Model based decision support for planning of road maintenance

J. M. Worm & A. van Harten
University Of Twente, School Of Management Studies, P.O. Box 217, 7500 AE Enschede, The Netherlands

In this article we describe a Decision Support Model, based on Operational Research methods, for the multi-period planning of maintenance of bituminous pavements. This model is a tool for the road manager to assist in generating an optimal maintenance plan for a road. Optimal means: minimising the Net Present Value of maintenance costs, while the plan is acceptable in terms of technical admissibility, resulting quality, etc. Global restrictions such as budget restrictions can also be imposed.

Adequate grouping of maintenance activities in view of quantity discounts is an important aspect of our model. Our approach is to reduce the complexity of the optimisation by hierarchical structuring in four levels. In the lowest two levels maintenance per lane sector is considered, first with an unbounded planning horizon and next with a bounded planning horizon and time-windows for maintenance. The grouping of maintenance activities for a specific road is the topic of the third level. At the fourth level, which we will not consider in this article, the problem of optimal assignment of the available maintenance budgets over a set of roads or road sections takes place. Here, some results are presented to demonstrate the effects of grouping and to show that this hierarchical approach gives rise to improvements compared with previous work. © 1996 Elsevier Science Limited.

1 INTRODUCTION

The necessity to make optimal use of available budgets for maintenance of roads has become a hot topic in the last few decades. Increasing pressure to carry out good planning of the maintenance of the Dutch road network, whose size amounts to 10,000 kilometres (highways and regional main roads), has been felt. This situation has been caused by fundamental shifts in economic factors, societal values and norms and governmental policy. Analogous effects are reported for road maintenance in other countries (see Golabi19).

The problem of optimisation of multi-period planning of the maintenance of bituminous pavements will be attacked by Operational Research methods. An optimal maintenance plan for a road (or a road section) is produced using an optimisation model in interaction with the planner. This maintenance plan is based on statistical and/or expert knowledge of the deterioration of the road. A choice is made between various possible maintenance actions, taking into account their effect on the future road conditions and their associated costs. Such a tool proposes to the planner answers to questions like: when should one take what maintenance action on which part of the road?

Specific to our approach is the fact that our final instrument is kept as flexible as possible. The key elements of our optimisation approach are:

1. the state dependent choice of actions with minimisation of the Net Present Value (NPV) of all present and future maintenance costs as the objective;
2. the best possible grouping of actions representing one project in view of 'economies of scale'. Unique to the optimisation model described in this article is the search for the possibility of cost reduction of maintenance actions by exploiting quantity discounts. This means that performing the same action on a larger connected part of a road is, relatively, cheaper per square metre;
3. optimisation within budget restrictions and/or global quality requirements on the future road conditions.

In all existing literature we know of, at least one of the above mentioned aspects is neglected. Nevertheless, comparison with our approach is useful. For a
general introduction to applications of O.R. in a maintenance context we refer to Refs 2, 28, 30 and 32.

Let us start with a description of IVON. For a certain planning period this system generates an optimal maintenance plan, comprising for every lane sector of a road an optimal action plus year of execution. Furthermore, this system yields estimates of expenditures for maintaining bituminous pavements over a number of years. This system concentrates on the aspects (1) and (2). With respect to aspect (1) another line of approach is used: ‘Minimisation of the so called Equivalent Annual Costs’ (see also Haas[1]). These EAC are a sort of measure of effectiveness: the action’s costs divided by the gain in lifetime as a result of the action. This criterion differs considerably from minimising the NPV of present plus future costs. This results in clear differences in the planning.

In Zwagemakers[38] we find a model which calculates an optimal maintenance strategy, for one specific lane sector, by minimising the NPV of maintenance costs over a planning period of infinite length. This minimisation uses Markov decision theory. A maintenance strategy is associated with each state a sector can be in. The necessary maintenance budgets per sector follow easily from this model. This work deals especially with aspect (1). In our model this work will be used as a building block and is extended with respect to aspects (2) and (3).

In international literature on road maintenance, aspect (3) gets a lot of attention. Often one uses assignment problems to allocate a budget within given restrictions while satisfying quality demands.

Hugo[16] uses a linear programming model to minimise the present value of the total costs of a number of possible rehabilitation projects. Based on a forecasting model a cost estimation is made and priorities are set for these projects.

In Ref. 3, Cook deals with the capital budgeting problem in pavement maintenance in two phases. In phase 1, a financial model is used as a means to determine appropriate budget levels, while fulfilling some global quality constraints. Phase 2 is used for optimal allocation of the available budget in such a way that desired service levels are achieved. In this phase the assignment of periodical regional budgets is tuned with the budgets asked for.

In Lytton[25] a set of mathematical models and computer programs is described, which can be used as tools for planning road maintenance and rehabilitation. The goal of the total system is the selection of projects so that the expected benefit is maximised. In this maximisation, constraints like the available resources of manpower, budget, materials and equipment are taken into account. Benefits are expressed in terms of reliability and traffic served. An analogous approach is found in Karan[18].

Compared with these references, our work results in a higher level of integration of technical and economic aspects of maintenance. Much more in line with our work is Golabi et al.[10] There the aspect of meeting global quality standards is very important (see also Refs 23 and 36). Their pavement management system is recommending the best maintenance action for every mile of the 7,400-mile network of highways. An important goal of their system is to yield defendable one-year and five-year budgets and to show the effect of budget cuts on road conditions. In their work, ‘economies of scale’ plays a minimal role.

Aspect (1) is stressed in Keane[20]. In this case a Markov model is used for describing the deterioration of the road. He works with a T-year time horizon and minimal quality standards, which are allowed to be violated with some prescribed probability. After a certain maintenance action has been carried out, the deterioration process can be based on a different transition matrix. This is done with certain pavement classification matrices. Using a dynamic programming approach a maintenance policy is derived, which minimises the expected discounted total costs over the planning horizon. In our work we prefer to use the Markov model from Ref. 38, since there is a better fit with the Dutch databases.

For more information on road maintenance we also refer to Refs 5, 8, 19, 22, 24, 27 and 29. Integration of the aspects (1), (2) and (3) appears to be specific for our work.

This paper deals with some of the main models constituting a prototype DSS for road maintenance. The DSS also includes a model dealing with budget restrictions. This model will be the subject of a later publication, cf. Ref. 34. The emphasis of this paper will be on the effects of ‘economies of scale’. Note that grouping of maintenance on adjacent lane sectors in the longitudinal direction leads to less traffic queuing, because of a relatively shorter throughput time than in the case of several smaller projects. From a governmental point of view, this can be treated as a quantity discount. In our instrument it is possible to take these secondary effects into account, cf. Ref. 1.

Mathematically, our optimisation is based on a four level hierarchical approach. Such a hierarchical approach is not common practice in maintenance planning, while in our view it has a tremendous potential scope. The various levels are related to well-recognisable steps in the planning process. We will come to this later. The mathematics of optimisation in levels 1, 2 and 3, varies in each level: from Markov decision theory, to dynamic programming, to a shortest route problem. Level 4 uses a different technique again, cf. Ref. 34.

In Section 3 we give a more precise formulation of the optimisation problem. Further we discuss the Operational Research techniques in greater detail and we present some results obtained by our instrument. It
should be mentioned that our instrument still has the status of a prototype. First we start with a discussion of the characteristics of road maintenance.

2 THE CHARACTERISTICS OF ROAD MAINTENANCE AND THE MAINTENANCE STRATEGY

By definition a maintenance plan consists of: a set of related activities directed to restoring or keeping a system (or object) in a desired condition, so that it will function within specified quality norms, compare Refs 6, 14 and 21.

Let us analyze this definition in more depth. To start with, there is the question of the structure of the system: what components have to be distinguished and what conditions are important in view of failure or malfunctioning. In the case of roads, physical deterioration processes will eventually lead to failure or malfunctioning of the system. The asphalt quality decreases monotonically until it passes a minimal acceptable level. This leads immediately to the question of describing the deterioration process with sufficient accuracy and also the norm for acceptable asphalt quality has to be established. Next, based on the deterioration behaviour, a decision has to be made as to what way the system is going to be maintained.

We shall now have a closer look at:

(a) the structure of the system,
(b) how to handle quality norms,
(c) what maintenance strategy is to be used,
(d) which philosophy should one take in dealing with the many degrees of freedom.

Consider a system of roads. For maintenance purposes, we put the following structure on the system. As elementary components we take traffic lane sectors and road sectors of 100 metres length, see Fig. 1. Further, higher aggregation levels will play a role in the planning: sections made up of a few hundred lane sectors or road sectors. Characteristic for this situation is, in the first place, the partition of the total system (a road) into a large number of parallel elementary subsystems (the lane sectors), with adjacency relations in longitudinal and transversal directions. The deterioration process of the lane sectors will be treated as uncorrelated. Even so, these adjacency relations play an important part in the planning of maintenance. One reason is that only certain combinations of maintenance actions are allowed on adjacent lane sectors in the transversal direction. An action can be applicable 'roadway' wide or on separate lanes. Separate lane actions can be planned on adjacent lane sectors without any problems. In the case of roadway wide actions, only actions belonging to the same so called top-layer group can be planned on adjacent sectors. Furthermore, by combining maintenance on adjacent sectors a cost reduction per square metre can be realised due to quantity discounts.

A second characteristic for the situation is that the size of the total system of roads is so large that for the planning a hierarchical ordering into sublevels between the total and the elementary level is necessary.

The state of the road surface is described in terms of appearing structural damage looking at features like crazing, ravelling, etc., in comparison with the ideal situation in case of a new bituminous pavement. We will come back to this later. The degree of damage per feature can be established through visual inspection and measurements. These inspections and measurements are performed on a year to year basis. For more information on inspection and measurements, we refer to Refs 4, 12 and 22. Quality standards are derived from the lowest acceptable value for these features. In a deterministic setting these standards are dealt with as being absolute. This can be seen as a strive for 'zero-defects'. In stochastic situations the maximal probability for excess is prescribed. In our approach of this maintenance problem we are dealing with structural damage only.

In the literature, various types of maintenance strategies suitable for various types of maintenance problems are distinguished (see Refs 6 and 9), in particular, the classification corrective vs preventive maintenance is often used.

Using a corrective maintenance strategy means that: failure or malfunctioning of the system gives rise to the initiation of maintenance. In the case of structural road surface deterioration, any violation of a quality norm can be considered as a failure. Surpassing these norms will lead to unacceptable safety risks for the road users. In this sense our maintenance strategy is a corrective one, because maintenance is primarily triggered by failures. Our strategy is state dependent, because the evolution of structural damage can be reasonably predicted and at the same time we have at hand actual information about the state of the road.
On each lane sector the initiating event for maintenance can be of two kinds:

1. end of technical lifetime of the asphalt,
2. economic depreciation of the road surface before its technical lifetime is over, which can only occur in the case of combined maintenance projects.

The keypoint of our approach is the design of a highly flexible planning procedure. This will lead to a higher degree of complexity of the planning procedure, but the underlying philosophy is that the gain in efficiency compensates for this effect.

We are dealing with an enormous freedom for maintenance planning. A normal planning problem comprises typically a 5–10 year period and deals with 15–20 possible actions per lane sector, from which the best choice has to be made. Considering a road section with about 100 sectors, it will be clear that we are facing an astronomical amount of possibilities for the planning of maintenance. There is some reduction in this amount due to the technical restrictions in transversal direction, but at the same time this makes the problem more complex. We can conclude here that solving the resulting optimisation problem is far from trivial.

As is known from systems theory, effective control requires several conditions to be satisfied: clear goal specification, an adequate model of the (evolution of the) system, enough information about the initial state of the system and sufficient options for control. Information and the variety of options have been discussed above. We shall now describe in more detail the objective of control and the evolution model.

The Net Present Value (NPV) of all maintenance costs (present and future) is the performance criterion we shall use. In a formula the NPV of a series of costs $C(t)$ made in years $t=1,2,3,...$ is: $NPV = C(1) + \alpha C(2) + \alpha^2 C(3) + ...$. Here $\alpha$ denotes the discount factor, $0 < \alpha < 1$. In fact the summation runs up to infinity, i.e. all future costs, also those after the planning period are taken into account. Hence the restvalue after the planning period plays a role. In this way we are avoiding maintenance plans generating little costs during the planning period, by postponing maintenance and associated costs until the planning period is over.

For the deterioration of the road surface, various models are used in the literature. The process of deterioration can be modelled deterministically or stochastically. In the stochastic case one either works on the basis of complete insight in the underlying probability density function or on the basis of complete uncertainty. In principle the model in our tool is of the Markov type. Hence we are dealing with a stochastic process of deterioration with known transition probabilities between various states, cf. Refs 15, 31 and 35. For the sake of simplicity, we restrict ourselves in this paper to a discussion of the deterministic case.

The deterioration process is dependent on variables such as the type of soil on which the road is founded, the traffic density and the type of pavement used (see Ref. 10). This forms the basis for distinguishing different road categories. Differences in deterioration due to climatic conditions are not of great importance in our situation.

The quality of a lane or road sector is described in terms of the amount of damage. Therefore a number of different damage aspects are distinguished, see Table 1 taken from Refs 4 and 22.

In total, damage is represented by a state vector, where each component of the vector corresponds with one of the characteristics. The value assigned to a characteristic damage variable has the interpretation of the most probable remaining lifetime for the feature in question. Note that this remaining lifetime indicates the length of time before the specified quality norm is reached for this specific feature only. This judgement is a combination of theoretical insight, practical experience and expert opinions. In Fig. 2 an example is given to show the relation between physical quality evolution and remaining lifetime.

The main advantage of representing structural damage by a vector of remaining lifetimes $i=(i_1,i_2,\ldots,i_F)$ is that the evolution becomes particularly simple. Note that $F$ stands for the number of features used in the state description and $i_f$ stands for the remaining lifetime of characteristic $f$, in case a lane sector has state $i$. If no maintenance action occurs in year 1 then $i$ evolves into $i_{next}$ with

$$i_{next} = i - (1,1)$$

(1)

The evolution of the state of the surface of the road depends also on the costs and quality effects of the maintenance actions at hand. The effect of a maintenance action is modelled as an improvement of the remaining lifetime-vector. In general actions will

<table>
<thead>
<tr>
<th>Table 1. A survey of groups of damage (see C.R.O.W.*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group of damage</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Texture</td>
</tr>
<tr>
<td>Evenness</td>
</tr>
<tr>
<td>Soundness</td>
</tr>
<tr>
<td>Marginal Strip</td>
</tr>
<tr>
<td>Miscellaneous</td>
</tr>
<tr>
<td>Remark: Features of damage marked with a (*) can be examined by detailed inspection or measurements.</td>
</tr>
</tbody>
</table>
have an improving effect on more than one feature. An example is given in Table 2 (cf. Ref. 38).

If maintenance action \( m \) is executed in year 1 then the effect is given by:

\[
i_{\text{next}} = E(i;m)
\]

where \( E \) is the probability function that transforms state \( i \) into \( i_{\text{next}} \) by executing \( m \). Denoting no maintenance by \( m = 0 \), (1) is a special case of (2). This deterioration model is the basis of the optimisation described in the next section.

Using our instrument, the road manager is offered a logical way and more possibilities than optimisation by hand. Let us now discuss the role of the road manager briefly. To start with he or she gets a role in defining the road sections to be considered in the optimisation, while the planning for other sections can be kept fixed. Another possibility is that he or she can exclude certain maintenance actions for certain types of lane sectors, e.g. because they have such a limited effect that in the short run a new action will be necessary. Also the road manager can ask to show the effects of postponement of a project to a next year in order to avoid excesses in technical depreciation of the road surface. These requirements will be incorporated in the concept admissible. Further the user can strongly influence every step in the optimisation process. A positive view of the user’s expertise is basic to our approach. The user can specify several important parameters like: the discount rate for calculating the NPV, costs and effects of maintenance actions, the minimal desired increase in lifetime that a maintenance action should achieve. Of course this has to be done in consultancy with higher management.

In the next chapter we start with a short overview of the total optimisation problem and then give some details of the structure of the optimisation module.

### 3 DESCRIPTION OF THE OPTIMISATION TECHNIQUE AND SOME RESULTS

The optimisation problem has the form:

\[
\text{Minimise } \sum_{r=1}^{R} \sum_{t=1}^{T} \alpha^{-1} \alpha' TC(MP(t, r))
\]

\[
t = 1, 2, \ldots, \quad r = 1, \ldots, R
\]

under the quality restrictions

\[
ST(t, r) \geq 0, \text{ for all } t \text{ and } r,
\]

and the global restrictions

\[
\sum_{r=1}^{R} TC(MP(t, r)) \leq B(t), \text{ for all } t
\]

In the above formulation \( ST(t, r) \) refers to the state of road \( r \) at the beginning of period \( t \), which is the combination of all lane sector states. \( MP(t, r) \) is a maintenance plan, which can be executed on state \( ST(t, r) \). The admissibility of \( MP(t, r) \) follows partly from (4), being the primary guarantee that necessary maintenance is carried out. In addition certain

<table>
<thead>
<tr>
<th>Action</th>
<th>Effect of action (in years)</th>
<th>Pricequilders/m²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ravelling</td>
<td>crazing</td>
</tr>
<tr>
<td>micro overlay</td>
<td>8</td>
<td>( t-1 )</td>
</tr>
<tr>
<td>overlay (cold)</td>
<td>8</td>
<td>( t-1 )</td>
</tr>
<tr>
<td>sanimat</td>
<td>10</td>
<td>( t+1 )</td>
</tr>
<tr>
<td>repave</td>
<td>12</td>
<td>( t+2 )</td>
</tr>
<tr>
<td>overlay</td>
<td>14</td>
<td>( t+5 )</td>
</tr>
<tr>
<td>milling + overlay</td>
<td>14</td>
<td>( t+11 )</td>
</tr>
<tr>
<td>Maximal remaining lifetime</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

The effect of an action being 14 means that in the next year the remaining lifetime is 14 years. An effect \( t+11 \) means that in the next year the remaining lifetime is \( t+11 \) years, in case the remaining lifetime is \( t \) years now. Note that this remaining lifetime has an upper bound as given in the last row.
In principle, a hierarchical approach will lead to a partial loss in flexibility of the planning. In the case where the hierarchical structure is properly linked to the technical and management concepts, this loss will usually be practically no problem. On the contrary, the more transparent optimisation structure is an advantage in terms of acceptability and workability. It enhances the understanding of the models by the user and at the same time gives the user an increased possibility of controlling the models. Hierarchical structuring of the original complex problem will often lead to an approximate solution for the original problem of higher quality than a direct approximation scheme applied to the full problem. The possibility of choosing a better optimisation technique for every subproblem can dominate this partial loss in flexibility completely. In our opinion this is the case in our problem.

Let us now have a closer look at the multi-period maintenance planning. The total optimisation problem is structured in the following way:

(A) one lane sector with infinite planning horizon;
(B) one lane sector with a finite planning horizon and with a given set of periods in which maintenance is allowed. By considering combinations of time-windows we are already anticipating the possibility that maintenance actions cannot be executed completely freely in time due to constraints set at a higher level. For every road (or road section) suitable sets of combinations of time-windows will be generated. A combination of time-windows $T_1 < T_2 < ... < T_D$ will be denoted by $W$. $T$ denotes the planning horizon and $D$ is the number of decision periods included in the set $W$:
(C) a road or road section consisting of $L$ road sectors on which maintenance is allowed only during a given combination of time-windows $W$. In this step there is a search for the best combination of maintenance actions in view of quantity discounts. The result will be a maintenance plan for this road or road section for each generated combination of time-windows;
(D) a road network existing of $R$ (sections of) roads each with various combinations of time-windows. Now a selection is made from the generated alternatives in such a way that global management demands with regards to quality and budgeting are fulfilled.

In the case where the fit between level (D) and the lower levels (A), (B) and (C) is not good enough it is impossible to satisfy the global managerial demands. In that case a new iteration is necessary. The user has possibilities as splitting up a road or road section into smaller sections, while keeping the segmentation of lane sectors intact. He can look at new combinations of time-windows for some roads or road sections, or have a discussion about relaxation of the global management demands. Another possibility is that the user allows for violation of the quality standards. This corresponds with introduction of an artificial action ‘violation of quality standards’ with certain penalty costs.

Before we continue with a discussion of the different levels of decision-making as distinguished, we want to mention that the multi-period planning of maintenance generated with our instrument has to be
updated every year, when new data become available. Hence, it is used as part of a rolling plan approach.

At decision level (A) we are solving a maintenance problem for one lane sector. The model is of Markov type with an unbounded planning horizon. Such problems have been extensively studied, cf. Refs 35, 38. Define $S$ to be the number of possible states a lane sector can be in and $A_i$ to be the set of admissible actions for asphalt in state $i$. The action ‘no maintenance’ is also treated as a possible maintenance action. Further, define $M+1$ as the number of possible maintenance actions and $C_i(m)$ as the costs per square metre associated with action $m$ ($m = 0,\ldots,M$) for a lane sector. The evolution is described by (2) and because we are only considering the deterministic case, it is useful to introduce

\[ p_{ij}(m) = 1 \text{ if } j = E(i,m) \\
= 0 \text{ in the other cases} \]

Let $V_i^*$ be the Net Present Value of the minimal maintenance costs starting with an initial state $i$, for a given discount factor $\alpha$. The optimal maintenance strategy can then be found from the following optimality equations:

\[ V_i^* = \text{Minimum}_{m \in A_i} \left[ C_i(m) + \alpha \sum_{j=1}^{S} p_{ij}(m) V_j^* \right] \]

These equations can be transformed to a linear programming problem and then they can be solved using an LP solver, cf. Refs 35 and 38.

In Fig. 3 an example is given of the output from level (A). As input we have: $S = 3600$, $F = 3$, $M = 6$ and $\alpha = 95\%$. For the maximal possible remaining lifetime for the used features, the action costs and effects we refer to Table 2. The computational time needed to solve this problem, when run on an Univac-1100, amounted to 841.38 seconds. The computing times will increase with increasing $F$ and $M$. In practical situations values $F \leq 4$ and $M \leq 15$ are sufficient and computing times for such values are not prohibitive, since real-time responses are not required. In Fig. 3 the optimal actions as function of the initial state are shown. The remaining life time of pavelling is 0 years. The $x$-axis refers to depth of ruts and the $y$-axis to crazing. The optimal action in a state $(0,x,y)$ is found at the point $(x,y)$ in the figure. The different shadings represent different actions. This figure shows that the best action is strongly state dependent. We assume that the deterioration mechanisms are the same for all lane segments, hence only one Markov model needs to be solved. Note that it is not correct to conclude from this figure: that if none remaining lifetime equals 0 years, then the no maintenance action is the optimal one.

At this level we need from the user information about: (1) the available actions with associated effects and costs, (2) the discount factor to be used and (3) the features of damage to be used for describing the state of the asphalt and per feature the maximal possible remaining lifetime.

At the following level (B) we are dealing with a planning horizon $T$ and a combination of time-windows $W$. The optimal present value of the costs after the planning horizon depends on the final state. They are calculated using step (A). It is of course possible to use at this place a restvalue calculated in another way. To calculate the NPV associated with the initial condition of the pavement we use backward dynamic programming (see Ref. 37).

Define $h_i(t)$ as the NPV at the beginning of period $t$ of the minimal costs (per square metre) over the periods $t$ up to and including $T$, when the lane sector is in state $i$ at the beginning of period $t$. Define $h_i(T+1) = V_i^*$; the restvalue found at level (A). Let $W = \{T_1, T_2, \ldots, T_P\}$ be the set of decision periods with $T_1 < T_2 < \ldots < T_P$ then we have for $1 \leq t \leq T$ the following formula:

\[ h_i(t) = \text{Minimum}_{m \in A_i(t)} \left[ C_i(m) + \alpha \sum_{j=1}^{S} p_{ij}(m) h_j(t+1) \right] \]

Here $A_i(t) = \{0\}$ if $t$ is not contained in $W$, where 0 denotes the ‘no maintenance’ action. In more general situations $A_i(t)$ can be chosen such that certain quality constraints, e.g. a certain minimum quality level for the pavement at the end of the time horizon, are taken care of.

At this level the same information as for level (A) is needed. At this place the user can also express his
opinion about what decision periods and what global quality constraints should be used. It is also possible to introduce at this place time dependent costs of maintenance actions.

A database containing the results of level (B) for all initial states and a database containing the state every lane sector is in at the beginning of period 1, are the input for the crucial level (C). At this level the optimisation over all allowed combinations in longitudinal and transversal direction are investigated. This is done by decoupling the grouping problem for a certain decision period from the grouping problem in later periods. Starting with the first decision period, a maintenance plan for that period is generated. Next, given the resulting expected state of every lane sector at the beginning of the second decision period, a maintenance plan for this period is generated, etc. It is possible to delay an identified maintenance project to a later decision period. Delaying a project is only allowed if the constraints are not violated. For a very general class of quantity discounts for combined maintenance on connected lane parts, the problem then reduces to a route problem in a multi-stage network. This O.R.-problem is well-known and it can be solved by very fast algorithms (see Refs 26 and 35).

This shortest route problem is created and solved for every decision period. Fig. 4 shows the transformation of a maintenance grouping problem into a shortest route problem for a one-lane road. In this figure an example is given for only a small part of a road, with \( L = 3 \) sectors. In general we can create \( \frac{1}{2}L(L+1) \) possible project, arcs in the network.

Let \( PR(x,y) \) denote group maintenance of road sectors \( x \) up to and including \( y \) (\( y > x \)). Then \( PR(x,y) \) is represented in the network by an arc starting at point \( x-1 \) and ending at \( y \), with a length \( CST_{x-1,y} \). An arc of this type is only included in the network if, for every lane part marked by the project \( PR(x,y) \), at least one common cost-reduction group can be identified, i.e. if for every lane sector of this lane part an allowable action in that reduction group can be found. In the case where actions belonging to the same so-called cost-reduction group are planned on adjacent lane sectors, they are treated as being the same for getting quantity discounts. This structure of the ‘economies of scale’ reduces the possibilities for relevant grouping of actions considerably, since only combined actions on connected road sectors are relevant.

Maintenance actions to be planned on a road containing more than one lane have to satisfy constraints in the transversal direction of the road. The actions should belong to the same top-layer group. Actions, including the ‘no maintenance’ action (\( m = 0 \)), which are not roadway wide are treated as belonging to the same top-layer group. This group is given number 1. From Table 3 one can directly see that the actions 0, 1, 2, 3, 4 and 9 are not roadway wide.

A possible project consists of actions which should belong to the same top-layer group, when looking in transversal direction, and per lane part, marked by this project, the actions should belong to the same reduction group. In the case where a road consists of two lanes, the user can choose from two possibilities to create a project on a road part:

1. Start with a lane part and determine the best project as for a one-lane road, which we will come to later. Note that with the best project, the project with minimal NPV is meant. Now the best reduction group and top-layer group for this lane part are known. For the adjacent lane part there is only freedom to choose the reduction group, because the top-layer group has now already been fixed.
2. Choose for every lane part a cost-reduction group. Next, given this combination of cost-reduction groups and the constraints in transversal direction, the best maintenance action for every lane sector is decided upon. After storage

---

**Fig. 4.** The transformation of candidate maintenance projects into a network representation.
of this project, this process is repeated for another combination of reduction groups. In case all possible combinations have been looked at, the best project out of this set of projects will then be chosen. Note that only combinations of reduction groups have to be considered, where at least one common top-layer group within both groups can be identified. From Table 3 one can see directly that, for example, it is possible to consider reduction group 6 on one lane and reduction group 7 on the other one, because the actions belong to the same top-layer group.

In our applications we usually prefer the second option, since in general it allows for a larger set of projects.

\( CST_{x-1,y} \) is defined as the minimal NPV of present and future maintenance costs of project \( PR(x,y) \) comprising road sectors \( x \) up to and including \( y \). In case (2) these future maintenance costs are calculated per road sector by using a weighted state for this sector. This weighted state is calculated using the states that the lane sectors belonging to that road sector, would probably be in, in the next decision period. Remember we are using the output of the first two levels as an input to this level. Because of computational limitations at levels (A) and (B), the technical constraints in transversal direction cannot be taken into account. Cost fluctuations in the NPV, calculated from the beginning of the first decision period, when going from one to the next decision period can be partly compensated by using this weighted state. Of course the user has to decide if a weighted state is used and how this weighted state is calculated.

Let us now consider the form of the quantity discount function. As a reference we take the discount coefficient of action \( m \) and \( z=\delta \). Applied to a surface \( \delta \text{SAR} \), with \( \delta \approx 1 \), then the costs are \( \delta \text{SAR} \cdot C_2(m) \cdot (1 - d(m) + d(m) \cdot \tau) \cdot \delta^{-1} \), with \( d(m) \) the discount coefficient of action \( m \) and \( \tau = 2,500/\text{SAR} \) (cf. Ref. 17). The user should also state if quantity discounts can be realised roadway wide or only per lane. In case the reduction group is not the same for every lanepart, cost reduction can only be realised per lanepart. If a \( q \% \) cost reduction can be realised by executing maintenance actions belonging to the same reduction group, to which action \( m \) also belongs, on a surface of 2,500 m\(^2\), then we have \( (100 - q)/100 \cdot C_i(m) = C_2(m) \). Hence \( (100 - q)/100 \cdot \delta \cdot \text{SAR} = 1 - d(m) + d(m) \cdot \tau \). If we denote by \( F(m,\delta) \) the costs for maintaining a surface \( \delta \text{SAR} \), then:

\[
\text{CST}_{x-1,y} = \text{Minimum admissible } m \left[ F(m, y - x + 1) + \alpha \cdot \delta \text{SAR} \cdot \sum_{z=x}^{y} \sum_{j=1}^{\delta} [p_{j,z}(m) \cdot h_z(2)] \right]
\]  

The first part of the formula is referring to the action costs with quantity discounts, to be made in period 1, in the case where action \( m \) is chosen. Note that the action costs are not dependent upon the beginning state of a lane sector. The second part is the formulation of the expected value of the NPV of future maintenance costs, calculated from the beginning of period 2, which is in this case the next period. As before, \( j \) is related to the initial condition \( i \) by (2) and for calculating the future costs \( h_z(2) \) we use the results from level (B). This formula for calculating \( \text{CST}_{x-1,y} \) is based on simple reduction groups with one action per group. An analogous formula holds for roadway wide maintenance, etc. In fact it is also used for generating a maintenance plan for a next decision period, with the only difference that use is made of an expected condition \( i \).

Finally, after creating all possible arcs with their respective lengths, the shortest route problem is solved. For computational reasons it can be necessary to limit the length of the road we are looking at. Now we have at hand the costs of maintaining the road or road section for the decision period we are in. Also, the maintenance projects are now known. If an arc starting in \( x_1 - 1 \) and ending in \( x_2 \) is in the optimal route, then we know that the maintenance of road sectors \( x_1 \) up to and including \( x_2 \) constitutes one

<table>
<thead>
<tr>
<th>Action number</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top-layer group</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Cost-reduction group</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>( C_2(m) )</td>
<td>0.00</td>
<td>14.02</td>
<td>8.50</td>
<td>2.13</td>
<td>2.33</td>
<td>6.37</td>
<td>7.22</td>
<td>21.67</td>
<td>7.22</td>
<td>12.75</td>
<td>13.60</td>
<td>19.12</td>
</tr>
</tbody>
</table>

Table 3. Action data
project. For every lane sector comprised by this project we then know the planned maintenance action.

In Table 4 the effect of quantity discounts on the maintenance plan is shown. The plan without quantity discounts is compared with the plan with quantity discounts. In both cases the constraints for adjacent lane sectors are taken into account and a discount factor $\alpha = 95\%$ has been used. To generate the results in Table 4 a realistic data set of a small road section with 2 road lanes of 1 kilometre length, i.e. 10 road sectors, has been used. The surface of one lane sector: SAR = 350m². To describe the state of a lane sector we used 4 characteristics. Further, the number of possible states is: $S = 30,184$, the planning period is $T = 5$ years and there is a choice between $M + 1 = 12$ maintenance actions. The quantity discounts are calculated as described above. For the sake of simplicity we assumed every maintenance action $m$ to belong to a different reduction group (see Table 3) and we took $d(m) = 0.028$ for every action. This value of $d(m)$ corresponds to a $\pm 15\%$ cost reduction for the 'same' maintenance action applied to 7 adjacent lane sectors with a total surface of 2,450m², compared with just one sector. For a certain decision period, the calculation of the restvalue per road sector was based on the mean state the lane sector states would be in, in the next decision period, in case the maintenance project we are looking at is executed.

Note that in Table 4 several actions occur. Action 0 is 'no maintenance', the actions 2 and 4 are applicable on one lane and the actions 6, 7, 8, 10 and 11 are roadway wide. The effect of quantity discounts on the planned maintenance actions is clear. It enhances more clustering.

Comparing the costs we observe that initially the costs are somewhat higher, in case quantity discounts can be realised. This can be explained by the fact that in case one or more neighbouring lane sectors are needing maintenance, this can result in accelerated maintenance on other neighbouring sectors which in a few years would also be in need of comparable maintenance. If, in the case of no quantity discounts, an action differing from the 'no maintenance' action is

\begin{table}[h]
\centering
\begin{tabular}{c|cccc}
\hline
Road sector & Planned actions in period 1 for a given set of initial conditions & & & \\
& Without quantity discounts & With quantity discounts & & \\
\hline
1 & 11 & 10 & 11 & 10 \\
2 & 7 & 6 & 11 & 10 \\
3 & 10 & 10 & 10 & 10 \\
4 & 2 & 2 & 2 & 2 \\
5 & 2 & 4 & 2 & 4 \\
6 & 8 & 8 & 8 & 8 \\
7 & 10 & 10 & 10 & 10 \\
8 & 0 & 0 & 4 & 4 \\
9 & 0 & 0 & 4 & 4 \\
10 & 4 & 0 & 6 & 6 \\
& lane 1 & lane 2 & lane 1 & lane 2 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{c|ccc}
\hline
Without quantity discounts & With quantity discounts & Present value (in year 1) & Present value (in year 1) \\
Year & & 65,882 & 69,575 & 69,954 \\
1 & & 2,750 & 3,334 & 3,024 \\
2 & & 6,270 & 11,286 & 9,749 \\
3 & & 9,675 & 3,334 & 2,743 \\
4 & & 0 & 84,577 & 86,380 \\
5 & & & & \\
Sum & & 231,351 & 221,634 & 308,014 \\
\hline
\end{tabular}
\end{table}

The effect of the planned actions on quality
- Sum of remaining lifetimes at the beginning of year 1: 469 years.
- Sum of remaining lifetimes at the end of the planning period in case quantity discounts are not allowed: 482 years.
- Sum of remaining lifetimes at the end of the planning period in case quantity discounts are allowed: 552 years.
planned on a certain lane sector, it is possible that another action in case of quantity discounts is chosen. In the Table an example is found for road sector 10. In case of no quantity discounts the combination: executing action 4 on lane sector (10,1) and action 0 on sector (10,2) is the solution, minimising the sum of the direct maintenance costs plus the restvalue. The combination: executing action 6 on both sectors is in this case more expensive. It is the other way around in case quantity discounts can be realised. Looking at the total costs we notice a 2.5% reduction in case of quantity discounts. Scaling up to the Dutch road network this is definitely a significant reduction. But there is more. In fact we should not only look at the costs but also look at the effect of the planned actions on the quality. In case of quantity discounts the sum of the remaining lifetimes is a good indication for the achieved mean level of quality. It is considerably higher in case of quantity discounts.

The user is given the possibility to demand that at most one maintenance action, different from the 'no maintenance' action, per lane sector during the planning horizon is allowed. In this case decoupling the grouping problem in a certain decision period from another decision period is not taking place. The shortest route problem is solved only once and the solution gives a maintenance plan for the whole planning period.

Finally the problem to satisfy the global constraints remains. It is considered in level (D). This leads to an integer programming problem which can be solved using Lagrangian relaxation (see van de Velde) and a Branch and Bound technique. This will be the subject of a later publication, cf. Ref. 34.

4 SUMMARY AND CONCLUSIONS

In this article, a short overview of existing models for road maintenance planning was given. Our approach to solving the maintenance problem, is different from those found in literature, since three aspects are clearly integrated:

1. minimisation of the NPV of all present and future maintenance costs,
2. the grouping of actions into one project in view of 'economies of scale' and
3. optimisation within budget restrictions and/or global quality requirements.

Solving the maintenance problem is done by a four-phased optimisation approach. For each phase a formulation in terms of an associated O.R.-model can be given. Given the initial state of a road or road section, various financial factors, technical constraints and some other user dependent input, an optimal maintenance plan over a certain planning period is generated. Some of the results calculated with the developed computer programs are presented. The results concern mainly the effects of 'economies of scale'. The following conclusions can be drawn:

1. 'Economies of scale' will lead to adjusted combinations of maintenance actions: it enhances more clustering.
2. By grouping of maintenance in view of realising quantity discounts, a total costs reduction of a few per cents can be realised. It is possible however that the initial costs are (perhaps) somewhat higher.
3. 'Economies of scale' has a positive effect on mean quality.

ACKNOWLEDGEMENT

We want to express our thanks to ICIM (Information Centre for Infrastructure and Environment) in Rijswijk (The Netherlands) for their help and financial support.

REFERENCES