

FDI OF ABRUPT FAULTS WITH COMBINED STATISTICAL DETECTION AND ESTIMATION AND QUALITATIVE FAULT ISOLATION

Eric-J. Manders, * Gautam Biswas ^{*,1}

* Department of Electrical Engineering and Computer Science
Vanderbilt University, P.O. Box 1679 Sta. B, Nashville, TN 37235
Email: {eric.j.manders,gautam.biswas}@vanderbilt.edu

Abstract: We present a novel approach for model-based FDI of abrupt faults in process components of continuous dynamic systems. Abrupt faults refer to parameter value changes that occur much faster than the nominal process dynamics, and component faults refer to faults that correspond to physical parameters in the bond graph model of a system. These faults cause transients in system behavior. We analyze this transient response by combining statistical detection and estimation with a model-based qualitative fault isolation engine. Detection uses the discrete wavelet transform in combination with a statistical decision function. Fault isolation is based on analysis of fault signatures in a qualitative framework. We demonstrate robust detection for small faults, and robust fault isolation that becomes more precise for larger faults. *Copyright © 2003 IFAC*

Keywords: Abrupt faults, qualitative fault isolation, transient detection

1. INTRODUCTION

In real applications, model-based approaches to fault detection and isolation (FDI) for continuous dynamic systems have to contend with uncertainties in the model structure and its parameter values, as well as noise in measurement data. An important focus of current research in model-based FDI is the design of *robust* FDI schemes, where the FDI result is made insensitive to uncertain information.

Research in model-based FDI of continuous dynamic systems has taken place largely in two communities. Control engineering approaches to FDI use analytical models in state-space and transfer function forms. Robust solutions are achieved through disturbance decoupling that explicitly represents model uncertainty as unknown inputs to the system (Chen and Patton, 1998), and statistical detection and estimation techniques that handle measurement uncertainty. A second strand of research originates in the artificial intelli-

gence community (Weld and de Kleer, 1990), and is based on constraint analysis techniques that link faults to deviations in behavior. Constraint models describe system behavior qualitatively, using symbolic equations or causal relations, thereby implicitly accommodating model uncertainty. However, analyzing system behavior with qualitative models requires a mapping of numerical measurement data to a symbolic representation. This mapping, the *signal-to-symbol transformation*, determines the ability to accommodate data uncertainty.

The success of analytical model-based FDI approaches in real world applications has been considerable. However, solutions are often strongly tailored to the specific system, and require careful tuning of the design parameters to achieve good performance (Isermann, 1997). Therefore, it is not always clear how well a design will generalize to other applications. Real applications of qualitative techniques for continuous dynamic systems are unknown (although good results have been shown for discrete systems). Qualitative FDI researchers have assumed that the symbolic input is error free, and *ad hoc* approaches to the signal-to-symbol transformation make it difficult to quantify

¹ This work was supported in part through grants from the NASA-IS program (Contract number: NAS2-37143) and the DARPA IXO program (Contract number: F30602-96-2-0227).

performance for small faults given measurement uncertainty.

This paper describes a novel approach for model-based FDI of abrupt faults in component parameters of a continuous dynamic system. In our approach, component parameters are the generic physical parameters of the bond graph of a system. These parameters appear in multiple terms of the state space model matrices, and, parameter changes directly affect system dynamics. In other words, these are multiplicative faults. Abrupt faults correspond to changes that occur at time scales much faster than the nominal dynamics of the system. We model abrupt faults as discrete and persistent changes in the value of component parameters. Note that this is strictly a temporal abstraction and does not imply a large fault magnitude (Basseville and Nikiforov, 1993). An abrupt fault in a component parameter results in transients in the system variables. Typically, the transient behavior vanishes after an interval, and for certain faults no evidence of the fault is observable in the measurements after some time. FDI schemes based on parameter and state estimation techniques to FDI have an inherent low-pass behavior (Chen and Patton, 1998) that can smooth such a transient response. As a result, the detection sensitivity for these faults may be reduced.

Our approach, named TRANSCEND, is based on the analysis of the fault transient. A model-based fault isolation scheme for qualitative analysis of the fault transients was developed by Mosterman & Biswas (Mosterman and Biswas, 1999). It is a hypothesize-and-test approach based on a prediction of the transient behavior immediately after the onset of the fault. To achieve robustness against model uncertainty and noisy measurements, we combine this qualitative fault isolation scheme with statistical detection and estimation techniques. This solution realizes a signal-to-symbol transformation component that is tailored to the detection of fault transients, and support the extraction of features that describe the transient dynamics.

2. ROBUST MODEL-BASED FDI BASED ON TRANSIENT DETECTION AND ANALYSIS

The TRANSCEND approach explicitly decouples the fault detection and fault isolation tasks. Fault detection is based on a numerical residual and coupled with a symbolic residual generator. The numerical residual is computed as the difference between observed and nominal system behavior. The output of the symbol generator is then input to a qualitative model-based fault isolation scheme (Fig 1). We present an outline of the fault isolation scheme. Details are presented in (Mosterman and Biswas, 1999).

2.1 Qualitative fault isolation from fault transients

The fault isolation engine follows the generate-and-test approach to residual evaluation. The model repre-

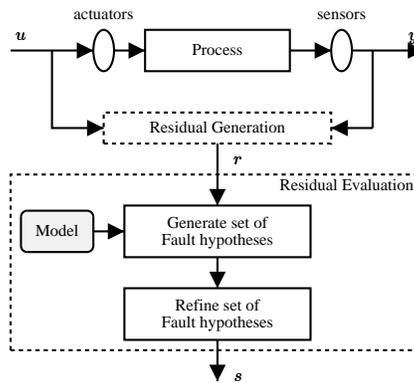


Fig. 1. TRANSCEND couples symbolic residual generation and model-based residual evaluation.

sentation is a *temporal causal graph* (TCG), a topological representation of the algebraic and temporal relations between variables in the system (Mosterman and Biswas, 1999). Vertices represent system variables and edge relations include the model parameters. Qualitative analysis is made possible by defining symbolic operations for the edge relations and symbolic values for the variables. The TCG can be derived automatically from the bond graph model for the system (Karnopp *et al.*, 1990; Mosterman and Biswas, 1999).

Qualitative transient behavior is expressed as a *fault signature* that describes the expected fault transient *immediately after the point of fault occurrence*. The signature corresponds to a qualitative interpretation of the Taylor series expansion of the residual around the point of fault occurrence (Manders *et al.*, 2000b). The order of the signature is defined by the highest derivative computed (a design parameter). Symbolic values for the elements of a signature are: ‘+’ for a positive or increasing value, ‘0’ for a zero or unchanged value, and ‘-’ for a negative or decreasing value. An unknown value is represented by ‘.’. The description of a fault signature in terms of the behavior around the point of fault occurrence is unique to the approach.

Fault isolation is triggered by the first non-zero magnitude symbol that is output by the signal-to-symbol generation module. This initial symbol reflects the magnitude deviation in the residual at the onset of the fault transient. The *hypothesis generation* step results in a set of fault hypotheses. A fault hypothesis consists of a candidate parameter with a direction of change for the parameter value and a fault signature for each of the measured variables. During *hypothesis refinement*, the signatures are compared with symbolic values computed from the measurements in a scheme called progressive monitoring, which will be illustrated in the example later. When a match fails, the candidate is dropped.

In previous work, the fault isolation scheme was implemented in a deterministic framework. Fault detection and the trigger for hypothesis generation were defined by the same event (Mosterman and Biswas, 1999). Computing the magnitude deviation symbol

was achieved with an instantaneous threshold comparison. Similarly, the derivative of the residual was computed with a discrete difference operator. The scheme was improved upon by using estimation techniques with noise suppression capabilities, or specific feature extraction techniques that detect a discontinuous onset of a fault transient (Manders *et al.*, 2000a).

2.2 Robust detection of fault transients

The transient signal detection problem can be formally stated in the hypothesis testing framework. Consider a (discrete time) signal x . Under the null hypothesis, H_0 , x consists only of noise, n , and under the alternative hypothesis, H_1 , x contains the signal, s , superimposed on the noise.

$$H_0 : x_k = n_k \quad H_1 : x_k = s_k + n_k.$$

A decision function that decides between the two hypotheses is ideally designed to exploit knowledge about the signal s . A key observation for the approach taken here is that a direct association of the decision function with the computation of symbols for fault isolation cannot fully exploit the fault transient response as an event of interest. Instead, the decision function should capture the knowledge about fault transients in a suitable *signal model*. Unfortunately, only minimal knowledge is available. Provided the fault does not result in an unstable system, the transient is a damped complex exponential signal with a possible discontinuous change at the onset. The component fault parameter, fault size, and time of fault occurrence are all unknown. Under these circumstances, defining a parametric signal model becomes unfeasible, and we resort to a non-parametric model instead. A suitable signal model for representing transient signals is a linear time-frequency (TF) transform. The decision function is then computed in the transform domain.

A framework transient detection based on this principle was first proposed in (Friedlander and Porat, 1989), using the discrete Gabor transform with an asymmetric damped exponential window function to match physical transient phenomena, and a generalized likelihood ratio test (GLRT) as a decision function. This scheme was evaluated in the context of TRANSCEND in (Manders and Biswas, 2001), and was found to have some limitations. The complex exponential basis functions in the Gabor transform do not easily allow discontinuities to be represented. Moreover, for transients whose location in the TF domain is unknown, the GLRT difficult to apply. The scheme described next addresses these problems.

2.3 A detection scheme in the time-frequency domain

The DWT is a linear transform that is very suitable to represent the non-stationary events in signals. The DWT has good localization properties of high frequency components, which is beneficial for

faults transients that exhibit a discontinuous change. We choose the 4-tap Daubechies wavelet for this study, and use the Fast Lifting Wavelet Transform (FLWT) (Sweldens, 1995) for our implementation. We compute the DWT in a sliding window over the data, to obtain fine grained tracking of the time-frequency evolution of the signal. A longer window results in higher detection sensitivity but at increased computational cost (which cannot be ignored in an on-line application). To determine a suitable window size, we exploit the fact that a fault transient is a damped signal, and most of the signal energy is present toward the onset of the transient. This allows us to choose a window that is smaller than the transient length.

For the decision function, we base our solution on a transient detection scheme developed by Wang & Willett (2001) The scheme is based on a DWT representation of the signal, and a decision function called a *power-law* statistic. If $C_{l,k}$ represents the squared coefficient of a DWT transform at index l of decomposition level k , the power-law statistic T_w for the DWT is defined as:

$$T_w(C) = \sum_{l=1}^L \sum_{k=1}^{2^L} (C_{l,k})^\nu, \quad (1)$$

where ν is a real valued exponent. To understand the role of ν , we consider two limit values. As $\nu \rightarrow \infty$, the detector selects the maximum coefficient, which is optimal if there is only one non-zero coefficient. This corresponds to a GLRT based detector that chooses the coefficient with the maximum power. For $\nu = 1$, the detector is essentially an energy detector in the transform domain. This would be optimal only if *all* coefficients are non-zero. The exponent ν lets us adjust the detector based on representation of the transient signal in the transform domain in an intuitive way. However, the optimal value for ν depends on the characteristics of the actual transient signals, and there is no analytic solution to determine ν . Wang & Willett (Wang and Willett, 2001) determined a range of values (through numerical analysis) where the decision function performs well for a variety of transient signals.

An enhanced version of the decision function is obtained by exploiting contiguity of real transient signals in both the time and frequency dimensions. This is accomplished by grouping neighboring coefficients. A grouping of three neighboring coefficients: $U_{l,k} = C_{l,k} + C_{l+1,2k-1} + C_{l+1,2k}$ was proposed, and is illustrated in Fig. 2.

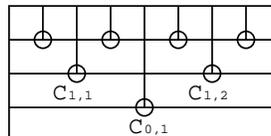


Fig. 2. Power-law contiguity for the DWT in a 4 level decomposition. Circles indicate groups of 3 neighboring coefficients.

A power-law decision function that exploits contiguity in the DWT coefficients, T_{wec3} , follows by replacing $C_{l,k}$ in Eq. 1 with $U_{l,k}$. This detector was found to have excellent performance for a wide range of transient signals (Wang and Willett, 2001).

2.4 Estimating the onset of the transient response

TRANSCEND analyzes the transient behavior of a fault immediately after the onset of the fault. Therefore, symbol generation must be initiated as close as possible to the fault onset. Given measurement noise, reliable fault detection occurs some time after the onset of the fault. Symbols computed around the time of detection will be a less accurate description of the transient dynamics at the point of fault occurrence. To improve on this we compute an estimated time of fault occurrence and align the symbol generation with that point in time. We follow a principle that is well known from abrupt change detection theory (Basseville and Nikiforov, 1993), where the change point is computed after detection by further analysis of the decision function. Our approach is illustrated in Fig. 3. Assume that the fault occurs at time t_f . As we slide the analysis window of length N over the transient signal the decision function increases, and crosses the threshold for detection h at time t_d . The decision function reaches a maximum value at t_p , which indicates that at that point, the alignment of the analysis window with the transient signal captures the most time-frequency energy in the signal.

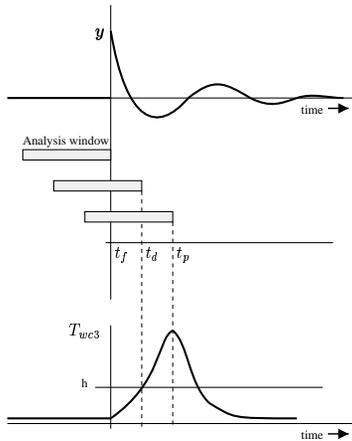


Fig. 3. Time line for detection with landmark points for alignment of the sliding window with the signal.

The maximum in the decision function provides the reference point that is used in the estimation of the time of fault onset, \hat{t}_f . We determine t_p for an observed fault transient, and compute \hat{t}_f as $\hat{t}_f = t_p - m$. The parameter m defines an offset, that will be discussed shortly. The goal is to compute \hat{t}_f so that $t_f < \hat{t}_f < t_d$. This avoids the case where $\hat{t}_f - t_f < 0$, which may result in generating the wrong symbol values. To determine m we perform a simulation study for the system. To avoid symbol generation errors we make

a conservative estimate and set $m = \min_{1 \leq i \leq N} (m_i)$, generated from N experiments. Note that when the threshold for detection is increased, the detection delay will increase also. Consequently, the relative benefit of computing \hat{t}_f increases also. For an increasing fault size, the detection delay decreases, and the gains from computing \hat{t}_f are diminished.

2.5 Symbol generation

We compute both the magnitude and the derivative estimation symbols using linear estimators. A basic sample mean estimator is used for the magnitude symbol. We select a short, length 5, estimator, to avoid excessive smoothing of a discontinuous onset for a transient. For the derivative estimation, a minimum variance unbiased estimator is used (Manders *et al.*, 2000a). For this estimator we select a long window to achieve maximum noise suppression. The combination of these two estimators is admittedly configured to perform especially well on transients that have a discontinuity at the onset.

Assuming that the system model is correct, threshold values for symbol generation are chosen to avoid fault isolation errors. This means that the initial hypothesis set should contain the correct candidate (although it will also contain spurious candidates), and that that candidate is not subsequently eliminated. While a true zero error probability is not practically possible, we can nevertheless set the thresholds to get a near zero-error rate under specific constraints. Thresholds are determined in a simulation experiment for all faults of interest and a fault size such that $PD=0.9$. Starting with low threshold values, thresholds are increased iteratively until a zero error rate is obtained for both the magnitude and the derivative symbols.

3. EXPERIMENTS WITH A DAMPED SPRING-MASS SYSTEM

We evaluate the TRANSCEND scheme using a simulated third-order damped spring-mass system shown in Fig. 3. The system consists of two masses, M_1 and M_2 , each connected to a damper, with friction coefficient R_1 and R_2 , respectively. A spring with stiffness parameter K connects the two masses. Our model parameters are generalized physical system phenomena, i.e., inertia parameters, I_1 and I_2 are used for masses M_1 and M_2 , respectively, a capacitance parameter, $C_1 = 1/K$ represents the spring, and resistances R_1 and R_2 represent the friction processes. The system has one input, a force F_1 acting on M_1 . The state vector for the system is $\mathbf{x} = (v_1, F_2, v_2)^T$. Fig. 5 shows the TCG, where the vertices corresponding to the state variables are f_3 , e_5 , and f_6 , respectively. The other vertices correspond to force and velocity values at different points in the system.

For this third order system we select a measurement vector (v_1, F_2, v_2) . For the nominal parameter values

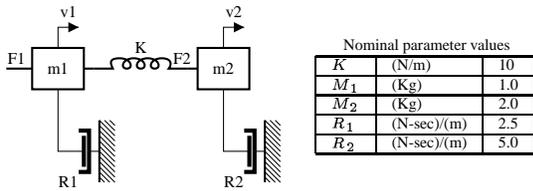


Fig. 4. Damped spring-mass system.

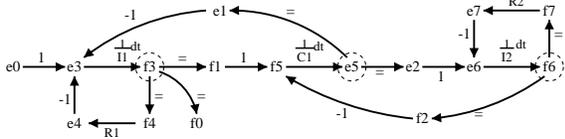


Fig. 5. Damped spring-mass system TCG.

(Fig. 3), the rise time of the system is approximately 0.2 (s). A suitable sampling rate, $t_s = 0.01$ (s), corresponds to a 5x oversampling rate. The input to the system is a unit step signal on F_1 . The nominal behavior is assumed to be the steady state response to the unit step input. We add zero-mean Gaussian noise with $\sigma = 0.01$ to all measured signals.

We design a DWT/power-law detector with an analysis window of $N=64$ samples, corresponding to a DWT decomposition of 6 levels ($\log_2(64)$). The optimal value of the power-law exponent for our system is determined in a simulation experiment, giving a value $\nu = 2.1$, which is in the (1.8–2.5) interval suggested in (Wang and Willett, 2001). We set the probability of false alarm, $PFA=10^{-4}$, and determine the value for the detection threshold, $h=3000$, by simulation. The offset m in the computation of \hat{t}_f was found to be 55 samples (zero errors in 100 realizations). As a reference for the time-frequency (TF) detector we use an energy detector (ED). The energy detector exploits no information about the signal and represents a lower bound for statistical detection performance.

As a fault scenario, consider an increase in the capacitance parameter, C_1^+ (physically this corresponds to a weakening of the spring). Fig. 6 shows the noise free residual for this fault. Table 1 shows performance

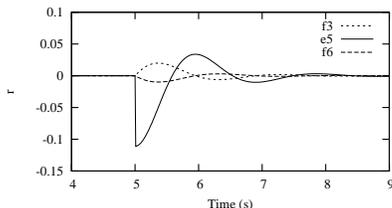


Fig. 6. Residual for fault C_1^+ and fault size 10%.

results (averages from 20 realizations). The time-frequency detector performs noticeably better than the energy detector, and for a fault size of 3% has excellent detection performance (probability of detection $PD=1$). For this fault size, $\hat{t}_f - t_d$ is -0.08 (s), which means that \hat{t}_f is 8 samples closer to t_f than t_d .

Fig. 7 shows the result for a simulation run with fault size 4%. The residual signal for measurement

e_5 , the DWT decomposition (absolute value of the detail coefficients as gray scale values), the power-law decision function, and the fault isolation sequence, are shown in top down sequence. For this realization, \hat{t}_f is slightly better than average for this fault size. The time $t_{hg} = 5.09$ (s), is the point where hypothesis generation is triggered (this is the *system time*, in real time, t_{hg} is delayed by the peak detection in the decision function). The delay between estimating the fault onset, and generating the first symbol, $t_{hg} - \hat{t}_f = 0.05$ (s), or 5 samples, the length of the magnitude estimation FIR filter.

Fault isolation occurs in two steps: at time step 0, which corresponds to t_{hg} , hypothesis generation is triggered by an observed magnitude deviation ('-') for measurement e_5 (F_2). Shown are the set of fault hypotheses for the generated candidates (we compute third order signatures). At time step 11, a '+' derivative symbol is computed for e_5 . The observed qualitative transient behavior is thus (-, +). The signature for fault C_1^+ for e_5 is (-, 0, +). This signature corresponds to a discontinuous change, because the first non-zero predicted derivative is opposite from the direction of the initial deviation. The matching algorithm indicates that the observed behavior matches this signature, because the progressive monitoring algorithm implies that this second order effect will propagate to an observed first order effect. No other candidate signature matches the observed behavior for e_5 so all other candidates are dropped. Consequently, the final hypothesis set contains only candidate C_1^+ . Note that the time from \hat{t}_f to time step 11 is 19 samples, the length of the FIR derivative estimator.

Table 2 shows the precision of the fault isolation results (the size of the final hypothesis set) for different fault sizes, with 10 experiments for each fault size. The 'count' column indicates the number of occurrences for each of the possible final hypothesis sets, grouped by fault size. For a fault size of 3%, fault isolation is not sufficiently precise, although we have seen that the probability of detection for this fault is very good. At a fault size of 5%, maximum fault isolation precision is consistently achieved. The true candidate is never dropped, regardless of fault size.

As a last illustration, Table 3 presents fault isolation results for fault, I_1^+ . The final hypothesis set may be one of three different sets over the range of fault sizes. The gradual reduction in the final hypothesis set is the result of symbols computed from different residual components.

size	SNR	PD_{TF}	PD_{ED}	$t_d - t_f$	$\hat{t}_f - t_f$
1.5%	14.0	0.7	0.65	0.19	0.07
2%	17.1	0.95	0.8	0.16	0.04
3%	20.5	1.0	0.9	0.11	0.03
4%	22.9	1.0	0.95	0.07	0.03
5%	24.8	1.0	1.0	0.03	0.02

Table 1. Fault detection performance for fault C_1^+ .

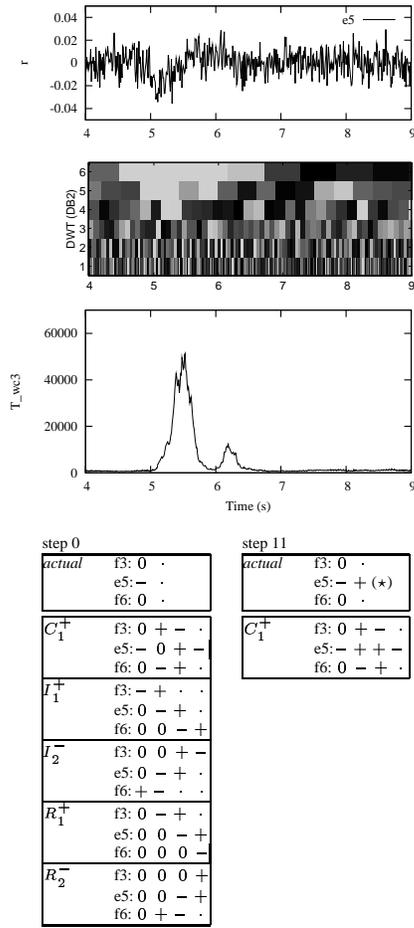


Fig. 7. TRANSCEND output for fault C_1^+ , fault size 4%.
 $t_d = 5.13$ (s), $\hat{t}_f = 5.03$ (s), $t_{hg} = 5.09$ (s).

size	SNR (dB)	fi nal hypothesis set	count
3%	20.5	no candidates generated	1
		$C_1^+, I_1^+, I_2^-, R_1^+, R_2^-$	7
		C_1^+	2
4%	22.9	$C_1^+, I_1^+, I_2^-, R_1^+, R_2^-$	6
		C_1^+	4
5%	24.8	C_1^+	10

Table 2. Fault isolation precision for fault C_1^+ .

size	SNR (dB)	fi nal hypothesis set	count
9%	23.2	$C_1^+, I_1^+, I_2^-, R_1^+, R_2^-$	10
10%	24.1	$C_1^+, I_1^+, I_2^-, R_1^+, R_2^-$	2
		I_1^+, R_1^+	8
12%	24.3	I_1^+, R_1^+	4
		I_1^+	6
20%	25.6	I_1^+	10

Table 3. Fault isolation precision for fault I_1^+ .

4. CONCLUSIONS AND DISCUSSION

We have developed a robust scheme for FDI of abrupt faults in components of continuous dynamic systems. The method is based on analysis of fault transients in the measurements. Statistical detection and estimation techniques are coupled with a qualitative model-based fault isolation scheme to achieve a robust solution that accomodates both measurement and model

uncertainty. By separating the detection and isolation tasks the performance of each task can be determined systematically. We exploit a signal model suitable for fault transients and a statistical decision function to achieve good detection performance. Good detection performance can be achieved for a fault size that is smaller than required for consistently precise fault isolation. Robust fault isolation shows that diminishing information results in lower fault isolation precision while maintaining accuracy.

REFERENCES

- Basseville, M. and I. V. Nikiforov (1993). *Detection of abrupt changes: theory and applications*. Prentice Hall, Englewood Cliffs, New Jersey.
- Chen, J. and R. J. Patton (1998). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic publishers, Boston, MA USA.
- Friedlander, B. and B. Porat (1989). Detection of transient signals by the Gabor representation. *IEEE Trans. Acoust., Speech Signal Processing* **37**(2), 169–180.
- Isermann, R. (1997). Supervision, fault-detection and fault-diagnosis methods – an introduction. *Control Engineering Practice* **5**(5), 639–652.
- Karnopp, D. C., D. L. Margolis and R. C. Rosenberg (1990). *Systems Dynamics: A Unified Approach*. second ed.. John Wiley and Sons, New York.
- Manders, E.-J. and G. Biswas (2001). Transient detection and analysis for diagnosis of abrupt faults in continuous dynamic systems. In: *Proc. IEEE Int. Workshop Intelligent Sig. Proc.*, Budapest, Hungary, pp. 15–20.
- Manders, E.-J., G. Biswas, P. J. Mosterman, L. A. Barford and R. J. Barnett (2000a). Signal interpretation for monitoring and diagnosis, a cooling system testbed. *IEEE Trans. Instrum. Meas.* **49**(3), 503–509.
- Manders, E.-J., S. Narasimhan, G. Biswas and P. J. Mosterman (2000b). A combined qualitative/quantitative approach for fault isolation in continuous dynamic systems. In: *Proc. 4th IFAC Symp. SAFEPROCESS*. Budapest, Hungary, pp. 1074–1079.
- Mosterman, P. J. and G. Biswas (1999). Diagnosis of continuous valued systems in transient operating regions. *IEEE Trans. Syst., Man Cybern. A* **29**(6), 554–565.
- Sweldens, W. (1995). The lifting scheme, a new philosophy in biorthogonal wavelet constructions. In: *Wavelet Applications in Signal and Image Processing III* (A. F. Laine and M. Unser, Eds.). pp. 68–79. SPIE.
- Wang, Z. and P. Willett (2001). All-purpose and plug-in power law detectors for transient signals. *IEEE Trans. Signal Processing* **49**(11), 2454–2465.
- Weld, D. and de Kleer, J., Eds.) (1990). *Readings in Qualitative Reasoning about Physical Systems*. Morgan Kaufmann, San Mateo, CA.