ROBUST MODEL IDENTIFICATION APPLICATION TO A TURBOFAN ENGINE

Andrés Marcos, [1] **Dinkar Mylaraswamy,** [2] **Gary J. Balas,** [3]

Abstract: In this paper we describe an application of a robust identification approach to a turbofan engine. The approach combines classical time-data system identification via prediction error models and the so-called model error model theory. The model error model approach provides guidelines for the characterization of the model uncertainty stemming from the nominal system identification and fits naturally in the robust fault detection and isolation framework for $\mathcal{H}_\infty$-based algorithms. *Copyright*© *2005 IFAC*

Keywords: system identification, model error model, turbofan engine

1. INTRODUCTION

A driving economic force in the aerospace industry, particularly commercial airlines, is the use of aircraft engine anomaly and fault detection & isolation (FDI) schemes to improve their reliability and optimize (reduce) their maintenance schedules. The current FDI state-of-the-art in the industry involves empirical techniques (e.g. neural networks and clustering (Uluyol, O. *et al.*, 2003; Kim, K. *et al.*, 2004)) and model-based techniques (least-squares estimation (Gorinevsky, D. *et al.*, 2002)). All these approaches provide analytical redundancy using a model of the system obtained by system identification or first-order principles. An important component in system identification is the estimation of model uncertainty descriptions and bounds.

Typically, the uncertainty characterization is performed in a heuristic manner using experimental data and knowledge of the system. An alternative approach is the so-called 'robust identification' which can be based on an statistical framework (correlation analysis of the residuals (Ljung, L., 1987; Reinelt, W *et al.*, 2002)) or a deterministic framework (Smith, R. and Doyle, J., 1992). The latter framework hinges in the representation of the model uncertainty and the noise error in terms of linear fractional transformations (LFT).

In this paper, an application of a robust identification approach to a Honeywell turbofan engine is presented. The robust engine model, to be used for FDI purposes using $\mathcal{H}_\infty$-optimization techniques, is obtained by applying prediction error model identification algorithms (Ljung, L., 1987) to a set of engine flight data provided by a Honeywell airline customer. The bias error, or model uncertainty, arising from the nominal system identification is initially characterized by the model error model approach (Ljung, 1999) in terms of linear weights.

2. THEORY

The identification approach used combines parametric identification for the nominal and fault models (based on classical prediction error theory) and uncertainty characterization (based on model error model theory). This combination of nominal and model error model identification is considered a robust model identification approach (Ljung, 1999; Reinelt, W *et al.*, 2002) and fits naturally in the robust fault detection and isolation framework of $\mathcal{H}_\infty$-based algorithms.

2.1 Nominal Model Identification

In the control field, system identification is understood to be a process by which a dynamic model is obtained from measured input and output data. The identified model is selected from a family of models and is

[1] Work performed as a Research Intern at Honeywell International Inc. Currently, Post-Doctoral Fellow at the Department of Engineering, Control & Instrumentation Group, University of Leicester, UK. e-mail: ame12@le.ac.uk, corresponding author.
[2] Principal Research Scientist, Honeywell International Inc., 3660 Technology Drive, Minneapolis, MN 55418, USA.
[3] Professor, Department of Aerospace Engineering and Mechanics University of Minnesota, 110 Union St. SE, MN, USA.

described by a finite set of parameters and a structure (the relation between the parameters). In order to obtain a good and reliable model, the designer must pay attention to several aspects: choice of an appropriate input/output data set, selection of the model structure and order, and thorough validation of the resulting model (Ljung, L., 1987).

Parametric models are selected in this paper for the structural identification stage while correlation and frequency response models (non-parametric models) are used as analysis tools to gain insight on the engine model. The nominal identification algorithm used is based in a *model-batch* evaluation approach: a set of candidate structures are selected and then evaluated using different orders and delays. The three main model structures chosen are: auto-regressive with exogenous input (ARX), output error (OE), and a general predictive model (PEM). Continuous-time, closed-loop engine data provided by a partner airline in Honeywell's Predictive Trend Monitoring (PTM) project is used for identification and validation purposes.

2.2 Model Error Model Identification

The concept of the model error model (MEM) introduced in reference (Ljung, 1999) allows a very defined separation of the residual into variance (noise error) and bias (modeling uncertainty) errors. The model error model provides as well with an alternative test to the classical residual validation approach. In references (Ljung, L., 2001; Reinelt, W *et al.*, 2002) this MEM concept is used to provide frequency-domain characterizations of the model uncertainty for the purposes of control-oriented model validation. The approach used in those references provide a natural fit with the robust control design framework. In this paper, this natural fit is used within the robust fault detection and isolation framework by using the model error model as an initial uncertainty weight.

There are no restrictions on the identification approach used for the nominal model or the model error model. Furthermore, the identification of the latter is not affected by the method or assumptions used to obtain the nominal model. The only information required for MEM identification is the residual estimates $\varepsilon(t)$ obtained using a (possibly different) set of output/input measurements. An algorithm and guidelines to calculate the model error models for an identified nominal system can be found in (Ljung, 1999; Reinelt, W *et al.*, 1999; Reinelt, W *et al.*, 2002).

As suggested in reference (Reinelt, W *et al.*, 2002), a simple approach to calculate an additive uncertainty weight from the uncertainty regions obtained using the nominal and MEM models is simply to calculate the lower error bound $\gamma(i\omega)$ that bounds the uncertainty region, $\Delta_a(i\omega)$, such that $\{G_o(i\omega) + \Delta_a(i\omega) \;/\; \Delta_a(i\omega) \leq \gamma(i\omega)\}$. Constant and first-order approximations are recommended to decrease the total order of the $\mathcal{H}_\infty$ FDI filter to be synthesized later.

3. TURBOFAN ENGINE MODEL

The turbofan engine under consideration is a two-spool, high bypass engine composed of fan section, gas producer module, combustor/turbine module and accessory gearbox.

A dynamic, open-loop, black-box nonlinear simulation model is available together with engine continuous-time closed-loop data. The simulation model has four main inputs: the first three can be considered as *disturbances* (total air temperature, *TAT* Celsius, Mach number *Mn*, and pressure altitude, *Alt* feet), and the fourth input is the command signal (fuel flow, Wf lbs/hr). There are three output channels: fan speed $N1$ in % r.p.m, high-pressure (core) speed $N2$ % r.p.m., and exhaust gas temperature EGT in Celsius. An electronic control unit (ECU) regulates the amount of fuel flow used to control the fan speed by means of regulation of $N2$.

Additional input channels to the simulation model account for faults and aircraft/engine bleed. The fault considered, HPT_f, characterizes a high pressure turbine (HPT) deterioration due to leading edge erosion or tip clearance change (i.e. distance between the turbine blades and the turbine encasing). It is measured as a dimensionless number.

4. NOMINAL AND FAULT MODEL IDENTIFICATION

This section describes the identification results for the nominal model and the fault model. The approach presented in Section 2 is used to identify an LTI Multiple-Input-Multiple-Output (MIMO) nominal model formed by three (one for each output) LTI Multiple-Input-Single-Output (MISO) models. This nominal MIMO model is augmented by a MIMO fault model (formed also by three MISO models) which characterizes the HPT_f fault effects on the engine outputs.

The operational envelope of interest is the cruise flight regime, hence the signals for identification and validation are obtained by combining four different cruise data sets for each signal. Figure 1 shows the input data for the identification signal.

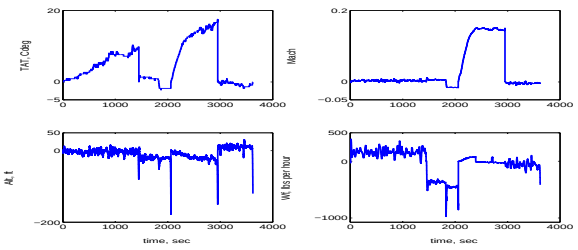

Fig. 1. Identification Signal Input Data.

The large deflections observed are mainly the result of the data sets connection. If these discontinuities were to be considered too severe for the identification, the data could be pre-filtered. In the present case, the discontinuities are left to artificially introduce some

nonlinear behavior that might arise from sudden reduction of speed during the cruise regime (i.e. the pilot can always maintain altitude while reducing speed by using the inter-play between the thrust and the elevators).

For each of the three prediction error models used (ARX, OE and PEM) different orders and delays are tested until a satisfactory fit is obtained. A primary concern is to obtain relatively low-order models since this influences the final order of the $\mathcal{H}_\infty$ FDI filter. The fit criterion is a combination of quantitative and qualitative measures. There are three quantitative measures used: the Akaike's Information Theoretic Criterion (AIC), the Final Prediction Error (FPE), and the percentage of output variation explained by the model simulation (PCT) (Ljung, L., 1987). Together with these criteria, the classical residual analysis, time-domain comparisons, and frequency-domain estimation, i.e. empirical transfer functions (ETFE) and spectral estimate (SPA), are used to qualitatively evaluate the models.

4.1 Nominal Model Identification Results

This subsection presents the identification of the three fault-free Multiple-Input-Single-Output (MISO) models using engine data obtained from the Predictive Trend Monitoring (PTM) database. The inputs for all the models are the engine disturbance inputs (TAT,Mn, and Alt) and the control engine input (Wf). The outputs for each of the three models are $N1$, $N2$ and EGT respectively.

Table 1 shows the structure, order of the coefficients, and delays for each of the MISO models. The terminology follows that of reference (Ljung, L., 2002). Numbers within brackets indicate the order of the four input channels for that coefficient (a single number means all channels have the same order).

Table 1. Structural nominal MISO models.

Model	Struc.	na	nb	nc	nd	nf	nk
MISO $N1$	OE	0	[1]	0	0	[1]	[0]
MISO $N2$	ARX	5	[5]	0	0	[0]	[0]
MISO EGT	OE	0	[1]	0	0	[1]	[1]

For the ARX model, a bandpass Butterworth filter with a frequency band of $[10e^{-5}\ \ 0.5]$ rad/s is used to focus in the low-mid frequency region. An example of the quantitative results obtained is given by the fit criteria for the MISO $N1$ model using the identification signal: $FPE = 1.90873e^{-01}$,$AIC = -1.65614$, and $PCT = 90.53235$ %, which indicates a very good fit. Similar results (with a minimum $PCT \approx 80\%$) are obtained for the other two models.

Figure 2 provides the time simulation comparison between the validation signal (solid) and the simulated output (dashed) for the MISO $N2$ model. The plot shows good agreement. Similar results are obtained for the other models with the EGT model showing some level of bias.

Figure 3 shows the frequency response comparison (i.e. ETFE and SPA obtained from the data and the

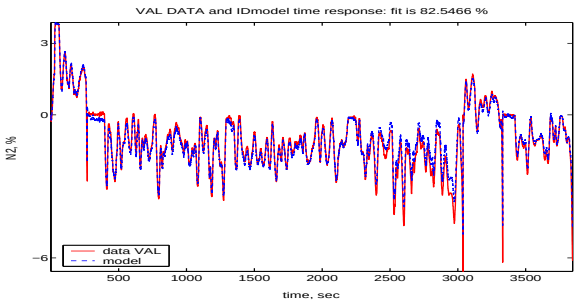

Fig. 2. VAL signal time resp. comparison, MISO $N2$.

freq. response of the identified model) for the MISO $N2$ model. One of the qualitative criteria is to obtain acceptable fit for the low-mid frequency region. The ETFE shows a large phase discrepancy around the one radian region. This discrepancy will have undesirable consequences on the achievable robustness of the diagnosis filter designed with this model.

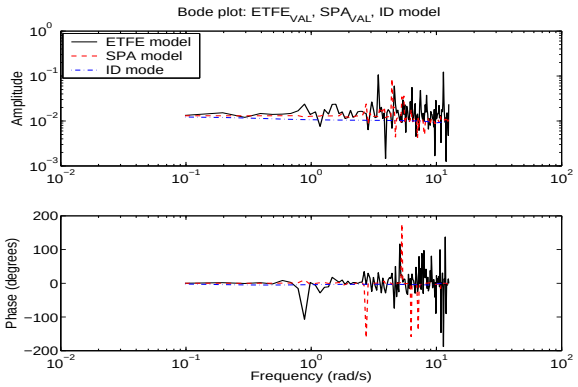

Fig. 3. Frequency analysis comparisons, MISO $N2$.

The classical residual analysis shows falsification in the MISO $N1$ model by the validation signal while the other models pass the residual test with only occasional small cross-correlation in the first few lags for the Alt and Wf channels.

Time simulations with forty-two different cruise flight data sets yields acceptable comparisons. Nevertheless, the EGT channel presents a larger bias - probably caused by control feedback in the data. This bias can be seen in Table 2, which gives the mean (MEAN) and standard deviation (STD) between the simulated LTI MIMO model, formed by the combinations of the three MISO models, and the data for all the flights.

Table 2. Statistical properties for 42 cruise flights.

	N1	N2	EGT
MEAN	0.03004	0.02706	1.98002
STD	0.15655	0.04701	2.28696

In summary, a 13 order LTI nominal engine MIMO model is identified with good time-domain and acceptable frequency-domain fits but with some cross-correlation problems and phase discrepancies.

4.2 Fault Model Identification Results

An identification similar to that for the nominal model above is used to obtain a MIMO LTI model of the

HPT fault effects (i.e. three MISO LFT models but in this case with five input channels: the previous four plus the fault input). Since no faulty data sets are available, the 'black-box' engine simulation model is used in combination with flight data sets to obtain the required input/output faulty measurements.

The identification and validation signals are each formed by using a different flight data set and a different 'pure' fault signal injected in the simulation model. The fault signals are selected to be an initial ramp followed by a pseudo-random signal in order to represent the slow degradation of the fault followed by abrupt breakage (i.e. incipient and abrupt faults). Figure 4 shows the 'pure' fault signal used for identification.

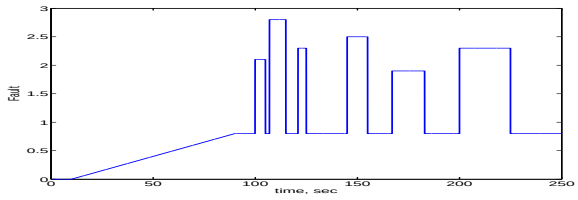

Fig. 4. Fault identification signal.

The same identification procedure as for the nominal model is followed except that only ARX and OE models are tested. Table 3 shows the final order and structure selected.

Table 3. Structural fault LTI MISO models.

Model	Struc.	na	nb	nf	nk
MISO *N*1	OE	0	[1 1 1 2 2]	[1 1 1 2 2]	[0]
MISO *N*2	ARX	1	[1]	[0]	[0]
MISO *EGT*	ARX	1	[1]	[0]	[0]

Good time simulation comparisons are obtained for the identified fault models, Figure 5 shows the most divergent case (i.e. the fault validation signal for the fault MISO *N*2 model). The plot shows the residual formed by subtracting the no-fault case for the same flight data, i.e. it shows only the fault effects.

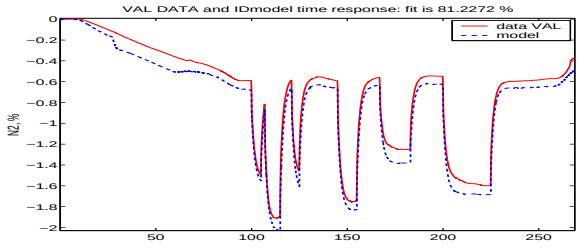

Fig. 5. Time response comparison, MISO HPT_f to $N2$.

As before, frequency and residual analyses were performed. It is desirable as well to obtain good fit for the low-mid frequency region while no emphasis is placed on the residual tests. The reason for not placing emphasis on residual tests is that closed-loop flight data and a designed fault signal are injected in an open-loop model (hence, we expect high levels of correlation between inputs).

The LTI MIMO fault model identified has nine states and, when augmented with the 13 states nominal MIMO model, good nominal and fault time simulations are obtained.

5. ROBUST MODEL IDENTIFICATION

In this section, the model error models (MEM) associated with the nominal models identified in the previous section are presented. For each identified MISO nominal model there exist one MEM (i.e. a total of three models corresponding to the number of outputs in the system). As established earlier, the purpose of these model error models is to provide an alternative test to the classical residual validation and to provide initial characterizations of the model uncertainty for the application of the robust residual generation algorithm.

5.1 Model Error Model

Model error models should be more flexible than the identified model on which they are based in order to capture the possible non-linearities and dynamics that are not adequately represented by the nominal model. A Box-Jenkins general parametric model provides this flexibility of design. Furthermore, due to the closed-loop nature of the data and the interdependencies of the inputs, a MISO identification (with the same four inputs as for the nominal model) is used. A SISO identification using the fuel flow as the driving input was also tried with much less success. The same identification and validation signals applied to the previous nominal model identification (see Section 4) are used to identify the MEMs.

Table 4 shows the order of the coefficients and delays for each of the MISO MEM models. The differences among the orders of the input channels for the same MEM are immediately noticeable. These differences illustrate the more difficult task that the MEM identification presents. Indeed the ouput percentage fit is around 20 % for most of the validation signals tested.

Table 4. Structural MISO MEM models.

Model	na	nb	nc	nd	nf	nk
MEM *N*1	0	[2 5 3 2]	4	4	[5 5 5 5]	[2 2 0 2]
MEM *N*2	0	[1 1 1 4]	1	1	[4 4 4 4]	[2 1 2 1]
MEM *EGT*	0	[1 1 1 4]	2	2	[1 1 1 4]	[1 1 1 0]

Figure 6 shows the *Mn* and *Alt* cross-correlation for the *EGT* MEM model using the validation signal. The case shown in the figure marginally satisfies the residual test for three standard deviations, but it is also the worst case obtained. The other channels and models pass the test with some cross-correlation stemming from Wf for the validation signal.

Now we turn our attention to the alternative nominal model validation test that the model error model provides. In Section 4.1 it was noted that the residual test was not passed for the nominal MISO *N*1 model using the validation signal, also some initial cross-correlation issues were identified in the Alt, WF channels for the other models. References (Ljung, 1999; Reinelt, W *et al.*, 2002) mention that the falsified nominal model is still of use if additional information with respect to the uncertainty is obtained. In Figure 7, the non-symmetric region (shaded band) and the

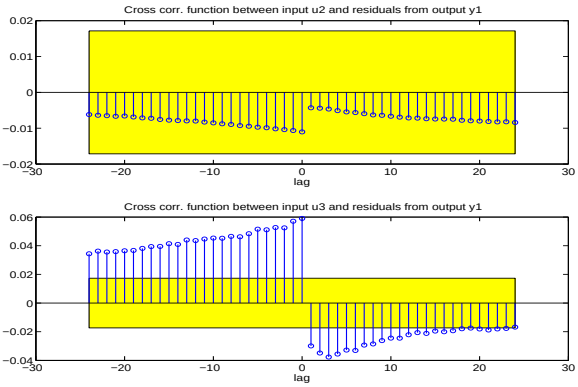

Fig. 6. Auto and cross-correlation, MEM EGT (band of 3 standard deviations).

nominal model (solid line) for the $N1$ model, specifically the *Alt* channel, is shown. It is clear that the nominal model is falsified since it is not completely included within the uncertainty region. Indeed, this happens for the *Alt* and *Mn* of all the models (except the *Alt* channel in the EGT model) which seems to indicate this test is more stringent than the classical residual test. The conclusion is that care should be exercised for these two channels at frequency regions approximately between $[0.03 \rightarrow 10]$ radians for the robust residual generation. Furthermore, for the $N2$ model, as was noted earlier, large phase discrepancies occur around the one radian region, see Figure 3. The MEM uncertainty characterization reflected this by using larger peaks of the uncertainty frequency band around the $[1 \rightarrow 15]$ radians region (specially in the *Mn* channel). This also highlights a limitation of the MEM uncertainty characterization whereby phase information is only indirectly considered in the definition of the uncertainty bands.

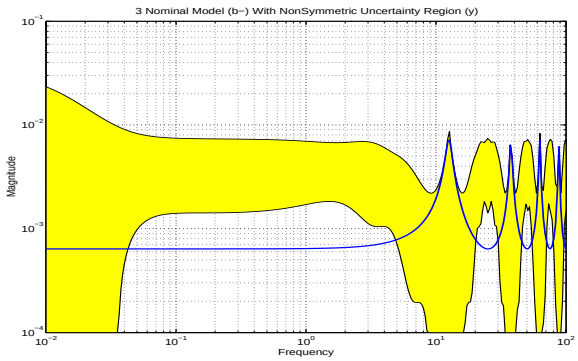

Fig. 7. Non-Symmetric uncertainty region (shaded band), MEM $N1$ channel *Alt*.

Figure 8 provides the symmetric region for the same model and channel. As expected by design, the non-symmetric uncertainty region and the nominal model are covered by the symmetric band. This region is more conservative than the non-symmetric, but it can now be used for uncertainty modeling in the frequency region.

5.2 Uncertainty Characterization

As discussed in Section 2.2, additive uncertainty weights are calculated for the uncertainty regions. The

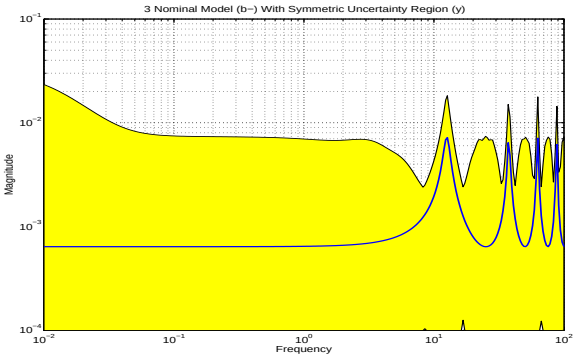

Fig. 8. Symmetric uncertainty region (shaded band), MEM $N1$ channel *Alt*.

$\mathcal{H}_\infty$-norm of the uncertainty region is used to obtain constant bounds for some of the channels, while first order fits (obtained using the command *drawmag* from (Balas, G.J. *et al.*, 1998)) are used for some other channels. These weights are used in the robust residual generation stage as an initial characterization of the uncertainty. Figure 9 provides an example of a first order fit for the fuel flow channel of the EGT model.

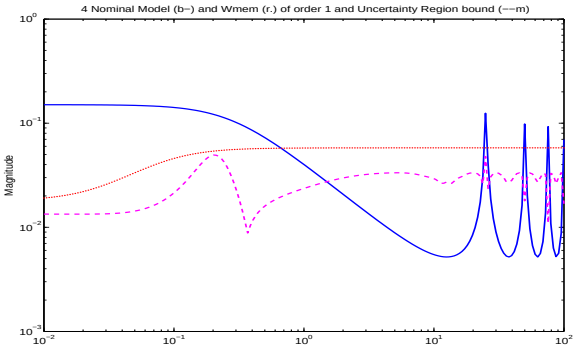

Fig. 9. First order fit symmetric uncertainty region, MEM EGT channel Wf.

Table 5 gives the values of the constant fits for each of the models and channels. Notice that the largest influences are registered in the Mach Number channel for all the three MISO MEM models. This is expected since this channel has the largest gain in the system. Also, note that the values for the core speed channel $N2$ are the smallest of each plant input row. This agrees with the time simulation and classical residual results from the nominal model identification and the MEM validation which showed that this model was seemingly the best of the three.

Table 5. Constant bound on MEM uncertainty region of nominal Model.

	N1	**N2**	**EGT**
TAT	0.2330	0.0219	0.7263
Mn	20.8582	6.6457	61.1725
Alt	0.0227	0.0045	0.0982
Wf	0.0049	0.0022	0.0495

6. $\mathcal{H}_\infty$ FDI SIMULATION RESULTS

In this section some results of the FDI $\mathcal{H}_\infty$ filter obtained using the LTI MIMO model and the previous

uncertainty characterization are shown, see reference (Marcos, A., 2004).

Figure 10 shows the time simulation performed using the complete (fault plus nominal) LTI MIMO model identified in this paper and the detection filter. A continuous-time flight data serie is used together with a simulated fault signal. It is observed that the fault is adequately detected even for relatively small (incipient) faults and for abrupt faults. The detection time is around 100 seconds, with remarkably good rejection properties.

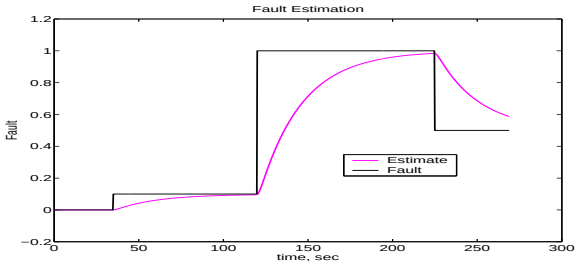

Fig. 10. Fault and estimation for LTI MIMO model and Flight 1.

Figure 11 shows the behavior of the detection filter when applied to the open-loop nonlinear simulation model of the engine. It is seen again that the $\mathcal{H}_\infty$ FDI filter is able to identify the high-pressure turbine fault, although there are obvious coupling issues with some of the disturbances. These results indicate the need for using a residual evaluation to provide bounds on the fault estimate in order to improve the false rate.

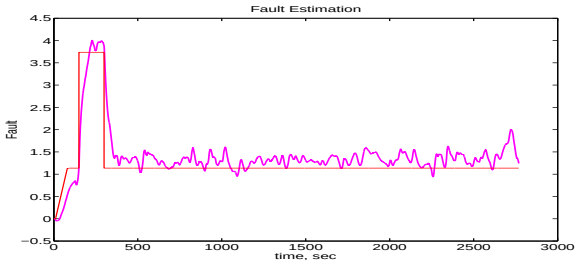

Fig. 11. Fault and estimation for nonlinear model and Flight 2.

7. CONCLUSIONS

In this paper, we presented an application of a robust identification approach to a Honeywell turbofan engine for the cruise regime. LTI MIMO parametric models were identified for the nominal and faulty engine cases. Recent developments on model validation concepts were also used, i.e. model error models were used to provide initial characterizations on the modelling uncertainty and to gain understanding of the shortcomings of the nominal identified models. The robust model was used to design an FDI $\mathcal{H}_\infty$ filter which has good robustness and performance properties for the LTI MIMO models but suffers from some disturbance rejection shortcomings for the nonlinear engine model due to phase uncertainty not included in the uncertainty characterization.

8. ACKNOWLEDGMENTS

The first author would like to thank Honeywell International Inc. for the opportunity to carry out this research as part of a Summer internship at the Vehicle Health Monitoring & Logistics Management Lab in Minneapolis. Special thanks to John Christoffel, Sunil Menon, Önder Uluyol and the rest of the team for their help and interest. We gladly acknowledge the help of Sachi Dash and Hanif Vhora for providing their knowledge and help as the designers of the engine simulation model.

REFERENCES

Balas, G.J., Doyle, J.C., Glover, K., Packard, A. and Smith, R. (1998). μ-Analysis and Synthesis Toolbox.

Gorinevsky, D., Dittmar, K., Mylaraswamy, D. and Nwadiogbu, E. (2002). Model-Based Diagnostics for an Aircraft Auxiliary Power Unit. In: *IEEE Conference on Control Applications*. Glasgow, Scotland.

Kim, K., Ball, C. and Nwadiogbu, E. (2004). Fault Diagnosis in Turbine Engines using Neural Networks Clustering Technique. In: *American Control Conference*. Boston, MA.

Ljung, L. (1987). *System Identification Theory for the User*. Information and System Sciences Series. Prentice-Hall, Inc.

Ljung, L. (1999). Model validation and model error modeling. In: *The Astrom Symposium on Control* (Wittenmark, B. and Rantzer, A., Eds.). Studentlitteratur. Lund, Sweden. pp. 15–42.

Ljung, L. (2001). Estimating Linear Time-Invariant Models of Nonlinear Time-Varying Systems. *European Journal of Control* **(2-3)**(7), 203–219.

Ljung, L. (2002). *System Identification Toolbox User's Guide*. The Mathworks, Inc.

Marcos, A. (2004). Aircraft Applications of Fault Detection and Isolation Techniques. PhD thesis. Department of Aerospace Engineering and Mechanics, University of Minnesota-Minneapolis. Minneapolis,MN.

Reinelt, W, Garulli, A. and Ljung, L. (2002). Comparing different approaches to model error modeling in robust identification. *Automatica* **38**(5), 787–803.

Reinelt, W, Garulli, A., Ljung, L., Braslavsky, J.H. and Vicino, A. (1999). Model error concepts in identification for control. In: *IEEE Conference on Decision and Control*. Phoenix, AZ, USA. pp. 1488–1493.

Smith, R. and Doyle, J. (1992). Model Validation: a Connection between Robust Control and Identification. *IEEE Transactions on Automatic Control* **37**(7), 942–952.

Uluyol, O., Kim, K., Menon, S. and Nwadiogbu, E (2003). Synergistic Use of Soft Computing Technologies for Fault Detection in Gas Turbine Engines. In: *IEEE International Workshop on Soft Computing in Industrial Applications*. Binghamton, NY.