceptable standards at which the roads are just adequate to meet their economic and social purposes.
2. Use an optimisation model to find the minimum present value of road agency costs consistent with providing these minimum standards in the absence of any budget constraint. The discount rate should be the government's borrowing rate. The result will be stream of annual maintenance costs, $f(t)$, and a residual value.

3. Use an optimisation model a second time to find the minimum present value of road agency costs subject to annual budget constraints for the years during which they are imposed while meeting minimum standards in all years. The analysis period should be long enough to encompass the time when additional spending is needed restore the network to minimum acceptable standards following the tightly constrained years. It may be necessary to adjust both the levels and durations of the short-term annual budget constraints if they are too tight to keep the roads above the minimum standards. The result will be a second stream of annual maintenance costs, $g(t)$, and a residual value.
4. Take the difference between the two streams of maintenance costs for each year, $f(t) - g(t)$, to obtain the annual cash outflows for the hypothetical loan that makes net cash flows from following the cost minimising maintenance strategy together with taking out the loan identical to net cash outflows under the deferred maintenance strategy.
5. Find the interest rate for the hypothetical loan by estimating the internal rate of return.

7.4 Marginal benefit–cost ratio with tight short-term budget constraints

When tight short-term budget constraints lead to maintenance deferral, an increase in annual maintenance budgets during the early years can cause the present value of agency costs to fall. The MBCR defined as $-\Delta PVU/\Delta PVA$, based on equation (9), is negative because ΔPVA is negative. A more general definition of MBCR is required.

$$MBCR = \frac{-\Delta PVTTTC + \Delta PVB}{\Delta PVB} = \frac{-\Delta PVU - \Delta PVA + \Delta PVB}{\Delta PVB} \quad (15)$$

where

- PVTTTC = present value of total transport costs
- PVU = present value of user costs
- PVA = present value of agency costs
- PVB = present value of budget constraints.

'Benefit' in the numerator of the standard BCR measure is a *gross* amount in the sense that the cost in the denominator has *not* been deducted. Negative ΔPVA is a *net* amount because ΔPVB has been deducted from it. For example, assume that increasing the maintenance budget in the constrained years by a present value of \$1 million causes no change in user costs ($\Delta PVU = 0$) but leads to PVA falling by \$4 million. The net gain to society is \$4 million = $-\Delta PVA = -\Delta PVTTTC$. The gross benefit after adding back the additional funds spent is \$5 million = $-\Delta PVA + \Delta PVB$. The MBCR = $[-\$0 - (-\$4m) + \$1m] / \$1m = \$5m/\$1m = 5$. The economic value of spending an additional \$1 million on maintenance would be the same as investing it in a capital project with a BCR of 5. Both yield a net gain of \$4 million after deducting the \$1 million spent.

Where the budget constraint is expressed as a present value, $\Delta PVA = \Delta PVB$ and equation (15) reduces to $MBCR = -\Delta PVU/\Delta PVA$.

8. OPTIMAL INCENTIVES IN MAINTENANCE CONTRACTS

The model of maintenance optimisation developed here fits neatly into the author's incentive regulation framework for commercial road supply. The scheme in Harvey (2013) applies to complete supply of road services by a public utility or private firm. We show here how it can be applied at a less ambitious level to outsourcing of maintenance activities to a contractor. This is often done with performance-based contracts in which maintenance standards are expressed in physical terms and financial penalties are imposed for failing to meet those standards. In Harvey's scheme, the sole performance measure is generalised user costs and the penalties for under-performance and bonuses for over-performance exactly equal their marginal social values. The model and parameters for estimating user costs as a function of road condition would form part of the maintenance contract. Profit maximising behaviour for the supplier then produces the welfare maximising outcome.

For simplicity, we initially assume an infinite time horizon. It is also assumed that the social and private discount rates are identical. Adapting Harvey's shadow toll formula, the maintenance contractor is paid in year t of the maintenance cycle, an annual sum of $p - [c(t) - c(0)]$ where p is a constant and $c(0)$ is user costs with a new pavement. The annual payment would be highest just after the pavement had been rehabilitated, at which time $c(t) - c(0) = 0$. Payments would reduce as the pavement became rougher causing $c(t)$ to rise above $c(0)$. Thus annual maintenance payments follow a sawtooth pattern over time, the mirror image of the pattern for user costs. The present value of the contractor's profits is:

$$PV\Pi = \frac{p}{r} - \left[PVU - \frac{c(0)}{r} \right] - PVM = \frac{[p - c(0)]}{r} - PVTTC \quad (16)$$

To maximise profits, the contractor would schedule maintenance treatments to set $dPV\Pi/dT = 0$. Since $[p - c(0)]/r$ is constant with respect to maintenance expenditure, the contractor will aim to set $dPVTTC/dT = 0$, which is the same as the condition for minimising social costs. The competitive tender process would ensure that p is bid down to the point where $[p - c(0)]/r - PVTTC = 0$ with r including a normal return on investment. The value of $c(0)$ is arbitrary. A different value would be exactly offset by a changed value of p . The value of p would be affected by pavement age at the start of the contract, being higher for an old pavement because future rehabilitations occur sooner.

Harvey's shadow toll formula includes a 'correction factor', ψ that modifies the incentive faced by the supplier in situations when the regulatory authority wishes to deliberately engender over- or under-investment or maintenance. In the case of a maintenance contractor, the annual payment becomes, $p - \psi[c(t) - c(0)]$ and the contractor's profit function

$$PV\Pi = \frac{p}{r} - \psi \left[PVU - \frac{c(0)}{r} \right] - PVM = \frac{[p - \psi c(0)]}{r} - \psi PVU - PVM \quad (17)$$

Profit maximisation requires the contractor to set

$$\frac{dPV\Pi}{dT} = \psi \frac{dPVU}{dT} - \frac{dPVM}{dT} = 0, \text{ which implies } \psi = \frac{dPVU}{dPVM} = \frac{1}{MBCR}$$

If the road agency wanted the road maintained at a below-optimal standard for budgetary reasons, ($MBCR > 1$), it would set ψ below one at the reciprocal of the MBCR that achieves the budget constraint. Reducing the financial reward for better maintenance, leads to a lower standard. The tender process would ensure that the value of p was lower than otherwise so the contractor's costs were exactly covered. Conversely, if the aim was to maintain the road at an above-optimal standard to meet a community service obligation, the road agency would set ψ above one.

If the contract is to terminate at a given future date, the contractor faces an incentive to defer the last treatment. The incentive can be removed by including in the contract an end-of-term adjustment whereby one party pays the other the difference between the actual residual value and an agreed residual value. For example, the contractor could pay the agency the value for a new pavement minus the actual residual value. The contractor is then penalised for returning a pavement to the agency in less-than-new condition. Alternatively the agency could pay the contractor the value of a pavement at a specified poor condition level, in which case, the contractor is rewarded for returning the pavement in good condition. The latter arrangement is better for the contractor but this advantage would be offset by a lower annual payment p compared with the former arrangement.

With the same contractor maintaining a large number of road segments at different stages of their life cycles, the fluctuations in payments that vary inversely with user costs would be smoothed out and the total residual value adjustment averaged out. Also, a considerable amount of the risk associated with the stochastic component of pavement deterioration would be diversified away.

9. IMPLICATIONS FOR ROAD FUNDING

Road maintenance optimisation modelling is usually undertaken to assist road agencies plan and budget for their maintenance activities. It can promote best use of available funds by helping determine the types and timings of treatments to be carried out. At a higher level, it can help determine appropriate levels of maintenance funding.

This paper has identified two distinct types of underfunding compared with the unconstrained optimum. The first concerns the present value of funds made available for maintenance, which can be measured by the 'maintenance deficit'. There is a variety of ways to define and measure the maintenance deficit, and different standards against which to compare actual funding levels — the unconstrained optimum, the level of spending required to achieve the same MBCR as for investment spending, and the level of spending required to attain predetermined minimum acceptable standards.

Economically efficient allocation of funds between maintenance and investment spending requires the MBCR for maintenance and the MBCR or cut-off BCR for investment spending to be the same. If they are not, the return to society from road expenditure can be improved by shifting funds from the budget with the lower MBCR to the budget with the higher MBCR. There is a strong case for efficient allocation of road expenditure between maintenance and investment.

The second type of underfunding identified in the paper occurs when maintenance deferral increases the present value of maintenance costs without any improvement in the user cost outcome. It occurs when tight short-term annual budget constraints force a departure from the most efficient spending profile over time. The most efficient profile can be estimated by an optimisation model set up to minimise total transport costs subject to a budget constraint expressed as a present value, or to minimise the present value of road agency costs subject to minimum standards constraints. If preventative maintenance that protects and preserves road pavements is deferred, components of the assets can be damaged, and the future cost of restoration can exceed the short-term saving, even after adjusting for the time value of money. Departing from the cost-minimising time path to achieve the lowest acceptable standards by deferring maintenance is equivalent to going into debt. This paper has proposed a way to use maintenance optimisation models to estimate an 'equivalent interest rate for maintenance deferral', which enables comparisons to be made between the cost of borrowing through maintenance deferral and conventional borrowing.

Appendix:
Road maintenance optimisation literature

	Description	Modelling platform	Road Deterioration Modelling			
			Deterministic		Probabilistic (Markov)	
			Single objective	Multi-objective	Single objective	Multi-objective
Prioritization (ranking)	Priority ranking according to certain criteria, including economic or engineering rules.		Hajek and Phang (1989) Fwa and Chan (1993) Sebaaly et al. (1996) Hawker and Abell (2000) Wong et al. (2003) Kulkarni et al. 2004; Bekheet et al. (2008)			
Enumeration	All possible solutions compared. Limited by the size of the solution space.	HDM-4	Morosiuk et al. (2006)		?	
Linear programming	Linear objective function and constraints with continuous decision variables.	CONNPAVE (Davis and Van Dine 1988)	Grivas et al. (1993) Männistö et al. (2001) de la Garza et al. (2011)		Davis and Van Dine (1988) Bekheet et al. (2008) Chou and Wang (2012)	
Non-linear programming	Non-linear objective function and constraints with continuous decision variables.				Abaza and Ashur (1999) Abaza et al. (2004) Abaza (2006)	Wu and Flintsch (2009)
Convex programming <i>Steepest descent method</i> <i>Conjugate gradient method</i>	Special case of the constrained nonlinear optimisation - concave objective function with convex constraints.	In conjunction with or within HDM-4	Tsunokawa and Ul-Islam (2003) Hiep and Tsunokawa (2005) Tsunokawa et al. (2006) Archondo-Callao (2008)			

	Description	Modelling Platform	Road Deterioration Modelling			
			Deterministic		Probabilistic (Markov)	
			Single objective	Multi-objective	Single objective	Multi-objective
Integer programming	Linear or non-linear objective function with constraints. Decision variables take integer values.		Fwa et al. (1988) Sirajuddin (1997) Ng et al. (2011)		Li et al. (1998)	
Dynamic programming	A method for dealing with large and complex optimisation problems by solving a sequence of smaller problems. There is no general-purpose computational procedure.		Abelson and Flowerdew (1975) Takeyama and Hoque (1995) Smadi (2001) Scheinberg and Anastasopoulos (2010) Yoo et al. (2008)		Feighan et al. (1987) Butt et al. (1994) Li et al. (1996)	
Genetic algorithm	An approach to finding 'good' approximate solutions with the benefit of savings in computing times. Based on natural selection and genetics. Suitable for solving combinatorial optimisation problems.	GENEPAV/HDM-4 (Ferreira and Queiroz 2012) GENETIPAV-D (Ferreira et al. 2002b) MODAT (Ferreira and Meneses 2011) NOS (Wang et al. 2007) PAVENET (Fwa et al. 1994) PAVENET-R (Chikezie et al. 2011) PAVMAN (Panagopoulou and Chassikos 2012) PLATO (Roper 2004)	Fwa et al. (1994) Chan et al. (1994) Fwa et al. (1996) Pilson et al. (1999) Zhang et al. (2001) Ferreira et al. (2002b) Roper (2004) Herabat and Tangphaisankun (2005) Chootinan et al. (2006) Maji and Jha (2007) Bosurgi and Trifiro (2007) Golroo and Tighe (2012) Ferreira and Queiroz (2012) Panagopoulou and Chassikos (2012)	Pilson et al. (1999) Fwa et al. (2000) Wang et al. (2003) Herabat and Tangphaisankun (2005) Ferreira and Meneses (2011) Golroo and Tighe (2012)	Ferreira et al. (2002a) Morcouc and Lounis (2005)	Di Mino et al. (2007) Wang et al. (2007) Wu et al. (2010)

BIBLIOGRAPHY

- Abaza K. 2006, 'Iterative linear approach for nonlinear non-homogenous stochastic pavement management models', *Journal of Transportation Engineering*, 132(3): 244-256.
- Abaza K. and Ashur S. A. 1999, 'Optimum decision policy for management of pavement maintenance and rehabilitation', *Transportation Research Record: Journal of the Transportation Research Board*, 1655(1): 8-15.
- Abaza K., Ashur S., Abu-Eisheh S. and Al-Khatib I. 2004, 'Integrated pavement management system with a Markovian prediction model', *Journal of Transportation Engineering*, 130(1): 24-33.
- Abelson P.W. and Flowerdew A.D.J. 1975, 'Models for the economic evaluation of road Maintenance', *Journals of Transport Economics and Policy*, 9(2): 93-114.
- Archondo-Callao R. 2008, 'Applying the HDM-4 model to strategic planning of road works', *Transport Papers (TP-20)*, The World Bank Group, Washington, D. C.
- Australian Transport Council (ATC) 2006, National Guidelines for Transport System Management in Australia, Volume 5, Background material.
- Austrroads 2002, Guidelines for community input in setting level of service and intervention standards for road networks, AP-R201-02, Austrroads, Sydney.
- Austrroads 2007, Guide to asset management, Part 5B: Roughness, Austrroads, Sydney.
- Austrroads 2008a, Guide to asset management, Part 5D: Strength, Austrroads, Sydney.
- Austrroads 2008b, Road Surface Characteristics and Crash Occurrence: A Literature Review, AP-T96/08, Austrroads, Sydney.
- Austrroads 2009, Guide to asset management, Part 2: Community and stakeholder requirements, Austrroads, Sydney.
- Bekheet W., Helali K., Li Y., Cheetham A., and Stanciu D. (2008), 'Comparison between probabilistic and deterministic pavement management analysis: a case study for Arisona DOT', *7th International Conference on Managing Pavement Assets*.
- Bennett C.R. and Greenwood I.D. 2004, *Modelling road user and environmental effects in HDM-4*, Highway development and management series, Volume 7, World Road Association (PIARC).
- Bosurgi G. and Trifiro F. 2007, 'Pavement design and maintenance via genetic algorithms', *4th International SIIV Congress*, Palermo, Italy.

- Butt A.A., Shalin M.Y., Carpenters S.H. and Carnahan J.V. 1994, 'Application of Markov process to pavement management system at network level', in *Proceedings for Third International Conference on Managing Pavements*, 2: 159-172.
- Burningham S. and Stankevich N. 2005, *Why road maintenance is important and how to get it done*, Transport note no. TRN-4, World Bank, Washington DC.
- Chan W.T., Fwa T.F. and Hoque K.Z. 2001, 'Constraint handling methods in pavement maintenance programming', *Transport Research Part C: Emerging Technology*, 9: 175-190.
- Chan W.T., Fwa T.F. and Tan Y. 1994, 'Road-Maintenance Planning Using Genetic Algorithms. I: Formulation', *Journal of Transportation Engineering*, 120(5): 693-709.
- Chikezie C.U., Olowosulu A.T. and Abejide O.S. 2011, 'Review of application of genetic algorithms in optimization of flexible pavement maintenance and rehabilitation in Nigeria', *World Journal of Engineering and Pure and Applied Science*, 1(3)68.
- Chootinan P., Chen A., Horrocks M.R. and Bolling D. 2006, 'A multi-year pavement maintenance program using a stochastic simulation-based genetic algorithm approach', *Transportation Research Part A: Policy and Practice*, 40(9): 725-743.
- Chou E. Y. and Wang S. 2012, *Benefit cost models to support pavement management decisions*, report prepared for Office of Research and Development, Ohio Department of Transportation and Federal Highway Administration, U. S. Department of transportation, State Job Number 134363, June.
- Davis C. F. and Van Dine C. P. 1988, 'Linear programming model for pavement management', *Transportation Research Record: Journal of the Transportation Research Board*, 1200: 71-5.
- de la Garza J.M., Akyildiz S., Bish D.R. and Krueger D. A. 2011, 'Network-level optimization of pavement maintenance renewal strategies', *Advanced Engineering Informatics*, 25(4): 699-712.
- Decker R. 1996, 'Applications of maintenance optimization models: a review and analysis', *Reliability Engineering and System Safety*, 51(3): 229-240.
- Di Mino G., Di Liberto C.M., Di Noto F. and Nigrelli J. 2007, 'A multi-objective approach based on the genetic algorithm technology for road pavement maintenance', *4th International SIIV Congress*, Palermo (Italy), 12-14 September.
- Feighan K.J., Shahin, M.Y. and Sinha K.C. 1987, 'A dynamic programming approach to optimization for pavement management systems', *2nd North American Pavement Management Conference*.
- Ferreira A. L. and Queiroz C. 2012, 'A deterministic section-linked optimization model for road network maintenance management', *3rd International Conference on Engineering Optimization*, Rio de Janeiro, Brazil 1-5 July.

- Ferreira A. L. and Meneses S. C. N. 2011, 'A multi-objective decision-aid tool for road network maintenance management', *5th International Conference on Bituminous Mixture and Pavements*, Thessaloniki, 1–3 June.
- Ferreira A., Antunes A. and Picado-Santos L. 2002(a), 'Probabilistic segment-linked pavement management optimisation model', *Journal of Transportation Engineering*, 128: 568-577.
- Ferreira A., Picado-Santos L. and Antunes A. 2002(b), 'A segment-linked optimisation model for deterministic pavement management system', *International Journal of Pavement Engineering*, 3(2): 95-105.
- Fwa F.T., Tan C.Y. and Chan W.T. 1994, 'Road maintenance planning using genetic algorithms II: analysis', *Journal of Transportation Engineering*, 120: 710-722.
- Fwa F.T., Sinha K.C. and Riverson J.D.N. 1988, 'Highway routine maintenance programming at network level', *Journal of Transportation Engineering*, 114(5): 539-554.
- Fwa F.T., Chan W.T. and Tan C.Y. 1996, 'Genetic-algorithm programming of road maintenance and rehabilitation', *Journal of Transportation Engineering*, 122: 246-253.
- Fwa T.F., Chan W.T. and Hoque K.Z. 2000, 'Multi-objective optimization for pavement maintenance programming', *Journal of Transportation Engineering*, 126(5): 367-374.
- Goldberg E.D. 1989, *Genetic algorithms in search, optimisation and machine learning*, Addison-Wesley Publishing Company, Menlo Park, California.
- Golroo A. and Tighe S.L. 2012, 'Optimum genetic algorithm structure selection in pavement management', *Asian Journal of Applied Science*, 5(6): 327-341.
- Grivas D.A., Ravirala V. and Schultz B.C. 1993, 'State increment optimization methodology for network-level pavement management', *Transportation Research Record: Journal of the Transportation Research Board*, 1397: 25-33.
- Harvey M.O. 2013, 'Commercial road supply with incentive regulation', unpublished manuscript under review. Available from author on request.
- Hawker L. and Abell R. 2000, 'Selection and prioritisation of maintenance works on major roads in England', *Proceedings of the 1st European Pavement Management Systems Conference*, Budapest, Hungary, CD Ed., 1-8.
- Hajek J. J. and Phang W.A. 1989, 'Prioritization and optimization pavement of preservation treatment', *Transportation Research Record*, 1216: 58-68, Transportation Research Board, Washington D. C.
- Herabat P. and Tangphaisankun A. 2005, 'Multi-objective optimization model using constraint-based generic algorithms for Thailand pavement management', *Journal of the Eastern Asia Society for Transportation Studies*, 6: 1137-1152.

- Hiep D. and Tsunokawa K. 2005, 'Optimal maintenance strategies for bituminous pavements: a case study in Vietnam using HDM-4 with gradient methods', *Journal of the Eastern Asia Society for Transportation Studies*, 6: 1123-1136.
- Holland J.H. 1975, *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, Michigan.
- Jawad D. and Ozbay, K. 2006, 'Probabilistic lifecycle cost optimization for pavement management at the project-level', *Transportation Research Board, 2006 Annual Meeting CD-ROM*, National Academy of Science, Washington, D.C.
- Konak A., Coit D.W. and Smith A.E. 2005, 'Multi-objective optimization using genetic algorithms: a tutorial', *Reliability Engineering and System Safety*, 91: 992-1007.
- Kulkarni R., Miller D., Ingram R., Wong C. and Lorenz J. 2004, 'Need-based project prioritization: alternative to cost-benefit analysis', *Journal of Transportation Engineering*, 130(2): 150-158.
- Li N., Haas R. and Xie W.C. 1997, 'Investigation of relationship between deterministic and probabilistic prediction models in pavement management', *Transportation Research Record: Journal of the Transportation Research Board*, 1592: 70-79.
- Li N., Xie W.C. and Haas R. 1996, 'Reliability-based processing of Markov chains for modelling pavement network deterioration', *Transportation Research Record: Journal of the Transportation Research Board*, 1524: 203-213.
- Li N., Xie W.C., Haas R. and Huot M. 1998 'Integer programming of maintenance and rehabilitation treatments for pavement networks', *Transportation Research Record: Journal of the Transportation Research Board*, 1629: 242-8.
- Li Y. and Madanat S. 2002, 'A steady-state solution for the optimal pavement resurfacing problem', *Transportation Research Part A: Policy and Practice*, 36(6): 525-535.
- Maji A. and Jha M.K. 2007, 'Modelling highway infrastructure maintenance schedules with budget constraints', *Transportation Research Record: Journal of the Transportation Research Board*, 1991: 19-26.
- Männistö V., Virtala P., Söderqvist M.K. and Kerali H.R. 2001, 'Network level analysis of pavements and bridges in Europe', *Proceedings of 5th International Conference on Managing Pavement*, 11-14 August, 2001, Seattle, Washington, USA.
- McLean J. and Foley G. 1998, *Road surface characteristics and condition: effects on road users*, ARRB Transport Research Ltd, Research report ARR314, Vermont South, Victoria.
- Morcous G. and Z. Lounis 2005, 'Maintenance optimisation of infrastructure networks using genetic algorithm', *Automation in Construction*, 14: 129-142.
- Morosiuk G., Riley, M and Odoki J.B. 2004, Modelling road deterioration and works effects, Highway development and management series, Volume 6, World Road Association (PIARC).

- Morosiuk G., Riley M. and Toole T. 2006, *Application Guide for HDM-4 Version 2.0*, Highway Development and Management Series, 2.
- Naudé C. Tsolakis D. and Toole T. 2008, 'An "accelerated" approach to funding road rehabilitation, 23rd ARRB Conference, Adelaide, Australia.
- Naudé C. and Toole T. 2012, 'Evaluation of the accelerated road rehabilitation program (ARRP)', 25th ARRB Conference, Perth, Australia.
- Ng M., Zhang Z. and Waller S.T. 2011, 'The price of uncertainty in pavement infrastructure management planning: An integer programming approach', *Transportation Research Part C: Emerging Technologies*, 19(6): 1326-1338.
- Newbery D.M. 1989, 'Road damage externalities and road user charges', *Econometrica*, 56(2): 295-316.
- Odoki J.B. and Kerali H.G.R. 2006, Analytical framework and model descriptions, Highway development and management series, Volume 4, World Road Association (PIARC).
- Opus 1999, *Review of VOC-pavement roughness relationships contained in Transfund's project evaluation manual*, Opus International Consultants Ltd, Central Laboratories Report 529277.00, Lower Hutt, New Zealand.
- Ortiz-Garcia J.J., Costello S.B. and Snaith M.S. 2006, 'Derivation of transition probability matrices for pavement deterioration modelling', *Journal of Transportation Engineering*, 132(2): 141-161.
- Panagopoupou M.I. and Chassiakos A.P. 2012, 'Optimization model for pavement maintenance planning and resource allocation', *TRB Transportation Research Circular E-C136*.
- Paterson W.D. 1987, *The Highway Design and Maintenance Standard Series - Road Deterioration and Maintenance Effects*, World Bank, John Hopkins University Press, Baltimore.
- Paterson W.D. and Attoh-Okine B. 1992, 'Summary models of paved road deterioration based on HDM-III', *Transportation Research Record* 1344: 99-105, Transportation Research Board, Washington, D.C.
- Pilson C., Hudson W.R. and Anderson V. 1999, 'Multi-objective optimization in pavement management by using genetic algorithms and efficient surfaces', *Transportation Research Record: Journal of the Transportation Research Board*, 1655: 42-48.
- Roper R.L. 2004, 'Using genetic algorithms to select optimum pavement treatment intervention strategies', 6th International Conference on Managing Pavements.
- Scheinberg T. and Anastasopoulos, P.C. 2010, 'Pavement preservation programming: a multi-year multi-constraint optimization methodology', paper prepared for presentation at the 89th Annual Meeting of the Transportation Research Board and publication in the *Transportation Research Record*.

- Sebaaly P.E., Hand A., Epps J. and Bosch C. 1996, 'Nevada's approach to pavement management', *Transportation Research Record* 1524: 109-117, Transportation Research Board, Washington, D.C.
- Semmens J. 2006, 'De-socializing the roads', in: Roth G. (ed.), *Street smart*, Transaction Publishers, New Brunswick.
- Sirajuddin A.M.Y. 1997, 'Highway maintenance fund allocation: tabulated manual procedure', *Journal of Transportation Engineering*, 123(5): 346-9.
- Smadi O. 2001, 'Knowledge based expert system pavement management optimisation', *Proceedings of 5th International Conference on Managing Pavements*.
- Small K.A and Winston C. 1988, 'Optimal highway durability', *American Economic Review*, 78(3): 560-569.
- Takeyama Y. and Hoque K.Z. 1995, 'Optimal maintenance decision for managing flexible pavement in Thailand', *Proceedings for Second International Conference on Road and Airfield Pavement Technology*, 1: 283-291.
- Tsunokawa K. Van Hiep D. and Ul-Islam R. 2006, 'True Optimization of Pavement Maintenance Options with What-If Models', *Computer-Aided Civil and Infrastructure Engineering*, 21: 193-204.
- Tsunokawa K. and Ul-Islam R. 2003, 'Optimal pavement design and maintenance strategy for developing countries: an analysis using HDM-4', *International Journal of Pavement Engineering*, 4(4):193-208.
- Wang F., Zhang Z. and Machemehl R. 2003, 'Decision making problem for managing pavement maintenance and rehabilitation projects', *Proceedings of the 82nd Transportation Research Board Annual Meeting (CD-ROM)*, TRB, Washington, D. C.
- Wang K.C.P., Nguyen V. and Zaniewski J.P. 2007, 'Genetic algorithm-based network optimization system with multiple objectives', *Transportation Research Record: Journal of the Transportation Research Board*, 2016: 85-96.
- Wong W., He G. and Luk S. 2003, 'Development of road management systems in China', *Proceedings of the Institution of Civil Engineers-Transport*, 156(4): 179-188.
- Wu Z. and Flintsch G. 2009, 'Pavement preservation optimization considering multiple objectives and budget variability', *Journal of Transportation Engineering*, 135(5): 305-315.
- Wu Z., Flintsch G., Ferreira A. and Picado-Santos L. 2010, 'Multi-criteria optimization applications in highway asset management', *Journal of Transportation Engineering*, Accepted for publication.
- Yoo J. and Garcia-Diaz A. 2008, 'Cost-effective selection and multi-period scheduling of pavement maintenance and rehabilitation strategies', *Engineering Optimization*, 40(3): 205-222.

Zhang X.J., Fwa F.T. and Chan W.T. 2001, 'Optimal priority setting for programming of pavement maintenance activities', in *Proceedings for 6th PES-CTR Symposium on Pavement Technology*.

Zietlow G.J. 2006, 'Role of the private sector in managing and maintaining roads' in Roth G. (ed.), *Street Smart*, Transaction Publishers, New Brunswick.

International Transport Forum

2 rue André Pascal

75775 Paris Cedex 16

itf.contact@oecd.org

www.internationaltransportforum.org
