Markov Decision Processes for Road Maintenance Optimisation

This paper primarily focuses on finding a policy for maintaining a road segment. The paper presents two methods for finding such a policy. The first one is using a probabilistic Markov Decision Process in order to determine the optimal maintenance policy. The second method is used by the Road and Hydraulic Engineering Division (DWW). That policy however is not optimal but does not require an MDP to be found. The paper will show that despite of finding an optimal policy via the first method this will not always be the best option to use.

Key words: Road Maintenance, Road Deterioration, Markov Decision Processes

Onne van der Weijde
ERASMUS UNIVERSITY ROTTERDAM
18-6-2013

Table of Contents

1. INTRODUCTION

When transporting goods from A to B, roads play an important role. Not only enable maintained roads a relative fast way for goods to be transported but they also allow for relative safe transportation. Inadequate maintained road severely affect the costs and speed of transportation (Mackie, Nellthorp and Laird, 2005).

Maintaining roads can be quite expensive (Rietveld, Bruinsma, & Koetse, 2007) and it is therefore important to find a cost efficient way to make sure roads are kept to a certain minimum quality. When considering the problem as a whole, finding a maintenance policy can be difficult. As a first step we will look in this thesis at a sub-problem considering small segment of road undergoing Road Deterioration (RD). Road inspection is assumed to take place periodically but is not taken into account when constructing the maintenance policy.

The purpose of this paper is to the investigate the use of a Markov Decision Process (MDP) for determining an optimal maintenance policy. The road deterioration process itself shall not be investigated in this paper, instead methods and data concerning road deterioration are used from Plasmeijer, (1999).

The Markov decision process eventually has to produce an optimal maintenance policy answering the following question: *which* maintenance action should be chosen when the road segment has reached a certain age and condition? Of course because of safety regulations the condition of the road must be kept at a reasonable condition. The optimal policy can be found by solving a linear program (Dekker, Nicolai, Kallenberg, 2007).

The results of the optimal policy are compared to another policy found by using the Equivalent Annual Cost Method (EAC) used by the Road and Hydraulic Engineering Division (DWW). This will give an indication of the reduction in expected average costs, using an MDP to find the optimal policy. The methods themselves are also compared to each other, discussing the computation time and easiness of finding the policies.

The paper is structured as follows: first a short literature overview on the subject is given in section 2. Then some aspects concerning road deterioration and maintenance will be discussed in section 3. The fourth section contains the road deterioration processes as is estimated in Plasmeijer, 1999. After introducing the road deterioration processes and the maintenance actions the Markov Decision Process can be solved using linear programming. Then another policy is constructed via the Equivalent Annual Cost method, of which the resulting policy and method are compared to the optimal policy and its process of finding it. Finally a conclusion can be given on the effectiveness of using the Markov decision processes for road maintenance.

2. LITERATURE OVERVIEW

This section surveys the road maintenance problem. There are roughly two approaches to formulate a model for this problem (Golabi, Kulkarni, and Way, 1982). One could use a model that gives least cost maintenance policies under the condition that the road is maintained at minimum standards, or develop a model that gives the best possible road conditions under budget constraints. In either case the road deteriorates over time and needs to be predicted. Both models can be based on formulating the problem as a constrained Markov decision process, and linear programming can be used to find

the optimal solution (Dekker, Nicolai, and Kallenberg, 2007). Below these subject will further be discussed.

2.1 Road Deterioration Process

The first problem when dealing with road maintenance is determining a way to estimate the road deterioration (RD) processes. In the paper of Martin and Kadar, 2012 a description is given how to estimate these RD processes. They mention four probabilistic approaches to RD modelling: survivor curves; Markov and semi-Markov approaches; continuous probability function and other probabilistic approaches. The advantage of probabilistic RD models over deterministic ones is that they can assign various probabilities to the future conditions of a pavement, unlike the deterministic models who provide an average estimate to a future condition which is not likely to be achieved.

2.1.1 Survivor Curves

Most survivor curves are based on historical records, location, condition and maintenance strategies of the road. These curves are easily computed but they do require reliable data in order to give proper predictions.
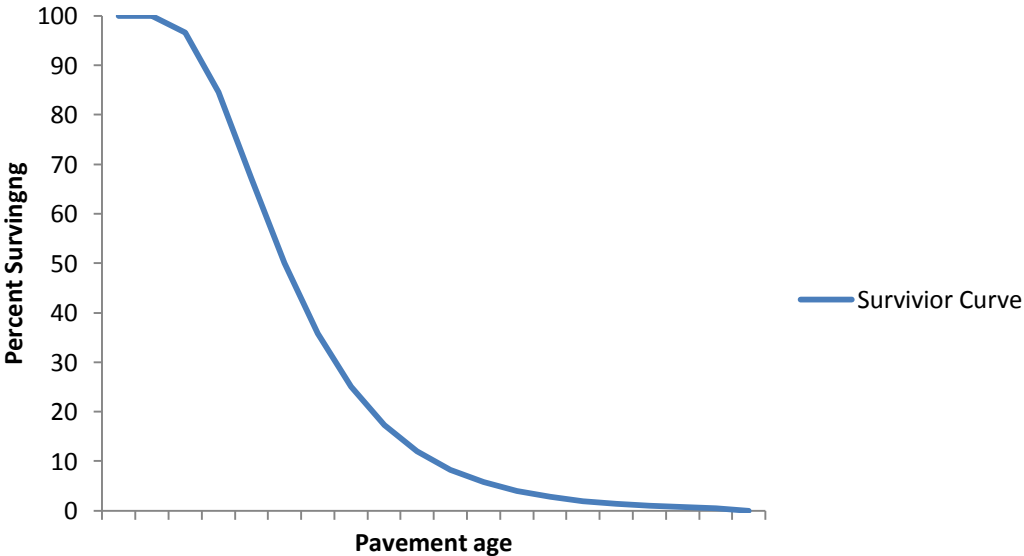
Figure 2.1: Survivor Curve

2.1.2 Markov Probabilistic Approach

The Markov probabilistic approach assumes that the future condition of the road only depends on its present condition. The advantage that this brings is that no prior information about the road condition is needed, especially useful when no historical records are available.

The determination of the probability when changing from state (transition probability) can be done by expert opinion or based on analysis like the earlier mentioned survivor curve, and thus do not directly relate to other variables , such as environment, traffic load, etc.

All the transition probabilities can be stored in a transition probability matrixes (TPM), which are different for different maintenance actions. An example of such a TPM is showed below where state 0 is 'new condition' and the transition probabilities are represented by $p_{ij} \geq 0$.

Table 2.1: Transition probability matrix

Condition state	0	1	2	3
0	p_{00}	p_{01}	p_{02}	p_{03}
1	p_{10}	p_{11}	p_{12}	p_{13}
2	p_{20}	p_{21}	p_{22}	p_{23}
3	p_{30}	p_{31}	p_{32}	p_{33}

Each row of the TPM should add up to 1: $\sum_j p_{ij} = 1$, and when dealing with a pure road deterioration process, so no maintenance is done, the road cannot improve on its own: $p_{ij} = 0$ for all $i > j$.

One could choose for more condition states, be this is not advisable when little or no performance data is available. Via a weighted average of observations of a particular deterioration, d_i, the condition states can be estimated as follows:

$$d_{iavg} = \frac{\sum_{i=1}^{n} L_i \cdot d_i}{\sum_{i=1}^{n} L_i},$$

where i is the current pavement section, n the number of pavement sections, L_i the length of section i, d_i deterioration on section i for given year, and d_{iavg} the weighted average deterioration d_i.

It is best to avoid to use TPM's to predict road condition when the variables are used out of their range of observation. TPM's are constructed by expert groups or survivor curves, so they predict performance without any explanatory power. The Markov probabilistic approach assumes independence of time when changing from road conditions. This can be dealt with using a Semi-Markov Probabilistic Approach. Here time between observations of the system is not fixed but also a random variable, with a distribution which depends on the state and action chosen.

2.1.3 Continuous Probabilistic Approach

The continuous probabilistic approach forecasts the future failure probabilities based on an continuous failure probability, usually derived for Bayesian models constructed with the help of observed data and expert groups using Bayesian regression techniques. They were originally used when only small quantities of poor quality observed data were available.

One could also use Markov chain Mote Carlo simulation to estimate the parameter distributions. This is done by using existing information and information from the performance data. Logit models could also be used to estimate various deterioration processes. The logit model used for estimating the probability of surface crack initiation would have a general form as follows:

$$p(crx) = \frac{1}{1 + e^{a \cdot x + \{(A,B) \; for \; crx = (0,1)\}}},$$

where, $P(crx)$ is the probability of a road segment being cracked, a vector containing logistical regression coefficients, x vector containing independent variables for surface age, traffic load, surface thickness and pavement/subgrade strength or other variables, and A, B are constants when their either is crack initiation or not.

2.2 Markov decision process

There are already a lot of publication regarding Markov Decision Processes (MDP's). One particularly interesting is Dekker, Nicolai and Kallenberg, 2007. In this paper a detailed explanation is given about MDP's . The main assumption of a Markov chain is that the present state includes all the information needed for future predictions, meaning that information from previous is not required. Because of this the transition probabilities P_{ij} can be defined as follows:

$$P_{ij} = P(X_{t+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_t = i) = P(X_{t+1} = j | X_t = i).$$

If the transition probabilities do not depend on t, then the Markov chain is stationary, but according to Plasmeijer, 1999 most of the road deterioration process are not stationary over time. To model road deterioration as a Markov chain one has to take into account that this will most likely result in a non-stationary process.

The Markov Decision Chain is an Markov chain that can be altered by actions, and where the optimal actions can be found. The chain is defined by E, A, P and r, where E denotes the state space, $A(i)$ the action set in state $i \in E$, $P(i,j)$, $i, j \in E$ the transition probabilities and $r_i(a), i \in E, a \in A(i)$, the immediate costs in state i when choosing action a.

By a policy $\boldsymbol{\beta}$ we mean a sequence of decision rules $(\beta_i(a), a \in A(i), i \in E)$ where $\beta_i(a)$ can be interpreted as the probability that action $a \in A(i)$ is chosen when in state $i \in E$. A policy is deterministic if all the decision rules are nonrandomized.

A criterion to find the best action a can be done by either optimizing the long-run expected average rewards (equation 1) or by finding the lowest expected total α-discounted cost (equation 2) by solving a set of equations.

$$g_i(R) = \liminf_{N \to \infty} \frac{1}{N+1} \sum_{k=0}^{N} \sum_{j \in E} P_{ij}^k(R) r_j(\pi^{k+1}) \qquad (2.1)$$

$$v^{\,\alpha}(R) = \sum_{k=0}^{\infty} \alpha^k P^k(R) r(\pi^{k+1}) \qquad (2.2)$$

Here π is a decision rule, $P(\pi)$ the matrix with $P_{ij}(\pi(i))$ as the i-th element, $r(\pi)$ vector with $r_i(\pi(i))$ as i-th element. $P^k(R) = P(\pi^k) \dots P(\pi^1)$ and $P^0(R) = I$, the identity matrix. If for equation 1 the state-space is finite then there is an average optimal policy, and in the case of equation 2 their exists an unique solution if $\alpha < 1$. A detailed procedure how to optimize can be found Dekker, Nicolai and Kallenberg, 2007.

2.2.1 Solution methods

There are three solution methods for finding optimal policy: policy improvement, value iteration and linear programming. the idea of policy improvement is to start with an initial policy and finds with each step for each state an improving action.

Value iteration repeatedly evaluates the optimality equation in order to find the solution. This algorithm is considered to be faster than the policy improvement method when the transition matrix is sparse and only few transitions are possible.

Linear programming requires you to formulate a LP for the optimal criterion. The advantage of this method is that standard LP solver can be used and one can easily set restrictions on the limiting probabilities of certain states.

2.2.2 Example

An example of practical use of Markov decision models is done in Golabi, Kulkarni, and Way, 1982. They developed a Network Optimisation System (NOS) for the State of Arizona , that finds optimal maintenance policies considering minimal road quality conditions, via linear programming. This NOS exists of a long term optimisation model, minimizing the yearly long run average maintenance cost and a short term model that minimises the total expected maintenance cost of a period of years. They considered four variables important for evaluating pavement performance: ravelling, present cracking, last-year change in cracking, and index tot the first crack. They claimed that in the first year 14 million dollars were saved and forecasted another 101 million in savings for the next four years. However Wang, and Zaniewski, 1996 reported that the steady state condition was never reached due to fluctuation in budgeting, pavement behaviour and transition probabilities.

2.2.3 Concerns Regarding MDP's

One of the main problems when dealing with Markov decision Models is that the state-space tends to be large, which consequently leads to high solving times . In the case of Dekker, Plasmeijer and Swart, 1998 for example, there are multiple deterioration processes affecting the road, causing the state space to be multi-dimensional. One way to deal with the high solving times is to use approximate dynamic Programming. van Roy, 2002 has given a good overview of Neuro-Dynamic Programming (NDP). The NDP algorithms are used to overcome the curse of dimensionality through use of parameterized function approximators that approximate the value function in a way similar to regression.

For an approximation of the value function to work two thing are necessary. First the parameterization that gives a good approximation needs to be chosen. Usually some experience or analysis that provides information about the shape of the function is required. Secondly, algorithms for computing parameter values are needed.

In the paper of Marbach and Tstisklis, 2001 a simulation-based algorithm for optimizing the average reward in a finite-state Markov decision process is discussed. To overcome the curse of dimensionality parametric representations are used. A class of policies is described in terms of a parameter vector θ. To estimate the gradient of the performance metric with respect to θ, is estimated using simulation, and the policy is improved by updating θ is a gradient direction. The authors in this paper claim the algorithm developed in their paper works very well, although they recommend it is tested further.

2.3 Road Maintenance with Deterministic Deterioration

The paper of Karabakal, Bean and Jack, 1994 shows us that one does not have to use a Markov decision process and probabilistic deterioration to schedule pavement maintenance. They formulate the problem of scheduling pavement maintenance over time to minimize cost under budget constraints as an integer program. The assumption is made that the pavement deterioration is known with certainty once the condition and maintenance action are known. The integer problem is solved via a heuristic: first they solve the problem without budget constraints after which the budget violation are minimized. The advantage of working with budget constraints is that governments work

with annual budgets. Markov decision models rarely work with budget constraints, and the outcome usually causes the expenditures to fluctuate widely over time. The disadvantage of using budget constraints is the option of solving each street segment's maintenance problem independently is lost, and the problem requires simultaneous consideration of all maintenance decisions in each time period. The authors claim that their heuristic works very well when with low budget violations.

3. ROAD ASPACTS

This section contains details on processes involving road deterioration and maintenance. The first section will shortly explain the road segment involved and which type of asphalt that road segment is assumed to be. Then a short introduction is given to the four main deterioration processes and there is briefly discussed when the road condition is acceptable, almost unacceptable and unacceptable. The last section is about the maintenance action that can be executed in order to preserve road quality.

3.1 Road Segment Features

The maintenance policy discussed in this paper focuses only on a segment of the road, so not on an entire road. A segment (lane-sector) is a part of a lane of 100 meters length as is shown in figure 3.1.

figure 3.1: configuration of a 2x2-lane road

The segment discussed in this paper is assumed to be a porous asphalt road with a permeable concrete structure. Although dense asphalt is also still used, porous asphalts are more and more used because of their less noisy nature, and reduction in splash and spray effects. The downside of this type of asphalt is that it is more expensive to maintain and is very susceptible to ravelling (Hagos, 2009).

Figure 3.2: Growth of porous asphalt surfaces on the main roads in the Netherlands

Source: Hagos, 2009

7
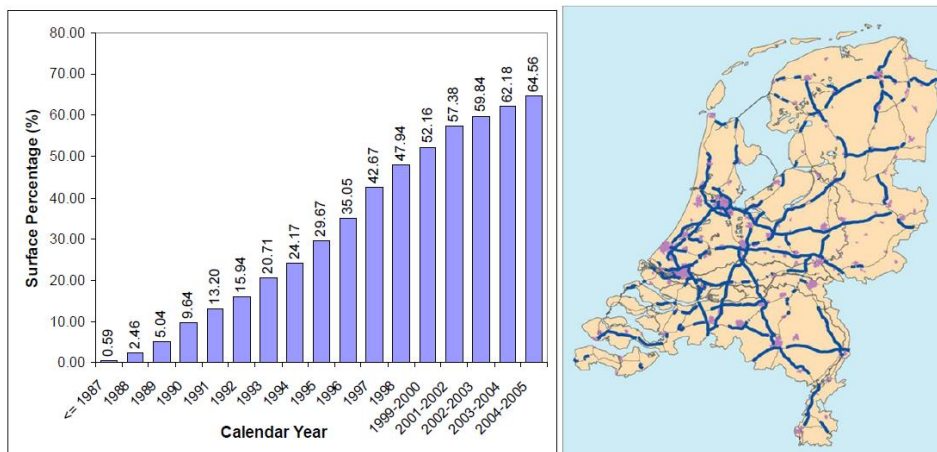
3.2 Deterioration of Roads

The road deterioration process involves four main damage features that will be discussed. First there is *cracking*, which covers all types of cracking that can take place on roads. The seriousness of cracking is measured as a percentage of the road length that is covered by cracks. The second damage feature is *ravelling*. This is the crumbling up of the asphalt layer as a result of the dislodgement of aggregate particles. The seriousness of ravelling is just like cracking measured as a percentage of the road suffering from ravelling. The third and fourth main damage features are *longitudinal-* and *transversal unevenness*. Longitudinal unevenness describes every unevenness along the length of the road while transversal unevenness describes the unevenness along the width of the road, mainly concerning rutting. The international roughness index (IRI) is used for measuring longitudinal unevenness, unlike transversal unevenness, that has no specific index, which is measured in terms of the average number of millimetres difference in height.

For simplification lets classify the seriousness of the damage in three categories: *Acceptable, Almost Unacceptable, Unacceptable*. Only when the seriousness of the damage is unacceptable is it obligated to repair the damage, but repairing the road in the other two categories is also allowed. The only state of interest is the Unacceptable state since the objective is to avoid reaching this state. The line between Acceptable and Almost Unacceptable is called the warning level, and the line between Almost Unacceptable and Unacceptable is the failure limit as is shown in Figure 3.3. Table 3.1 specifies the warning level and failure limit for each damage feature.

figure 3.3: classification of the damage
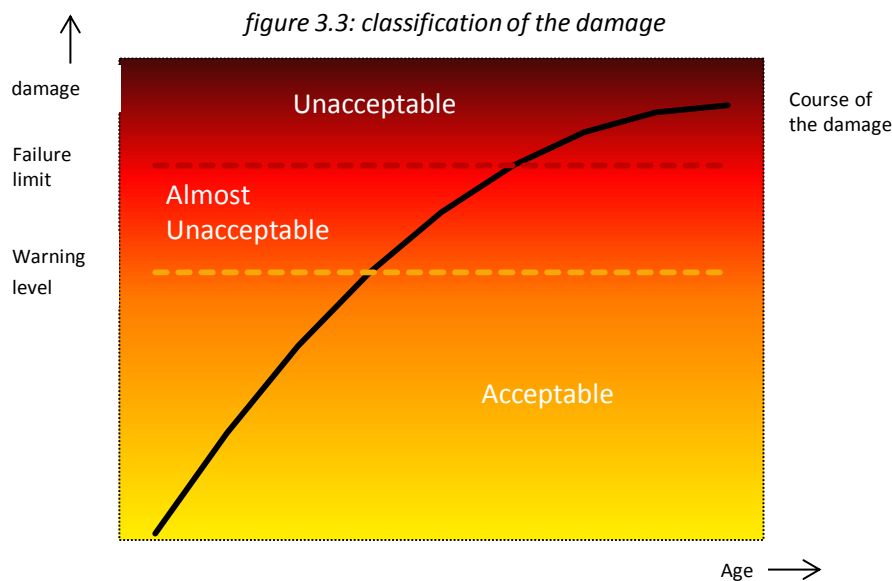
Table 3.1: warning level and failure limit per damage feature

Damage feature	Acceptable	Almost Unacceptable	Unacceptable
Cracking	$\leq 10\%$	$10 - 25\%$	$\geq 25\%$
Ravelling	$\leq 15\%$	$15 - 30\%$	$\geq 30\%$
Longitudinal unevenness	$\leq 2\frac{1}{2}$ IRI	$2\frac{1}{2} - 3\frac{1}{2}$ IRI	$\geq 3\frac{1}{2}$ IRI
Transversal unevenness	≤ 15 MM	$15 - 18$ MM	≥ 18 MM

Source: Van der Horst, et al. 1991

Each type of damage has its own deterioration process, which shall be discussed in section 4. For simplification let us assume that there is no correlation between the different processes, and that

there is also no correlation between the deterioration of adjacent segments. This seems like a harsh assumption but it keeps the problem simple for now. The following tables give a better picture of the deterioration process, the data was assembled by Plasmeijer, 1999 from road experts of the Road and Hydraulic Engineering Division (DWW).

Table 3.2 shows the failure probabilities of porous asphalt after 5, 10, 15 and 20 for the four main damage features. Table 3.3 shows how the process of deterioration of the four main damage features develops in the course of time. Table 3.4 outlines the minimal, mean and maximal duration before the damage reaches the failure limit if the road is in perfect condition.

Table 3.2: porous asphalt failure probabilities

Damage feature	5	10	15	20
Cracking	0.00	0.00	0.10	0.50
Ravelling	0.00	0.40	0.90	1.00
Longitudinal unevenness	0.00	0.10	0.20	0.30
Transversal unevenness	0.00	0.00	0.25	0.70

Table 3.3: porous asphalt speed of deterioration

Damage feature	Speed
Cracking	Increasing
Ravelling	Increasing
Longitudinal unevenness	Constant
Transversal unevenness	First decreasing, then increasing

Table 3.4: porous asphalt estimated lifetimes

Damage feature	min	mean	max
Cracking	15	>15	>15
Ravelling	8	12	>15
Longitudinal unevenness	7	>15	>15
Transversal unevenness	12	>15	>15

The data speak mostly for themselves with the only remarkable things being the susceptibility of porous asphalt to ravelling, and that the transversal unevenness speed deterioration changes over time.

2.3 Maintenance Actions

When facing road deterioration several action can be taken to repair the road. The possible actions that can be taken are listed below with the exception of conservation because of safety issues.

- Regeneration: pavement or remix of the toplayer
- Replacement: milling and inlay of the toplayer
- Overlaying: addition of a new layer of asphalt
- Rut filling: addition of emulsion-concrete to eliminate the ruts

- Profile correction: adjustment of the road profile by milling and levelling

Table 3.5 gives an overview of which action can be chosen to repair a certain damage feature, and table 3.6 shows the impact on the life expectancy of the road per damage feature, when a certain maintenance package is chosen. The latter originated from the Dutch Directorate General of Public Works and Water Management (RWS), which is supported by the Road and Hydraulic Engineering Division (DWW).

Table 3.5: types of maintenance actions and their impact on the road

Damage feature	Regeneration	Replacement	Overlay	Rut filling	Profile correction
Cracking	x	x	x		
Ravelling	x	x	x		
Longitudinal unevenness			x		x
Transversal unevenness	x	x	x	x	

Table 3.6: possible maintenance packages and their effects on the remaining lifetime

Package	Cracking	Ravelling	Longitudinal unevenness	Transversal unevenness
1. 50u/i(100%)	17	t	t	18
2. 40u/i(5%)+50STA+PAC	17	t+4	t+3	16
3. 60u/i(75%)+50STA+PAC	17	11	t+3	20
4. Milling and levelling+PAC	t	t	18	16
5. Milling+100STA+PAC	17	t+8	18	20
6. 150u/i(100%)	17	t+5	t	20

Source: DWW-RWS, Delft; PAC = porous asphalt concrete, STA = gravel asphalt; t = expected residual lifetime before maintenance was taken place ; α u/i ($\beta\%$): milling and inlay of α MM asphalt layer with $\beta\%$ new asphalt.

Table 3.6 has two types of road quality improvement, nominal effects and relative effects. We assume that maintenance actions only affect the state of the road segment not the age, meaning that if maintenance has taken place the road will be in an better condition but will still follow the deterioration process at the time were it left off.

The maintenance policy only applies for an area of $350\ m^2$ ($3.5m * 100m =$ road segment). No additional cost are taken into account (road blocks, etc.), and every maintenance package is assumed to be executable on a road segment. Off course this is not realistic as some packages can only be executed on a whole carriage way, not just a single segment, but again it keeps the problem simple.

The cost are calculated as suggested by Plasmeijer, 1999 (formula 3.1) with unit prices from the DWW-RWS. As for the discount coefficient a percentage of 10% is taken what will have a surcharge effect for the relative small area we are maintaining. The results are presented in table 3.7. An important thing to notice is that the costs are assumed state independent, meaning the condition of the road does not play are role when determining the maintenance costs.

Table 3.7: cost per maintenance package

Package	€ / m^2
1.	4125
2.	15345
3.	23570
4.	17780
5.	21660
6.	9420

Source: DWW-RWS, Delft; august 1996

$$costs = (unit\ price\ per\ m^2) \cdot area \cdot \left(1 - dicount\ coefficient \cdot \left(1 - \frac{standard\ area}{area}\right)\right) \quad (3.1)$$

4. ROAD DETERIORATION PROCESS

This paper uses a probabilistic modelling approach to predict the road deterioration process. Deterministic road deterioration models often underestimate or overestimate the road deterioration process (Chua, et al. 1993), mainly because they use an average condition of the road to predict the road behaviour in the future while a probabilistic approach can assign different probabilities to each state in the future.

4.1 General Deterioration Models

The deterioration process for cracking, ravelling and longitudinal unevenness are modelled by the following Brownian motion, as is done in Plasmeijer, 1999 :

$$D_t = \mu t^q + \sigma \sqrt{t} \cdot U, \qquad t \geq 0\,, \qquad (4.1)$$

where D_t is the damage after t years, with D_0 = 0, q must be larger than 0 and represents the shape parameter of the deterioration process. If q is 1 then the process is stationary, else the deterioration speed is either increasing ($q > 1$) or decreasing ($q < 1$) over time. The parameter μ represents the trend parameter and σ is the volatility parameter. Then not only D_t is normally distributed with mean μt^q and variance $\sigma^2 t$ but also $W_{u,t} = D_u - D_t$, $u > t$, is normally distributed with mean $\mu(u^q - t^q)$ and variance $\sigma^2(u - t)$, as a result the transition probabilities can recursively be calculated using function 4.2.

$$\Phi\left(\frac{y - x - \mu(u^q - t^q)}{\sigma\sqrt{u - t}}\right), \quad u > t. \qquad (4.2)$$

4.2 Time-Varying Shape Parameter

As is shown in table 3.3, transversal unevenness follows a different deterioration speed pattern after a certain time period. To coop with this a time-varying shape-parameter is used. The model then becomes:

$$D_t = \mu t^{q(t;\mu)} + \sigma \sqrt{t} \cdot U, \qquad t \geq 0\,, \qquad (4.3)$$

where

$$q(t;\mu) = \frac{q}{\sqrt{e^{1-\left(2*t*\left(\frac{F}{\mu}\right)^{-\frac{1}{q}}-1\right)^2}}}. \qquad (4.4)$$

Here F is the failure rate and from $t = 0$ onwards the damage speed decreases overtime till $t = \frac{1}{2} * \left(\frac{F}{\mu}\right)^{\frac{1}{q}}$, after which the deterioration speed increases over time. Again the deterioration process D_t for transversal unevenness is normally distributed with mean $\mu t^{q(t;u)}$ and variance $\sigma^2 t$ and function 4.2 can be reused by only substituting q for $q(t;\mu)$. Clearly this process is non-stationary.

4.3 Function Parameters and Transition Probability Matrices

The used parameters for the deterioration functions D_t are displayed in table 4.1.

Table 4.1: parameters

Damage feature	μ	q	σ
Cracking	0.75	7/6	0.75
Ravelling	1.5	7/6	2.1
Longitudinal unevenness	0.175	1	0.42
Transversal unevenness	6.66	1/3	0.9

Source: Plasmeijer, 1999.

As is shown in table 4.3 most deterioration processes are non-stationary meaning that the policy is state and time dependent. The non-stationarity also results in having for each time t a different transition probability matrix.

To construct the transition probability matrices a few choices were made. To translate the deterioration curve into different states the deterioration curve has to be divided into intervals with same length. The difficult part here, is to choose 'good' intervals: choose the intervals too small and one will end up with a very large state-space, choose the intervals too large and there are simply too few states for any practical use. In this paper there are three different state descriptions used. This will help see the impact on expected average reward and computation time when the state-space changes. We have chosen for interval lengths of $i \in (2, 2.5, 4)$ years which represent one state. For the state descriptions see table A.1-A.3 in the appendix. As an example the deterioration curve of the Ravelling process is shown in figure 4.1.

Figure 4.1: Ravelling process (deterministic part)
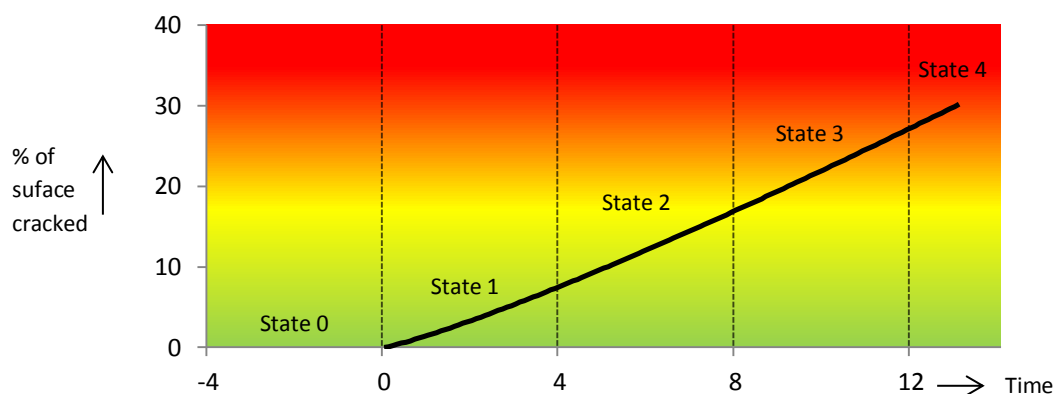
That leaves us with two more problems. First the possibility of the road quality to improve is removed by adding all the probabilities of previous states to the state at time t, so a diagonal matrix will appear. With Brownian process it is possible just like stocks to fluctuate, and that is not desirable when modelling road deterioration therefore any 'improvement' of the road is seen as 'not deteriorating'. The second problem concerns determining the final state. Not very much information is available about states beyond the failure limit and it is therefore difficult to predict the deterioration behaviour of these states. So the last state includes all states beyond the failure limit as if the road after the failure rate was not operational any more. An example of a transition probability matrices for the ravelling process is show in table 4.2.

Table 4.2: Transition Probability Matrix for the Ravelling Process at time t = 1

state	0	1	2	3	4	5	6	7	8
0	0.1869	0.5756	0.2341	0.0034	0	0	0	0	0
1	0	0.1869	0.6790	0.1336	0.0005	0	0	0	0
2	0	0	0.1869	0.7143	0.0986	0.0002	0	0	0
3	0	0	0	0.1869	0.7345	0.0785	0.0001	0	0
4	0	0	0	0	0.1869	0.7481	0.0650	0	0
5	0	0	0	0	0	0.1869	0.7578	0.0553	0
6	0	0	0	0	0	0	0.1869	0.7653	0.0478
7	0	0	0	0	0	0	0	0.1869	0.8131
8	0	0	0	0	0	0	0	0	1

5. MARKOV DECISION PROCESS

As mentioned earlier in this paper the problem consists of finding an optimal maintenance policy for a road segment. The policy must try to minimize both the time an road is in an unacceptable state and the costs involved. In order to make a decision which maintenance action is suited best considering the future condition of the road must been know. The future state of the road can be predicted with the road deterioration process explained in section 4. The idea is to use a Markov decision model that find a policy describing which maintenance action should be taken given the current state of the road.

5.1 Problem formulation

Consider the four deterioration processes Cracking, Ravelling, Longitudinal Unevenness and Transversal Unevenness that are observed at discrete time points to be in any one of C, R, LU and TU possible states respectively, which are numbered by $1, 2, \dots, D$, $D \in \{C, R, LU, TU\}$. Then the road segment can be in any of the M possible states that are defined as (i, j, k, l) with $i \in 0, \dots, C$, $j \in 0, \dots, R$, $k \in 0, \dots, LU$, $l \in 0, \dots, TU$. The total number of states, M, can be calculated by $C * R * LU * TU$. For ease of notation let us give each state (i, j, k, l), an unique number $q \in 1, \dots, M$. Let U be the set of states when at least one of the four main damage features reaches an unacceptable state according to table 3.1. After road inspection at time t the state of the process is observed, and a maintenance action is chosen from the set A. The set A contains all maintenance packages displayed in table 3.6 with addition of an action where no maintenance takes place.

If the process is in state q and action a is chosen then the next state of the system is determined according to the transition probabilities $P_{q,r}(a)$. The transition probabilities can extracted from the

transition probability matrices discussed in section 4.3. If we let X_n denote the state of the process at time n and a_n the action chosen at time n, then $P_{q,r}(a)$ becomes:

$$P\{X_{n+1} = (m, n, o, p)|\ X_n = (i, j, k, l), a_n = a\}$$
$$= P\{C_{n+1} = m|C_n = i, a_n = a\} * P\{R_{n+1} = n|R_n = j, a_n = a\}$$
$$* P\{LU_{n+1} = o|LU_n = k, a_n = a\} * P\{TU_{n+1} = p|TU_n = l, a_n = a\}, \qquad (5.1)$$

where $P\{C_{n+1} = m|C_n = i, a_n = a\}$, $P\{R_{n+1} = n|R_n = j, a_n = a\}$, $P\{LU_{n+1} = o|LU_n = k, a_n = a\}$ and $P\{TU_{n+1} = p|TU_n = l, a_n = a\}$ are the transition probabilities for respectively Cracking, Ravelling, Longitudinal Unevenness and Transversal Unevenness.

The objective is to use the Markov decision model in order to find a policy $\boldsymbol{\beta}$ that minimises the long run expected cost per unit time, under the restriction that the road should never reach an unacceptable state. The policy is a rule for choosing a maintenance action at a certain state.

5.2 Linear Programming

According to Ross, 2010 the linear program needed to solve for acquiring the optimal policy is as follows:

$$\min_{\pi\ =(\pi_{qa})} \sum_q \sum_a \pi_{qa} R(q, a)$$

subject to:

$$\sum_q \sum_a \pi_{qa} = 1,$$

$$\sum_a \pi_{ra} = \sum_q \sum_a \pi_{qa} P_{qr}(a), \quad for\ all\ r$$

$$\sum_{\delta \in U} \sum_a \pi_{\delta a} = 0$$

$$0 \leq \pi_{qa} \leq 1, \quad for\ all\ q, a \qquad (5.2)$$

Here π_{qa} is the limiting (or steady-state) probability, and $R(q, a)$ the cost when the process will be in state q and action a will be executed if policy $\boldsymbol{\beta}$ is chosen. $\boldsymbol{\beta} = \{\beta_q(a), a \in A, q = 1, \dots, M\}$ is a randomized policy which is defined by:

$$\beta_q(a) = \frac{\pi_{qa}}{\sum_a \pi_{qa}}, \quad for \sum_a \pi_{qa} > 0$$

where $\beta_q(a)$ can be interpreted as the probability that maintenance action a is chosen in state q. When the sum over all limiting probabilities of a certain state is equal to zero the maintenance policy for that state is undefined. However it can also be shown that there is a π^*_{qa} minimizing the objective function, that is zero for all $a \in A$ except one, meaning the policy is non-randomized.

The advantage of linear programming is that one can easily ad restrictions on the fraction of time the process spends in certain states. This is already done with the linear program (5.2) where the last

equation represents the time spent in U, (the set bad states), must be equal to zero. It is also possible to add inequality constraints to regulate road-quality. One could simply add constraints similar to following:

$$\sum_{\delta \in U_b} \sum_a \pi_{\delta a}, \leq \alpha$$

where U_b is the set of states you wish to regulate, and α the maximum fraction of time you are willing to spent in these states. Of course when adding these constraints the optimal policy will most likely choose multiple actions a, for some states q. This because the added restrictions prevents the optimal policy from fully choosing the cheaper maintenance action, and instead mixes the cheaper maintenance action with more expensive action to satisfy the constraints. So when using inequality constraints most of the time it will not be possible to convert the optimal policy into a non-randomized policy unless an increase in average-cost is accepted.

This linear program can be solved by the simplex method, but because the number of states M can become large very fast we first try to rewrite the linear program (5.2). Clearly the last equation of the linear program can be rewritten as $\pi_{\delta a} = 0$ for all $\delta \in U, a \in A$, as every $\pi_{qa} \geq 0$. Now these variables are fixed and can be left out of the picture when solving the linear program. Also the maintenance costs are state independent so $R(q, a) = R(a)$. Define U^c as the complement of all the bad states (the set of all Acceptable and Almost Unacceptable states) the linear program can be defined as follows:

$$\min_{\pi = (\pi_{qa})} \sum_{q \in U^c} \sum_a \pi_{qa} R(a)$$

subject to:

$$\sum_{q \in U^c} \sum_a \pi_{qa} = 1,$$

$$\sum_a \pi_{ra} = \sum_{q \in U^c} \sum_a \pi_{qa} P_{qr}(a), \quad for\ all\ r$$

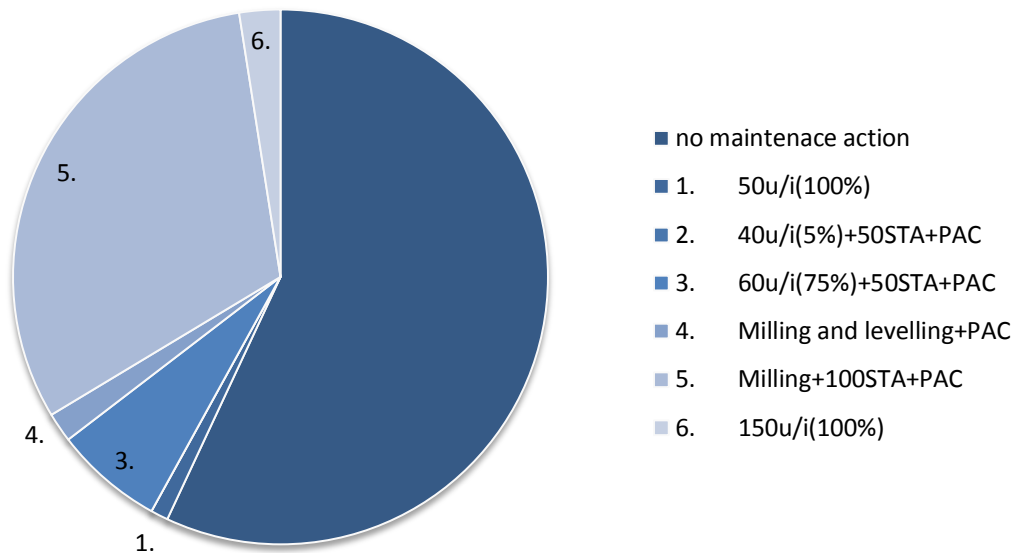$$0 \leq \pi_{qa} \leq 1, \quad for\ all\ q \in U^c, a \in A \qquad (5.3)$$

5.3 Results
There are three policies $\in (MDP_2, MDP_{2.5}, MDP_4)$ presented, one for each state-space determined in section 4.3 and time is assumed to be constant at $t = 1$ for all policies. First the policy with the smallest state-space is determined. With the information obtained from that policy the computation time of the other policies can be reduced.

For the first state-space the deterioration curve was split into intervals of four years. This meant that the number of states for the four main deterioration processes Cracking, Ravelling, Longitudinal Unevenness and Transversal Unevenness became respectively 7, 5, 7 and 7. The total amount of states for the multidimensional problem quickly becomes very large, in this case consisting of $7^3 * 5 = 1,715$. The amount of unacceptable states sum up to $1,715 - 6^3 * 4 = 851$ states total. There are seven possible maintenance actions that can be chosen, this includes the maintenance

actions displayed in table 2.6 plus the possibility of executing no maintenance action. To solve the linear program Cplex, 2012 was used on a Windows 7 64-bit computer, 3.60 GHz CPU, 16 GB RAM, and was solved within a reasonable computation time of under 10 seconds. Policy MDP_4 contains too many results to display in this paper and only a diagram of the fractions of time spent choosing the maintenance actions is shown in figure 5.1.

Figure 5.1: distribution of maintenance actions for policy MDP_4

- no maintenace action
- 1. 50u/i(100%)
- 2. 40u/i(5%)+50STA+PAC
- 3. 60u/i(75%)+50STA+PAC
- 4. Milling and levelling+PAC
- 5. Milling+100STA+PAC
- 6. 150u/i(100%)

The yearly expected average costs of policy MDP_4 are € 8891.09. The first thing we immediately notice is that the fraction of time spent executing no maintenance action is over 50%. This is a good thing, meaning that most of the time there are no cost incurred. The second largest fraction of time is spent executing maintenance action 5. Despite its high cost maintenance action 5 is clearly favourite, probably because it repairs all four damage features. The most remarkable observation is that maintenance action 2 is not chosen at all for any of the recurrent states. Apparently there is always another maintenance action that is either cheaper and/or more efficient than maintenance action 2. Another remark that cannot be seen from figure 5.1 is that maintenance actions 1-6 are only chosen in a state where a transition probability (tp) > 0 to an unacceptable state exists, called action states. Intuitively this is not very remarkable since the target is to avoid the unacceptable states, but one could also think conversely: in states with tp's to all unacceptable states equal to zero, one should not execute any maintenance action, called non-action states. With prior knowledge of these patterns one could considerably reduce the amount variables when solving the linear program, at the cost of a slightly increase in the expected average costs.

Next a larger state-space is used to obtain policy $MDP_{2.5}$. The four main deterioration processes Cracking, Ravelling, Longitudinal Unevenness and Transversal Unevenness, have respectively 10, 7, 10 and 10 states. The size of the state space consists of $10^3 * 7 = 7,000$. The amount of unacceptable states sum up to $7,000 - 9^3 * 6 = 2,626$ states total. Initially there were seven possible maintenance actions that could be chosen, but the policy of MDP_4 suggested that maintenance action 2 might not be 'optimal to use'. So to save computation time, this maintenance action is left out when formulating the linear program. Let us also process the idea of action states, and non-action states in the linear program, so one could only choose a maintenance action in the

action states. This results in a substantial variable reduction for the linear program. At first there were *(state-space size) * (maintenance actions)* $= 7,000 * 7 = 49.000$ variables. When rewriting LP formulation 4.2 to 4.3 the Unacceptable States were removed from the LP reducing the amount of variables to $(7,000 - 2.626) * 7 = 30.618$. Now with the removal of maintenance action 2 this becomes $(7,000 - 2,626) * 6 = 26,244$, and with the introduction of action and non-action states the number of variables further reduces to (*non-action states * maintenace action) + (action states-Unacceptable states) * maintenance actions* $= 797 * 1 + (6,203 - 2,626) * 5 = 18,682$. The variables are now reduced with approximately 62%, also reducing computation time to solve the LP considerably to only 58 seconds. Of course it is very unlikely that the optimal policy is found but the policy found is most likely to be near optimal. Just like policy MDP_4, this policy has also too many results to display and instead a diagram is shown in figure 5.2.

Figure 5.2: distribution of maintenance actions for policy $MDP_{2.5}$
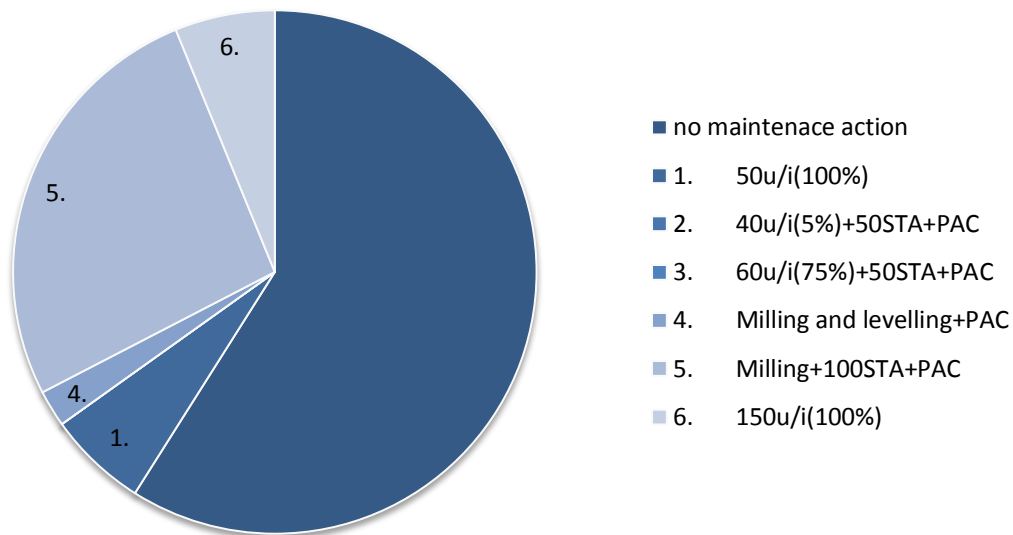
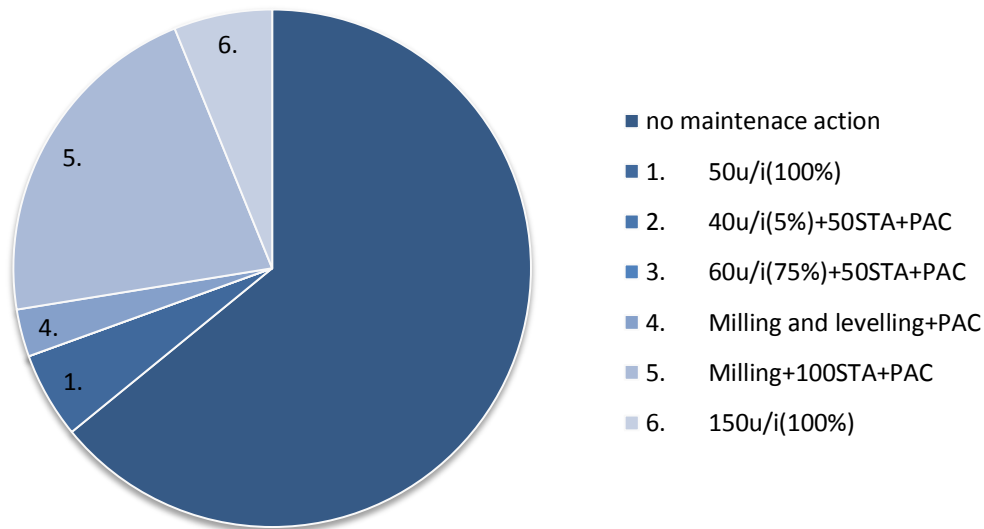

The expected average costs of policy $MDP_{2.5}$ are €6962.27, considerably less than that of policy MDP_4. This probably has to do with the shift in time spent executing the two most expensive maintenance actions (3 and 5) to cheaper actions. Maintenance action 3 even became obsolete. The previous made assumed assumptions however, could have enabled this.

The final state space used has respectively 12, 8, 12, 12 states for the four main damage processes. This results into 13,824 states total of which 2,064 are non-action states, and thus 11,760 action states where 4,507 of them are Unacceptable states. Similar assumptions are made as with policy $MDP_{2.5}$, only this time also excluding maintenance action 3. This would most certainly lead to an even further deviation from the optimal policy, but hopefully not by too much. By doing so the amount of variables is reduced from 96,768 to a stunning 31,076, roughly one-third of the initial amount. As expected the computation time has still increased significantly to 627 seconds, although still reasonable.

The diagram of policy MDP_2 , displayed in figure 5.3, is not surprisingly very similar to that of policy $MDP_{2.5}$, with only a slight decrease in time spent executing maintenance action 5 causing the fraction of time spent executing no maintenance action to increase, therewith explaining the decrease in yearly expected average costs to €5965.17.

Figure 5.3 distribution of maintenance actions for policy MDP_2

- no maintenace action
- 1. 50u/i(100%)
- 2. 40u/i(5%)+50STA+PAC
- 3. 60u/i(75%)+50STA+PAC
- 4. Milling and levelling+PAC
- 5. Milling+100STA+PAC
- 6. 150u/i(100%)

5.4 Pros and Cons

First let us start with the most obvious flaw when using a Markov decision process to find an maintenance policy. Clearly when the state-space becomes large the computation time goes up. Since we deal with a multidimensional state space, it can become very large very fast causing computation time to be unreasonable. There are three ways to deal with the large state-space of which we already saw two. One could, like is done for policy MDP_4, limit the state-space per deterioration process in order to reduce the multidimensional state-space. The advantage of this method is that one would still end up with an optimal policy, but as we clearly saw, when using smaller state-spaces the expected average costs tend to be much higher than when using larger state-spaces. Another method is to reduce the variables involved when solving the LP. This was done for policies MDP_2 and $MDP_{2.5}$. The policies obtained are unlikely to be optimal anymore and there are not really any rules for reducing variables other than using intuition and common sense. The final method used to find the maintenance policy is by using approximate dynamic programming. This mainly includes Neuro Dynamic Programming or Reinforcement Learning. However we do not go into this subject, and instead recommend reading Bertsekas and Tsitsiklis, 1996 for more information on this matter.

Less troublesome problems are the undefined states, which are defined as states where $\sum_a \pi_{ra} = 0$ for $q \in M$. These states are expected never to be reached (unacceptable states and some almost unacceptable states) or passed by only once (early states, states were most deterioration process have done little damage yet). This leaves room open for discussion what to do when one finds himself in these states. For the latter category it is quite simple: one should execute no maintenance action. For the first category it is somewhat more complicated but since the long term expected average cost are not affected when such an event occurs the only concern is to 'get back on track' as soon as possible. One way to accomplish this is to use the EAC method (explained in section 6).

The advantages of using a Markov decision process are also quite obvious, as when the linear program is solved, one ends up with an 'optimal' policy. Also it is fairly easy to control road condition

as this is done by simply adding constraints to LP. For the smaller problems the LP can be solved using standard LP solving software.

6. COST-EFFECTIVE MAINTENANCE

Now (near) optimal maintenance policies have been found a comparison can be made with a more simple approach used by the DWW. First a short introduction on how to find a policy using this approach is given, after which results are given and a comparison is made.

6.1 Equivalent Annual Cost Method

A more simple an intuitively logic way to determine a maintenance policy is to repair the road when it has reached an Almost Unacceptable or Unacceptable state and choose the maintenance action that results in the lowest cost for one expected residual life year.

One of the advantages of this method is that policies for each state can be calculated individually, fixing computation time. The impact of this is that state-space does not play a role any more for computation time, allowing the state space to be very large, so large even that states are not defined by intervals but by their exact damage value. Of course this is only useful when road inspections can give these exact damage values. The down side of this method is that one cannot control the fraction of time spent in certain states, causing to road to possibly end up in Unacceptable states for a period of time.

The Acceptable states are defined as states where the damage values of all the four deterioration process are below their warning level. All the Acceptable states have the same maintenance policy and that is to execute 'no maintenance action'. For the other states where at least one of the deterioration processes has exceeded the warning level maintenance policy $a \in A$ is chosen for which

$$EAC_q(a) = \frac{costs\ of\ execution\ maintenace\ packace\ a}{expected\ gain\ in\ years\ of\ action\ a\ when\ in\ state\ q}, \qquad (6.1)$$

is minimised, where the expected gain in years of action a when in state q is the difference between the expected residual lifetime right after execution of action a, and the expected residual lifetime right before execution of action a.

This method is clearly a greedy algorithm for finding a maintenance policy, and therefore it is unlike to find an optimal solution. It makes a locally optimal choice at each decision point by selecting the maintenance action that gives the 'best bang for the buck' not taking into account anything else.

5.2 Results

Instead of defining the Almost Unacceptable states as is done in section 3.2, they are defined as states in which a transition probability > 0 exists to an unacceptable state, the so-called action-states. Although this complicates the EAC method by involving a Markov process, it should not differ too much since these states overlap the Almost Unacceptable states and ensures a more equal comparison with the linear program.

For the comparison that is done later with the Markov decision process there are also three policies, EAC_2, $EAC_{2.5}$, EAC_4, constructed, instead of just one. The computation time for all the policies is per state constant and neglectable, when the action-states are predetermined.

For equal comparison we shall also compute the yearly expected average cost and the fraction of time spent in Unacceptable states. This is done by constructing transition probability matrices for the three maintenance policies according to the Markov process, after which the limiting probabilities could be calculated.

The distribution of maintenance actions for policy EAC_4 is displayed in figure 6.1. The figure clearly shows that this policy prefers the use of cheap maintenance actions at the cost of a relative small fraction of time spent executing no maintenance action. Unlike policy MDP_4 who prefers a high fraction of time spent executing no maintenance action at the expense of the using more expensive maintenance actions. Despite the fact that the fraction of time spent executing maintenance action 3 is zero, does not mean it is not chosen in any of the states. On the contrary maintenance action 3 is chosen only in transient states, but they do not play a role in the yearly expected average costs. For this policy these costs are € 9050.18, only slightly higher than that of policy MDP_4. If we look at the fraction time spent in unacceptable states we see that this adds up to 0.0107, not that very high, but what does strikes is that only the longitudinal unevenness process exceeds its failure limit. It is also in states where this damage process in combination with other deterioration processes reaches serious damage levels that the policy EAC_4 deviates from policy MDP_4, choosing maintenance actions that only deal with the other damage processes and not the longitudinal unevenness. Policy EAC_4 also prefers maintenance action 6 over 3 when dealing with serious damage levels of cracking, ravelling and transversal unevenness, in contrary to policy MDP_4. If it was not for these factors the two policies would almost look identical.

Figure 6.1: distribution of maintenance actions of policy EAC_4
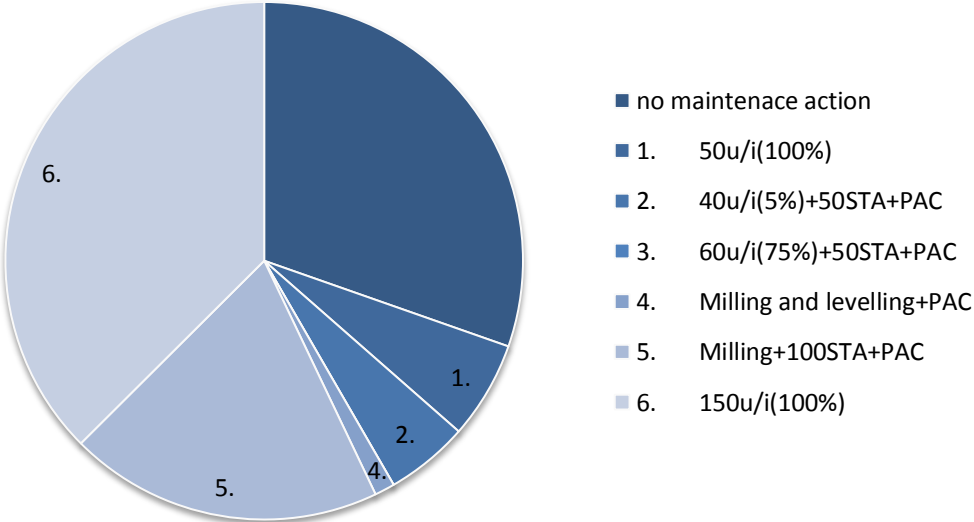
As was expected the yearly expected average costs of policy $EAC_{2.5}$ have dropped substantially to €7034.30. An explanation might lie in figure 6.2. Here we see a significant drop in fraction of time spent executing maintenance action 6, and an increase in fraction of time spent doing nothing and executing the cheaper maintenance action 1. The fraction of time spent in unacceptable states has

also dropped to 0.0003, and again only the longitudinal unevenness process exceeds its failure limit. When compared to policy $MDP_{2.5}$ we see the same problem with the longitudinal unevenness we saw earlier when comparing policy EAC_4 with MDP_4. Further comparisons are hard to make since policy $MDP_{2.5}$ has left out maintenance action 2, but they do share that maintenance action 3 is not executed in any of the recurrent states.

Figure 6.2: distribution of maintenance actions of policy $EAC_{2.5}$
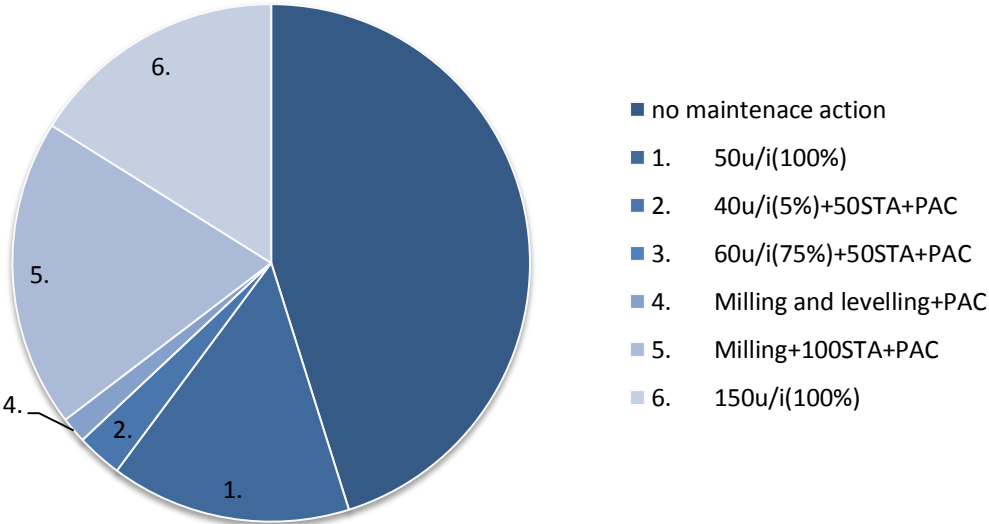
The final policy EAC_2 has yearly expected average costs of €6550.07, again a reduction in cost when compared to policy $EAC_{2.5}$. This is caused by a slight reduction of fraction of time spent executing maintenance actions 5 and 6, as is shown in figure 6.3. However when comparing these costs with those of policy MDP_2 they are not as close anymore as was the case with policies EAC_4 and MDP_4, and $EAC_{2.5}$ and $MDP_{2.5}$. Policy EAC_2 is approximately 10% more expensive than its counterpart policy MDP_2.

Figure 6.2: distribution of maintenance actions of policy $EAC_{2.5}$
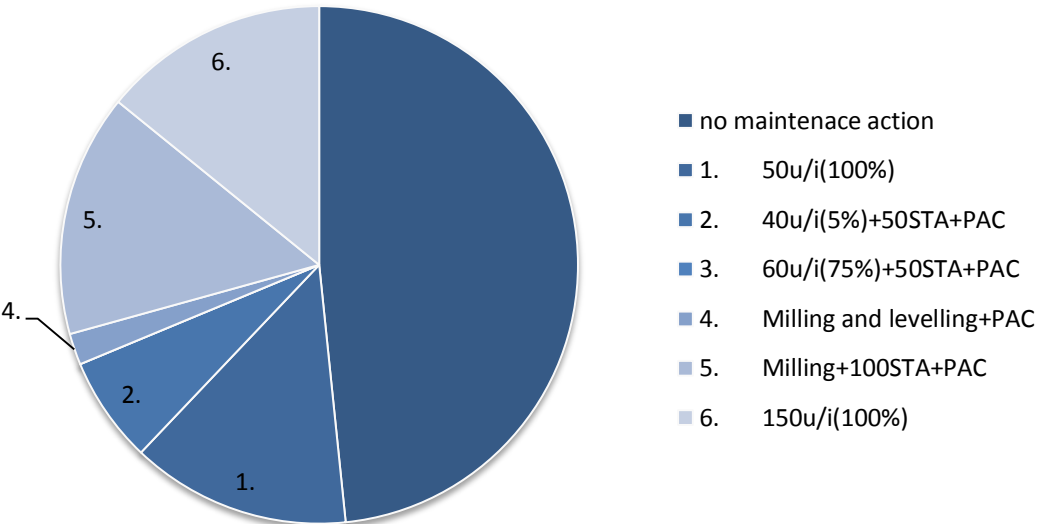
The real issues of the EAC_2 is that the policy does not deal efficiently when multiple deterioration processes have reached advanced damage levels and do not share the same residual life times. For example when reaching state (3,1,6,7) the EAC method only focuses on the last deterioration process (transversal unevenness) because it can be dealt with either maintenance action 1 or 2 instead of using a more expensive maintenance action 4 which tackles both deterioration processes. Only to find out later that one of these more expensive maintenance action is necessary to coop with the longitudinal unevenness. This especially applies when longitudinal unevenness reaches critical levels together with other deterioration processes, as we saw earlier with the two other EAC policies. The method seems to forget the longitudinal unevenness and resorts to one of the cheapest maintenance actions to get only one or two years of residual life time. This eventually leads to a fraction of time spent in Unacceptable states of 0.0005.

5.3 Comparison

When we look at the yearly expected average cost the MDP policies are only slightly cheaper than EAC policies, with the exception of policy $\boldsymbol{MDP_2}$. It is when the largest of the three state spaces is used that the reduction in yearly expected average costs is noticeable when compared with the EAC method.

Clearly the MDP policies have the advantage over EAC policies when it comes to yearly expected average cost, but not so much when it comes to computation times. For the EAC policies these are constant, since one does not need to know all the policies for all the states at a certain point in time. One only needs the know the policy of the current state. For the MDP policies computation time is a serious issue. We earlier saw computation time increasing severely when larger state-spaces were used, from 10 second to 10 minutes. So when considering even larger state-spaces the computation time will become problematic. That brings us back to the expected average costs. We already established a negative relation between the size of the state-space and the expected average cost, meaning it would be cheaper to choose a larger state-space when possible. For the MDP policies that relation is extended with computation time, and thus eventually limiting the size of the state-space, unlike EAC policies where the state-space can be infinitely large and thus costs can be brought further down.

There is one thing however the EAC policies do not account for and that is the restriction on fraction of time spent in unacceptable states. For the MDP policy that fraction is equals zero but for the EAC policies these fractions are respectively 0.0107, 0.0003, 0.0005. Although not equal to zero these fractions are very small.

7. CONCLUSION

To sum it up: first this paper showed how a Markov Decision Process was formulated and solved via linear programming. The results were then compared to the Equivalent Annual cost method, used by the DWW, to see if solving the MDP via linear programming had any practical use.

Solving the linear program following for the MDP certainly gives either an optimal, or near optimal policy and it is easy to regulate road condition via restrictions, but these advantages come at a price, for multiple deterioration processes causes the state-space to be multi-dimensional. One way to deal with this, is to limit the state-space per deterioration processes and thus the keeping the multi-

dimensional state-space small, but this will eventually lead to higher expected average costs. The other way to deal with the curse of dimensionality is to use Neuro-dynamic Programming explained by Bertsekas and Tsitsiklis, 1996. This however is not investigated in this paper and will only lead to a near optimal policy.

The method currently used by the DWW has the advantage of being a very simple method, requiring not much data and has very low computation times since one only needs to calculate the policy for the desired state. The downside of this method however is that one cannot regulate road condition as well as the with the MDP. Because the EAC-method is not limited by the size of the state-space one could take an infinitely large state-space, suffering no increased yearly expected average costs.

As the comparison in section 6.3 showed us, the MDP policies trump EAC policies on both expected average costs and the fraction of time spent in the unacceptable states, but it also saw that an increase in state-space causes expected average costs decline for the EAC-method making this method possibly cheaper when using a larger state-space, with trade-off a very low fraction of time spent in the unacceptable states.

APPENDIX

Table A.1: state description for the four main damage features with one state representing four years

States / Damage Feature	Cracking in (%)	Ravelling (in %)	Longitudinal Unevenness (in IRI)	Transversal Unevenness (in MM)
0	< 0	< 0	< 0	< 0
1	0 - 3.78	0 - 7.56	0 - 0.70	0 - 9.31
2	3.78 - 8.49	7.56 - 16.97	0.70 - 1.40	9.31 - 10.22
3	8.49 - 13.62	16.97 - 27.24	1.40 - 2.10	10.22 - 11.14
4	13.62 - 19.05	> 27.24	2.10 - 2.80	11.14 -13.14
5	19.05 - 24.71		2.80 - 3.50	13.14 - 18.57
6	> 24.71		> 3.50	> 18.57

Table A.2 state description for the four main damage features with one state representing two and a half years

States / Damage Feature	Cracking in (%)	Ravelling (in %)	Longitudinal Unevenness (in IRI)	Transversal Unevenness (in MM)
0	< 0	< 0	< 0	< 0
1	0 - 2.18	0 - 4.37	0 - 0.44	0 - 8.52
2	2.18 - 4.90	4.37 - 9.81	0.44 - 0.88	8.52 - 9.62
3	4.90 - 7.87	9.81 - 15.74	0.88 - 1.31	9.62 - 10.13
4	7.87 - 11.01	15.74 - 22.02	1.31 - 1.75	10.13 - 10.61
5	11.01 - 14.28	22.02 - 28.56	1.75 - 2.19	10.61 - 11.30
6	14.28 - 17.67	> 28.56	2.19 - 2.63	11.30 - 12.46
7	17.67 - 21.15		2.63 - 3.06	12.46 - 14.53
8	21.15 - 24.71		3.06 - 3.50	14.53 - 18.57
9	> 24.71		> 3.50	> 18.57

Table A.3: state description for the four main damage features with one state representing two years

States / Damage Feature	Cracking in (%)	Ravelling (in %)	Longitudinal Unevenness (in IRI)	Transversal Unevenness (in MM)
0	< 0	< 0	< 0	< 0
1	0 - 1.68	0 - 3.37	0 - 0.35	0 - 8.07
2	1.68 - 3.78	3.37 - 7.56	0.35 - 0.70	8.07 - 9.31
3	3.78 - 6.07	7.56 - 12.13	0.70 - 1.05	9.31 - 9.85
4	6.07 - 8.49	12.13 - 16.97	1.05 - 1.40	9.85 - 10.22
5	8.49 - 11.01	16.97 - 22.03	1.40 - 1.75	10.22 - 10.61
6	11.01 - 13.62	22.03 - 27.24	1.75 - 2.10	10.61 - 11.14
7	13.62 - 16.30	27.24 - 32.60	2.10 - 2.45	11.14 - 11.92
8	16.30 - 19.05	> 32.60	2.45 - 2.80	11.92 - 13.14
9	19.05 - 21.85		2.80 - 3.15	13.14 - 15.13
10	21.85 - 24.71		3.15 - 3.50	15.13 - 18.57
11	> 24.71		> 3.50	> 18.57

REFERENCES

Bertsekas D.P. & Tsitsiklis J.N. (1996). Neuro-Dynamic Programming, *Athena Scientific*, Belmont, Massachusetts.

Chua, K. H., Monismith, C. L., & Crandall, K. C. (1994). Mechanistic performance models for pavement management. In *Conference Proceeding of the Third International Conference on Managing Pavements*, Vol. 1, pp. 194-254.

IBM ILOG CPLEX Optimization Studio (64 bits) version 12.5*, 2012, computer software, IDM Corp.

Dekker, R., Nicolai, R.P. and Kallenberg, L.C.M. (2007). 'Maintenance and Markov Decision Models', in *Encyclopedia of Statistics* in Quality and Reliability, Ruggeri, F., Kennett, R. and Faltin, F.W. (eds.), John Wiley & Sons Ltd, Chichester, UK, pp. 993-1000, (2007).

Dekker R, Plasmeijer R, Swart J. (1998). Evaluation of a new maintenance concept for the preservation of highways, *IMA Journal of Mathematics applied in Business and Industry, 9: 109-156*

Golabi, K., Kulkarni, R. B., and Way, G.B. (1982). A Statewide Pavement Management System. *Interfaces* **12**, December 1982 (pp. 5-21).

Hagos, E.T. (2008) *The Effect of Aging on Binder Properties of Porous Asphalt Concrete,* Delft University of Technology, The Netherlands.

van der Horst, R., et al. (1991). Basisboek Groot Onderhoud aan Snelwegen, *Road and Hydraulic Engineering Division, Directorate General of Public Works and Water Management, Delft, The Netherlands.* (written in Dutch).

Karabakal, N., Bean, J.C. and Lohmann, J.R. (1994). *Scheduling pavement maintenance with deterministic deterioration and budget constraints Karabakal,* Technical Report 94-16, University of Michigan.

Mackie, P., Nellthorp, J. and Laird, J. (2005). 'Treatment of Maintenance', *Transport Note No. TRN-13*, Transport Economics, Policy and Poverty Thematic Group, The World Bank, January.

Marbach, P. Tstisiklis, J.N. (2001). Simulation-Based Optimization for Markov Reward Processes*, IEEE transactions on automatic control,* vol. 46, no. 2, February 2001, pp. 191-209.

Martin, T., & Kadar, P. (2012). *The scoping and development of probabilistic road deterioration (RD) models* (No. AP-T201-12).

MATLAB version 8.0.0, 2012, computer software, The MathWorks Inc., Natick, Massachusetts.

Plasmeijer, R. P. (1999). *Maintenance Optimisation Techniques for the Preservation of Highways,* Tinbergen Institute research series (No. 213).

Rietveld, P., Bruinsma, F.R. and Koetse, M.J.J. (2007). 'Infrastructure maintenance costs: A comparison of road, rail and inland navigation, and implications for user charges', in OECD (Ed.), *Estimation and Evaluation of transport costs*, pp. 49-76. Paris: OECD.

Ross, S. M. (2010). 'Markov Decision Processes', in *Introduction to Probability Models 10th Edition,* pp.265-268, Los Angeles, California.

van Roy B. (2002). Handbook of Markov Decision Processes*, International series in Operations Research & Management Science,* Volume 40, pp. 431-459.

Wang, K.C.P. and Zaniewski, J.P. (1996). 20/30 Hindsight: The New Pavement Optimisation in the Arizona State Highway Network. *Interfaces* **26,** May-June 1996 (pp. 77-89).