Title
Pavement systems reconstruction and resurfacing policies for minimization of life-cycle costs under greenhouse gas emissions constraints

Permalink
https://escholarship.org/uc/item/8jz5k6xq

Authors
Lee, J
Madanat, S
Reger, D

Publication Date
2016-11-01

DOI
10.1016/j.trb.2016.08.016

Peer reviewed
Pavement systems reconstruction and resurfacing policies for minimization of life-cycle costs under greenhouse gas emissions constraints

Jinwoo Lee, Samer Madanat, Darren Reger

Abstract

Pavement management systems, designed to minimize total lifecycle costs, will need to evolve to meet the needs of the future. Environmental concerns are likely to add an additional consideration for the state DOTs when allocating their financial resources. Transportation agencies will be concerned with determining maintenance, resurfacing and reconstruction policies for pavement segments in their systems while also addressing the environmental impact of these activities.

In this paper, we propose an efficient solution to solve for pavement resurfacing and reconstruction policies that minimize societal (agency and user) costs under a Greenhouse Gas (GHG) emissions constraint. The main methodological contribution of this work relative to the state of the art is that we formulate the problem to include multi-dimensional pavement segment states and heterogeneous management activities. It allows for a more realistic representation of the majority of current pavements in the world. For example, the assumption that pavements are perpetual, i.e., do not need reconstruction during their lifetime, can be relaxed. A case study using California roads is performed; we find that, for that specific group of pavement segments, the optimal policies to minimize societal costs do not vary greatly from the policies that minimize GHG emissions. An agency can use these results to determine what GHG emission budgets are feasible for the highway system that it manages.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Managing pavement assets is an important responsibility for state departments of transportation (DOTs). This stems from the fact that poor pavement condition has direct effects on road users, increasing their vehicle wear and tear as well as their fuel consumption. Many DOTs currently use pavement management systems (PMS) to determine when to maintain, rehabilitate or reconstruct its pavements. Commonly used PMS, which seek to minimize total costs (user and agency), may soon face additional constraints such as reducing environmental impacts, such as greenhouse gas (GHG) emissions. Such constraints may be driven by a state-wide goal (e.g. AB32 in California) or potentially as a necessary part of a carbon budget or cap and trade system. It has been shown that effective management of highway assets can bring a marked reduction in GHG emissions (Santero et al., 2011; Santero et al., 2011a, b).

* Corresponding author.
E-mail address: jinwoo.lee@nyu.edu (J. Lee).

http://dx.doi.org/10.1016/j.trb.2016.08.016
0191-2615/© 2016 Elsevier Ltd. All rights reserved.
Pavement management with multiple objectives is a topic that has received increasing attention in the transportation research community. Multi-criteria optimization in Pavement Management Systems (PMS) has been used to incorporate heterogeneous objectives simultaneously (Fwa et al., 2000; Mbwana, 2001; Li and Sinha, 2004; Bai et al., 2015).

Recent research has addressed the problem of minimizing GHG emissions as one of multiple objectives in pavement management. Gosse et al. (2013) considered three objectives: total agency costs; GHG emissions; and system performance, i.e. average of pavement surface condition, but did not capture the interdependencies between the three objectives. They adopted Genetic Algorithm as a solution method to solve an unconstrained tri-objective problem. Lidicker et al. (2012) studied the problem for a single pavement segment with two objectives: minimize discounted life cycle costs and undiscounted GHG emissions over an infinite time horizon. In their work, the relationship between pavement surface condition and users’ GHG emissions was captured. GHG emissions were directly related to vehicles’ fuel efficiency, and the influence of pavement surface condition on fuel efficiency was accounted for, using results from the literature (Watanatada et al., 1987; National Research Council, 2006; Evans et al., 2009; Zaabar and Chatti, 2010).

Other approaches have been used, where GHG emissions are minimized as a single objective, or used as a constraint in total cost minimization problems. Zhang et al. (2010, 2013) solved the segment-level and network-level problems to minimize GHG emissions or total costs. Reger et al. (2014) extended the work of Lidicker et al. (2012) to the system-level problem, minimizing the total life cycle costs under a GHG budget constraint. They used the bottom-up solution methodologies developed by Sathaye and Madanat (2011 and 2012). Reger et al. (2015) also solved a different problem, where the objective is to minimize GHG emissions subject to an agency financial budget constraint.

One important limitation in the work of Lidicker et al. (2012) and Reger et al. (2014, 2015) is that it considered only one type of maintenance, namely resurfacing of varying thicknesses. Therefore, their work is limited to the case of “perpetual pavements”, which are pavement that will not need to be reconstructed in the future.

The two potential interventions considered in this paper are reconstruction and resurfacing, although the methods are generalizable to any treatment for which models for condition and rate of deterioration exist. Resurfacing is the most common type of activity performed by an agency. It consists of removing the top layer of the wearing course and replacing it with new asphalt concrete. This results in improvements in the condition of the pavement, but is not as effective when the damage has reached the underlying layers. In this case, the entire pavement structure may be replaced, an activity known as reconstruction. Reconstruction improves the pavement condition back to its original condition state. Reconstruction is more resource-consuming than resurfacing: it takes considerably longer to perform and comes with much larger expenses and environmental impacts.

In the literature, the problem of joint resurfacing and reconstruction of highway pavements has usually been formulated as follows: for a given pavement section, an agency must decide 1) how often to reconstruct 2) how many resurfacings to perform between reconstructions and 3) the timing and intensity of those resurfacings. See for example Rashid and Tsunokawa (2012) and Lee and Madanat, (2014a,b).

In this paper, we develop and demonstrate a methodology to solve for an optimal set of policies for managing pavement system under an environmental constraint, where both reconstruction and resurfacing are available alternatives. The paper is organized as follows: Section 2 addresses the problem formulation and proposes a bottom-up solution methodology. A case study and conclusions are presented in Section 3 and Section 4 respectively.

2. Problem formulation and solution methodology

2.1. Problem formulation

The scope of the problem is a system of pavement segments that are managed over an infinite time horizon. This problem is formulated as a discrete time Markov Decision Processes (MDP) with finite actions sets, and with multi-dimensional pavement segment states. The objective, shown in (1a), is to minimize the discounted total cost-to-go for the system, which is the sum of the user costs and the agency costs for a system comprised of N segments that are numbered n = 1, ..., N. The control, x, is the pavement management policy including various interventions with a range of effects and costs, and X is the set of all possible x.

The state in year t, S(t), is an augmented multi-dimensional vector consisting of the set of all segment-level states, S_n(t). The segment-specific attributes are categorized into two groups: (i) attributes that influence the deterioration process of the segment, and are not affected by management interventions (e.g., annual traffic load on the segment, precipitation and temperature, etc.); and (ii) attributes affected by management interventions (e.g., structural condition, serviceability, history-dependent variables such as time from last performed interventions, i.e., age). In the remainder of the paper, S_n(t) will include two elements: condition and age, where age is the time since the last reconstruction. Both groups of attributes determine the deterioration process of a segment (the evolution of S_n(t) over time).

The undiscounted average annual GHG emissions along the planning time horizon are constrained, as shown in (1b). The value of emissions is not discounted because it is unknown whether emissions in the future will be more or less important than in the current time; this is consistent with the literature (Sedjo and Marland, 2003).

\[
J^*(S(0)) = \min_{x \in X} J(S(0), x) = \min_{x \in \{x_1, ..., x_k\}} \sum_{n=1}^{N} J_n(S_n(0), x_n)
\] (1a)

\[
\text{such that } \sum_{n=1}^{N} \rho_n J_n(S_n(0), x_n) \leq \sum_{n=1}^{N} \rho_n \gamma_n S_n(0, x_n) \leq \gamma \sum_{n=1}^{N} \rho_n \gamma_n J_n(S_n(0), x_n)
\] (1b)
s.t.  
\[ Q(S(0), x) = \sum_{n=1}^{N} Q_n(S_n(0), x_n) \leq B \]  

(1b)

where  
\[ S(t): \text{state in year } t, S(t) = \{S_1(t), \ldots, S_N(t)\}, S(t) \in \mathcal{S}, S_n(t) \in \mathcal{S}_n; \]
\[ J(S(0), x): \text{Discounted annual average cost-to-go associated with } S(0) \text{ and a management policy, } x = \{x_1, \ldots, x_N\}; \]
\[ J_n(S_n(0), x_n): \text{Discounted annual average cost-to-go for segment } n \text{ associated with segment-level initial state, } S_n(0), \text{ and management policy, } x_n; \]
\[ Q(S(0), x): \text{Undiscounted annual average GHG emission (metric tons/yr) under policy } x; \]
\[ Q_n(S_n(0), x_n): \text{Undiscounted annual average GHG emission for segment } n \text{ under policy } x_n; \]
\[ B: \text{Annual GHG emission limit for the system.} \]

More details on the cost and emission functions in (1a) and (1b) are presented in Section 3.1.

2.2. Solution methodology

We propose a bottom-up solution methodology, which allows us to account for the segment-specific features in the system-level optimization. This section describes both steps in the bottom-up methodology: the system-level solution and the segment-level solution.

2.2.1. System-level solution method

We define the Lagrangian dual of the optimization problem (1) as \( D(\cdot) \). The optimal \( D(\cdot) \) of the Lagrangian dual is:

\[ D^*(S(0)) = \sup_{\Lambda} \left\{ \min_x J(S(0), x) + \Lambda \left[ Q(S(0), x) - B \right] \right\} \]  

(2)

\( \Lambda \) is the non-negative Lagrangian multiplier. According to the concept of weak duality, the optimal solution, \( \Lambda^* \), of the Lagrangian dual (2) is a lower bound on \( J(S(0)) \) of (1a). \( D(S(0)) \) is not guaranteed to be differentiable with respect to \( \Lambda \), so it may be impossible to find the optimal solution analytically. However, it is possible to find \( \Lambda^* \) of (2) numerically. To find the numerical solution, we refer to the Lagrangian function as:

\[ D(S(0)|\Lambda) = \min_x J(S(0), x) + \Lambda \left[ Q(S(0), x) - B \right] \]  

(3)

We introduce a new function \( V(S(0), x|\Lambda) \) that is sum of \( J(S(0), x) \) and \( \Lambda \cdot Q(S(0), x) \), expressed as:

\[ V(S(0), x|\Lambda) = \sum_{n=1}^{N} V_n(S_n(0), x_n|\Lambda) = \sum_{n=1}^{N} [J_n(S_n(0), x_n) + \Lambda \cdot Q_n(S_n(0), x_n)]. \]  

(4)

To simplify the problem, we assume that the management strategies are independent across segments and that the costs and emissions for each segment are independent from the management strategies of other segments.

**Assumption 1.** For each \( n \), let \( n' \in \{1, \ldots, N\}/\{n\} \). The segment-level control, \( x_n \), segment-level discounted costs, \( J_n(S_n(0), x_n) \), and segment-level emissions, \( Q_n(S_n(0), x_n) \), are independent of \( x_{n'} \).

We do not constrain the problem to preserve connectivity (i.e., making sure that when a given segment is closed for reconstruction, the links that provide an alternate route are not also under construction). We also do not consider the potential economies of scale from rehabilitating adjacent facilities, such as the reduction in the fixed costs from equipment mobilization. These two assumptions, taken together, mean that we ignore network interdependencies in our optimization. As shown in Medury and Madanat (2013), accounting for these interdependencies across links complicates the problem significantly. In any case, our assumptions are realistic if the construction durations are relatively short in comparison to the whole lifecycle length, and there are not multiple closures in the network.

Based on Assumption 1, we can derive that, for a given \( \Lambda \) value, the optimal \( x \) in Eq. (3) is the set of segment-level optimal solutions \( x_n \) in Eq. (5), which is a weighted bi-objective optimization problem for all \( n \).

\[ \arg \min_x J(S(0), x) + \Lambda \left[ Q(S(0), x) - B \right] = \left\{ x_n = \arg \min_{x_n} V_n(S_n(0), x_n|\Lambda), \forall n \right\} \]  

(5)

According to the above separable property, the optimization problem (3) is decomposed into \( N \) segment-level problems, and the optimal solution for each segment-level problem (5) is included in the optimal solution of (3). Afterword, the optimization problem (2) can be solved by numerical methods based on the optimal result of (3).
In the minimization of \( V(S(0), x|\Lambda) \) in (4), the resulting curve of \((Q_2(S(0)|\Lambda), J_2(S(0)|\Lambda)) \) where \( Q^*(S(0)|\Lambda) \in [Q^*(S(0)|\Lambda) = 0] \) is equivalent to the Pareto Frontier of the multi-objective minimization of \( f(S(0), x) \) and \( Q(S(0), x) \).

If \( \Lambda = 0 \) in (3), the problem is equivalent to the problem of minimizing the discounted lifetime cost only. Here, the optimal solution is denoted by \( x^*_{\Lambda=0} \). We define the lower bound to \( J(S(0), x) \) as \( J(S(0), x^*_{\Lambda=0}) \). This lower bound (to \( J(S(0), x) \)) coincides with the upper bound to \( Q(S(0), x) \), which we define as \( Q(S(0), x^*_{\Lambda=0}) \) or \( Q^*(S(0)) \).

If \( \lim \Lambda \rightarrow \infty \) in (3), the problem is equivalent to the problem of minimizing the annual average emission only. The optimal solution is denoted by \( x^*_{\Lambda=\infty} \). We define the lower bound to \( Q(S(0), x) \) as \( Q(S(0), x^*_{\Lambda=\infty}) \) or \( Q^*(S(0)) \). This lower bound coincides with the upper bound to \( J(S(0), x) \), which we define as \( J(S(0), x^*_{\Lambda=\infty}) \) or \( f^*(S(0)) \).

The optimal solution of (2) is denoted by \( \Lambda^0 \) and \( x^0 \). If \( \Lambda^0 \) and \( x^0 \) satisfy \( \Lambda^0 Q(S(0), x^0) = B \), \( x^0 \) is the optimal solution of (1) denoted by \( x^* \). There are four possible cases, of which three are trivial:

(i) If \( B < Q^*(S(0)) \), the problem is infeasible.
(ii) If \( B = Q^*(S(0)) \), the optimal solution of (1) is \( x^*_{\Lambda=\infty} \).
(iii) If \( B > Q^*(S(0)) \), the optimal solution of (1) is \( x^*_{\Lambda=\infty} \).
(iv) If \( Q(S(0)) < B < Q^*(S(0)) \), the solution is not trivial.

In case (iv), finding an exact solution can be NP-hard, so efficient numerical methods should be applied to solve the problem with a lower computational complexity. As an example, based on the results of a finite number of \( \Lambda \), it is possible to estimate the approximate form of \( J(S(0), x^*) \) with respect to \( B \). If the approximate form has a convex shape, a subgradient method can be used, which is an iterative method commonly used to solve convex minimization problems. Another possible example is a bisection method that has been applied in the infrastructure management literature (e.g., Hu et al., 2015). In the first step of the bisection method, an initial range of \( \Lambda \) is chosen, from 0 to a certain value where \( Q(S(0)|\Lambda) \) is less than \( B \), and \( \Lambda \) is updated iteratively.

The following section describes the solution methodology for the segment-level weighted bi-objective optimization problem in (5).

2.2.2. Solution methodology for the segment-level problem

To solve the system-level problem formulated in (1), we have shown that it is necessary to solve the lower-level problem for each segment as shown in (5). The objective is to minimize the weighted sum of the discounted total cost-to-go for the segment \( n, J_n(S_n(0), x_n) \), annualized by the factor \((1 - e^{-r})\), as shown in Eq. (6), and the annual average GHG emissions, \( Q_n(S_n(0), x_n) \), defined in (7). These two terms have different units: one is the monetary value, and the other is the GHG emissions.

The segment-level policy is denoted by \( x_n \), and is stationary (i.e. \( x_n = \{x_n(S_n) \mid \forall S_n \in S_n \} \)). That is, this policy is applicable to the steady state problem, for an infinite planning horizon.

\[
J_n(S_n(0), x_n) = \lim_{T \to \infty} (1 - e^{-r}) \cdot E \left\{ \sum_{t=0}^{T} [U_n(S_n(t), x_n(S_n(t)))] + M_n(S_n(t), x_n(S_n(t)))]e^{-rt} \right\}
\]

\[
Q_n(S_n(0), x_n) = \lim_{T \to \infty} \frac{1}{T} E \left\{ \sum_{t=0}^{T} [W_n(S_n(t), x_n(S_n(t)))] + A_n(S_n(t), x_n(S_n(t)))] \right\}
\]

Where:

\( S_n(t) \): segment state in year \( t \), \( S_n = \{s_1, s_2, \ldots, s_n(t)\} \);
\( U_n(S_n(t), x_n(S_n(t))) \): User costs given segment state in year \( t \) under policy \( x_n(S_n(t)) \);
\( M_n(S_n(t), x_n(S_n(t))) \): Sum of agency costs and user delay costs due to interventions in year \( t \) under policy \( x_n(S_n(t)) \);
\( r \): Discount rate, a positive scalar, \( 0 < r \leq 1 \);
\( W_n(S_n(t), x_n(S_n(t))) \): User GHG emissions given segment statement in year \( t \) under policy \( x_n(S_n(t)) \);
\( A_n(S_n(t), x_n(S_n(t))) \): Sum of agency GHG emissions and user delay emissions due to interventions in year \( t \) under policy \( x_n(S_n(t)) \).

Note that the segment-level discounted total costs in (6) depend on the current state \( S_n(0) \). However, the annual average GHG emissions are independent of \( S_n(0) \). This independence is explained next.

If there exists a management policy, \( x_n \), such that \( S_n \) eventually reaches \( S_{\text{new}} \) where \( S_{\text{new}} \) stands for the state after reconstruction, i.e.

\[
\exists x_n, \text{ s.t. } (K_n(S_n, S_{\text{new}}|x_n)) < \infty
\]

where \( K_n(S_n, S_{\text{new}}|x_n) \) is the elapsed time to reach \( S_{\text{new}} \) starting from \( S_n \) under policy \( x_n \), the following result holds (Bertsekas, 2011):
\[
\lim_{T \to \infty} \frac{1}{T} E \left\{ \sum_{t=0}^{T} \left[ W_n(S_n(t), x_n(S_n(t))) + A_n(S_n(t), x_n(S_n(t))) \right] \right\} \\
= \lim_{T \to \infty} \frac{1}{T} E \left\{ \sum_{t=0}^{T} \left[ W_n(S_n(t), x_n(S_n(t))) + A_n(S_n(t), x_n(S_n(t))) \right] \right\}
\]

(9)

Most pavements are not designed to have perpetual lives, and thus reconstruction is inevitable. Therefore, \( E[K_n(S_n, S_{\text{new}}(x_n))] \leq H_{\text{max}} < \infty \) for all \( S_n \) under any realistic \( x_n \), where \( H_{\text{max}} \) refers to the maximum allowable length of a pavement’s lifecycle (defined as the number of years between reconstructions).

Therefore, the average GHG emission, \( Q_n'(S_n(0), x_n) \), is equal to \( Q_n(S_{\text{new}}, x_n) \), which means that the average annual emissions over the planning horizon \( Q_n'(S_n(0), x_n) \) is independent of the current condition \( S_n(0) \) for all \( S_n(0) \in \mathbb{S} \). We define \( Q_n'(x_n) \) as the annual average GHG emission of the pavement life cycle, which is independent of \( S_n \), and we have \( Q_n(S_n, x_n) = Q_n'(x_n), \forall S_n \in \mathbb{S} \).

The segment-level optimization problem in (5) can be converted into the recursive Eq. (10), where \( x_n(S_n) \) is a set of all maintenance and reconstruction activities applicable to segment \( n \) if current state is \( S_n \). At most one intervention is carried out on each segment in a year, and we assume that it is applied at the start point of the year. \( S_n^* \) is the state at the end of the year which is influenced by the deterioration process and \( x_n(S_n) \).

\[
V_n^*(S_n|\Lambda) = J_n^*(S_n|\Lambda) + \Lambda \cdot Q_n^*(S_n|\Lambda) \\
= \lim_{T \to \infty} \min_{x_n(S_n)} E \left\{ \left(1 - e^{-\rho T} \right) \cdot \left[ U_n(S_n, x_n(S_n)) + M_n(S_n, x_n(S_n)) \right] + e^{-\rho T} J_n^*(S_n'|\Lambda) + \right\} + \Lambda \cdot \left(1 - e^{-\rho T} \right) \cdot \left[ W_n(S_n, x_n(S_n)) + A_n(S_n, x_n(S_n)) \right] + \frac{T-1}{T} \cdot Q_n^*(S_n'|\Lambda) \right\}
\]

(10)

This problem is a mixed-discounted-and-undiscounted Markov decision model. There is a rich literature on how to solve DP problems for either discounted or undiscounted problem, but problems that contain both discounted and undiscounted costs have not been solved in the literature. A mixed-discounted problem in infinite time horizon with positive discount rates of value less than one has a stationary optimal solution (Feinberg and Schwartz, 1994). Therefore, the above problem can be solved by replacing the undiscounted term by a discounted term with arbitrary discount rate \( \rho, 0 < \rho < r \).

\[
V_n^*(S_n|\Lambda, \rho) = J_n^*(S_n|\Lambda, \rho) + \Lambda \cdot Q_n^*(S_n|\Lambda, \rho) \\
= \min_{x_n(S_n)} E \left\{ \left(1 - e^{-\rho T} \right) \cdot \left[ U_n(S_n, x_n(S_n)) + M_n(S_n, x_n(S_n)) \right] + e^{-\rho T} J_n^*(S_n'|\Lambda, \rho) + \right\} + \Lambda \cdot \left(1 - e^{-\rho T} \right) \cdot \left[ W_n(S_n, x_n(S_n)) + A_n(S_n, x_n(S_n)) \right] + e^{-\rho T} Q_n^*(S_n'|\Lambda, \rho) \right\}
\]

(11)

In (11), \( Q_n'(S_n(0), x_n|\rho) \) is the optimal solution of \( Q_n'(S_n(0), x_n|\rho) \) for given \( \rho \). Here, \( Q_n'(S_n(0), x_n|\rho) \) is the approximate function of \( Q_n(S_n(0), x_n|\rho) \) in (7), which is defined in (12).

\[
Q_n'(S_n(0), x_n|\rho) = \lim_{T \to \infty} \left(1 - e^{-\rho T}\right) \cdot E \left\{ \sum_{t=0}^{T} \left[ W_n(S_n(t), x_n(S_n(t))) + A_n(S_n(t), x_n(S_n(t))) \right] e^{-\rho t} \right\}
\]

(12)

Note that, because of the approximation factor \( \rho \), \( Q_n'(S_n(0), x_n|\rho) \) is not equal to \( Q_n(S_{\text{new}}, x_n|\rho) \). However, it is two-sided bounded as in (13) for a given \( \rho \) value, and if \( \lim \rho \to 0^+ \), \( Q_n'(S_n(0), x_n|\rho) \) would hold if emissions occurred only at the end of each lifecycle (an extreme underestimation). The lower bound is calculated as the sum of discounted emissions during a cycle; this is the sum of an infinite geometric series where the start term is \( e^{-\rho H_{\text{max}}} \cdot H_{\text{max}} \cdot Q_n(x_n) \) and the common ratio is \( e^{-\rho H_{\text{max}}} \), multiplied by the annualizing factor, \( (1 - e^{-\rho}) \). The inequality in (15) would hold if the total emissions for each lifecycle occurred at the start of each lifecycle (an extreme overestimation). The upper bound is also calculated in a similar way where the start term is \( H_{\text{max}} \cdot Q_n(x_n) \) and the common ratio is \( e^{-\rho H_{\text{max}}} \).

\[
\left(1 - e^{-\rho}\right) \cdot e^{-\rho H_{\text{max}}} \cdot \frac{Q_n(x_n)}{1 - e^{-\rho H_{\text{max}}}} \leq Q_n'(S_n(0), x_n|\rho)
\]

(14)
\[
Q_i^t(S_{\text{new}}, x_n|\rho) \leq (1 - e^{-\rho}) \cdot \frac{H_{\text{max}} \bar{Q}_n(x_n)}{1 - e^{-\rho H_{\text{max}}}} \tag{15}
\]

From (13), (14) and (15), conservative error bounds for the ratio between \(Q_i^t(S_n(0), x_n|\rho)\) and \(\bar{Q}_n(x_n)\) can be obtained for a given \(\rho\) as shown in (16). It can be seen that \(Q_i^t(S_n(0), x_n|\rho) = \bar{Q}_n(x_n)\) if \(\lim \rho \to 0^+\). This proves that the introduction of the approximation factor should not affect the accuracy of the results as long as \(\rho\) is arbitrarily small.

\[
\frac{(1 - e^{-\rho}) \cdot e^{-\rho H_{\text{max}}} H_{\text{max}} \bar{Q}_n(x_n)}{1 - e^{-\rho H_{\text{max}}}} \leq Q_i^t(S_n(0), x_n|\rho)
\leq A_{\rho}^n \cdot (1 - e^{-\rho}) \cdot \frac{H_{\text{max}} \bar{Q}_n(x_n)}{1 - e^{-\rho H_{\text{max}}}} \tag{16}
\]

**Solution algorithm (value iteration).** The \(i\)th value iteration step for (11) is shown in (17). The algorithm terminates if the solution converges.

\[
TV^i + 1(S_n|\Lambda, \rho) = T_J^{i+1}(S_n|\Lambda, \rho) + \Lambda \cdot TV^{i+1}(S_n|\Lambda, \rho)
= \min_{x_n(S_n)} \left\{ \frac{(1 - e^{-\rho})}{\Lambda} \left\{ U_n(S_n, x_n(S_n)) + M_a(S_n, x_n(S_n)) \right\} + e^{-\rho} T_J^i(S_n|\Lambda, \rho) \right\} \tag{17}
\]

where

- \(i\): iteration index;
- \(TV_n^i(S_n|\Lambda, \rho)\): optimal objective value for the \(i\)th iteration, with state \(S_n\) for given \(\Lambda\) and \(\rho\);
- \(T_J_n^i(S_n|\Lambda, \rho)\): optimal cost for the \(i\)th iteration, with state \(S_n\) for given \(\Lambda\) and \(\rho\);
- \(TV_n^{i+1}(S_n|\Lambda, \rho)\): optimal emission for the \(i\)th iteration, with state \(S_n\) for given \(\Lambda\) and \(\rho\).

The entire segment-level Dynamic Programming algorithm follows:

**Algorithm**

0. Set \(i = 0\);
1. Set \(T_J_n^0(S_n|\Lambda, \rho)\) and \(TV_n^0(S_n|\Lambda, \rho)\) for all \(S_n \in S_n\);
2. Set \(i = 1\);
3. Compute \(TV_n^i(S_n|\Lambda, \rho), x_n^i, T_J_n^i(S_n|\Lambda, \rho)\) and \(TV_n^{i+1}(S_n|\Lambda, \rho)\) for all \(S_n \in S_n\).
4. If a termination condition is satisfied, then stop;
5. \(i \leftarrow i + 1\), and go to step 3.

The approximate policy \(x_n^i|\Lambda, \rho(S_n)\) is defined as (18). The optimal undiscounted emission value \(\bar{Q}_n(x_n^i|\Lambda, \rho)\) can be computed to calculate \(D(S(0), \Lambda)\) (see Appendix A).

\[
x_n^i|\Lambda, \rho(S_n) = \arg \min_{x_n(S_n)} \left\{ J_n(S_n, x_n(S_n)) + \Lambda \cdot Q_i^t(S_n(0), x_n|\rho) \right\} \tag{18}
\]

### 3. Case study

We consider an AC pavement highway system consisting of 50 pavement segments from Caltrans (California Department of Transportation) District 4, where each segment is one-directional 1 km long. The system includes rural and urban highways, and has a wide range of attribute values such as number of lanes, structural number, traffic volumes for all vehicles and trucks (AADT and AADTT) and traffic loading (ESALs), etc. (Caltrans 2014). The overall complexity of the solution method is proportional to the number of segments, so the size of \(N\) does not matter computationally. The solution algorithms are programmed in MATLAB, and the optimization problems are solved on Mac OS with a 2.5 GHz processor and 16GB RAM.

We start by specifying the component models included in the objective function and constraints.

#### 3.1. Deterioration, cost and emission models

Two management interventions are available: resurfacing with full overlay thickness and reconstruction.

The structural design of a pavement influences the deterioration process and the reconstruction costs. The stationary segment-level decision variables for state \(S_n\) are denoted by \(x_n(S_n) \in \{\text{do-nothing}, \text{resurfacing}, \text{reconstruction}\}\). We assume ‘mill-and-fill’ resurfacing that maintains the structural numbers constant. Also, structural design is not considered as a decision variable, so we use the same structural numbers for future reconstructions.

Routine maintenance such as crack sealing and patching, which are known to slow down the deterioration process, are assumed to be applied regularly between interventions, and their costs and emissions are not considered in the optimization problem. We assume that 80% of truck traffic loading is applied on the right-most lane of multi-lane roadways, and the surface condition of the right lane is the relevant measure of condition in the optimization. The discount rate \(r\) is 0.07, and the maximum lifecycle length is 60 years, i.e. \(H_{\text{max}} = 60\). The details of the pavement deterioration, cost and emission models are presented in Appendix B.
3.1. Deterioration and improvement models

The segment-level condition state \( s_n \) is defined as a two-dimensional vector with continuous surface roughness, \( s_n \), in International Roughness Index (IRI) units of m/km and age, \( h_n \), defined as the elapsed time from the latest reconstruction or construction. The deterioration model of the AC pavements is adopted from Paterson (1987). This is a deterministic model, where future condition is dependent on current surface condition, time, traffic loading and structural number. We adopt the improvement models resulting from resurfacing from Lee and Madanat (2014b).

3.1.2. Cost models

User costs and emissions are mainly due to reduced fuel efficiency and the additional vehicle maintenance from poor surface condition (increased pavement roughness). User costs and emissions are also caused by congestion due to reconstruction. Night-time partial closure resurfacing is assumed, so additional user costs and emissions resulting from resurfacing activities can be ignored for multi-lane roadways. We assume partial closure during reconstruction. Additional user delays during reconstruction are computed through a queuing theory analysis using hourly traffic data. Traffic rerouting in the network is not considered.

For consistency with previous related work, we adopt the same models for user costs, \( U_n(\cdot) \), and agency costs, \( M_n(\cdot) \), of Lee and Madanat (2015). The vehicle operating costs, which is the sum of fuel, tire, maintenance and depreciation costs, are negatively related to surface condition. The increase in vehicle operating costs is a linear function of IRI, and dependent on vehicle type (Zaabar and Chatti, 2010). The travel time delay is due to partial roadway closures during reconstruction. Resurfacing costs are a linear function of the number of lanes and overlay thickness. Reconstruction costs are comprised of labor costs, disposal costs, material costs and so on. Those factors depend on the number of lanes and the structural design (represented in this paper by the structural number).

3.1.3. Emission models

The emission models from Reger et al. (2015) are used for both user emissions, \( W_n(\cdot) \), and resurfacing emissions, \( A_n(\cdot) \). Reconstruction emissions are estimated using the Pavement Life-cycle Assessment Tool (PaLATE) for Environmental and Economic Effects (PaLATE, 2013). The emissions are linearly related to thickness of each structural layer with a corresponding parameter dependent on the materials; this parameter includes the disposal emissions.

3.2. Case study results

The results for the selected highway system are shown in Fig. 1, with the GHG emission constraint on the x-axis, and the optimal total annual average discounted costs on the y-axis. The approximation factor, i.e., the arbitrary discount rate for emissions, is set as \( \rho = 0.002 \). All segments are assumed to be in the best condition at the start of the planning horizon, i.e. \( s_n = (s_{new}, 0 \text{ yr}) \). \( \forall n \). Because the optimal stationary policy is independent of the initial condition, there is no loss of generality from this assumption.

Fig. 1 can be understood as the Pareto Frontier of the equivalent multi-criteria problem; its slope is the shadow price of the constraint, \( \lambda \), which can be interpreted as the optimal price of carbon. The left end-point of the curve is the minimum GHG emission constraint, beyond which the problem is infeasible. The y-value of the right end-point is the value of the unconstrained cost minimization problem; higher GHG emission constraints will not be binding at optimality after this point. As the level of the emission constraint decreases by 5.6% from the right-end point to the left-end point, the total
costs increase by 11.3%. If the agency is currently minimizing lifecycle costs only (i.e., it is operating at the right end of the curve), large reductions in GHG emissions can be achieved for minimal increases in lifecycle costs. The price of carbon in this part of the curve is small (around $20 per ton of CO\textsubscript{2}), and therefore such reductions are politically feasible.

Fig. 2 demonstrates how optimal policies for resurfacing and reconstruction activities are affected by the GHG constraint. Each segment has its own stationary policy, so we plot the average lifecycle length, resurfacing frequency and trigger roughness for resurfacing for the 50 segments. Note that, for a segment that has multiple resurfacing activities in one life cycle, the trigger roughness levels are different because the segment policy is stationary in terms of the state vector, which includes age.

As shown in Fig. 2, more frequent resurfacings are necessary, and their average threshold decreases, with a more restrictive emission constraint. This, in turn, is expected to reduce users’ GHG emissions due to improved roadway surface condition. On the other hand, Fig. 2 shows that the reconstruction period is not very sensitive to the GHG constraint. Two opposite trends influence the reconstruction period as the emission constraint becomes more restrictive. On one hand, if the number of resurfacings in a lifecycle is unchanged, the time between reconstructions should decrease to reduce the user operating emissions. On the other hand, if the resurfacing frequency per life cycle increases, a longer reconstruction period is optimal.

The interdependency between resurfacing and reconstruction policies is illustrated in Fig. 3, where the optimal policy is presented for a rural highway segment in Route 680 in Contra Costa County. The figure shows that, as the system-level GHG constraint decreases, a shorter lifecycle length is optimal while the number of resurfacings per cycle is constant (two per cycle). At a certain point near the left-end of the figure, it becomes optimal to apply 3 resurfacings per cycle, and the lifecycle length increases from 50 years to 58 years. The average trigger roughness values for 2 or 3 resurfacings per cycle are represented by the gray line, which decreases with decreasing emission constraint.
By examining the trend of the average lifecycle length for the system, we can infer that it is efficient to maintain a constant number of resurfacings per cycle, and to reduce the lifecycle length and the trigger roughness if the emission constraint is moderate. For tighter emission constraints, it becomes optimal to perform more resurfacings per cycle, leading to longer lifecycle lengths.

Even if traffic rerouting is not assumed in this paper, traffic delays caused by lane closures during construction has a significant impact on the results. Fig. 4 shows how the objective value and the management strategies are influenced by different traffic delay levels ranging from 80% to 120% of the original value under the same constraint that $B$ is 4600 tons per year. The optimal results for the original problem are: the average lifecycle length is 55.67 year; the average number of resurfacing is 1.66; the average threshold of resurfacing is 2.67 IRI; and the total discounted annual average costs are 715,658 $ per year.

As can be seen in Fig. 4, the total discounted annual average cost, i.e. the objective value, increases with traffic delays. The optimal lifecycle lengths are shown to be positively related to travel time delay. As GHG emissions for each reconstruction activity increase, to satisfy the GHG emission constraint, fewer reconstructions should be performed. Note that no additional traffic congestion occurs as a result of resurfacing activities, because it is assumed that resurfacing is performed during night time hours. It is also shown in Fig. 4 that more resurfacings with lower threshold roughness levels need to be carried out. It has been noted that, in a longer lifecycle, more frequent resurfacing activities are necessary, according to the literature (e.g. Lee and Madanat, 2014b). We can observe that the resurfacing policy is more sensitive to travel time delay than the reconstruction policy. This is because the proportions of resurfacing in both the total costs and the total GHG emissions are lower than those of reconstructions. This traffic delay analysis shows that the reconstruction strategy is robust to the uncertainties of the construction-related traffic delays.

Recall that the original segment-level problem in Eq. (10) was approximated by introducing the discount rate $\rho > 0$ to transform the mixed-discounted-and-undiscounted problem into a mixed-discounted problem in Eq. (11). We showed that the approximated emissions converge to the true emissions as $\rho$ goes to zero in (16). In Appendix C, we show how the actual optimal results in the case study are influenced by the approximation discount rate.

For $\rho \leq 0.003$, $x_n^{*,\Lambda,\rho}$ for all $S_n \in \mathcal{S}_n$, is constant for each segment. Therefore, this optimal policy can be regarded as the close-to-optimal policy. The stationary policy converges because $|X|$ is finite for $|S_n(S_n)| < \infty$ and $|X_n(S_n)| < \infty$. The Pareto Curves obtained from selected $\rho$ values (0.015, 0.008 and 0.003) are shown in Fig. 5. As expected, a smaller $\rho$ yields better results. In other words, for a given GHG emission limit $B$, the discounted lifetime costs $J(S(0), x_n^*)$ increases with $\rho$. In the case of $\Lambda = 0$, the optimal results are independent of $\rho$, so the three curves meet at this point (the right-end point of all curves). We can also see that, if the life cycle costs are same, the GHG emissions are lower with lower $\rho$.

4. Conclusion

In this paper, we have presented an optimization problem to minimize the total lifetime cost for pavement systems under a GHG emission constraint. The problem is applicable to various kinds of infrastructures, heterogeneous treatments and multi-dimensional condition states. The main contribution of this research relative to the state of the art (in the field of multi-objective minimization of lifetime costs and emissions) is that the problem is formulated to include multi-dimensional
pavement segment states vector and heterogeneous management activities. It allows for a more realistic representation of the majority of current pavements in the world. For instance, the assumption that pavements are perpetual, i.e., do not need reconstruction during their lifetime, can be relaxed.

We have proposed a bottom-up solution method to find the optimal system-level policy satisfying the GHG emission constraint, which is a two-step approach, starting at the segment-level and ending at the system-level. The segment-level problem is approximated by a mixed-discounted problem and solved using dynamic programming. Numerical methods are used to solve the system-level problem. The sensitivity analysis indicates that the approximation discount rate does not markedly affect the final results.

As a case study, we chose 50 pavement highway segments in California District 4. Two kinds of interventions were considered, resurfacing and reconstruction. The case study results indicate that the total discounted costs increase and it is necessary to perform more resurfacings per cycle as the emission budget decreases. If the number of resurfacings in one lifecycle is unchanged, the reconstruction frequency should increase to reduce emissions. Optimal resurfacing and reconstruction policies are dependent on each other, and influenced by the emission constraint. The tradeoff between the GHG emissions and the total costs can be inferred from the Pareto Frontier. For the pavement segments used in this analysis, it can be seen that the possible reduction in GHG emissions is 5.6% of the maximum level for an increase in lifecycle costs of 11.3%. While this cost to carbon reduction ratio of 2 does not look encouraging, it is reasonable to assume that the agency is currently minimizing lifecycle costs only (i.e., it is operating at the right end of the curve). From that point, large reductions in GHG emissions can be achieved for minimal increases in lifecycle costs because the slope of the Pareto Frontier in that region is closer to 0.

The main limitation of our work is the assumption that the parameters of the cost and emission models, as well as traffic demand, are constant, when in fact, they may be highly variable in the near future due to innovation in automobile technology. Fuel efficiency is being improved, and user costs and vehicles’ GHG emissions will be reduced as a consequence. Moreover, electrical vehicles and hybrid vehicles are becoming more popular, so new emission and cost models are necessary. Freight logistics and car-sharing systems are also evolving, so we can expect traffic demand for both passenger vehicles and trucks to change.

The second limitation of our work is that network interdependencies are not considered. Ignoring interdependencies among network links allowed us to separate the system-level problem into a sum of segment-level problems, which produces a computationally efficient solution methodology, but some important aspects, such as traffic redistribution and resurfacing coordination across links, are not included. In future work, it is possible to include these effects by including appropriate nonlinear constraints, and using Approximate Dynamic Programming to solve the problem.

Lastly, we only considered the GHG emissions in this research. We believe that it should extended to other key air pollutants such as small particles, produced by the exhausts of diesel vehicles, and nitrogen oxides (NOx), responsible for about 75,000 premature deaths each year in the 40 countries in Europe (WHO, 2015).

**Acknowledgment**

Funding for this research was provided by the NYUAD research funds of the second author.
Appendix A. Calculations of the undiscounted average emissions

Suppose \( x_n \) is given. The GHG emissions with relative annual emissions between annual emissions and annual average emissions are denoted by \( \zeta_n(S_n(0), x_n) \), and it is defined as (A1).

\[
\zeta_n(S_n(0), x_n) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} [W_n(S_n(t), x_n(S_n(t))) + A_n(S_n(t), x_n(S_n(t))) - Q_n(S_{\text{new}}, x_n)]
\]  

(A1)

The relative value iteration is presented as (A2) (Bertsekas, 1998).

\[
\zeta_{n+1}^i(S_n, x_n) = E\left[ W_n(S_n, x_n(S_n)) + A_n(S_n, x_n(S_n)) + \zeta_n^i(S_n', x_n) \right] \\
- E\left[ W_n(S_{\text{new}}, x_n(S_{\text{new}})) + A_n(S_{\text{new}}, x_n(S_{\text{new}})) + \zeta_n^i(S_{\text{new}}, x_n) \right]
\]  

(A2)

where \( i \) is the iteration index and \( S_{\text{new}} \) is the next condition from \( S_{\text{new}} \) under \( x_n \).

If the above relative value iteration converges to a certain vector at the \( i \)th iteration, the undiscounted average emission \( \tilde{Q_n}(x_n) \) is computed by (A3).

\[
\tilde{Q_n}(x_n) = E\left[ W_n(S_{\text{new}}, x_n(S_{\text{new}})) + A_n(S_{\text{new}}, x_n(S_{\text{new}})) + \zeta_n^i(S_{\text{new}}, x_n) \right]
\]  

(A3)

Appendix B. Details of cost, emission, and deterioration models for case study

We adopt the pavement roughness model developed by Paterson (1987), which is deterministic and history dependent. The initial condition of a pavement of the unit period is \( S_0 = [s_0, h_0] \), and \( S_i^n = [s_i^n, h_i^n] \) refers to the condition after intervention is performed. If 'do-nothing' is selected as an intervention, \( S_n = S_i^n \), \( M_n(\cdot) = 0 \) and \( A_n(\cdot) = 0 \). Eq. (B1) states the continuous deterioration, \( F(S_i^n, u) \), during a unit period, i.e. one year, and it indicates that the process is influenced by the structural number, \( SN_n \), and the traffic loading, \( I_n \). In B1, if \( u = 1 \), \( S_n(u) \) is \( S_0 \) that is defined in (10). As necessary, we can set the unit time period to any time length instead of one year used in this paper. The model parameters in (B1) are estimated as: \( a = 725; b = 0.03 \) that is higher than the value estimated by Paterson (1987); and \( q = -4.99 \).

\[
F(S_i^n, u) = \left\{ s_i^n e^{bu} + a \cdot u \cdot (SN_n + 1)^9 \cdot I_n \cdot e^{b(h_i^n + u)} \cdot h_i^n + u \right\}, \forall u \in (0, 1] \]  

(B1)

If a resurfacing is carried out, the improvement model is shown as (B2), the cost model is formulated as (B3), a function of overlay thickness \( w(s_n) \) that is given as a function of initial roughness level as formulated in (B4), and given lane numbers, \( D_n \). Overlay thickness can be also a decision variable, but we do not include it in the optimization based on Lee and Madanat (2015), \( m_1 = 10.491 \) \$/lane-km-inch and \( m_2 = 33.012 \) \$/lane-km. \( m_3 \) is additional user costs resulted from construction process, and it is set as 0 in this paper by the assumptions made in 3.1. The scaling factor \( \Omega \) is multiplied to consider the relative weight between user cost and agency cost, but we set \( \Omega \) to one for the consistency to the previous literature. The emission model (B5) is linear to material used in resurfacing, where \( e_1 = 225 \) kg-CO\(_2\)/lane-km-mm. The recent related literature (Lidicker et al., 2012; Reger et al., 2015) set \( e_2 \) to 0, but we conservatively set this value to 3375 kg-CO\(_2\)/lane-km.

\[
S_i^n = \min (s_n, \max (s_0, (1 - \mu_1) \cdot s_n)), h_i^n \]  

(B2)

\[
M_n(S_n, \text{rehabilitation}) = D_n \cdot w(s_n) \cdot m_1 + D_n \cdot m_2 + \Omega \cdot m_3 \]  

(B3)

\[
w(s_n) = \frac{\mu_2 + \mu_3}{\mu_1} \cdot \left( s_n - s_0 \right) \]  

(B4)

\[
A_n(S_n, \text{rehabilitation}) = D_n \cdot w(s_n) \cdot e_1 + D_n \cdot e_2 \]  

(B5)

Reconstruction improvement model is represented in (B6). In this paper, we select a fixed value of \( s_{\text{new}} \) as the best achievable level after reconstruction. The cost model (B7) is linear to thicknesses of all layers, \( \rho^j \), where \( j = 1 \) for hot mix asphalt (HMA), \( j = 2 \) for aggregate sub-base, and \( j = 3 \) for aggregate base. For the unit cost for each layer, \( m_1^j = 3179.6 \) \$/lane-km-in, \( m_2^j = 1011.7 \) \$/lane-km-in, and \( m_3^j = 794.9 \) \$/lane-km-in. \( m_4 = 57.380 \) \$/lane-km, and \( m_5 \) is calculated based on the traffic volumes in peak-hours while a road is partially closed. Traffic redistribution is not considered. In the emission model presented in (B8), \( e_3^1 = 164.37 \) tons-CO\(_2\)/lane-ft and \( e_3^2 = 83.79 \) tons-CO\(_2\)/lane-ft. Typically, user emissions due to traffic delay is about 10–20% of the total emissions from construction, so scaling factor \( Y_n \) is multiplied.

\[
S_i^n = \{ s_{\text{new}}, 0 \} \]  

(B6)

\[
M_n(S_n, \text{reconstruction}) = \sum_{j=1.2.3} D_n \cdot \rho^j \cdot m_4^j + D_n \cdot m_5 + \Omega \cdot m_6 \]  

(B7)
\[ A_n(S_n, \text{reconstruction}) = Y_n \cdot \left[ \sum_{j=1,2,3} D_n \cdot \rho^j \cdot a_j^3 \right] \] (B8)

Additional user vehicle operating costs due to roughness in a unit time period, \( U(S_n, x_n(S_n)) \), is equivalent to \( U(S_u^+) \) because \( S_u^+ \) is determined by \( S_n \) and \( x_n(S_n) \) as shown in (B2) and (B6). As noted in (B9), \( U(S_n, x_n(S_n)) \) is a function of pavement roughness, with following parameters: \( m_7 \) is the marginal cost of a car and \( m_8 \) is the marginal cost of a truck and \( 0.004080$/IRI-km-truck \). AADT is the annual average traffic volume, and \( AADT_n \) is the annual average truck traffic volume. The additional user emissions due to pavement roughness, \( W(S_n, x_n(S_n)) \), is represented as (B10), where \( e_4 = 0.0028917 \text{kg-CO}_2E/IRI-km-car \) and \( e_5 = 0.0075516 \text{kg-CO}_2E/IRI-km-truck \).

\[
U(S_n, x_n(S_n)) = Y_n \cdot \int_0^1 (m_7 \cdot (AADT_n - AADTT_n) + m_8 \cdot AADTT_n) \cdot (F(S_n^+, u) - S_{\text{new}}) \, du
\] (B9)

\[
W(S_n, x_n(S_n)) = \int_0^1 (e_4 \cdot (AADT_n - AADTT_n) + e_5 \cdot AADTT_n) \cdot (F(S_n^+, u) - S_{\text{new}}) \, du
\] (B10)

**Appendix C. Approximation error in the emission function**

If the optimal strategy obtained in the approximated problem with positive \( \rho \) converges as \( \rho \) converges to zero, the strategy with non-zero \( \rho \) is close-to-optimal, where the global optimality is for \( \rho = 0 \).

We use different values of \( \rho \) to obtain the optimal policy, \( x_{n,\Lambda,\rho}^* \), and calculate the original value of \( Q_n(\cdot, x_{n,\Lambda,\rho}^*) \) (see Appendix A), where the GHG emission constraint, \( B \), is fixed for each value of \( \rho \). We compute the following measure to compare how the approximated average emissions converge to the undiscounted emission.

\[
\text{Error Measure} = \frac{\sum_{n=1}^N \left| Q_n(S_{\text{worst}, x_{n,\Lambda,\rho}^*}, \rho) - Q_n(S_{\text{new}, x_{n,\Lambda,\rho}^*}, \rho) \right|}{\sum_{n=1}^N Q_n(x_{n,\Lambda,\rho}^*)}
\] (C1)

where \( S_{\text{worst}} \) is the condition such that \( Q_n(S_{\text{worst}, x_{n,\Lambda,\rho}^*}, \rho) \) has the maximum value.

Fig. C1 has two axes: \( \rho \) and the error measure. We observe that \( \rho \) and the error measure are directly proportional, and this proportionality can be induced from (16) as \( \rho \) approaches zero.

![Fig. C1. Relation between the error measure and the approximation discount rate for GHG emissions.](image-url)
References


Rashid, M., Tsunokawa, K., 2012. Trend curve optimal control model for optimizing pavement maintenance strategies consisting of various treatments.


