



A time series analysis framework for transportation infrastructure management

Pablo L. Durango-Cohen *

*Department of Civil and Environmental Engineering and Transportation Center, Northwestern University,
2145 Sheridan Road, A335, Evanston, IL 60208-3109, United States*

Received 2 June 2005; accepted 29 August 2006

Abstract

We present an integrated framework to address performance prediction and maintenance optimization for transportation infrastructure facilities. The framework is based on formulating the underlying resource allocation problem as discrete-time, stochastic optimal control problem with linear dynamics and a quadratic criterion. Facility deterioration is represented as a time series which provides an attractive and rigorous approach to specify and estimate performance models. The state and decision variables in the framework are continuous which allows the framework to overcome important computational and statistical limitations that do not allow existing optimization models to address various problems that arise the management of transportation infrastructure. To illustrate the advantages of the proposed approach, we conduct a numerical study where we examine the case of multiple technologies being used simultaneously to collect condition data. Specifically, we illustrate how the framework can be used to quantify the effect of the capabilities of inspection technologies, i.e., precision, accuracy and relationships, on life-cycle costs. This information can be used to compute the operational value of combining technologies, and thus, to guide in their selection based on economic criteria.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Transportation infrastructure; Inspection technologies; Maintenance optimization; Stochastic optimal control; Kalman filter; Time series

1. Introduction

Transportation infrastructure management refers to the process of making decisions concerning the allocation of resources for the preservation, i.e., maintenance and repair (M&R), of the facilities that comprise transportation systems, e.g., pavement and bridge networks. In developed countries, where much of the transportation infrastructure is mature and portions are nearing the end of their service lives and need to be replaced, M&R decisions are increasingly important. This is due to both the far-reaching and serious negative impacts of deficient infrastructure, as well as the magnitude of M&R expenditures. In the United States, for

* Tel.: +1 847 491 4008; fax: +1 847 491 4011.

E-mail address: pdc@northwestern.edu

example, annual M&R expenditures are on the order of tens of billions of dollars.¹ M&R decisions trade off user costs, which depend on facility condition and correspond to a fraction of the costs associated with travel time, fuel consumption, vehicle depreciation and maintenance, with M&R costs. As facilities deteriorate, the rate at which user costs accrue increases. M&R costs are incurred to improve condition, and thus, reverse the effects of deterioration. Models to support infrastructure management evaluate both the short and long-term economic consequences associated with M&R decisions. This evaluation involves processing data related to current infrastructure condition and forecasting the effect of M&R decisions on future condition.² The economic consequences of M&R decisions are then estimated with a function that maps condition forecasts to costs.

The importance of transportation infrastructure management has, over the last 40 years, motivated a great deal of research to address both the development and estimation of statistical performance models to support condition forecasting, as well as the formulation and analysis of optimization models to support M&R decision-making. Optimization models to support M&R decisions have been developed from different perspectives to address numerous applications. Extensive surveys of M&R models and applications appear in Barlow et al. (1965), McCall (1965), Pierskalla and Voelker (1976), Sherif and Smith (1981), Valdez Flores and Feldman (1989) and Dekker (1996). With few exceptions, discrete-time M&R optimization models, consistent with the periodic review nature of infrastructure management, are formulated as finite (state and action) Markov decision processes (MDPs). Bellman (1955), Dreyfus (1960), Derman (1962) and Klein (1962) were first to propose the MDP formulations for M&R problems. Golabi et al. (1982) were first to adapt the methodology to the management of transportation infrastructure.

An important assumption in the MDP framework is that the state variables, representing facility condition or its proxies, are discrete. This seemingly innocuous assumption explains an important divergence in the infrastructure management literature. On one hand, there is the development and estimation of statistical models for condition forecasting (cf. Humplick (1992); Ben-Akiva and Ramaswamy (1993); Ben-Akiva and Gopinath (1995)). These models assume that facility condition is represented by continuous variables.³ On the other hand, there is the development and estimation of transition probabilities that are used for performance prediction in a manner consistent with the MDP framework, i.e., where condition or its proxies are represented by variables defined over discrete (and ordinal) sets (cf. Madanat et al. (1997) and Mishalani and Madanat (2002)). These sets are constructed by partitioning the variables' state-spaces into mutually exclusive and collectively exhaustive sets. The partitioning process introduces forecasting errors and uncertainty, and thus, explains why using continuous variables is not only intuitively appealing, but has also been shown in empirical studies to be superior.

In addition to introducing forecasting errors and uncertainty, the use of discrete state variables leads to computational and statistical limitations that make the MDP framework unattractive to support the management of transportation infrastructure. This is because both the number of parameters that require estimation to specify the transition probabilities for the model, and the computational effort to obtain optimal M&R policies increase exponentially with the number of variables in the model. These difficulties are well-known and referred to as "the curse of dimensionality". Moreover, these problems are statistically and practically significant because there are situations where having the flexibility to add variables to a model may be desirable, e.g., in cases where multiple technologies are used simultaneously to collect condition data, or to add explanatory variables to a facility deterioration model. The former is of practical importance as the use of multiple technologies (e.g., satellite imaging, video, radar, laser and sensors) to evaluate and measure distresses on transportation infrastructure is increasingly common. Relaxing the Markovian assumption (by add-

¹ Source: Federal Highway Administration, *Highway Statistics*, tables FA3, SF2, LGF2, SF4B for various years. In 2003, the most recent year for which statistics are available, federal, state and local maintenance expenditures on highways and roads exceeded \$30 billion.

² Data related to infrastructure condition are obtained by collecting distress measurements. Examples of distresses in pavement management include roughness, type and extent of cracking, rut depth and profile, and extent of surface patching. Condition forecasts are generated with a deterioration model which is a statistical expression that relates condition to a set of explanatory variables such as design characteristics, traffic loading, environmental factors, and history of M&R activities.

³ Condition is assumed to be continuous to reflect the fact that its degradation is caused by continuous physical and chemical processes that take place at the microscopic level. Examples of these processes include fatigue due to loadings, expansion and contraction due to temperature changes, and corrosion of reinforcing steel bars in reinforced concrete structures.

ing lagged dependent variables or by adding exogenous variables) are examples of situations where the flexibility to add explanatory variables to a performance model may be desirable. The motivation for considering this case is that empirical studies such as Ben-Akiva and Ramaswamy, 1993 in pavement management, and Madanat et al. (1997) and Mishalani and Madanat (2002) in bridge management have shown that physical deterioration of transportation infrastructure *may not be* Markovian.

To address the above limitations, we present an integrated framework to simultaneously address performance prediction and M&R decision-making for transportation infrastructure facilities. The framework is based on formulating the underlying resource allocation problem as a discrete-time, stochastic optimal control problem with linear dynamics and a quadratic criterion, i.e., a linear-quadratic (LQ) problem. Facility deterioration is represented as a time series with continuous state variables. This, in turn, makes the proposed framework consistent with models in the literature, and provides an attractive and rigorous approach to specify and estimate performance models. We describe the framework in detail in Section 2. In Section 3, we exploit the computational versatility of the proposed framework to examine the case of multiple technologies being used simultaneously to collect condition data. Specifically, we use numerical examples to illustrate how the framework captures the capabilities of inspection technologies, i.e., precision, accuracy and relationships, and therefore, can be used as a tool to guide the selection of inspection technologies based on economic criteria. Concluding remarks are presented in Section 4.

2. Model formulation

In this section, we present a LQ formulation for the problem of allocating resources for M&R of transportation infrastructure. The framework consists of two components: a state-estimation problem in which arrays of condition data are processed and used to generate condition forecasts; and an optimization problem that yields M&R policies. We begin this section by introducing notation and discussing the assumptions that we use to formulate and solve the proposed optimization model.

2.1. Notation and assumptions

We introduce the model as an alternative to the latent MDP formulation to support M&R decision-making for transportation infrastructure proposed by Madanat and Ben-Akiva (1994). In particular, we assume, without loss of generality, that the deterioration, measurement-error, and cost functions are Markovian and stationary. The other assumptions are as follows:

- (1) The state and decision variables in the model are defined over continuous spaces, i.e., $X_t \in \mathbb{R}^n, Z_t \in \mathbb{R}^m, t = 1, 2, \dots, T + 1$, and $A_t \in \mathbb{R}^p, t = 1, 2, \dots, T$. The n and m -dimensional vectors X_t and Z_t represent facility condition and the set of distress measurements collected at the start of stage t . The vector A_t represents the set of resource allocation decisions selected for stage t .
- (2) Following the latent performance modeling approach of Ben-Akiva et al. (1993), we assume that facility condition can be succinctly expressed in terms of characteristics such as: functional performance, structural fitness, safety, and aesthetics. The ambiguity that exists in defining and measuring these characteristics is captured by representing condition as a vector, X_t , consisting of latent/unobservable components. Henceforth, to simplify the presentation but without loss of generality, we assume X_t and A_t are unidimensional (i.e., $n = p = 1$).
- (3) The relationship between facility condition and the distress measurements is given by a measurement error model of the form: $Z_t = \alpha + \beta X_t + \xi_t$. The vector ξ_t is assumed to follow a Gaussian distribution with finite covariance matrix, R .
- (4) Facility deterioration can be represented by a stochastic linear system of the form: $X_{t+1} = gX_t + hA_t + \epsilon_t$. In time series analysis, this type of model is referred to as an AutoRegressive Moving Average with eXogenous input (ARMAX) model. We assume that ϵ_t follow a normal distribution with mean $\bar{\epsilon}_t$ and finite variance $\sigma_{\epsilon_t}^2$.
- (5) The period cost function can be represented (or approximated) by a second order polynomial, i.e., $g(X_t, A_t) = aX_t^2 + bX_t A_t + cA_t^2 + dX_t + eA_t + f$.

- (6) The salvage value function can be represented (or approximated) by a second order polynomial, i.e.,
- $$s(X_{T+1}) = -p_{T+1}X_{T+1}^2 - q_{T+1}X_{T+1} - r_{T+1}.$$

Prior to discussing a solution procedure for the problem, we note that the above assumptions are not overly restrictive. Specifically:

- The assumption of continuous state variables is consistent with state-of-the-art statistical performance models for infrastructure facilities. Furthermore, it is attractive because it eliminates forecasting errors and uncertainty introduced in the process of obtaining discrete measurement scales that are consistent with MDP formulations. The nature of these problems is documented in Madanat et al. (1997) and Mishalani and Madanat (2002).
- The decision variables in the framework, A_t , $t = 1, 2, \dots, T$, can be interpreted as investment levels or as maintenance rates. This assumption has been used before in M&R optimization models for transportation infrastructure (cf. Friesz and Fernandez (1979)), and it is appropriate for tactical and strategic models that are used for budget allocation decisions. This is in contrast with operational models used for sequencing and scheduling M&R activities. An important practical criticism of assuming that the decision variables are continuous is that different M&R decisions can only differ in their intensity. While this assumption might be a good approximation when considering activities such as overlays of different thicknesses, it is probably not appropriate when comparing across fundamentally different activities, e.g., routine maintenance versus reconstruction. This criticism can be mitigated by specifying condition as a (multi-dimensional) vector, e.g., $X_t = [x_t^f, x_t^s]$ (with x_t^f and x_t^s representing the facility's functional performance and structural fitness at the start of stage t , respectively), and adjusting the deterioration and measurement-error models accordingly. It is also possible to define a decision vector, for each period, e.g., $A_t = [a_t^f, a_t^s]$, with components that impact one or both of the condition types. This type of model can be used to evaluate trade-offs between M&R activities that differ in their effect and costs.
- The linear specifications assumed for the measurement-error models, along with the assumptions of continuous variables and that error terms are Gaussian, is compatible with statistical performance models for infrastructure facilities presented in the literature (cf. Hudson et al. (1987), Humplick (1992), Ben-Akiva and Ramaswamy (1993) and Ben-Akiva and Gopinath (1995)). The vector of measurements, Z_t , is modeled as a manifestation of the latent condition, X_t , which is corrupted by systematic and random errors, captured with the parameters α and β , and ξ_t , respectively. The components, α^i , β^i and ξ_t^i are assumed to be properties of the technology,⁴ i.e., α^i and β^i describe technology i 's accuracy and the variance of ξ_t^i , $\mathbb{V}(\xi_t^i) = \sigma_{\xi_t^i}^2$, is its precision. Often, distresses corresponding to different physical characteristics are collected and the number of measurements (i.e., the dimension of Z_t) is larger than the number of latent condition variables (i.e., the dimension of X_t). In such models, α , β and ξ account for the fact that the measurements are imperfect proxies of the latent condition variables.
- ARMAX models provide a convenient, flexible and rigorous framework to formulate and estimate infrastructure deterioration models; however, the assumption that they are appropriate requires (empirical) validation. While preliminary statistical analysis and results are encouraging (Chu and Durango-Cohen, submitted for publication),⁵ the assumptions embedded in ARMAX models may not be universally valid, e.g., that the effect of M&R actions is linear and additive. Even though there is flexibility to (partially) address some of the limitations, it is important to realize that the ARMAX framework may be inadequate in certain situations. Thus, the use of ARMAX models should be interpreted as analogous to the use of linear programming, even though most realistic problems probably do not exhibit the linear structure that is assumed.

⁴ The parameters in the model may also depend on factors such as location, time period, equipment operator, or the true value of the underlying distress.

⁵ State-space specifications of ARMAX models are used to formulate and estimate deterioration and measurement-error models for transportation infrastructure. The models are estimated using deflection and pressure measurements generated by sensors embedded in an asphalt pavement.

- The assumptions about costs are not overly restrictive because, for example, it may be possible to obtain optimal M&R policies for general classes of continuous cost functions by solving a sequence of problems. In each problem the cost function is approximated by a second-order Taylor Series expanded about a different point. The procedure is analogous to the Newton–Raphson method for solving systems of equations/optimization problems. The procedure is described further in Dreyfus (1977) and the references therein.

2.2. Optimization problem

With the assumptions discussed in the preceding section, the M&R optimization problem can be formulated as a dynamic program. The fact that condition is unobservable means that when selecting a course of action at the start of stage t , a decision-maker can only rely on the sequence of applied actions and measurements collected by the start of t . This information corresponds to the “state of the system” and can be represented with the information vector $I_t \equiv \{Z_1, A_1, \dots, Z_{t-1}, A_{t-1}, Z_t\} = \{I_{t-1}, A_{t-1}, Z_t\}$. The optimal objective function, $v_t(I_t)$ is defined as the minimum expected discounted cost from the start of t until the end of the horizon given the information available at the start of t , I_t . Mathematically, the model can be expressed with the following recurrence relation and boundary condition:

$$v_t(I_t) = \min_{A_t \in \mathbb{R}} \{ \mathbb{E}_{X_t|I_t} [aX_t^2 + bX_tA_t + cA_t^2 + dX_t + eA_t + f] + \delta \mathbb{E}_{I_{t+1}|I_t} [v_{t+1}(I_{t+1})] \}^6 \tag{1}$$

$$v_{T+1}(I_{T+1}) = \mathbb{E}_{X_{T+1}|I_{T+1}} [p_{T+1}X_{T+1}^2 + q_{T+1}X_{T+1} + r_{T+1}] \tag{2}$$

The first term in Eq. (1) corresponds to the expected costs incurred in stage t ; the second term corresponds to the minimum discounted (with a discount factor δ) sum of costs from the start of stage $t + 1$ to the end of the planning horizon (start of stage $T + 1$). The boundary condition (Eq. (2)) is used to assign the residual value at the end of the planning horizon.

With the assumptions presented in the previous section, the optimal policy for the above dynamic program can be expressed as a closed-form linear function of $\mathbb{E}_{X_t|I_t}[X_t]$. The parameters can be computed recursively with the following formulas:

$$\mu_t(I_t) = - \left[\frac{b + 2\delta p_{t+1}gh}{2c + 2\delta p_{t+1}h^2} \mathbb{E}_{X_t|I_t}[X_t] + \frac{2\delta p_{t+1}h\bar{e} + \delta q_{t+1}h + e}{2c + 2\delta p_{t+1}h^2} \right], \quad t = T, \dots, 1 \tag{3}$$

$$p_t = a + \delta p_{t+1}g^2 - \frac{(b + 2\delta p_{t+1}gh)^2}{4[c + \delta p_{t+1}h^2]}, \quad t = T, \dots, 1 \tag{4}$$

$$q_t = d + 2\delta p_{t+1}\bar{e}g + \delta q_{t+1}g - \frac{[b + 2\delta p_{t+1}gh][e + 2\delta p_{t+1}h\bar{e} + \delta q_{t+1}h]}{2[c + \delta p_{t+1}h^2]}, \quad t = T, \dots, 1 \tag{5}$$

$$r_t = f + \delta p_{t+1}(\bar{e}^2 + \sigma_e^2) + \delta q_{t+1}\bar{e} + \delta r_{t+1} - \frac{[e + 2\delta p_{t+1}h\bar{e} + \delta q_{t+1}h]^2}{4[c + \delta p_{t+1}h^2]}, \quad t = T, \dots, 1 \tag{6}$$

where $\mu_t(I_t)$ represents the optimal decision for period t , and is expressed as a function of the parameters p_t, q_t, r_t . The formulas are obtained by induction as shown in Appendix A. The equations are evaluated recursively noting that $p_{T+1}, q_{T+1}, r_{T+1}$ are the parameters that define the salvage value function.

Using the solution to the above system of equations allows us to obtain the following closed form expression for the optimal objective value function:

$$v_t(I_t) = p_t \mathbb{E}_{X_t|I_t}[X_t^2] + q_t \mathbb{E}_{X_t|I_t}[X_t] + r_t. \tag{7}$$

⁶ $\mathbb{E}_{Y|W}[\cdot]$ and $\mathbb{V}_{Y|W}(\cdot)$ are the conditional expectation and conditional variance operators. The expectation/variance is taken with respect to Y given W .

The computational effort to obtain an optimal M&R policy, $\mu_t(I_t)$, $t = 1, 2, \dots, T$, does not depend on the size of the vector Z_t . However, we note that to implement the optimal policy and to evaluate the optimal objective value function it is necessary to compute the conditional expected state given the set of information in each period. This step is referred to as the state-estimation problem and it is discussed further in the following section. We note that the key to processing distress measurements generated simultaneously by multiple technologies is to compute these expectations efficiently.

2.3. The state estimation problem

The state estimation problem consists of finding the conditional expected state given the set of information in each period, $\mathbb{E}_{X_t|I_t}[X_t]$, $t = 1, 2, \dots, T$. Under the assumptions discussed earlier, the expectation can be computed with a recursive algorithm known as the Kalman filter. The algorithm is presented below:

Kalman filter Algorithm

Repeat at the start of each period:

Given $\mathbb{E}_{X_{t-1}|I_{t-1}}[X_{t-1}]$, $\mathbb{V}_{X_{t-1}|I_{t-1}}(X_{t-1})$, A_{t-1} , and $Z_t = \bar{z}_t$

Define: $\hat{X}_{t-1} \leftarrow \mathbb{E}_{X_{t-1}|I_{t-1}}[X_{t-1}]$,

$P_{t-1} \leftarrow \mathbb{V}_{X_{t-1}|I_{t-1}}(X_{t-1})$, and

$I_t = \{I_{t-1}, A_{t-1}, \bar{z}_t\}$

Time Update:

$\hat{X}_t^- = g\hat{X}_{t-1} + hA_{t-1}$

$P_t^- = g^2P_{t-1} + \sigma_\epsilon^2$

Measurement Update:

$K_t = P_t^- \beta' (P_t^- \beta \beta' + R)^{-1}$

$\mathbb{E}_{X_t|I_t}[X_t] \leftarrow \hat{X}_t^- + K_t(\bar{z}_t - \beta\hat{X}_t^- - \alpha)$

$\mathbb{V}_{X_t|I_t}(X_t) \leftarrow (1 - K_t\beta)P_t^-$

The **Time Update** step uses the system equation to project the estimates of the conditional expectation and variance, i.e., the first two moments of the state distribution (which is Normal under the assumptions presented earlier). The **Measurement Update** step revises (with Bayes' Law) the conditional expectation and variance taking into account the new set of measurements obtained at the start of period t , \bar{z}_t . The computational effort of the Kalman filter, and hence the effort of obtaining and implementing an optimal policy using the proposed framework, increases polynomially with the dimension Z_t . This is because the algorithm consists of basic linear algebra operations whose complexity increases polynomially with the dimensions of the vector β and the matrix R . The dimensions of β and R increase linearly and quadratically, respectively, with the dimension of Z_t . As a result, the framework does not suffer from the shortcomings of the latent MDP approach.

3. Numerical study

In this section we present numerical examples to illustrate how the capabilities of (multiple/simultaneous) inspection technologies are captured in the time series framework, and, in turn, how the framework can serve to guide the selection of inspection technologies based on economic criteria. Specifically, we use the time series framework to:

- (1) Examine the effect of uncertainty, both in the deterioration process and in the data-collection process, on the optimal costs of managing transportation infrastructure facilities;

- (2) Analyze the impact of systematic measurement errors on optimal M&R decisions and the ensuing costs of managing transportation infrastructure facilities; and to
- (3) Quantify the value of combining inspection technologies for condition assessment.

We consider instances of the management of a facility over a planning horizon of 100 periods and assume that the period cost function, deterioration model, and measurement-error model are as follows:

$$g(X_t, A_t) = X_t^2 + A_t^2 - 700X_t + 121,597.75, \quad (8)$$

$$X_{t+1} = X_t - A_t + 30 + \epsilon_t; \text{ where } \epsilon_t \sim N(0, \sigma_\epsilon^2), \quad t = 1, 2, \dots, T. \quad (9)$$

$$Z_t = X_t + \zeta_t; \text{ where } \zeta_t \sim N(0, \sigma_\zeta^2), \quad t = 1, 2, \dots, T + 1. \quad (10)$$

The parameters presented above were selected to simulate measurements consistent with the data used in the studies by [Humplick \(1992\)](#) and [Ben-Akiva et al. \(1993\)](#). These studies consider different technologies to measure alligator cracking on a set of in-service pavements. The salvage value of the facility is set to zero (i.e., $s(X_{T+1}) = 0$); the discount rate to 5% ($\delta = 1/1.05$); and the initial condition, X_1 , to 350.5 (its optimal long-term steady state). To initialize the Kalman filter in the simulations, we set estimates of the first two moments of the state distribution $\mathbb{E}_{X_0|I_0}[X_0]$ and $\mathbb{V}_{X_0|I_0}(X_0)$ to 350.5 and 10,000, respectively.

In [Humplick \(1992\)](#) and [Ben-Akiva et al. \(1993\)](#), the true values of the distress, extracted from the measurements, ranged from 200 to 500 square feet. In the numerical examples that follow, we assume that the latent condition variable is defined over the interval $\mathcal{S} \equiv [200, 500]$. The deterioration process is such that, in the absence of M&R actions, it would take a facility 10 periods on average to deteriorate from the best condition to the worst. The parameters in the period cost function (8) were chosen to set the long-term steady state at the midpoint of \mathcal{S} , and to set the optimal costs of managing a deterministic system with a perfect technology to \$0. The initial facility condition and the length of the planning horizon were chosen to minimize their impact on the results.

Prior to presenting the results of the numerical study, we emphasize that our objective is to develop qualitative insights about how parametric changes affect the optimal costs of managing infrastructure facilities. Thus, the parameters used in the study are not representative of any particular facility, although the situation considered was “inspired” by the management of a pavement section whose condition is inspected on a yearly basis. In practice, some parameters, e.g., length of the review period, depend on managerial decisions, while others can be estimated using time series analysis. In particular, the specification and estimation of deterioration, measurement-error and user cost models that are consistent with the framework is an ongoing research area.

3.1. Effect of uncertainty on optimal costs

We use the time series framework to study the effect of uncertainty, both in the deterioration process and in the data-collection process, on the optimal costs of managing infrastructure facilities. Random measurement errors associated with the precision of a given technology constitute an important source of uncertainty in the data-collection process. Therefore, this study illustrates how the proposed framework can serve to quantify the value of using inspection technologies of different precisions.

In this framework, uncertainties in the deterioration and in the data-collection processes are captured by the parameters σ_ϵ and σ_ζ , respectively. To understand the impact of these parameters on the costs of managing infrastructure facilities, we consider their effect on the optimal objective function value for the first stage, $v_1(I_1)$, which represents the minimum expected costs from the start of the first period until the end of the planning horizon for a given I_1 . First, note that $v_1(I_1)$ varies linearly with the variance of the deterioration process, σ_ϵ^2 (see Eq. (6)). The effect of uncertainty in the data-collection process, i.e., measurement uncertainty, captured by σ_ζ is more subtle because it does not appear directly in $v_1(I_1)$. Intuitively, however, we know that this parameter determines the variance in the conditional state estimate, $\mathbb{V}_{X_1|I_1}(X_1)$, i.e., noisy measurements increase the uncertainty in the state estimate. By recalling that $\mathbb{E}_{X_1|I_1}[X_1^2] = \mathbb{V}_{X_1|I_1}(X_1) + \mathbb{E}_{X_1|I_1}^2[X_1]$ we observe that the optimal objective value function, $v_1(I_1)$, varies linearly with $\mathbb{V}_{X_1|I_1}(X_1)$. Unfortunately, the effect of σ_ζ

on $\mathbb{V}_{X_t|I_t}(X_1)$ can only be ascertained experimentally (through successive iterations of the Kalman filter), because the effect depends on the sequence of measurements, which also depends on σ_ϵ .

Hence, we conducted a simulation study with $\sigma_\epsilon \in \{0, 2.5, 5, 7.5, 10\}$ and $\sigma_\xi \in \{0, 25, 50, 75, 100\}$ in order to study the impact σ_ϵ and σ_ξ further. In the simulation, $\sigma_\epsilon = 0$ and $\sigma_\epsilon = 10$ denote a deterministic deterioration process and a highly variable deterioration process, respectively. Similarly, $\sigma_\xi = 0$ represents a perfect inspection technology and $\sigma_\xi = 100$ represents a highly imprecise technology. For the remaining 24 combinations of the two parameters besides the pair $(\sigma_\epsilon, \sigma_\xi) = (0, 0)$, we simulated 1000 instances of the deterioration and inspection process described by (8)–(10). The average total discounted costs of applying the optimal M&R policy are presented in Fig. 1.

From the figure we observe that, as expected, the costs to manage facilities increases as the uncertainty in the deterioration process grows, and that imprecise data collection technologies result in increased costs. To understand the effect of σ_ξ on $\mathbb{V}_{X_t|I_t}(X_t)$ and on costs, we consider random instances of the simulations described earlier with $\sigma_\epsilon = 5$. Fig. 2 shows how the Kalman filter updates the second moment of the state distributions for the four technologies considered, $\sigma_\xi = 25, 50, 75, 100$. We observe that the variance in the state distribution drops rapidly in the initial periods. The asymptote and the convergence rate are properties of a given technology’s precision. Precise technologies reduce the uncertainty in the conditional distribution of X_t given I_t , (i.e., $\mathbb{V}_{X_t|I_t}(X_t)$ is smaller). This leads to more efficient/appropriate M&R decisions and to reduced costs over the planning horizon. An important observation is that the conditional variance of the state distribution is well within the precision of each technology, i.e., the procedure filters out the random error/noise in the measurements. For example, the variance in the state distribution when measurements are collected with the technology $\sigma_\epsilon^2 = 10,000$ converges to approximately 1000.

3.2. Effect of systematic measurement errors on life-cycle costs

In this section, we consider the effects of these systematic measurement errors, i.e., additive and multiplicative biases, on the optimal costs of managing infrastructure facilities. Note that the **Measurement Update** step in the Kalman filter corrects for these biases in the process of estimating the first two moments of the

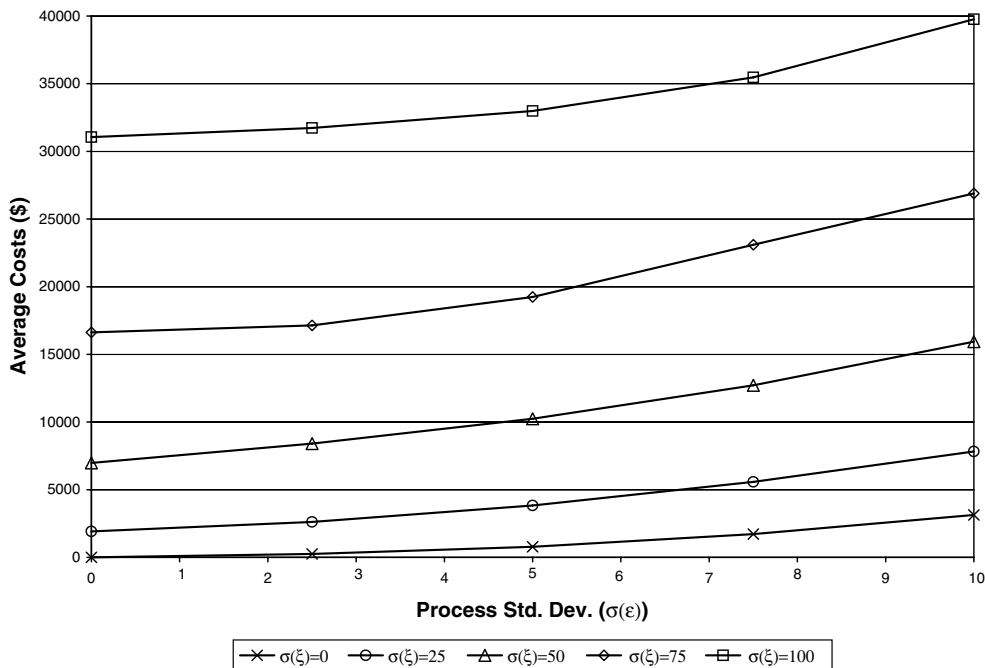


Fig. 1. Costs versus deterioration process standard deviation – Effect of technology precision.

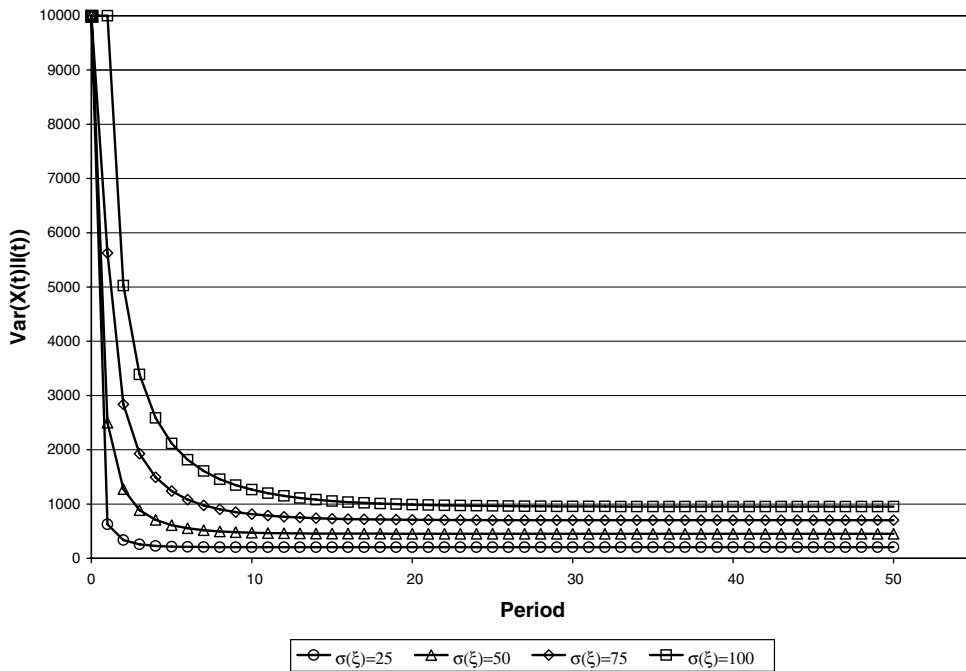


Fig. 2. Updated state distribution: second moments.

conditional state distribution for the given array of measurements. Prior studies in the literature argue that in the event that biases can be corrected for, technologies with higher precisions (lower values of σ_ξ) are preferable (cf. [Humplick \(1992\)](#) and [Ben-Akiva et al. \(1993\)](#)). We show that this is not entirely correct because a multiplicative bias changes the state of uncertainty about the distribution of the underlying latent variable. That is, given a measurement-error model that describes the capability of a technology, the variance of the distribution of X_t is $\frac{\sigma_\xi^2}{\beta}$.

To illustrate the effect of multiplicative bias on the costs of managing infrastructure facilities we consider instances of the process described by (8)–(10) with $\sigma_\epsilon \in \{0, 2.5, 5, 7.5, 10\}$. We assume the inspection technology is described with $\sigma_\xi = 50$ and $\beta \in \{0.5, 1, 1.5, 2\}$ and simulated 1000 instances for each of the 20 possible combinations of the parameters. The average total discounted costs are presented in Fig. 3. The results show how technologies that lead to more precise estimates of the latent condition, X_t , lead to lower life-cycle costs.

3.3. Effect of combining multiple technologies for condition assessment

In this section we show how the time series framework can be used to quantify the value of combining different technologies. In particular, we show that decisions related to adopting and deploying inspection technologies should not be dictated solely by precision but rather should consider how different technologies and the measurements they produce relate to each other.

We consider an inspection process that yields two distress measurements that are unbiased indicators of the single-dimensional underlying condition. That is, the measurement error model is:

$$Z_t = X_t + \xi_t. \tag{11}$$

The vectors ξ_t are assumed to follow a Gaussian distribution with finite covariance matrix. We assume that each of the technologies produces highly imprecise measurements with error standard deviation $\sigma_\xi = 50$. The relationship between the technologies is captured by the correlation between the measurements and we consider cases where $\rho = 0, 0.25, 0.5, 0.75, 1$ ($\rho = 0$ corresponding to independent technologies/measurements

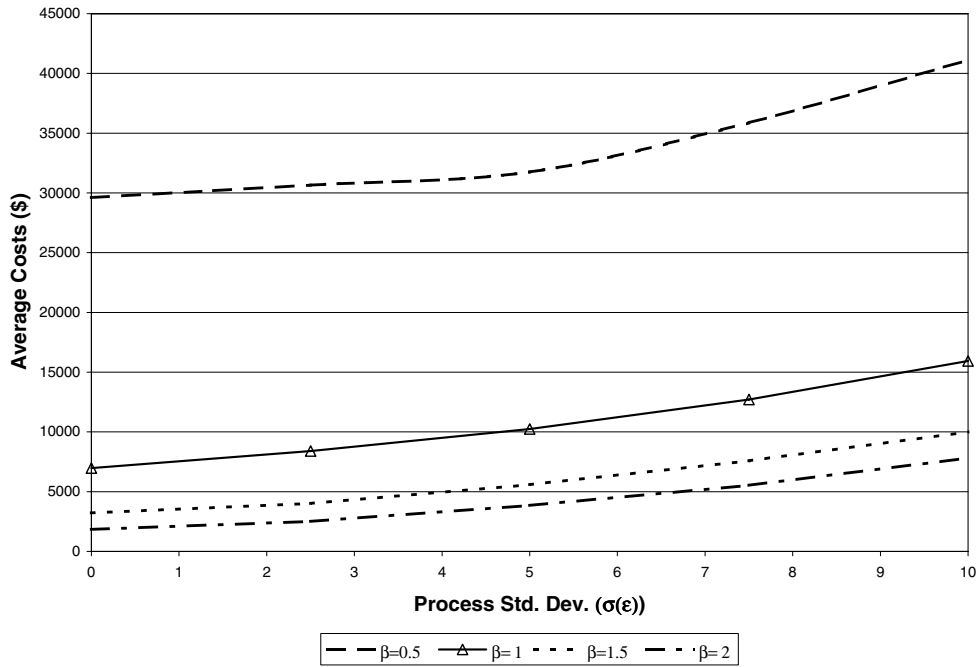


Fig. 3. Costs versus deterioration process standard deviation – Effect of multiplicative biases.

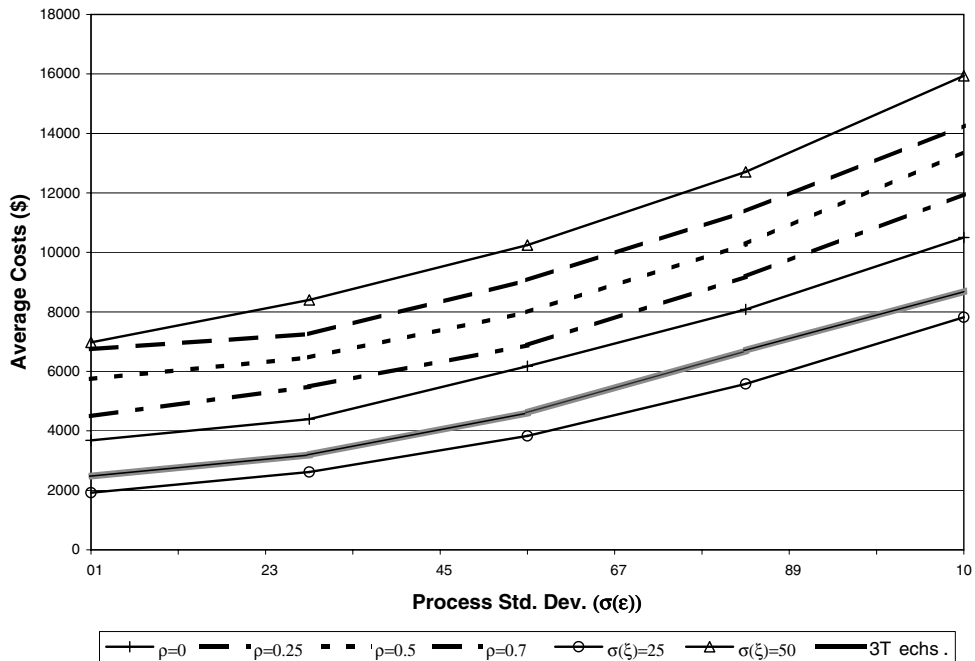


Fig. 4. Costs versus deterioration process standard deviation – Combining technologies for condition assessment.

and $\rho = 1$ corresponding to perfectly correlated technologies). The results, presented in Fig. 4, are for the average costs over 1000 instances for each of the values of ρ . The figure also includes the average costs when the facility is monitored with a single technologies with $\sigma_{\xi} = 50$ and $\sigma_{\xi} = 25$.

We observe that for the cases of two or three independent technologies ($\rho = 0$) the costs incurred are close to those incurred when the facility is monitored using a single technology with $\sigma_\xi = 25$ (i.e., a much more precise technology). The implication is that combining imprecise technologies can lead to substantial savings. This is particularly important because highly precise technologies tend to be much more expensive to adopt, particularly when they are first introduced. We also observe that in cases when the measurements are perfectly correlated ($\rho = 1$) the costs are identical to the costs incurred when the facility is managed using a technology with $\sigma_\xi = 50$. In this case there is no information gained by collecting a second or further distress measurements.

4. Conclusions

In this paper, we present an integrated framework to simultaneously address condition and cost forecasting and M&R decision making for transportation infrastructure facilities. The framework is based on formulating the underlying resource allocation problem as a discrete-time, stochastic optimal control problem with linear dynamics and a quadratic criterion. Facility deterioration is represented as a time series, which provides an attractive and rigorous approach to specify and estimate deterioration (and cost) models for the framework.

An important assumption in the proposed framework is that the state variables are defined over continuous sets. This makes the framework compatible with statistical models for performance prediction. Moreover, this representation allows the framework to overcome computational and statistical limitations of existing M&R optimization models based on the MDP approach.

We present numerical examples to showcase the versatility of the proposed framework. The examples are motivated by the proliferation of inspection technologies to support the management of transportation facilities. Specifically, we use the framework to analyze the impact of (i) uncertainty in the deterioration and data-collection/inspection processes, (ii) systematic measurement errors, and (iii) combining inspection technologies on the optimal costs of managing infrastructure facilities. The main observation that follows from the numerical results is that technology precision and accuracy, as well as the relationships between technologies or measurements all contribute to the capabilities of the inspection process, and thus, to the life-cycle costs of managing transportation facilities. This differs from earlier studies suggesting that inspection technologies can be selected solely on the basis of their precision. Our examples show that, even in the event that biases can be corrected for, accuracy (i.e., multiplicative bias) impacts the capability of the inspection process and the ensuing costs. We also show that combining imprecise technologies can improve the capabilities of the inspection process, and can lead to substantial operational savings. This is particularly important because highly precise technologies tend to be much more expensive to adopt, particularly when they are first introduced.

Acknowledgement

This work was partially supported by the National Science Foundation through Grant 0547471.

Appendix A. Optimal solution

We proceed by induction, as is done in Dreyfus (1977) or Bertsekas (2000), to show that the optimal policy, $\mu_t(I_t)$, and optimal objective value function, $v_t(I_t)$, are as given in expressions (3)–(7).

First, note that for stage $T + 1$, the optimal objective value function is given by the boundary condition as follows:

$$v_{T+1}(I_{T+1}) = \mathbb{E}_{X_{T+1}|I_{T+1}} [p_{T+1}X_{T+1}^2 + q_{T+1}X_{T+1} + r_{T+1}] = p_{T+1}\mathbb{E}_{X_{T+1}|I_{T+1}} [X_{T+1}^2] + q_{T+1}\mathbb{E}_{X_{T+1}|I_{T+1}} [X_{T+1}] + r_{T+1}$$

Next, we evaluate the following expectation assuming that $v_{t+1}(I_{t+1}) = p_{t+1}\mathbb{E}_{X_{t+1}|I_{t+1}} [X_{t+1}^2] + q_{t+1}\mathbb{E}_{X_{t+1}|I_{t+1}} [X_{t+1}] + r_{t+1}$:

$$\begin{aligned}
\mathbb{E}_{I_{t+1}|I_t}[v_{t+1}(I_{t+1})] &= \mathbb{E}_{I_{t+1}|I_t}[\mathbb{E}_{X_{t+1}|I_{t+1}}[p_{t+1}X_{t+1}^2 + q_{t+1}X_{t+1} + r_{t+1}]] = \mathbb{E}_{X_{t+1}|I_t}[p_{t+1}X_{t+1}^2 + q_{t+1}X_{t+1} + r_{t+1}] \\
&= \mathbb{E}_{X_t|I_t}[\mathbb{E}_{X_{t+1}|I_t, X_t}[p_{t+1}X_{t+1}^2 + q_{t+1}X_{t+1} + r_{t+1}]] \\
&= \mathbb{E}_{X_t|I_t}[\mathbb{E}_\epsilon[p_{t+1}(gX_t + hA_t + \epsilon)^2 + q_{t+1}(gX_t + hA_t + \epsilon) + r_{t+1}]] \\
&= \mathbb{E}_{X_t|I_t}[\mathbb{E}_\epsilon[p_{t+1}(g^2X_t^2 + h^2A_t^2 + \epsilon^2 + 2ghX_tA_t + 2gX_t\epsilon + 2hA_t\epsilon) + q_{t+1}(gX_t + hA_t + \epsilon) + r_{t+1}]] \\
&= p_{t+1}g^2\mathbb{E}_{X_t|I_t}[X_t^2] + p_{t+1}h^2A_t^2 + p_{t+1}(\sigma_\epsilon^2 + \bar{\epsilon}^2) + 2p_{t+1}gh\mathbb{E}_{X_t|I_t}[X_t]A_t + 2p_{t+1}g\bar{\epsilon}\mathbb{E}_{X_t|I_t}[X_t] \\
&\quad + 2p_{t+1}h\bar{\epsilon}A_t + q_{t+1}g\mathbb{E}_{X_t|I_t}[X_t] + q_{t+1}hA_t + q_{t+1}\bar{\epsilon} + r_{t+1}.
\end{aligned}$$

Thus, the optimal objective value function for stage t is as follows:

$$\begin{aligned}
v_t(I_t) &= \min_{A_t \in \mathbb{R}} \{ \mathbb{E}_{X_t|I_t}[aX_t^2 + bX_tA_t + cA_t^2 + dX_t + eA_t + f] + \delta \mathbb{E}_{I_{t+1}|I_t}[v_{t+1}(I_{t+1})] \} \\
&= \min_{A_t \in \mathbb{R}} \{ a\mathbb{E}_{X_t|I_t}[X_t^2] + b\mathbb{E}_{X_t|I_t}[X_t]A_t + cA_t^2 + d\mathbb{E}_{X_t|I_t}[X_t] + eA_t + f \\
&\quad + \delta [p_{t+1}g^2\mathbb{E}_{X_t|I_t}[X_t^2] + p_{t+1}h^2A_t^2 + p_{t+1}(\sigma_\epsilon^2 + \bar{\epsilon}^2) + 2p_{t+1}gh\mathbb{E}_{X_t|I_t}[X_t]A_t \\
&\quad + 2p_{t+1}g\bar{\epsilon}\mathbb{E}_{X_t|I_t}[X_t] + 2p_{t+1}h\bar{\epsilon}A_t + q_{t+1}g\mathbb{E}_{X_t|I_t}[X_t] + q_{t+1}hA_t + q_{t+1}\bar{\epsilon} + r_{t+1}] \} \tag{12}
\end{aligned}$$

To find the optimal decision for (t, I_t) we consider:

$$\begin{aligned}
\frac{\partial v_t(I_t)}{\partial A_t} &= b\mathbb{E}_{X_t|I_t}[X_t] + 2cA_t + e + \delta(2p_{t+1}h^2A_t + 2p_{t+1}gh\mathbb{E}_{X_t|I_t}[X_t] + 2p_{t+1}h\bar{\epsilon} + q_{t+1}h) = 0 \\
\mu(I_t) &= - \left[\frac{b + 2\delta p_{t+1}gh}{2c + 2\delta p_{t+1}h^2} \mathbb{E}_{X_t|I_t}[X_t] + \frac{2\delta p_{t+1}h\bar{\epsilon} + \delta q_{t+1}h + e}{2c + 2\delta p_{t+1}h^2} \right] \tag{13}
\end{aligned}$$

We also note that the second order/sufficiency condition is satisfied when:

$$\frac{\partial^2 v_t(I_t)}{\partial A_t^2} = 2c + \delta p_{t+1}h^2 \geq 0$$

The expressions for parameters p_t , q_t and r_t can be obtained by substituting the result obtained for $\mu(I_t)$ in (13) for A_t in expression (12). The results are presented in Eq. (4) through (6).

References

- Barlow, R., Proschan, F., Hunter, L., 1965. *Mathematical Theory of Reliability*. Wiley, New York, NY.
- Bellman, R.E., 1955. Equipment replacement policy. *Journal of the Society of Industrial and Applied Mathematics* 8 (3), 133–146.
- Ben-Akiva, M., Gopinath, D., 1995. Modeling infrastructure performance and user costs. *ASCE Journal of Infrastructure Systems* 1 (1), 33–43.
- Ben-Akiva, M., Ramaswamy, R., 1993. An approach for predicting latent infrastructure facility deterioration. *Transportation Science* 27 (2), 134–153.
- Ben-Akiva, M., Humplick, F., Madanat, S., Ramaswamy, R., 1993. Infrastructure management under uncertainty: The latent performance approach. *ASCE Journal of Transportation Engineering* 119 (1), 43–58.
- Bertsekas, D., 2000. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA.
- Chu, C., Durango-Cohen, P., 2005. Estimation of latent infrastructure performance models using time series analysis. *Transportation Research Part C*.
- Dekker, R., 1996. Applications of maintenance optimization models: A review and analysis. *Reliability Engineering and System Safety* 51, 229–240.
- Derman, C., 1962. On sequential decisions and markov chains. *Management Science* 9 (1), 16–24.
- Dreyfus, S., 1960. A generalized equipment replacement study. *Journal of the Society of Industrial and Applied Mathematics* 8 (3), 425–435.
- Dreyfus, S., 1977. *The Art and Theory of Dynamic Programming*. Academic Press, San Diego, CA.
- Friesz, T., Fernandez, E.J., 1979. A model of optimal transport maintenance with demand responsiveness. *Transportation Research Part B* 13 (4), 317–339.
- Golabi, K., Kulkarni, R., Way, G., 1982. A statewide pavement management system. *Interfaces* 12 (6), 5–21.
- Hudson, W., Elkins, G., Uddin, W., Reilley, K., 1987. Improved Methods and Equipment to Conduct Pavement Management Distress Surveys. Final Report, FHWA, Report No. FHWA-TS-87-213.

- Humplick, F., 1992. Highway pavement distress evaluation: Modeling measurement error. *Transportation Research Part B* 26 (2), 135–154.
- Klein, M., 1962. Inspection–maintenance–replacement schedules under markovian deterioration. *Management Science* 9 (1), 25–32.
- Madanat, S., Ben-Akiva, M., 1994. Optimal inspection and repair policies for infrastructure facilities. *Transportation Science* 28 (1), 55–61.
- Madanat, S., Karlaftis, M., McCarthy, P., 1997. Probabilistic infrastructure deterioration models with panel data. *ASCE Journal of Infrastructure Systems* 3 (1), 4–9.
- McCall, J., 1965. Maintenance policies for stochastically failing equipment: A survey. *Management Science* 11 (5), 493–524.
- Mishalani, R., Madanat, S., 2002. Computation of infrastructure transition probabilities using stochastic duration models. *ASCE Journal of Infrastructure Systems* 8 (4), 139–148.
- Pierskalla, W., Voelker, J., 1976. Survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly* 23 (3), 353–388.
- Sherif, Y., Smith, M., 1981. Optimal maintenance models for systems subject to failure – a review. *Naval Research Logistics Quarterly* 28 (1), 47–74.
- Valdez Flores, C., Feldman, R., 1989. A survey of preventive maintenance models for stochastically deteriorating single unit systems. *Naval Research Logistics Quarterly* 36 (4), 419–446.