Available online at www.sciencedirect.com

SCIENCE DIRECT

Control Engineering Practice 14 (2006) 635–644

CONTROL ENGINEERING PRACTICE

www.elsevier.com/locate/conengprac

Model-free actuator fault detection using a spectral estimation approach: the case of the DAMADICS benchmark problem

F. Previdi[a], T. Parisini[b,*]

[a]Department of Management and Information Science DIGI, University of Bergamo, 24044 Dalmine (BG), Italy
[b]Department of Electrical, Electronic and Computer Engineering-DEEI, University of Trieste, 34127 Trieste, Italy

Received 4 March 2004; accepted 4 April 2005
Available online 13 June 2005

Abstract

This paper presents the application to the DAMADICS benchmark fault detection problem of a model-free fault detection technique based on the use of a specific spectral analysis tool, namely, the *Squared Coherency Functions* (SCFs). The detection of a fault is achieved by on-line monitoring the estimate of the squared coherency function, which is sensitive to the occurrence of significative changes in the plant dynamics. The alarm threshold are based on the estimates of the confidence intervals of the SCF. Results on both simulation and real data of the DAMADICS benchmark (which is developed to approximate the industrial process in a sugar factory located in Lublin, Poland) are outlined.

Keywords: Model-free fault detection; Spectral methods; Squared coherency function; Actuator fault detection; DAMADICS project

1. Introduction

This work is focused on the development of a fault detection method for application to an actuator benchmark specifically designed for industrial fault diagnosis (FD) studies. The Damadics benchmark, described in Bartys, Patton, Syfert, de las Heras, and Quevedo (2005), fills the lack of well-defined benchmarks for FD testing of industrial fluid actuators. In fact, it is strongly focused on application to real industrial actuator valves and has a number of features that makes it an ideal test bench for FD methods. Specifically, its main characteristics are:

(a) The Damadics benchmark is independent on the fault detection methods/algorithms. This is possible since it has been based on a careful study of the physical phenomena that can give origin to faults in the actuator system.
(b) The Damadics benchmark clearly defines the process and the data sets; the fault scenarios are univocally set and standardized. This is done in view of industrial applicability of the proposed FD solutions, to cut off those methods/algorithms that have no practical feasibility.
(c) The Damadics benchmark specifies the way of reporting results, allowing a clear performance comparison of different methods.

Specifically, in this paper, a FD method based on *Squared coherency function* (SCF) estimation is investigated and results about the application to the Damadics benchmark are reported. Since early 1970s, a lot of approaches to fault diagnosis have been developed that may be roughly classified as (i) *model-based* FD approaches and (ii) *model-free* FD approaches (see Willsky, 1976; Isermann, 1984; Gertler,1988,1998; Frank, 1990; Patton, Frank, & Clark, 1989; Chen & Patton, 1999 and references cited therein). In this work an algorithm is proposed which can be considered a model-free technique, since it is not necessary to resort

*Corresponding author: Tel.: +39 040 5587138; fax: +39 040 5583460.
E-mail address: parisini@units.it (T. Parisini).

0967-0661/$ - see front matter
doi:10.1016/j.conengprac.2005.04.001

to any explicit I/O model of the plant (at least in time domain).

In general, model-free methods are considered those that do not require the direct estimation of a time-domain model of a plant to generate symptoms. So, they could provide powerful solutions in the industrial-oriented setup defined by the Damadics benchmark. Roughly speaking, a model-free method mainly consists in addressing the FD problem from the control-board operator's point of view. The basic idea is to monitor on-line the measurements of the control system variables without the need for defining explicit dependence laws in time-domain among them. By analysis of the measured variables the operator can decide about the operating mode of the plant and raise alarms. In this context, sometimes the problem can be assimilated to a pattern-recognition problem (see, for instance, Guglielmi, Parisini, & Rossi, 1995 and the references cited therein).

The use of the SCF, computed using carefully chosen I/O measurements from the plant, seems to fit this point of view. In fact, it is sensitive to changes in the plant dynamics, which is a typical consequence of a fault occurrence. For instance, assume that the fault-free dynamic behavior of the plant is linear. This situation is quite common in industrial processes where the nominal behavior of the plant can be regarded as a linear one due to the presence of the regulation system. One of the main (and undesired) effects of the occurrence of a fault is the deviation from the nominal healthy mode of operation. So, the effect of a fault occurring on the plant is simply assumed to be a nonlinear dynamical perturbation to the above linear one. Then, FD can be achieved by designing a fault detector sensitive to the arising of nonlinear components in the plant dynamics. Such a detector can be based on the computation of the SCF using I/O measurement signals from the plant, without resorting neither to a model of the plant, nor to a model of the fault dynamics. In fact, it is well known that the SCF could be effectively used as a nonparametric tool for identification of nonlinear dynamics in I/O data (Haber, 1985). As an example, in Young and Patton (1990), the SCF is used to assess the accuracy of an identification as a function of frequency, during model identification of helicopter dynamics. Similarly, in (Previdi, Lovera, & Mambretti, 1999) the SCF is used to test for the presence of nonlinear dynamics in I/O data measured in a urban drainage network. In this work, we will use these ideas to develop a FD algorithm specifically tailored for the Damadics benchmark process. The analysis will regard only those faults whose primary nature is abrupt. This choice was done to focus the work on the evaluation of the basic performance parameters of the FD algorithm, i.e. the quality of the detection (in terms of false alarms and missed detection) and the quickness of the algorithm in detecting the fault.

The paper is organized as follows: In Section 2, some preliminary definitions of variables commonly used in spectral analysis are given. This is useful in view of dealing with some features of the spectral estimates on which the FD method is based. Then the SCF-based methodology for FD is presented. Specifically, in Section 3, the "concept" at the basis of the FD method is described and a (possible) extension for application to the present case is proposed. Then, Section 5 is devoted to the outline of the algorithm proposed in this work. This Section is preceded by a short description of the main features of the benchmark definition about the time schedule of the fault scenarios and the criteria for performance evaluation (Section 4). The application to the Damadics benchmark problem is described in Section 6, which contains extensive simulation results, showing the effectiveness of the proposed FD technique. Finally, in Section 7 application to the detection of a fault using experimental process measurements with artificially generated faults.

2. Preliminary definitions

First of all, for the reader's convenience, some definitions of variables widely used in spectral analysis of stationary stochastic processes are recalled (Jenkins & Watts, 1968; Ljung, 1999; Papoulis, 1973).

Let us consider a real continuous stationary process $\xi(t)$. The *power spectrum* $\Gamma_{\xi\xi}(f)$ is defined as the Fourier transform of the autocovariance function $\gamma_{\xi\xi}(\tau)$, i.e.

$$\Gamma_{\xi\xi}(f) = \int_{-\infty}^{+\infty} \gamma_{\xi\xi}(\tau)\mathrm{e}^{-j2\pi f\tau}\,\mathrm{d}\tau. \tag{1}$$

When considering the problem of estimating the power spectrum from a finite realization of length T, the following *power spectrum estimator* can be used:

$$C_{\xi\xi}(f) = \int_{-T}^{T} c_{\xi\xi}(\tau)\mathrm{e}^{-j2\pi f\tau}\,\mathrm{d}\tau, \tag{2}$$

where $c_{\xi\xi}(\tau)$ is an estimator of the autocovariance function.

It is well-known that the estimator in Eq. (2) has poor variance properties. Specifically, its variance does not approach zero as T increases, namely the estimator is *inconsistent*. A possible solution to overcome inconsistency is to use the following *smoothed power spectrum estimator*:

$$\bar{C}_{\xi\xi}(f) = \int_{-\infty}^{+\infty} w(\tau)c_{\xi\xi}(\tau)\,\mathrm{e}^{-j2\pi f\tau}\,\mathrm{d}\tau. \tag{3}$$

The function $w(\tau)$ is called *lag window*. It must satisfy the following conditions:

(i) $w(0) = 1$,
(ii) $w(\tau) = w(-\tau)$,
(iii) $w(\tau) = 0, |\tau| > T$.

These are widely used standard conditions for the lag window. Specifically, conditions (i)–(ii) mean that $w(\tau)$ must be an autocorrelation function, so that the obtained estimate is a true power spectrum. Condition (iii) is needed because only finite realization with length T of ξ are considered for estimation. In practice, the data set with length T is divided into N sub-series each of length M and the smoothed estimator is computed as the sample mean of the N estimates obtained for each sub-series. So, condition (iii) can be replaced by

(iii) $w(\tau) = 0, |\tau| \geqslant M,\ \ M < T$

since covariances need to be computed only up to lag M.

Using the convolution property, Eq. (3) can be written as follows:

$$\bar{C}_{\xi\xi}(f) = \int_{-\infty}^{+\infty} W(y) C_{\xi\xi}(f-y)\,\mathrm{d}y, \tag{4}$$

where the *frequency window* $W(f)$ is defined as the Fourier transform of the lag window:

$$W(f) = \int_{-\infty}^{+\infty} w(\tau) \cdot \mathrm{e}^{-j2\pi f\tau}\,\mathrm{d}\tau. \tag{5}$$

The definitions given above can be extended to deal with two continuous stationary processes $\xi(t)$ and $\eta(t)$. Specifically, the *cross power spectrum* $\Gamma_{\xi\eta}(f)$ can be defined as the Fourier transform of the cross covariance function $\gamma_{\xi\eta}(\tau)$, i.e.

$$\Gamma_{\xi\eta}(f) = \int_{-\infty}^{+\infty} \gamma_{\xi\eta}(\tau)\mathrm{e}^{-j2\pi f\tau}\,\mathrm{d}\tau. \tag{6}$$

The corresponding *cross power spectrum estimator* is

$$C_{\xi\eta}(f) = \int_{-T}^{T} c_{\xi\eta}(\tau)\mathrm{e}^{-j2\pi f\tau}\,\mathrm{d}\tau, \tag{7}$$

where $c_{\xi\eta}(\tau)$ is an estimator of the cross covariance function.

Smoothing is a commendable procedure even for cross spectra and, similarly as above, the *smoothed cross power spectrum estimator* can be defined as

$$\bar{C}_{\xi\eta}(f) = \int_{-\infty}^{+\infty} W(y) C_{\xi\eta}(f-y)\,\mathrm{d}y. \tag{8}$$

Finally, the SCF (Stoica & Moses, 1997; Jenkins & Watts, 1968) is introduced as

$$\kappa^2_{\xi\eta}(f) = \frac{|\Gamma_{\xi\eta}(f)|^2}{\Gamma_{\xi\xi}(f)\Gamma_{\eta\eta}(f)} \tag{9}$$

together with the corresponding smoothed estimator

$$\bar{K}^2_{\xi\eta}(f) = \frac{|\bar{C}_{\xi\eta}(f)|^2}{\bar{C}_{\xi\xi}(f)\bar{C}_{\eta\eta}(f)}. \tag{10}$$

Smoothed estimates of the SCF are at the basis of the FD algorithm proposed in the present paper.

3. A "concept" for the development of a FD algorithm

In this section, the *concept* at the basis of the proposed FD algorithm is described. Specifically, the basic idea of the algorithm is shortly outlined in an "ideal" context.

It is well known that, given a linear system in continuous time with input $u(\cdot)$, output $y(\cdot)$ and frequency response $G(j2\pi f)$, in the absence of disturbances, the following relations are true:

$$\Gamma_{yu}(f) = G(j2\pi f)\Gamma_{uu}(f), \tag{11}$$

$$\Gamma_{yy}(f) = |G(j2\pi f)|^2 \Gamma_{uu}(f). \tag{12}$$

By substituting Eqs. (11), (12) into Eq. (9), it is possible to obtain immediately

$$\kappa^2_{uy}(f) = \frac{|G(j2\pi f)|^2\Gamma^2_{uu}(f)}{\Gamma_{uu}(f)|G(j2\pi f)|^2\Gamma_{uu}(f)} = 1, \quad f \in \Re. \tag{13}$$

Now, suppose that, in the absence of disturbances, a fault is affecting the system at the time $t > t_{from}$. Suppose that the fault modifies the system dynamics, *introducing observable nonlinear dynamics in the system output* (see Vemuri & Polycarpou, 1996). If so, the following is rigorously true:

- if $t < t_{from}$, then $\kappa^2_{uy}(f) = 1,\ f \in \Re$,
- if $t \geqslant t_{from}$, then $\kappa^2_{uy}(f) < 1\ f \in \mathscr{F} \subseteq \Re$.

In qualitative terms, the meaning of the above expressions is the following: prior to the occurrence of a fault, the SCF κ^2_{uy} is identically equal to one, since the fault-free plant has linear dynamics. After the occurrence of a fault κ^2_{uy} is less than one for some frequency, since the plant affected by the fault exhibits nonlinear dynamics. So, a FD scheme can be conceived and based on a statistical test on the smoothed SCF estimator, at given nominal significance level α, under the hypotheses

- H_0: $\bar{K}^2_{uy}(\cdot) = 1$,
- H_1: $\bar{K}^2_{uy}(\cdot) < 1$.

The possibility that this specific "ideal" framework could be used in FD on a controlled plant is motivated by practical considerations: whilst it is very common to suppose the plant nominal mode of operation to be approximately *linear* (e.g. the plant is linearized about a given steady state trajectory by the feedback controller), the same cannot in general be ensured when a faulty mode of operation is considered. In this respect, a fault is regarded as a *nonlinear perturbation* of the *linear* closed loop plant dynamics. Consequently, in very qualitative terms, the detection of a fault can be achieved by means of a quantity, in this case the SCF, that is sensitive to the occurrences of events introducing nonlinearity in the plant dynamics (Haber, 1985). This is exactly the case that have been originally studied in (Previdi & Parisini, 2001) and applied to a nonlinear plant which is feedback linearized by a controller.

However, in the present case, devoted to the Damadics benchmark, it is necessary to consider that the plant *has not perfectly linear dynamics even in fault-free conditions*. So, the proposed ideal concept cannot be used as it is. But it is reasonably supposed that the controlled plant can be characterized by a *nominal fault-free estimate* $\bar{K}^2_{uy,ff}(f)$, which can be regarded as a kind of "*model reference*" for the fault-free plant. So, an extended FD algorithm can be designed based on real-time computation of the SCF estimate. The FD scheme must be able to react to deviations of the current SCF estimate from the fault-free SCF, as a consequence of the occurrence of a fault. To do so, it is necessary to define, in a statistical sense, a couple of *limit functions*: if the current SCF estimate is outside these limits it must be considered different from the fault-free SCF. The advantage of using a spectral based procedure is that the *confidence interval functions* (at a given significance α) of the fault-free SCF estimate $\bar{K}^2_{uy,ff}(f)$ are natural candidates to develop the limit functions to be used as alarm thresholds.

4. Damadics benchmark definition

In the present section, a brief outline of the Damadics benchmark definition is given. Specifically, the testing conditions for the FD algorithm are given, both for simulation data and for real process measurements with artificially generated faults. In particular, the benchmark definition (Bartys & Syfert, 2002) needs the use of specific performance indicators for algorithm evaluation and it clearly defines the time schedule of the tests (e.g. how long is the setup time, how long is the fault development,...).

The benchmark considers only nineteen possible fault scenarios (among the unlimited number of possible faults), regarded as the most significant for FD algorithm testing. The following definitions are given referring to simulation data, but they are applied with no changes also to experimental process measurements with artificially generated faults.

4.1. Definition of the time development of a fault scenario

For all scenarios two time zones are defined (see Fig. 1):

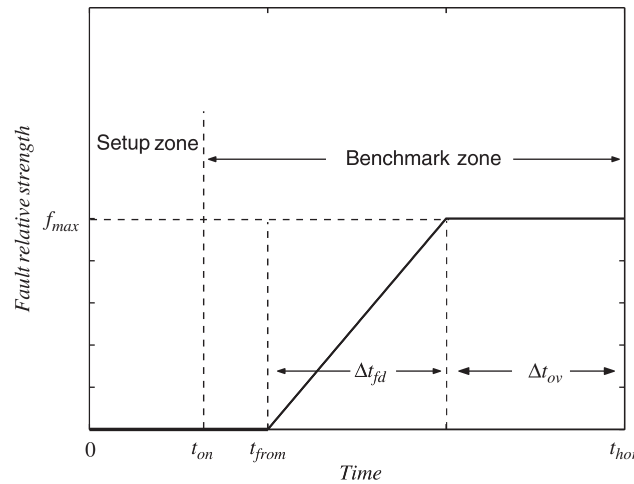

Fig. 1. Time development of a fault scenario. The value $f_{\max}$ is the maximum relative strength of the fault. When an abrupt fault is occurring, $\Delta t_{fd} = 0$. So, the time profile is a step.

(i) The *Setup zone*, used for proper tuning of the FD algorithm before testing. This zone is limited by the time t_{on}.
(ii) The *Benchmark zone*, used for FD algorithm testing and performance evaluation. This zone begins at t_{on} and it is limited by the time t_{hor}. This time interval is split into two parts:
   - from t_{on} to t_{from} the FD algorithm is active but no faults are occurring on the plant. So, the algorithm is expected not to rise any alarm.
   - from t_{from} to t_{hor} the FD algorithm is active and a fault is affecting the plant. A fault could be *incipient*, i.e. it gradually develops increasing its intensity from t_{from} for a time interval of duration Δt_{fd} or it could be *abrupt*, i.e. it is suddenly applied at the maximum of its intensity (i.e., $\Delta t_{fd} = 0$). The last time slice, from the end of the interval Δt_{fd} to t_{hor} is the observation time interval Δt_{ov}.

All results (figures, performance indexes, etc.) are referring to these time zones.

4.2. Definition of the performance indicators for evaluation of a FD algorithm

As already said, the present work is focused only on Fault Detection and not on Isolation. So, in the following, the indicators will be reported only for FD algorithm evaluation. The complete performance indicator set, is specified in Bartys and Syfert (2002).

First, the *detection signal* $D(t)$ is introduced. It is a binary signal defined for $t_{on} \leqslant t \leqslant t_{hor}$. It can assume only two values: $D(t) = 0$ if the FD algorithm is not detecting a fault; $D(t) = 1$ if the FD algorithm is raising an alarm.

Then, the *detection time* t_{dt} is defined: it is the length of the time interval from the beginning of the fault occurrence (at t_{from}) up to the last leading edge of the detection decision signal $D(t)$ (the time instant when $D(t)$ assumes definitely the value 1). So, the detection time is an indicator of how long a time the FD algorithm takes to definitely detect the fault.

Finally, two performance indicators are introduced:

(i) the *true detection rate*

$$r_{td} = \frac{\int_{t_{from}}^{t_{hor}} D(\tau)\,d\tau}{t_{hor} - t_{from}}. \tag{14}$$

It is defined so that $0 \leqslant r_{td} \leqslant 1$. It represent the time fraction of the benchmark zone during which the FD algorithm has *correctly* detected the fault: the fault is acting on the plant and the FD algorithm has revealed it.

(ii) the *false detection rate*

$$r_{fd} = \frac{\int_{t_{on}}^{t_{from}} D(\tau)d\tau}{t_{from} - t_{on}}. \tag{15}$$

Also the false detection rate is defined so that $0 \leqslant r_{fd} \leqslant 1$. It represent the time fraction of the benchmark zone during which the FD algorithm has *wrongly* detected the fault: the fault is not acting on the plant and the FD algorithm is raising a false alarm.

5. Outline of the FD algorithm

In this section, the outline of the proposed FD algorithm is given. The algorithm is based on two main steps.

The first one is a preliminary *off-line* stage performed on fault-free data. It is devoted to the choice of the parameters for SCF estimation.

The second step is the *on-line* testing of the algorithm, according to the time schedule described in Section 4.1 and to the performance evaluation criteria described in Section 4.2.

5.1. Preparation of the tools for spectral estimation

First of all, the user must choose the I/O variables u and y to be used for the estimation of the SCF. This choice must be "physically driven", such as to say, the algorithm must be used with those measurements whose cause/effect relationship is significantly affected by a fault occurrence.

Then, the problem of the choice of a window for smoothing the spectral estimates can be considered. Notice that this step can be skipped if the user has a priori knowledge about the window function useful for effective smoothing. Otherwise, in Previdi and Parisini (2001) an algorithm for a low-bias estimation of the SCF has been proposed. The algorithm needs a set of the chosen I/O fault-free data. It provides a window function $w_o^M(\tau)$, among a finite number of user-defined windows. The width value M_o is estimated by numerical minimization of an approximate expression of the bias $B[\bar{K}_{uy}^2(f)]$ of the smoothed SCF estimator, which expression can be found in Lohnberg (1978).

Results about the application of this procedure to the Damadics benchmark data are described in Section 6.1.

5.2. FD algorithm

In the following, the FD algorithm is described.

Specifically, the algorithm needs an *initial setup* where, using the fault-free data from 0 to t_{on} (the *Setup zone*), the alarm thresholds are computed.

Then, the algorithm is tested using the data from t_{on} to t_{hor} (the *Benchmark zone*). The algorithm raises alarms subject to the verification of assigned conditions.

5.2.1. FD algorithm initial setup (definition of the alarm thresholds)

First of all, a time t_{st} is defined: the data in the interval $[0, t_{st}]$ are used to obtain the first estimate $\bar{K}_{uy,t_{st}}^2(f)$ of the SCF (see Remark 1). Then, for each $t \in (t_{st}, t_{on}]$ do the following.

(1) Compute a smoothed estimate of the *nominal fault-free value* of the SCF, namely $\bar{K}_{uy,t}^2(f)$ using the data in $[0, t]$. The smoothing is performed using the window family resulting from Section 5.1.

(2) Assign a *significance level* α and, for the current estimate $\bar{K}_{uy,t}^2(f)$, compute the $100(1-\alpha)\%$ *confidence interval functions* $[k_{uy,t,\alpha}^-(f), k_{uy,t,\alpha}^+(f)]$.

The computation of the confidence intervals for the SCF estimate is performed by a standard approach in statistics based on the so-called "variable stabilizing transformations", whose aim is to obtain an expression for the variance of an estimator which is independent by the mean. In the present case, the variance smoothed (not squared) coherency estimator can be transformed into the following one:

$$\bar{Y}_{uy} = \frac{1}{2}\ln\frac{1 + |\bar{K}_{uy}|}{1 - |\bar{K}_{uy}|} \tag{16}$$

which has variance approximately independent of frequency, i.e.

$$\mathrm{var}[\bar{Y}_{uy}] \approx \frac{I}{2T} \tag{17}$$

and for which is reasonable to assume Normal distribution. So, approximate confidence intervals can be computed. In Eq. (17), $I = \int_{-\infty}^{\infty} W(f)\,df$.

(3) Based on the confidence interval functions obtained at the previous step, compute the *upper* and *lower alarm threshold functions* according to the following:
 (i) *Upper Alarm Threshold Function*:

$$\mathscr{A}^{+}(f) = \sup_{t \in (t_{st}, t_{on}]} k^{+}_{uy,t,\alpha}(f) \quad \forall f \in \mathscr{F} \tag{18}$$

 (ii) *Lower Alarm Threshold Function*:

$$\mathscr{A}^{-}(f) = \inf_{t \in (t_{st}, t_{on}]} k^{-}_{uy,t,\alpha}(f) \quad \forall f \in \mathscr{F} \tag{19}$$

So, the alarm thresholds are computed on the basis of the confidence intervals using a "worse case" criterion: at each time t during the setup, compare the confidence interval of the current SCF estimate, at each frequency, with those estimated at the previous time and choose the larger.

Summarizing, the results of the initial setup of the algorithm are the alarm threshold functions $\mathscr{A}^{-}(f)$ and $\mathscr{A}^{+}(f)$. These functions are computed on the basis of the confidence interval functions of the SCF estimates at a fixed significance level α, which directly affects the fault detectability. In fact, it is expected that, as α increases, the false alarm rate should decrease and the time interval between the fault occurrence and its detection (namely t_{dt}) should increase. In fact, for larger α values the confidence intervals will be larger and so also the alarm threshold. Further comments on this topic and the results about simulation data are shown in Section 6.

5.2.2. FD algorithm testing

The testing of the algorithm in the *Benchmark zone* (Section 4.1) is straightforward: for each $t \in (t_{on}, t_{hor}]$ an estimate $\bar{K}^2_{uy,t}(f)$ is computed, using the data in $[0, t]$. Then, if $\bar{K}^2_{uy,t}(f) > \mathscr{A}^{+}(f)$ or if $\bar{K}^2_{uy,t}(f) < \mathscr{A}^{-}(f)$ for any $f \in \mathscr{F}$, an alarm is generated.

Remark 1. The time instant t_{st} is defined in the interval $[0, t_{on}]$ by practical reasons. It is used to obtain the first estimate of the coherency function. In fact, being the coherency function a spectral variable, a finite batch of data is needed to obtain an estimate, namely the data in the time interval $[0, t_{st}]$.

6. Application to simulation data

The data used in this section have been generated using the software simulator available in the Simulink library for the Damadics project according to the standardized procedure described in Bartys and Syfert (2002) and Bartys et al. (2005).

Table 1
Fault scenarios considered in the present work

Fault	Component	Strength	Time development
f_1	Control valve	Big	Abrupt
f_2	Control valve	Big	Abrupt
f_7	Control valve	Big	Abrupt
f_8	Servomotor	Big	Abrupt
f_{10}	Servomotor	Big	Abrupt
f_{11}	Servomotor	Big	Abrupt
f_{12}	Positioner	Big	Abrupt
f_{14}	Positioner	Big	Abrupt
f_{15}	Positioner	Big	Abrupt
f_{17}	External	Big	Abrupt
f_{18}	External	Big	Abrupt
f_{19}	External	Big	Abrupt

As already said, the analysis is focused on those faults whose primary nature is abrupt (i.e., $\Delta t_{fd} = 0$ with reference to Fig. 1).

So, 12 fault scenarios have been considered of the 19 possible, choosing three scenarios for each component of the actuator, so that all the components of the actuator are examined (see Table 1). The characteristics of each fault are described in Bartys et al. (2005).

All the experiments are performed with sampling time $T_s = 1$ s. The time schedule used in all the experiments is the following (see Section 4.1 and Fig. 1):

(1) $t_{on} = 500$ s is the end time of the setup zone (computation of the alarm thresholds);
(2) $t_{from} = 900$ s is the time instant at which each fault occurs;
(3) $t_{hor} = 1800$ s is the time instant at which each experiment ends.

The length of the batch of data used to compute the first SCF estimate in the setup zone is $t_{st} = 256$ s. In the following, results coming from the application of the algorithm described in Section 5 to the 12 fault scenarios of Table 1 are described following the same outline of Section 5.

6.1. Preparation of the tools for spectral estimation

First of all, the I/O variables to be used for the estimation of the SCF have been chosen. The value of the *rod displacement measurement* X' and the *medium flow* F, generated by simulations including measurement noise, have been used for the detection of all the considered faults, except f_{17}: in this case, due to the physical nature of the fault, the *pressure on the control valve inlet* P_1 and the *pressure on the control valve outlet* P_2 have been used. These choices have been driven by the physics of the system, but they are not the unique possible choices. As a matter of example, detection of f_7

could be achieved by using also the *rod displacement measurement* X' and the *fluid temperature* T_1.

Then, using the algorithm shortly described in Section 5.1, the following window for smoothing the spectral estimates has been chosen (Papoulis, 1973):

$$w_o^M(\tau) = \frac{1}{\pi}\left|\sin\frac{\pi\tau}{M}\right| + \left(1 - \frac{|\tau|}{M}\right)\cos\frac{\pi\tau}{M}, \quad |\tau| \leqslant M. \tag{20}$$

The frequency domain counterpart of Eq. (20) is

$$W_o^M(f) = \frac{4M}{\pi^2}\frac{1+\cos(2M\pi f)}{(4M^2 f^2 - 1)^2}. \tag{21}$$

The estimated width value is $M_o = 4$. So, using Eq. (20) the window is $w_o^4(\tau)$ and it is depicted in Fig. 2. Similarly, using Eq. (21) the window is $W_o^4(f)$ and it is depicted in Fig. 3.

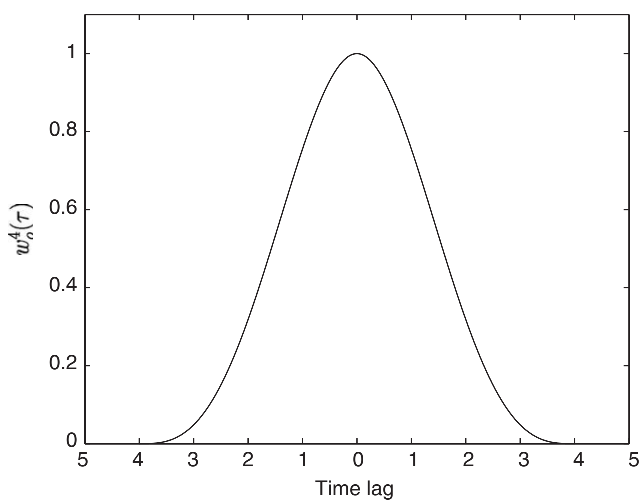

Fig. 2. Plot of the smoothing window in time domain described in Eq. (20) with $M = 4$.

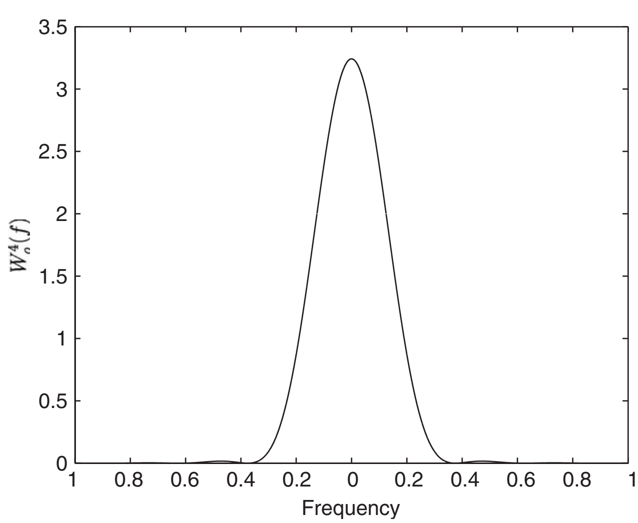

Fig. 3. Plot of the smoothing window in frequency domain described in Eq. (21) with $M = 4$.

6.2. FD algorithm

6.2.1. FD algorithm initial setup (definition of the alarm thresholds)

As pointed out in Section 5.2, the alarm thresholds $\mathscr{A}^-(f)$ and $\mathscr{A}^+(f)$ depend on the significance level α used in the computation of the confidence interval functions. The algorithm has been tested using increasing values for α, specifically $\alpha = 90\%, 95\%, 99\%$. It is expected that, as α increases, the false alarm rate should decrease. In fact, increasing α will introduce more conservativeness in the algorithm: as a consequence, the time instant t_{dt} at which the fault is detected will be larger.

In Fig. 4, the alarm thresholds $\mathscr{A}^-(f)$ and $\mathscr{A}^+(f)$ computed according to the proposed algorithm using the medium flow F and the noisy rod displacement X' are shown for different values of α.

6.2.2. FD algorithm testing

An example of how the FD algorithm works during testing in the *Benchmark zone* is shown in Figs. 5, 6. This is a typical example of how a fault is detected. Specifically, the detection of fault f_{17} has been considered. In fault-free conditions, at each time instant, a SCF estimate is produced which is completely inside the alarm thresholds (see Fig. 5). An alarm is raised as soon as the SCF current estimate crosses, in any point, one of the alarm thresholds. This situation is depicted in Fig. 6.

As a second example, a plot concerning detection of fault f_2 is shown in Fig. 7. This picture represents the alarm thresholds for $\alpha = 90\%$ and the SCF estimate at time $t = t_{from} + t_{dt}$ (which is $t = 1593$ s, being $t_{from} = 900$ s and $t_{dt} = 693$ s in this case (see Table 2)).

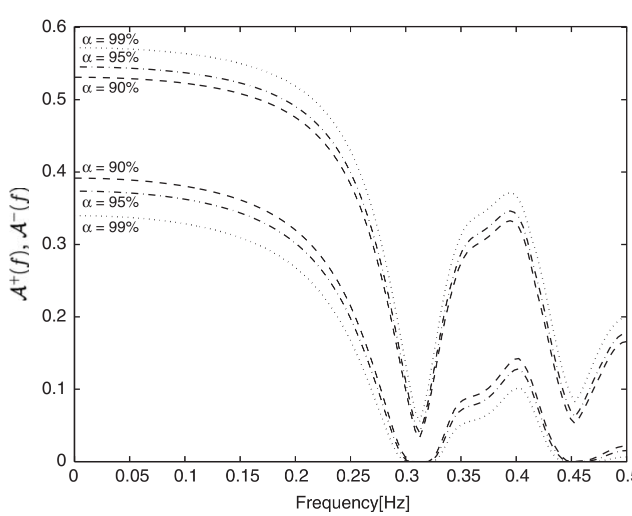

Fig. 4. Plot of the alarm thresholds with $\alpha = 90\%$ (dashed), $\alpha = 95\%$ (dash-dotted), $\alpha = 99\%$ (dotted) for fault detection using SCF of X' and F.

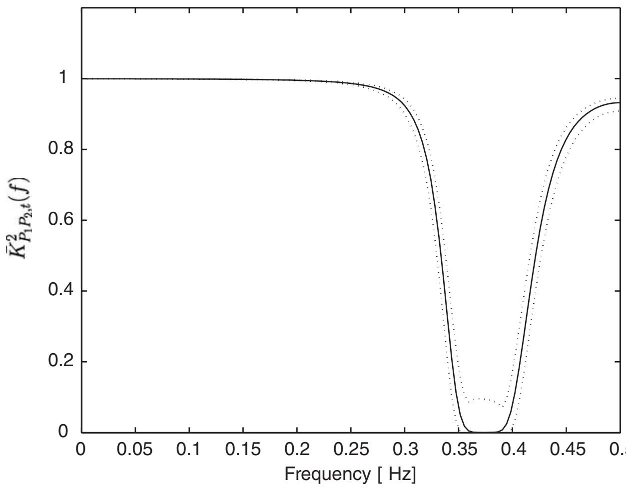

Fig. 5. Detection of the fault f_{17} with $\alpha = 95\%$. The SCF estimate $\bar{K}^2_{P_1P_2,t}(f)$ is completely inside the alarm thresholds (dotted line): so, no alarm is raised.

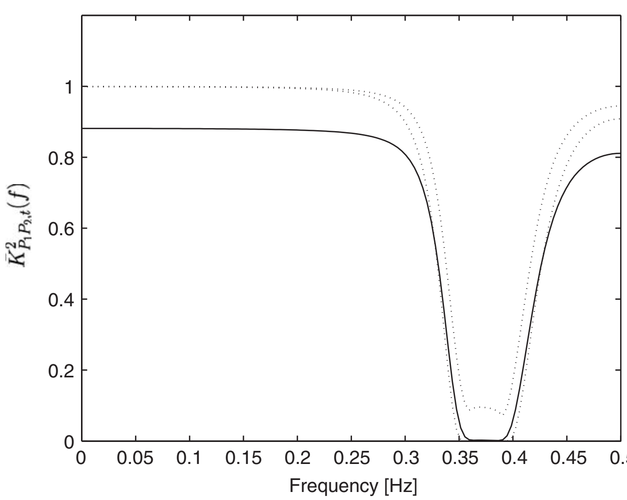

Fig. 6. Detection of the fault f_{17} with $\alpha = 95\%$. The SCF estimate $\bar{K}^2_{P_1P_2,t}(f)$ has crossed the lower alarm threshold (dotted line): so, an alarm is raised.

Similar plots can be provided for all the fault scenarios considered in this work and they are not reported for the sake of brevity.

To evaluate the performance of the FD algorithm according to the rationale depicted in Section 4.2 the Detection Decision signal $D(t)$ is evaluated for $t_{on} < t < t_{hor}$, i.e. in the *Benchmark zone*. Remember that $D(t) = 1$ for $t_{on} < t < t_{from}$ means that the algorithm is raising a *false alarm*. Similarly, if $D(t) = 0$ for $t_{from} < t < t_{hor}$ this means that the FD algorithm is *missing an alarm*. As an example, in Fig. 8, $D(t)$ is shown in the case of the fault scenario f_{15} for three values of α.

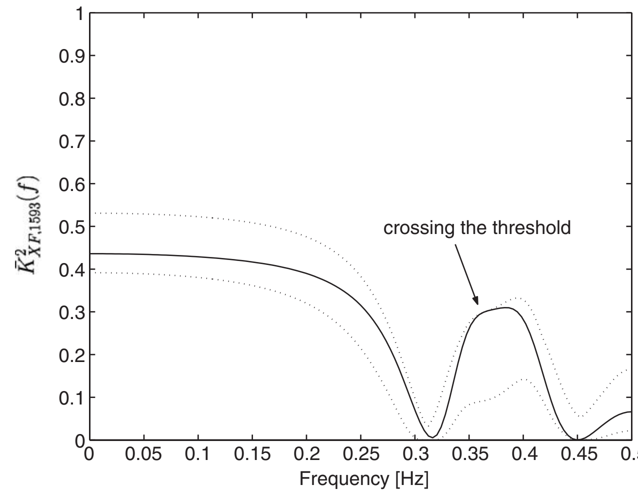

Fig. 7. Detection of the fault f_2 with $\alpha = 90\%$. The SCF estimate $\bar{K}^2_{XF,t}(f)$ at the time $t = 1593$ s is represented. Notice that the SCF estimate crosses the upper alarm threshold (dotted line): so, an alarm is raised.

Table 2
Results for $\alpha = 99\%$

	Variables used	t_{dt} (s)	r_{fd}	r_{td}
Fault 1	X, F	405	0	0.55
Fault 2	X, F	693	0	0.23
Fault 7	X, F	324	0	0.60
Fault 8	X, F	261	0	0.71
Fault 10	X, F	243	0	0.73
Fault 11	X, F	144	0	0.84
Fault 12	X, F	657	0	0.27
Fault 14	X, F	261	0	0.71
Fault 15	X, F	135	0	0.85
Fault 17	P_1, P_2	15	0	0.98
Fault 18	X, F	162	0	0.82
Fault 19	X, F	252	0	0.72

It is clear from Fig. 8 that α is a crucial tuning parameter since it regulates how tight are the alarm thresholds and, consequently, how fast is the algorithm in detecting a fault after its occurrence and if false alarms are raised or not.

A complete report of the results is provided by Tables 2–4.

From Tables 2–4 a general property of the proposed algorithm can be deduced: the detection time t_{dt} is longer the larger is the value of α. Accordingly, the rate of false detection r_{fd} is lower the larger is the value of α. This means that, for large α values the FD algorithm needs more data to detect a fault, i.e. the algorithm is slower in detection, but virtually no false alarms are experienced. On the contrary, by using small α values, false alarm could be evidenced, but the FD algorithm can promptly detect the occurrence of a fault. Notice

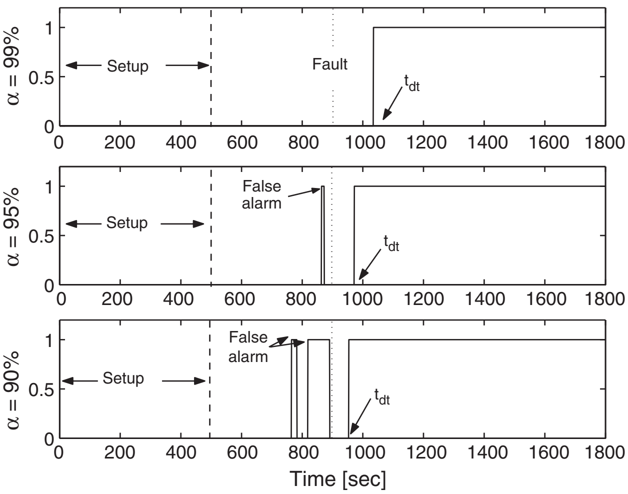

Fig. 8. Detection decision signal for different values of α in fault scenario f_{15}.

Table 3
Results for $\alpha = 95\%$

	Variables used	t_{dt} (s)	r_{fd}	r_{td}
Fault 1	X, F	153	0.02	0.83
Fault 2	X, F	171	0.02	0.81
Fault 7	X, F	153	0.02	0.83
Fault 8	X, F	153	0.02	0.83
Fault 10	X, F	234	0.02	0.74
Fault 11	X, F	10	0.02	0.99
Fault 12	X, F	306	0.02	0.66
Fault 14	X, F	153	0.02	0.81
Fault 15	X, F	72	0.02	0.92
Fault 17	P_1, P_2	9	0	0.99
Fault 18	X, F	63	0.02	0.93
Fault 19	X, F	63	0.02	0.78

Table 4
Results for $\alpha = 90\%$

	Variables used	t_{dt} (s)	r_{fd}	r_{td}
Fault 1	X, F	72	0.22	0.92
Fault 2	X, F	63	0.22	0.93
Fault 7	X, F	72	0.22	0.92
Fault 8	X, F	18	0.22	0.89
Fault 10	X, F	135	0.22	0.79
Fault 11	X, F	9	0.22	0.99
Fault 12	X, F	81	0.22	0.85
Fault 14	X, F	18	0.22	0.89
Fault 15	X, F	54	0.22	0.94
Fault 17	P_1, P_2	9	0	0.99
Fault 18	X, F	36	0.22	0.96
Fault 19	X, F	54	0.22	0.94

that the highest value obtained for r_{fd} is 0.22 that can be interpreted as a 22% of false alarms in the worst case here considered. On the contrary the best value for r_{fd} is 0 that means that no false alarms at all are raised at any time when the fault detection algorithm is on.

Similar considerations can be applied to the rate of true detection r_{td}. Its value is higher for small α values and vice versa. Its peak value is 0.99: this means that, in the best case, the fault occurrence is detected almost immediately and the alarm is kept on for all the time the fault is acting on the actuator. On the other side the smallest value is 0.23. Notice that this value is not due to wrong detection, but only to a delayed detection (fault scenario f_2 when detection is done at $t = 693$ s).

7. Application to experimental process measurement with artificially generated faults

In this section, results from the application of the proposed FD method to real process data are described. Specifically, the problem of detection of fault f_{17} (unexpected pressure drop across the valve) on Actuator 1 (thin juice inflow control) has been addressed.

The artificial fault has been generated on the plant on November 20, 2001 at time $t = 37,780$ s from midnight and it lasts until time $t = 38,400$ s. The fault is abrupt as evident from the sudden drop in the valve outlet pressure P_2 (see Fig. 10). The alarm thresholds have been calculated using fault-free data on November 18, 2001 and November 19, 2001 with $\alpha = 99\%$. The SCF has been computed at each time t using the last 1024 available data. In Fig. 9 the SCF estimate together with the alarm thresholds are shown, immediately after the detection of the fault. The fault is detected at $t = 37,801$ s. So, the detection time is $t_{dt} = 21$ s. The figure shows that the occurrence of the fault significantly modifies the SCF. In Fig. 10 the outlet valve pressure P_2 and the Detection Decision signal $D(t)$ are shown. No false alarms have been registered on the day November

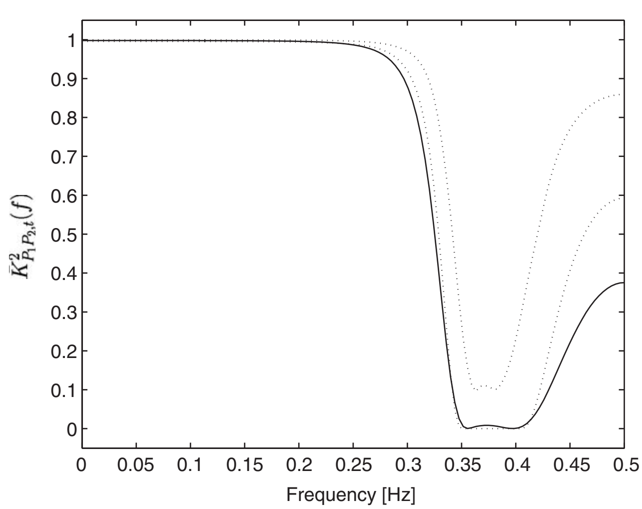

Fig. 9. Detection of the fault f_{17} with $\alpha = 99\%$. SCF estimate $\bar{K}^2_{P_1P_2,t}(f)$ for $t = 37,801$ s together with the alarm thresholds (crossed).

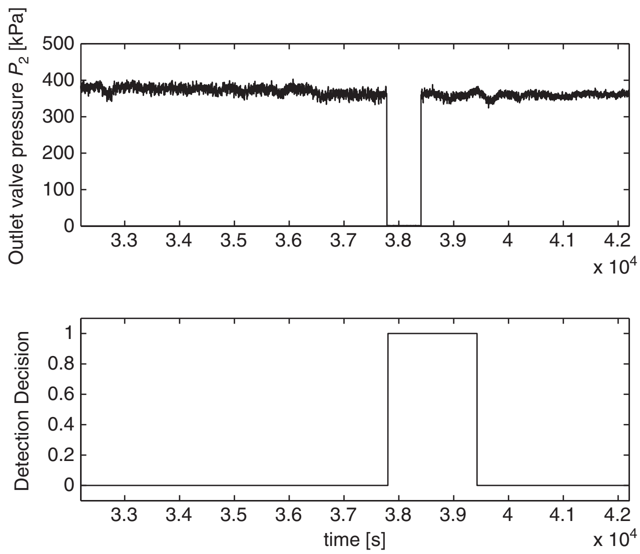

Fig. 10. Upper: Outlet valve pressure in Actuator 1. Lower: Detection Decision signal for the detection of fault f_{17} with $\alpha = 99\%$.

20, 2001 from midnight to the time of the fault. From Fig. 10 it is also possible to see how long the algorithm takes to switch off the alarm, after that the fault is no more affecting the plant. In fact, the alarm is switched on (such as to say, the SCF current estimate is out of the alarm thresholds) up to the time $t = 39,450\,\mathrm{s}$, i.e. the alarm is switched off after $\Delta t = 1050\,\mathrm{s}$. This time interval is in agreement with the fact that the real-time computation of the SCF estimate is done by using the last 1024 available measurements.

8. Conclusions

In this work, faults whose primary nature is abrupt have been analyzed for the Damadics benchmark problem. In this case, the proposed algorithm provides effective fault detection. In particular, for any value of the design parameters, the algorithm presents a high true detection rate. In general, it has been observed that the detection of a fault is faster when low values of the tuning parameter α are used. On the contrary, the lower the value for α, the higher the rate of false alarms.

It is worth pointing out that, in the simulated case, the signal used to excite the system, i.e. the control variable CV, is a sinusoidal wave. It must put in evidence that such a signal is not the best choice for spectral estimation based methods. However, this fact has not seriously limited the power of the proposed algorithm, at least in the discussed application.

Finally, results are presented for the detection of the fault f_{17} on Actuator 1 using real process data. The algorithm is successful in a fast (21 s) fault diagnosis.

Acknowledgements

The authors acknowledge funding support under the EC RTN contract (RTN-1999-00392) DAMADICS. Thanks are expressed to the management and staff of the Lublin sugar factory, Cukrownia Lublin SA, Poland for their collaboration and provision of manpower and access to their sugar plant.

References

Bartys, M., Patton, R., Syfert, M., de las Heras, S., Quevedo, J. (2005). Introduction to the DAMADICS actuator FDI benchmark study. *Control Engineering Practice*, this issue.

Bartys, M., Syfert, M. (2002). Benchmark definition file. http://diag.mchtr.pw.edu.pl/damadics

Chen, J., & Patton, R. J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Boston, Dordrecht, London: Kluwer Academic Publishers.

Frank, P. M. (1990). Fault diagnosis in dynamic system using analytical and knowledge-based redundancy. A survey and some new results. *Automatica*, *26*, 459–474.

Gertler, J. J. (1988). Survey of model-based failure detection and isolation in complex plants. *IEEE Control System Magazine*, *8*, 3–11.

Gertler, J. J. (1998). *Fault detection and diagnosis in engineering systems*. New York: Marcel Dekker.

Guglielmi, G., Parisini, T., & Rossi, G. (1995). Fault diagnosis and neural networks: a power plant application, Keynote Paper. *Control Engineering Practice*, *3*, 601–620.

Haber, R., 1985. Nonlinearity tests for dynamic processes. *Proceedings of the IFAC identification and system parameter estimation*, York, UK, pp. 409–413.

Isermann, R. (1984). Process fault detection based on modeling and estimation methods. A survey. *Automatica*, *20*, 387–404.

Jenkins, G. M., & Watts, D. G. (1968). *Spectral analysis and its applications*. San Francisco: Holden-Day.

Ljung, L. (1999). *System identification*. Englewood Cliffs, NJ: Prentice-Hall.

Lohnberg, P. (1978). Improved approximation of bias in squared coherency estimates for weakly smooth spectra. *IEEE Transactions on Acoustic, Speech, and Signal Processing*, *ASSP-26*, 172–174.

Papoulis, A. (1973). Minimum-bias windows for high-resolution spectral estimates. *IEEE Transactions on Information Theory*, *IT-19*(1), 9–12.

Patton, R. J., Frank, P. M., & Clark, R. N. (Eds.). (1989). *Fault diagnosis in dynamic systems theory and applications*. Englewood Cliffs, NJ: Prentice-Hall.

Previdi, F., Lovera, M., & Mambretti, S. (1999). Identification of the rainfall-runoff relationship in urban drainage networks. *Control Engineering Practice*, *7*(12), 1489–1504.

Previdi, F., & Parisini, T. (2001). Model-free fault detection: a spectral estimation approach based on coherency functions. *International Journal of Control*, *74*(11), 1107–1117.

Stoica, P., & Moses, R. (1997). *Introduction to spectral analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Vemuri, A. T., Polycarpou, M. M., 1996. On-line approximation based methods for robust fault detection. *Proceedings of the 13th IFAC World congress* (pp. 319–324).

Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, *12*, 601–611.

Young, P., & Patton, R. J. (1990). Comparison of test signals for aircraft frequency domain identification. *Journal of Guidance, Control and Dynamics*, *13*(3), 430–438.