

PAPER • OPEN ACCESS

Pavement maintenance optimization model using Markov Decision Processes

To cite this article: P Mandiartha *et al* 2017 *J. Phys.: Conf. Ser.* **890** 012104

View the [article online](#) for updates and enhancements.

Related content

- [The Application of FAHP in Decisions of Pavement Maintenance](#)
Zhaorong Wu
- [Design of Road Pavement Using Recycled Aggregate](#)
Eva Remišová, Martin Decký, Milan Mikolaš *et al.*
- [Road pavement condition mapping and assessment using remote sensing data based on MESMA](#)
Y Pan, X Zhang, X Jin *et al.*

Pavement maintenance optimization model using Markov Decision Processes

P Mandiartha¹, C F Duffield², I S b M Razelan¹ and A b H Ismail¹

¹Faculty of Civil Engineering and Earth Resources, Universiti Malaysia Pahang, 26300 Gambang, Kuantan, Pahang, Malaysia

²Department of Infrastructure Engineering, The University of Melbourne, Melbourne, Parkville, 3010, Australia

Email: putumandiartha@ump.edu.my

Abstract. This paper presents an optimization model for selection of pavement maintenance intervention using a theory of Markov Decision Processes (MDP). There are some particular characteristics of the MDP developed in this paper which distinguish it from other similar studies or optimization models intended for pavement maintenance policy development. These unique characteristics include a direct inclusion of constraints into the formulation of MDP, the use of an average cost method of MDP, and the policy development process based on the dual linear programming solution. The limited information or discussions that are available on these matters in terms of stochastic based optimization model in road network management motivates this study. This paper uses a data set acquired from road authorities of state of Victoria, Australia, to test the model and recommends steps in the computation of MDP based stochastic optimization model, leading to the development of optimum pavement maintenance policy.

1. Introduction

The stochastic model chosen in this paper is the Markov Decision Process (MDP) in conjunction with the use of the Markov Chains as the performance prediction model. MDP, some called it as Discrete Stochastic Dynamic Programming, consists of elements: decision epoch, states $S \equiv \{1, 2, \dots, N\}$, actions $a \in A_s$, transition probabilities $p_t(\cdot/s, a)$, and rewards $r_t(s, a)$ [1, 2, 3].

The expected value at decision epoch t may be evaluated by computing:

$$r_t(s, a) = \sum_{j \in S} r_t(s, a) p_t(j/s, a), \quad (1)$$

$$\sum_{j \in S} p_t(j/s, a) = 1 \quad (2)$$

Thus, the collection of objects of

$$\{T, S, A_s, p_t(j/s, a), r_t(s, a)\} \quad (3)$$

is denoted as the MDP.

In this paper, the MDP is utilized in the process of obtaining an optimum maintenance policy by considering the performance prediction (using Markov Chains) analysis and cost functions as inputs.



The objective of the MDP is to find the optimum choice of action based on minimizing or maximizing an objective function. In this paper, the objective is to minimize the total maintenance cost. There are three types of solution method for the MDP: value iterations, policy iterations, and linear programming (LP).

1.1. MDP solutions

The solutions for MDP can be classified into two types which are an *average cost method* and a *discounted cost method* [1]. The difference lies in the process of obtaining the expected total reward. Both of the solutions can be formulated using dual LP. The disadvantage of using dual LP in the *discounted cost method* is that the choice of $\alpha(j)$ only influences the value of the objective function (Z) but not the outcome of the optimal policy [3]. In other words, the discounted cost results in only different values of the objective function without being able to obtain the difference in the decision variables. This problem can be solved when using the *average cost method*. Therefore, it is preferable to use the *average cost method* as solution method for pavement maintenance optimization since the final objective is to find maintenance action for each road condition state, represented by the decision variable.

Generally, all solutions are obtained for unconstrained MDP, meaning that there are no additional constraints associated with the formulation are added. However, there are some additional constraints that are considered important to better predict the real situation since there are many economic and political constraints influence the decision of the road authority. In the road pavement area, another requirement of a model is the capability to accommodate the strategic policy of the road authority [4, 5, 6]. At the strategic level, models should be able to cope with budget constraints and performance constraints. In general, Tijms [7] has formulated a theory of the integration of additional constraints in the dual LP of MDP solution.

1.2. Policy

Based on the result of dual LP, then the optimal policy of the optimization can be obtained. The policy R can be viewed as a rule that prescribes decision $d_i(R)$ whenever the system in state i , for each $i = 0, 1, 2, \dots, M$. Thus, R is characterized by the values $\{d_0(R), d_1(R), \dots, d_M(R)\}$. Policy R can be characterized by matrix in equation (4).

$$\begin{matrix}
 0 \\
 1 \\
 \vdots \\
 S
 \end{matrix}
 \begin{bmatrix}
 D_{01} & D_{02} & \cdots & D_{0A} \\
 D_{11} & D_{12} & \cdots & D_{1A} \\
 \cdots & \cdots & \cdots & \cdots \\
 D_{S1} & D_{S2} & \cdots & D_{SA}
 \end{bmatrix}
 \tag{4}$$

For a deterministic policy, D_{sa} is equal to 1 and 0 otherwise. While for a randomized policy D_{sa} is determined by $0 \leq D_{sa} \leq 1$ and each row sums to 1.

The decision variables (denoted here by x_{sa}) for a linear programming model are defined as follows. For each $s = 0, 1, 2, \dots, S$ and $a = 1, 2, \dots, A$, let x_{sa} be the steady-state unconditional probability that the system is in state s and decision a is made; i.e., $x_{sa} = P \{ \text{state} = s \text{ and decision} = a \}$. Each x_{sa} is closely related to the corresponding D_{sa} since, from the rules of conditional probability, as presented in equation (5).

$$x_{sa} = \pi_s D_{sa} \tag{5}$$

Where π_s is the steady-state probability that the Markov Chains is in state s , represented by equation (6).

$$\pi_s = \sum_{a=1}^A x_{sa} \tag{6}$$

So, the optimal policy D_{sa} can be obtained through equation (7).

$$D_{sa} = \frac{x_{sa}}{\pi_s} = \frac{x_{sa}}{\sum_{a=1}^A x_{sa}} \quad (7)$$

Some literatures discuss in detail of this policy formulation [1, 3, 8]. In this paper, the policy means a set of selected maintenance actions for each condition states of the road.

1.3. Recommended MDP formulation

The recommended formulation of MDP for application in pavement maintenance consists of two steps, as follows:

- Step 1. Dual LP formulation.

The method described here uses the average cost method MDP solved through dual LP with additional constraints, modelled in equations (8 – 12). Microsoft Excel's Solver program can be used for solving the dual LP with constraints. The objective function is to minimize the total cost.

The dual LP model is to choose x_{sa} to:

$$\text{Minimize: } Z = \sum_{s=0}^S \sum_{a=1}^A C_{sa} x_{sa} \quad (8)$$

Subject to:

$$(i) \quad \sum_{s=0}^S \sum_{a=1}^A x_{sa} = 1 \quad (9)$$

$$(ii) \quad \sum_{a=1}^A x_{ja} - \sum_{s=0}^S \sum_{a=1}^A x_{sa} p(j/s, a) = 0 \quad ; \text{ for } j = 0, 1, 2, \dots, S \quad (10)$$

$$(iii) \quad x_{sa} \geq 0 \quad ; \text{ for } s = 0, 1, 2, \dots, S \text{ and } a = 1, 2, \dots, A \quad (11)$$

$$(iv) \quad \sum_{s \in S} \sum_{a \in A(s)} \alpha_{sa}^{(t)} x_{sa}(\pi) \leq \beta^{(t)} \quad (\text{The additional constraints}) \quad (12)$$

In comparison, the policy obtain using pure solution (MDP without additional constraints) is also presented in this paper to depict the effect of additional constraint.

- Step 2. Policy development.

A policy is determined by using equation (7). The policy R can be characterized by the matrix in equation (4).

2. Illustration of the computation of MDP for pavement maintenance optimization

2.1. Inputs of MDP optimization for road network pavement maintenance

In this presented paper, the types of flexible pavement analysed in chip seal surfaced flexible pavement roads with high volume road (AADT > 3000 vehicles/day). In Australia, where the data is obtained, chip seal is used as surfacing especially for rural road.

2.1.1. Actions and state space

There are four actions (i.e. actions A-D) as a combination of routine (denoted as **R**), periodic (**P**) and rehabilitation (**Rehab**) maintenance (table 1). Meanwhile, the road conditions are divided into 11 condition states (states 1- 11). The objective of the optimization is to minimize the average total cost subjected to some additional constraints which are imposed directly into the optimization problem, and then the problem is solved via the dual LP program.

2.1.2. Transition probabilities Matrix (TPM) as output of Markov Chains analysis

The main determinant of Markov Chains is the transition probability. As a consequence of choosing an action in certain state *i*, there are probabilities that the state condition remain in the same state or move to next state *j*. This is denoted as a transition probability. The following figure 1 presents the TPM for high volume chip seal roads. The detail of development process of these TPMs (Markov Chains analysis) can be found in [9].

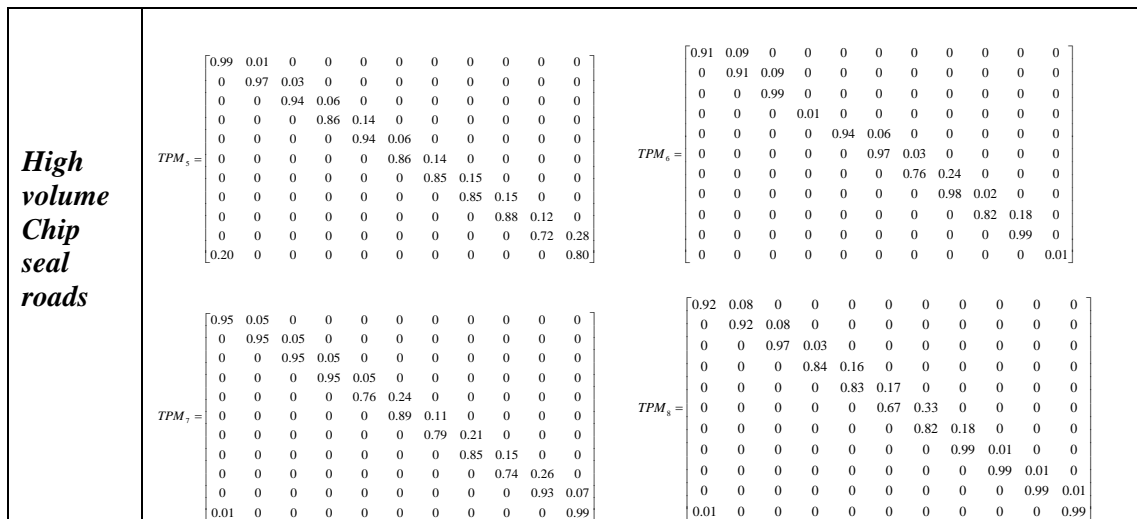


Figure 1. TPM for high volume chip seal road.

2.1.3. Costs function

The cost function is developed by analysing the historical data or could be based on the engineering judgement. As consequence of selection certain action in a particular road condition state, the cost is incurred. The costs are varied depending on type action and condition state. The typical cost function inputs developed in this paper analysis are presented table 1.

Table 1. Chip Seal high volume roads cost (\$/km-lane).

Action	Action A (R + No P + No Rehab)	Action B (R + P + No Rehab)	Action C (R + No P + Rehab)	Action D (R + P + Rehab)
State 1	1422.17	7674.38	3737.15	8496.50
State 2	1422.17	7674.38	3737.15	8496.50
State 3	1541.79	3508.74	3737.15	6435.46
State 4	1733.35	4632.09	4054.62	8397.03
State 5	1374.86	4088.84	3965.81	7130.37
State 6	1249.19	6043.85	3519.32	7010.16
State 7	1255.91	7669.21	4922.35	8498.54
State 8	1253.62	20721.82	3247.80	11464.76
State 9	1249.19	6580.87	11116.18	13911.13
State 10	1292.99	21777.39	23123.33	21777.39
State 11	1272.95	51370.56	22985.81	51370.56

2.1.4 Additional constraint MDP formulation

An additional constraint is imposed onto the dual LP formulation of MDP namely performance/proportion constraint α_{sa} (i.e. performance constraint at a state, s , with an action a).

Another constraint which is budget constraint $\beta^{(t)}$ (budget constraint at time t) can also be introduced but not discussed in this paper. Scenarios were developed in this analysis which are pure solution and scenario 1 (performance constraint). The resulting maintenance policies were then compared between scenarios.

2.2. Output: maintenance policy

The final objective of the MDP optimization is to find the optimal policies R based on scenarios developed. The resulting policy in this case refers to the optimal maintenance action for each road condition state based on a scenario selected by the decision maker. Based on the Solver results, the policies are determined for high volume chip seal roads by using equation (7). These policies are the main output of the MDP program. An important matter to note is that these policies are subjected to costs which mean that when implemented the associated costs must also be utilized.

The overall results is presented in Table 2. It is noted that there is no linear correlation between the actions and states. This is explainable since the optimization model and the associated objective function is based on cost minimization. In the worst state, all of the results suggest to use only action A, which means that no matter action chosen, the roads will still move to next worst condition state. The suitable action beyond state 11 is the reconstruction, however this action is not considered in the model.

2.3. Policy results for scenarios (pure solution and scenario 1)

For high volume chip seal roads, the policies obtained for the scenarios are as follow (see table 2):

- The pure solution suggests using action A (routine maintenance only) in most condition states except in state 1. In state 1, the policy suggests choosing action B (note: actually all actions are the same in state 1 in terms of cost due to limited data available in this conditions state).
- In scenario 1, since performance constraints are added, higher type of policy is obtained. In state 1, the policy suggests using action A and action B with probability of 0.2 and 0.8, respectively. For state 3, actions A and action B are also suggested with the probability of 0.19 and 0.81, respectively. In state 5, the scenario 1 suggests choosing only action D. Similarly, in state 9, only action C is suggested. Action A is implemented in other states than those.

Table 2. Maintenance policy obtained.

State	IRI	Chip Seal Rods							
		Pure solution				Scenario 1			
		Weight	Action 1	Weight	Action 2	Weight	Action 1	Weight	Action 2
State 1	< 1.3	1.00	B	-	-	0.20	A	0.80	B
State 2	1.4 – 1.6	1.00	A	-	-	1.00	A	-	-
State 3	1.7 - 1.9	1.00	A	-	-	0.19	A	0.81	B
State 4	2.0 – 2.2	1.00	A	-	-	1.00	A	-	-
State 5	2.3 – 2.5	1.00	A	-	-	1.00	D	-	-
State 6	2.6 – 2.8	1.00	A	-	-	1.00	A	-	-
State 7	2.9 – 3.1	1.00	A	-	-	1.00	A	-	-
State 8	3.2 – 3.4	1.00	A	-	-	1.00	A	-	-
State 9	3.5 – 3.7	1.00	A	-	-	1.00	C	-	-
State 10	3.8 – 4.0	1.00	A	-	-	1.00	A	-	-
State 11	4.1 <	1.00	A	-	-	1.00	A	-	-

3. Conclusion and recommendations

This paper presents the steps and calculation processes of the use of MDP for road network pavement maintenance optimization model. This method uses an average cost MDP solved through dual LP as a minimization problem. The required inputs for the computation of MDP have also been presented which form very important aspects in the model. An additional constraint has also been introduced to reflect constraints faced by the road authorities in managing road network. The final output of the MDP optimization in the form of maintenance policy has also been presented along with its explanation and interpretation. A maintenance policy in some condition states may consist of combination of maintenance actions in the form of probability. This will provide justification for the road authorities to prioritize maintenance on section that requires higher types maintenance which consequently at higher cost.

The most challenging, if not the most difficult, phase of works when using MDP for pavement maintenance optimization is when developing the inputs, consisting of Markov Chain's transition probabilities matrix and cost of actions. Ideally, both inputs have to be related to each other based on the actual data. The effect of a specific maintenance action has to be recorded along with its cost. However, significant amount of historical data is required which often not available, especially with regard to historical maintenance costs. Otherwise, all inputs will rely heavily on the engineering judgement. The development of performance prediction model of roads after specific maintenance action in this case using Markov Chains with costs recorded and the computation of MDP in order to obtain optimal policy are two separate works. The MDP results are highly dependence on the quality of the inputs.

Acknowledgments

The authors would like to thank RDU 160363 (UMP Pekan smart traffic research project) of University Malaysia Pahang for providing funding for the paper publication.

4. References:

- [1] Puterman M L 1994 *Markov Decision Processes: Discrete Dynamic Programming* (New York: John Wiley & Sons, Inc.)
- [2] Winston W L 1994 *Operations Research: Application and Algorithms. Belmont* (California: Wadsworth, Inc.)
- [3] Hillier F S and G J Lieberman 2001 *Introduction to Operation Research* (New York: The McGraw Hill Company)
- [4] Saraswatula S R and S N Ammirkhanian 1992. *Pavement Management System - State of The Art* ed Holt F B and W L Gramling (Philadelphia: ASTM)
- [5] Dekker R 1996 Application of Maintenance Optimization Models: a Review and Analysis *Reliability Engineering and System Safety* vol 51 pp 229-240
- [6] Worm J M and A v Harten 1996 Model Based Decision Support for Planning of Road Maintenance *Reliability Engineering and System Safety* vol 51 pp 305-316
- [7] Tijms H C 1994 *Stochastic Models: An Algorithmic Approach* (Chichester, UK: John Wiley & Sons Ltd)
- [8] Derman C 1970 *Finite State Markovian Decision Process* (New York: Academic Press)
- [9] Mandiartha P, Duffield C F, Thompson R G and M R. Wigan 2017 Measuring pavement maintenance effectiveness using Markov Chains analysis, *Structure and Infrastructure Engineering* vol 13 (London: Taylor & Francis) pp. 844-854