DESIGNING MODEL-BASED FAULT ESTIMATOR FOR A SEPARATION COLUMN

Dimitry Gorinevsky[1] and Dinkar Mylaraswamy[2]
Honeywell Laboratories

Abstract. This paper considers a practical approach to model-based fault diagnostics. The fault estimation is performed by a delayed FIR filter that is designed as a noncausal Wiener filter. The approach is demonstrated using a case study of a process upset in a separation column in a petrochemical plant. The designed fault estimator demonstrates early and reliable detection of the fault. It presents a conclusive evidence of the fault more than 250 minutes earlier than the action the operator took during the upset based on direct sensory data.

1 Introduction

This paper discusses fault diagnostics in the context of a case study centered around a real life incident in a petrochemical plant. An ethylene manufacturing unit is a typical example of tightly coupled separation columns, optimized to minimize energy consumption and increase throughput. In each separation column, it is very common to find a 2–3 layered control strategy that maximizes heat utilization; each layer receives measurements from several sensors upstream and downstream to find an optimum setting for the cascading controller. With their maze of cascading controls such processes can often be very vulnerable to single point faults – a minor sensor failure can lead to disastrous domino effects in apparently unrelated process parts.

The problem started out as a simple sensor failure that was part of a three layered control strategy. The feedback action masked this problem from the operator for several hours before it manifested itself. After about 6 hours, the operator responded to the first alarm, without realizing the root cause, and aggravated the situation. Within 15 minutes he realized the gravity of the situation and cut the feed. The unit operated in this low feed mode for several hours (translates into several thousands of dollars loss) before the problem was identified and fixed. It is believed that if the operator was aware of the faulty sensor, he could have intervened immediately by breaking the cascade control loop and containing the problem to a manageably small section of the process. This builds up a case for an automated monitoring of the critical process variables and relating their behavior to one or more fault factors.

The problem of fault diagnosis has been discussed in the technical literature on control as well as in industrial practice over the past several years. The two distinct schools of thought in addressing this problem are: model-based [18, 10, 6, 7] and data-driven [2, 4, 9]. A comparative discussion of these approaches for a closely related industrial application can be found in [14] and [13]. In 1995, the Abnormal Situation Management (ASM^{TM}) Consortium was created by biggest petrochemical process companies and controls vendors, such as Honeywell, to address fault diagnostics and recovery in the process control environment. Most of the early ASM activity was concentrated on data-driven diagnostic methods because of their relative ease of development and deployment. A recently completed ASM research concluded that model-based diagnostics is necessary for providing practical sustainable decision support systems with low false alarm rates [12].

The main issue with practical application of model-based diagnostic methods is the heavy cost of developing and maintaining good first principles models for nominal and faulty behaviors. This issue has always been a barrier in the industrial acceptance of model-based diagnostic methods. While the issue of modeling effort cannot be avoided altogether, the approach presented in this paper lends itself to easy to application.

Industrial plant models usually have to be split into interconnected simpler subsystems. They might include complicated nonlinear heat and mass transfer equations that defy analytic solution and are solved numerically. Often, the modeling is done with the help of specialized modeling and simulation software packages. At the same time, many of the fault diagnostics methods discussed in the literature require the use of detailed analytical models, which makes them difficult to implement and deploy with practical simulation models. This paper demonstrates a model-based fault diagnostics methodology that is well suited to work with simulation models and uses easily available data only for setup.

In this paper, as in many other model-based fault diagnostics studies, the fault is modeled as an unknown input to a plant. The problem is to estimate the unknown fault input from the noisy data which is given by the residual error of the model prediction. Such a problem of estimating an unknown input (also called fault intensity) from noisy data was discussed in the literature. In the model-based fault detection and identification (FDI) literature, the approaches such as robust residual filters [3, 16, 20] are most relevant to this study. However, these approaches are not optimal and computationally difficult. The same problem of estimating an unknown input of a dynamic system from the noisy output data was rigorously studied in the signal

[1]Honeywell CA35-5272H, One Results Way, Cupertino, CA; email: Dimitry.Gorinevsky@honeywell.com
[2]Honeywell MN65, 3660 Technology Dr., Minneapolis, MN; email: Dinkar.Mylaraswamy@honeywell.com

processing literature where it is known as the *deconvolution* problem. The deconvolution problem has been discussed in a number of papers with respect to different applications including the problem of fault estimation [1, 8, 21]. All the above mentioned papers propose different formulations and solutions of optimal (or suboptimal) filtering problem, assuming that the unknown input needs to be estimated from the data using a given fixed delay. In the FDI papers, such as [3, 16, 20], the delay is ignored (considered to be zero). The main drawbacks of these solutions and formulations is that they (i) are conceptually complicated multi-step mathematical solutions that are difficult to implement in practical applications, (ii) provide only a limited accuracy of the unknown input estimation because all of the output information beyond a fixed estimation delay is lost. Further, the choice of the delay parameter, which is a very important implementation detail, is not discussed. Yet, in practice the output response to the unknown (fault) input is often delayed and choosing an appropriate estimation delay is critical for the estimation accuracy.

This paper uses a much simpler optimal Wiener filter solution to the deconvolution problem. The Wiener filter design assumes that all future output information is available for estimation of the unknown input at any point in time. In this paper the designed Wiener filter is truncated and implemented in the form of a fixed delay FIR filter with practically the same accuracy. This automates the choice of the estimation delay. The solution discussed in this paper is conceptually simpler and much easier to implement in practical applications compared to the above referenced papers. The delayed FIR filter used herein is related to Kernel regression methods, e.g. see [5] and references thereof. Unlike the filter design presented in this paper, Kernel regression is usually applied as an empirical approach.

2 Case Study

The layout of a typical separation column is shown in Figure 1. The objective of such a column is to separate *light* components (overhead product) from the *heavy* components (bottoms product). This is done by supplying vapors at the bottoms through the reboiler and cold liquid (reflux) at the top through the condenser. For additional details on distillation columns see reference [11, 19]. The following

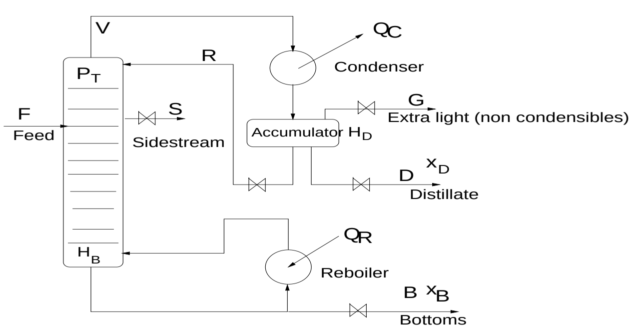

Figure 1: Layout of a typical separation column

key variables characterize the performance of the separation process: x_B – the composition of the heavy key in the bottoms product expressed as a fraction; F_h – the flow of the hot vapors to the reboiler, often F_h is translated to the actual heat supplied Qrs using standard heat transfer correlations; Qrd – a target for the amount of heat that needs to be supplied at the reboiler; Pov – the tower pressure; Rov – the reflux flow. These variables are available in the process control system either directly from respective sensors, or as values calculated by the control system based on the sensor data and standard correlations. The trends and instantaneous values of these variables are displayed by the control system user interface and observed by an operator.

As mentioned in the Introduction, this paper is centered around a relatively significant upset that happened in a petrochemical plant. The traces of some key variables (normalized values) during this upset are shown in Figure 2. During this upset, a temperature sensor T_{up} used in the calculation of the reboiler heat demand setpoint Qrd developed a bias around $t = 450$ (Figure 2, upper left). This almost froze the calculated value of the reboiler heat demand Qrd (Figure 2, upper right). As a result, flow of the heating medium (F_h, Figure 2, lower right) did not respond to a deteriorating bottom composition (x_B, Figure 2, lower left). The operator detected the loss of bottoms composition around sample 850 and since it was too late, the feed to the unit was cut as a preventive measure. The unit operated in this reduced feed mode for several hours before the heat balance was fully restored. For this unit, this translates to several thousand dollars of production loss. This is a classical event in which the control system masks an incipient fault. According to an experienced operator: "The controller setpoint, Qrd and the controller output, F_h are consistent. The setpoint is not changing, and therefore the F_h is not changing and since x_B is within normal range, nothing is wrong with the unit". This is one of the primary reasons why this event went undetected till things started to go really bad and resulted in loss of production. It is important to get an early warning to an operator in incidents like this one. This paper describes a design of fault diagnostics system that would have, with a high confidence given an early indication of the problem. For this particular incident, if an early warning was provided to the operator around $t = 600$, that clearly indicated the error in the calculation of Qrd, the operator would have responded correctly by breaking the cascade and containing the problem. It is believed that such a response would have avoided the expensive correction the operator was forced to take.

A general idea of model-based fault diagnostic systems, such as one considered herein, is to compare model-based prediction of the plant output variables to the experimentally observed values. The residual error of such prediction is then used to evaluate the fault condition. This is possible by using a model of how the fault influences the plant outputs. The design of the model-based diagnostics scheme in this paper goes through the following sequence of steps: **(1)** Develop a model for nominal process operation; **(2)** Develop a model for the fault; **(3)** Design a fault estimation scheme that allows evaluating fault intensity fromthe residual error of the model-based prediction.These steps are discussed in the next two sections of

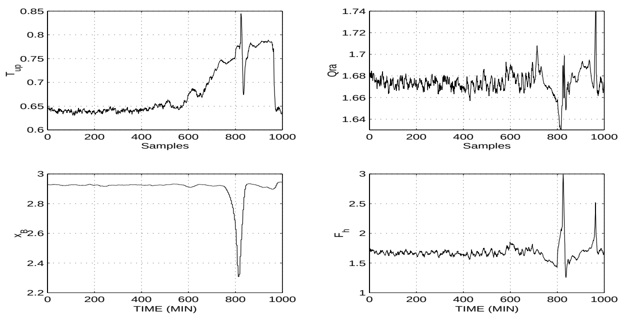

Figure 2: Variable traces in the incident

the paper. The last section presents an application of the designed diagnostics system to the incident data of the case study. There is potentially an additional step that involves decision-making regarding when to warn an operator, based on statistical properties of the fault estimate. This last step is not discussed in any detail in this paper and the decision is nade based on simple thresholding of the fault estimate.

3 Process and Fault Models

A model was developed that simulates the separation column under constant separation factors and floating pressure conditions. Detailed thermodynamic models were developed for the condensation and vaporization in the condenser and the reboiler. The distillation column itself is modeled as a ideal binary separation column with pressure dependent relative volatility $\alpha(P_T)$. The feed F is modeled as a measured disturbance, while the composition is an unmeasured disturbance. For additional details, see the report by Mylaraswamy [15]. The models outlined above can be summarized as a set of differential algebraic equations of the form

$$\begin{aligned} \dot{\hat{x}} &= \phi\left(x(t),\, u(t),\, d(t),\, z(t)\right), && (1)\\ 0 &= \psi\left(x(t),\, u(t),\, d(t),\, z(t)\right), && (2)\\ \hat{y}(t) &= g(x(t)), && (3) \end{aligned}$$

where $\hat{x}(t)$ is a state vector, $\hat{y}(t)$ is the output vector, $u(t)$ is the control input vector, $d(t)$ is the external disturbance vector and $z(t)$ is an auxiliary variable vector. These vectors are described below. The 12-dimensional vector of auxiliary variables $z(t)$ is defined as an implicit function of $\hat{x}(t)$, $u(t)$, and $d(t)$ by a 12-dimensional vector of algebraic equations (2). As it is usually the case, the developed nominal plant model is available in the form of simulation software rather than in an analytic state space form (5).

The incident discussed in this paper corresponds to the failure of the sensor that measures T_{up} used in the calculation of the reboiler heat demand calculation, Qrd. The fault was modeled as:

$$T_{up}^{fault} = T_{up} + b_{T_{up}}, \qquad (4)$$

where $b_{T_{up}}$ is the bias in the measurement of the temperature. Equation (4) describes a modification of the model (1)–(3) that is needed to reflect the fault influence. By replacing T_{up} in (1)–(2) with T_{up}^{fault} (4), then solving (2) for $z(t)$, and finally substituting $z(t)$ expressed through $\hat{x}(t)$, $u(t)$, and $d(t)$ into (1), the simulation model of the plant can be presented in the form

$$\dot{\hat{x}} = \varphi(\hat{x}, v, f), \qquad \hat{y} = g(\hat{x}), \qquad (5)$$

where $\hat{x} \in \Re^6$ is the plant state vector, $v = \left[u^T \; d^T\right]^T \in \Re^7$ is the vector of measured external inputs to the plant including control inputs and disturbance variables, and $f = [\, b_{T_{up}} \,] \in \Re$ is the fault variable.

In the absence of the fault (for $f = 0$), the nonlinear function $\varphi(\cdot, \cdot, 0)$ is an implicit function defined by (1) and (2). Vector $\hat{y} \in \Re^6$ is the vector of the plant output variables. The plant inputs collected in the vector v are provided by a number of different cascaded controllers and include operator settings and measured upstream disturbances. These inputs are available in addition to the output variables for diagnostics purposes, but cannot be directly influenced by the fault diagnostics software.

The fault modeling and diagnostics analysis will further assume that (i) the plant is operating close to a steady state regime and (ii) the plant operation change caused by the fault is small. These two assumptions hold in our case study and allow linearizing the plant model around a nominal regime. These two assumptions hold in most industrial process plant fault diagnostics problems: (i) holds because the steady state regime is an optimized regime of the production and large deviations from this regime are avoided in plant operation; (ii) the focus of the fault diagnostics is on early detection of fault conditions before they cause large changes in the plant operation.

Under normal conditions, because of the approximations and the uncertainty in the model, the departure of the measurement variables, such as Qrd and Qrs, from a nominal value will not be identically zero. The challenge is to be able to identify significant departures caused by the fault as quickly as possible without generating too many false alarms. model can be assumed to describe this deviation.

$$\dot{x} = \varphi(x, v; f) + \xi, \qquad y = g(x) + \eta, \qquad (6)$$

where ξ and η are external disturbances (process variation) sequences introduced to account for the deviation between the model prediction and real process measurements. Statistical properties of these disturbance sequences are discussed further on. As mentioned above, the parameter vector f in (6) is normally zero. The value of f gives a measure of the fault condition intensity. By subtracting the nominal model (5) with $f = 0$ from (6), and linearizing the system by retaining only first order difference terms, the following fault residual model can be obtained

$$\dot{\tilde{x}} = A\tilde{x} + Bf + \xi, \qquad \tilde{y} = C\tilde{x} + \eta, \qquad (7)$$

where

$$A = \frac{\partial \varphi}{\partial x}(\hat{x}, v, 0), \quad B = \frac{\partial \varphi}{\partial f}(\hat{x}, v, 0), \quad C = \frac{\partial g}{\partial x}(\hat{x}),$$

The model (7) is in continuous-time. At the same time data collected from the process and used in the signal processing algorithms below is sampled at a certain rate—the sampling rate. In what follows, a sampled-time form of the fault residual model will be used

$$\tilde{y}(t) = F(z^{-1})f(t) + D(z^{-1})e(t), \qquad (8)$$

where $\tilde{y} = y - \hat{y}$ is the model prediction residual, z^{-1} is a unit delay operator and $e(t)$ is a unit white noise sequence. The signature filter $F(z^{-1})$ is obtained by sampling the continuous-time transfer function $C(Is - A)^{-1}B$, and the shaping filter $D(z^{-1})$ defines the disturbance dynamics. The disturbances ξ and η in (7) were not defined and they are assumed such that (8) holds.

The fault residual model (8) will be further used for design of fault detection algorithms. Variables and parameters in this model can be conveniently characterized from the available information on the nominal model and the process data. In particular, $\tilde{y}$ can be obtained by running the simulation model and subtracting the model prediction from the actually observed data. The shaping filter $D(z^{-1})$ can be identified from the plant normal operation data, where $f(t) \equiv 0$. Finally, the fault signature filter $F(z^{-1})$ can be obtained by linearizing the nominal simulation model (5). In this work it was conveniently done by running the simulation twice to obtain the response of the predicted process output with respect to a small step change in the fault intensity f. Responses of the output variables $\tilde{y}$ to a step input in the sensor fault considered in this case study are shown in Figure 3. These responses define the fault signature filter $F(z^{-1})$ in (8).

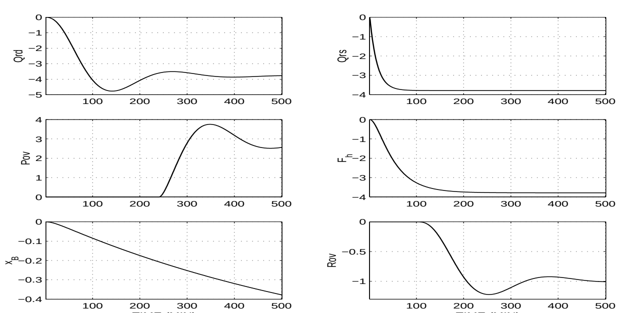

Figure 3: Step responses of the model outputs to a two-unit step change in the fault variable $f = b_{T_{up}}$

The fault estimation problem to be considered in the remainder of the paper is as follows. Consider the system (8), where the signal $\tilde{y}$ is available, $e(t)$ is a white noise sequence, filters $F(z^{-1})$ and $D(z^{-1})$ are known, and $f(t)$ is an unknown sequence. The problem is to estimate the fault sequence $f(t)$. This estimate can be further used for decision-making, e.g., by comparing the estimated fault intensity with a threshold.

Estimating fault intensity $f(t)$ from the model prediction error $\tilde{y}$ in (8) requires specifying what an *a priori* information on the estimated signal $f(t)$ is available and what part of the signal $\tilde{y}$ is accessible at time t.

In this work the estimated fault time sequence is modeled as an output of a first order coloring filter driven by a white noise independent of other signals in the system.

$$f(t) = \Phi(z)\zeta(t), \qquad \Phi(z) = \frac{f_0}{1 - a_f z^{-1}}, \qquad (9)$$

where $\Phi(z)$ is the first-order shaping filter, the parameter f_0 has a physical meaning of the probable amplitude of the fault and the parameter a_f has a meaning of the time constant for the fault development. The two parameters f_0 and a_f are further viewed as engineering tuning knobs in the fault estimator design.

4 Wiener Filtering Approach

The problem of estimating the unknown input sequence $f(t)$ in the system (8) from the output sequence $\tilde{y}(t)$ is known in the literature as a *deconvolution* problem. When solving the problem, a fixed delay d of the estimation is typically assumed, such that the output data up to time $t+d$ are used in estimating $f(t)$. Solutions of the deconvolution problem differing in complexity of computations and formulation are presented in the papers [1, 8, 21]. All these solutions are conceptually complicated and require deriving and solving several simultaneous Diophantine equations. To the best of the authors knowledge these solutions are not implemented in standard software packages, such as the Matlab Signal Processing Toolbox.

This paper pursues a non-causal Wiener filter solution to the deconvolution problem that does not specify the estimation delay explicitly and assumes that all future information is available. It turns out that the filters obtained as a result of such design use future information with rapidly decaying weight and for all practical purposes can be truncated to have a finite delay. In this case study, as well as in many other applications, the delay in the fault estimation resulting from such a design is practically acceptable. A big advantage of the filter design in this paper is the simplicity of both the concept and the computational implementation.

4.1 Filtering problem

Different authors might mean different by a Wiener filter name. Some authors consider a *causal* filter obtained as a result of Wiener-Hopf factorization (projection) of a noncausal least-square optimization equations. In other literature, as well as in this paper, Wiener filter is a *noncausal* least-square optimal filter.

Consider the following problem of optimal smoothing

$$\hat{f} = Ly, \qquad L = \arg\min E(\|\hat{f} - f\|^2), \qquad (10)$$

where $\hat{f}$ is an estimate of the unknown input f, L is a sought linear operator, and E denotes mathematical expectation. The optimal operator L can be found by differentiating the expectation of the quadratic error expressed as a contour integral of an appropriate transfer function

$$L(z) = V(z)F^*(z^{-1})(W(z) + F(z^{-1})V(z)F^*(z^{-1}))^{-1}, \quad (11)$$

where, $V = \Phi\Phi^*$, $W = DD^*$, and L, F, Φ, D are transfer function matrices corresponding to the respective operators and the superscript $*$ denotes the complex conjugate transpose. The argument z has been omitted for these transfer functions. The unity matrix I in (11) has size corresponding to the dimension of f. In our case study this is a 1×1 matrix and $I = 1$. It is assumed that none of the transfer functions L, F, Φ, and D have poles on the unit circle.

As mentioned above, the operator L is a noncausal operator. In accordance with this, the transfer function (11) has poles both inside and outside of the unit circle. The transfer function $L(z)$ can be considered as a 2-sided z-transform of a corresponding non-causal pulse response kernel $l(t)$. A background on 2-sided z-transform theory can be found in the textbook [17]. The operator L can be implemented as a convolution with the kernel $l(t)$.

$$(Ly)(t) = \sum_{\tau=-\infty}^{\infty} l(\tau)y(t-\tau)d\tau \qquad (12)$$

4.2 FIR filter

Since $L(z)$ in (11) has no poles on the unit circle, the kernel $l(t)$ in (12) decays as $t \to \infty$ and $t \to -\infty$. Let $r > 1$ be the largest number such that $L(z)$ is analytical inside the ring $r^{-1} < |z| < r$. Then, in accordance with the properties of the 2-sided z-transform, the kernel $l(t)$ asymptotically decays at least as fast as $r^{-|t|}$. For any practical purpose the infinite exponentially decaying 'tails' of the kernel can be truncated and the kernel implemented as a delayed FIR operator. The parameter r defines the degree of the kernel localization and, thus, the delay of the unknown input estimator L.

Let the truncated FIR kernel be $l(t)$ defined for $t_{\min} \leq t \leq t_{\max}$ and zero outside this interval. Then, the operator (12) can be presented as a delayed convolution operator with the delay of $-t_{\min}$ in the form

$$(Ly)(t + t_{\min}) = \sum_{\tau=0}^{t_{\max}-t_{\min}} \Lambda(\tau)y(t-\tau)d\tau, \qquad (13)$$

where $\Lambda(t) = l(t - t_{\min})$ is a FIR convolution kernel.

Let us now discuss a computational procedure for calculating the convolution kernel $\Lambda(t)$. This is done by computing an inverse of the Fourier transform for the operator. The transfer function $L(z)$ (11) is analytical on the unit circle (for $|z| = 1$) and can be computed on the unit circle as

$$L(e^{i\omega}) = \Phi_{ff}F^* \left(\Phi_{yy} + F\Phi_{ff}F^*\right)^{-1}, \qquad (14)$$

where $F = F(e^{i\omega})$, and $F^* = F^*(e^{-i\omega})$. In accordance with (11), $\Phi_{ff} = V(e^{i\omega})$ and $\Phi_{yy} = W(e^{i\omega})$ have meaning of the power spectra for the operators Φ in (9) and D in (8) respectively.

The practical considerations in computing transfer function (14) on the unit circle are as follows. First, in (14), Φ_{ff} has a meaning of the spectral power of the fault signal. This spectral power is defined by the two tuning knob parameters f_0 and a_f in the model (9). The values of these parameters can be estimated from the available application-level knowledge of the possible fault development scenario. Second, the Fourier transform $F(e^{i\omega})$ of the fault signature pulse response in (8) can be computed directly from the pulse or step response data without even reverting to a dynamical model for the fault signature operator F. Third, the function Φ_{yy} in (14) has a meaning of the spectral power of the process variation. This spectral power can be estimated directly from the model prediction residual data $\tilde{y}(t)$ collected for a normal operation of the plant in the absence of the faults, i.e., for $f(t) = 0$.

Once the complex-valued function (14) is available, the kernel $l(t)$ can be computed as an inverse Fourier transform. This can be efficiently implemented by applying the DFT (Discrete Fourier Transform) algorithm. To do this, $L(e^{i\omega})$ is computed on a regular spaced grid of the frequencies ω. Such DFT computation is inherently approximate. Its accuracy would be acceptable as long as the number of grid points is sufficiently larger than the support interval for the truncated FIR kernel $l(t)$.

5 Fault Detection: Application Results

A Wiener filter fault estimator designed using the methodology described in the previous section was applied to the case study data shown in Figure 2. As the main step in the estimator design, the Fourier transform (14) of the fault estimator kernel was computed on a grid of 2048 frequency points. This was done in accordance with the steps outlined in the previous section. First, the spectral power $\Phi_{ff}(\omega) = \Phi(e^{i\omega})\Phi^*(e^{i\omega})$ of the fault signal was computed in accordance with (9). The two tuning knob parameters in the model (9) were set to the values $f_0 = 0.015$ and $a_f = 0.985$. Second, the Fourier transform $F(e^{i\omega})$ of the fault signature pulse response in (8) was computed from the step response data in Figure 3. To do this, the sampled-time step responses in Figure 3 were differentiated to obtain the pulse responses. Then $F(e^{i\omega})$ was obtained by computing an FFT of these 2048 point pulse responses. Third, the spectral power of the process variation $\Phi_{yy}(\omega)$ in (14) was estimated from the model prediction residual data collected during normal operation of the plant in the absence of the faults. The standard Matlab Signal Processing Toolbox function EPA was used to obtain an estimate of the spectrum at 256 frequency points. The spectral power estimates between these points were obtained by linear interpolation. The spectral power estimates for Qrd and Qrs are illustrated in Figure 4.

The complex sequence $L(e^{i\omega})$ (14) computed at the 2048 frequency points was used in computing an inverse FFT to obtain an estimate of the fault estimator convolution kernel. This estimate was further truncated by removing the kernel weights 'tails' with values less than $5 \cdot 10^{-4}$ of the maximum value. At first, the optimal fault estimator kernel $l(t)$ (11), (12) was computed as described above

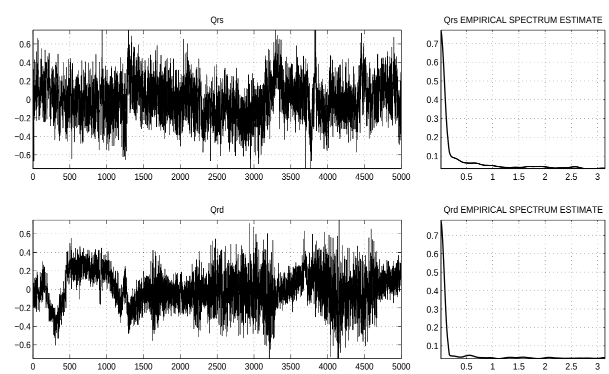

Figure 4: Residuals (left) and spectrum estimates (right) for Qrd and Qrs. Normal plant operation

based on the model (8). Since in (8) the output vector $\tilde{y}$ is a six dimensional vector, $l(t) \in \Re^6$. In other words, the kernel $l(t)$ includes six kernels, each applied to one of the outputs $\tilde{y}$ and the results are added up. The multivariable design of the optimal fault estimator kernel gives significantly different magnitudes for different components of the kernel. This is because the external disturbance has very different intensities (with respect to amplitude and delay) for different outputs, as seen in Figure 3.

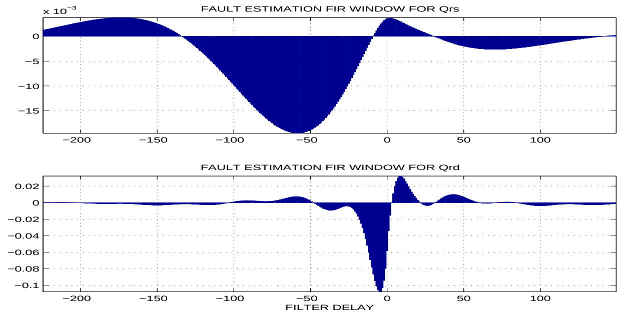

Figure 5: FIR fault estimator kernels for Qrs and Qrd

Therefore, in addition to optimal fault estimator using complete information about the outputs, a suboptimal estimator was designed by using only two of the six available outputs: Qrs and Qrd. The choice of the outputs Qrs and Qrd was dictated by the fact that they provide small delay in the output response to the fault. As a result the fault estimator convolution kernels for these outputs allow obtaining the estimate with small delay. In order to ensure timely detection of the fault, this delay must not be too large. These outputs also provide a good ratio of the fault response amplitude to the process variation (disturbance) intensity. The output F_h in Figure 3 also has small delay, it was, however, discarded because of the modeling errors that can incidentally cause the residual for F_h to be large in the absence of faults.

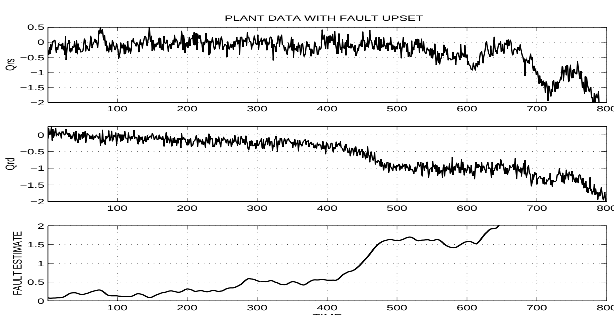

Figure 6: Fault estimation for the case study with designed FIR filter kernels for Qrs and Qrd.

The two components of the suboptimal fault estimator convolution kernel corresponding to the model prediction residuals for Qrs and Qrd are illustrated in Figure 5. Using these kernels requires a delay of about 230 min in the fault estimation. Such delay is reasonable in the case study in question and still allows early detection of the fault. Because of the optimized filtering properties of the designed fault estimator kernels such a fault estimate is highly accurate and allows avoiding false alarms.

The result of applying the designed fault estimator to the process upset data is illustrated in Figure 6. The three plots are the computed residuals Qrs and Qrd and the sensor fault intensity estimate $\hat{f}(t)$. The conclusive evidence of the fault is available when $\hat{f}(t)$ becomes greater than one at $t = 450$. Note that because of the estimation delay, the estimate $\hat{f}(t = 450)$ is available at time $t = 674$. This is still more than 250 minutes earlier than the action taken by the operator during the upset.

References

[1] Ahlen, A., Sternrad M., "Wiener filter design using polynomial equations," *IEEE Trans. on Signal Processing*, vol. 39, no. 11, 1991, pp. 2387–2399

[2] Cinar, A. and Undey, C. "Statistical process and controller performance monitoring. A tutorial on current methods and future directions," *American Control Conf.*, pp. 2625 - 2639, San Diego, CA, June 1999,

[3] Ding, X. and Guo, L. "Time domain optimization of fault detection systems," *IEEE Conf. on Decision and Control*, pp. 616 - 621, December 1996

[4] Dunia, R., and Qin, J.C. "Joint diagnosis of process and sensor faults using principal component analysis," *Control Engineering Practice*, vol. 6, 1998, pp. 457–469.

[5] Fenu, G., Gorinevsky, D., and Parisini, T. "Nonparametric kernel smoothing and FIR filtering for model-free symptom generation," *IEEE Conf. on Decision and Control*, Phoenix, AZ, December 1999

[6] Frank, P.M. "Fault diagnostics in dynamic systems using analytical and knowledge-based redundancy – A survey and some new results," *Automatica*, vol.26, no.3, 1990, pp.459–474.

[7] Gertler, J. "Survey of model-based fault isolation and detection in complex plants," *IEEE Contr. Syst. Magazine*, 1988

[8] Grimble, M.J. "H_2 inferential filtering, prediction, and smoothing with application to rolling mill gauge estimation," *IEEE Tr. on Signal Processing*, vol.42,no.8,1994,pp.2078–2093

[9] MacGregor, J.F., Kourti, T., Nomikos, P. "Analysis monitoring and fault diagnostics of industrial processes using multivariate statistical methods," *1996 IFAC World Congress*, pp. 145-150, San Fracisco, CA

[10] Isermann, R. and Balle, P "Trends in the application of model based fault detection and diagnosis of technical processes," *1996 IFAC World Congress*, San Francisco, CA, July 1996.

[11] McCabe, W.L. and Smith, J.C. *Unit Operations in Chemical Engineering.* McGraw Hill, New York, 1976.

[12] Mylaraswamy, D. and Bullemer, P. "Fielding a multiple state estimator platform," *NPRA Computer Conf.*, Nov 2000.

[13] Mylaraswamy, D. and Venkatasubramanian,V. "A hybrid framework for large scale process fault diagnosis," 30^{th} *European Symposium on Process Systems Engineering*, vol. 21, pp. 935–940, 1997.

[14] Mylaraswamy, D. *DKit: A Blackboard-based, Distributed, Multi-Expert Environment for Abnormal Situation Management.* Ph.D. thesis, School of Chemical Eng., Purdue University, 1996.

[15] Mylaraswamy, D. "Focused state estimators: Design," *Technical Report MS-100-600-130.2*, Honeywell Technology Center, Minneapolis, 2000.

[16] Nikoukhah, R., "Innovation generation in the presence of unknown inputs: Application to robust failure detection," *Automatica*, vol. 30, no. 12, 1994, pp. 1851–1867.

[17] Oppenheim, A.V., Schafer, R.W., and Buck, J.R. *Discrete-Time Signal Processing*, Prentice Hall, 1999

[18] Patton, R., Frank, P., Clark R., *Fault diagnosis in dynamic systems.* Prentice Hall, 1989.

[19] Perry, R.H. and Green, D.W. *Perry's Chemical Engineering Handbook.* McGraw Hill, New York, 1984.

[20] White, J.E. and Speyer, J.L., "Detection filter design: Spectral theory and algorithms," *IEEE Trans. on Automatic Control*, vol. AC-32, no. 7, 1987, pp. 593–603.

[21] Zhang H.-S., Liu, X.-J., Chai, T.-Y., "A new method for optimal deconvolution," *IEEE Trans. on Signal Processing*, vol. 45, no. 10, 1997, pp. 2596–2599