Fault Diagnosis for a Turbine Engine[1]

Yixin Diao and Kevin M. Passino
Department of Electrical Engineering, The Ohio State University
2015 Neil Avenue, Columbus, OH 43210

Abstract

Fault detection and diagnosis for jet engines is complicated by the presence of engine-to-engine manufacturing differences and engine deterioration during normal operation, the complexity of an accurate engine model, and our inability to directly measure certain engine variables. Here, we work with a sophisticated component level model (CLM) simulation of a turbine engine (XTE46) that can simulate the effects of manufacturing and deterioration differences, in addition to a variety of failures. To develop a fault diagnosis system we begin by using the CLM to generate data that is used by a Levenberg-Marquardt method to train a Takagi-Sugeno fuzzy system to represent the engine. The multiple copies of this nonlinear model, each representing a different failure, are then used to generate error residuals by comparing them to the engine output. In fact, we manage the composition of the set of models with a "supervisor" that ensures the appropriate models are on-line, and that processes the error residuals to detect and identify faults. The robustness of the approach is analyzed and several simulations are conducted to illustrate the effectiveness of the method.

1 Introduction

Stimulated by the growing demands for improving the reliability and performance of systems, many fault diagnosis methods, in particular, model-based analytical redundancy methods, have been developed in the last two decades, which have the capability of detecting the occurrence of faults and determining the types (or locations) of the faults. Survey papers by Frank [1] and Patton [2] present excellent overviews of advances in model-based fault diagnosis methods. Some methods have been applied in complex applications. In [3] the authors studied fault detection of jet engine sensor systems using eigenstructure assignment and in [4] the authors detected sensor and actuator faults of flight control systems using multiple model adaptive estimation. Recently, methods such as fuzzy systems, neural networks and expert systems have received increasing attention and been applied in the field of fault diagnosis, (e.g., see the discussion in [5]). These methods have the potential to "learn" the plant model from input-output data or "learn" fault knowledge from past experience, and they can be used as function approximators to construct the analytical model for residual generation, or as supervisory schemes to make the fault analysis decisions [6, 7, 8].

This paper presents a fault diagnosis method that utilizes fuzzy modeling and an expert supervisory scheme. The analytical model of the plant, which is in the form of Takagi-Sugeno fuzzy systems, has been developed using nonlinear system identification (Section 2). In the fault diagnosis system that we describe in Section 3, a bank of multiple models are used, and the supervisory scheme determines the proper model bank to generate residuals and analyzes the residuals to detect and identify faults. We also study robustness of the approach, and then in Section 4 the proposed method is applied to a turbine engine (XTE46) and its effectiveness is demonstrated via simulation examples.

2 Model Development Using Fuzzy Systems

In the course of developing fault diagnosis schemes the use of analytical redundancy implies that a mathematical model of the system is used to describe the inherent relationship (or redundancy) contained among the system inputs and outputs which may be used to generate the residuals for fault diagnosis. The resulting approaches are usually referred to as "analytical redundancy based" fault diagnosis or "model-based" methods. This is the approach we take here.

2.1 The Takagi-Sugeno Fuzzy System

Developing mathematical models for nonlinear systems can be quite challenging. However, Takagi-Sugeno fuzzy systems are capable of serving as the analytical model for nonlinear systems due to its universal approximation property, that is, any desired approximation accuracy can be achieved by increasing the size of the approximation structure and properly defining the parameters of the approximator [9]. A Takagi-Sugeno fuzzy system can be defined by

$$y = F_{ts}(x,\theta) = \frac{\Sigma_{i=1}^{R} g_i(x)\mu_i(x)}{\Sigma_{i=1}^{R}\mu_i(x)} \quad (1)$$

$$g_i(x) = a_{i,0} + a_{i,1}x_1 + \cdots + a_{i,n}x_n \quad (2)$$

$$\mu_i(x) = \prod_{j=1}^{n} \exp\left(-\frac{1}{2}\left(\frac{x_j - c_j^i}{\sigma_j^i}\right)^2\right) \quad (3)$$

[1]This work was supported by the NASA Glenn Research Center under grant NAG3-2084. Please address all correspondence to K. Passino (passino@ee.eng.ohio-state.edu, (614) 292-5716).

where y is the output of the fuzzy system, $x = [x_1, x_2, \ldots, x_n]^\top$ holds the n inputs, $i = 1, 2, \ldots, R$ represent R different rules, and $j = 1, 2, \ldots, n$ represent n different inputs. The shapes of the membership functions are chosen to be Gaussian, and center-average defuzzification and product are used for the premise and implication in the structure of the fuzzy system. The $g_i(x)$ are called consequent functions of the fuzzy system, where $a_{i,j}$ are linear parameters. The premise membership functions $\mu_i(x)$ are assumed to be well defined so that $\Sigma_{i=1}^{R}\mu_i(x) \neq 0$. The parameters that enter in a nonlinear fashion are c_j^i and σ_j^i, which are the centers and relative widths of the membership functions for the j^{th} inputs and i^{th} rules.

2.2 Fuzzy Modeling for the XTE46 Engine

The CLM for the XTE46 aircraft engine is a complicated multiple-input multiple-output nonlinear system (involving schedules, look-up tables, and partial differential equations.) We assume (as pointed out by GE Aircraft Engines) that its fundamental dynamic characteristics, however, can be represented by a single-input single-output system in the form

$$\dot{x} = f(c, p, x, u) \quad (4)$$
$$y = g(c, p, x, u) \quad (5)$$

where $x = [\mathrm{XNL}, \mathrm{XNH}]^\top$ is the state vector, $u = \mathrm{WF36}$ is the input variable, $y = \mathrm{XN2}$ is the output of the engine, $c = [\mathrm{ALT}, \mathrm{XM}, \mathrm{DTAMB}, \mathrm{PC}]^\top$ represents the operating condition of the engine, $p = [\mathrm{ZSW2}, \mathrm{SEDM2}, \mathrm{ZSW7D}, \mathrm{SEDM7D}, \mathrm{ZSW27}, \mathrm{SEDM27}, \mathrm{ZSW41}, \mathrm{ZSE41}, \mathrm{ZSW49}, \mathrm{ZSE49}]^\top$ represents the quality parameter vector, $f(\cdot)$ denotes the unknown function representing the nonlinear characteristics of the engine, and $g(\cdot) = \mathrm{XNL}$ because the output variable XN2 is just the measurement of the state variable XNL.

Given a specific operating condition and fixed values of quality parameters, the analytical model for the XTE46 engine can be obtained using Takagi-Sugeno fuzzy systems, which is in the form

$$\begin{aligned}
\widehat{\mathrm{XNLdot}}(k) &= F_{ts}^1(\widehat{\mathrm{XNL}}(k), \widehat{\mathrm{XNH}}(k), \mathrm{WF36}(k), \theta^1) \\
\widehat{\mathrm{XNHdot}}(k) &= F_{ts}^2(\widehat{\mathrm{XNL}}(k), \widehat{\mathrm{XNH}}(k), \mathrm{WF36}(k), \theta^2) \\
\widehat{\mathrm{XNL}}(k+1) &= \widehat{\mathrm{XNL}}(k) + T_s\widehat{\mathrm{XNLdot}}(k) \\
\widehat{\mathrm{XNH}}(k+1) &= \widehat{\mathrm{XNH}}(k) + T_s\widehat{\mathrm{XNHdot}}(k) \\
\widehat{\mathrm{XN2}}(k+1) &= \widehat{\mathrm{XNL}}(k+1)
\end{aligned}$$

where two multiple-input single-output (MISO) Takagi-Sugeno fuzzy systems, F_{ts}^1 and F_{ts}^2, are specified with corresponding parameters θ^1 and θ^2, respectively. The variables $\widehat{\mathrm{XNL}}(k)$ and $\widehat{\mathrm{XNH}}(k)$ denote the estimated values of $\mathrm{XNL}(k)$ and $\mathrm{XNH}(k)$, $\widehat{\mathrm{XN2}}(k)$ is the estimated value of $\mathrm{XN2}(k)$, and $\widehat{\mathrm{XNLdot}}(k)$ and $\widehat{\mathrm{XNHdot}}(k)$ are the outputs of the fuzzy systems. The variable T_s is the sample time, which is 0.02 second.

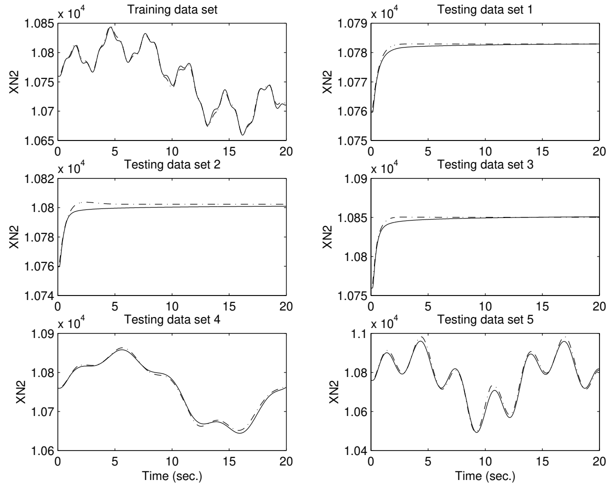

Figure 1: Performance of "medium fan fault model."

The fuzzy systems are trained using engine data generated by the transient driver of the component level model simulator of the XTE46 engine (provided by GE Aircraft Engines). One thousand engine input-output data pairs are collected which reflect the transient performance of the engine for 20 seconds (sampled every 0.02 second) at a specific "node" (of operating conditions and quality parameters).

For simplicity we only consider two fault types: fan fault and compressor hub fault, and each may have three different fault levels: small, medium and large. Using Takagi-Sugeno fuzzy systems we have obtained 16 models including one nominal model (no fault), six single fault models (small fan fault, medium fan fault, etc.), and nine double fault models (small fan fault with small compressor hub fault, small fan fault with medium compressor hub fault, etc.). The performance of the models are reasonably good. For instance, the training and testing performance of one model (medium fan fault) is shown in Figure 1, where the solid lines represent the system response of the real engine and the dashed lines represent the response of the analytical model.

3 Fault Diagnosis by Interfacing Multiple Models with a Supervisory Scheme

3.1 Residual Generation via Multiple Models

The engine may run in many situations: no fault or different types of faults, possibly with manufacturing differences and deterioration. Comparing to the method which only uses a single model of the plant and discriminates between different faults by setting different thresholds for different residuals, the multiple model method may be preferable. The multiple model approach requires a bank of models to identify different situations by comparing the residuals generated these models. Assume that there are N possible fault sit-

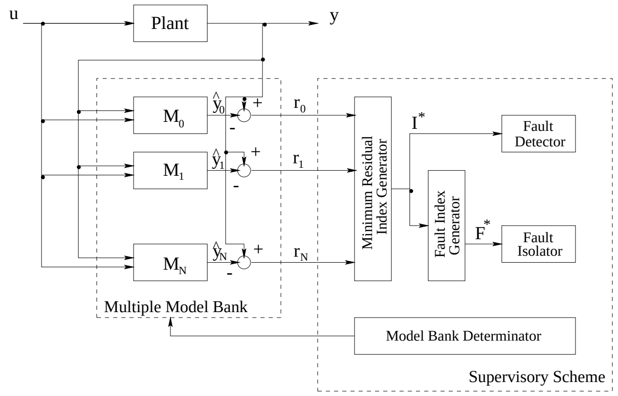

Figure 2: Structure of the fault diagnosis scheme.

uations. Using nonlinear system identification with Takagi-Sugeno fuzzy systems (as explained in the previous section), we may have $N+1$ models $\{M_j\}_{j=0}^N$, where the nominal model M_0 corresponds to the situation where there is no fault, and the fault models M_j, $j = 1, 2, \ldots, N$ represent the j^{th} fault situations. By running a bank of models on-line, as shown in Figure 2, the residuals r_j can by generated by $r_j = y - \hat{y}_j$, where y is the output of the plant and $\hat{y}_j$ is the output of model M_j, (which we think of as an estimate of the plant output.)

3.2 Expert Supervisory Scheme

The evaluation of the residual signals generated by the models is performed using an expert supervisory scheme. The heuristic knowledge of faults and processing experience can be incorporated into the expert system in the form of rules easily, and thus its advantages are the transparency of operation and simple integration of *a priori* knowledge. Basically, the rule-based expert supervisory system performs two functions:

1. Determine the bank of models to generate the residuals: Due to our understanding of the plant, we already have some knowledge of faults, i.e., the possible types of faults and the possible combinations of different faults. According to this, $N+1$ models have been established to describe the characteristics of these possible situations in which the plant may operate. Besides, we may know that there are some relationships among these fault situations. For instance, suppose fault i and fault j are two different types of faults, if fault i and only fault i is happening at present, then the next possible situation may be the existence of fault j together with fault i, or no fault, but not the situation where only fault j exists under the assumption that no two faults can happen or extinguish simultaneously. Therefore, at each instant only a subset of $N+1$ situations may happen and thus only a subset of models may be required to generate residuals, which can be summarized as the rule: *IF (fault situation i is currently in existence) THEN (the subset S_i of models $\{M_j\}_{j=0}^N$ will be used as the model bank to generate residuals afterwards).* In particular, the rules may be

IF (there is no fault) THEN (the model bank consists of no fault model and all possible single fault models), or, *IF (there is a single fault) THEN (all single fault models except the one that has been isolated will be removed from the model bank and the multiple fault models including the isolated fault and another possible fault will be added).* The initial situation is usually assumed to be no fault. Using this type of knowledge the computation complexity can be reduced and the possibility of false alarms can be decreased as well.

2. Fault diagnosis by decision logic: Once the residuals are generated, the fault decision logic is used to diagnose the faults. A sequence of minimum residual indices is first generated by

$$I^*(k) = arg\min(r_i(k)) \tag{6}$$

where $I^*(k)$ denotes the index of the model whose residual is minimum, that is, the most appropriate model to represent the plant at instant k. Note that when the plant situation changes, the input-output characteristics of the system will change as well. Therefore, the change of index, i.e., the change of most suitable model, may be used to indicate the change of plant situations, i.e., the occurrence of new faults. However, the new index after the change does not mean that the corresponding fault situation has occurred. This is because during this transient phase the residuals may change drastically and some of them may happen to be very small for a short time and become large later since they are not the proper fault situations. In order to isolate the fault situations correctly, some logic rules are used to guarantee that a fault will be isolated only if it lasts at least for T_0 seconds. The results of isolation can be represented by a fault index F^*, as shown in Figure 2. This strategy may affect fault sensitivity to some extent but it is effective to isolate faults accurately which is more important.

3.3 Robustness Analysis

The robustness of fault diagnosis refers to the ability to prevent false alarms in the presence of modeling uncertainties, that is, if the system is in the i^{th} fault situation, the fault diagnosis system should indicate the i^{th} fault situation rather than the j^{th} fault situation where $j \neq i$. The robustness of the proposed fault diagnosis system is achieved by using the time delay term T_0 in the fault isolation scheme.

Assume that the nonlinear dynamic system in the i^{th} fault situation can be described by

$$\begin{aligned} \dot{x} &= f(c, p_i^0, x, u) + \eta_i(c, p_i, x, u) & (7) \\ y &= Cx & (8) \end{aligned}$$

where $f(\cdot)$ and C represent the known model dynamics which are characterized by the developed Takagi-

Sugeno model with "nominal" quality parameter vector p_i^0 corresponding to the i^{th} fault situation but without considering initial engine variation and engine deterioration, and $\eta_i(\cdot)$ represent modeling uncertainties in the system, where p_i are unknown quality parameter vectors with manufacturing differences and deterioration. We assume that the modeling errors are additive, unstructured, and bounded, so that the accumulation effect of modeling uncertainties may be described by

$$\left| C \int_0^t \eta_i(c, p_i, x, u) d\tau \right| \leq \eta_i^0 \tag{9}$$

The model corresponding to the j^{th} fault situation can be represented by

$$\begin{aligned} \dot{\hat{x}}_j &= f(c, p_j^0, \hat{x}_j, u) && (10) \\ \hat{y}_j &= C\hat{x}_j && (11) \end{aligned}$$

where $\hat{x}_j$ is the estimated state vector of the j^{th} model and $\hat{y}_j$ is the estimated output.

We assume that the i^{th} model can represent the "nominal" system dynamics in the i^{th} fault situation accurately enough comparing to modeling uncertainties, that is,

$$\left| C \int_0^t (f(c, p_i^0, x_i, u) - f(c, p_i^0, \hat{x}_i, u)) d\tau \right| \leq \bar{\alpha}\eta_i^0$$

and the difference between the j^{th} model ($j \neq i$) and the system dynamics in the i^{th} fault situation is quite large in the measure of modeling uncertainties, that is

$$\max_{t-T_0 \leq t' \leq t} \left| C \int_0^{t'} (f(c, p_i^0, x_i, u) - f(c, p_j^0, \hat{x}_j, u)) d\tau \right| \geq \underline{\alpha}\eta_i^0$$

where $\bar{\alpha}$ and $\underline{\alpha}$ are parameters which satisfy the inequality $\underline{\alpha} \geq \bar{\alpha} + 2$ (they are used in the proof later) and we use max to represent the idea that $|C \int_0^{t'} (f(c, p_i^0, x_i, u) - f(c, p_j^0, \hat{x}_j, u)) d\tau|$ may only be smaller than $\underline{\alpha}\eta_i^0$ for no longer than T_0 seconds, which may happen when the estimated output of a wrong model approaches the output for a short period of time and then leaves again. Note that the effects of faults on quality parameters are usually larger than those of initial engine variation and engine deterioration, and the inaccuracy that arises from the accumulation of drifting errors is usually smaller than that from modeling uncertainties; hence, the above assumptions may be satisfied in real applications.

Under above definitions and assumptions the robustness problem of the fault diagnosis system is that in the i^{th} fault situation, the residuals of the j^{th} models ($r_j = y - \hat{y}_j, j \neq i$) cannot be smaller than the residual of the i^{th} model ($r_i = y - \hat{y}_i$) for a period of T_0 time, i.e.,

$$\max_{t-T_0 \leq t' \leq t} |r_i(t')| \leq \max_{t-T_0 \leq t' \leq t} |r_j(t')| \tag{12}$$

so that the fault diagnosis system will still indicate "i" as the current fault situation. The following analysis ensures robustness of the proposed fault diagnosis scheme. Related analysis, but for a different problem, may be found in [6].

For the i^{th} model, defining the state estimation error to be $e_i = x - \hat{x}_i$, the error dynamics are described by

$$\begin{aligned} \dot{e}_i &= f(c, p_i^0, x, u) - f(c, p_i^0, \hat{x}_i, u) + \eta_i(c, p_i, x, u) \\ r_i &= Ce_i \end{aligned}$$

By processing the differential equation we obtain

$$\begin{aligned} |r_i(t)| &\leq \left| C \int_0^t \left(f(c, p_i^0, x, u) - f(c, p_i^0, \hat{x}_i, u) \right) d\tau \right| \\ &\quad + \left| C \int_0^t \eta_i(c, p_i, x, u) d\tau \right| \\ &\leq \bar{\alpha}\eta_i^0 + \eta_i^0 = (\bar{\alpha} + 1)\eta_i^0 \end{aligned}$$

and thus

$$\max_{t-T_0 \leq t' \leq t} |r_i(t')| \leq (\bar{\alpha} + 1)\eta_i^0$$

In addition, the error dynamics of the j^{th} model are

$$\begin{aligned} \dot{e}_j &= f(c, p_i^0, x, u) - f(c, p_j^0, \hat{x}_j, u) + \eta_i(c, p_i, x, u) \\ r_j &= Ce_j \end{aligned}$$

so that

$$\begin{aligned} &\max_{t-T_0 \leq t' \leq t} |r_j(t')| \\ \geq &\max_{t-T_0 \leq t' \leq t} \left| C \int_0^{t'} \left(f(c, p_i^0, x, u) - f(c, p_j^0, \hat{x}_j, u) \right) d\tau \right| \\ &- \max_{t-T_0 \leq t' \leq t} \left| C \int_0^{t'} \eta_i(c, p_i, x, u) d\tau \right| \\ \geq &\underline{\alpha}\eta_i^0 - \eta_i^0 = (\underline{\alpha} - 1)\eta_i^0 \end{aligned}$$

Therefore, $\max_{t-T_0 \leq t' \leq t} |r_i(t')| \leq \max_{t-T_0 \leq t' \leq t} |r_j(t')|$ when $\underline{\alpha} \geq \bar{\alpha}+2$, which means the robustness of the fault diagnosis system is guaranteed and no false alarms will be generated.

4 Fault Diagnosis Simulation Results

First, we study the case where only the small fan fault happens in the "takeoff" region and there is no compressor hub fault in the "climb" region. The initial bank of multiple models is determined by the expert supervisory scheme to be composed of six single fault models and one no fault model. The output (XN2) of the engine and the estimated outputs (and residuals) of multiple models are shown in Figure 3 (where the wide solid line represents the plant output, the thin solid line represents the output or residual of the nominal model, the thin dotted line represents that of the small fan fault model, the thin dash-dotted line represents that of the medium fan fault model, the thin dashed line represents that of the large fan fault model, the wide dotted line represents that of the small compressor hub fault model, the wide dash-dotted line represents that of the medium compressor hub fault model and the wide dashed line represents that of the large

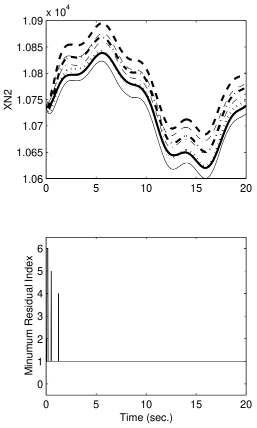

Figure 3: Small fan fault situation.

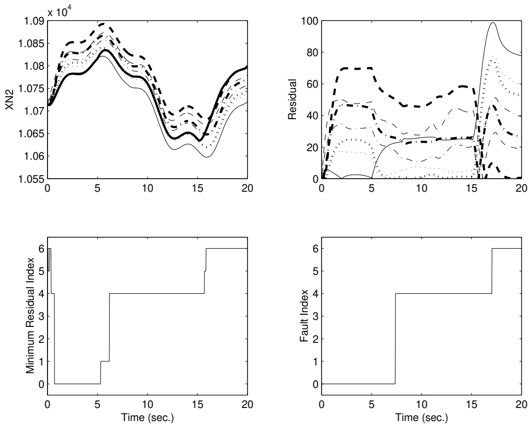

Figure 4: Small compressor hub fault followed by large compressor hub fault.

compressor hub fault model). As shown in Figure 3, after the original oscillations, which are absorbed by the isolation function in the supervisory scheme (where $T_0 = 1.2$ seconds), the fault is isolated at 2.44 seconds (where fault index 1 indicates small fan fault, index 2 indicates medium fan fault, index 3 is large fan fault, index 4 is small compressor hub fault, etc.).

For multiple fault situations, we study a case where there is no fault for the first 5 seconds, a small compressor hub fault occurs and lasts for 10 seconds, and at 15 seconds the size of the compressor hub fault changes from small to large. As shown in Figure 4, using six single fault models and one no fault model as the initial bank of multiple models, a fault is detected at 5.3 seconds, which is a small fan fault. However, this is just the transient phenomenon. After it lasts for 0.86 seconds, the minimum error index changes to 4, which indicates a small compressor hub fault. At 7.38 seconds the small compressor hub fault is claimed to be isolated, as shown in the fault index plot. Afterwards, at 15.7 seconds a fault is declared to be detected from the minimum error index and at 17.04 seconds the large compressor fault is isolated. Clearly, a fault can be detected shortly after it occurs and be isolated correctly a bit longer.

Note that to keep the sensitivity of fault detection, any changes in the minimum residual index (i.e., the changes of the most suitable model) will be considered to indicate the occurrence of new faults and thus the minimum residual index is used to detect the fault promptly. In our simulation examples, the detection delay is less than 1 second. Furthermore, considering the presence of modeling uncertainties, a time delay term T_0 is included to achieve the robustness of the fault isolation scheme. In general, there is a tradeoff between robustness and sensitivity of fault isolation. Improving the robustness to modeling uncertainties may cause the fault isolation scheme to be insensitive. Here, T_0 is chosen to be 1.2 seconds by *a priori* knowledge on modeling errors and the isolation delay in our simulation is less than 2.5 seconds.

Acknowledgment: We would like to thank S. Adibhatla at General Electric Aircraft Engines for his assistance and advice on many aspects of this project. Moreover, we would like to thank T.H. Guo at NASA Glenn Research Center for his support.

References

[1] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - a survey and some new results," *Automatica*, vol. 26, pp. 459–474, 1990.

[2] R. J. Patton, "Fault-tolerant control: The 1997 situation," in *IFAC Fault Detection, Supervision and Safety for Technical Processes, Kingston Upon Hull, UK*, pp. 1029–1051, 1997.

[3] R. Patton and J. Chen, "Robust fault detection of jet engine sensor systems using eigenstructure assignment," *Journal of Guidance, Control and Dynamics*, vol. 15, no. 6, pp. 1491–1497, 1992.

[4] T. Menke and P. Maybeck, "Sensor/actuator failure-detection in the vista F-16 by multiple model adaptive estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, pp. 1218–1229, 1995.

[5] P. J. Antsaklis and K. M. Passino, eds., *An Introduction to Intelligent and Autonomous Control.* Norwell, MA: Kluwer Academic Publishers, 1993.

[6] A. Vemuri and M. Polycarpou, "Robust nonlinear fault diagnosis in input-output systems," *International Journal of Control*, vol. 68, pp. 343–360, 1997.

[7] K. M. Passino and P. J. Antsaklis, "Fault detection and identification in an intelligent restructurable controller," *Journal of Intelligent and Robotic Systems*, vol. 1, pp. 145–161, June 1988.

[8] E. G. Laukonen, K. M. Passino, V. Krishnaswami, G.-C. Luh, and G. Rizzoni, "Fault detection and isolation for an experimental internal combustion engine via fuzzy identification," *IEEE Transactions on Control Systems Technology*, vol. 3, pp. 347–355, 1995.

[9] K. M. Passino and S. Yurkovich, *Fuzzy Control.* Menlo Park, CA: Addison Wesley Longman, 1998.