Adaptive Load-Allocation for Prognosis-Based Risk Management

Brian Bole [1], Liang Tang [2], Kai Goebel [3], George Vachtsevanos [1]

[1] *Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA.*

bbole3@gatech.edu
gjv@ece.gatech.edu

[2] *Impact Technologies, LLC, Rochester, NY 14623, USA.*

Liang.Tang@impact-tek.com

[3] *NASA Ames Research Center, Moffett Field, CA 94035, USA.*

kai.goebel@nasa.gov

ABSTRACT

It is an inescapable truth that no matter how well a system is designed it will degrade, and if degrading parts are not repaired or replaced the system will fail. Avoiding the expense and safety risks associated with system failures is certainly a top priority in many systems; however, there is also a strong motivation not to be overly cautious in the design and maintenance of systems, due to the expense of maintenance and the undesirable sacrifices in performance and cost effectiveness incurred when systems are over designed for safety. This paper describes an analytical process that starts with the derivation of an expression to evaluate the desirability of future control outcomes, and eventually produces control routines that use uncertain prognostic information to optimize derived risk metrics. A case study on the design of fault-adaptive control for a skid-steered robot will illustrate some of the fundamental challenges of prognostics-based control design.

1. INTRODUCTION

Some form of risk management can be seen in virtually every decision that human beings make. Typically, the desirability of future outcomes can be objectively evaluated; however, evaluating the best present control decision is complicated by uncertainty in estimating the future effects of control actions. In the case of controlling a system with incipient faults, the design objective is to obtain a system with high performance, low maintenance cost, and low failure rates. The effects of decisions regarding the design, maintenance, and operation of a system on its future performance, maintenance cost, and failure rates are commonly estimated by using empirical data or expert knowledge to assess probable outcomes.

The fault analysis process typically starts with the identification of potential failure modes and the quantification of the severity and likelihood of each, based on expert knowledge and historical data. The Failure Modes, Effects, and Critically Analysis (FMECA) is one of the most widely applied *a priori* fault analysis methods; it is currently nearly universally applied in industrial automation (Gergely, Spoiala, Spoiala, Silaghi, & Nagy, 2008), automotive (SAE, 1994), and aerospace (Saglimbene, 2009) industries. Fault Tree Analysis (FTA), Event Tree Analysis (ETA), Reliability Block Diagrams (RBD), and other fault analysis techniques that utilize historical failure rates will continue to play an ever more prominent role in the design of hardware specifications and contingency management policies.

In addition to the established practice of utilizing historical fault data to manage failure risks, there is also a growing push to develop technologies for online fault identification and fault growth prediction to improve system operation and maintenance. Online anomaly detection and diagnostic routines are enabling an increased use of condition based maintenance and control (CBMC) policies (Rao, 1998). Pseudo-inverse (Caglayan, Allen, & Wehmuller, 1988), model predictive control (MPC) (Monaco, Ward, & Bateman, 2004), and H_2 and H_∞ robust control theory (Doyle, Glover, Khargonekar, & Francis, 1987) are commonly used methods to recover controllability of a system after a known fault mode is detected. Further improvements in performance and safety are expected if the diagnostic information used by CBMC routines is supplemented with prognostic routines that predict the growth fault modes as a function of future use; however, the development and use of prognostic information is typically an extremely challenging proposition due to significant uncertainty in predicting future fault growth. Prudent methods for modeling fault diagnostic and prognostic uncertainty should be selected on a case-by-case basis; particle Filtering and Bayesian Reasoning are commonly used for estimating fault magnitudes and predicting future growth based on uncertain measurements and physical modeling (Arulampalam, Maskell, Gordon, & Clapp, 2002; Orchard, Kacprzynski, Goebel, Saha, & Vachtsevanos, 2008; Saha & Goebel, 2008; Sheppard, Butcher, Kaufman, & MacDougall, 2006).

The analytical approach to fault-adaptive control design that is introduced in this paper will assume that a non-empty space of current control actions to maintain system stability is known, and the controller must attempt to select control actions from that space to best manage the risk posed by degrading components. The future effects of control actions will be represented by generic probability distributions and control actions for best risk management will be derived by attempting to optimize an objective function that quantifies the relative aversion to the risk posed by further degrading components and the risk of degrading future system performance. Candidate metrics for evaluating risk from uncertain prognostic estimates may be drawn from the growing body of publications on vehicle health management (IVHM) (Srivastava, Mah, & Meyer, 2008); although, nearly all current studies in this area consider only end of life predictions in risk calculations and ignore data regarding short term fault growth, which will not be ideal in many cases. Literature on risk management in finance and actuarial science contain a rich array of tools that facilitate flexible risk-reward analysis on a continuous scale over a finite horizon. For example, Black-Scholes models (Lauterbach & Schulz, 1990) and value at risk (VaR) (Venkataraman, 1997) are prolific financial risk management tools that are also promising candidates for analyzing prognostic predictions (Schreiner, Balzer, & Precht, 2010).

This paper will explore the fundamental principles behind the derivation, verification, and validation of controls for optimal risk management on systems with incipient faults that grow in severity with increased component loading. The utility of various VaR based risk metrics for evaluating risk over a prognostic horizon, will be explored in a case-study on the use of prognostics-based load-allocation control for an unmanned ground vehicle (UGV).

2. Prognostics for Risk Management

The risk analysis process should begin with the definition of an analytical expression to evaluate the desirability of future control outcomes. In practice some form of scenario analysis should be used to derive and validate evaluation metrics though empirical studies (Abhken, 2000). Evaluation functions for future control outcomes represent the relative value of preserving nominal system performance and minimizing component degradations or failures for given scenarios.

A general form of an outcome evaluation function is

$$J_M\left(x\left(t\right)\right) + J_d\left(d_i^T\right),\ \{t = t_0..T\},\ \{i = 1, 2, ..N\} \quad (1)$$

where $x\left(t\right)$ represents the system state at time t, d_i^T represents the amount that component i has been degraded at the end of the mission, $J_M\left(x\left(t\right)\right)$ evaluates how well the system conformed to mission specifications and mission priorities over a mission that starts at $t = t_0$ and ends at $t = T$, and $J_d\left(d_i^T\right)$ evaluates the cost associated with the final state of degradation for each of the N components. The problem of specifying control actions to maximize this evaluation function will be referred to as the intrinsic optimization problem.

Due to uncertainty in the way faults grow with component loading and uncertainty regarding external operating conditions, it is generally impossible to design a controller that solves the intrinsic optimization problem directly; however, any control technique that claims to manage or mitigate the risk posed by load dependent fault modes can be viewed as being implicitly derived based on the optimization of an intrinsic cost function. The development of analytical tools that utilize knowledge of the intrinsic cost function in the design of prognostics-based fault-adaptive controllers will facilitate an understanding of the benefits of proposed approaches, as well as their fundamental limitations.

3. Component Load-Allocation

In a broad variety of systems the performance of the system and the growth of potential faults can be viewed as being direct functions of component loads. In general, the fault adaptive control problem can be fully understood in terms of a search for optimal performance and risk metrics that are evaluated on the space of allowable component loads over a given prognostic horizon. The space of allowable component load-allocations over a given prognostic horizon and the methods used to derive that space, will vary from one application to the next; however, many aspects of the fundamental searching problem will be invariant across a range of applications, facilitating the development of widely applicable analysis and control techniques.

The domain of allowable component load allocations over a given prognostic horizon will be defined at each control timestep using available system modeling and prognostic information to translate tolerances on performance degradations and fault growth risks into the component control domain. If the system is overactuated then the search for optimal component load allocations can be decomposed into two reduced-order sub-problems. An output control effort optimization will search for the optimal net system output control effort, and a restricted component load optimization will utilize any inherent overactuation in the system to find load allocations that minimize component degradations while providing the system output force requested by the output control effort optimization routine.

The separation of the component loading and system output regulation tasks will be shown for a generic nonlinear system,

$$\dot{x} = A(x) + B(x)\,\boldsymbol{u} \tag{2}$$

where $A(x) \in \mathbb{R}^n$, $B(x) \in \mathbb{R}^{n\times m}$, $\boldsymbol{x}(t) \in \mathbb{R}^n$, is the state, and $\boldsymbol{u}(t) \in \mathbb{R}^m$ is the control effort or load on each of the m components in the system. If $B(x)$ does not have full column rank, i.e, rank$\{B(x)\} = k < m\ \forall x$, then the system is overactuated, and $B(x)$ can be factorized as:

$$B(x) = B_\nu(x)\,B_u(x) \tag{3}$$

where $B_\nu(x) \in \mathbb{R}^{n\times k}$ and $B_u(x) \in \mathbb{R}^{k\times m}$ both have rank k. Now the system can be rewritten as:

$$\begin{aligned} \dot{\boldsymbol{x}} = &\quad A(x) + B_\nu(x)\,\boldsymbol{\nu} \\ \boldsymbol{\nu} = &\quad B_u(x)\,\boldsymbol{u} \end{aligned} \tag{4}$$

where $\boldsymbol{\nu}(t) \in \mathbb{R}^k$ can be interpreted as the net control effort produced by the m system components.

Because $B_\nu(t)$ has full column rank, a desired system output will uniquely determine the net output control effort, $\boldsymbol{\nu}(t)$ (using the pseudo inverse); however, since $B_u(x)$ has a nullspace of dimension $m-k$ there are available degrees of freedom in assigning component loads, $\boldsymbol{u}(t)$, for a given $\boldsymbol{\nu}$. Then component loads for best risk management can effectively be expressed as a function of $\boldsymbol{\nu}$, where any inherent redundancies in actuation, identified by the null space of B_u, are used to minimize component damages while still resulting in the net control effort commanded.

Practical applications of control allocation are currently found in aerospace (Gokdere, Bogdano, Chiu, Keller, & Vian, 2006; Karpenko & Sepehri, 2005) and automotive vehicles (Hattori, Koibuchi, & Yokoyama., 2002). A survey of efficient methods for determining the optimal control allocation for general linear and nonlinear systems is discussed in (Oppenheimer, Doman, & Bolender, 2006). Proof of the equivalence of this type of control allocation and optimal control is given in (Harkegard & Glad, 2005), for nonlinear systems with quadratic cost functions.

3.1 Load-Allocation as a Bounded Optimization

The objective of the general fault-adaptive control problem is to select the current component load allocations in an attempt to optimize the system's intrinsic cost function. In this work, component loads are allocated at the current control time-step by attempting to optimize an objective function that uses system modeling and fault prognostic information to quantify the expected trade-off between system performance and fault risk over a specified prognostic horizon. Constraints on allowable system performance and fault growth risk over a prognostic horizon will be enforced in the domain of allowable component load allocations in an attempt to satisfy minimum remaining-useful-life requirements for failing components.

The analysis presented in this document will use a fault risk metric of the following generic form:

$$\begin{gathered} f\Big(\tilde{d}_i(t+\tau)\Big),\ \text{Pr}\Big(d_i(t+\tau) > \tilde{d}_i(t+\tau)\Big) = \alpha, \\ \{i = 1, 2, ..N\},\ \text{given}\ \{\boldsymbol{u}_i(t) ... \boldsymbol{u}_i(t+\tau)\} \end{gathered} \tag{5}$$

where τ is the length of the prognostic horizon, $d_i(t)$ is the estimated degradation of component i at time t, $\tilde{d}_i(t+\tau)$ is a VaR estimate for component damage at the prognostic horizon, and $f(\tilde{d}_i(t+\tau))$ represents a risk metric that penalizes VaR estimates. VaR estimates are defined as the threshold damage such that the probability of the actual damage exceeding a given magnitude at a given future time equals α. Published literature contains relatively few examples of VaR being employed to manage the risk posed by incipient fault models; however, VaR is a standard risk assessment tool in finance, and it is powerful and widely applicable tool for risk management in systems with degrading components (Schreiner, Balzer, & Precht, 2008; Schreiner et al., 2010; Venkataraman, 1997).

A general form of the cost function used to represent the relative aversion to the risk of degrading future system performance and the risk posed by degrading components is

$$g(|\boldsymbol{\nu} - \boldsymbol{r}|)\,|_t^{t+\tau} + f\left(\tilde{d}_i(t+\tau)\right) \tag{6}$$

where $\boldsymbol{r}$ represents the desired net output control effort required for nominal performance and $g(|\boldsymbol{\nu} - \boldsymbol{r}|)$ penalizes performance degradation over a given prognostic horizon.

In published literature on prognostics for risk management there is a nearly ubiquitous use of expected remaining useful life (RUL) or expected time to failure (TTF) estimates to assess risk; however, in general, the methodology used to assess risk from fault prognosis information should be tailored to the system's expected use, its maintenance costs, the danger of potential failure modes, and the growth of uncertainty over a prognostic horizon. In this paper, the length of the prognostic horizon and the utility of various metrics for quantifying risk from prognostic predictions will be explored as a design choice. In cases where RUL or TTF based risk metrics are deemed most appropriate they can be realized as a special case of finite horizon prognosis, in which the prognostic horizon is extended until component failure is assured.

Constraints on allowable system performance are defined either in terms of a maximum deviation from commanded system states or a maximum deviation from the desired nominal system output force at a given time,

$$|\mathbf{y}_c - \mathbf{y}_o| \le \boldsymbol{\Delta}(t) \tag{7}$$

$$|\boldsymbol{\nu} - \mathbf{r}| \le \tilde{\boldsymbol{\Delta}}(t) \tag{8}$$

A finite horizon prognosis constraint will place an upper-bound on the probability that a component will become damaged by more than a specified amount over the prognostic horizon. This constraint is written as follows:

$$\text{Pr}\left(d_i(t+\tau) > \gamma_i(t+\tau)\,|\boldsymbol{u}_i(t)\right) \le \beta \tag{9}$$

where $\gamma_i\,(t+\tau)$ is the maximum allowable fault dimension at time $t+\tau$ and β is the upper bound on the probability that the fault dimension of component i is larger than its maximum allowable value at time $t+\tau$.

3.2 Verifying Constraint Feasibility

If the future performance requirements are known in advance, then the existence of feasible solutions to the optimal component load allocation problem can be verified by first finding the minimum allowable net output control effort needed to satisfy the performance constraints;

$$\tilde{\boldsymbol{\nu}} = \min_{\boldsymbol{u}_i} \{\boldsymbol{\nu}\}\,, \quad \text{s.t. } |\boldsymbol{\nu} - \mathbf{r}| = \tilde{\boldsymbol{\Delta}}\,(t)\,,\ \forall t \tag{10}$$

where $\tilde{\boldsymbol{\nu}}$ is the minimum allowable net output control effort under the performance constraint. Feasible solutions to the optimal load-allocation problem exist if there exists a distribution of component loads that result in $\tilde{\boldsymbol{\nu}}$ and do not violate the prognostic constraint at the end of the mission. This condition is written as follows:

$$\begin{aligned} &\Pr\left(d_i\,(T) > \gamma_i\,(T)\,|\boldsymbol{u}_i\,(t)\right) \le \beta, \\ &\qquad\qquad \text{s.t. } \tilde{\boldsymbol{\nu}} = B_u\,(x\,(t))\,\boldsymbol{u}, \quad t \in [t_0, T] \end{aligned} \tag{11}$$

4. UGV Application Example

Simulation studies for optimal load-allocation on a skid-steered UGV will demonstrate some of the fundamental properties of the proposed control methods. As shown in Figure 1, each of the wheels in a skid steered vehicle are fixed to the frame and are pointing straight forward. The system is over-actuated, as the four motors of the four-wheeled UGV are linked through their mutual contact with the ground. Assuming that all of the robot's wheels are getting approximately the same traction, then a skid-steered wheeled vehicle will behave much like a treaded vehicle. In the presented simulation studies the UGV's modeling is simplified by treating it as a treaded vehicle. The net output control effort of the modeled UGV is defined as follows:

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_f \\ \nu_{\dot{\phi}} \end{bmatrix} = \begin{bmatrix} T_1 + T_2 + T_3 + T_4 \\ T_1 + T_2 - T_3 - T_4 \end{bmatrix} = \begin{bmatrix} T_L + T_R \\ T_L - T_R \end{bmatrix} \tag{12}$$

where ν_f represents the net motive torque applied in the direction of travel, $\nu_{\dot{\phi}}$ represents the net turning torque, T_L is the sum of the motor torques on the left side of the robot, and T_R is the sum of the motor torques on the right side of the robot.

The UGV model is

$$\begin{aligned} M\dot{x} &= \ -C\,(x) + B \cdot u \\ y &= \begin{bmatrix} \frac{r}{2} & \frac{r}{2} \\ \frac{-r}{\alpha W} & \frac{r}{\alpha W} \end{bmatrix} x \\ x &= \begin{bmatrix} w_l \\ w_r \end{bmatrix} = \begin{bmatrix} w_1 \\ w_3 \end{bmatrix} = \begin{bmatrix} w_2 \\ w_4 \end{bmatrix} \\ u &= \begin{bmatrix} T_1 & T_2 & T_3 & T_4 \end{bmatrix}^T,\ y = \begin{bmatrix} \mathrm{v} \\ \dot{\phi} \end{bmatrix} \end{aligned} \tag{13}$$

$$\begin{aligned} M &= \begin{bmatrix} \frac{mr^2}{4} + \frac{r^2 I}{\alpha W^2} & \frac{mr^2}{4} - \frac{r^2 I}{\alpha W^2} \\ \frac{mr^2}{4} - \frac{r^2 I}{\alpha W^2} & \frac{mr^2}{4} + \frac{r^2 I}{\alpha W^2} \end{bmatrix} \\ B &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{aligned} \tag{14}$$

where the coefficients in this model are defined in Table 1. Note that the UGV model is linear except for a possibly non-linear fictional force, $C\,(x)$, and the system is overactuated, because the B matrix does not have full column rank.

In simulations linear kinetic friction will be used,

$$C\,(\hat{x}) = \begin{bmatrix} k/2 & k/2 \\ k/2 & k/2 \end{bmatrix} \tag{15}$$

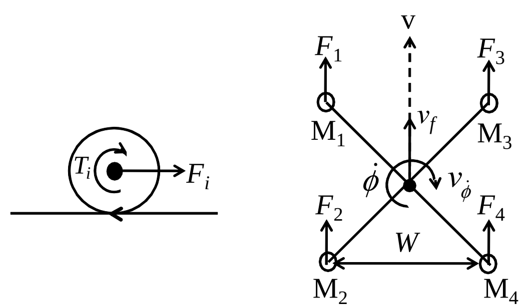


Figure 1. Visualization of motor torque allocation for a UGV

Symbol	Description	Units	Value
r	Wheel radius	m	0.1
W	Vehicle width	m	0.4
I	Wheel rotational inertia	kg·m^2	0.1
m	Vehicle mass	kg	1
$C\,(x)$	Frictional force	N	-
w_i	Wheel speed of motor i	rad/s	-
w_l	Left side wheel speed	rad/s	-
w_r	Right side wheel speed	rad/s	-
T_i	Torque produced by motor i	N·m	-
v	Vehicle speed	m/s	-
$\dot{\phi}$	Vehicle angular velocity	rad/s	-
α	Terrain-dependent parameter	-	-

Table 1. Definitions of symbols used in the UGV model

4.1 Prognostic Modeling

Winding insulation breakdown is a primary failure mechanism for the UGV's motors. The following model is used to estimate winding insulation lifetimes as a function of temperature,

$$L_N\,(t) = \alpha e^{-\beta T_W(t)} \tag{16}$$

where L_N is the expected remaining useful life (RUL) for new insulation in seconds and $T_W(t)$ $(^\circ C)$ is the winding temperature at time t (Montsinger, 1930).

The RUL estimate for a motor winding at any given time is calculated using:

$$L(t) = L_N(t) \cdot \left(1 - \frac{d(t)}{100}\right) \tag{17}$$

where $d(t)$ is the percentage of insulation lifetime used prior to time t,

$$d(t) = \int_0^t \frac{d\tau}{L(\tau)} \tag{18}$$

A probability distribution is added to the α coefficient in Eq. (16) to capture uncertainty in the prognostic model. Figure 2 shows the resulting probabilistic insulation life versus temperature model, where the pdf's mean corresponds to $\alpha = 10^5\ (s)$, and standard deviations are given by $\alpha = 10^5 \pm 1.5 \times 10^4$. The β coefficient in Eq. (16) is set to $0.035\,(°C^{-1})$.

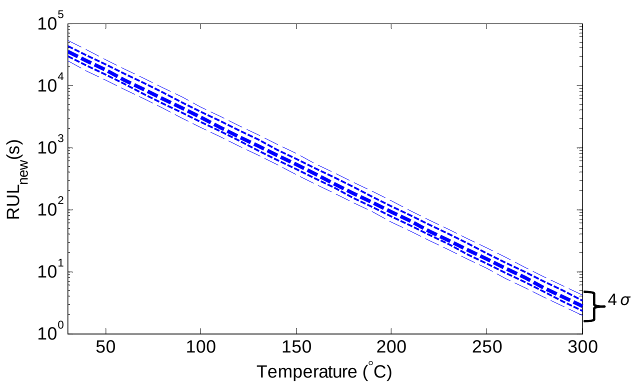


Figure 2. Addition of an uncertainty pdf to the insulation breakdown model.

Thermal Model

A first order thermo-electrical model, shown in Figure 3, is used to track the winding-to-ambient temperature as a function of copper losses,

$$\dot{T}_{wa} = -\frac{T_{wa}(t)}{R_{wa}C_{wa}} + \frac{P_{loss}(t)}{C_{wa}} \tag{19}$$

where T_{wa} is winding-to-ambient temperature, P_{loss} is power loss in the copper windings, C_{wa} is thermal capacitance, and R_{wa} is thermal resistance.

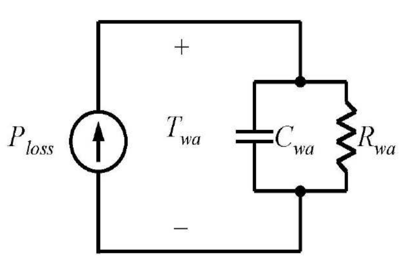


Figure 3. Thermal model for motor windings

5. SIMULATION STUDIES

In simulation studies of the UGV load-allocation problem, the intrinsic optimization problem, introduced in Eq. (1), is defined using the following performance and component degradation penalties:

$$J_M(\boldsymbol{\nu}) = \frac{1}{T}\int_0^T \exp|\phi_c(t) - \phi(t)| + K_{p1} \tag{20}$$

$$J_d(\boldsymbol{u}) = \max_i \left[\exp\left(\tilde{d}_i(T)\right) \cdot \frac{4}{3}\right] + K_{p2} \tag{21}$$

where $\boldsymbol{\phi}_c(t)$ and $\boldsymbol{\phi}(t)$ represent waypoints for the desired and actual path followed by the UGV respectively. K_{p1} and K_{p2} are penalty functions that effectively enforce constraints on the maximum acceptable path error and the maximum acceptable VaR estimate at the end of a mission. Performance and component degradation constraints are defined as follows:

$$|\boldsymbol{\phi}_c(t) - \boldsymbol{\phi}(t)| < 1, \quad \forall t \in [0, .., T] \tag{22}$$

$$\tilde{d}_i(T) < 90\,\% \tag{23}$$

The performance and prognostic penalties introduced in Eq. (5) and Eq. (6) are defined as follows:

$$g(|\boldsymbol{\nu} - \boldsymbol{r}|) = \int_t^{t+\tau} \left|\exp\left(\begin{bmatrix} r_f - \nu_f \\ r_{\dot{\phi}} - \nu_{\dot{\phi}} \end{bmatrix}\right)\right| dz \tag{24}$$

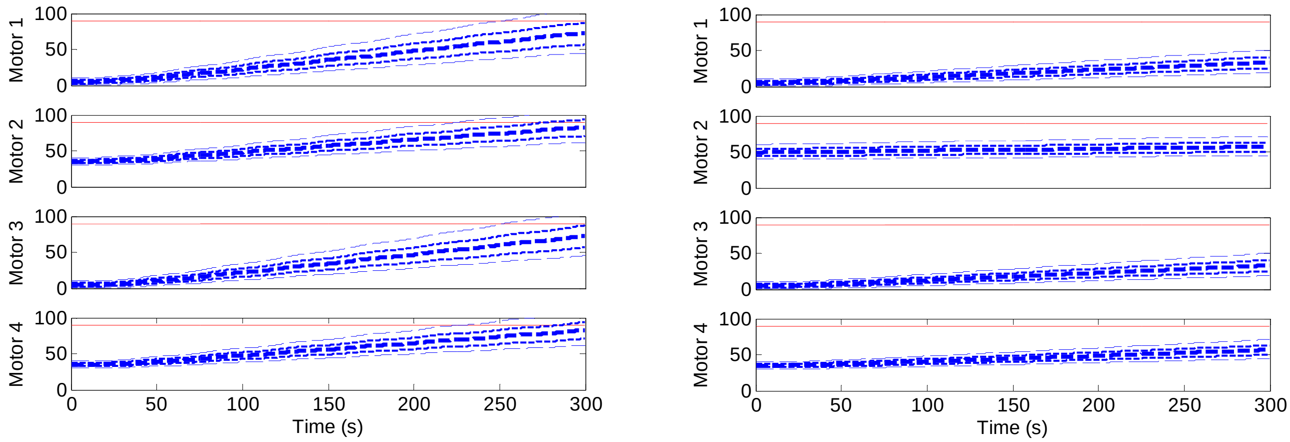
$$f\left(\tilde{d}_i(t+\tau)\right) = \lambda \cdot \sum_{i=1}^{4} \Big[\exp\left(\tilde{d}_i(t+\tau) - \gamma_i(t+\tau)\right) + \exp\left(\tilde{d}_i(t+\tau)\right) + C_p\Big] \tag{25}$$

$$\Pr\left(d_i(t+\tau) > \tilde{d}_i(t+\tau)\right) = 2\% \tag{26}$$

where λ represents the relative value of maximizing performance and minimizing component degradations, $\gamma_i(t+\tau)$ is an upper-bound on the 98% confidence VaR estimates at time $t+\tau$, and C_p is an additional penalty that effectively disallows controls that cause the upper VaR bound to be exceeded (if other solutions exist). Simulation studies presented later in this paper will explore the effect of varying τ and λ on the system's intrinsic evaluation function. In the reported simulation studies, $\gamma_i(t+\tau)$ is defined using a linear interpolation from $\tilde{d}_i(t)$ to the maximum allowable degradation at the end of the mission. The effect of varying the formulation of $\gamma_i(t+\tau)$ on load-allocation in a triplex redundant electro-mechanical, was explored in a previous publication (Bole et al., 2010).

In simulations, the cost of possible motor load allocations is evaluated by assuming that the current demands on the system and the current component load allocations are constants over the prognostic horizon. The space of feasible motor-load allocations to be searched over is defined by the following performance constraint:

$$\begin{array}{lcl} 0.8 \cdot r_L & \leq T_L \leq & 1.2 \cdot r_L \\ 0.8 \cdot r_R & \leq T_R \leq & 1.2 \cdot r_R \end{array} \tag{27}$$



(a) Desired UGV path

(b) Min performance UGV path

Figure 4. Plots of winding insulation degradation estimates (mean and $\pm$ 2 standard deviations) on the desired UGV path (a) and the minimum allowable performance path (b), using λ & τ such that $J_d\left(\boldsymbol{u}\left(t\right)\right)|_{\lambda,\tau} = \min_{\lambda,\tau}[J_d\left(\boldsymbol{u}\left(t\right)\right)]$

where r_L and r_R are desired net control effort outputs from the left-hand and right-hand motors respectively. The desired torque output from the UGV at a given time instant is defined by the following proportional control law:

$$\boldsymbol{r} = \begin{bmatrix} r_L + r_R \\ r_L - r_R \end{bmatrix} = \boldsymbol{r}_{\text{ref}}\left(t\right) + \begin{bmatrix} p_1 \cdot \cos\left(\phi_e\right) \cdot e_d \\ p_2 \cdot \sin\left(\phi_e\right) \end{bmatrix} \quad (28)$$

where $\boldsymbol{r}_{\text{ref}}\left(t\right)$ is output control effort that would be used at time t if the vehicle followed the reference path exactly, p_i are the proportional control coefficients, ϕ_e is the vehicle's heading error with respect to the reference path, and e_d is the vehicle's position error with respect to the reference path. Component load allocations for best risk management are found at each time-step by evaluating the objective function on a sufficiently dense uniform grid over the space of all component load allocations satisfying the performance constraints.

5.1 Verifying Mission Feasibility

In the simulation studies discussed here, the four-wheeled UGV is commanded to follow a figure-8 type path. By design, the commanded path is so demanding that following it exactly will yield no solutions to the load-allocation problem that satisfy the final VaR constraint (defined in Eq. (23)). The existence of solutions to the load-allocation problem that will not violate the performance and VaR constraints for the given mission is proven by verifying that using the minimum allowable UGV performance over the mission will allow all motors to end the mission with adequate health. Figure 4 shows simulation results for load-allocation controls that minimize the final VaR evaluation metric (defined in Eq. (21)) on the minimum allowable UGV performance path and the desired UGV path. As shown in the figure, the motors on each side of the vehicle are initialized at different levels of degradation in order to observe discrimination in the allocation motor loads based on their relative healths. The simulation results, shown in Figure 4, prove that although following the desired UGV path exactly is guaranteed to result in violation of the final VaR constraint, the load-allocation problem does have feasible solutions satisfying both the performance and VaR constraints.

5.2 Control with Foreknowledge of the Mission and the Fault Growth Model

Due to the fact that in simulation studies the desired path for the UGV and a fault growth model are known in advance, the optimal load allocations over the given mission can be approximated without the need for prognosis. Analysis of the direct optimization of the system's intrinsic cost function over a known mission will provide substantial insight into the development of prognostics-based risk-management controllers. Optimization routines will specify candidate UGV paths over a mission by defining a set of waypoints and using a third order spline to interpolate between those points.

The search space for the path planning routines is the set all adjustments to given waypoints that will not violate the performance constraint, given in Eq. (22). The net output control effort output required to follow a given path is found by inverting the modeled UGV dynamics given in Eq. (13),

$$\begin{bmatrix} T_L\left(t\right) \\ T_R\left(t\right) \end{bmatrix} = f^{-1}\left(\boldsymbol{\phi}_p\left(t\right)\right), \quad \forall t \in [0,..,T] \quad (29)$$

where $\boldsymbol{\phi}_p\left(t\right)$ is the (x,y) position of the UGV at time t.

Individual motor load allocations are derived using the following expression for splitting load proportionately among

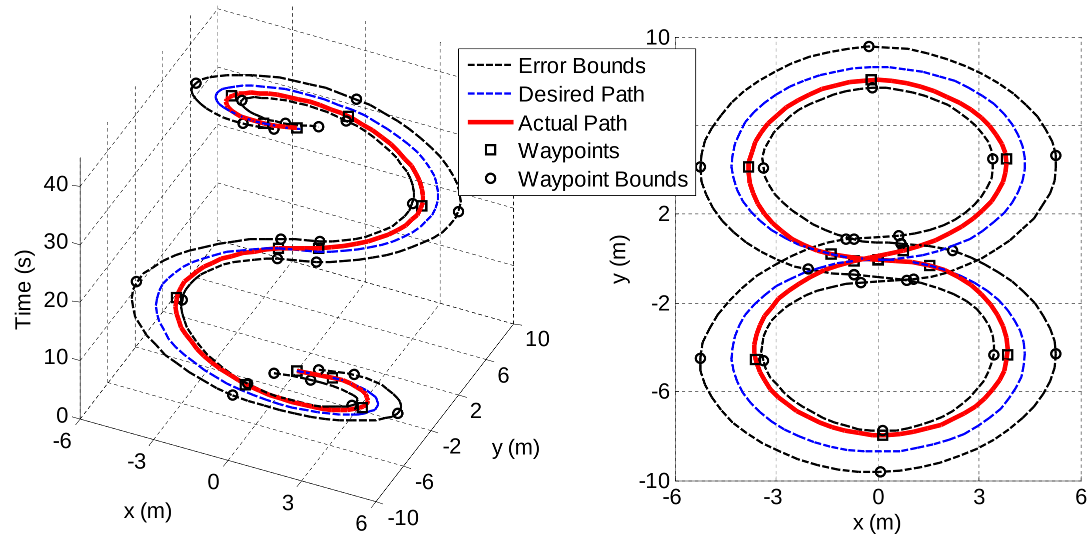


Figure 5. Results of search for optimal figure-8 path

	$\tilde{d}_1(T)$	$\tilde{d}_2(T)$	$\tilde{d}_3(T)$	$\tilde{d}_4(T)$	$J_d\|_0^T$	$J_M\|_0^T$	$(J_M + J_d)\|_0^T$
min allowable performance path	49%	73%	49%	73%	2.76	2.16	4.93
load-allocation with future knowledge	64%	64%	64%	64%	2.52	1.69	4.21
prognosis based load-allocation	70%	86%	70%	86%	3.17	1.58	4.76

Table 2. Results of simulation studies

the two motors on each side of the vehicle:

$$\begin{aligned} T_1(t) \cdot k_1 + T_2(t) = T_L(t) \\ T_3(t) \cdot k_2 + T_4(t) = T_R(t) \end{aligned} \tag{30}$$

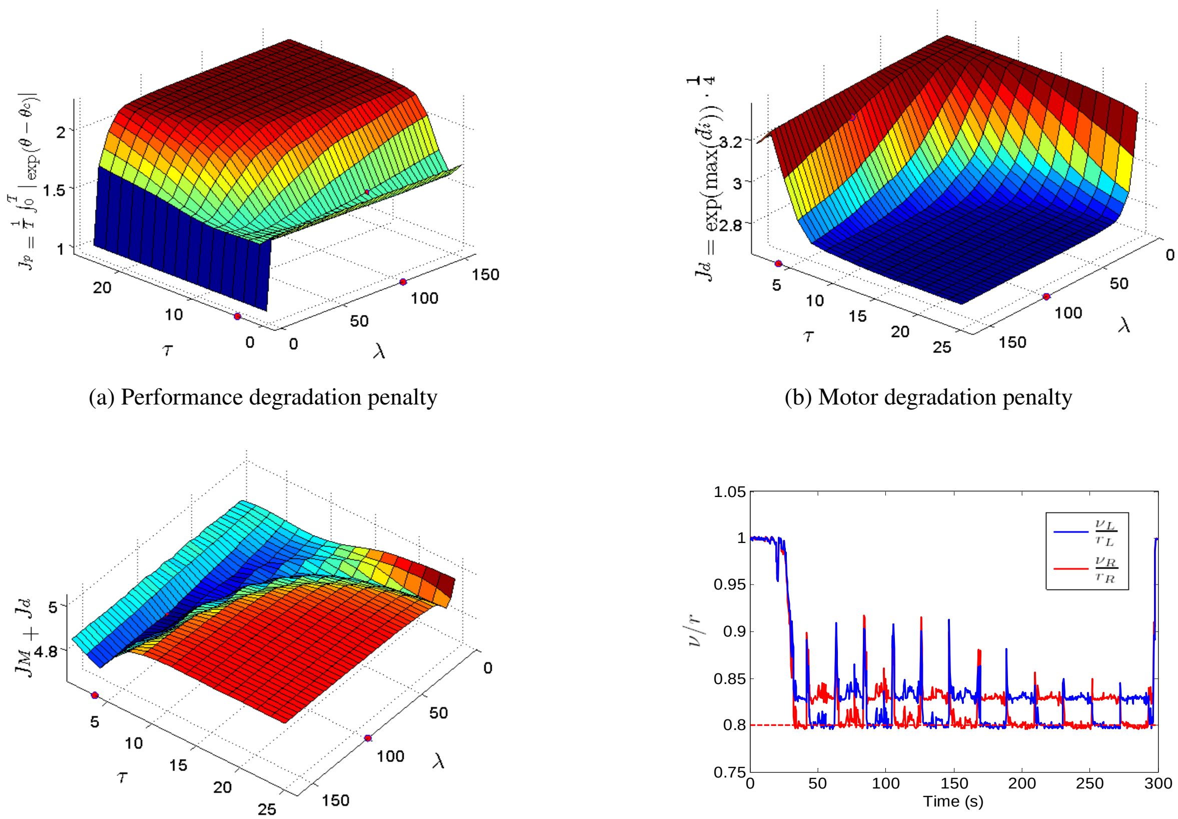
Optimal motor load allocations for a given path are derived by evaluating Eq. (1) over sufficiently dense uniform grid on k_1 and k_2, and selecting the value resulting in minimum cost.

Figure 5 shows plots of the desired UGV path, the bounds on allowable path error, and an approximation of the optimal UGV path, for one cycle of the commanded figure-8 maneuver. The simulated mission consists of eight repetitions of this figure-8 maneuver. A nested optimization is used to estimate the optimal motor load allocations over the given mission. An outer-loop optimization routine uses a gradient descent search over the space of allowable adjustments to a set of waypoints, where the space of allowable adjustments to each waypoint is shown in Figure 5 as the linear region between the black circles. An inner loop optimization routine finds the net output torque from the left-hand and right-hand motors required to follow eight repetitions of a given figure-8 path, and then searches for the optimal proportional load split among the motors on each side of the vehicle using the uniform grid method described earlier.

Estimates of the optimal VaR metrics for the winding insulation degradations and the control evaluation costs for the given mission are shown in Table 2. Note that the estimated optimal motor load-allocations will result in final winding VaR estimate being nearly equal for all four motors, due the fact that the control evaluation function is defined to penalize only the highest motor degradation. Also, note that the error between the commanded and the estimated optimal path is greatest in the extreme upper and lower regions of the figure-8 path because introducing an error in those regions results in the greatest reduction in the total distance traveled by the UGV. Both of these results are expected when the future commanded UGV path and the future fault grown model are known in advance; however, in general, it will be very difficult to match those results with controllers that rely on uncertain predictions of future states.

5.3 Prognostics-Based Control

At each control time-step, a prognostics-based controller will allocate motor loads to best manage the risk posed by uncertain estimates of future system performance and fault prognosis. In simulation, motor load-allocations for best risk management are derived by evaluating Eq. (6) on a sufficiently dense uniform grid over the space of all motor loads satisfying the performance constraint. Fundamentally, the prognostics-based control problem is to specify risk-reward evaluation metrics, of the form given in Eq. (6), that will result in the derived controls coming as close as possible to matching the minimum control evaluation metric achievable using foreknowledge. Figure 6 shows plots the intrinsic evaluation



(a) Performance degradation penalty

(b) Motor degradation penalty

(c) Net evaluation metric

(d) Net control effort outputs for best computed λ & τ

Figure 6. UGV simulation results for $\lambda = [0, 3, 6, ..., 160]$ and $\tau = [.4, 1.6, 2.8, ..., 24.4]$; optimal value at $\lambda = 100$ & $\tau = 3.6$s

metrics over a range of values for the prognostic horizon, τ, and the weighting factor, λ. The intrinsic evaluation metrics, shown in the plots, were obtained by computing the optimal motor load-allocations at each control time-step, after substituting τ and λ into the evaluation functions for predicted future component degradations and system performance, which were defined in Eq. (24) and Eq. (25). In general, as the prognostic horizon is increased the increased uncertainty in fault growth predictions will result in a greater perceived risk, and thus a more conservative control. Also, increasing the weighting factor, λ, on the prognostic penalty will tend to result in solutions with higher path errors and less component degradations. Plots of the intrinsic evaluation metrics versus λ and τ show these general trends. The trough seen in Figure 6 (c) indicates a domain of τ and λ values corresponding to controls that are neither overly conservative nor overly aggressive. The best computed intrinsic control evaluation costs and the corresponding winding insulation degradation VaR's at the end of the mission are given in Table 2. Future work will continue to explore the analytical relationships between the metrics used to evaluate risk from prognostic estimates and their resultant performance on example systems.

6. Conclusion

Any control technique that claims to manage or mitigate the risk posed by load dependent fault modes can be viewed as being implicitly derived based on the risk-reward optimization that was explicitly addressed in this work. The paper introduced a methodology for deriving and validating prognostics-based fault-adaptive control routines that began with the derivation of an expression for evaluating the desirability of future control outcomes, and eventually produced control routines that sought to optimize derived risk metrics using uncertain prognostic information. A case study on the design of fault-adaptive control for a skid-steered robot demonstrated some of the challenges associated with deriving risk metrics that will minimize the risk of component failures without becoming overly conservative and unnecessarily sacrificing performance. Future work will introduce more sophisticated methods for utilizing stochastic prognostic information to associate risk with a given distribution of component loads; more sophisticated methods for solving the resulting stochastic optimization problems will also be explored.

ACKNOWLEDGMENT

This work was supported by the United States National Aeronautics and Space Administration (NASA) under STTR Contract #NNX09CB61C. The program is sponsored by the NASA Ames Research Center (ARC), Moffett Field, California 94035. Dr. Kai Goebel, NASA ARC, Dr. Bhaskar Saha, MCT, NASA ARC and Dr. Abhinav Saxena, RIACS, NASA ARC, are the Technical POCs and Liang Tang is the Principal Investigator and Project Manager at Impact Technologies.

REFERENCES

Abhken, P. (2000). An empirical evaluation of value at risk by scenario simulation. *Journal of Derivatives*, *7*, 1074-1240.

Arulampalam, S., Maskell, S., Gordon, N., & Clapp, T. (2002, Feb.). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, *50*(2), 174-188.

Bole, B., Brown, D., Pei, H.-L., Goebel, K., Tang, L., & Vachtsevanos, G. (2010, Oct.). Fault adaptive control of overactuated systems using prognostic estimation. In *International conference on prognostics and health management (PHM).*

Caglayan, K., Allen, S., & Wehmuller, K. (1988). Evaluation of a second generation reconfiguration strategy for aircraft flight control systems subjected to actuator failure surface damage. In *Proceedings of the IEEE national aerospace and electronics conference* (p. 520-529).

Doyle, J., Glover, K., Khargonekar, P., & Francis, B. (1987). State-space solutions to standard h_2 and h_∞ optimal control problems. *IEEE Transactions on Automatic control*, *33*, 831-847.

Gergely, E., Spoiala, D., Spoiala, V., Silaghi, H., & Nagy, Z. (2008). Design framework for risk mitigation in industrial PLC control. In *IEEE international conference on automation, quality and testing, robotics* (Vol. 2, p. 198-202).

Gokdere, L., Bogdano, A., Chiu, S., Keller, K., & Vian, J. (2006). Adaptive control of actuator lifetime. In *IEEE aerospace conference.*

Harkegard, O., & Glad, T. (2005, Jan.). Resolving actuator redundancy - optimal control vs. control allocation. *Automatica*, *41*(1), 137-144.

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.*

Karpenko, M., & Sepehri, N. (2005, January). Fault tolerant control of a servohydraulic positioning system with crossport leakage. *IEEE Transactions on Control Systems Technology*, *13*(1), 155-161.

Lauterbach, B., & Schulz, P. (1990). Pricing warrants: An empirical study of the black-scholes model and its alternatives. *Journal of Finance*, *45*, 1181-1209.

Monaco, J., Ward, D., & Bateman, A. (2004, Sept.). A retrofit architecture for model-based adaptive flight control. In *AIAA 1st intelligent systems technical conference.*

Montsinger, V. M. (1930). Loading transformers by temperature. *Transactions of the American Institute of Electrical Engineers*, *32*.

Oppenheimer, M., Doman, D., & Bolender, M. (2006). Control allocation for over-actuated systems. In *14th mediterranean conference on control and automation.*

Orchard, M., Kacprzynski, G., Goebel, K., Saha, B., & Vachtsevanos, G. (2008). Advances in uncertainty representation and management for particle filtering applied to prognostics. In *International conference on prognostics and health management PHM.*

Rao, B. (1998). *Handbook of condition monitoring* (A. Davies, Ed.). Chapham and Hall.

SAE. (1994). *Potential failure mode and effects analysis in design (design FMEA) and potential failure mode and effects analysis in manufacturing and assembly processes, reference manual* (Tech. Rep. No. J1739).

Saglimbene, M. (2009). Reliability analysis techniques: How they relate to aircraft certification. In *Reliability and maintainability symposium* (p. 218-222).

Saha, B., & Goebel, K. (2008). Uncertainty management for diagnostics and prognostics of batteries using bayesian techniques. In *IEEE aerospace conference.*

Schreiner, A., Balzer, G., & Precht, A. (2008). Risk analysis of distribution systems using value at risk methodology. In *Proceedings of the 10th international conference on probabilistic methods applied to power systems.*

Schreiner, A., Balzer, G., & Precht, A. (2010). Risk sensitivity of failure rate and maintenance expenditure: application of var metrics in risk management. In *15th IEEE mediterranean electrotechnical conference* (p. 1624-1629).

Sheppard, J., Butcher, S., Kaufman, M., & MacDougall, C. (2006). Not-so-naïve bayesian networks and unique identification in developing advanced diagnostics. In *IEEE aerospace conference* (p. 1-13).

Srivastava, A., Mah, R., & Meyer, C. (2008, Dec.). *Integrated vehicle health management automated detection, diagnosis, prognosis to enable mitigation of adverse events during fight* (Tech. Rep.). Version 2.02, National Aeronautics and Space Administration.

Venkataraman, S. (1997). Value at risk for a mixture of normal distributions: the use of quasi-bayesian estimation techniques. *Economic Perspectives*, 2-13.

Brian M. Bole graduated from the FSU-FAMU School of Engineering in 2008 with a B.S. in Electrical and Computer Engineering and a B.S. in Applied Math. He received a M.S. degree in Electrical Engineering from the Georgia Institute

of Technology in 2011, and he is currently pursuing a Ph.D. Brian's research interests include stochastic optimization, robust control, fault prognosis, and risk management. Brian is currently instigating the use of risk management and stochastic optimization techniques for optimal adaptation of active component load allocations in robotic and aviation applications. In a previous project, Brian work with the Georgia Tech EcoCar team to develop a stochastic-optimization based controller for optimizing fuel economy on a charge sustaining hybrid electric vehicle.

Kai Goebel received the degree of Diplom-Ingenieur from the Technische Universität München, Germany in 1990. He received the M.S. and Ph.D. from the University of California at Berkeley in 1993 and 1996, respectively. Dr. Goebel is a senior scientist at NASA Ames Research Center where he leads the Diagnostics & Prognostics groups in the Intelligent Systems division. In addition, he directs the Prognostics Center of Excellence and he is the Associate Principal Investigator for Prognostics of NASA's Integrated Vehicle Health Management Program. He worked at General Electric's Corporate Research Center in Niskayuna, NY from 1997 to 2006 as a senior research scientist. He has carried out applied research in the areas of artificial intelligence, soft computing, and information fusion. His research interest lies in advancing these techniques for real time monitoring, diagnostics, and prognostics. He holds ten patents and has published more than 100 papers in the area of systems health management.

Liang Tang is a Lead Engineer at Impact Technologies LLC, Rochester, NY. His research interests include diagnostics, prognostics and health management systems (PHM), fault tolerant control, intelligent control, and signal processing. He obtained a Ph.D. degree in Control Theory and Engineering from Shanghai Jiao Tong University, China in 1999. Before he joined Impact Technologies, he worked as a post doctoral research fellow at Intelligent Control Systems Laboratory, Georgia Institute of Technology. At Impact Technologies he is responsible for multiple DoD and NASA funded research and development projects on structural integrity prognosis, prognostics and uncertainty management, automated fault accommodation for aircraft systems, and UAV controls. Dr. Tang has published more than 30 papers in his areas of expertise.

George J. Vachtsevanos is a Professor Emeritus of Electrical and Computer Engineering at the Georgia Institute of Technology. He was awarded a B.E.E. degree from the City College of New York in 1962, a M.E.E. degree from New York University in 1963 and the Ph.D. degree in Electrical Engineering from the City University of New York in 1970. He directs the Intelligent Control Systems laboratory at Georgia Tech where faculty and students are conducting research in intelligent control, neurotechnology and cardiotechnology, fault diagnosis and prognosis of large-scale dynamical systems and control technologies for Unmanned Aerial Vehicles. His work is funded by government agencies and industry. He has published over 240 technical papers and is a senior member of IEEE. Dr. Vachtsevanos was awarded the IEEE Control Systems Magazine Outstanding Paper Award for the years 2002-2003 (with L. Wills and B. Heck). He was also awarded the 2002-2003 Georgia Tech School of Electrical and Computer Engineering Distinguished Professor Award and the 2003-2004 Georgia Institute of Technology Outstanding Interdisciplinary Activities Award.