

Controlling Tracking Performance for System Health Management - A Markov Decision Process Formulation

Brian Bole¹, Kai Goebel², and George Vachtsevanos³

^{1,2} NASA Ames Research Center, Moffett Field, CA, 94040

brian.bole@nasa.gov

kai.goebel@nasa.gov

³ Georgia Institute of Technology, Atlanta, GA, 30332

gfv@ece.gatech.edu

ABSTRACT

After an incipient fault mode has been detected a logical question to ask is: How long can the system continue to be operated before the incipient fault mode degrades to a failure condition? In many cases answering this question is complicated by the fact that further fault growth will depend on how the system is intended to be used in the future. The problem is then complicated even further when we consider that the future operation of a system may itself be conditioned on estimates of a system's current health and on predictions of future fault evolution. This paper introduces a notationally convenient formulation of this problem as a Markov decision process. Prognostics-based fault management policies are then shown to be identified using standard Markov decision process optimization techniques. A case study example is analyzed, in which a discrete random walk is used to represent time-varying system loading demands. A comparison of fault management policies computed with and without future uncertainty is used to illustrate the limiting effects of model uncertainty on prognostics-informed fault management policies.

1. INTRODUCTION

Diagnostic routines give an operator or supervisory controller an indication of component malfunction so that fault management (FM) actions can be taken prior to more serious failures. Because it is in many cases not possible or cost effective to replace components at the first sign of malfunction, additional prognostic models are desired to estimate how long degraded components may continue to be used before failures occur.

After detecting a fault or other anomalous system behavior, it may first be necessary to determine whether or not system stability and performance can be maintained. A survey of various methods used to maintain high performance in the presence of incipient faults or failures is found in (Zhang & Jiang, 2008). If the system can continue to operate in the

presence of a fault, then we can consider the prediction of fault growth and eventual system failure with continued use. It generally holds that higher system performance will result in higher loads on system components, and high loads will increase the risk of further component degradation. This sets the stage for FM policy that seeks to trade off reduced system performance for a reduced risk of component degradation and failure.

The prognostics-informed FM problem may be generally described in terms of the following parts:

1. A space of available FM actions that may be taken at present or future decision making epochs
2. Potentially uncertain models used to evaluate the cost and risk of available FM actions
3. A searching technique to seek a balance between the cost and risk of available FM actions

The *cost* and *risk* terms above are used to distinguish between the upfront operational cost of an FM action, and the estimated effect on mission safety respectively. The contribution of this paper is to present and examine a novel Markov decision process (MDP) formulation for prognostics-informed FM. The action space considered here is a range of acceptable, although potentially undesirable, deviations from the nominal tracking performance of a system. The FM formulation presented here incorporates explicit stochastic models for future system output demands, and component degradation as a function of applied load. Dynamic programming is shown to solve for an optimal finite horizon FM policy, after formulating the FM problem as an MDP. Finally, a multivariate stochastic system example, originally introduced in (Bole et al., 2012b), is used to demonstrate the formulation and solution of the FM problem as an MDP.

The FM action space discussed here is similar to that described in (Gokdere et al., 2006), which considered the prognostics-based adaptation of weighting factors in a lin-

ear quadratic regulator. The control optimization approach described here builds on our previous work regarding prognostics-informed component load allocation (Bole et al., 2011). A generalized Markov process formulation of component fault growth dynamics, originally described in (Bole et al., 2012a), is also adapted for use here.

The MDP formulation that is presented here is significant, because MDP optimization tools are widely used to solve cost and risk balancing problems, but there are currently few examples of their use in the area of prognostics-informed FM. Some examples of MDP for FM have been published in the areas of scheduled maintenance (Smilowitz & Madanat, 1994), health care (Sonnenberg & Beck, 1993), and autonomous mission replanning (Balaban & Alonso, 2013; Agha-mohammadi et al., 2014). A formal description of fault growth modeling and remaining useful life estimation in terms of Markov process models can be found in (Banjevic & Jardine, 2006). Established usage of MDP optimization methods for sequential decision making in the presence of stochastic modeling information is currently found in areas such as economics (Hauriea & Moresino, 2006), supply chain management (Parlara et al., 1995), and robotics (Cassandra et al., 1996).

This paper is organized as follows. Section 2 introduces a generalized representation of component degradation dynamics in terms of a multivariate Markov process. Section 3 describes an MDP formulation of the tracking performance planning FM problem, and the use of dynamic programming to identify optimal finite horizon FM policies. Section 4 introduces a case study example, in which uncertainty in fault growth physics models is represented by a uniformly distributed random process, and uncertainty in future exogenous loading demand models is represented by a discrete random walk. Concluding remarks are given in Section 5.

2. BUILDING A MARKOV PROCESS MODEL FOR FAULT GROWTH DYNAMICS

The eventual failure of individual components within a multi-component system is represented here in terms of component fault modes that will grow in severity until they cross a threshold, after which they are considered no longer viable. Fault magnitudes are assumed to be represented by a positive real number, corresponding to a measurable physical property such as crack length, spall width, or pitting depth; although, in many cases, faults cannot be directly measured *in situ* and diagnostic routines are needed to approximate current fault magnitudes based on the secondary effects observed in available sensor measurements (Feldman et al., 2010).

A generic function is introduced here to represent the dynamics of a particular component failure mode in discrete-time;

$$s^q(k+1) = f(s^q(k), u^q, \xi) \quad (1)$$

Here, $s^q(k)$ represents a fault magnitude for the q^{th} compo-

nent in a system at time-index k , u^q represents a load applied to component q , and ξ is a random variable representing uncertainty in this fault growth model.

The component load terminology is used here as a stand-in for pressure, force, torque, or a wide variety of other stressors that drive component deterioration. Component loads are assumed to be dictated partly by the dynamics of the system's operating environment, and partly by available supervisory FM actions that may be taken in response to online estimates of environmental states and component fault magnitudes.

The stochastic component degradation process given in Eq. (1) is formulated in terms of a discrete Markov process as:

$$\begin{aligned} p_{i,j}^q(u^q) &= p(s^q(k+1) = s_j | s^q(k) = s_i, u^q) \\ &= \sum_{\xi \in \Xi} p(\xi(k) = \xi) \cdot p(f(s_i, u^q, \xi) = s_j | u^q, \xi), \\ s_i, s_j &\in S^q, u^q(k) \in U^q(k), k \in \mathbb{N}, i, j \in \{0, 1, \dots, n\} \end{aligned} \quad (2)$$

$$\sum_{j=0}^n p_{i,j}^q = 1, \quad \forall i \in \{0, 1, \dots, n\} \quad (3)$$

Here, $p_{i,j}^q(u^q)$ represents the probability of transitioning from damage state s_i to damage state s_j , given a particular component loading, u^q . The $f(s_i, u^q, \xi)$ term, represents the fault growth model introduced in Eq. (1). The S^q , U^q , and Ξ terms, represent quantized state spaces for s^q , u^q , and ξ respectively. Equation (3) specifies that the sum of all transition probabilities defined at each system state must always be equal to one.

A mandate of monotonically increasing component fault modes is subsequently incorporated into the Markov process notation given in Eq. (2) as:

$$p_{i,j}^q = 0, \quad \text{if } j < i \quad (4)$$

This constraint will be problematic for other fault growth modeling techniques that represent process uncertainty with an analytical distribution that lacks an explicit lower bound. For example, in the case of Kalman filtering or Gaussian process models of fault growth, an assumption of unbounded Gaussian uncertainty would introduce some probability that the fault mode will be smaller in the future than it was known to be in the past. It would be necessary, in such cases, to assure that the probability attributed to non-realizable outcomes, ($p(s^q(\tau) < s^q(t))$ for $\tau > t$), will be acceptably small.

Sensor noise and feature mapping uncertainties will often result in significant uncertainty in estimates of present fault magnitudes. It is common practice for such diagnostic estimates to be reported in terms of a probability distribution over the potential fault magnitudes that could correspond to a given set of observations. The incorporation of uncertain beliefs about the present state of a system at fixed decision

making epochs can be found in publications on partially observable Markov decision processes; see the survey paper by Lovejoy for more information (Lovejoy, 1991). The additional notation necessary to include state estimation uncertainty in the FM problem is omitted from this paper in order to promote clarity in this initial work.

The Markov process notation given here may be used to describe stochastic fault growth process models in which the following assumptions are satisfied:

- *Assumption 1*: The fault growth dynamics are taken to be memoryless; i.e., the conditional probability distribution for future states depends only on the present state of the process, and not the past. This assumption is commonly referred to as the Markov assumption.

Should it be the case that a fault growth process of interest is not completely memoryless, but future states only depend on a finite number, m , of previous states, then the Markov process notation given here could be extended to satisfy the Markov assumption by defining the state space of the process to be the ordered m -tuple of the current state and the m previously visited states (Wang & Chang, 1996).

- *Assumption 2*: State transition probabilities are considered to be time invariant; although, it may be the case that fault growth models are not precisely known a priori and must be adapted online using techniques such as particle filtering (Orchard et al., 2008) or Bayesian learning (Saha et al., 2009).
- *Assumption 3*: At all discrete time-steps, the state space, the action space, and the space of environmental and other exogenous inputs to the system are adequately represented by a finite quantization of these spaces.

In the event that fault growth must be modeled as a continuous time process, a representation of fault growth modeling similar to that given here may be expressed in terms of a continuous time Markov process (Serfozo, 1979) or a semi-Markov process (Dong & He, 2007).

If all assumptions are satisfied, then the state transition probabilities defined in Eq. (2), are directly derivable from Eq. (1), given a model for the statistics of the random variable, ξ . The notation given in Eq. (2) does not explain however how component loads, represented by u^q , will be expected to vary in the system of interest.

If component loads are assumed to be directly controllable, then the prognostics-based control problem may be viewed as a component load allocation problem. Practical applications of control in terms of instantaneous component load allocations are currently found for aircraft (Boikovic & Mehra, 2002), spacecraft (Shertzer et al., 2002), and automobiles (Hattori et al., 2002). Two previous publications demonstrated the consideration of fault diagnostic updates and uncertain prognostic estimates to optimize component load allocation in an electro-mechanical actuator (Bole et al., 2010) and an unmanned ground vehicle (Bole et al., 2011). This

paper takes a new approach to the optimization of component load in response to diagnostic and prognostic updates. Here, the FM problem is formulated in terms of the component load reduction resulting from the degradation of a system's nominal tracking performance.

2.1. Fault Prediction in Terms of Performance Allocation

The FM problem is considered here in the context of a trade-off between competing desires to reduce loads on degraded system components, while also minimizing deviation from a system's nominal tracking performance. This section introduces a novel notation for describing the FM action space in terms of its effect on a system's nominal tracking performance. This new notation will be shown to beneficially simplify the FM optimization description, which will be introduced in the following section.

First, the loads that would be exerted by a nominal control policy on system components are represented by the following generic function:

$$\mathbf{u} = g(\mathbf{w}) \quad (5)$$

where \mathbf{w} represents a vector of environmental or other exogenous disturbances to a system, \mathbf{u} represents a component loading vector, and $g(\mathbf{w})$ represents the loading response of a nominal control policy.

Evaluation of system tracking performance is represented as:

$$J_T = \int_0^N h(\mathbf{u}(t), \mathbf{w}(t)) dt \quad (6)$$

where J_T represents a tracking performance score and $h(\mathbf{u}, \mathbf{w})$ represents a deterministic function used to evaluate system tracking performance as a function of \mathbf{u} and \mathbf{w} .

Next, the tracking performance of a system under nominal control is defined as J_N , and the tracking performance of a system after an FM action is taken is defined as:

$$J_T = J_N \cdot \rho \quad (7)$$

where ρ represents an induced reduction in the tracking performance of a nominally controlled system. Here, $\rho = 1$, indicates no change from the nominal control policy, and $\rho < 1$ represents FM actions that degrade a system's nominal tracking performance.

This paper does not address the real-world implementation of FM actions that would effect the u and ρ parameters. The discussion in this paper proceeds under the assumption that available FM actions, whatever they may be, can be expressed in abstracted form presented here.

A modification of Eq. (5) is introduced next to represent component loads for control policies with tracking performance

$$J_N \cdot \rho. \quad \mathbf{u} = \hat{g}(\mathbf{w}, \rho) \quad (8)$$

Substituting Eq. (8) into the state transition model in Eq. (2) gives:

$$\begin{aligned} p_{i,j}^q(\rho(k)) &= p(s^q(k+1) = s_j | s^q(k) = s_i, \rho = \rho(k)) \\ &= \sum_{\mathbf{w} \in \mathbf{W}} \sum_{\xi \in \Xi} p(\mathbf{w}(k) = \mathbf{w}) \cdot p(\xi(k) = \xi) \cdot \\ &\quad p(f(s_i, \hat{g}(\mathbf{w}, \rho)^q, \xi) = s_j | \mathbf{w}, \xi, \rho), \quad (9) \end{aligned}$$

$$\text{where } s_i, s_j \in S$$

This notation is now only missing a representation of the dynamics for the random variables, ξ and \mathbf{w} that were used to denote process uncertainty and future demand uncertainty respectively. If the random variables, \mathbf{w} or ξ , are independent and identically distributed (i.i.d.), then a single probability mass function (pmf) will describe $p(\mathbf{w}(k) = \mathbf{w})$ or $p(\xi(k) = \xi)$. However, if the random variables are not i.i.d., then additional dependencies may need to be included to identify the variable's pmf at time-index k . The work presented here will make the assumption that ξ is i.i.d., while \mathbf{w} will be allowed to be non-i.i.d. The additional notation needed to incorporate a Markov process representation for \mathbf{w} into Eq. (9) is covered in the next section.

2.2. Allowing Non-i.i.d. Exogenous Input Modeling

The exogenous input term, \mathbf{w} , is considered in only one dimension at this point in order to simplify the stochastic process notation. This notation could be extended later to handle multiple exogenous input sources as need arises. A generic Markov process representation for w is:

$$p_{l,m}(w) = p(w(k) = w_m | w(k-1) = w_l) \quad w_l, w_m \in W \quad (10)$$

where $p(w(k) = w_m | w(k-1) = w_l)$ represents the probability of w transitioning to state w_m at time-index k , given that it was in state w_l at time-index $k-1$.

The Markov process model of the degrading system can now be expressed as a four dimensional matrix that incorporates the Markov process model for w .

$$\begin{aligned} p_{(i,j),(l,m)}^q(\rho) &= p_{l,m}(w) \cdot \sum_{\xi \in \Xi} p(\xi(k) = \xi) \cdot \\ &\quad p(f(s_i, \hat{g}(w_m, \rho)^q, \xi) = s_j | w_m, \xi, \rho), \quad (11) \end{aligned}$$

$$\text{where } s_i, s_j \in S, w_l, w_m \in W$$

where, $p_{(i,j),(l,m)}^q(\rho)$ represents the probability of component q transitioning from fault state s_i and exogenous loading demand w_l to fault state s_j and exogenous loading demand w_m .

3. FAULT MANAGEMENT AS A FINITE HORIZON MDP

Next, the Markov process model developed in Eq. (11) is used to formulate the FM problem as a Markov Decision Process (MDP). This requires formalizing the space of all control actions that may be selected by a FM routine when the system is in any particular state.

At a given time-index the system will be in some state, s , and the FM system may select any control action that is defined to be available in state s . At the next time-index the system will move to state s' , and a new control action will be selected. The combination of the FM action and resulting state transition is assigned a cost within the MDP that is referred to as the state transition cost.

Over the past several decades much has been published on the theory of encoding various forms of risk aversion into the specification of MDP state transition costs (Hernandez & Marcus, 1996; Ruszczyrski, 2010). This task is similar in nature to the expected utility maximization problems that have come into popular use in the practice of decision theory (Schoemaker, 1982).

An FM policy is considered here to be mapping of all system states to corresponding actions, with the objective of minimizing the accumulation of state transition costs incurred. If an FM policy is evaluated over an infinite horizon, then the optimal mapping of system states to control actions will not vary based on the time-index. This is referred to as a stationary control policy. However, if FM policies are evaluated over a finite horizon, then the optimal mapping will change based on the time-index. This is referred to as a non-stationary control policy.

This paper focuses on a finite horizon formulation of the FM problem. This type of optimization may be applicable to FM problems in which a machine will be taken out of service after fixed time-intervals for maintenance. The time-varying nature of optimal finite horizon MDP policies is highlighted with an example in the following section.

The space of available FM actions is supposed here to be represented by a domain of ρ values available at all system states. Where, ρ as defined in Section 2.1, represents the system's tracking performance. Possible FM policies are denoted: $\pi = \{\mu_0(s(0), w(0)), \dots, \mu_{N-1}(s(N-1), w(N-1))\}$. Where, FM actions are assumed to be taken at each time-index in the interval $k = [0, N-1]$, and μ_k represents a function used at time-index k to map observations of s and w into a FM action;

$$\rho(k) = \mu_k(s(k), w(k)) \quad (12)$$

The expected cost of enacting a particular FM policy, given

initial values for s and w at time-index 0, is denoted:

$$J_{\pi}(s(0), w(0)) = E \left\{ c_N(s(N)) + \sum_{k=t_p}^{N-1} c_k(s', w', s, w, \rho) \right\} \quad (13)$$

Here, $c_k(s', w', s, w, \rho)$ denotes a state transition cost assigned to possibility of transitioning from one component fault state vector and one exogenous loading demand, (s, w) , to another, (s', w') , given a supervisory control action, ρ . A terminating cost, denoted by $c_N(s(N))$, penalizes the total component degradation over a simulated time window. Cost discounting and average cost formulations, used in the formulation of infinite horizon MDPs, are not considered here.

An optimal FM policy is defined as a one that minimizes J_{π} ;

$$J_{\pi^*}(s(0), w(0)) = \min_{\pi \in \Pi} J_{\pi}(s(0), w(0)) \quad (14)$$

where π^* represents an optimal FM policy, and Π represents the space of all possible FM policies.

After stating the FM problem as an MDP, optimized policies may be identified using well studied techniques such as backwards induction for finite horizon policies, and linear programming, value iteration, and policy iteration for discounted and average-reward infinite horizon policies.

The well known dynamic programming algorithm uses backwards induction to identify an optimal FM policy over the time window $k \in \{N-2, N-1\}$, and then for $k \in \{N-3, N-2, N-1\}$, and so on until the optimal policy is found over the entire time-window of interest. A detailed description of backwards induction algorithms can be found in standard texts, such as (Bertsekas, 1995).

The computational burden of this solution method is: $O(mn^2N)$. Here, N represents the time horizon to be optimized over. The variables m and n represent the cardinalities of the discrete action space for ρ and the state space of (s, w) respectively.

While this is a great improvement over the computational burden of an exhaustive search, which is $O(m^{n^N})$, the reader should note that the cardinalities of the state space and action space used in the MDP will grow exponentially with the dimensionalities of s , w , and ρ . Therefore, the dynamic programming method quickly becomes computationally infeasible for higher dimensional problems. Approximate dynamic programming algorithms must then be used to search for near-optimal solutions in higher dimensional systems, where it would be infeasible to identify optimal solutions with dynamic programming (Powell, 2007).

Note that while the discovery of an optimizing MDP policy through finite horizon dynamic programming may be computationally challenging, the optimal policy is computed offline,

and requires no online optimization as long as the Markov process model used to generate the policy is still applicable. If online updates to the Markov process modeling of environmental loading and fault growth dynamics were considered, then the optimizing policy would need to be recomputed.

4. CONSIDERATION OF A MULTIVARIATE STOCHASTIC SYSTEM CASE STUDY

The following discrete time component damage accumulation model is considered here:

$$s(k+1) = f(s(k), u(k), \xi(k)) = \begin{cases} s(k) - \lambda \cdot |u(k)| \cdot \xi(k), & s > 0 \\ 0 & \text{else} \end{cases} \quad (15)$$

where $s \in [0\%, 100\%]$ represents system health as a percentage between 100%, indicating perfect health, to 0%, indicating failure. Component load is represented by u , and process uncertainty in component health deterioration modeling is represented by ξ . This model defines the rate of damage accumulation to be proportional to the magnitude of applied load, u . Where, process uncertainty, ξ , and a constant proportional factor, λ , are included as multiplicative terms in the relationship.

Equation (8) introduced the function $u = \hat{g}(w, \rho)$ to describe component loads as a function of a stochastic vector, w , representing exogenous demands on the system, and a variable, ρ , representing induced deviation from a nominal tracking performance. Here, $\hat{g}(w, \rho)$ is assumed to take the form,

$$\hat{g}(w, \rho) = w \cdot \rho \quad (16)$$

Substitution of Eq. (16) into Eq. (15) yields:

$$s(k+1) = f(s(k), \rho(k) \cdot w(k), \xi(k)) = \begin{cases} s(k) - \lambda \cdot |\rho(k) \cdot w(k)| \cdot \xi(k), & s > 0 \\ 0 & \text{else} \end{cases} \quad (17)$$

Process uncertainty, $\xi(k)$, in the example fault growth model is taken to be represented by a uniform distribution over the set $\{.7, .8, .9, 1.1, 1.2, 1.3\}$;

$$p(\xi(k) = \xi) = \begin{cases} \frac{1}{6} & \xi \in \{.7, .8, .9, 1.1, 1.2, 1.3\} \\ 0 & \text{else} \end{cases} \quad (18)$$

The following discrete random walk process, is considered for the modeling of future exogenous inputs:

$$p_{l,m}(w) = \begin{cases} \frac{1}{3} & w_m - w_l \in \{-1, 0, 1\} \\ 0 & \text{else} \end{cases}, w_l, w_m \in \mathbb{N} \quad (19)$$

Figure 1 shows box plots for the parameters $|w|$, ξ , and $|w| \cdot \xi$ that were generated from 100 simulations of the stochastic

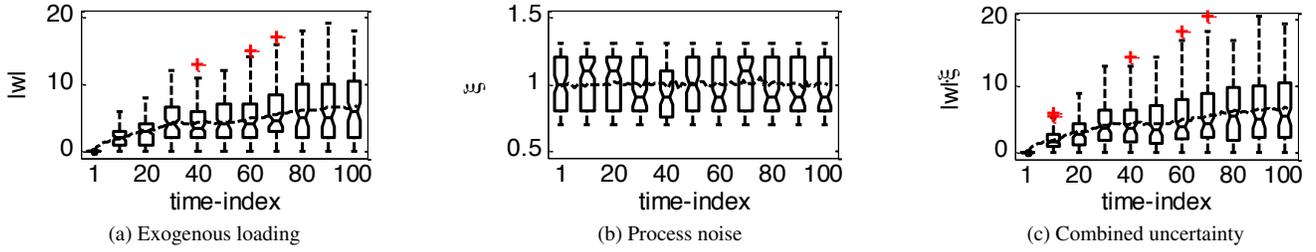


Figure 1. Box plots of the w (a), ξ (b), and $w \cdot \xi$ (c) parameters over 100 simulations of the stochastic system

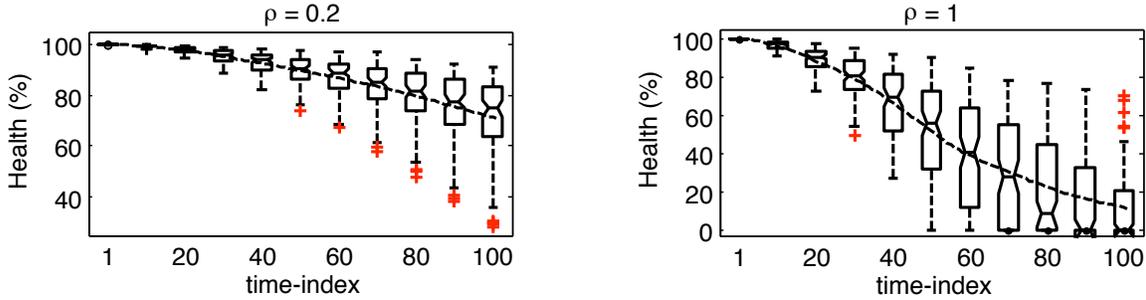


Figure 2. Box plots for 100 simulations of the example fault growth process with $\rho = 0.2$ (left), and with $\rho = 1$ (right).

system. The box plots shown in Figure 1 provide a convenient means of representing the statistics of the stochastic variables as observed over repeated simulations. The top and bottom of the boxes plotted in Figure 1 represent the first and third quantiles of the simulation data at a given time-index. The notch in each box represents the median of the data points. The dashed line represents the mean value observed at each time-index. Finally, the whiskers in the box plots extend to the most extreme points falling within the range,

$$q_1 - 1.5 \cdot (q_3 - q_1) < d_i < q_3 + 1.5 \cdot (q_3 - q_1) \quad (20)$$

where q_1 and q_3 are the first and third quantiles of the data respectively, and d_i represents a datapoint. Points falling outside of this range are considered outliers and are denoted in the plots with red crosses.

Figure 2 shows the results of 100 repeated simulations of Eq. (17) using $\lambda = \frac{1}{30}$ and two sample values of ρ : $\rho = .2$ and $\rho = 1$. Setting $\rho = .2$ over the 100 time-index simulation corresponds to an 80% reduction of the system's nominal tracking performance. Setting $\rho = 1$ corresponds to no deviation from a system's nominal tracking performance.

It can be observed from the sample results shown in Figure 2 that enacting the 80% reduction in nominal tracking performance over the time-window shown would result in very little risk of the component's health deteriorating beyond 40%. This control policy is perceived to be very 'safe', but likely overly conservative for many cases. On the other hand, enacting no deviation from the nominal tracking performance

would likely be unacceptably 'risky', as it is seen to result in component failure in many of the simulation runs.

4.1. Optimal FM With and Without Future Uncertainty

The Markov process representation of this example system is expressed as a single dimensional version of Eq. (11).

$$p_{(i,j),(l,m)}(\rho) = p_{l,m}(w) \cdot \sum_{\xi \in \Xi} p(\xi(k) = \xi) \cdot p(f(s_i, \hat{g}(w_m, \rho), \xi) = s_j | w_m, \xi, \rho), \quad (21)$$

State transition costs are designated as:

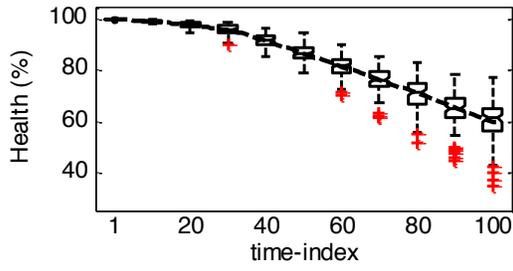
$$c_k(s', w', s, w, \rho) = 1 - \rho \quad (22)$$

Recall from Section 2.1, that $\rho = 1$ corresponds to no deviation from a system's nominal tracking performance. Correspondingly, we see in Eq. (22) that $\rho = 1$ is assigned a state transition cost of zero.

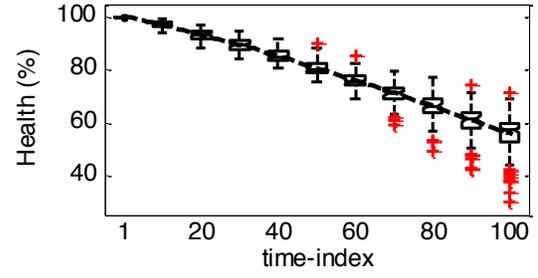
The terminating cost for this example is designated to be inversely proportional to the square of component health percentage at time-index N .

$$c_N(s(N)) = (100 - s(N))^2 \quad (23)$$

The domain of feasible FM policies is considered here to allow the choice of nominal performance reductions from 0% to 80% at each simulated time-index. The domain of allowable ρ values that may be enacted at each simulated time-



(a) Health degradation for optimal FM computed with uncertain modeling



(b) Health degradation for optimal FM computed with future knowledge

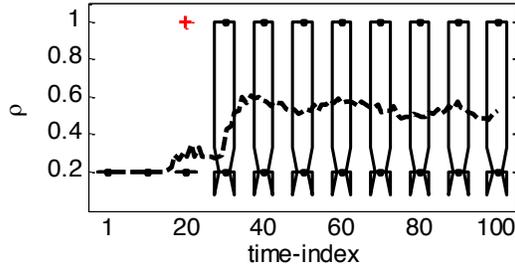
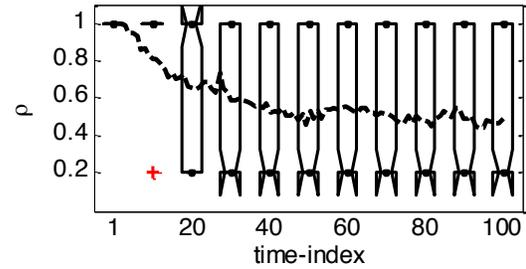
(c) Tracking performance deviation coefficient, ρ , for optimal FM computed with uncertain modeling(d) Tracking performance deviation coefficient, ρ , for optimal FM computed with future knowledge

Figure 3. Plots of component health percentage (top) and ρ (bottom) over 100 simulations of the optimal FM policy computed using a stochastic model of the system (left) and computed using future knowledge of the particular evolution of random variables observed in each simulation of the simulated system (right).

Table 1. Data Units, Sources, and Dates

	$\mu(\sum_k c_k)$	$\sigma(\sum_k c_k)$	$\mu(c_N)$	$\sigma(c_N)$	$\mu(J_\pi)$	$\sigma(J_\pi)$
Sample FM policy: $\rho = 1$	0	0	81.6	26.7	81.6	26.7
Sample FM policy: $\rho = .2$	78.4	0	9.9	10.5	88.3	10.5
Optimal FM policy computed using stochastic modeling	55.1	18.2	16.8	7	71.8	23
FM policies computed using perfect future knowledge	40.7	20.1	20.4	7.3	61.1	25.7

index is defined here as $\rho \in \{.2, .3, \dots, 1\}$.

As described in Section 3, optimal finite-horizon MDP policies may be found using dynamic programming. The optimal FM action at time-index k will be dependent upon the current component health state, the exogenous input state, and the future variation that is estimated for the random variables included in the system model. The effect that model uncertainty has on the optimization of FM is quantified here by comparing the optimal FM actions identified using stochastic models for w and ξ , with the optimal FM actions identified using deterministic knowledge of the values taken by w and ξ over each simulation run. The optimal FM policy computed using deterministic knowledge of the profiles taken by w and ξ over a simulation run represents the typically unrealizable case of policy optimization with perfect future knowledge.

The difference in total cost for FM policies computed with and without future knowledge is referred to as regret (Jacquet & Szpankowski, 2004).

Figures 3a and 3c show the distributions of component health percentage and ρ values observed over 100 repeated simulations of the optimal FM policy calculated using the stochastic models for w and ξ described earlier. Figures 3b and 3d show the distributions of component health percentage and ρ values observed over repeated simulations of an optimal FM policy that is computed using prior knowledge of the profiles to be taken by w and ξ in each simulation run.

Table 1 shows the sample mean and standard deviation of the state transition costs ($\sum_k c_k$), the terminating cost (c_N), and the total control cost ($J_\pi = c_N + \sum_k c_k$) for the four control policies discussed in this paper. Comparison of the control

costs given in Table 1 show that the two optimal FM policies computed with and without future knowledge clearly score lower control costs than the two sample FM policies that were used to generate Figure 2. The optimal FM policy computed with future knowledge is also clearly seen to outperform the optimal FM policy computed using stochastic modeling information.

A visual comparison of the health deterioration plots given in Figure 3, shows a somewhat slower degradation of component health early in the mission for the FM policy lacking perfect future knowledge. The slower degradation of component health for the FM policy lacking future knowledge corresponds to more conservative FM actions, resulting in larger induced reductions to nominal tracking performance. Comparison of the ρ plots given in Figure 3, shows that the optimal control computed with uncertain modeling starts by commanding maximum degradation of nominal system tracking performance ($\rho = 0.2$), while the optimal control computed with future knowledge starts by commanding no degradation of the nominal system performance ($\rho = 1$).

Higher conservatism at the beginning of a mission is expected for the control policy lacking future knowledge, given that prognostic uncertainty and assessed risk will be highest early in the mission. The mean ρ command (represented by the dashed lines in Figure 3) is seen to converge to approximately $\rho = 0.5$ for both control policies as the end of the mission draws near. The similarity of the optimal policies computed with and without future uncertainty towards the end of the simulated mission is also expected, given that future uncertainty is decreasing as the end of the mission approaches.

The results shown here are highly dependent on the control costs, defined in Eq. (22) and Eq. (23). Different definitions for Eq. (22) and Eq. (23) would effect the optimal FM policies identified, and would thus be likely to result in a different set of observed behaviors. The design of control cost definitions to best reflect performance and safety assurance goals is outside of the scope of this paper, although this would be essential for practical applications of this control approach.

The primary objective of the analysis presented here was to demonstrate the incorporation of stochastic models for component degradation and exogenous demands into an optimal FM policy. The demonstration of the optimal FM policies identified with and without future uncertainty was presented to illustrate the effect of uncertainty on optimal control solutions.

5. CONCLUSIONS

A generalized Markov process representation of fault dynamics was developed for the case that available modeling of fault growth physics and available modeling of future environmental stresses may be represented by two independent Markov process models. A metric was introduced to represent the magnitude of nominal tracking performance reduction to be

caused by a given set of fault management (FM) actions.

A Markov decision process (MDP) formulation of the FM problem was provided for a system with multiple degrading effectors. Dynamic programming was shown to solve for the optimal MDP policy over a finite time-window. A multivariate stochastic process example was considered to illustrate the effects of compounding uncertainties in physics of failure and exogenous demand modeling.

Problems still to be tackled using the notational tools described in this paper include: multi-component system applications, comparative analysis to evaluate the effect of using various prognostic horizon lengths in the formulation of supervisory FM, and the utilization of discounting and average reward cost functions for infinite horizon optimizations.

ACKNOWLEDGMENT

The project support by AvSafe/SSAT and OCT/ACLO are respectfully acknowledged

REFERENCES

- Agha-mohammadi, Ali-akbar, Ure, N. K., How, J. P., & Vian, J. (2014). Health aware stochastic planning for persistent package delivery missions using quadrotors. *IEEE Intelligent Robots and Systems*.
- Balaban, E., & Alonso, J. (2013). A modeling framework for prognostic decision making and its application to UAV mission planning. In *Annual conference of the prognostics and health management society*.
- Banjevic, D., & Jardine, A. (2006). Calculation of reliability function and remaining useful life for a Markov failure time process. *IMA Journal of Management Mathematics*, 17, 115-130.
- Bertsekas, D. (1995). *Dynamic programming and optimal control*. Athena Scientific.
- Boikovic, J. D., & Mehra, R. K. (2002). Control allocation in overactuated aircraft under position and rate limiting. In *Proceedings of the american control conference* (p. 791-796).
- Bole, B., Brown, D. W., Pei, H.-L., Goebel, K., Tang, L., & Vachtsevanos, G. (2010, Oct.). Fault adaptive control of overactuated systems using prognostic estimation. In *Annual conference of the prognostics and health management society*.
- Bole, B., Goebel, K., & Vachtsevanos, G. (2012a). Markov modeling of component fault growth over a derived domain of feasible output control effort modifications. In *Annual conference of the prognostics and health management society*.
- Bole, B., Goebel, K., & Vachtsevanos, G. (2012b). A stochastic optimization paradigm for the development and validation of fault risk assessment and risk management techniques. In *AIAA infotech conference*.

- Bole, B., Tang, L., Goebel, K., & Vachtsevanos, G. (2011). Adaptive load-allocation for prognosis-based risk management. In *Annual conference of the prognostics and health management society*.
- Cassandra, A., Kaelbling, L., & Kurien, J. (1996). Acting under uncertainty: Discrete Bayesian models for mobile-robot navigation. In *IEEE/RSJ international conference on intelligent robots and systems*.
- Dong, M., & He, D. (2007). Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis. *European Journal of Operational Research*, 178, 858-878.
- Feldman, A., Kurtoglu, T., Narasimhan, S., Poll, S., Garcia, D., Kleer, J., ... Gemund., A. (2010). Empirical evaluation of diagnostic algorithm performance using a generic framework. *International Journal of Prognostics and Health Management*, 1(2).
- Gokdere, L. U., Bogdanov, A., Chiu, S. L., Keller, K. J., & Vian, J. (2006, March). Adaptive control of actuator lifetime. In *IEEE aerospace conference*.
- Hattori, Y., Koibuchi, K., & Yokoyama., T. (2002, Sept.). Force and moment control with nonlinear optimum distribution for vehicle dynamics. In *Proc. of the 6th international symposium on advanced vehicle control*.
- Hauriea, A., & Moresino, F. (2006). A stochastic control model of economic growth with environmental disaster prevention. *Automatica*, 42(8), 1417-1428.
- Hernandez, D., & Marcus, S. (1996). Risk sensitive control of Markov processes in countable state space. *Systems & Control Letters*, 29, 147-155.
- Jacquet, P., & Szpankowski, W. (2004). Markov types and minimax redundancy for markov sources. *IEEE Transactions on Information Theory*, 50, 1393-1402.
- Lovejoy, W. S. (1991). A survey of algorithmic methods for partially observed Markov decision processes. *Annals of Operations Research*, 28, 47-66.
- Orchard, M., Kacprzynski, G., Goebel, K., Saha, B., & Vachtsevanos, G. (2008). Advances in uncertainty representation and management for particle filtering applied to prognostics. In *Annual conference of the prognostics and health management society*.
- Parlara, M., Wang, Y., & Gerchak, Y. (1995). A periodic review inventory model with Markovian supply availability. *International Journal of Production Economics*, 42(2), 131-136.
- Powell, W. (2007). *Approximate dynamic programming: Solving the curses of dimensionality*. John Wiley & Sons.
- Ruszczyriski, A. (2010). Risk-averse dynamic programming for markov decision processes. *Mathematical Programming*, 125, 235-261.
- Saha, B., Goebel, K., Poll, S., & Christophersen, J. (2009). Prognostics methods for battery health monitoring using a Bayesian framework. *IEEE Transactions on Instrumentation and Measurement*, 58(2), 291-296.
- Schoemaker, P. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, 529-563.
- Serfozo, R. F. (1979). An equivalence between continuous and discrete time Markov decision processes. *Operations Research*, 27, 616-620.
- Shertzer, R. H., Zimpfer, D. J., & Brown., P. D. (2002, Aug.). Control allocation for the next generation of entry vehicles. In *AIAA guidance, navigation, and control conference*.
- Smilowitz, K., & Madanat, S. (1994). Optimal inspection and repair policies for infrastructure facilities. *Transportation Science*, 28, 55-62.
- Sonnenberg, F., & Beck, R. (1993). Markov models in medical decision making. *Medical Decision Making*, 13(4), 322-338.
- Wang, H. S., & Chang, P.-C. (1996). On verifying the first-order Markovian assumption for a Rayleigh fading channel model. *IEEE Transactions on Vehicular Technology*, 45(2), 353-357.
- Zhang, Y., & Jiang, J. (2008). Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control*, 32, 229-252.