Certification Analysis for a Model-Based UAV Fault Detection System

Bin Hu[*] and Peter Seiler[†]

Department of Aerospace Engineering & Mechanics

University of Minnesota, Minneapolis, MN, 55455, USA

Model-based fault detection algorithms can be used to improve the reliability of unmanned aerial vehicles (UAVs) while still satisfying their restrictive size, power, and weight requirements. However, the use of model-based algorithms introduces new failure modes that do not exist in physically redundant architectures. Hence a certification process is needed for such systems that incorporates analysis tools, high fidelity simulations, and flight test data. This paper focuses on one aspect of such a process: the use of flight test data to validate theoretical analysis results. Specifically, this validation is performed to assess the false alarm probability of a simple, model-based UAV fault detection system. This example highlights the main certification issues that arise due to limited flight data and stringent reliability requirements. In addition, the flight test data shows non-Gaussian statistical behavior that leads to some discrepancies with the analysis results. Further discussions are presented for this observed behavior.

Nomenclature

P_F	Probability of false alarm
P_D	Probability of detection
v	Lateral velocity, m/s
p	Roll rate, rad/s
r	Yaw rate, rad/s
ϕ	Bank angle, rad
ψ	Yaw angle, rad
δ_{ail}	Aileron deflection, rad
δ_{rud}	Rudder deflection, rad

I. Introduction

Safety-critical aerospace systems must be designed for extremely high levels of reliability. In addition, there is a need to certify the reliability of the system design with aviation authorities, e.g. the Federal Aviation Administration (FAA) in the United States or the European Aviation Safety Agency. The system reliability and safety requirements for commercial flight control electronics are typically on the order of no more than 10^{-9} catastrophic failures per flight hour.[1,2] The aircraft industry uses designs that are based almost exclusively on physical redundancy, e.g. the Boeing 777 flight control system has multiple redundant computing processors, actuators, and sensors.[3,4] In a physically redundant configuration, a failed component is detected by directly comparing the behavior of each redundant component. Hence, these schemes tend to detect faults accurately, and their performance is relatively straightforward to certify using fault trees.[5,6]

The Modernization and Reform Act of 2012 requires the FAA to integrate unmanned aircraft in the national airspace. In particular, Section 332 of HR658 requires the FAA to "provide for the safe integration

[*] PhD Candidate, huxxx221@umn.edu

[†] Assistant Professor, seile017@umn.edu

of civil unmanned aircraft systems into the national airspace system as soon as practicable, but not later than September 30, 2015." This creates new design challenges as UAVs typically cannot afford the full payload associated with physically redundant architectures due to their more restrictive size, power, and weight requirements. One alternative to physical redundancy is model-based fault detection and isolation (FDI).[7–9] Model-based methods would enable some physically redundant components to be replaced with model-based fault detection thus leading to a dramatic reduction in the system size, weight, and power consumption. There have already been some efforts to implement analytical redundancy on commercial manned aircraft, e.g. the oscillatory monitors on the Airbus A380.[10] These considerations also motivated the recent AddSafe project in Europe[11] to assess the suitability of more advanced fault detection and isolation (FDI) methods for manned commercial aircraft.

Model-based FDI is an important technology to enable safe integration of UAVs within the national airspace. However, analytically redundant systems must rigorously demonstrate the required levels of reliability to certification authorities before model-based FDI finds wide acceptance for UAV applications. The use of models introduces new failure modes that do not exist in physically redundant architectures, e.g. the system may fail due to incorrect detection of faults arising due to model uncertainty. Thus the standard approaches to assess the reliability of physically redundant architectures, e.g. fault trees, can not be directly applied for systems that use analytical redundancy. A certification process is needed that incorporates theoretical analyses, high fidelity nonlinear simulations, and flight tests. Flight tests are commonly used to assess the system-level design prior to aircraft entry into service. However, flight tests alone are insufficient to assess the reliability of a safety-critical systems because catastrophic failures are extremely rare events by design. Moreover, the flight data is typically limited due to cost and time constraints. It is important to complement flight test data with high fidelity simulations and theoretical analyses. One approach is to linearize the system at many different trim conditions within the flight envelope and validate the reliability of the FDI system at each trim condition using theoretical linear analysis tools. High fidelity Monte Carlo simulations can be performed to complement the linear analyses and investigate the performance during transient (non-trim) conditions. The high fidelity simulations and the linear analyses both assume certain underlying models for the aircraft dynamics, hardware failure modes and the environmental conditions. The limited flight data can be used to validate these assumptions and the results obtained by simulations and linear analyses. This approach is similar to the existing procedure to validate the robustness and performance of flight control laws.[12]

This paper focuses on the use of limited flight test data to validate the results of linear analyses of FDI performance. In particular, linear probabilistic analyses can provide estimates of key FDI performance metrics including false alarm and detection probabilities. This paper uses experimental flight test data to validate a theoretical false alarm analysis and discusses the gaps between theory and experiments. Section II describes the flight test experiment, sets up a simple parity equation fault detection system, and defines the false alarm and detection probability metrics for quantifying FDI performances. This section also describes the major issues in using flight test data to validate the linear analysis results. Section III reviews theoretical bounding methods for false alarm analyses of linear FDI systems and then introduces a framework to validate the linear analyses based on flight test data. Section IV applies the validation framework to the UAV FDI system introduced in Section II and provides both results and discussions. Gaps between the theoretical results and flight data including the heavy tail phenomenon are discussed.

II. Problem Formulation

This section formulates the validation problem. First, UAV flight test data is presented with and without simulated faults in the aileron and roll rate sensor. Next, a simple parity-equation based FDI algorithm is described to detect the roll rate or yaw rate sensor fault. The performance metrics for the FDI system are defined in terms of the false alarm and detection probabilities. A framework to validate the performance of the FDI design using theoretical analysis, nonlinear simulations and limited flight test data is proposed. This paper focuses on the use of flight test data to validate theoretical analyses.

A. Flight Test Experiment

The University of Minnesota (UMN) UAV Research Group[13] has developed several low-cost experimental platforms. The flight test experiment described in this paper is performed with an Ultra Stick 25e UAV

(Figure 1). This UAV, referred to as Thor in the remainder of the paper, is a commercially available, fixed-wing, radio-controlled aircraft. Thor has a wing span of 1.27m, mass of 1.9kg, cruise speed of 17m/s, and endurance of 15-20min. This aircraft is one of primary flight test vehicles operated by the UMN UAV research group and additional details on this research infrastructure can be found in survey papers.[14–16] There are two important aspects that deserve further comments. First, the research group supports open-source development with all design information and flight test data available online through the UMN UAV Research Group website:[13] `www.uav.aem.umn.edu`. Second, a high fidelity simulation model has been developed for Thor with an aerodynamic model derived using frequency domain system identification techniques based on flight test data.[17,18]

Figure 1. University of Minnesota Ultra Stick 25e UAV.

Figure 2 shows experimental flight test results performed using Thor. The figure shows the response of the aircraft lateral dynamics (bank angle ϕ, roll rate p, yaw rate r) to a series of bank angle step commands for nominal (unfaulted) and faulted scenarios. The faulted flight test scenario consists of simulated aileron and roll rate faults. The aileron fault is an additive ramp starting at $t = 7sec$ and increasing to a maximum fault value of $5deg$ at $t = 20sec$. The roll rate sensor fault is a step bias of -80 deg/sec injected at $t = 12sec$. The flight control computer operates at 50 Hz corresponding to a sampling time of 0.02 sec. Both flight tests used the baseline lateral controller developed by the UMN UAV research group.[14] This baseline lateral controller consists of yaw rate feedback to the rudder, roll rate feedback to the aileron, and an outer-loop bank angle tracking controller. The details of the baseline controller are included in the appendix. The baseline controller trimmed Thor near a trim condition so that the closed-loop system of the aircraft can be approximated as a linear system. The unfaulted flight test scenario was repeated 6 times and the faulted scenario was repeated 3 times.

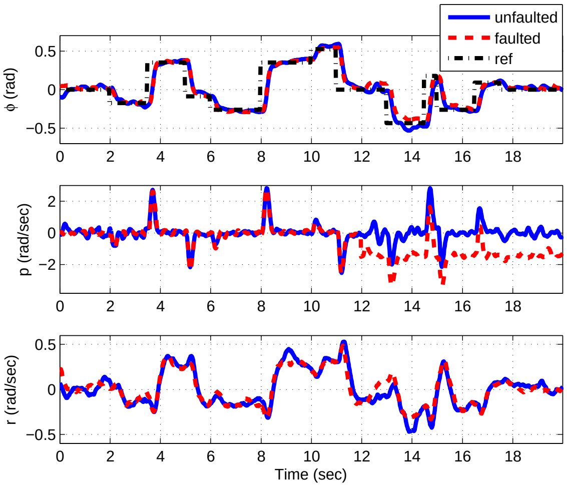

Figure 2. Flight test data for unfaulted and faulted scenarios.

The flight test results for the faulted scenario indicate that the aileron ramp fault injected at $t = 7sec$ has minimal impact due to the compensation by the baseline lateral controller. The roll rate sensor fault injected at $t = 12sec$ is apparent in the second subplot of Figure 2 but it also has a small impact on the bank angle tracking performance. This is again due to the compensation by the baseline lateral controller. Although the baseline controller is robust to both injected faults it is still important to detect the failures.

The controller is robust to constant sensor faults in the sense that these faults do not affect stability or the steady-state tracking as long as the system could be approximated as linear systems at trim conditions. However, a sufficiently large fault could cause a large deviation from the trim condition so that the nonlinear dynamics may dominate the aircraft and even destabilize the closed-loop system. Moreover, the closed-loop system may not be robust to transients faults rather than constant bias faults. Due to the high risk it is difficult to perform flight test with the faults which could potentially destabilize the aircraft. The current faulted scenario fits the purpose of FDI research. Study on the current flight data provides insight into the fault detection problem with more complicated faults. The next subsection describes a simple closed-loop parity equation fault detection system to detect the roll rate or yaw rate sensor fault.

B. Parity Equation Fault Detection System

Figure 3 shows a parity-equation fault detection scheme to detect the roll rate sensor fault discussed in the previous subsection. G_θ denotes the monitored closed-loop system which includes the lateral/directional dynamics and the baseline lateral control system. This paper considers an analysis in discrete-time. Hence the subscript $\theta(k) \in \{0, 1\}$ denotes the status of the system at time k: $\theta(k) = 0$ if the roll rate sensor is operational and $\theta(k) = 1$ if a sensor fault has occurred. The closed-loop dynamics may also be uncertain as denoted in the figure by the linear fractional dependence on Δ. The input to the closed-loop system is the bank angle reference command, ϕ_{ref}. The output of G_θ, denoted y, refers to the roll rate p or yaw rate r. The measured output, y_{meas}, includes the effects of sensor noise as well as any stochastic disturbances acting on the aircraft, e.g. wind gusts. The combined stochastic effects of sensor noise and disturbances can be modeled at the output of the closed-loop system through appropriate block diagram manipulation. Specifically, the stochastic effects are modeled with an additive output signal of a linear system M driven by an independent and identically distributed (IID) Gaussian process n. M will introduce time correlations into the noise process, n. The dashed box in Figure 3 represents the entire picture of the closed-loop lateral/directional dynamics control system including model uncertainty, stochastic disturbances, and fault. A parity equation fault detection scheme is used to monitor the status of G_θ. The fault detection scheme is compromised of two parts: a filter that generates a residual carrying the information of the occurrence of the fault, and a decision function that generates a logic signal indicating the status of the monitored system. The residual, denoted e_y, carries the information regarding the occurrence of the injected sensor fault.

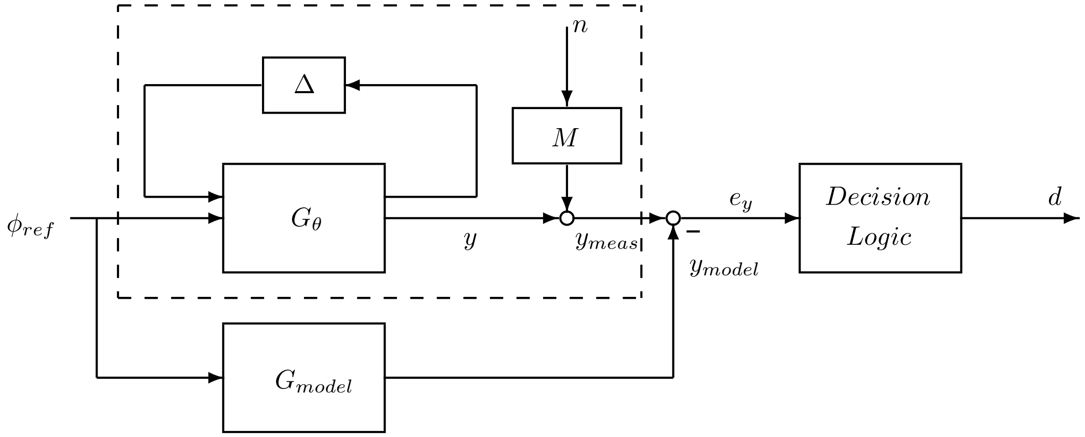

Figure 3. A Parity Equation FDI System.

The residual is generated using the analytical (model-based) relationship between ϕ_{ref} and y. Specifically, an estimate of y, denoted y_{model}, is generated using the dynamic model of the closed-loop system. Comparing this estimate with the measured value, y_{meas} gives the following residual:

$$e_y = y_{meas} - y_{model} \tag{1}$$

y_{model} is directly computed from ϕ_{ref} and G_{model}. G_{model} is the nominal closed-loop model of Thor and is included in the appendix. The residuals computed from the flight data for both $y = p$ and $y = r$ are shown in Figure 4. e_r will not be able to detect the roll rate sensor fault while e_p immediately indicates the onset of the roll-rate sensor fault at $t = 12$sec. Conversely, it is expected that faults in the yaw rate sensor would appear in e_r but not in e_p.

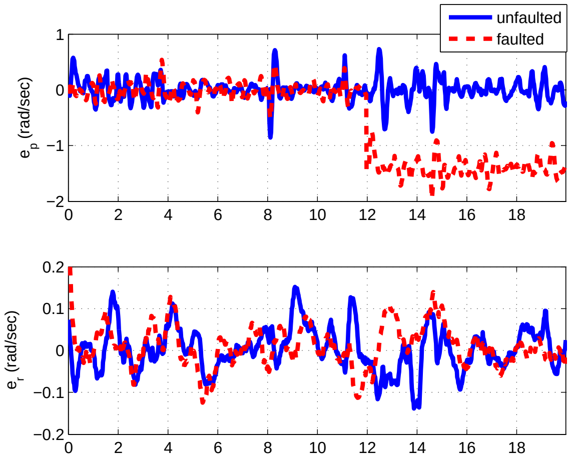

Figure 4. Residuals for unfaulted and faulted scenarios.

The residual generator is typically designed in a way such that e_y is small when $\theta(k) = 0$, and large when a fault occurs. Based on e_y, the decision logic generates a signal $d(k)$ to indicate the status of G_θ, i.e. $d(k) = 1$ if a fault has been detected and $d(k) = 0$ otherwise. There are many approaches for designing decision function logic, such as thresholding, statistical testing methods, and fuzzy logic.[8,9] Our analysis focuses on constant thresholding:

$$d(k) := \begin{cases} 0 & \text{if } |e_y(k)| \leq T \\ 1 & \text{else} \end{cases} \tag{2}$$

A fault is declared when $e_y(k)$ exceeds the threshold T. Thresholding is widely used in industrial applications due to its simplicity. The restriction to constant thresholds can also be viewed as a steady-state approximation for time-varying thresholds. This paper focuses on the validation method rather than the fault detection system design itself. There are many more advanced fault detection techniques than the simple parity-equation approach described here.[7–9] The parity-equation approach described here is used to simplify the design. The focus of this paper is on certification issues and validation using flight test data. The validation approach described in this paper could also be used for more advanced fault detection designs.

C. Performance Metrics

Commercial aircrafts achieve high levels of reliability almost exclusively through the use of hardware redundancy. In a physically redundant configuration, a failed component is detected by directly comparing the behavior of each redundant component. Hence, these schemes tend to detect faults accurately, and their performance is relatively straightforward to certify using fault trees.[5,6] In contrast, the use of model-based FDI schemes will introduce new failure modes associated with the imperfect detection capabilities of the model-based algorithm. These new failure modes must be considered in order to certify the performance of a system designed with a model-based FDI algorithm. As an example, the extended fault tree technique characterizes false alarms and missed detections as basic events that are incorporated into a fault tree.[19,20] The extended fault tree technique ties the system level failure probabilities to the FDI detection performance as well as the hardware component failure rates. In this approach, the probability of false alarm and probability of missed detection connect to the system failure rate. Hence these two quantities are used as the probabilistic metrics for quantifying the performance of the FDI system. It is noted that missed detection probability could also be equivalently expressed in terms of a detection probability. The probabilities of false alarm and detection have been used in the literature on FDI, e.g.[8,21] In this paper, the probability of detection is used in place of the probability of missed detection.

System level reliability requirements for aircraft are typically specified per hour, e.g. flight control systems certified by the FAA are required to have less than 10^{-9} catastrophic failures per hour.[2] Thus it is reasonable to expect that such system level requirements would be decomposed into per-hour requirements on the fault detection false alarm and detection probability. The performance requirements are specified over a specified N- step window in discrete-time domain. This motivates the following definitions.

Definition 1 *The probability of false alarm, denoted $P_F(N)$, is defined as the conditional probability that $d(k) = 1$ for some k in $1 \leq k \leq N$ given that $\theta(k) = 0$ for all k in $1 \leq k \leq N$.*

Definition 2 *The probability of detection, denoted $P_D(N)$, is defined as the conditional probability that $d(k) = 1$ for some k in $1 \leq k \leq N$ given that $\theta(k) = 1$ for all k in $1 \leq k \leq N$.*

These two metrics can be connected back to a system level failure rate in a rigorous way using the law of total probability.[19, 20, 22] Given a specified sample rate the per-hour false alarm probability can be converted to the discrete-time window size N. For example, a system with a 50Hz sample rate has $N = 1.8 \times 10^5$ samples per hour. The detection probability window size is typically much smaller. For example, a fault that is required to be detected within $1sec$ corresponds to a detection window of $N = 50$ for a system with a 50Hz sample rate. It is also important to note that the system dynamics and filters introduce time-correlations into the fault detection residuals. Thus it is not possible, in general, to accurately assess the FDI performance over N-step windows by using single frame probabilities. For example, if the probability of a false alarm in a single frame is given by P_0 then, due to the time correlations in the residual, the N-step false alarm probability will not be simply $1 - (1 - P_0)^N$. It is important to include the effects of these time correlations in the analysis.

D. Certification

A certification process for UAV FDI systems should incorporate theoretical probabilistic analyses, high fidelity Monte Carlo simulations and flight test validation. These methods have complementary benefits and shortcomings that address the issues related to model fidelity and stringent reliability requirements. Flight tests provide the highest fidelity as they validate the system performance in the actual operating conditions. However, flight tests are typically limited because they are expensive and time consuming. High fidelity nonlinear simulations can be used to assess the system reliability over a larger set of conditions. In particular, the aircraft can operate at many different flight conditions and in different modes. Given the stringent reliability requirements it can be costly to perform simulation for all possible situations. Theoretical probabilistic analysis can be used to complement flight tests and high fidelity simulations. Theoretical analysis can be used to assess performance at many flight conditions but under more restrictive modeling assumptions. For example, the probabilities of false alarm and detection can be efficiently and accurately calculated when the fault detection system is described as a linear time invariant system driven by Gaussian noise (see Section III.A).

A basic framework to assess the FDI performance using analysis, simulations, and flight tests is as follows. First, linearize the aircraft dynamics at a trim condition in the flight envelope. Next, apply the linear analysis tools to rigorously assess the FDI performance at this trim condition under a variety of assumptions on the environmental conditions (e.g. wind gusts) and model uncertainty. This linear analysis can be repeated at a grid of trim operating conditions in the flight envelope. Next, high fidelity Monte Carlo simulations can be performed to complement the linear analyses. The Monte Carlo simulations and the linear analyses both assume certain underlying probability models for failure modes and environmental disturbances. Thus finally, flight test data can be used to validate these assumptions and the results obtained by simulations and linear analyses. The objective is to use the flight test data to validate the linear analyses for several different typical flight conditions. If the results match then this gives further confidence in the linear analyses performed for the remaining (non-flight tested) operating conditions. This certification approach is similar to the framework used to evaluate flight control laws.[12, 23]

The remainder of the paper focuses on the use of flight test data to validate theoretical probabilistic analyses. For the flight test described in Section II.A, Thor was trimmed by the lateral directional controller and hence the entire closed-loop FDI system can be approximated by a linear time-invariant system driven by noise. Section III.A summarizes a theoretical method to compute the probability of false alarm $P_F(N)$ for such systems. The theoretical analysis assumes the noise driving the linear time-invariant system is Gaussian and IID. The theoretical analysis also neglects the model uncertainty and the time-varying dynamics. Thus it is important to use flight test data to validate the strong assumptions that underly the linear theoretical analysis. It is also noted the fault detection residuals are correlated in time. Both the linear analysis and the flight test validation will address these time correlations. The main challenges of the validation are due to the limited flight data and the stringent reliability requirements. For example, for a system sampling at 50 HZ there are $N = 1.8 \times 10^5$ samples per hour. Moreover, the fault detection system should be designed such that the false alarm probability $P_F(1.8 \times 10^5)$ is very small. If $P_F(1.8 \times 10^5) \approx 0.01$ then hour-long experiments should be run for at least 1000 times in order to draw statistically meaningful conclusions. In practice the costs and flight endurance limits prohibit such an extensive flight test campaign. For example, the endurance limit for Thor is approximately 20 minutes. This complicates the validation of the linear analyses. The goal of this paper is to find a reasonable validation framework to at least partially address these difficulties. The objective is to assess the quality of the linear analyses of false alarm probabilities within the constraints of limited flight data. The validation of linear analyses of probability of detection $P_D(N)$ introduces additional difficulties that are discussed at the end of this paper.

III. Validation Approach

Section III.A summarizes a theoretical technique to bound the false alarm probability for linear time invariant systems driven by Gaussian noise. Next, the approach to validate the linear analyses using flight test data is described in Section III.B.

A. False Alarm Analysis of Linear FDI systems

Consider a discrete-time state-space system of the form:

$$\begin{aligned} x(k+1) &= Ax(k) + Bn(k) \\ e_y(k) &= Cx(k) + Dn(k) \end{aligned} \tag{3}$$

Here, $x(k) \in \mathbb{R}^h$, $n(k) \in \mathbb{R}^l$, and $e_y(k) \in \mathbb{R}$ with the state matrices (A, B, C, D) having compatible dimensions. $n(k)$ is an IID Gaussian process with $n(k) \sim \mathcal{N}(0, \Sigma)$. This is reasonable to model the fault detection system shown in Figure 3 if the nominal dynamics of the system ($\theta = 0$ and $\Delta = 0$) are given by the model G_{model} used in the parity equation. In this case all discrepancies between the measured and model-based outputs is lumped in the stochastic errors driven by n.

For false alarm analysis, it it reasonable to assume the stochastic system is in the steady state. In this case the residual $e_y(k)$ is a strictly stationary zero-mean Gaussian process.

Define $E_N = \begin{bmatrix} e_y(1) & e_y(2) & e_y(3) & \cdots & e_y(N) \end{bmatrix}^T$ as the vector of residuals over the N-step window and let Λ_N denote the covariance matrix of E_N. This covariance matrix can be computed as follows. First, the steady state covariance matrix Σ_x of the random vector $x(k)$ can be solved from the Lyapunov equation:

$$\Sigma_x = A\Sigma_x A^T + B\Sigma B^T \tag{4}$$

Σ_x can be accurately and efficiently computed for given state-space data (A, B), e.g. using the `Matlab` function `dlyap`.[24] The covariance matrix of E_N is then given by the following Toeplitz matrix:

$$\Lambda_N(i,j) = \begin{cases} C\Sigma_x C^T + D\Sigma D^T & \text{if } i = j \\ CA^{|i-j|}\Sigma_x C^T + CA^{|i-j|-1}B\Sigma D^T & \text{else} \end{cases} \tag{5}$$

The probability density function of the residual vector E_N thus has the form:

$$f_N(E_N) = \frac{1}{\sqrt{(2\pi)^N |\Lambda_N|}} e^{-\frac{1}{2}E_N^T \Lambda_N^{-1} E_N} \tag{6}$$

Based on Definition 1, the probability of false alarm over the N-step window, $P_F(N)$, can be expressed in terms of this density function. Specifically, this probability is given formally in terms of the residuals as $P_F(N) = P[\cup_{k=1}^N \{|e_y(k)| > T\}]$. Using basic probability rules for mutually exclusive events, this can be expressed as

$$P_F(N) = 1 - P[\cap_{k=1}^N \{|e_y(k)| \leq T\}] = 1 - \int_{-T}^{T} f_N(E_N) dE_N \tag{7}$$

Let $Q(N) := P[\cap_{k=1}^{N}\{|e_y(k)| \leq T\}]$ denote the probability of no false alarm over N steps and note that $Q(N) := 1 - P_F(N)$.

The main theoretical analysis result applied in this paper is the following theorem providing bounds on $P_F(N)$ using product type inequalities:[25, 26]

Theorem 1 *Suppose e_y is a stationary zero-mean Gaussian process and denote $Q(0) := 1$. Then:*

$$\gamma_N^{(k)} = 1 - \left[\frac{Q(k)}{Q(k-1)}\right]^{N-k} Q(k) \tag{8}$$

$$P_F(N) \leq \gamma_N^{(k)} \tag{9}$$

where $1 \leq k \leq N$ and $\gamma_N^{(k)}$ is decreasing in k.

The important aspect of this result is that N is typically large for false alarm analysis and hence this makes it computationally intractable to directly evaluate $P_F(N)$. Since $\gamma_N^{(k)}$ for $1 \leq k \leq 3$ only depends on low dimensional integrals $Q(1)$, $Q(2)$, and $Q(3)$, this theorem bounds the false alarm probability also in terms of $Q(1)$, $Q(2)$, and $Q(3)$. $Q(1)$ is a one-dimensional Gaussian integral that can be accurately computed from the error function, e.g. `erf` in `Matlab`. $Q(2)$ and $Q(3)$ correspond to two and three dimensional Gaussian integrals, respectively. These integrals can also be efficiently computed to within machine (double) precision using the `Matlab` function `mvncdf`.[27] The sequence of bounds can be extended to include higher dimensional integrals ($Q(k)$ for $k \geq 3$) at the expense of additional computation. It can be proven under some technical assumptions that these bounds converge to $P_F(N)$[25, 28] and hence the bounds themselves can be used as estimates of $P_F(N)$.

B. Validation of Linear Analysis

The linear analyses given in Theorem 1 provide theoretical bounds for the probability of false alarm. Flight tests can also be used to compute sample estimates of the false alarm probability. The sample estimates obtained from flight test data can be used to validate the theoretical bounds computed from linear analysis. However, accurate calculation of sample estimates require a large amount of flight test data. This places restrictions on the validation of $P_F(N)$ for large values of N. For example, only 6 unfaulted data sets each of length 20seconds were performed on Thor. The system sample rate is 50Hz and hence the entire dataset consists of $50 \times 20 \times 6 = 6000$ samples. Additional data can be collected but basic limits on flight data will always exist.

Given these limits, the flight data will be used to validate the linear analysis of $P_F(N)$ for small N. If the theory and the flight data do not agree for small values of N then it is not likely that the linear analyses will be accurate for large N. For concreteness, the flight data will be used to validate the linear analyses for $N = 5$. To perform a linear analysis of $P_F(5)$, a stochastic model for the residuals is needed in the form provided in Equation 3. As discussed previously, Thor is trimmed during the flight test and the FDI system can be approximated as a linear system. It is assumed that the nominal dynamics of the closed-loop system ($\theta = 0$ and $\Delta = 0$) are given by the model G_{model}. In this case the residual generated by the simply parity-equation architecture in Figure 3 is simply given by the output of the model M driven by noise n. Thus the residuals e_y are given in the form Equation 3 where the state-space system represents the dynamics of M. One set of experimental data is used to identify an autoregressive moving average (ARMA) model M whose output is the residual $e_y(k)$. The model selection is based on autocorrelation function (ACF) and partial autocorrelation (PACF) function. The model identification is performed using the function `sarima` in `R` studio, which is a commonly used software in statistics community. ARMA models are identified for both the roll rate ($y = p$) and yaw rate ($y = r$) outputs. Both ARMA models can be recast into an equivalent state space model described by Equation 3. The linear analyses can then be performed to estimate the probability $P_F(5)$ based on Equations 4, 5 and Theorem 1.

Since one set of experimental data is used to fit a stochastic model for the FDI residual, only 5 datasets remain for validation. This corresponds to a toal of 5000 time samples. The sample estimates of $P_F(5)$ are computed from the data as follows. First, the data is binned into 1000 groups and a Bernoulli random variable $\mathbf{B}(i)$ is assigned to each group for $i = 1, 2, \cdots, 1000$. For the i^{th} group of data, $\mathbf{B}(i) = 1$ if the magnitude of at least one residual in the group exceeds the threshold T . Otherwise set $\mathbf{B}(i) = 0$. To avoid time correlations, the results for odd bins are neglected, i.e. $\mathbf{B}(i)$ is only used for i even. This enables the application of the law of large numbers to estimate the true value of $P_F(5)$ based on the sample estimate $P_F(5)$. The final sampled estimate of $P_F(5)$ is given by

$$P_F(5) \approx \frac{\sum_{i=1}^{500} \mathbf{B}(2i)}{500} \tag{10}$$

The sample estimate of $P_F(5)$ is compared with the results from the linear analysis to determine the accuracy of the linear analysis. It is noted that the use of 500 groups places a fundamental bound on the accuracy of the sample false alarm probability and for $P_F(5)$. In particular, a false alarm in no bin gives the sample estimate $P_F(5) \approx 0$ while a false alarm in a single group gives the estimate $P_F(5) \approx \frac{1}{500}$. Thus the resolution of the sample false alarm estimate is $\frac{1}{500}$ and validation of the linear analysis can not be performed for false alarm probabilities below this level.

IV. Results and Discussions

This section applies the theoretical method discussed in Section III.A to assess the false alarm probability. The theoretical analysis is validated using the method described in Section III.B.

As described in Section II.A, there are a total of six unfaulted flight tests available for the validation of the false alarm analysis. Given the modeling assumptions in Section III.B, the residuals are generated by the stochastic system M which is driven by Gaussian noise n. One set of experimental flight data is used to fit ARMA models for FDI residuals. There are many techniques for identifying time-series models.[29,30] The ARMA model identification is only briefly described since this step is not the main focus of the paper. Additional details of the model selection are included in the appendix. The ARMA models for the residuals e_p and e_r are identified using the basic procedure provided in Example 3.39 in Shumway's textbook.[29] First, the ACF and PACF provide qualitative information about the order of the ARMA model to fit. Next, the function `sarima` in `R` toolbox is used to fit ARMA model to FDI residuals. The ACF of the innovation term n and Ljung-box statistic p-value are used to check the independence of n. Moreover, a normal quantile-quantile (Q-Q) plot is used to check the normality of the innovation term. Finally, Akaike's information criterion (AIC), Bayesian information criterion (BIC) and bias corrected AIC (AICc) are used to pick up the best ARMA models among all the candidates passing the diagnostics.

For the roll rate residual e_p the final identified ARMA model is:

$$e_p(k) = 1.0592e_p(k-1) + 0.2379e_p(k-2) - 0.4585e_p(k-3) + n_p(k) + 0.8141n_p(k-1) + 0.0787n_p(k-2) \tag{11}$$

where $n_p(k)$ is an IID Gaussian process with $n_p(k) \sim \mathcal{N}(0, 1.193 \times 10^{-3})$. Equation 11 can be recast into an equivalent state space model described by Equation 3 with the following state-space data (A, B, C, D):

$$\begin{aligned} A &= \begin{bmatrix} 1.0592 & 0.2379 & -0.4585 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ B &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \\ C &= \begin{bmatrix} 1.0000 & 0.8141 & 0.0787 \end{bmatrix} \\ D &= 0 \end{aligned} \tag{12}$$

The driving noise for this state space system is the IID process n_p defined above. Theorem 1 provides a sequence of bounds $\gamma_5^{(k)}$ on the false alarm probability over a five-step window, $P_F(5)$. The first three bounds $\gamma_5^{(k)}$ $(k = 1, 2, 3)$ are evaluated using the covariance matrix of the residual e_p. This covariance matrix of the residual e_p is computed using Equations 4 and 5. The sample estimate of $P_F(5)$ is computed from Equation 10 with the remained five sets of flight data. Figure 5 shows the sampled estimate of $P_F(5)$ and the theoretical estimates of $P_F(5)$ as functions of the threshold value T. The top subplot shows $P_F(5)$ on a log scale while the bottom subplot shows $P_F(5)$ on a linear scale. The third bound $\gamma_5^{(3)}$ is almost identical to $\gamma_5^{(2)}$ and hence it is omitted from the figure for clarity. For threshold values smaller than 0.5, the sampled estimate of $P_F(5)$

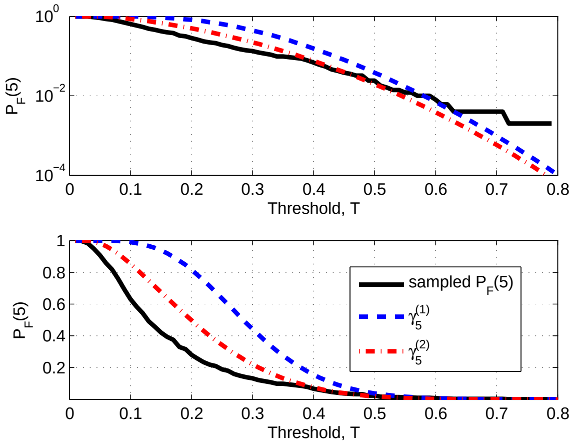

Figure 5. FAR $P_F(5)$ vs T for roll rate residual e_p.

is bounded by $\gamma_5^{(2)}$. But for larger threshold values, the sampled $P_F(5)$ exceeds the theoretical bound and shows heavy tail behavior. Further discussion of these results is given later in this section.

For the yaw rate residual e_r the final identified ARMA model for the FDI residual is:

$$e_r(k) = 1.7840e_r(k-1) - 0.7997e_r(k-2) + n_r(k) - 0.3563n_r(k-1) \tag{13}$$

where $n_r(k)$ is an IID Gaussian process with $n_r(k) \sim \mathcal{N}(0, 4.132 \times 10^{-5})$. Equation 13 can also be recast into an equivalent state space model described by Equation 3 with the following state-space data (A, B, C, D):

$$\begin{aligned} A &= \begin{bmatrix} 1.7840 & -0.7997 \\ 1 & 0 \end{bmatrix} \\ B &= \begin{bmatrix} 1 & 0 \end{bmatrix}^T \\ C &= \begin{bmatrix} 1 & -0.3563 \end{bmatrix} \\ D &= 0 \end{aligned} \tag{14}$$

The driving noise for this state space system is the IID process n_r defined above. The sampled estimates and the theoretical bounds on $P_F(5)$ are shown in Figure 6 as functions of the threshold T. The top subplot again shows $P_F(5)$ on a log scale and the bottom subplot is shown on a linear scale.

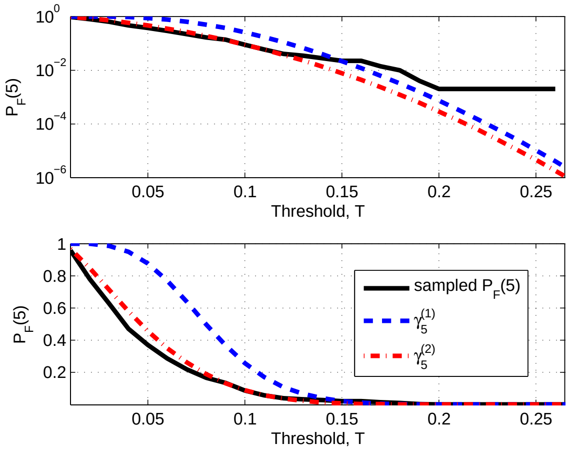

Figure 6. FAR $P_F(5)$ vs T for yaw rate residual e_r.

For both e_p and e_r the sampled estimates of $P_F(5)$ are bounded by the theoretical $\gamma_5^{(2)}$ for small thresholds T. For both cases, there exist non-Gaussian heavy tail behavior for larger thresholds. This heavy tail behavior is also expected based on the the normal Q-Q plot in Figure 8 which is included in the appendix. Comparing Figure 5 and 6, it appears that the theoretical bounds are more accurate for the yaw rate fault residual e_r. This agrees with the normal Q-Q plots shown in Figure 8 which demonstrate that e_r has a much better agreement with a Gaussian distribution than e_p. The main cause of the heavy tail behavior could be the time varying nature of the system during aggressive bank angle maneuvers. These aggressive bank angle commands result in a non-constant variance of the driving noise process n. The sample ACF and PACF of the squared innovation series show correlations which is consistent with this hypothesis. The time-varying nature of the variance of e_p is also observed in Figure 4. It is possible that the linear analysis tools will only provide good estimates when the aircraft is not performing aggressive maneuvers. To verify this hypothesis, further flight tests will be performed in the future. In particular, the bank angle reference signal ϕ_{ref} should be kept near 0 so that the UAV will remain near trim. The hypothesis will be confirmed if the variance of the residual remains time-invariant during these flight tests.

As mentioned before, the validation of the linear analyses can not be performed for false alarm probabilities below the resolution of the sampled false alarm probability. The current analysis only consists of 500 groups of flight data and hence this limits the resolution of the sampled false alarm probability to $\approx \frac{1}{500}$. More flight data is required to validate the linear analysis for lower false alarm probabilities. Hence an important question is how to balance the trade-off between the resolution of validation approach and the costs of flight test experiments.

For future research, the central limit theorem can be a useful tool to rigorously analyze the resolution of the sampled false alarm probability. A typical procedure can be found in Chapter 1.1 of Rubino and Tuffin's textbook.[31] The difficulty here is due to the correlations of the sampled data. Proper justifications must be made to ensure the validity of the central limit theorem.

The experimental validation of the linear analyses of probability of detection requires repeatedly trimming the aircraft and injecting the faults. This operation is more difficult due to the transient nature of this kind of experiments. Future flight test should be designed to address this problem.

V. Conclusion

This paper discusses a process to certify a model-based fault detection system using theoretical analysis, nonlinear simulations and flight test data. The main focus of the paper is a method to use limited flight test data to validate theoretical, linear analyses of the false alarm probability. This validation approach is applied to a simple UAV fault detection system for sensor faults. The sampled false alarm probability and the theoretical bounds agree for small threshold values. However, the results diverge for larger thresholds. A possible cause of this non-Gaussian heavy tail behavior is the aggressive bank angle maneuvers used in the flight tests. Future research will include flight tests in trim conditions to confirm or deny this hypothesis.

Appendix

A. Closed-Loop Lateral/Directional System

This section provides the dynamic model of the closed-loop lateral directional system for completeness. A block diagram of this control system is shown in Figure 7. $AC_{lat/dir}$ represents the lateral dynamics of the aircraft with inputs aileron (δ_{ail}) and rudder (δ_{rud}) deflections. The units of all the signals are included in the nomenclature section. The aircraft dynamics are linearized at 17m/sec and the linearized dynamics are described by the following five state model:

$$\frac{d}{dt}\begin{bmatrix} v \\ p \\ r \\ \phi \\ \psi \end{bmatrix} = \begin{bmatrix} -0.6288 & 0.68 & -16.37 & 9.794 & 0 \\ -2.095 & -13.53 & 3.861 & 0 & 0 \\ 1.002 & -0.451 & -2.466 & 0 & 0 \\ 0 & 1 & 0.04961 & 0 & 0 \\ 0 & 0 & 1.001 & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \\ \psi \end{bmatrix} + \begin{bmatrix} -1.552 & 2.242 \\ -130 & 2.066 \\ -1.933 & -20.9 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_{ail} \\ \delta_{rud} \end{bmatrix} \tag{15}$$

The outputs used for feedback are bank angle (ϕ), roll rate (p), and yaw rate (r). The aileron and rudder actuators are modeled as

$$Act(s) = \frac{50.27^2}{s^2 + 80.43s + 50.27^2}\tau(s) \tag{16}$$

$\tau(s)$ is a sixth order Pade approximation for a $0.05sec$ time delay.

The control system has an inner/outer loop structure to achieve tracking of bank angle reference commands, ϕ_{ref}. The inner-loop roll rate damper (K_p) uses proportional gain to reject disturbances in turbulent conditions:

$$K_p(s) = -0.07 \tag{17}$$

An inner-loop yaw damper (K_r) applies a proportional gain to increase damping in the Dutch roll mode, and a washout filter to avoid an adverse yaw effect during turns:

$$K_r(s) = \frac{0.065s}{s+2} \tag{18}$$

Finally, an outer-loop proportional-integral controller (K_ϕ) is applied to track the bank angle reference command:

$$K_\phi(s) = -0.64 - \frac{0.2}{s} \tag{19}$$

The closed-loop model from bank angle reference commands to roll and yaw rate outputs can be completely determined from the models for the aircraft, actuators and control laws.

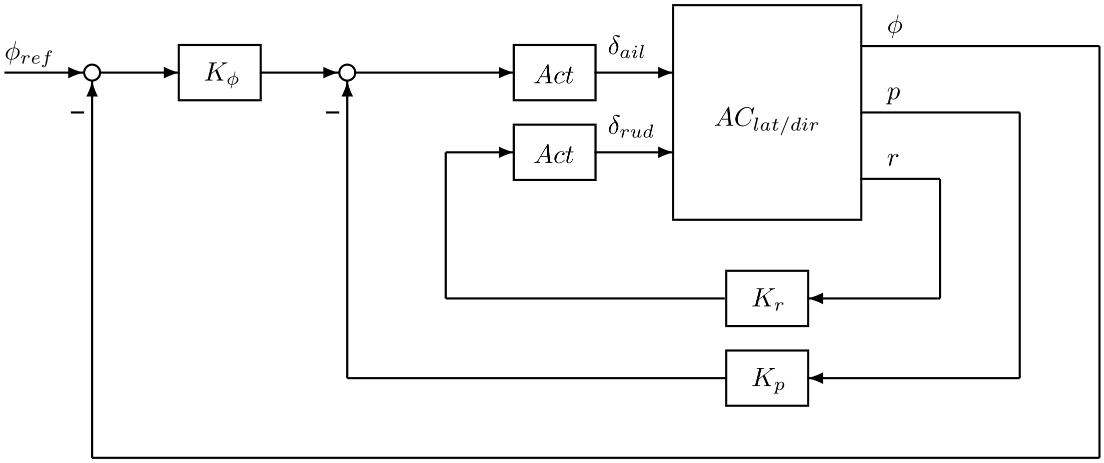

Figure 7. Closed-loop Lateral/directional System.

B. ARMA Model Selection

This section briefly summarizes the process to identify ARMA models for the roll rate and yaw rate residuals. First, the roll rate residual e_p was computed as the difference between the roll rate measurement and the estimate response computed using the closed-loop model described in Appendix A. This residual is assumed to be entirely due to a stochastic system, denoted as M in Figure 3, driven by an IID Gaussian noise process n. The sampled ACF and PACF of e_p was used to identify an appropriate order for an ARMA model of M. The sampled PACF dropped off after a lag of 3 and this indicates the suitability of a ARMA model with AR order 3. Next, the function `sarima` in the `R` toolbox was used to fit the ARMA model to the residual e_p. This function generates several diagnostic plots that can be used to assess the model. Finally, several criterion (e.g. AIC, BIC, and AICc) are used to pick up the best ARMA models among all the candidates passing the diagnostics. For the roll rate residual e_p, the final selected model is the ARMA(3,2) process given in Equation 11. Overall the model fit was reasonable and several diagnostics were used to further evaluate the quality of this model. The ACF of the innovation term n_p as well as the Ljung-box statistic both imply the independence of this innovation term. However, the Q-Q plot shown in the top subplot of Figure 8 shows a deviation of the data (circles) from the expected Gaussian distribution (solid line). This implies that the distribution of the innovation has a heavy tail. Similar steps were used to identify the model for the yaw rate residual. The result is the ARMA(2,1) process given in Equation 13. The ACF of the innovation term n_r as well as the Ljung-box statistic again both imply the independence of this innovation term. In addition, the Q-Q plot shown in the bottom subplot of Figure 8 shows the innovation n_r has a much better agreement with a Gaussian distribution except several outliers and does not have the strong heavy tail phenomenon observed with n_p.

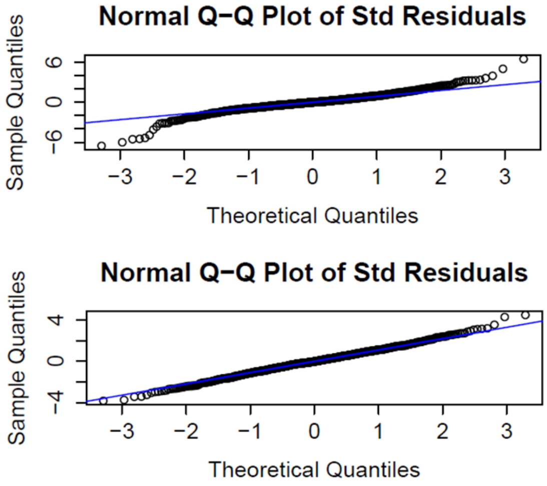

Figure 8. Normal Q-Q plots for innovation terms n_p and n_r.

Acknowledgments

The authors would like to thank Brian Taylor and Andrei Dorobantu for providing support with the experimental data and aircraft modeling. This work was partially supported by the National Science Foundation under Grant No. NSF-CMMI-1254129 entitled CAREER: Probabilistic Tools for High Reliability Monitoring and Control of Wind Farms. It was also partially supported the NASA Langley NRA Cooperative Agreement NNX12AM55A entitled Analytical Validation Tools for Safety Critical Systems Under Loss-of-Control Conditions, Dr. Christine Bel-castro technical monitor. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF or NASA.

References

[1] Bleeg, R., "Commercial jet transport fly-by-wire architecture considerations," *AIAA/IEEE Digital Avionics Systems Conference*, 1988, pp. 399–406.

[2] Collinson, R., *Introduction to Avionic Systems*, Kluwer, 2003.

[3] Yeh, Y., "Triple-triple redundant 777 primary flight computer," *Proceedings of the 1996 IEEE Aerospace Applications Conference*, 1996, pp. 293–307.

[4] Yeh, Y., "Safety critical avionics for the 777 primary flight controls system," *Proceedings of the 20th Digital Avionics Systems Conference*, 2001, pp. 1.C.2.1–1.C.2.11.

[5] Krasich, M., "Use of fault tree analysis for evaluation of system-reliability improvements in design phase," *IEEE Proc. Reliability and Maintainability Symposium*, 2000, pp. 1–7.

[6] Lee, W., Grosh, D., Tillman, A., and Lie, C., "Fault tree analysis, methods, and applications: a review," *IEEE Trans. on Reliability*, Vol. 34, No. 3, 1985, pp. 194–203.

[7] Chen, J. and Patton, R., *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Kluwer, 1999.

[8] Ding, S., *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*, Springer-Verlag, 2008.

[9] Isermann, R., *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*, Springer-Verlag, 2006.

[10] Goupil, P., "Oscillatory failure case detection in the A380 electrical flight control system by analytical redundancy," *Control Engineering Practice*, Vol. 18, No. 9, 2010, pp. 1110–1119.

[11] "ADDSAFE: Advanced Fault Diagnosis for Sustainable Flight Guidance and Control," http://addsafe.deimos-space.com/, 2012, European 7th Framework Program.

[12] Renfrow, J., Liebler, S., and Denham, J., "F-14 flight control law design, verification, and validation using computer aided engineering tools," *IEEE Conference on Control Applications*, 1994, pp. 359–364.

[13] "University of Minnesota UAV Research Group," http://www.uav.aem.umn.edu/, 2013.

[14] Dorobantu, A., Johnson, W., Lie, F. A., Taylor, B., Murch, A., Paw, Y. C., Gebre-Egziabher, D., and Balas, G., "An Airborne Experimental Test Platform: From Theory to Flight," *Proceedings of the American Control Conference*, 2013.

[15] Paw, Y., *Synthesis and Validation of Flight Control for UAV*, Ph.D. thesis, University of Minnesota, 2009.

[16] Murch, A., Paw, Y., Pandita, R., Li, Z., and Balas, G., "A Low Cost Small UAV Flight Research Facility," *Advances in Aerospace Guidance, Navigation and Control*, 2011, pp. 29–40.

[17] Dorobantu, A., Murch, A., Mettler, B., and Balas, G., "Frequency Domain System Identification for a Small, Low-Cost, Fixed-Wing UAV," *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, 2011.

[18] Dorobantu, A., Murch, A., Mettler, B., and Balas, G., "System Identification for Small, Low-Cost, Fixed-Wing Unmanned Aircraft," *AIAA Journal of Aircraft*, 2013.

[19] Aslund, J., Biteus, J., Frisk, E., Krysander, M., and Nielsen, L., "Safety analysis of autonomous systems by extended fault tree analysis," *International Journal of Adaptive Control and Signal Processing*, Vol. 21, No. 2-3, 2007, pp. 287–298.

[20] Gustafsson, F., Aslund, J., Frisk, E., Krysander, M., and Nielsen, L., "On threshold optimization in fault-tolerant systems," *IFAC World Congress*, 2008.

[21] Willsky, A. S. and Jones, H. L., "A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems," *IEEE Transactions on Automatic Control*, Vol. 21, 1976, pp. 108–112.

[22] Hu, B. and Seiler, P., "A probabilistic method for certification of analytically redundant systems," *Submitted to 2nd International Conference on Control and Fault-Tolerant Systems*, 2013.

[23] Buus, H., McLees, R., Orgun, M., Pasztor, E., and Schultz, L., "777 Flight Controls Validation Process," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 33, No. 2, 1997, pp. 656–666.

[24] Hammarling, S. J., "Numerical solution of the stable, non-negative definite Lyapunov equation," *IMA J. Numer. Anal*, Vol. 2, 1982, pp. 303–323.

[25] Hu, B. and Seiler, P., "Probability Bounds for False Alarm Analysis of Fault Detection Systems," *Submitted to 51st Annual Allerton Conference on Communication, Control, and Computing*, 2013.

[26] Glaz, J. and Johnson, B., "Probability Inequalities for Multivariate Distributions with Dependence Structures," *Journal of the American Statistical Association*, Vol. 79, No. 386, 1984, pp. 436–440.

[27] Genz, A., "Numerical computation of rectangular bivariate and trivariate normal and t Probabilities," *Statistics and Computing*, Vol. 14, 2004, pp. 251–260.

[28] Glaz, J. and Johnson, B., "Boundary Crossing for Moving Sums," *Journal of Applied Probability*, Vol. 25, No. 1, 1988, pp. pp. 81–88.

[29] Shumway, R. and Stoffer, D., *Time Series Analysis and Its Applications: With R Examples*, Springer, 3rd ed., 2011.

[30] Ljung, L., *System Identification: Theory for the User*, Prentice-Hall, 2nd ed., 1999.

[31] Rubino, G. and Tuffin, B., *Rare Event Simulation using Monte Carlo Methods*, Wiley, 2009.