Fault Detection on Pulp Pressure Screens

Leoncio W. Estévez-Reyes

B.Eng. Instituto Universitario Politécnico de Barquisimeto, Venezuela, 1982.

M.Eng. McGill University, Montreal, 1991.

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF ELECTRICAL ENGINEERING

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June 1995

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of ELECTRICAL ENGINEERING

The University of British Columbia
Vancouver, Canada

Date DEC. 22, 1995

Abstract

In this thesis the development of a strategy for the detection and isolation of faults particular to the pressure screens commonly used in mechanical pulp mills is presented. After exploring several non-model and model-based approaches for fault detection and isolation (FDI), a method is chosen. The method relies on the simultaneous identification of the states and process coefficients of a dynamic model for said screens. The criterion used in developing such a model is explained and the concept of "dynamical" friction coefficients as fault detection indicators is introduced. A technique called the Singular Pencil Matrix (SPM) for simultaneous identification of states and process coefficients is discussed and some simulation and industrial results using SPM are shown. As part of the validation of the dynamic model a reduced version of it is proposed and tested. The problem of modelling the process noise when applying SPM is discussed, some options are studied, and an alternative scheme based on cumulative functions for confirmation of screen faults is presented.

Table of Contents

List of Tables

List of Figures

Acknowledgments

A friend of mine says that the acknowledgments are the only part of an engineering book that is easily understood at a first glance. Having endured the reading of many well and not so well written engineering books during all my years in school, I tend to agree with his assertion. I also believe that it is the only place in a mathematical book in which one can crack a joke, or make a very personal comment, without getting in deep trouble. I do not plan to crack many jokes here but, having written over a hundred pages full of the driest graphs and formulas anyone could imagine, would like to include a few lines showing one of the human qualities that make us slightly different from an EPROM or a Neural Network: appreciation. Those lines should also help me establish that the next pages are the brainchild of someone capable of having emotions and not of a gravedigger coldly passing mathematical equations from some tombs to another. So, here I go.

Firstly, I would like to thank my thesis supervisor, Dr. Guy A. Dumont, for giving me the opportunity to work under his guidance. Despite his many obligations, I always benefited from the overall direction he provided. I am also grateful for his financial support.

I would like to extend the warmest thanks to the people I have met at The Pulp and Paper Centre during the course of my studies: Ye Fu, Jahan Ghofraniha, Ivar Jonsson, Kristinn Kristinsson, Navid Mohsenian, George Wang, and others, for their cooperation and encouragement. Sometimes they have shared knowledge some other times just moral support but, one way or the other, they have always been most helpful.

I must also express my gratitude to Bruce Allison, David Browning, Robert Gooding and John Hoffmann of the Vanlab's Screen Pilot Plant at Paprican for allowing me to play with their toys and for letting me bring my own ones into their playground.

And last but not least, I would like to thank the members of my family: Pilar, Alfonso and Diego, who gracefully endured these years of voluntary exile in Vancouver. We can move to sunnier places now, unless they don't want to.

Chapter 1

Introduction

1.1 Pulp Screening Systems

Screening and cleaning systems are found in all pulp and paper industry plants. The purpose of these systems is to remove certain unwanted constituents –such as coarse fibres and inorganics– from a pulp slurry, so that the suspension becomes more suitable for the product in which it will be used: paper, cardboard, etc. [30]. The removal of particles in screens is somewhat dependent on particle size and shape, whereas in centrifugal cleaners it is based on the principle of particle weight [43].

Screens are employed in all pulping processes, mechanical and chemical. Their physical location can differ. Thus, we find them placed between the cooking and bleaching stages or after the bleach plant in chemical pulp mills, and after or between the refining stages in thermo-mechanical pulp mills. Screens are also installed ahead of the paper machine. In all cases, the principles of operation and control are the same, although the nature of the constituents to be removed is different. Generally speaking, we can divide the screening systems in two classes: coarse and fine [47]. As the name suggests, coarse screening involves the removal of large or coarse materials from the slurry, whereas fine screening is designed to deal with smaller contaminants.

There are several types of screens currently in use in mills. The major types are flat, rotary, centrifugal, and pressure [60].

The focus of this research was aimed at the screening systems used in mechanical pulping. In this environment, fine screening is predominant, and pressure screens are utilized in virtually all operations.

1.2 Fundamentals of Pressure Screens

The mechanisms by which screens remove the unwanted materials from the pulp suspension are not fully developed as a mathematical corpus. Nevertheless, there is general agreement with the idea that separation in pulp screens is accomplished by the combination of two fundamental

1

processes, barrier screening and probability screening [20]. In the former, the screen plate orifices are smaller than the size of the contaminants and the screen behaves much as a sieve. In the latter, the openings are larger than the unwanted materials but some physical factors prevent their passage. Among the probability screening mechanisms most cited in the literature are the angle at which the particles approach the screen plate apertures (critical angle), the force created by the difference in fluid velocities inside the screen (shear force), the movement of fibres inside the screen (fiber alignment), the shortening of screen plate apertures due to the rotation of the plate (apparent hole size), and the interaction of neighboring fibres (fiber mat) [29], [30].

To measure how well a screen is performing it is necessary to have an indication of how much undesired material has been removed from the incoming pulp stream. One performance parameter widely used is called the *debris reject efficiency*, E_r. This measures the contaminants removed as a percentage of the mass of contaminants in the incoming pulp, according to the formula:

$$E_r = \frac{S_r}{S_i} R_w \qquad (1.1)$$

where

R_w : Mass Reject Rate [non-dimensional]

S_r : Debris concentration in rejected pulp [in %] $\qquad (1.2)$

S_i : Debris concentration in incoming pulp [in %]

The *mass reject rate* R_w is defined as

$$R_w = \frac{T_r}{T_i} 100 \qquad (1.3)$$

where

T_r : Mass flow of rejected pulp [in Tons per day]

$\qquad (1.4)$

T_i : Mass flow of incoming pulp [in Tons per day]

[47]. The classical method used in industry to evaluate screen performance is the shive removal efficiency (SRE_r) vs. reject rate (R_w) curve illustrated in Figure 1.1. This graph, as the name suggests, measures the screen reject efficiency as defined in (1.1) but regarding shives. A shive is an intact fiber bundle or fibrous mass having a contrasting color to the pulp fibres and having dimensions greater than some arbitrarily set minimum [61] and, as such, is a particular type of debris. The plot
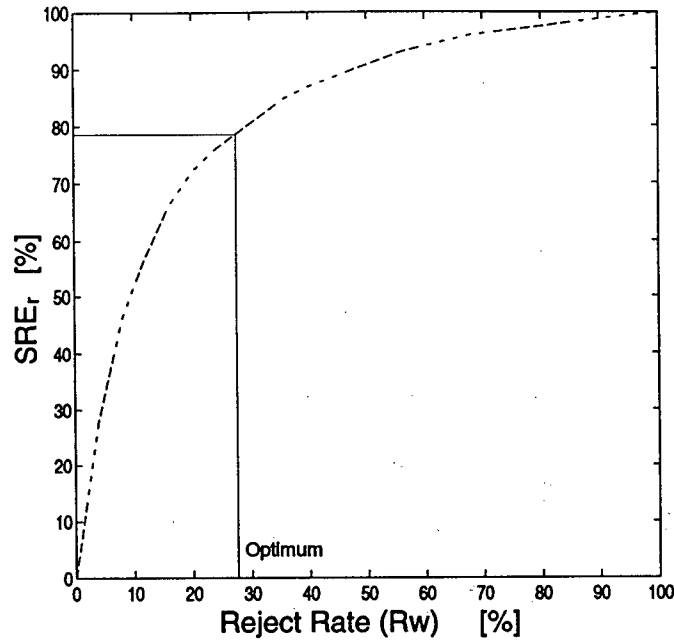
Figure 1.1: Efficiency vs. Reject Rate Curve [14]

displays the ratio between the volumetric flow of rejected shives and the total flow of incoming shives, as a function of R_w. As shown by the curve, the higher the percentage of flow going through the reject line, the higher the percentage of debris removed. Complete removal of the contaminants in the pulp can only be achieved if all flow goes to the reject line, i.e., no pulp is accepted. This is obviously not practical. With approximately 25% of the total incoming fibres going to the rejects (the proportion most often found in mills), around 80% of the shives present in the pulp are expected to be removed[1]. If all feedstock flow is accepted (0% reject rate) no removal will take place.

In order to avoid the need for a complete mass and debris balance, which is a difficult and time consuming task, some formulas have been proposed. The best known uses the *screening quotient* Q which is defined as

$$Q = 1 - \frac{S_a}{S_r} \qquad (1.5)$$

---

[1] The real figure varies for every pressure screen.

3

Rotor

Screen
Plate

Reject
Line

Feed Line

Accept Line

Dilution Line
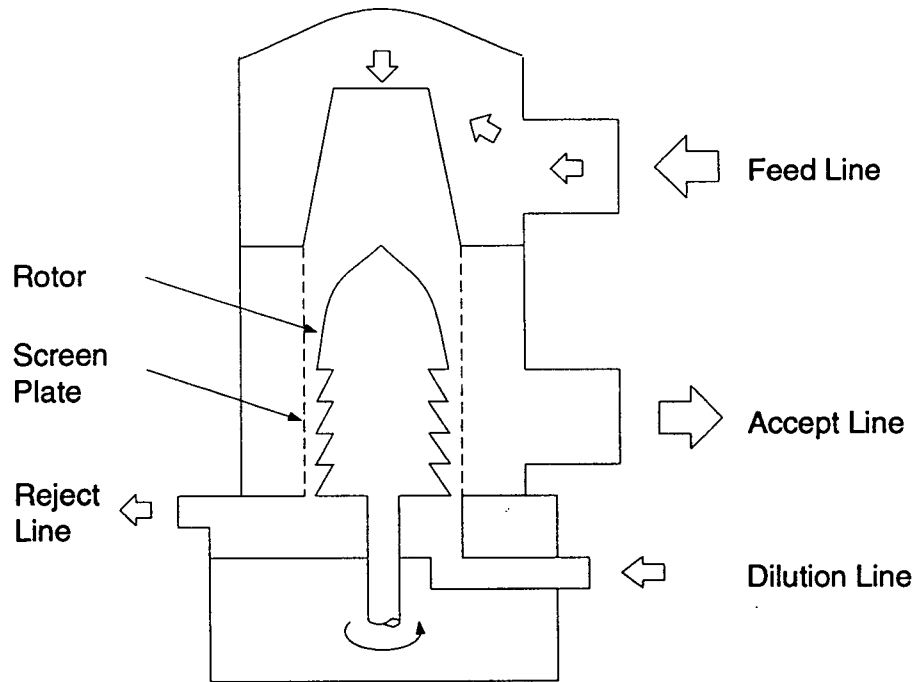
Figure 1.2: Typical pressure screen

where S_a is the debris content by weight in the accepted pulp (in percentage units) , and leads to a different expression for the relationship between E_r and R_w, namely,

$$E_r = \frac{R_w}{1 - Q + R_w Q}$$ (1.6)

[50]. Q is proposed as a constant that establishes screen performance independently of reject rate because, if Q is known, the screen efficiency can be characterized only with measurements of the accepts and rejects shive concentrations. This formula, however, has limitations [30].

Another important parameter to measure screen performance, albeit one with several meanings in the literature, is *capacity*. A good definition is the one given by Gooding and Craig [21]. For them, capacity is the volumetric accept flow rate of a screen at a given pressure differential.

1.3 Pressure Screen Design

There are many configurations of pressure screens in use today, but most of them are built with a cylindrical screen plate. In a typical model, as the one seen in Figure 1.2, incoming stock is tangentially fed into the inside of the cylinder. The pressure will force the passage of the suspension

through the openings in the plate to an accept compartment, from where it is tangentially discharged into the accept piping[2]. The rejected material passes down the inner surface of the plate into a reject section, whence it is discharged into the reject piping. Dilution water –which may or may not be used– can be added into the reject compartment and also into the screening zone. An electric motor which drives the rotor is utilized as a means of preventing the screen from plugging and to accelerate the pulp suspension.

The most important feature of any screen is the configuration of its plate. Depending on the application, the screen plate can have a smooth incoming surface or a grooved one. Also, its openings can either be round holes or slots.

1.4 Control Goals for Pulp Screens in the Mill

The variables that affect the performance of any pressure screen are the following:

1. Screen plate design, i.e., shape and size of the openings and type of surface.

2. Rotor and screen housing design.

3. Feed stock flow rate.

4. Feed stock pressure.

5. Pressure drop across the screen.

6. Retention time of fibres.

7. Reject flow rate.

8. Accept flow rate.

9. Stock characteristics, i.e., amount of debris, type of fibres.

10. Feed stock consistency.

11. Dilution flow to screen.

12. Feed stock temperature.

13. Rotor speed.

---

[2] In some pressure screens the incoming-accept flow is reversed. Incoming stock will be fed to the outside and accepts will go out from the inside of the screen plate.
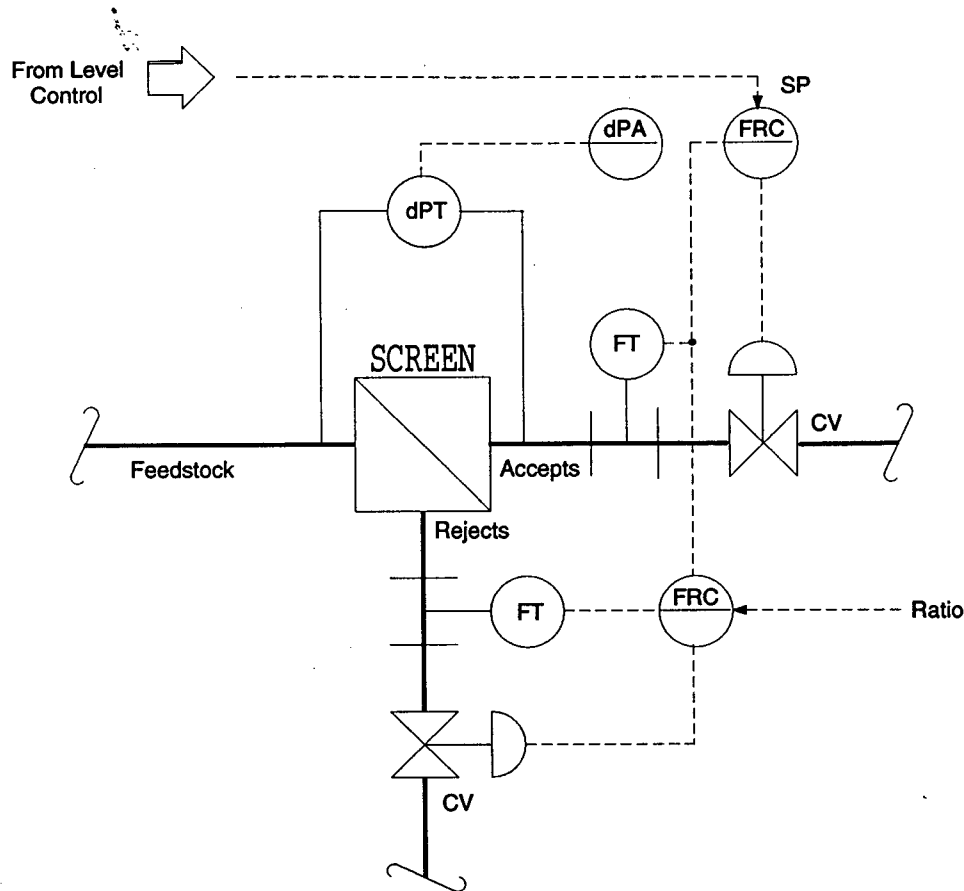
Figure 1.3: "Classical" control scheme for a pressure screen

Some of these variables are strongly related. For example, retention time is dependent on the stock flow rate and the reject rate, and pressure drop is a direct function of the flow through the screen plate.

The most important variable –albeit one that can not be modified during operation– is the screen plate design. In practice, the "classical" control strategy calls for three parameters to be continuously controlled by the plant operators: reject rate, feed consistency and internal dilution [30]. The stock flow rate and the rotor speed are seldom modified. The remaining variables are either secondary, or out of reach for an on-line controller. Figure 1.3 shows the typical control schematic for a pressure screen.

The ideal screen control would allow the extraction of all the debris in the feed stock without removal of any acceptable pulp [58]. This objective cannot be directly achieved, among other reasons, because there are no reliable on-line measurements of pulp quality in industry to date. Therefore, the best possible control is one which would optimize the relationship between the amount of debris that

will go with the accept flow and the amount of good fiber that will be rejected, the screen capacity, and the economic and operational aspects of the pulp processing plants, through manipulation of indirect variables. Unfortunately such control must be tailored to satisfy the specifications of the product being made. In practice, a general controller most likely does not exist.

1.5 Literature Review

Even though pressure screens were introduced to the pulp and paper industrial community over thirty years ago, not much has been published about their control and even less on their faults. The literature on screening seldom dwells on this operational aspect, dealing mainly with the fundamental physical principles that govern the operation of these devices. Nevertheless, screening control has been acknowledged as a promising prospect, expected to develop very fast in the coming years [39].

The 'classical' control strategy for pressure screens has called for either or both of two requirements to be met: a stable operating pressure and a stable stock throughput, measured as a flow rate, to achieve optimum screening efficiency [48]. The standard technique for screening control shown in Fig. 1.3 operates as follows: the accepts valve is used to control the screen throughput and the desired operating point to maximize efficiency is controlled through the ratio of flows –accept vs. reject. The pressure drop is indirectly controlled by this ratio and, under normal conditions, is expected to remain constant. However, if it increases above a preset level, the accepts valve will be fully closed and the rejects valve will be fully opened to purge the screen. Usually, an alarm is provided to warn the operator in the event of such abnormality [49], [64], [36], which is the only type of screen fault addressed by this scheme. An alternative calls for differential pressure control through a feed flow valve instead of the accepts valve [55], with no changes made for the alarm.

On-line control based on shive content of the pulp has been suggested by J. Hill et al. [25], [26]. In this approach, the mass reject rate R_w and consequently the flows in the stock and rejects lines are adjusted in accordance with the shive content and size distribution in the accepts slurry. Stock *freeness*, defined as the resistance of fibres to the flow of water when pulp drains [60], has also been used to control the reject rate in screens [7]. Both alternatives suffer from sensor problems, and do not address the issue of faults.

7
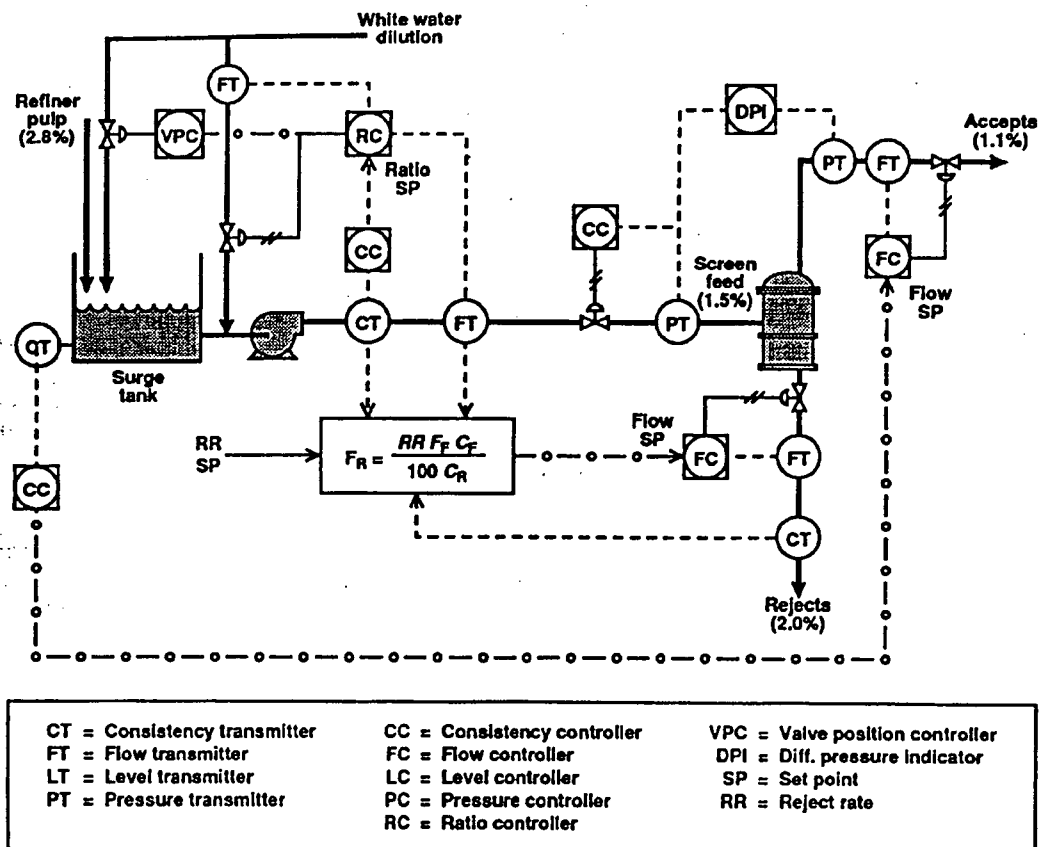
Figure 1.4: Daishowa America mill pressure screen control system (from reference [14])

A recent implementation of screen control uses a more sophisticated strategy, based on one supervisory loop and seven regulatory loops for each screen, as seen in Figure 1.4 [14].

The purpose here is to put controller set-points in terms of actual reject rates and shive removal efficiencies rather than in terms of flows and consistencies. This amounts to the imbedding of the efficiency vs. reject rate curve in the control strategy, instead of the mere maintenance of a constant differential pressure across the screen. Three categories of tasks are achieved by the control loops: rejects control, feed control and inventory control. The first one maintains the percentage of fibres flowing through the reject line constant. The second one controls feed stream pressure and consistency, and the last one prevents the screen feed tank from running empty or from overflowing through normal operation.

8

It is important to mention that despite the sophistication of this control scheme, the detection of screen faults is achieved by the same means used in more modest schemes: measurement of the pressure differential between the feed line and the accepts line. Moreover, the only fault monitored is screen blinding.

1.6 Objective

The objective of this thesis is to develop a strategy for fault detection of pulp pressure screens which can overcome the shortcomings of the approach currently used in industry. As said before, that method relies on monitoring of the pressure differential between the feed line and the accept line, and on setting off an alarm when the differential exceeds a certain pre-fixed value. It will be shown that this pressure differential alarm can be misleading as a fault detection tool at best, and can severely curtail the throughput and efficiency of a screen.

The current approach to screen fault detection has also the following weaknesses which shall be addressed:

- It is too slow, as it shows failure of the screen after it has occurred. In reality, that means that no true fault detection scheme is applied because not much can be done to prevent screen failure.
- It does nothing to detect screen faults different from gradual plugging of the screen plate.
- It leads to an inflexible control scheme, as it does not easily allow for changes in the feed rate of the screen.
- It does not lend itself to being used with varying consistencies in the pulp, as this condition affects the pressure differential alarm limit.
- It cannot be linked to the quality of the pulp.

A true fault diagnosis algorithm for screens had never been attempted in the past. The existing methods used to control pulp screens in industry do not prevent screens from plugging with all the unwanted consequences that this problem produces: process interruptions, unsafe working conditions, economic losses, etc. Even the newest strategies used for screen control consider the screen as a system with one or more SISO loops with no coupling between them [40], [14]. This assumption

9

does not seem realistic and most probably has an influence on the poor control performance seen in industry, in general. The research proposed here shall improve this situation by using a MIMO approach better suited to the reality of the dynamics of the screen. It also has the potential to open the door to sophisticated MIMO control techniques which cannot be used today.

Finally, one could also identify at least three kinds of potential benefits to be obtained by pursuing the research presented in this thesis: improvements on existing industrial control methods, establishment of a fault diagnosis technique suitable for industrial environments in general, and advancements in on-line estimation and control techniques.

1.7 Contribution of this Work

To the best of my knowledge, no work has ever been published on fault detection and isolation (FDI) for pressure screens. Having said that, I believe that the major contributions of the present research are the following:

- The introduction of faults in pulp pressure screens as a separate subject of study.
- The derivation of the first dynamic model in the literature to characterize pressure screen behavior.
- The development of the first model-based approach for fault detection on pressure screens.
- The use of the Singular Pencil Matrix estimation technique in a true practical application.
- The enhancement of knowledge on pressure screen operation and faults, in general, and the gain of valuable physical insight for better screen control.

At the same time, the strategy of using a dual identification method to obtain simultaneously the state variables and the process parameters of a system introduces more flexibility into the general universe of fault detection and Isolation (FDI). Having both quantities estimated at once might help to overcome the weaknesses of the fault detection methods based in only one of these groups of process variables, namely, the need for an extremely precise model of the system under identification, and the need to have as many working equations as parameters in the system to achieve full fault identification.

10

The testing of the proposed algorithm for FDI in a pilot plant also brought the potential to develop results suitable for the industrial environment, where noise and perturbations are the rule rather than the exception.

1.8 Outline of the Thesis

The thesis is organized as follows. In Chapter 1 the basic concepts related to pressure screens and screen control are presented. Chapter 2 provides a look at background material on fault detection in general. It also focuses on the issue of faults and failure in pressure screens, the current method used in industry to prevent its occurrence and its shortcomings. Chapter 3 describes several attempts at using non-parametric methods for the purpose of detecting screen faults, and their results. Some conclusions are then derived. Chapter 4 shows the derivation of a mathematical dynamical model of a pressure screen needed to use the parametric fault detection methods. The different techniques available for on-line identification of the model are presented. Chapter 5 presents the environment in which the testing of the parametric-based methods took place and it shows the results. Then, some conclusions are drawn. Chapter 6 explores some alternative techniques for improving the efficacy of the model-based fault detection schemes. Chapter 7 concludes the thesis summarizing the overall results and mentioning avenues of further research.

Chapter 2

Faults, Screen Faults, and their Detection

2.1 Faults, Failure and Fault Detection

2.1.1 Some definitions

Following Himmelblau's definitions, when talking about some equipment it must be implicit that the terms *fault* and *malfunction* mean "departure from an acceptable range of an observed variable or calculated parameter associated with the equipment" [27]. In this context, a fault implies a minimum degradation of the normal performance. The term *failure* will be used as "an indication of complete inoperability of equipment or the process" [27]. From these definitions, it is obvious that the criteria for establishing the presence of a fault is a subjective matter which depends on the characteristics of the process under observation. Failure, on the other hand, has no ambiguity.

The structure of a *fault detection and isolation* (FDI) system is similar to a closed loop control system. Figure 2.5 shows the block diagram proposed by Isermann to characterize the different supervisory stages of a typical FDI system [33].

The first step is called *fault detection*. As the name implies, it involves the detection of a fault. For this, some process variable, measurable or estimated, is tested against a certain range of a reference value. In the second stage the location of the fault is established and its most probable cause
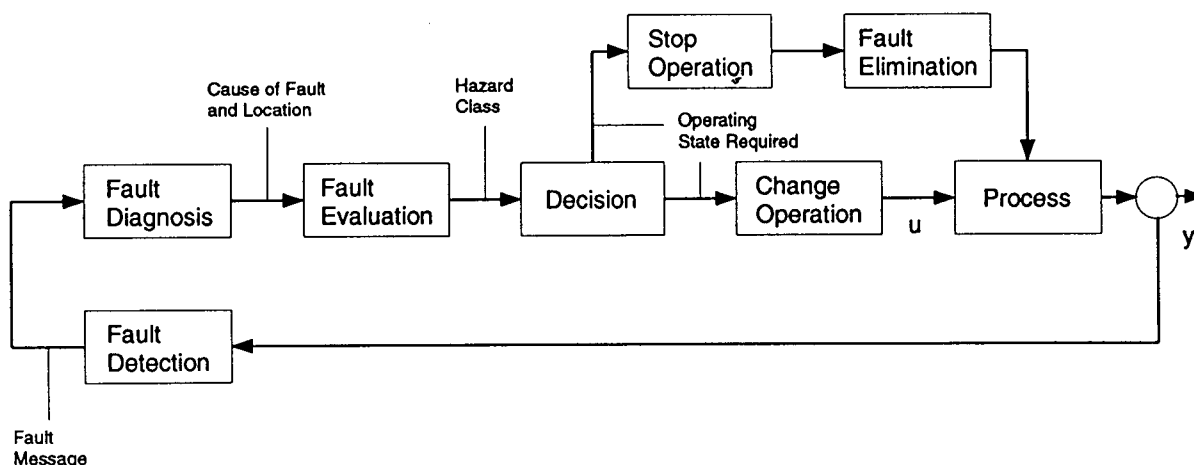
Figure 2.5: Supervision loop under appearance of a fault (from reference [33])

is determined. This is known as *fault diagnosis*. Third, an assessment of how much the process will be affected has to follow. Its name is *fault evaluation*. Finally, a *decision* regarding the required control action can be made. Depending on whether the fault has been deemed tolerable, conditionally tolerable or intolerable (failure), the plant operation must continue, or be modified, or be stopped.

Several classifications for fault detection and identification (FDI) have been suggested in the literature. Some of them are based on the type of possible malfunctions to be identified, some others are based on the methodology used, and yet some others on whether the information gathered is redundant in terms of the hardware (*hardware redundancy*) or in terms of the software (*analytical redundancy*) [27], [33], [18], [53].

From a systems identification point of view, the fact that a system may or may not have an analytical representation seems to be the most suitable characteristic to be taken into account when attempting a classification of FDI. Further refining will lead us into a scheme such as the one shown below,

$$
\text{FDI Systems}
\begin{cases}
\text{Non-model Based}
\begin{cases}
\text{Measurable signals U,Y} \\
\text{Characteristic Quantities } \eta = f(U,Y) \\
\text{Knowledge Based}
\end{cases} \\
\\
\text{Model Based}
\begin{cases}
\text{State Variables X} \\
\text{Characteristic Quantities } \eta = g(U,Y,X,\theta) \\
\text{Process Parameters } \theta
\end{cases}
\end{cases}
\tag{2.7}
$$

where the definitions are as follows:

- *Non-Model Based System:* FDI schemes do not contemplate a model representation of the system under supervision. Nevertheless, they might include mathematical models of the system signals.
- *Measurable Signals:* Process faults are detected and singled out with the aid of measurable input $U(t)$ and output $Y(t)$ plant signals.

- *Knowledge-Based System:* Fault detection is based on some stored knowledge of the plant (usually of qualitative nature) and inference mechanisms that allow for the making of comparisons and deductions between the plant behavior and this stored data.

- *Model Based System:* The fault detection schemes are based on a mathematical process model of the system under supervision or of parts of it.

- *State Variables:* Faults are detected with the aid of partially measurable and partially non-measurable internal state variables of the supervised system.

- *Process Parameters:* FDI schemes are based on constants or time-dependent coefficients which appear in the mathematical model used to represent the supervised plant.

- *Characteristic Quantities:* Malfunctions of a system are diagnosed with the aid of a combination of information sources: state variables, process parameters and modelling of plant signals (inputs and outputs).

2.1.2 Non-Model-Based Methods

As defined before, all fault detection methods which do not require a parsimonious model of the process or system under measurement are called non-model based, or more loosely, non-parametric.

FDI approaches which involve the use of *measurable signals* and *non-model based characteristic quantities* generally require the performance of three operations: observing the records of the signals or quantities in time, computing simple statistics of those records, and carrying out some tests to determine the presence of a fault. The results are usually presented in graphical form, as a function of sample sequence or time. A fault is deemed to have occurred whenever a non-random change brings the observed variable out of the pre-established statistical limits. The best known methods are the Shewhart Control Chart, which involves the use of mean and sample range values [59]; the Cumulative Sum Control Chart, which makes use of sums of a function of a random process variable[3]; the Geometric Moving Average Control Chart, which involves the use of smoothed values of the sample mean [27]; and Multivariate Control Charts [37], which use statistics calculated from values of many variables which might have joint probability distributions.

---

[3] The function can be the variable itself, the difference between the variable and a target value, the sample mean, etc.

Himmelblau [27], Pau [54], Willsky [69] and Wise et al. [70], provide diverse examples of those FDI techniques falling under the category of *measurable signals* and *non-model based characteristic quantities* systems.

Knowledge-based FDI methods are based on the use of artificial intelligence techniques. They involve the use of a computer program (which usually emulates the behavior of a highly experienced human operator), and qualitative and/or quantitative techniques based on existing knowledge of the system, to determine the presence and nature of faults.

The key issues in implementing a *knowledge-based* FDI system involve the choice of: a symbolic representation for the knowledge on the system, the rules which allow for the symbolic reasoning, and a graphic interface to provide the representation and reasoning.

S. Tzafestas in the book by Patton et al. [53], Y. Ishida [35], Kramer and Finch [42], and Kramer and Leonard [41] are only a few references on the growing area of FDI *knowledge-based* techniques.

2.1.3 Model-Based Methods

Whenever the fault detection schemes are based on the use of a mathematical process model of the system, they are called model-based or parametric.

Process parameters are defined as coefficients or constants that appear in the mathematical process model of a plant. These coefficients reflect, directly or indirectly, the physical parameters like friction, mass, viscosity, length, etc. of the system under supervision and, therefore, many faults can be determined by examining their behavior [32]. An FDI system based on *process parameters* has a block representation as the one shown in Figure 2.6.

The mathematical model of the plant can be either static or dynamic. Provided that we have a SISO system, the former might have the form

$$Y(U) = \beta_0 + \beta_1 U + \beta_2 U^2 + \ldots \tag{2.8}$$

and the latter, for processes with lumped parameters might have the form

$$y(t) + a_1\dot{y}(t) + a_2\ddot{y}(t) + \ldots + a_n y^{(n)}(t) = b_o u(t) + b_1\dot{u}(t) + b_2\ddot{u}(t) + \ldots + b_m u^{(m)}(t) \tag{2.9}$$
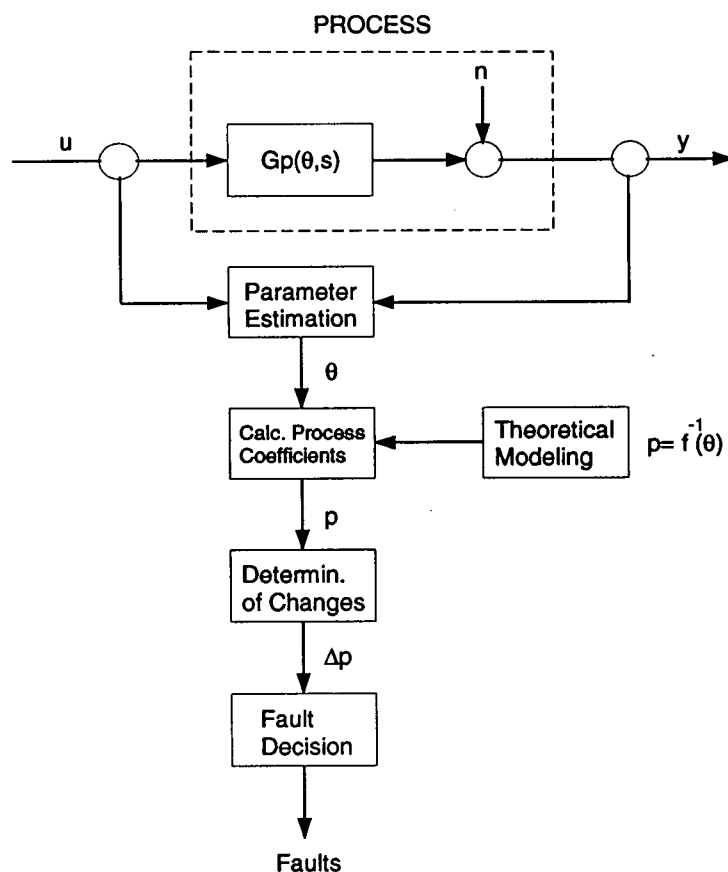
PROCESS

Figure 2.6: Fault detection scheme based on parameter estimation (from reference [33])

which is usually, for simplicity, a linearization about an operating point.

The general procedure for fault detection and isolation using *process parameters* will include the following steps:

1. Determination of the equation used for process modelling, $Y(t) = f\{U(t), \theta\}$

2. Determination of relationship between the model parameters θ_i and the physical parameters p_j as $\theta = f(p)$.

3. Estimation of the model parameters.

4. Determination of the physical parameters, $p = f^{-1}(\theta)$.

5. Calculation of variations in physical parameters Δp_j.

6. Determination of type and location of fault.

16

In an industrial environment one would expect steps 1 and 2 to be performed off-line and the remaining to be executed on-line.

Since changes in the physical parameters are the main concern, the model of the 'normal' process, including its tolerance range, must be known accurately at all times.

Two parameter estimation methods for continuous-time models have attracted the bulk of attention in the FDI literature : Least-squares and Instrumental Variables. In the first one, our time-invariant linearized equation

$$a_0 y_u(t) + a_1 \dot{y_u}(t) + a_2 \ddot{y_u}(t) + \ldots + y_u^{(n)}(t) = b_o u(t) + b_1 \dot{u}(t) + b_2 \ddot{u}(t) + \ldots + b_m u^{(m)}(t) \quad (2.10)$$

has the measured output $y(t)$ contaminated by a stationary stochastic noise $n(t)$, therefore,

$$y(t) = y_u(t) + n(t) \quad (2.11)$$

and substituting for $y_u^{(i)}(t)$ in terms of its measurements will lead to

$$y^{(n)}(t) = \psi^T(t)\theta + e(t) \quad (2.12)$$

with $e(t)$ being the equation error, and ψ, θ being the regressors and the unknown parameters, respectively. Once input and output signal measurements have been made and the derivatives determined at discrete times $t = kT_0$, with $k = 0, 1, 2, \ldots N$ and T_0 the sampling time, N+1 equations will result with the form

$$y^{(n)}(k) = \psi^T(t)\hat{\theta} + e(k) \quad (2.13)$$

where $e(k)$ is again an equation error. These can be expressed as the vector equation

$$y^{(n)} = \Psi\hat{\theta} + e \quad (2.14)$$

Having a cost function of the form

$$V = \sum_{k=0}^{N} e^2(k) = e^T e \quad \text{and} \quad \frac{dV}{d\theta} = 0 \quad (2.15)$$

will get the well-known least squares estimate of the parameter vector

$$\hat{\theta} = \left[\Psi^T \Psi\right]^{-1} \Psi^T y^{(n)} \quad (2.16)$$

17

It must be remembered that, since the parameters are biased in the presence of colored noise, this method is not well-suited for high noise-to-signal ratio cases. For the equations above, the time derivatives of the input and output signals can be obtained by several methods: backward differences, spline interpolation, Newton interpolation or state variable filtering [33]. The state variable filtering method has the advantage of providing the derivatives and filtering the noise without differentiation [71]. When dealing with systems that have a high noise-to-signal ratio, the instrumental variables parameter estimation provides better results than least-squares. Basically, the technique consists of modifying the least-squares solution to include a vector of instrumental variables $\hat{x}$ which are chosen to be highly correlated with the noise-free output of the system y_u, but totally uncorrelated with the noise on the measurement of the system variables [71].

An FDI system based on *state variables* has the block representation shown in Figure 2.7. In general, the dynamic relationships that exist in a system can be expressed in state representation as

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$$

$$(2.17)$$

assuming that the system has been linearized around the operating point. The influence of a fault can be evaluated in the *residuals*, which are functions of the state variables that are accentuated by the malfunction. There are three preferred ways to generate residuals in the literature: the parity space approach, dedicated observer schemes and fault detection filters [18]. All of them share the presence of some kind of state observer.

The rationale behind the *parity space* approach is to check out the consistency of the state variables mathematical representation when confronted with the actual plant measurements. When the difference between the theoretical and the real values exceeds a certain range, a fault is deemed to be present. It must be stressed that partial consistency, i.e., relations that reflect only part of the overall model can be used advantageously. This concept was generalized by Lou et al. [46]. Assuming that the system is given by

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k)$$
$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k)$$

$$(2.18)$$

18
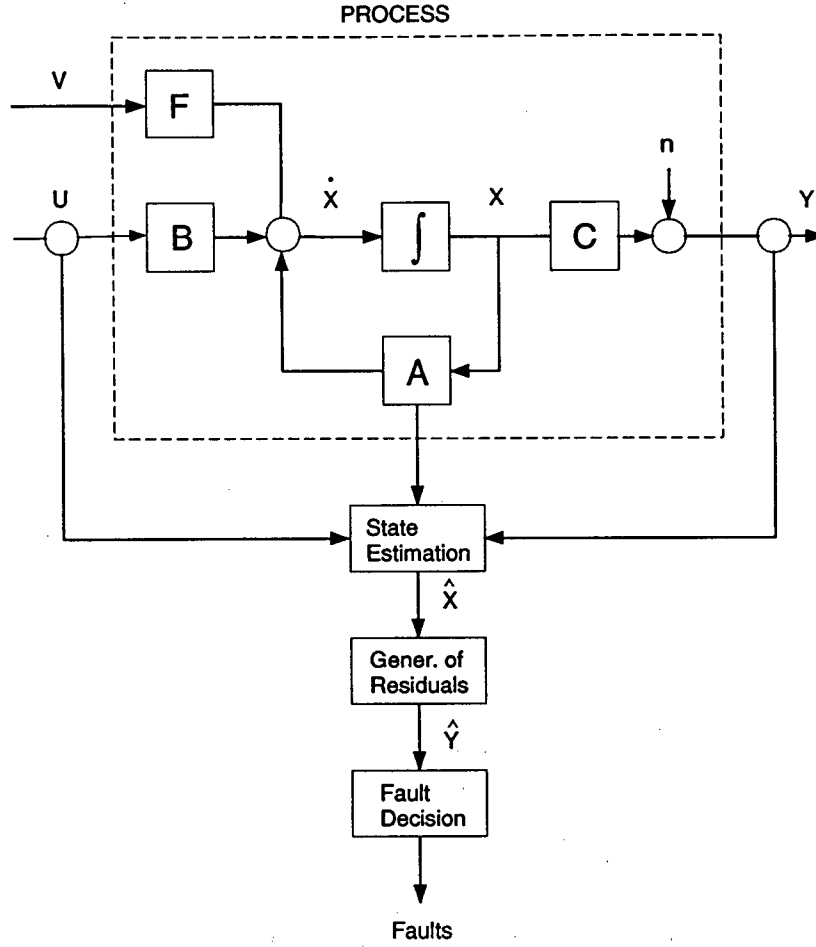
Figure 2.7: Fault detection scheme based on state variables estimation (from reference [33])

the output equation can be expressed by substitution as

$$\mathbf{y}(k+1) = \mathbf{CAx}(k) + \mathbf{CBu}(k) + \mathbf{Du}(k+1) \tag{2.19}$$

If we extend the formulation for any interval s such that $s > 0$ it turns into

$$\mathbf{y}(k+s) = \mathbf{CA^s x}(k) + \mathbf{CA^{s-1}Bu}(k) + \ldots + \mathbf{CBu}(k+s-1) + \mathbf{Du}(k+s) \tag{2.20}$$

Putting together all the equations for $s = 0 \ldots n'$; with $n' \leq n$ and shifting by n' we obtain the representation

$$\mathbf{Y}(k) = \mathbf{Rx}(k) + \mathbf{HU}(k) \tag{2.21}$$

19

with

$$Y(k) = \begin{bmatrix} y(k - n') \\ y(k - n' + 1) \\ y(k - n' + 2) \\ \vdots \\ y(k) \end{bmatrix} \qquad (2.22)$$

$$R = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n'} \end{bmatrix} \qquad (2.23)$$

$$H = \begin{bmatrix} D & \cdots & & & 0 \\ CB & D & & & \vdots \\ CAB & CB & D & & \\ \vdots & & & \ddots & \\ CA^{s-1}B & \cdots & \cdots & CAB & CB & D \end{bmatrix} \qquad (2.24)$$

$$U(k) = \begin{bmatrix} u(k - n') \\ u(k - n' + 1) \\ u(k - n' + 2) \\ \vdots \\ u(k) \end{bmatrix} \qquad (2.25)$$

If we define a subspace of $(s + 1)q$ dimensional vectors v such that $v^T R = 0$ (which is called the parity space of order s), and apply it to equation 2.21, we obtain the expression

$$v^T Y(k) = v^T H U(k) \qquad (2.26)$$

A parity check can be performed by using the different vectors v at any time. The residuals $r(k)$ can now be defined as

$$r(k) = v^T \left[\begin{bmatrix} y(k - s) \\ \vdots \\ y(k) \end{bmatrix} - H \begin{bmatrix} u(k - s) \\ \vdots \\ u(k) \end{bmatrix} \right] \qquad (2.27)$$

and if we introduce the state equations 2.18 in the expression for $r(k)$, it will become

$$r(k) = v^T \begin{bmatrix} C \\ CA \\ \vdots \\ CA^s \end{bmatrix} x(k-s) \tag{2.28}$$

with $r(k) = 0$ if no fault occurs. It is evident from all the formulations above that the vectors v must be in the left null-space of matrix

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^s \end{bmatrix} \tag{2.29}$$

but, apart from this, can be chosen freely. This leads to a variety of parity relations and, as mentioned above, these relations are tests in the consistency of parts of the input-output dynamics of the system as opposed to tests for the whole system model. Whenever a particular fault is present, the consistency is altered and that fact is shown in the residuals. The concept leads to an observer called the *dead-beat observer* for SISO and MIMO systems [18], which can be explained using the following line of reasoning. A state-space model can be described in input-output form as

$$y(k) = S(z)u(k) \tag{2.30}$$

where each element of the $S(z)^4$ matrix is a transfer function. If we use an observer to estimate the states of a particular system with $D = 0$, the equations we obtain are

$$x(k+1) = (A - KC)x(k) + Bu(k) + Ky(k)$$
$$r(k) = y(k) - Cx(k) \tag{2.31}$$

where r represents the innovations and K the observer feedback matrix. Following the equivalence between the state-space and the input-output representations, we could also express $r(k)$ as a function of the observables $u(k)$ and $y(k)$ as

$$r(k) = Q(z)y(k) + P(z)u(k) \tag{2.32}$$

---

4   $S(z) = C(zI - A)^{-1}B + D$

where $\mathbf{Q}(z)$ and $\mathbf{P}(z)$ are transfer functions including $\mathbf{K}$. For the innovations to become FDI-usable residuals, they must be zero if the system is under no-fault and the observer $\mathbf{K}$ must have been chosen adequately, that is

$$\mathbf{Q}(z)\mathbf{y}(k) + \mathbf{P}(z)\mathbf{u}(k) = 0 \qquad (2.33)$$

As $\mathbf{y}(k) = \mathbf{S}(z)\mathbf{u}(k)$, that implies

$$[\mathbf{Q}(z)\mathbf{S}(z) + \mathbf{P}(z)]\mathbf{u}(k) = 0 \qquad (2.34)$$

or

$$\mathbf{P}(z) = -\mathbf{Q}(z)\mathbf{S}(z) \qquad (2.35)$$

Equation 2.32 then becomes

$$\mathbf{r}(k) = \mathbf{Q}(z)\mathbf{y}(k) - \mathbf{Q}(z)\mathbf{S}(z)\mathbf{u}(k) \qquad (2.36)$$

If we recall the form of the residuals in the *parity space* approach, namely,

$$\mathbf{r}(k) = \mathbf{v}^{\mathbf{T}}\mathbf{y}(k) - \mathbf{v}^{\mathbf{T}}\mathbf{H}\mathbf{u}(k) \qquad (2.37)$$

the relationship is evident [19].

The main idea behind the *dedicated observer* approach is to use the estimation error obtained with the use of system observers (linear or non-linear, full or reduced order) or Kalman filters, as the residual for fault detection. For a linearized system defined by

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{E}\mathbf{d}(t) + \mathbf{K}\mathbf{f}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{F}\mathbf{d}(t) + \mathbf{G}\mathbf{f}(t) \qquad (2.38)$$

where

$\mathbf{x}$: nx1 state vector

$\mathbf{u}$: px1 input vector

$\mathbf{y}$: qx1 output vector

$\mathbf{A}, \mathbf{B}, \mathbf{C}$: known matrices

$\mathbf{Ed}$: term for unknown inputs

$\mathbf{Kf}$: term for actuator and component faults

$\mathbf{Fd}$: term for unknown inputs to the sensors

$\mathbf{Gf}$: term for sensor faults

$$(2.39)$$

22

a full order observer will render the following state $\hat{x}$ and output $\hat{y}$

$$\dot{\hat{x}} = (A - HC)\hat{x} + Bu + Hy$$

$$\hat{y} = C\hat{x}$$

(2.40)

The choice of the feedback gain matrix H will determine the performance of the observer. The observer state and output estimation errors are

$$\underline{\epsilon} = x - \hat{x}$$

$$e = y - \hat{y}$$

(2.41)

and by substitution, will become

$$\dot{\underline{\epsilon}} = (A - HC)\underline{\epsilon} + Ed + Kf - HFd - HGf$$

$$e = C\underline{\epsilon} + Fd + Gf$$

(2.42)

The output estimation error e can then be used as the residual. When no faults are present, the residual will only depend on the unknown input. When a fault appears the residual will increase, therefore, detection of the increment in e will determine the presence of faults.

The *fault detection filter* approach is nothing but a dedicated observer with a special choice for the feedback gain matrix H. In this case, the system is modelled as

$$\dot{x}(t) = Ax(t) + Bu(t) + k_i f_i(t)$$

$$y(t) = Cx(t) + k_j f_j(t)$$

(2.43)

with $f(t)$ a scalar function of time and k_i and k_j vectors of dimension $n \times 1$ and $q \times 1$, respectively, used to model faults. By design, k_i can be chosen to represent the fault directions of the actuators and components in the system. k_j represents the directions and modes of sensor faults. With $i = 1, 2, \ldots r$, r is the number of fault directions. j is associated with the number of sensors in the system. The observer is then represented by

$$\dot{\hat{x}} = (A - HC)\hat{x} + Bu + Hy$$

$$\hat{y} = C\hat{x}$$

(2.44)

and the residual vector $r = y - \hat{y}$ has now directional properties associated with the possible faults. The estimation errors will reflect either an actuator or component fault as

$$\dot{\underline{\epsilon}} = (A - HC)\underline{\epsilon} + k_i f_i$$

$$r = C\underline{\epsilon}$$

(2.45)

23

or a sensor fault as

$$\dot{\underline{\epsilon}}_j = (\mathbf{A} - \mathbf{HC})\underline{\epsilon}_j + \mathbf{h_j}\bar{f}_j$$

$$\mathbf{r_j} = \mathbf{C}\underline{\epsilon}_j + \mathbf{k_j}f_j$$

(2.46)

where $\mathbf{h_j}$ is the j-th column of $\mathbf{H}$. From these equations it can be seen that the residual can be unidirectional, for component and actuator faults, or could lie in a plane, in the case of a sensor fault. The proper choice of $\mathbf{H}$ will make the residual of a particular fault f_i to be constrained either in a single direction or into a plane, in the residual space. Therefore, any large residual projections along the known fault directions or fault plane will mean that a fault is present, and will tell which kind of fault it is.

Needless to say, the *parity space*, *dedicated observer*, and *fault detection filter* approaches require very precise system modelling.

2.2 Faults and Failure of Pressure Screens

Abiding by our definitions, failure of a screen is produced mainly by two causes: plugging (either line plugging or plate plugging), and plate mechanical failure. These conditions demand the screen to be shut down to allow for off-line repairs.

All screen faults can be grouped in five categories:

1. Blinding. Fibres accumulate within or adjacent to the screen plate apertures (or in the reject piping). Low screen capacity will result from this condition and, in extreme cases, will lead to plugging and stoppage of the accept pulp flow.

2. High loss of good fiber into the rejects. Either by improper feed consistency, volume of dilution water or plate perforation [47].

3. Loss of accept quality. The presence of an undesired percentage of contaminants in the treated pulp would result as a consequence of high rotor speeds, low consistencies or wrong reject rate.

4. Motor overloads.

5. Flow instability, caused by high flow velocities and cavitation.

In general, the main objective of screening is the steady removal of debris at adequate capacity. Whenever this objective is not fully accomplished, there will be a faulty operating condition. As with the existing technology most variables associated with the quality of the pulp can not be measured on-line, a true FDI system should focus on detecting the remaining types of faults. In this research, the emphasis is put into the operational faults: blinding and motor overloads.

How long it will take to blind a screen depends on its operating conditions and on the severity of the blinding agents. Fast blinding can have a duration of a few seconds after the critical accept flow has been met. Slow blinding can develop over a period of 30 seconds or more. Consistency[5], rotor speed, length of fibres and flexibility of the fibres, are among the factors which can influence this phenomenon.

2.3 Present Standards and Practices

The "classical" technique for screening control has been shown in Fig. 1.3. As said previously, when it is implemented, the pressure drop across the screen is maintained constant. The accept valve is used to control the screen throughput and the differential pressure is controlled through the ratio of flows: accept vs. reject. However, if the pressure drop increases above a preset level, screen blinding is deemed to have occurred. Then, the accept flow is temporarily stopped and the reject valve will be fully opened to purge the screen, using the back-flushing effect of the screen rotor. Usually, an alarm will go off to warn the operator that screen plugging has happened [49].

The main shortcoming of this approach is that there are many other factors, besides blinding, which could cause the differential pressure to raise. Increased throughput is one of them. As a consequence, a healthy screen operating at high capacity could be stopped and flushed. This weakness is brought to light when one looks at the findings of Craig and Gooding [22], which were supported by the findings in this research, and which are reflected in Figure 2.8. The relationship between screen pressure differential and the square of the accept flow rate is clearly linear and positive when the screen operation is normal. This behavior is not maintained when the screen fails. In such case, after a critical accept flow is reached, the pressure differential will rise sharply without any increase in

---

5   Consistency is defined as the mass or weight percentage of bone dry fiber in a stock [61].
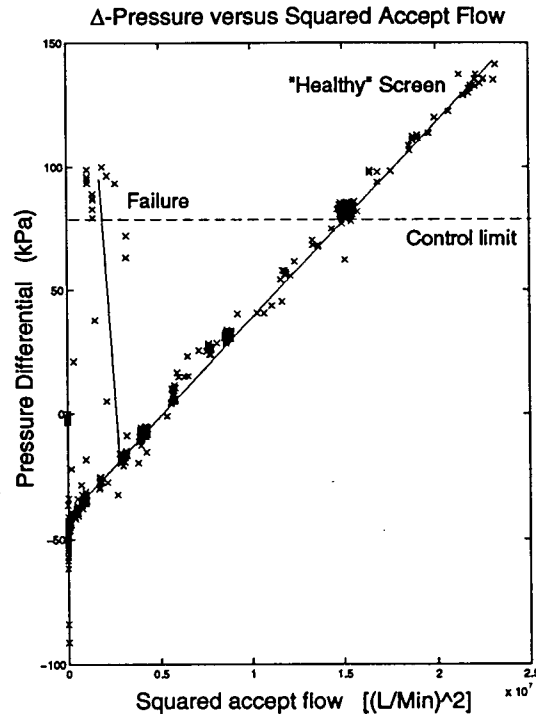
Figure 2.8: Pressure differential vs. squared accept flow relationships for a pulp that blinds and for a pulp that does not

accept flow. What this tells us is very simple: the use of pressure differential to detect faults implies an inefficient use of the screen, as it curtails the throughput. Obviously, the heavily used pressure differential technique is not satisfactory. Additionally, the fact remains that if the screen is under the effect of blinding, this fault is only detected after the fault has become a failure. This means that a fault detection scheme capable of giving warnings of incipient operational deviations is lacking.

2.4 Shortcomings

A true fault diagnosis scheme for pressure screens had never been attempted in the past. Neither the "classical" method nor the alternative schemes used to control pulp screens in industry prevent screens from plugging. Even the newest strategies used for screen control consider the screen as a system with one or more SISO loops with no coupling between them [40], [14]. This assumption does not seem realistic and most probably has an influence on the poor control performance generally seen

26

in industry. That translates into interruptions in the pulp-making process, unsafe working conditions and economic losses.

The current approach to screen fault detection has also the following weaknesses which shall be addressed:

- It is too slow, as it shows failure of the screen after it has occurred. In reality, that means that no true fault detection scheme is applied because not much can be done to prevent screen failure.

- It does nothing to detect screen faults different from screen plate blinding.

- It leads to an inflexible control scheme, as it does not easily allow for changes in the feed rate of the screen.

- It does not lend itself to being used with varying consistencies in the pulp, as this condition affects the pressure differential alarm limit.

- It cannot be linked to the quality of the pulp if future on-line measurements become available.

The present research shall improve this situation by using a MIMO approach better suited to the reality of the dynamics of the screen. It also has the potential to open the door to sophisticated MIMO control techniques which cannot be used today.

2.5 Summary

In this chapter, the fundamentals of Fault Detection and Isolation (FDI) systems and the principles of their application on pulp pressure screens have been described. After delving into some definitions like *fault*, *failure*, and the different steps in a typical FDI system, a classification of such systems has been introduced. The criteria used for said classification is the existence of an analytical representation of the process plant under control. Details of the fundamental ideas behind the use of non-model or model based FDI approaches have been provided. The state of the art of Fault Detection on pressure screens and its shortcomings have been presented. Finally, the need for a new FDI method has been substantiated.

Chapter 3

Non-Parametric Fault Detection of Screens

3.1 The Experimental Setting

Process and equipment faults are undesirable events in a process plant, and failure is to be avoided if at all possible. Developing a fault detection strategy in an environment where product deviations and equipment malfunctions are unwelcome is not a simple task and usually becomes a lengthy process, since not much can be done to expedite progress, or to study particular alternatives. On the other hand, the use of theoretical tools alone to form a fault detection scheme for equipment in industry is of limited use. In any case, direct knowledge of the real system under study, and an understanding of what constitutes normal and faulty operations, are required. So, how can one build an FDI system for pulp mill equipment? It seems that the best approach for such task should be one which combines the flexibility of the theoretical tools with the depth of knowledge gained in the industrial operations.

Fortunately, in the course of this research, access to Paprican's screening research pilot plant, located in Vancouver, B.C. was gained. This plant, whose *Process and Instrumentation* (P&ID) diagram can be seen in Figure 3.9, comprises three storage tanks, a centrifugal pump, two industrial-scale pulp screens, and the instrumentation required to measure and control all important process variables of the equipment operation. It allows for circulation of pulp flows up to 10,000 liters per minute (L/Min) with a consistency of up to 5%. In this setting, it was possible not only to operate the screens and observe their behavior, but most kinds of faults could be induced at will and their effects documented. This pilot facility is as close as possible to the industrial reality while maintaining the rigor of a research environment.

Two main sources were used to gather the information regarding the screen process variables: asynchronous transfer of data from a *Distributed Control System* (DCS) Bailey Infi 90 which governs the automatic operation of the pilot plant, and a data acquisition personal computer (PC) capable of extremely fast sampling. In a period of several weeks, cabling, junction boxes, terminals,
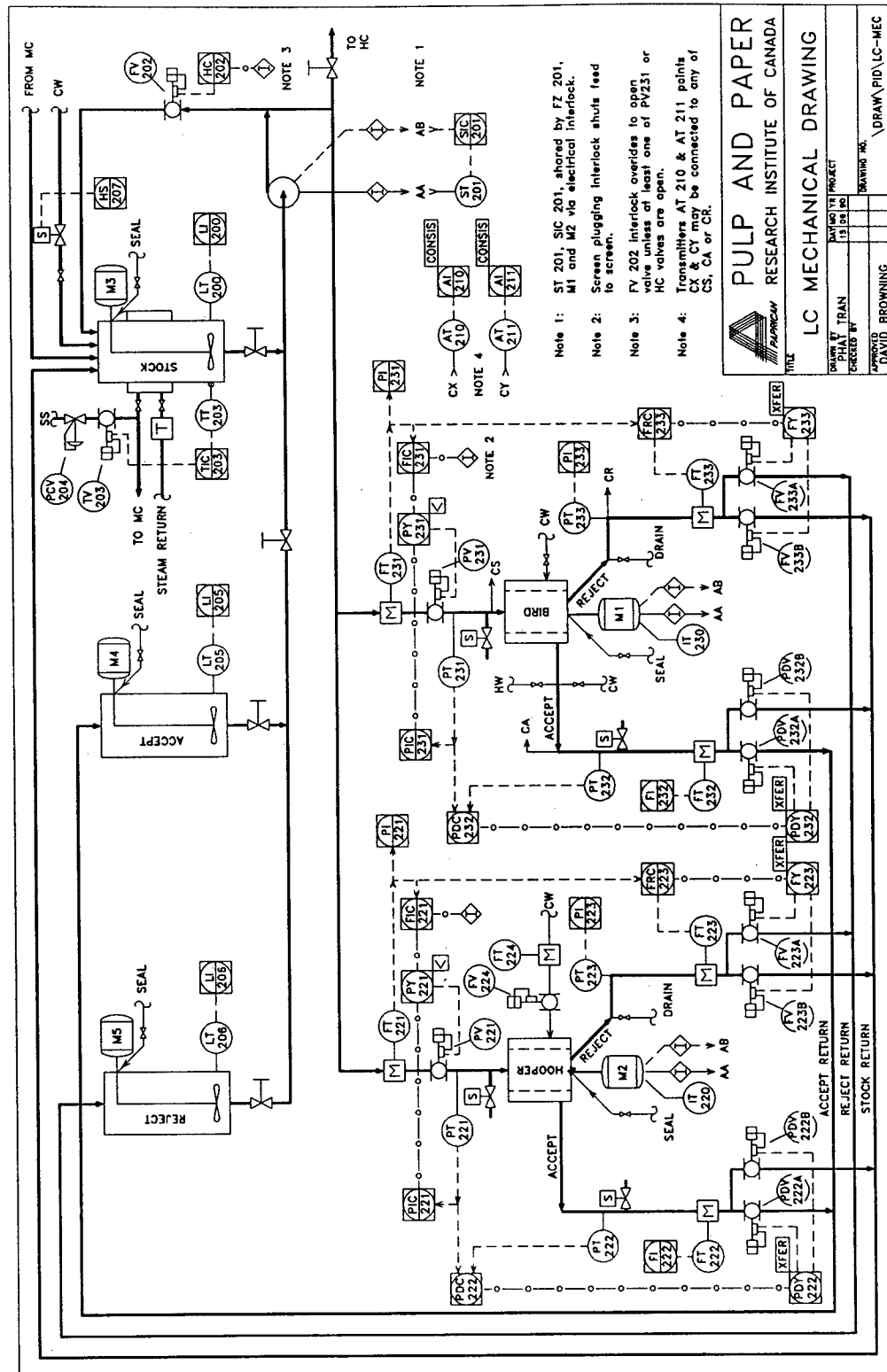
Figure 3.9: Paprican's Van-lab Screen Pilot Plant. PID diagram
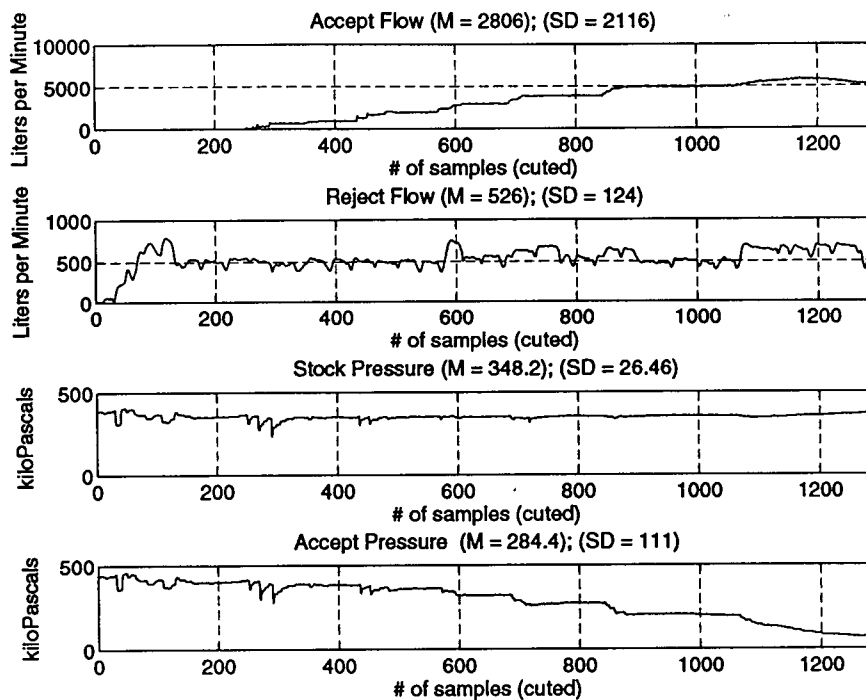
Figure 3.10: Typical Slow Sampled Data. Unfiltered

signal conditioning circuits and anti-aliasing filters were designed, built, and connected to the DCS instrumentation inputs. Having two "independent" sources allowed for the study of process data taken simultaneously at different rates of sampling, as can be seen in Figures 3.10, and 3.11.

The typical fault trial would include the following steps:

1. One of the two screens is chosen for the trial. This includes plate selection.

2. The pulp consistency is adjusted by mixing dry pulp and water in the stock tank. This has a capacity of 11,000 liters and a side-mounted mixer.

3. The reject valve is set at 75% open. The stock valve is set at 20% open and the accept valve is closed.

4. The reject flow rate for the trial is chosen. Several accept flow rates are selected, in incremental steps. The stock pressure is to be maintained at 350 kPa.

5. The pump and screen motors are started.

6. The feedstock and reject valves are controlled manually using the DCS operator's console until the starting process conditions, with zero accept flow, are met.

30
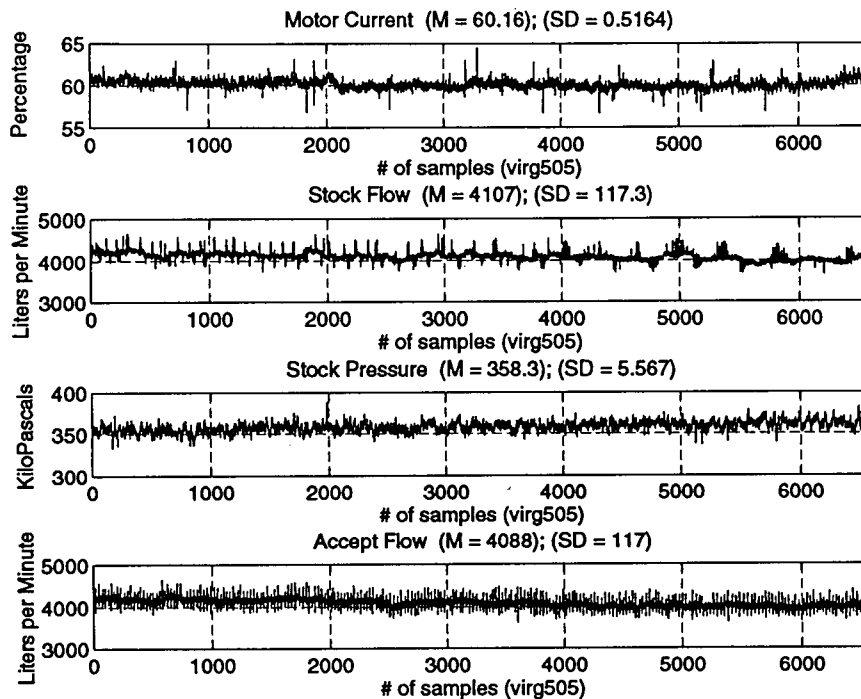
Figure 3.11: Typical Fast Sampled Data. Unfiltered

7. Slow data sampling is started.

8. The accept valve is opened to reach the first selected accept flow rate and this setting is kept until steady state is achieved. Fast sampling data is collected by the PC for periods of up to thirty seconds.

9. If failure is not present, the valves are manipulated to achieve the next higher accept flow rate previously selected, and step 7 is repeated. If failure occurs, the settings remain untouched until the differential pressure between the feedstock and the accept lines triggers a system shut–down mechanism. Fast sampling data is then gathered at equally spaced accept flow rates until the whole operation comes to a full stop.

The behavior of the accept flow when compared with the opening of the accept valve during a typical fault trial can be seen in Figure 3.14. During the whole operation, the data gathered by the DCS is downloaded to the PC. As mentioned, this information has a lower sampling rate and is transmitted through one of the DCS asynchronous ports to the personal computer.
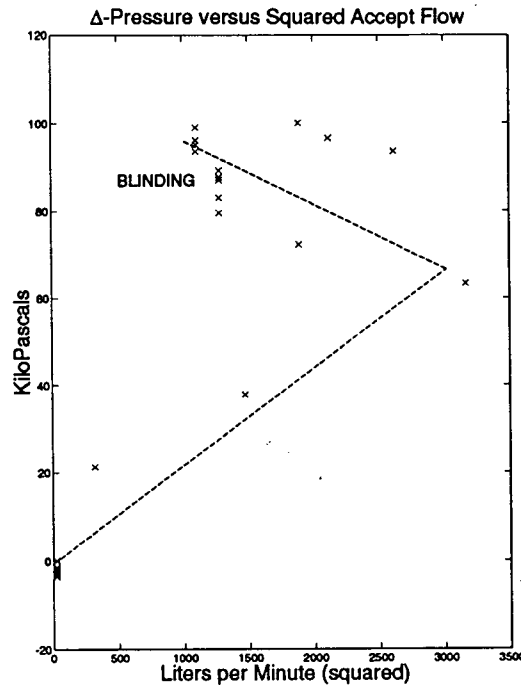
Figure 3.12: Screen Operated with Rotor Off

3.2 The Motor Load

It has been determined experimentally that velocity, acceleration, and convergence have a strong influence on whether the fibres inside a screen will pass through the screen plate or not [30]. Aside from the flows and pressures, one of the factors which strongly affect the velocity of the pulp are the pulsations imparted by the rotor. Therefore, when thinking of alternatives to replace the current method of fault detection, one variable which seems indicated for closer analysis is the screen motor load. In this research, that variable was measured using an amperage transformer calibrated for phase correction.

The influence of the rotor on the screen operation was demonstrated during the course of several experimental trials. When the screen is operated without turning on the rotor, its throughput is dramatically reduced, as can be seen in Figure 3.12, and failure occurs quite rapidly. The screen shows no sign of malfunction when the accept flow rate remains below 50 L/Min (or 2500 $[L/Min]^2$). If the accept valve is opened further, the flow gets close to 65 L/Min and then the pressure differential
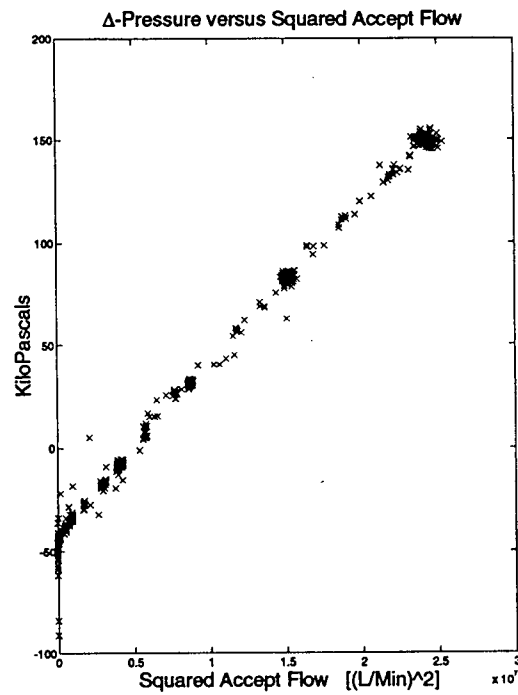
Figure 3.13: Screen Operated with Rotor On
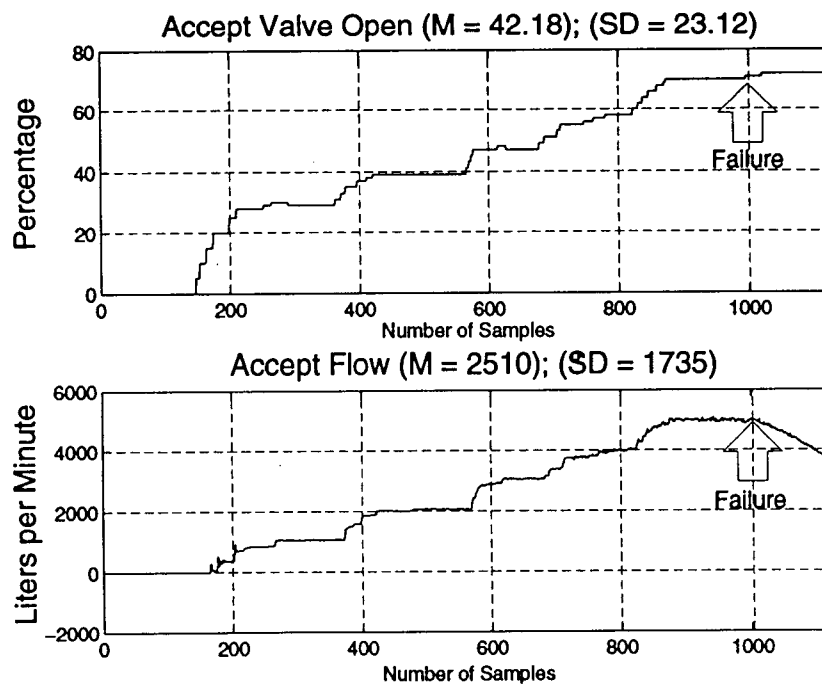
Figure 3.14: Typical Experimental Trial. Accept Flow versus Accept Valve Opening

Scr. Motor Current (M = 60.27); (SD = 3.418)

Failure

Accept Flow (M = 2510); (SD = 1735)
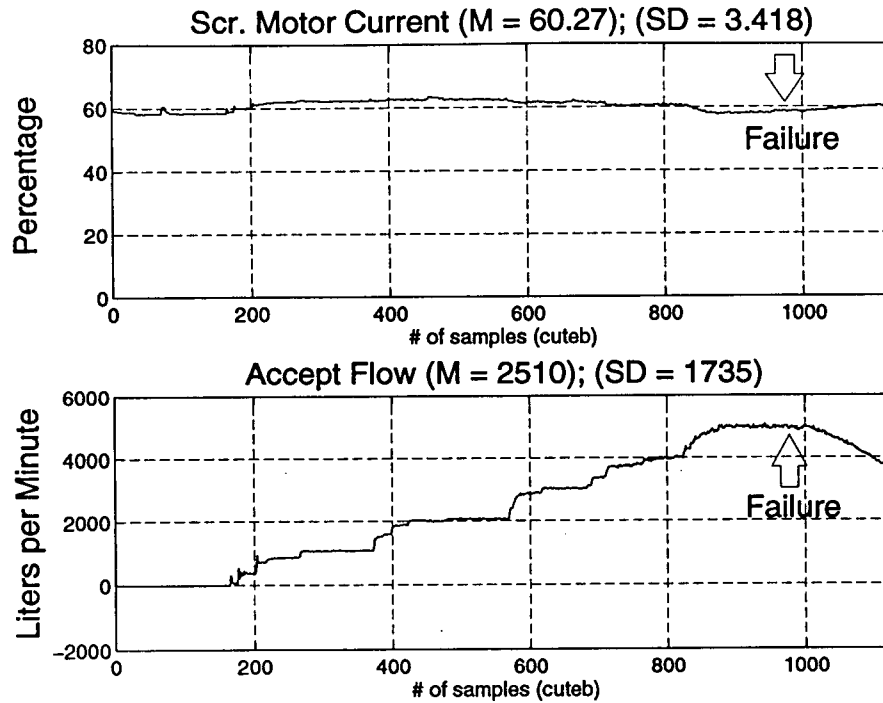
Failure

Figure 3.15: Screen Motor Load. Example 1

Scr. Motor Current (M = 59.33); (SD = 7.171)

Failure

Accept Flow (M = 4261); (SD = 709)
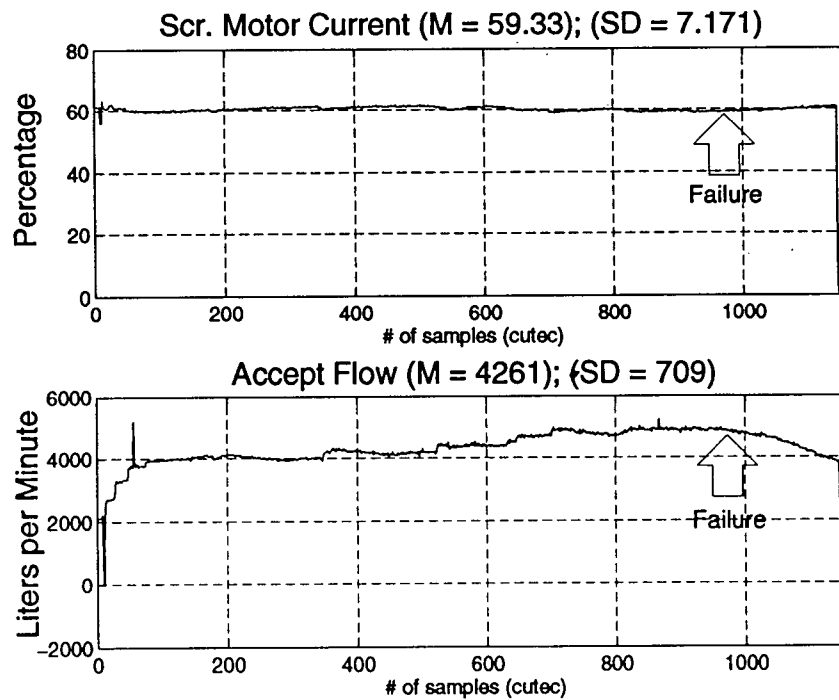
Failure

Figure 3.16: Screen Motor Load. Example 2

34
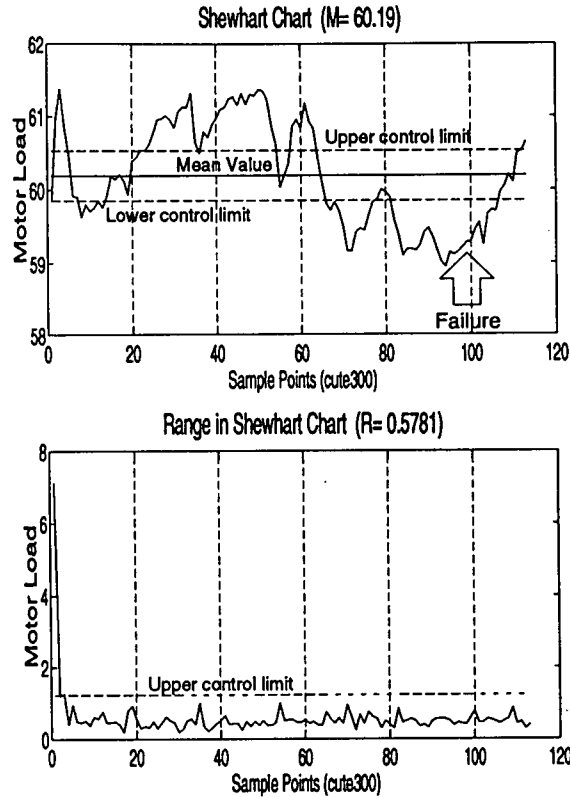
Figure 3.17: Shewhart Chart with $A_2 \hat{R}$ limits

suddenly starts to increase and the accept flow diminishes: clear signs of failure due to plugging. The same screen can function without problems with flows above 5,000 L/Min (or 2.5 x 10^7 [L/Min]^2) when operated with a fully operating rotor. See Figure 3.13.

Screen motor load was also mentioned by people from industry as a variable which would be looked at by operators, to guarantee the absence of malfunctions.

Despite the indubitable influence that the rotor has in the prevention of screen plate blinding, in all tests conducted in which the screen was made to fail, the screen motor load did not show any significant trends or signal levels which would help to account for the presence (or absence) of a failure, let alone of a fault.

As can be seen in Figures 3.15 and 3.16, when the throughput of the screen is increased, the motor load signal does not register any dramatic changes, even when the behavior of the differential
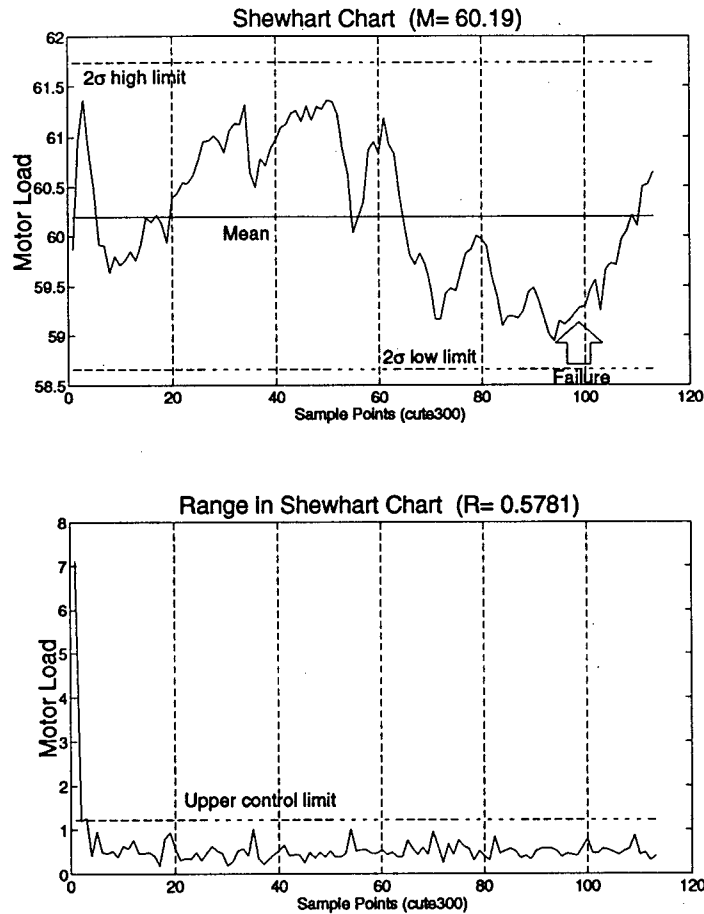
Figure 3.18: Shewhart Chart with 2σ limits

pressure and the accept flow tell us that failure has already occurred. If we prepare control charts based on the motor load data sampled in the tests, the results are equally poor. As seen in Figure 3.17, when the range and the arithmetic mean of the signals are used to create the upper and lower statistical control limits, many false alarms appear in the chart but the appearance of a true fault is not detected. When the common value of twice the standard deviation is used to set the upper and lower control limits (see Figure 3.18), no point falls out of statistical control. Clearly, both alternatives render the chart useless.

Judging from all the results obtained in the pilot plant trials, it does not look as if direct use of the screen motor load signal, or any statistical manipulation of it, could be a better option for fault
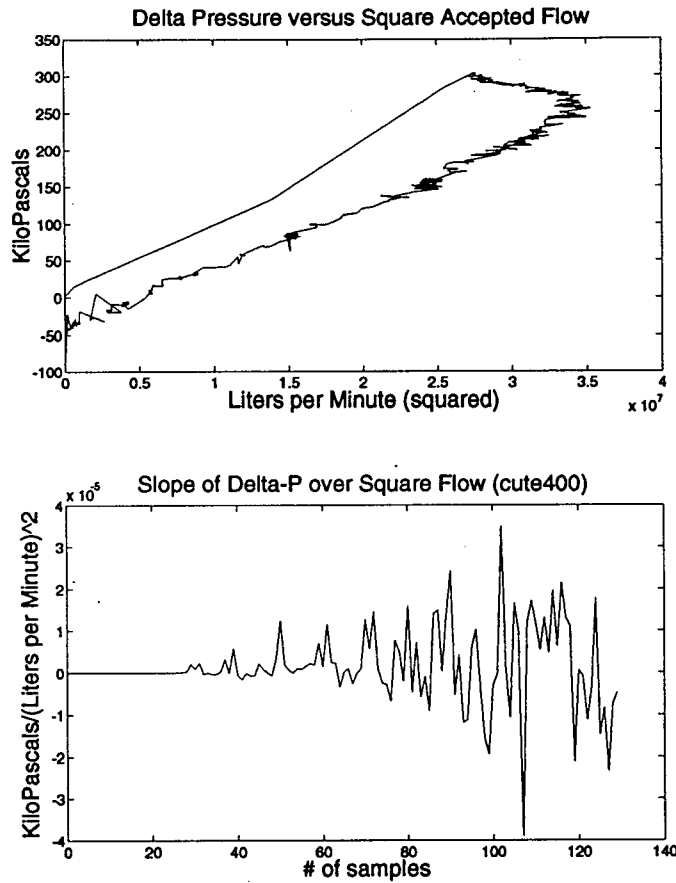
Figure 3.19: Differential Pressure versus Square of Accepts Flow (Raw Data)

detection than the current high differential pressure scheme used in industry. Quite to the contrary. If the latter can be deemed as imprecise, the former seems to be of no help at all.

3.3 Square of Flow versus Differential Pressure

It is a proven fact that screen blinding is determined by the decrease of flow through the accept line while the pressure differential between the feed line and the accept line increases [21]. The current approach to fault detection focuses only on the increase in pressure differential, paying no attention to what happens to the flow. Therefore, a logical next step would be to find an index or indicator which would correct for that omission. The first candidate that comes to mind is pointed by the hydrodynamic theory: the plot of one against the square of the other [22].

37
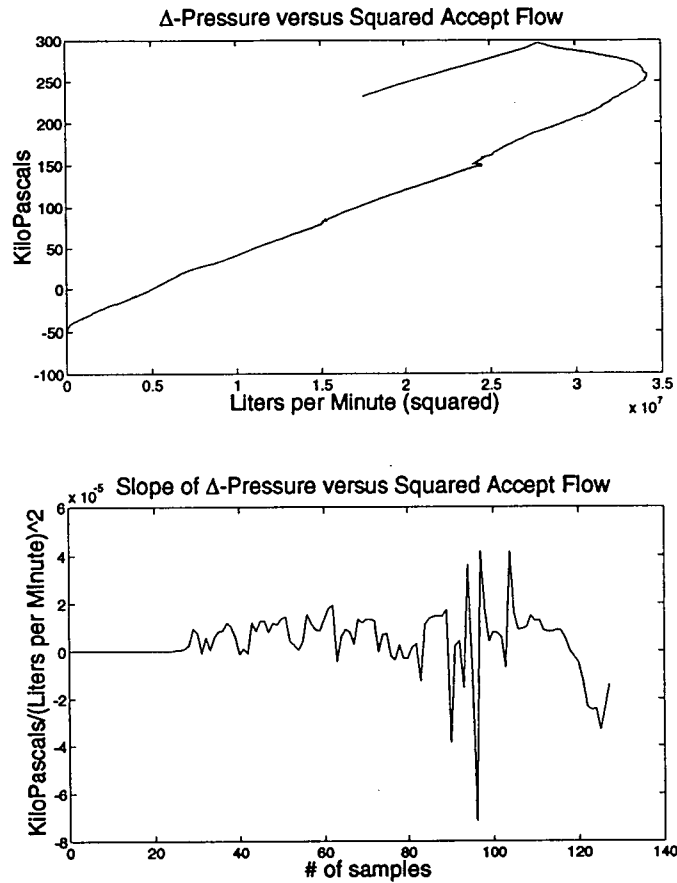
Figure 3.20: Differential Pressure versus Square of Accepts Flow (Filtered Data)

While running the screens to failure in the pilot plant trials, one feature immediately became obvious: when plotting the relationship between the square of the accept flow and the pressure differential, the slope of the graph would change dramatically depending on whether the screen was operating normally or was deemed as having failed. A screen operating normally would have a positive slope, while a faulty screen would have a negative one. The slope of the graphic, then, appeared as a promising candidate for a fault indicator.

The upper part of Figures 3.19 and 3.20 shows that, when taken as a batch, the trial data concerning the slope reveals unequivocally whether a fault has occurred or not. FDI-wise, this is a very important finding. Nonetheless, the information does not establish any time references. This means that the failure, though clearly present, might have occurred at any time during the test. To

38

overcome this deficiency, instead of gathering data and plotting curves at the end of a trial, one could think of looking at the data while it is being collected, say once every minute. This would give the operator some sense of timing. Obviously, if implemented as a cumulative algorithm developing in time, as demonstrated by Figures 3.21 through 3.22, such information is better than the current method of fault detection, as it would allow the plant operator to know if a high differential pressure alarm is due to blinding or to other factors. One big weakness remains: the information would be post-mortem. Faulty conditions would not be identified in their early stages but after a failure has been declared.
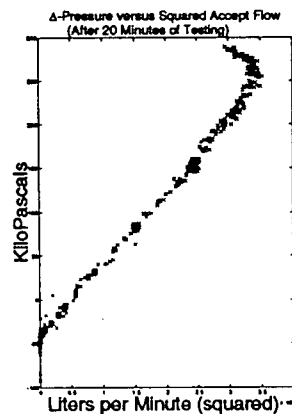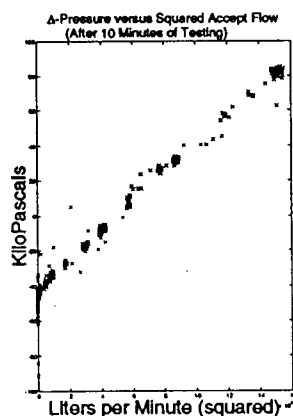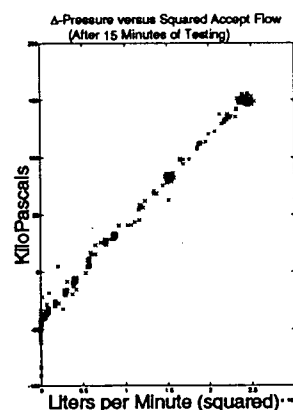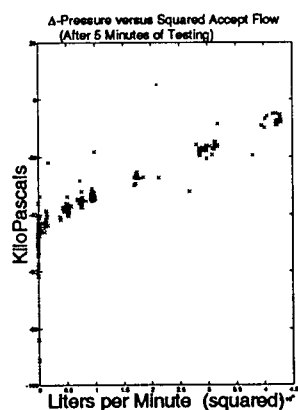
Figure 3.21: Cumulative slope index. Parts 1 and 2

Figure 3.22: Cumulative slope index. Parts 3 and 4

Going further, one could try to examine the slope of the graphic recursively, adding every new value which is acquired by the process sampling system. The lower part of Figures 3.19 and 3.20 shows that such approach is not practical for industrial use. Due to the presence of process noise, even if the data is heavily filtered, the calculated value of the slope oscillates in such a way that the index is rendered useless. If the average of several points is examined, the presence of a faulty condition could be determined, but being forced to average many points in time will bring us back to square one: such fault would be declared after it has become a failure. That would not constitute true FDI.

The problems caused by the presence of noise become more acute if we sample the data faster. As it can be seen in Figures 3.23 and 3.24, faster sampling not only does not help in making the slope index more reliable (lower graph) but it even disables the possibility of looking at a clear graphic of the slope. The batch data (upper graph) now looks as an amorphous blob from which no conclusions can be drawn, no matter how heavily we filter it, and the recursive slope index is still inconclusive, except if used for averaging.

Despite the lack of promise shown by the slope of the pressure differential versus the square of the accepts flow as a true fault detection index, two useful conclusions were obtained from the screen trials:

1. The slope of the differential pressure versus squared accept flow graphic is an improvement over the "classical" control approach used in mills to determine if a true failure is present.

2. When filtering screen process data, averaging produces smoother results for the variables than low pass filtering.

3.4 Heuristic Rules

An alternative for detecting faults on a plant which model is not known is to use all the existing knowledge of the system to extract useful conclusions. First, one has to look at the process conditions which can be associated with the normal and abnormal operation of the plant. The second step is to choose a symbolic representation of those conditions. Then, some inferential or heuristic rules are laid out. These rules, made by comparisons and deductions, will help determine whether a fault can

40
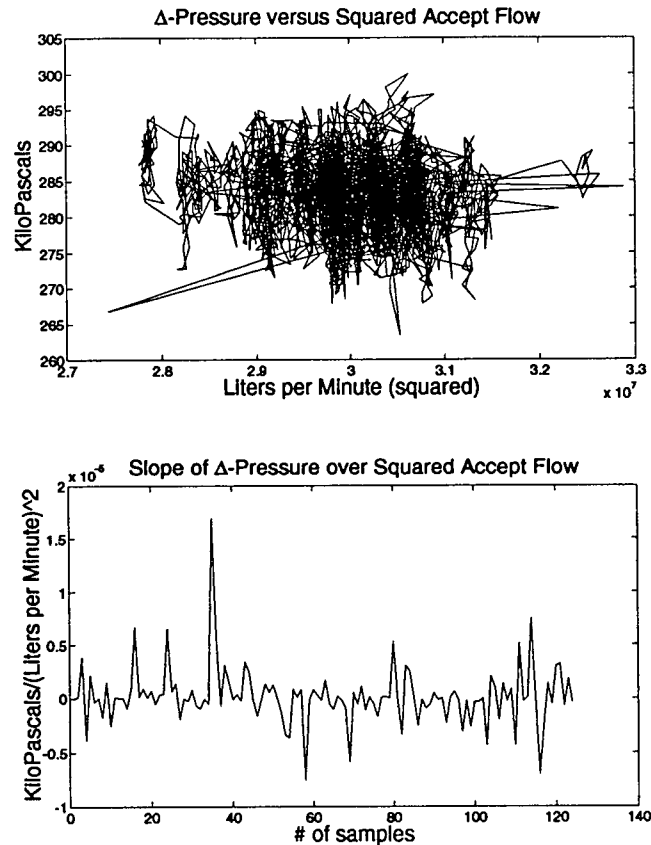
Figure 3.23: Differential Pressure versus Square of Accepts Flow (Raw Data)

be deemed as present or not. Finally, some graphic means to display the reasoning and the results must be designed [65].

Based on the existing literature on screens and on our findings, a fault detection mechanism for the screens based on heuristic rules does not seem to hold much promise. So far, there is no evidence of any observable variable or set of variables which can be used to make educated assumptions on whether a fault is present or not. Motor load, flows, pressures, etc., do not offer much insight about incipient screen malfunctions. What can be established without doubt is the presence of failure. That, in itself, is an advance over the "classical" screen control method commonly used in mills, but does not represent a true FDI method.

All in all, it does not seem that developing the next steps for a rule-based FDI system will add
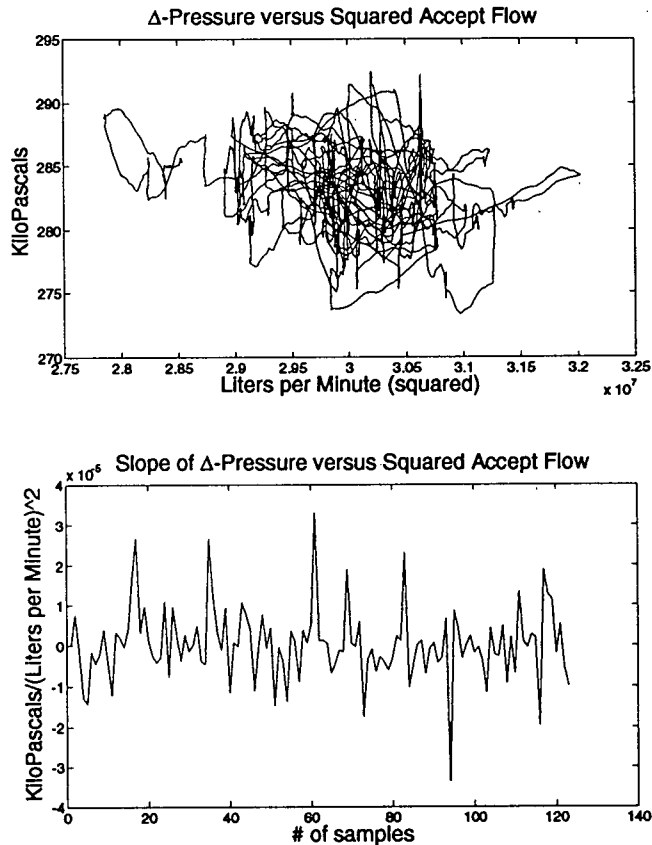
Figure 3.24: Differential Pressure versus Square of Accepts Flow (Filtered Data)

anything useful. Most important, by setting an expert system to deal with screen faults not much is gained over the use of the differential pressure versus squared accept flow method. The expense in computing memory and programming is, therefore, obviously unwarranted. A different approach is in order.

3.5 Summary

In this chapter the experimental setting and the results obtained when testing the non-model based FDI approaches on pressure screens have been presented. After describing the pilot plant used throughout the different experiments, a detailed explanation of a typical fault trial has been given. The rationale used in choosing several signals (motor load, and squared accept flow versus

delta pressure) for the application of some non-model based approaches has been explained and the results obtained with such techniques have been shown. Using those results, the limitations of the non-model based FDI techniques on pressure screens have been established and the need for a more sophisticated model-based approach has been justified.

Chapter 4

Model-Based Fault Detection of Screens

4.1 The Need for a Model

As it was mentioned in Chapter 2, from a systems identification point of view it is possible to distinguish two classes of FDI systems: model-based and non-model-based. The main difference is that the former relies on the existence of a mathematical model for the plant, while the latter does not.

With the exception of knowledge-based systems, FDI methods which do not require a model tend to be simpler than those in which a model is required. Their simplicity, however, does not come for free. Usually, the information they provide does not discriminate between possible causes for the faults, or leads to very conservative detection thresholds prone to produce false alarms [19], or does not detect the faults fast enough [33]. The non-model-based approaches attempted in the present research have shown some of these drawbacks, and that made them unsuitable for true fault detection. Due to limitations in the knowledge of the screen fault signatures, diagnostic methods built around concepts of artificial intelligence did not hold much promise, either. Therefore, a strategy based on developing a model of the screens seemed indicated.

In general, a mathematical model for the process or plant provides the basis for obtaining better information towards fault detection. Measurable coefficients, also called features, become available with the model. Changes in these features with regard to their normal values can be used to detect fault symptoms and to identify the cause of the faults. This leads to early detection of incipient malfunctions and to a deeper level of diagnosis. The price to pay for these improvements is the development of the model. In this research, basic physical laws and other well-established relationships were explored and used to postulate the pressure screen model. The end result can be defined as of the type known as a grey box [45].

44

4.2 Modelling of a Pressure Screen

4.2.1 Pulp Screen Mass Balances

As the process that takes place inside a screen is non-reactive, three material balances can be written: one for the water in the slurry, another for the pulp and a third for both components [16]. The first two will be expressed as:

$$\frac{d(m_{H_2O})}{dt} = \dot{M}_{F_{H_2O}} + \dot{M}_{\text{dilution}} - \dot{M}_{A_{H_2O}} - \dot{M}_{R_{H_2O}}$$

$$\frac{d(m_{sol})}{dt} = \dot{M}_{F_{sol}} - \dot{M}_{A_{sol}} - \dot{M}_{R_{sol}}$$

$$(4.47)$$

where the subscripts refer to water (H_2O), solids in the pulp slurry (sol), *incoming feedstock* (F), *reject* (R), and *accept* (A) mass flows. In industrial environments total volumetric flows, rather than individual mass flows, are commonly measured. Therefore, in order to determine the mass of both phases in the slurry, *consistency* measurements and volume-to-mass equivalences are required.

Consistency (C) was defined earlier as the weight of oven-dry fibre in 100g of the pulp-water mixture [48]. However, this definition introduces the problem of determining the amount of non-fibrous additives in the stock. To overcome this difficulty the definition of *consistency* is usually modified to mean weight of solids in 100g of the pulp-water stock.

On the other hand, the relationship between total volumetric flow F and individual mass flows $\dot{M}$ can be obtained from the relationship between volume (V), mass (M) and density ρ

$$\rho V_{\text{Total}} = M_{H_2O} + M_{\text{solids}} \qquad (4.48)$$

Since

$$M_{\text{solids}} = C\left(M_{\text{solids}} + M_{H_2O}\right) \qquad (4.49)$$

then

$$M_{\text{solids}} = \rho C V_{\text{Total}} \qquad (4.50)$$

and

$$M_{H_2O} = \rho(1 - C)V_{\text{Total}} \qquad (4.51)$$

Assuming that neither the consistencies nor the densities change too fast with time, the balances become

$$\frac{d(V_{H_2O})}{dt} = \frac{1}{\rho_{H_2O}}[a_1 F_F - a_2 F_A - a_3 F_R] + F_{\text{dilution}}$$

$$\frac{d(V_{sol})}{dt} = \frac{1}{\rho_{sol}}[a_4 F_F - a_5 F_A - a_6 F_R]$$

(4.52)

where

$$a_1 = \rho_F(1 - C_F)$$

$$a_2 = \rho_A(1 - C_A)$$

$$a_3 = \rho_R(1 - C_R)$$

(4.53)

$$a_4 = \rho_F C_F$$

$$a_5 = \rho_A C_A$$

$$a_6 = \rho_R C_R$$

Due to the nature of the components, all densities are approximately equal. Now, because the volume of the screen is constant and water is an incompressible fluid, fibres can only accumulate at the expense of water, that is,

$$\frac{d(V_{\text{Total}})}{dt} = \frac{d(V_{H_2O})}{dt} + \frac{d(V_{sol})}{dt} = 0$$

(4.54)

Adding both expressions for the mass accumulated in the system, we obtain the final mass balance equation

$$\frac{d(V_{\text{Total}})}{dt} = \frac{1}{\rho}\left(F_F - F_A - F_R + F_{\text{dilution}}\right)$$

(4.55)

and, recalling that the Volume (V) inside the screen does not change,

$$0 = F_F - F_A - F_R + F_{\text{dilution}}$$

(4.56)

4.2.2 Pulp Screen Energy Balances

As the screen is an open system where no phase changes or reactions occur, the starting point for an energy balance is the expression

$$\frac{d[U + E_K + E_P]}{dt} = \sum_{\text{input streams}} \dot{M_i}\left[\frac{v_i^2}{2g_c} + \frac{g}{g_c}h_i + \underline{U_i}\right] -$$

$$\sum_{\text{output streams}} \dot{M_o}\left[\frac{v_o^2}{2g_c} + \frac{g}{g_c}h_o + \underline{U_o}\right] + Q + W$$

(4.57)

[27] with

$$U : \text{Internal Energy}$$

$$\underline{U} : \text{Specific Internal Energy}$$

$$E_K : \text{Kinetic Energy}$$

$$E_P : \text{Potential Energy}$$

$$V : \text{Volume}$$

$$v : \text{Velocity}$$

$$g : \text{Gravity} \tag{4.58}$$

$$h : \text{Height}$$

$$g_c : \text{Conversion factor}$$

$$W : \text{Work}$$

$$Q : \text{Heat}$$

$$\dot{M} : \text{Mass Flow}$$

$$T : \text{Temperature}$$

The work done on the system can be divided into hydraulic ($W_h = PV$) and shaft work (W_s). Since the process is adiabatic, the equation becomes

$$\frac{d[\text{Energy}]}{dt} = \sum_{\text{input streams}} \dot{M}_i \left[\frac{v_i^2}{2g_c} + \frac{g}{g_c} h_i + \underline{U}_i + \frac{P_i}{\rho} \right] - \\ \sum_{\text{output streams}} \dot{M}_o \left[\frac{v_o^2}{2g_c} + \frac{g}{g_c} h_o + \underline{U}_o + \frac{P_o}{\rho} \right] + W_s \tag{4.59}$$

leading to the non-linear expression

$$\frac{d[\text{Energy}]}{dt} = e_1 F_F^3 + e_2 F_F + F_F P_F + e_3 F_{\text{Dil}}^3 + e_4 F_{\text{Dil}} + F_{\text{Dil}} P_{\text{Dil}} \\ -e_5 F_A^3 P_A - e_6 F_A - F_A P_A - e_7 F_R^3 - F_R P_R + W_s + k_A + k_R \tag{4.60}$$

where k refers to the frictional losses inside the screen [11] and, again, the subscripts refer to *incoming feedstock* (F), *rejects* (R), *accepts* (A), and *dilution water* (Dil). More details can be found in Appendix A where the coefficients e_i are defined. The above relationship, after applying some

47

simplifying

$$\frac{dT}{dt} = f_1 F_F^3 + f_2 F_F + f_3 F_F P_F + f_4 F_{\text{Dil}} + f_5 F_{\text{Dil}} + f_3 F_{\text{Dil}} P_{\text{Dil}}$$
$$-f_6 F_A^3 - f_7 F_A - f_3 F_A P_A - f_8 F_R^3 - f_3 F_R P_R + f_3 W_s + f_3 k_A + f_3 k_R$$

(4.61)

where the f_i's are coefficients defined in Appendix A. After linearization and the fact that $\frac{dT}{dt}$ varies very slowly when compared with the flows and pressures, this becomes

$$0 = g_1 F_F' + g_2 P_F' + g_3 F_{\text{Dil}}' + g_4 P_{\text{Dil}}'$$
$$-g_5 F_A' - g_6 P_A' - g_7 F_R' - g_8 P_R' + f_3 W_s' + f_3 k_A' + f_3 k_R'$$

(4.62)

with

$$g_1 = \frac{3\rho \bar{F}_F^2 + 2A_F^2 \rho g (h_F - h_R) + 2A_F^2 g_c \bar{P}_F}{2A_F^2 \rho g_c V C_v}$$

$$g_2 = \frac{3\bar{F}_{\text{Dil}}^2 + 2A_{\text{Dil}}^2 g (h_{\text{Dil}} - h_R) + 2A_{\text{Dil}}^2 g_c \bar{P}_{\text{Dil}}}{2A_{\text{Dil}}^2 \rho g_c V C_v}$$

$$g_3 = \frac{3\rho \bar{F}_A^2 + 2A_A^2 \rho g (h_A - h_R) + 2A_A^2 g_c \bar{P}_A}{2A_A^2 \rho g_c V C_v}$$

$$g_4 = \frac{3\rho \bar{F}_R^2 + 2A_R^2 g_c \bar{P}_R}{2A_R^2 \rho g_c V C_v}$$

(4.63)

$$g_5 = \frac{\bar{F}_F}{\rho V C_v}$$

$$g_6 = \frac{\bar{F}_{\text{Dil}}}{\rho V C_v}$$

$$g_7 = \frac{\bar{F}_A}{\rho V C_v}$$

(4.64)

$$g_8 = \frac{\bar{F}_R}{\rho V C_v}$$

$$f_3 = \frac{1}{\rho V C_v}$$

and the P''s and F''s are deviation variables from the set points $\bar{P}'$ and $\bar{F}'$.

4.2.3 Mechanical Losses

From the mechanics of incompressible fluids, it is known that the energy losses associated with friction in pipelines are proportional to the square of the fluid velocity. This can be expressed as

$$k = f_p v^2$$

(4.65)

48

and it is the underlying principle of Fanning's and D'Arcy-Weisbachs's friction equations [9], [57]. Although f_p is normally a constant, it can be theorized that it will vary with the accumulation of fibres on the screen plate. Thus, considering both f_p and v as time-varying, we differentiate D'Arcy-Weisbach's equation, namely,

$$\frac{dk}{dt} = \frac{d\left(f_p v^2\right)}{dt}$$ (4.66)

We then obtain the following expression

$$\frac{dk}{dt} = v^2 \frac{df_p}{dt} + f_p \frac{d\left(v^2\right)}{dt}$$ (4.67)

which, when using the fact that $k = f_p v^2$ gives us the result

$$\frac{dk}{dt} = v^2 \frac{df_p}{dt} + \frac{k}{v^2} \frac{d\left(v^2\right)}{dt}$$ (4.68)

This shows that a dynamic expression for the losses inside the screen must include a term dependent on the square of the fluid velocity and another term dependent on the losses themselves. Following this line of reasoning, an expression is proposed to reflect the dynamic behavior of the losses inside the screen:

$$\dot{k} = d_1 k + d_2 v^2$$ (4.69)

The constant

$$d_1 = \frac{1}{v^2} \frac{d\left(v^2\right)}{dt}$$ (4.70)

should be fairly sensitive to variations in the speed and, by implication, to flow changes. As the flow becomes smaller one should expect to see it grow in magnitude, fairly rapidly. The constant d_2 would reflect the rate of change of the parameter f_p due to accumulation of fibres on the screen plate. If blinding starts, the dynamic characteristic of the formulation will reflect this phenomenon: the appearance of blinding increases the velocity of the flow and the frictional losses, which in turn increase blinding, and so on.

As the velocity across the screen plate orifices is very difficult to measure, it is proposed that the flows be used in the equation. It is also proposed that the same equation be applied to account

49
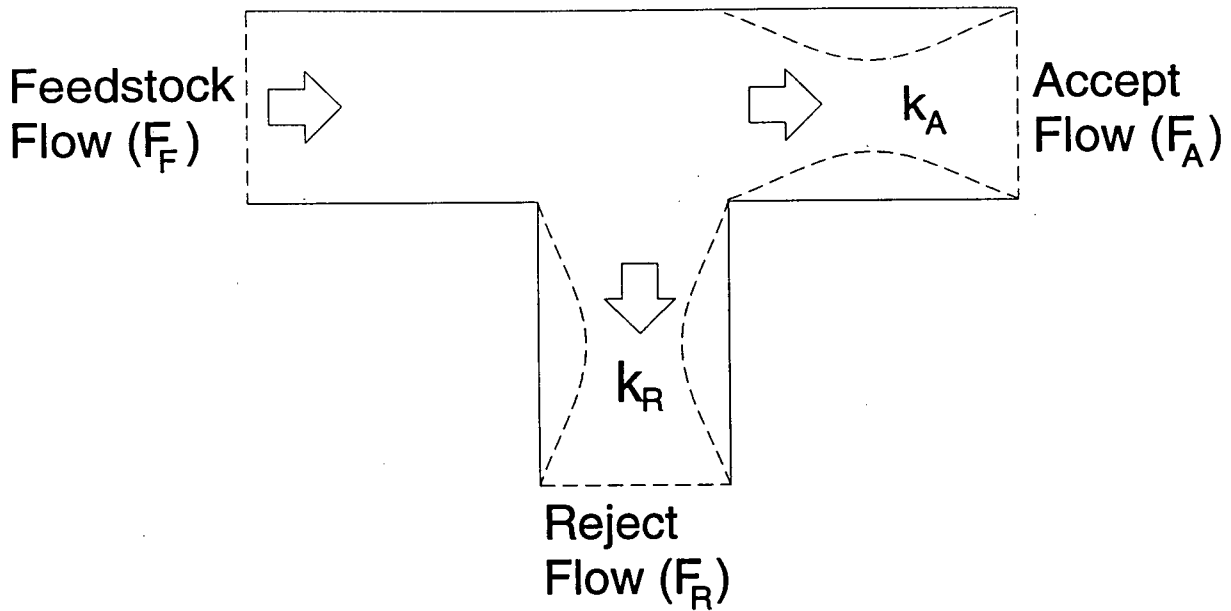
Figure 4.25: The screen as a dynamic "T"

for the two different streams of pulp inside the screen. This amounts to saying that the screen is considered as some sort of "T" pipe fitting with different dynamic friction coefficients: one set for the accepts and another for the rejects, as it is shown in Figure 4.25

Accordingly, after linearization and the fact that $F'_F = F'_A + F'_R - F'_{Dil}$, the dynamic expressions become

$$\frac{dk'_A}{dt} = d_1 k'_A + d_2 F'_F + d_2 F'_{Dil} - d_2 F'_R$$

$$\frac{dk'_R}{dt} = d_3 k'_R + d_4 F'_R$$

(4.71)

with the d's being the *dynamic loss coefficients*, the k's referring to the energy losses, and other measurable variables being as described before.

4.2.4 The Proposed Mathematical Model

Theoretically, the mass balances have the potential to give us information about the accumulation of fibres inside the screen. However, due to the state of the art of the existing measuring devices, attempts at determining the weight of fibres involved in blinding do not seem realistic. Another concern is that, due to the amount of fibres involved, such measurements might not convey enough information on the dynamics of this process. Therefore, the proposed mathematical model for screen

50

diagnosis relies directly on the balance of energy and the equation of losses for incompressible fluids and only indirectly on the mass balances. It is believed that this set of equations forms a reasonable basis towards a screen operational model. As the degrees of freedom allowed for the energy balance and the losses relations are equal to three, the model can be expressed in discrete state space form as

$$
\begin{bmatrix} zk_A' \\ zk_R' \end{bmatrix} = \begin{bmatrix} j_1 & 0 \\ 0 & j_2 \end{bmatrix} \begin{bmatrix} k_A' \\ k_R' \end{bmatrix} + \begin{bmatrix} j_3 & j_3 & j_3 \\ 0 & 0 & j_4 \end{bmatrix} \begin{bmatrix} F_F' \\ F_{Dil}' \\ F_R' \end{bmatrix} + \underline{0}
$$

$$
y = \begin{bmatrix} h_8 & h_8 \end{bmatrix} \begin{bmatrix} k_A' \\ k_R' \end{bmatrix} + \begin{bmatrix} h_1 & h_2 & -h_3 \end{bmatrix} \begin{bmatrix} F_F' \\ F_{Dil}' \\ F_R' \end{bmatrix} +
$$

$$
\begin{bmatrix} h_4 & h_5 & -h_6 & -h_7 & h_8 \end{bmatrix} \begin{bmatrix} P_F' \\ P_{Dil}' \\ P_A' \\ P_R' \\ W_s' \end{bmatrix}
$$

(4.72)

As mentioned before, the states k represent the energy losses and the output y the accept flow. The derivation of the parameters h and j is explained in Appendix B.

This system is <u>reachable</u>, i.e., by using a realizable sequence of control signals it is possible to drive it from an initial state to any final state[6], and <u>observable</u>, i.e., any state can be determined from a finite sequence of input and output signals[7]. That makes it ideal for identification. More details can be found in Appendix A.

4.2.5 Pulp Quality

Although research on the subject has advanced significantly in the last years [23], [17], modelling of the effects of screening on pulp quality can not yet be achieved through the use of mechanistic or physical principles. Results have to rely on empirical relations obtained through testing and experiments. The tools to use in search of empirical relationships are the *screening coefficient Q*, and the *debris reject efficiency E_r*, previously mentioned. In the proposed model, the quality aspect

---

[6] Provided the parameter $j_4 \neq 0$ (which seems true in all cases).

[7] Provided $j_1 \neq j_2$ (which is also true in the majority of cases)

of pulp screening has not been included. Nonetheless, the path chosen for identification (the *singular pencil model*, to be explained in the next sections) allows for easy incorporation in the model of future theoretical formulations which might seem pertinent.

4.3 Identifying the Proposed Model

4.3.1 What to Identify?

Two approaches have concentrated the attention in the literature on model-based fault diagnosis: process parameters and state variables. Although somewhat different, both share the same phases when implemented:

1. <u>Data processing.</u> Measured signals of the process are gathered and made suitable for fault detection.

2. <u>Fault Detection.</u> Some characteristic process features are extracted and compared with their expected values. Deviations are used as indications of fault occurrence.

3. <u>Identification.</u> A classification of faults is prepared and any subsequent fault is compared with this fault catalogue to determine type, location, size and probable cause.

The first two characterize the detection stage of an FDI system and the last one constitutes the diagnosis.

Fault detection procedures must be insensitive to any changes in the process not produced by faults, such as modelling errors, noise, changes in operating points, etc.[8], for if not, its practical use is severely compromised: no plant operator will trust a system plagued with false alarms.

Unfortunately, the process parameters and the state variables methods can not always reach decoupling between the effects of the unknown inputs and the effects of faults. Therefore, redundant methods capable of precise identification are attractive to the researcher. Based on the fact that process parameters and state variables share the same procedural stages, a strategy that combines both approaches is therefore proposed. To the best of my knowledge, this has not been attempted previously.

---

[8] These latter are termed *unknown inputs*.

4.3.2 How to Identify It?

Instead of selecting measurable states to calculate the model parameters, or assuming perfect knowledge of the system to calculate the system states, the dynamic model that has been proposed includes the two quantities as unknowns. Therefore, both the process parameters and the system states must be estimated.

The problem of combined estimation of parameters and states was originally treated as a nonlinear problem and the Extended Kalman filter was proposed to solve it [51], [3], [6]. However, although this method has been proven feasible, the estimates often diverge. Building a methodology for fault detection on the basis of an identification approach that does not guarantee convergence of the parameter estimates does not seem very sound. Instead, an approach based on using a special canonical model known as the *singular pencil* model is proposed. When the noise statistics are known, the singular pencil model allows to solve the simultaneous estimation of state variables and system parameters as an optimal linear filtering problem and its convergence properties are excellent [5].

4.4 The Singular Pencil Approach

4.4.1 From an ARMAX Model

Assume that we have an ARMAX model for a linear, discrete, multivariable, time-invariant system with input variable $u_k \epsilon R^m$, output variable $y_k \epsilon R^p$, and noise input $e_k \epsilon R^p$ assumed as a zero mean white Gaussian sequence with covariance $\Lambda (\Lambda > 0)$, such as:

$$A\left(z^{-1}\right)y_k = B\left(z^{-1}\right)u_k + C\left(z^{-1}\right)e_k \tag{4.73}$$

where z^{-1} is the *shifting* operator, and the matrices $A\left(z^{-1}\right)$, $B\left(z^{-1}\right)$, and $C\left(z^{-1}\right)$ are of the form

$$A\left(z^{-1}\right) = A_o + A_1 z^{-1} + \ldots A_v z^{-v}$$

$$B\left(z^{-1}\right) = B_o + B_1 z^{-1} + \ldots B_v z^{-v} \tag{4.74}$$

$$C\left(z^{-1}\right) = I + C_1 z^{-1} + \ldots C_v z^{-v}$$

v being the maximum of the degrees of the polynomials in $A\left(z^{-1}\right)$, $B\left(z^{-1}\right)$, and $C\left(z^{-1}\right)$. These matrices have q rows. Introducing auxiliary variables in a vector x_k with $n = vq$ we can represent

the ARMAX model as

$$
\begin{bmatrix} I_n \\ 0_{q \times n} \end{bmatrix} z x_k = \begin{bmatrix} 0_{q \times q} & & & 0 \\ I_q & & & \\ & \ddots & & \\ 0 & & & I_q \end{bmatrix} x_k - \begin{bmatrix} A_v \\ \vdots \\ A_1 \\ A_0 \end{bmatrix} y_k + \begin{bmatrix} B_v \\ \vdots \\ B_1 \\ B_0 \end{bmatrix} u_k + \begin{bmatrix} C_v \\ \vdots \\ C_1 \\ I_q \end{bmatrix} e_k \tag{4.75}
$$

with I_α and $0_{\alpha \times \beta}$ denoting an $\alpha \times \alpha$ identity and a $\alpha \times \beta$ null matrix, respectively. And $v = \max\{n_i\}$, where n_i is the largest polynomial degree of z^{-1} in row i of $[A(z^{-1}), B(z^{-1}), C(z^{-1})]$. For simplicity, this can be written as

$$
z x_k = E_* x_k - A_* y_k + B_* u_k + C_* e_k
$$
$$
0 = E_0 x_k - A_0 y_k + B_0 u_k + e_k \tag{4.76}
$$

Assuming that this representation is canonical or has been modified to be canonical [8], it can be put as

$$
P(z) \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = \begin{bmatrix} E_* - zI & -A_* & B_* & C_* \\ E_0 & -A_0 & B_0 & I \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = 0 \tag{4.77}
$$

with $P(z)$ a singular pencil of matrices and x_k satisfying the definition of the system state vector. This model is called a *singular pencil matrix* (SPM) model.

A pencil of matrices is a first-order polynomial in an indeterminate D with matrix coefficients, of the form $M+DN$, where M and N are $m \times n$ matrices. The pencil is called *singular* if M and N have different order or if the determinant $|M+DN|$ is equal to zero [67].

4.4.2 From a State Space Model

If we start modeling the singular pencil from a state-space model of the form

$$
x_{k+1} = A x_k + B u_k
$$
$$
y_k = C x_k + D u_k \tag{4.78}
$$

with the matrices $A(z)$, $B(z)$, $C(z)$, and $D(z)$ being $n \times n$, $n \times p$, $m \times n$, and $m \times p$, respectively, the implicit equation is obtainable by inspection as:

$$
\begin{bmatrix} I \\ 0 \end{bmatrix} z x_k = \begin{bmatrix} A \\ C \end{bmatrix} x_k + \begin{bmatrix} 0 & B \\ -I & D \end{bmatrix} \begin{bmatrix} y_k \\ u_k \end{bmatrix} \tag{4.79}
$$

where the I is our well known identity matrix. This initial expression has to be transformed into the implicit canonical form

$$P(z) \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = \begin{bmatrix} E_* - zI & -A_* & B_* & C_* \\ E_0 & -A_0 & B_0 & I \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = 0 \qquad (4.80)$$

to guarantee that the identification is unique[9]. That task is achieved by using linear transformations T_1, T_2, and T_3, as required, on the state-space expression and adding the matrices

$$C_* \text{ and } I \qquad (4.81)$$

which help account for the process noise [5]. Transformation T_1 is obtained by selecting a chained basis from the matrix

$$M = \begin{vmatrix} C(A)^n \\ \vdots \\ CA \\ C \end{vmatrix} \qquad (4.82)$$

and it leads to

$$\begin{bmatrix} T_1 A T_1^{-1} - zI & T_1 0 & T_1 B \\ C T_1^{-1} & -I & D \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ u_k \end{bmatrix} = \begin{bmatrix} E_*' - zI & -A_*' & B_*' \\ E_0' & -A_0' & B_0' \end{bmatrix} \begin{bmatrix} x_k' \\ y_k \\ u_k \end{bmatrix} = 0 \qquad (4.83)$$

Transformation T_2 is obtained from the form

$$T_2 = I_n + \begin{vmatrix} f(E_*'^{11}) & \cdots & f(E_*'^{1n}) \\ \vdots & \ddots & \vdots \\ f(E_*'^{n1}) & \cdots & f(E_*'^{nn}) \end{vmatrix} \qquad (4.84)$$

where E_*^{ii} is an n_i-square companion matrix and, given

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} \qquad (4.85)$$

---

[9]   Note that matrices A, A_*, and A_0; B, B_*, and B_0; and C and C_* are all different.

$f(A)$ is a matrix function that produces the following result:

$$f(A) = \begin{vmatrix} 0 & -a_{11} & -a_{12} & \cdots & -a_{1(n-1)} \\ 0 & 0 & -a_{11} & \cdots & -a_{1(n-2)} \\ 0 & 0 & 0 & \ddots & \\ \vdots & & & & \vdots \end{vmatrix} \tag{4.86}$$

and transformation T_3 is derived using matrices L and R such that

$$\begin{bmatrix} I & L \\ 0 & R \end{bmatrix} P'(z) = \begin{bmatrix} E_* - zI & G_* \\ E_0 & G_0 \end{bmatrix} \tag{4.87}$$

with

$$G_* = \begin{vmatrix} -A_* & B_* \end{vmatrix}$$
$$G_0 = \begin{vmatrix} -A_0 & B_0 \end{vmatrix} \tag{4.88}$$

[5]. After applying those three linear transformations in the order above, we obtain the reduced upper-right row echelon form required for the singular pencil, and the model becomes

$$P(z) \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = \begin{bmatrix} E_* - zI & G_* & C_* \\ E_0 & G_0 & I \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = 0 \tag{4.89}$$

with

$$E_* = \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix}$$
$$E_0 = \begin{vmatrix} 0 & 1 \end{vmatrix} \tag{4.90}$$

$$G_* = \begin{bmatrix} -j_1 j_2 & j_2(h_1 j_1 - h_8 j_3) & h_8(j_2 j_3 - j_1 j_4) - h_3 j_1 j_2 & j_2(h_2 j_1 - h_8 j_3) & 0 & \cdots & 0 \\ j_1 + j_2 & h_8 j_3 - h_1(j_1 + j_2) & h_8(j_4 - j_3) + h_3(j_1 + j_2) & h_8 j_3 - h_2(j_1 + j_2) & 0 & \cdots & 0 \end{bmatrix}$$
$$G_0 = \begin{bmatrix} -1 & h_1 & -h_3 & h_2 & h_4 & h_5 & -h_6 & -h_7 & h_8 \end{bmatrix} \tag{4.91}$$

and C_* to be chosen from previous knowledge of the process noise. More details and explanations can be found in Appendix B.

4.4.3 Singular Pencil Identification

Once the canonical representation is achieved, E_* and E_0 are determined uniquely. As y_k and u_k are measurable and therefore known, this means that the simultaneous state and parameter estimation problem can be solved as a linear filtering problem, in contrast to the nonlinear estimation problem encountered when using the state space model approach to attempt the simultaneous estimation of state and process parameters.

Making vectors $w_k^T = [y_k^T, u_k^T]$, $J = [I_p | 0_{p \times m}]$, and r (which contains the non-pivot parameters a_{ijk} and b_{ijk} in the matrices A_*, A_0, B_*, and B_0) our singular pencil equation can be expressed as

$$x_{k+1} = E_* x_k + \tilde{G}_*(w_k)r + C_* e_k$$
$$0 = E_0 x_k + \tilde{G}_0(w_k)r - J w_k + e_k$$

(4.92)

with

$$\tilde{G}_*(w_k)r = [-A_* | B_*] w_k$$
$$\tilde{G}_0(w_k)r = [-A_0 | B_0] w_k + J w_k$$

(4.93)

As these matrices are not unique, they must be built in a way that ensures that the elements of matrices A and B, i.e., a_{ijk} and b_{ijk}, are isolated into the column vector r. Since the system is time-invariant, $r = r_k = r_{k+1}$, and further refinements are possible in our representation of the system, namely

$$s_{k+1} = F_k s_k + D_* e_k$$
$$y_k = H_k s_k + e_k$$

(4.94)

with

$$s_k = \left[x_k^T, r_k^T\right]^T$$
$$H_k = \left[E_0 \tilde{G}_0(w_k)\right]$$
$$F_k = \begin{bmatrix} E_* & \tilde{G}_*(w_k) \\ 0 & I_l \end{bmatrix}$$
$$D_* = \begin{bmatrix} C_* \\ 0_{l \times n} \end{bmatrix}$$

(4.95)

and l is the dimension of the vector r.

This last expression is nothing but the representation of a linear stochastic system in state space form with state vector s. As such, if C_* and the covariance matrix Λ of the noise are known, it is

possible to use the Kalman filter algorithm to obtain the optimal linear estimate of the augmented state vector s_k, which in turn, will lead to the vectors x_k and r_k, rendering the states and the process parameters of the system.

The recursive equations to use will then be

$$\hat{s}_{k+1} = F_k \hat{s}_k + K_k(y_k - H_k \hat{s}_k)$$

$$K_k = (F_k P_k H_k^T + S)(F_k P_k H_k^T + \Lambda)^{-1} \tag{4.96}$$

$$P_{k+1} = F_k P_k F_k^T + Q - K_k(F_k P_k H_k^T + \Lambda)K_k^T$$

where

$$S = \begin{bmatrix} C_* \Lambda \\ 0_{l \times p} \end{bmatrix}$$

$$Q = \begin{bmatrix} C_* \Lambda C_*^T & 0_{n \times l} \\ 0_{l \times n} & 0_{l \times l} \end{bmatrix} \geq 0 \tag{4.97}$$

with initial conditions

$$\hat{s}_0 = E[s_0]$$

$$P_0 = E\left[(s_0 - E[s_0])(s_0 - E[s_0])^T\right] \geq 0 \tag{4.98}$$

and $E[(.)]$ is the expected value of $(.)$

Once the process parameters and state variables under normal operating conditions are obtained, they will be stored in memory. From this point, the physical coefficients in the model can be recovered, if necessary, applying reverse linear transformations T_1^{-1}, T_2^{-1} and T_3^{-1}. Successive on-line values can then be compared. The presence of swings or deviations will suggest the existence of faults, which lead us to the next stage of the FDI system: the diagnosis.

If C_* is not known (which is very likely in practice) there are several possible options, two of which are: the use of an extended Kalman filter algorithm to estimate simultaneously the noise and the measurement parameters, or the application of non-parametric noise estimation techniques to determine its characteristics, before attempting the singular pencil strategy. This problem will be addressed in the coming chapters.

4.4.4 Application of Equations

To offer a feel for the practical application of the proposed singular pencil approach for identification, let us look at some graphs. Figure 4.26 shows some input/output information obtained

from the pressure screen pilot plant trials conducted along this research. Although a detailed discussion of these trials will not be attempted until next chapter, a few things can be said here to illustrate the practical results obtained when using our mathematical derivations.
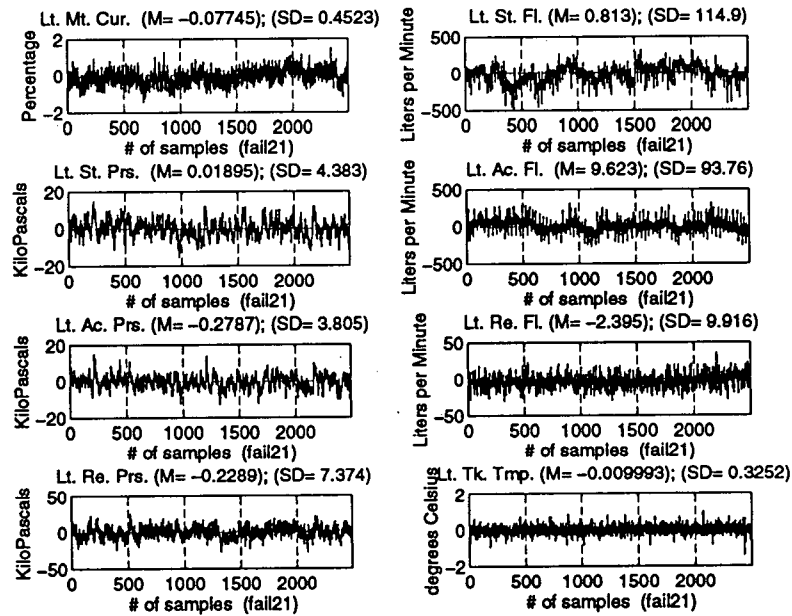
Figure 4.26: Simulated measurements of pressure screen

To begin with, the typical data batches shown in the graph correspond to the screen process variables measured by the sensors and then gathered by the fast sampling data acquisition system. This data clearly corresponds to deviation values and it is the data fed to the singular pencil algorithm. These measurements are then processed using the estimation sequence described in the previous sections to obtain the parameters shown in Figures 4.27 and 4.28. They are the parameters included in the augmented states vector s_k.

As our initial discrete screen model includes two states, the first two parameters in s_k (seen on top in graph 4.27) correspond to those states x_k. One could use them directly for diagnosis or apply similarity transformations to them to recover the original states representing the friction losses (k_A and k_R). Either way, their shape is not expected to change by much. At this point it is worth underscoring one important fact: as the graphs reveal, in general, convergence of the estimation is

Acc. fric. m= 0.0002437 sd= 0.006253

Rej. fric. m= 0.0004837 sd= 0.01608

coef. l11 m= -0.001578 sd= 0.0003794

coef. l12 m= 0.0006228 sd= 0.0005197

coef. l13 m= 0.0022 sd= 0.0008495

coef. l21 m= -0.03088 sd= 0.01239

coef. l22 m= 0.01442 sd= 0.005503

coef. l23 m= 0.0453 sd= 0.01788
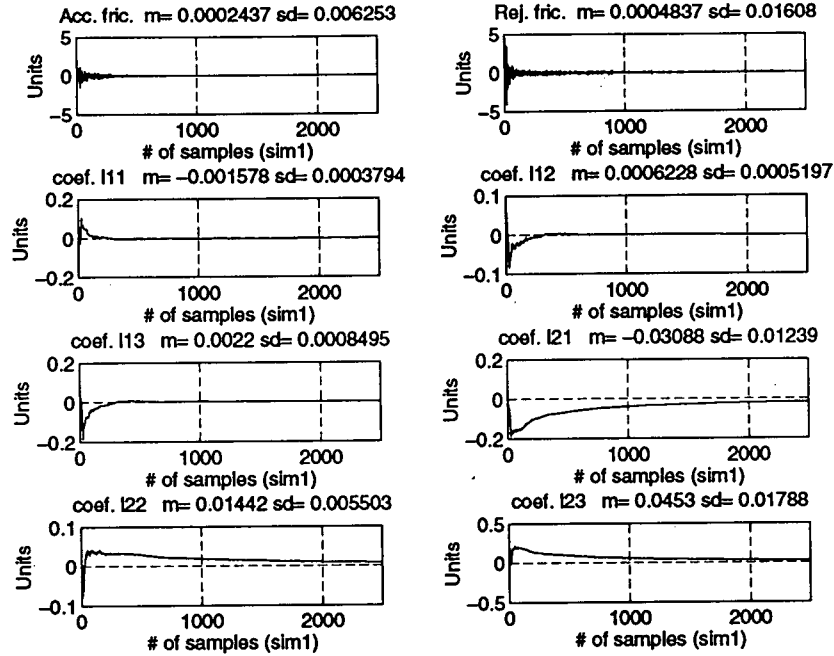
Figure 4.27: Estimated parameters in the augmented state vector

coef. h1 m= 0.9793 sd= 0.009581

coef. h3 m= -0.9027 sd= 0.04067

coef. h4 m= -0.008663 sd= 0.002494

coef. h6 m= 0.002984 sd= 0.00145

coef. h7 m= -0.003191 sd= 0.002882

coef. h8 m= -0.004283 sd= 0.000685
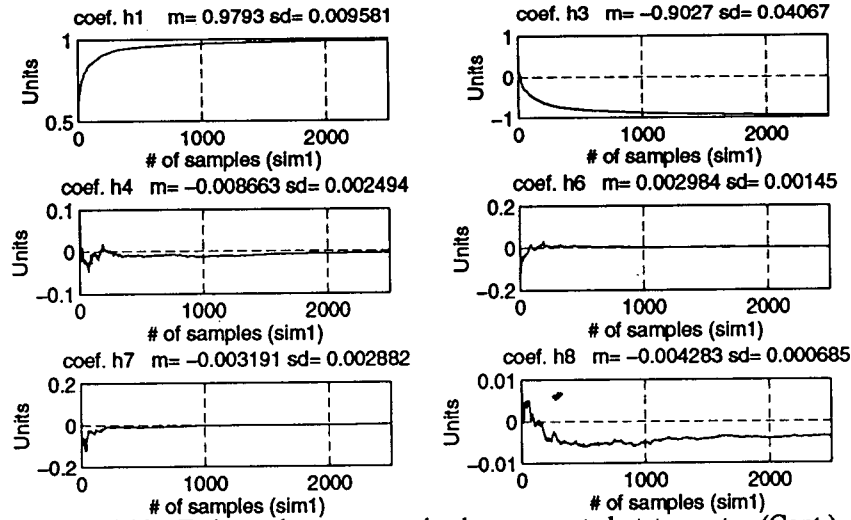
Figure 4.28: Estimated parameters in the augmented state vector (Cont.)

good. This comes as no surprise because the convergence properties of the SPM method are well documented in the literature [8].

The remaining parameters in the vector have to be transformed back by using the T_1^{-1}, T_2^{-1} and T_3^{-1} linear transformations. When doing this, they render the d coefficients in the discrete model which are associated with the friction losses inside the screen, and which are

60
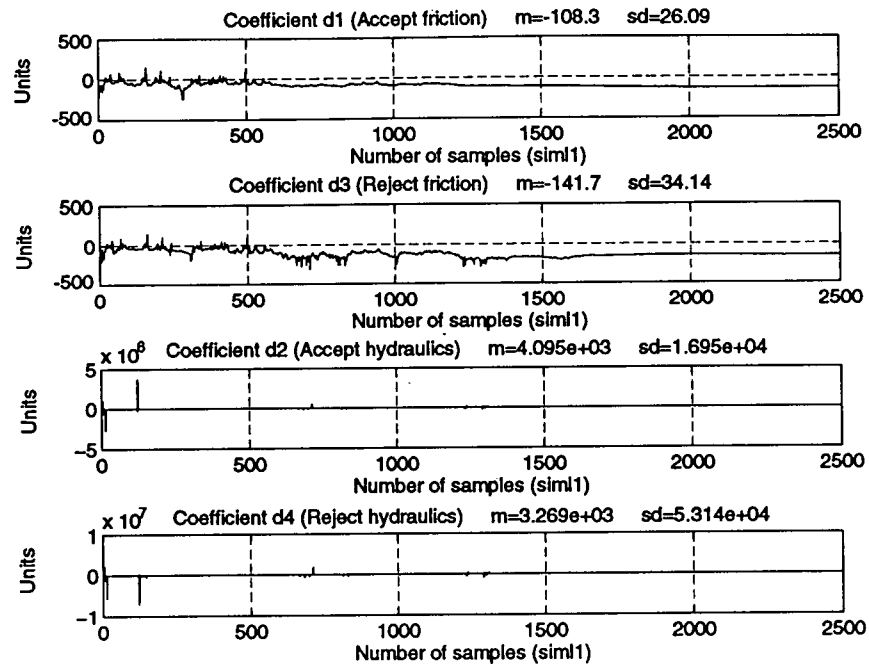
Figure 4.29: Estimated friction coefficients in the screen model

of interest for the fault detection process due to their physical meaning. The typical estimated d coefficients are shown in Figure 4.29. As expected, after convergence is achieved, the accept and reject friction coefficients (d_1 and d_3) are negative.

With this information and proper judgement, one can attempt to detect the incipient operational problems most commonly affecting industrial pressure screens, as it will be demonstrated in subsequent chapters.

4.5 Summary

In this chapter a mathematical model of the typical pressure screen and a procedure for its use as an FDI method have been presented. The model is of the grey-box type and it has been developed using mechanistic principles and an expression analogous to the D'Arcy-Weisbach formula to account for the dynamics inside the screen. The problem of identifying the parameters and states in the proposed model and its non-linear nature has been explored and an alternative for breaking this

61

non-linearity has been proposed. The technique, called the Singular Pencil Matrix (SPM) has been presented, and its implementation for use on pressure screens has been covered in detail.

Chapter 5

Testing and Experimental Results

5.1 Model Validation

5.1.1 FDI and model checking

Having built a tentative model for pulp pressure screens and suggested a suitable identification procedure, it is then appropriate to ask if the model is adequate for fault detection and isolation and, if not, what kind of inadequacies it might have. This is known as the model validation or model checking problem [45].

Most methods for model validation found in the literature on system identification rely on statistical lack of fit tests applied to the model errors or residuals. Model-based FDI is also built around changes in the model parameters or residuals, thus, it may be seriously affected by modelling errors [31]. In fact, when identifying for FDI, the main question that has to be answered when detecting lack of fit in the model is the following: is the deviation due to a model inadequacy or is it the signature of a fault? That is the problem the "robust" fault detection techniques try to solve. Unfortunately, it has no complete solution in the literature [19]. Moreover, the evidence suggests that it is not possible to improve simultaneously the fault sensitivity and the immunity to model errors of any FDI technique.

In practice, despite any formal aspects which might be raised, the main criterion for model validation must be whether the proposed model is good enough for the purpose at hand, or not [45]. Therefore, the paramount concern in this research is to determine the "goodness" of the proposed model in detecting incipient pressure screen faults. Or, in other words, its ability to display parameter deviations in the presence of faults in such a way that fault diagnosis is made possible. All other issues must be subordinated to that one.

63

5.1.2 The validation procedure

In the singular pencil model proposed for estimation in the previous chapter, the matrix C_* has to be fixed according to the *a priori* knowledge of the noise. In practice, this means that testing the statistical lack of fitness of the model helps to determine the choice of C_* and, in turn, the choice of C_* affects the statistical fit of the model, but this selection muddies the question of the model appropriateness from a fault-wise perspective. Does a better fit guarantee good fault discrimination? Not necessarily. In our case, perhaps good fit might simply indicate that the choice of C_* is masking the presence of a fault. Therefore, instead of using the standard validation techniques and simply testing the statistical properties of the proposed model (including the matrix C_*), one has to address the issue of model "goodness" for fault detection from a broader perspective. In doing so, some questions have to be asked: Do the estimated model coefficients show similar values under similar operating conditions? Are the values of the estimated coefficients different for different conditions? How far between them are the values corresponding to normal and faulty operation?

To obtain the required answers while taking all important facts into consideration, a three step validation procedure for the model-based scheme has been proposed:

1. Determination of parameters repeatability through simulation.

2. Evaluation of FDI-"goodness" in pilot plant trials.

3. Evaluation of statistical fit of the FDI-"good" models.

Thus, the singular pencil model approach is to be tested with simulated data first and, once its ability to offer repeatable results for different operating conditions is established (and, therefore, its suitability for the task at hand), pilot plant trials are to be used to assess its FDI-"goodness" and statistical fit.

5.2 Simulations and Repeatability

In instrumentation, repeatability is defined as the *"closeness of agreement among a number of consecutive measurements of the output for the same value of the input under the same operating conditions, approaching from the same direction, for full range traverses"* [1]. In the process plant, a very repeatable but not too accurate instrument is often better than a more accurate one with poorer

ACCURATE SHOOTING

Normal   Faulty
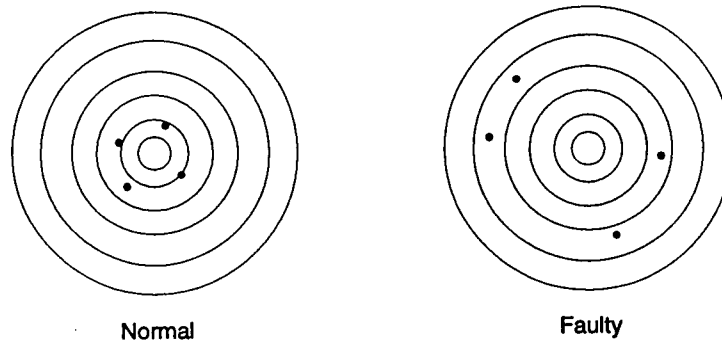
REPEATABLE SHOOTING

Normal   Faulty
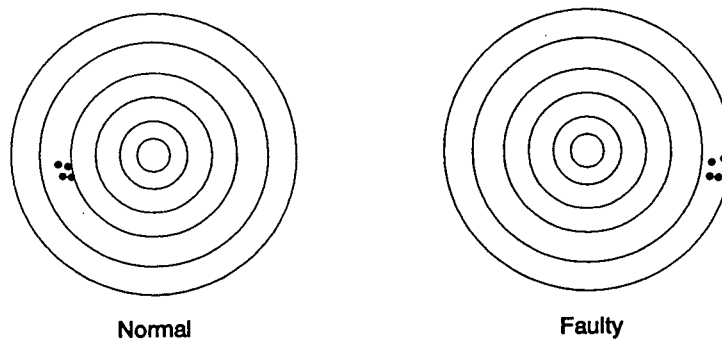
Figure 5.30: Shooting targets to illustrate accuracy and repeatability

repeatability. The reason is obvious: poor repeatability leads to uncertain measuring results. Good repeatability, even under the presence of measurement bias, provides consistent information.

As illustrated by the shooting targets in Figure 5.30, one can immediately establish an analogy with the field of model-based Fault Detection and Isolation: a model which shows consistent results for "normal" and "faulty" behavior of a plant is preferable to another which is more accurate (statistically or otherwise) in reflecting the "true" screen behavior but less repeatable. Accordingly, the first property of the FDI model proposed in this research which should be investigated is its repeatability. To fulfill that pre-requisite, a simulation algorithm was implemented.

The algorithm basically creates batches of data representing all the measurable process variables of a typical pulp pressure screen: pressures, flows, temperature and rotor motor load. It calculates the states and process coefficients included in the screen model and it extracts and plots the "*d*"

	CASE 1	CASE 2	CASE 3
Process Conditions	Motor Current: 57% ± 0.29 Accept Pressure: 320 kPa ± 2.88 Accept Flow: 1500 LPM ± 30 Reject Pressure: 367 kPa ± 3.3 Reject Flow: 500 LPM ± 10 Feed Flow = Accept + Reject Feed Pressure: 350 kPa ± 3.15 Tank Temperature: 41 deg.C ± 0.21	Same	Same
Perturbation	None	Increasing accept flow (Ramp at half of batch with 0.5% slope)	Decreasing accept flow (Ramp at half of batch with -0.5 % slope)
Noise	Random with uniform distribution	Same	Same

Table 5.1 Screen Simulations. Process Conditions

coefficients associated with the friction inside the screen. Noise can be added to all signals to make the simulation stochastic.

The process data can be made to correspond with clearly distinct operating conditions: normalcy and failure. It also allows for the superposition of several disturbances on all variables related to either condition. Thus, one can have as many data batches as desired with the only difference among them being the perturbation or the added noise. This allows for the testing of very similar data sets. For these sets, the states and the coefficients in the model are estimated and the repeatability of their values is investigated.

Unfortunately, as it has been shown in previous chapters, the precise relations between the different screen process variables (particularly pressures and flows) which might indicate the presence of a fault are not well understood. Empirical knowledge on the screen friction losses is also lacking. That combination of facts means that manipulation of these variables to accurately simulate incipient faults is out of reach[10], but one can test the operational extremes of 'normal' and 'faulty' behavior of a screen suffering from blinding and make them follow one another. If good estimation repeatability is achievable during simulations of normal operation and failure, one might reasonably argue that the model should give repeatable results in the intermediate ranges that precede the two extremes, i.e., in the presence of incipient faults. Such argument can be tested against data showing normal operating conditions followed by failure and ultimately, against real data. For now, let us concentrate on the simulations.

	CASE 4	CASE 5	CASE 6
Process Conditions	Same as in case 1	Same	Same
Perturbation	Decreasing accept flow and increasing reject flow (Ramps at half of batch with ± 1% slope)	Fast decreasing accept flow (Step at half of batch with -2% gain)	Decreasing accept flow (Ramp at beginning of batch with -0.05 % slope)
Noise	Random with uniform distribution	Same	Same

Table 5.2 Screen Simulations. Process Conditions (Cont.)

In order to test the repeatability of the proposed model-based strategy, six different pressure screen operational cases were evaluated using the simulation program:

1. Screen operated in absence of perturbations (normally operated with change in feed).

---

10 If that simulation was possible, the whole idea of the screen model as a fault detection tool would be moot as one would simply look at the "right" variables for indication of the known faulty behavior.

2. Screen with sudden increase in accept flow (normalcy with change in throughput).

3. Screen with sudden decrease of accept flow (normally operating screen suddenly brought to failure).

4. Screen with sudden decrease in accept flow and increase in reject flow (another option for a screen made to fail).

5. Screen with step changes affecting the screen flow (normally operating screen followed by fast occurring failure).

6. Failed screen.

The characteristics of this cases, including process conditions, perturbations, and process noise, are shown in Tables 5.1 and 5.2. It is pertinent to mention that each simulation case has faster dynamics than the precedent, except for case 6 which obviously has slower dynamics than most of the others.

Cases 1, 2 and 6 correspond to what have been termed as the extremes of pressure screen behavior: normalcy and failure. The remaining three cases included normalcy followed by conditions associated with failure. To illustrate for the reader, the process variables for cases 4 and 5 can be seen in Figures 5.31 and 5.32.

At this point, as establishing the statistical fit of the model is not relevant, the model noise matrix C_* used in the simulations was selected arbitrarily with coefficient values which were deemed "reasonable"[11] and kept the same for all cases. Since blinding is the underlying agent of the simulated failures, tracking the behavior of the accept losses coefficient is of particular interest[12]. After running the algorithm used to extract the values of the screen losses coefficients, the results for d_1 shown in Tables 5.3 and 5.4 were obtained. Those same results can be observed in graphic form in Figure 5.33.

In the graphic, the center lines represent the average value obtained in the different trials tested for each case. The upper and lower values represent the average plus or minus $t.\sigma$ of the samples, which for a Gaussian distribution implies a confidence factor of approximately 95% [12]. The coefficient t is Student's "t" and has a value of 2 for a sampling universe larger than 30.

---

[11] That is, corresponding to a system which is causal and realizable.

[12] As the screen blinds, the passage of fibres from the feedstock to the accept line becomes more difficult. This suggests a change in k_A and subsequently on the d's.
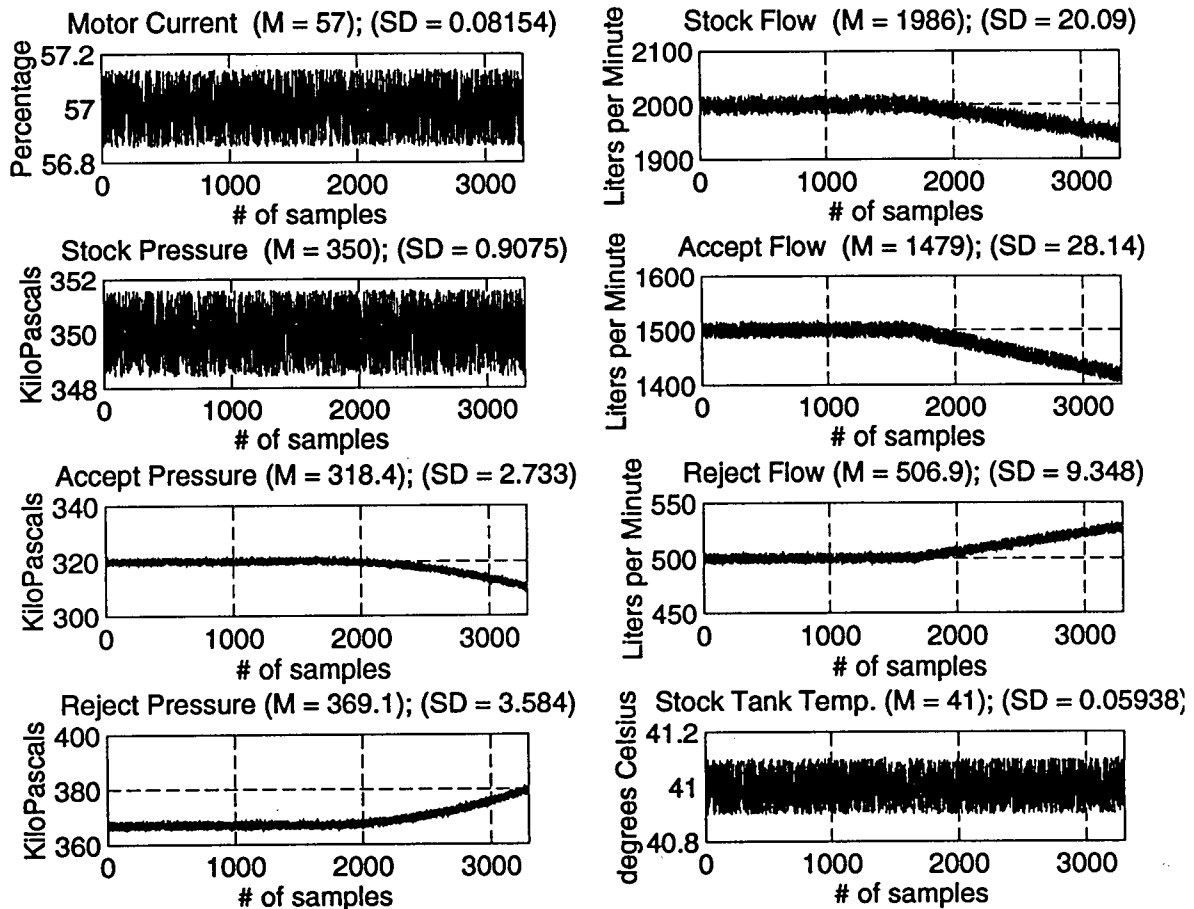
Figure 5.31: Simulated data for pressure screen which begins to fail (Case 4)

Variations of cases 1,2, 4 & 6 were also tested to evaluate the behavior of the reject friction coefficient when under simulated pipe plugging. This time, the main manipulated flows and pressures were the reject. For instance, in case 2, instead of increasing the accept flow, the variable dispensed with was the reject flow. Cases 1, 4 & 6 were treated analogously. The results obtained are shown in table 5.5, and can be observed graphically in Figure 5.34.

From the information on the tables and graph the conclusion is obvious: the proposed model-based strategy assures that most cases which are significantly different from an operational point of view will produce repeatable and discriminating results for the estimated process coefficients. This is equivalent to saying that it has the potential to be used for fault identification and diagnosis: the increase in the magnitude of the d coefficients (particularly d_1 and d_3) above a safe value can be
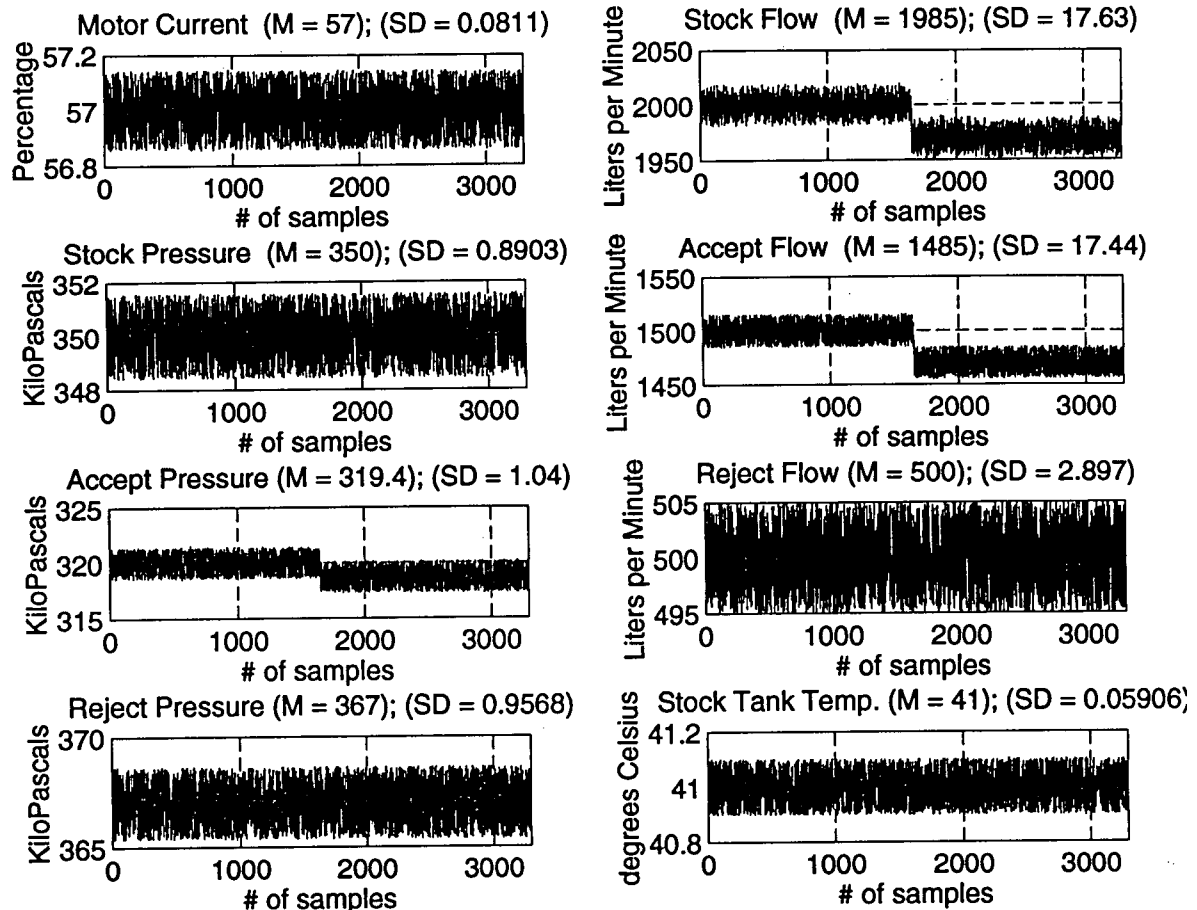
Figure 5.32: Simulated data for pressure screen which goes into fast failure (Case 5)

	CASE 1	CASE 2	CASE 3
d1 **Statistics**	Mean: 108.45 σ: 8.14 Range: 26	Mean: 156.17 σ: 3.67 Range: 10.18	Mean: 195.48 σ: 13.54 Range: 37.01

Table 5.3 Screen Simulations. Results

used as an indication that a fault has occurred. The main question that arises from the simulation is the following: why are the values for the accept coefficient of a failed screen similar to those of a 'healthy' screen? And for that, the answer is simple: because they reflect similar rate of changes in the accept friction losses of the screen. Let us take a more detailed look into this statement.

In the model proposed in this research, the accept losses coefficient d_1 represents the inverse of

70

	CASE 4	CASE 5	CASE 6
d1 *Statistics*	Mean: 193.06 σ: 13.36 Range: 31	Mean: 257.44 σ: 3.53 Range: 9.78	Mean: 125.16 σ: 6.87 Range: 19.05

Table 5.4 Screen Simulations. Results (Cont.)
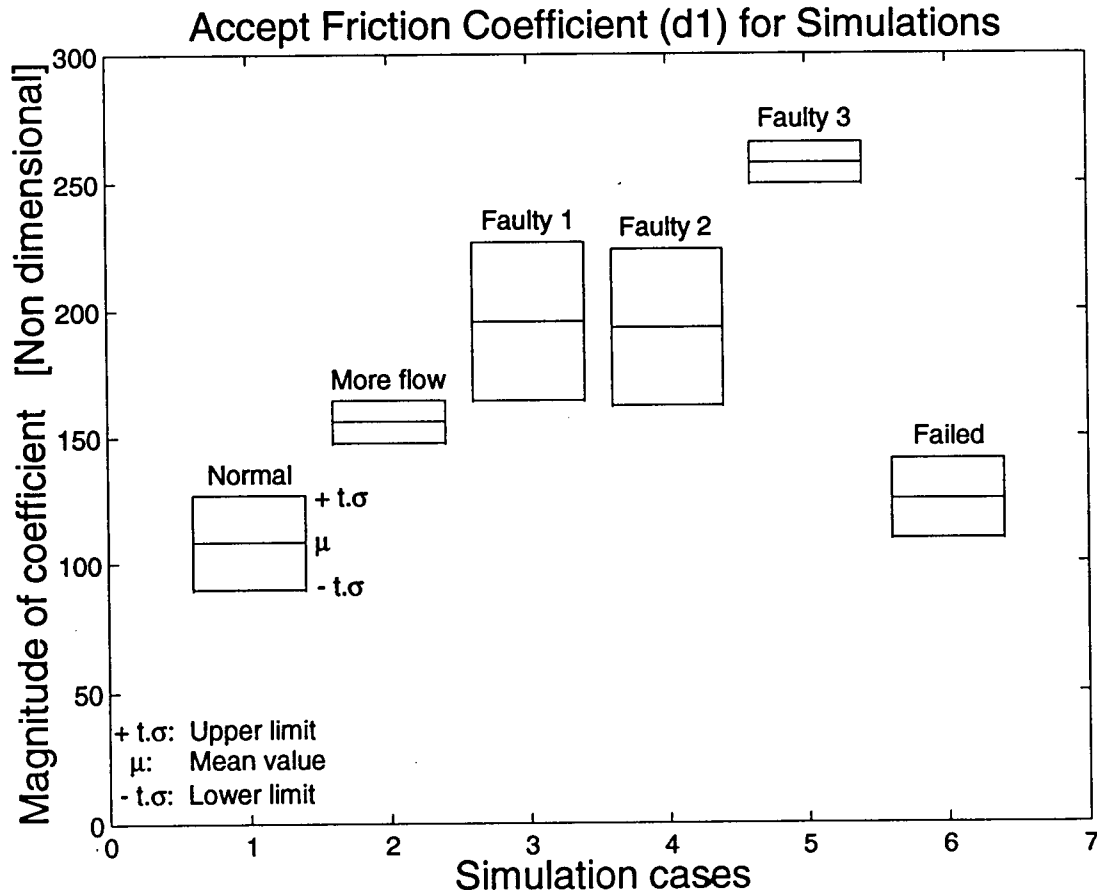
Figure 5.33: Accept Friction Coefficient in Simulations

	CASE 1	CASE 2	CASE 4	CASE 6
d3 *Statistics*	Mean:143.99 σ: 3.12 Range: 7.59	Mean:162.17 σ: 2.39 Range: 7.5	Mean:169.72 σ: 5.58 Range: 17.08	Mean: 145.38 σ: 6.12 Range: 19.83

Table 5.5 Screen Simulations. Results (Cont.)

Reject Friction Coefficient (d3) for Simulations

Figure content showing a box plot chart with y-axis "Magnitude of coefficient [Non dimensional]" ranging from 0 to 200, and x-axis "Simulation cases" ranging from 0.5 to 6. Box plots labeled: Normal, More flow, Faulty 2, Failed.

Legend annotations:
+ t.σ
μ
- t.σ

+ t.σ: Upper limit
μ: Mean value
- t.σ: Lower limit
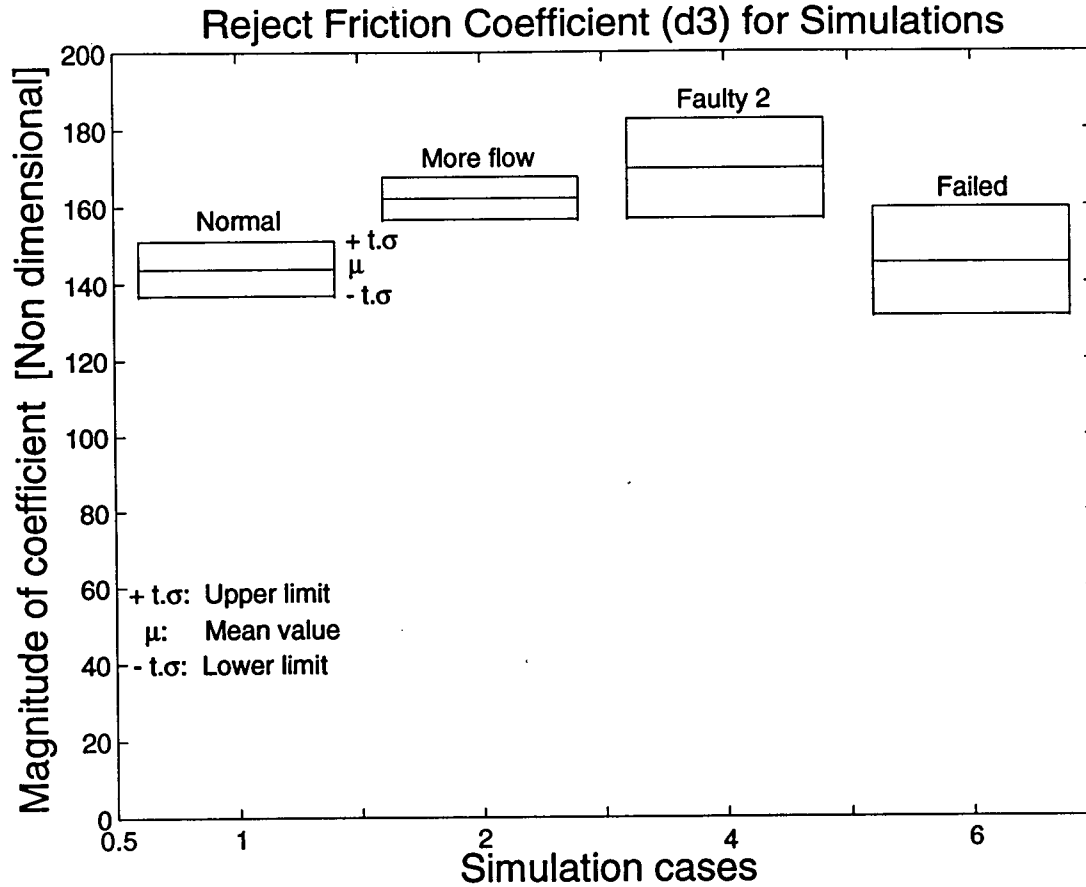
Figure 5.34: Reject Friction Coefficient in Simulations

one of the time constants governing the behavior of the system. From the modified D'Arcy-Weisbach expression used to represent the screen friction losses

$$\dot{k}_A = d_1 k_A + d_2 v_A^2 \qquad (5.99)$$

assuming the velocity v constant, and reordering, the equation becomes

$$\frac{dk_A}{dt} - d_1 k_A = \alpha d_2 \qquad (5.100)$$

which gives a weighting function of the form

$$k_A(t) = k_0 e^{d_1 t} \qquad (5.101)$$

It is easy to see then that d_1 (which must be negative if plugging occurs) represents the rate of change of the accept friction losses. The faster the variation in the losses, the bigger the magnitude

of d_1 becomes, and vice versa. In a well operating screen the plate orifices are mostly clear. A failed screen shall have those orifices mostly plugged. Although one would expect to see a normal screen show less friction losses than a plugged (failed) screen, in both cases the rate of change of those losses should be similar: fairly low, and the value of the coefficient d_1 should reflect that. On the contrary, a screen which is suffering from blinding would display a higher absolute value of d_1 until it gets totally plugged.

5.3 Experiments in the Pilot Plan

5.3.1 The setting

To further explore the validity of the FDI model-based strategy proposed for the screens, test trials were conducted at Paprican's screening research pilot plant. As described in Chapter 3, this setting made possible not only to operate and gather data on typical pressure screens but it also allowed for the induction of most kinds of faults. This is as close as it is possible to get to the industrial reality while maintaining the rigor of a research environment. Special emphasis was put on exploring blinding, among all screen faults, as it is the most common of faults and also the one which causes the most frequent operational perturbations of these devices in mills. Again, the typical fault trial for detection of blinding would include the following steps:

1. A screen is chosen for the trial. This includes plate selection.
2. Pulp consistency is adjusted.
3. The screen valves are set at starting values.
4. The reject flow rate for the trial is chosen. Several accept flow rates are selected, in incremental steps. The stock pressure is to be maintained at 350 kPa.
5. The pump and screen motors are started.
6. Samples of pulp for determination of initial consistency are taken.
7. The pipes are tested for air entrapments and, if detected, purged.
8. The valves are controlled manually until the starting process conditions, with zero accept flow, are met.

9. Slow data sampling through the DCS is started.

10. The accept valve is opened to reach the selected accept flow rate and this setting is kept until steady state is achieved. Fast sampling data is collected by the PC for periods of up to thirty seconds.

11. If failure is not present, the valves are manipulated to achieve the next higher accept flow rate previously selected, and step 8 is repeated. If failure occurs, the settings remain untouched until the differential pressure between the feedstock and the accept lines triggers a system shut–down mechanism. Fast sampling data is then gathered at equally spaced accept flow rates until the whole operation comes to a full stop.

12. Samples of pulp for determination of final consistency are taken.

Figure 5.35 shows a schematic diagram of the screen instrumentation used to conduct all tests. The interconnection DCS-PC was designed and installed and all programs for the data acquisition and calculations were written in LabView® and Matlab®.
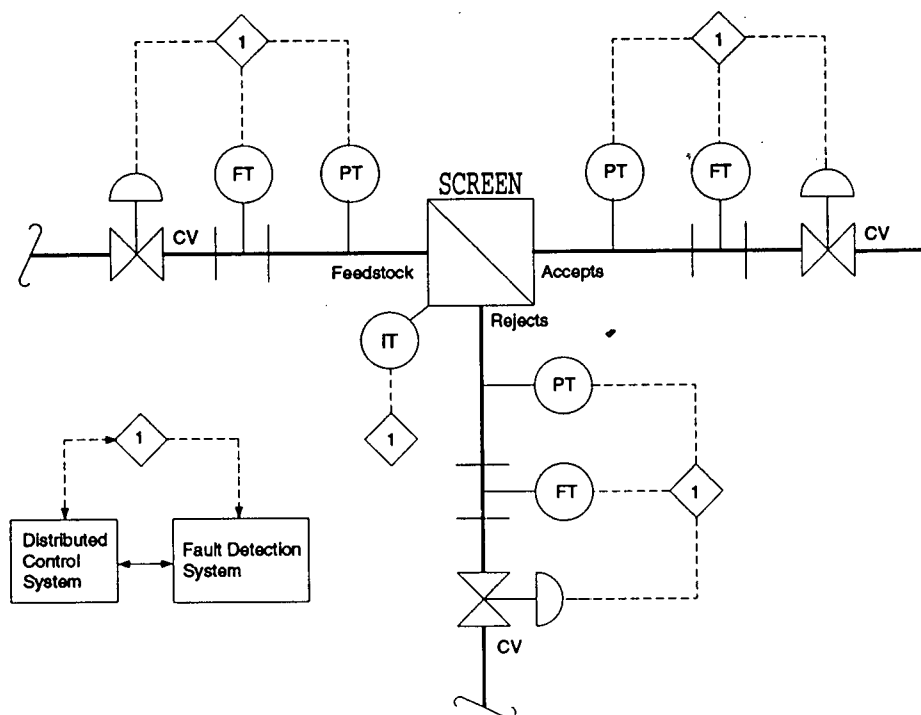

Figure 5.35: Instrumentation scheme for pressure screen tests

The pulp used was CTMP unbleached softwood, composed by 10% aspen, 45% spruce, and 45% fir. Its freeness was 100 ml CSF and all trials were run at consistencies ranging from 1.1% up to 2.8%. Coarseness of the pulp was not available.

The sampling rate for the fast data acquisition was chosen to be 300 Hz and the slow sampling was selected at 1 Hz. In the first case, the frequency was singled out to comply with the time constants of the expected phenomena, the bandwidth of the screen sensors, Shannon's theorem and the bandwidth of the anti-aliasing filters. The second frequency was picked to comply with the DCS speed limits.

First, several rounds of trials with varying consistencies were conducted to gather general information on the screen friction coefficients, and to determine the accept flow values at which failure occurred. Later on, trials starting at accept flow ranges close to failure were studied in more detail. For that purpose, the amount of increases in the flow rate mentioned in step 9 was made small but all other provisions remained the same.

5.3.2 Some Results

As it has been mentioned before, there is a correspondence between the coefficients d_1 and d_3 (which have been termed *accept losses* and *reject losses* coefficients) and the time constants in the model. Under the light of the simulation results, it seemed logical to concentrate the attention on the magnitude of those coefficients as potential indicators of faults. This assumption was proved correct.

The typical set of tests in the pilot plant to determine the fault detection capabilities of the proposed model would produce results for the model losses coefficients as seen in Figure 5.36. These graphs correspond to a single operating point during a given trial. When all the results corresponding to a trial in which the screen is gradually brought to failure are put together, the graph corresponding to the magnitude of the *d* coefficients (Figure 5.37) reflects the same pattern observed during the simulations: under normalcy the magnitude of the *d* coefficients is small. When an increase in the screen throughput is suspected of causing blinding, d_1 begins to grow faster and faster, until it reaches a maximum. This point coincides with the external evidence gathered by the sensors in the pilot plant that tells us that failure has indeed occurred. After the screen effectively fails, the value
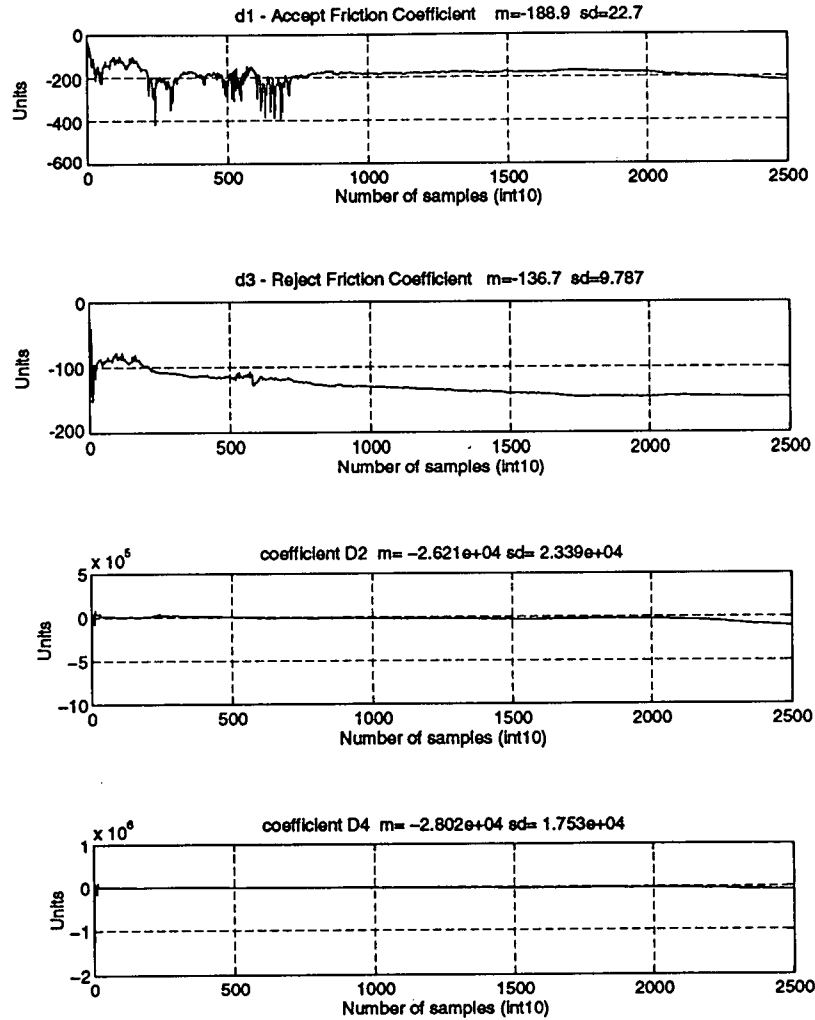
Figure 5.36: Screen friction coefficients

of d_1 diminishes to a range close to the one associated with normalcy. The variation in d_3 is only

noticeable when the maximum blinding is achieved. Since d_3 is referred to the reject flow, one might

theorize that it is only at the point of maximum blinding that the impact on the reduced accept flow

starts to have an impact on the reject piping friction conditions.

As failure of the screen is deliberately caused by blinding, the pattern observed in the results

agrees with the physical meaning of the proposed model, namely, that the coefficients reflect the rate

of change of the friction losses inside the screen.

d1 - Accepts Physical Coefficient (Rej. Flow = 1000 L/min) (Noise= 1.5, 0.75, 1)

d3 - Rejects Physical Coefficient (Rej. Flow = 1000 L/min) (Noise= 1.5, 0.75, 1)
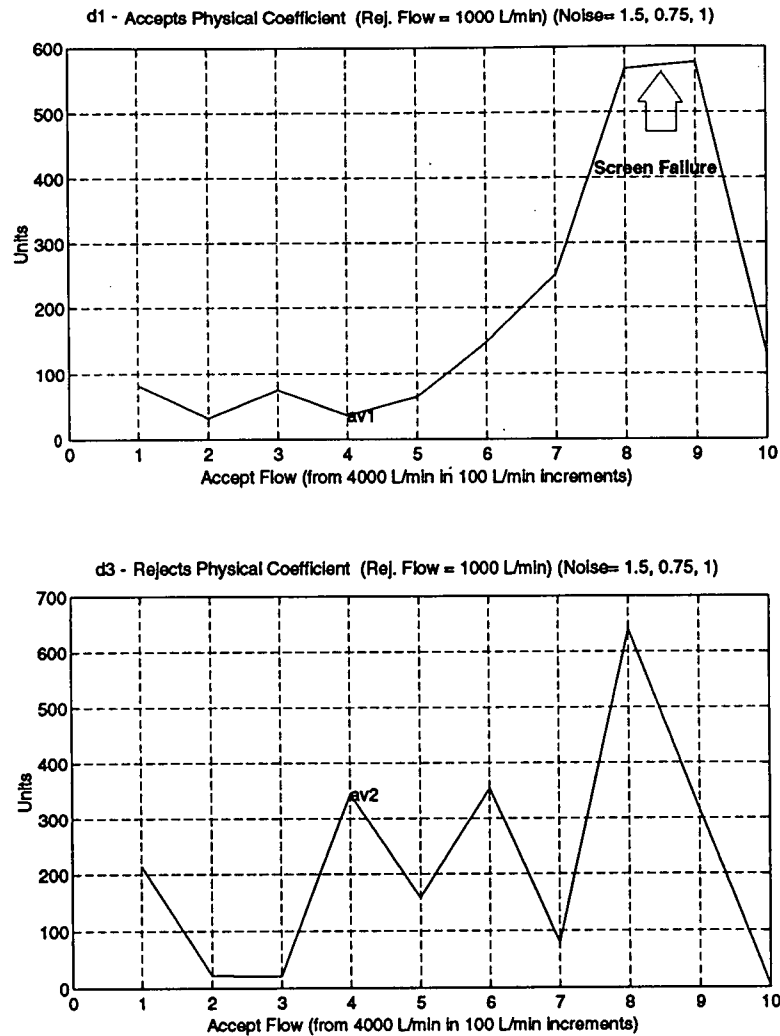
Figure 5.37: Screen friction coefficients

Let us look at the same results from a different angle. As shown in Fig. 5.40, when the accept flow is between 4100 and 4500 L/Min the magnitude of the accept losses coefficient d_1 remains below 100 non-dimensional units. When the flow moves beyond 4500 L/Min, the value of the coefficient begins to grow rapidly, suggesting a more difficult passage of the pulp through the screen plate openings, presumably due to the appearance of blinding. The value keeps growing until the screen begins to show the effects of plugging, and reaches its maximum shortly thereafter this condition is firmly present. Although at the present time there are no mechanisms to measure the
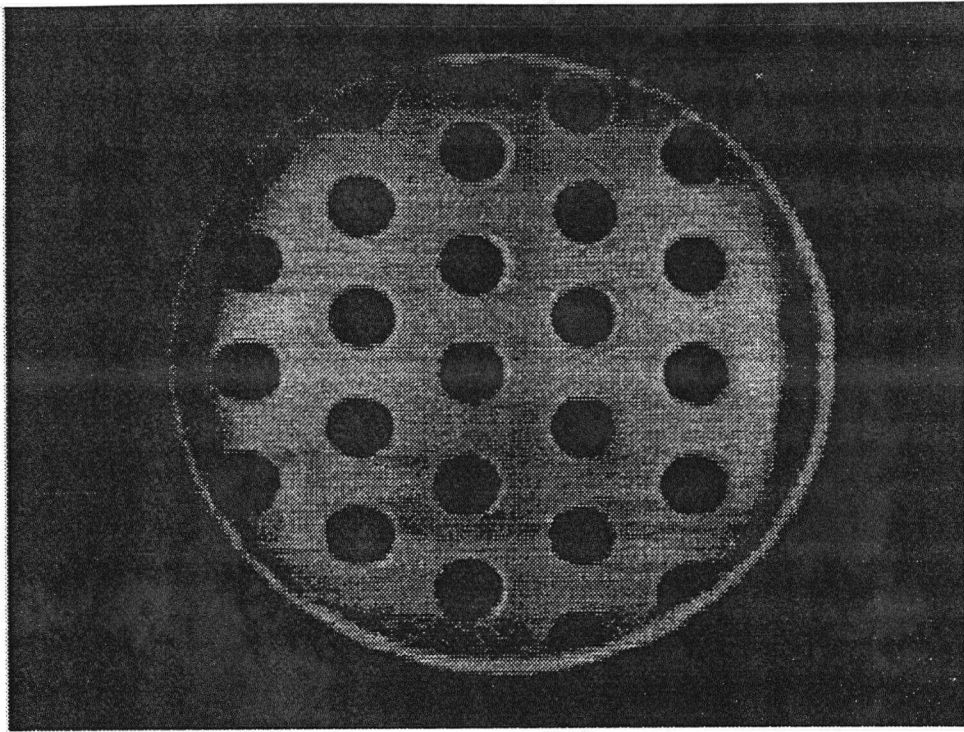
Figure 5.38: Snapshot of screen plate orifices under normal operating conditions
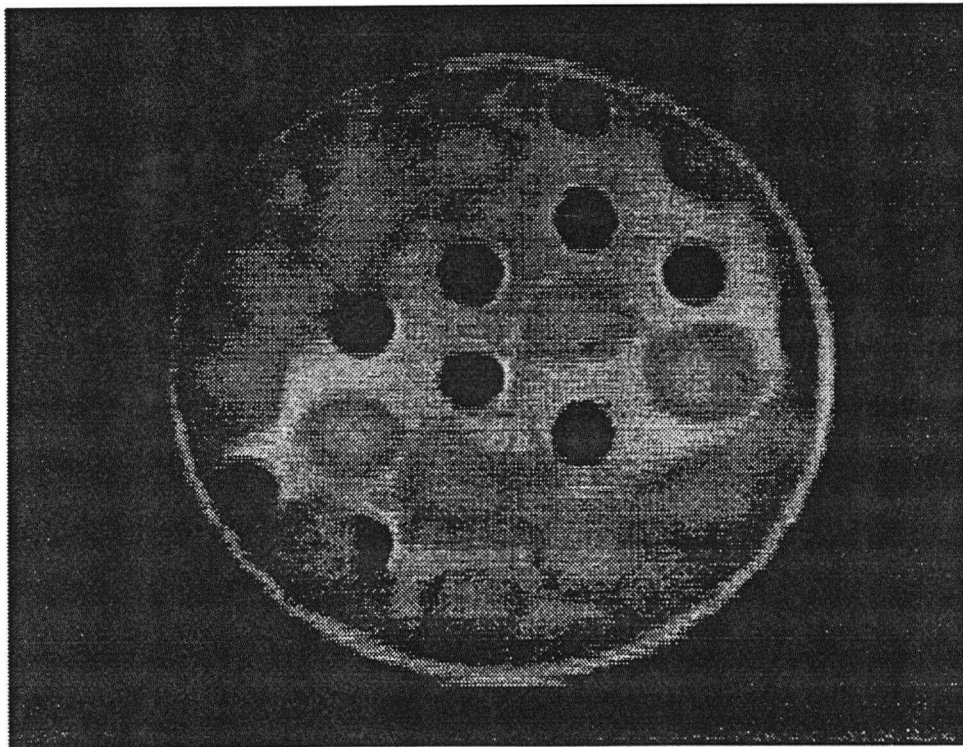
Figure 5.39: Snapshot of screen plate orifices under blinding
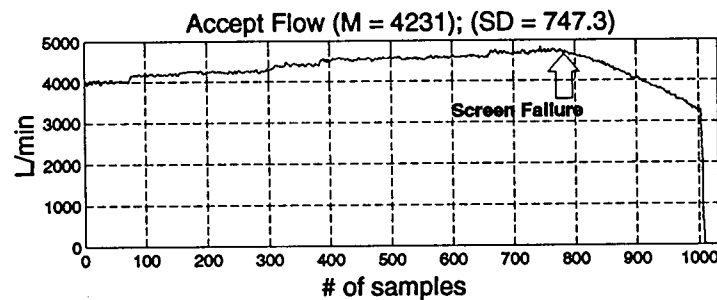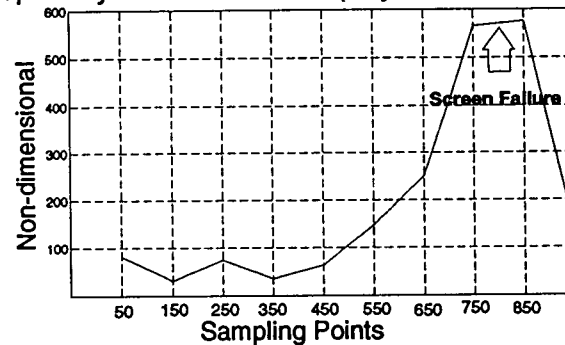
Figure 5.40: Accept friction coefficient and Accept Flow

accumulation of fibres inside the screen plate, observations through the use of a high speed video camera (See Figures 5.38 and 5.39) that shows some portions of the interior of the screen plate, seem to confirm this theory. It is easy to see then that the magnitude of d_1[13], which represents the rate of change of the accept friction losses, grows larger when the rate of accumulation of fibres increases. The faster the variation in the losses, the larger the magnitude of d_1 becomes. When the screen is totally plugged and, therefore, failure is declared, the rate of change of the losses falls back to a value similar to the ones obtained for a "healthy" screen. This is what would be expected, following the theoretical predictions of the proposed model. The natural conclusion is that by monitoring the averaged coefficients imbedded in said model, failure of the screen due to blinding could be predicted before it happens [15]. If the magnitude of the coefficients grows steadily and exceeds twice the

---

[13] The value of d_1 is always negative, as plugging is underway.

value obtained in the normal operating region, one could say fairly confidently that the screen is under a faulty condition.

During the different trials some interesting observations were made:

1. For accurate results, fast uniform sampling and anti-aliasing filtering are required. Some commercial DCS do their sampling using a mechanism called 'report by exception'. This implies that the control software does not poll the incoming field signals at regular time intervals but whenever its computing cycles allow for it, or when the value of the signal exceeds a pre-set minimum percentage of full range. This practice makes impossible to use the strategy proposed here. In such case, a dedicated PC can easily handle the sampling and filtering requirements.

2. High noise spikes in the process data can affect the convergence of the Kalman filter, effectively derailing the estimation of the friction coefficients. Therefore, they must be filtered out. Figure 5.41 shows the effects caused by the appearance of this problem which is caused by the data acquisition electronics. The spikes at samples 125, 340, 800, 1800, 2600, etc. in the pressure and motor load signals make the accepts friction value change suddenly, making almost impossible to assess its "true" value.

3. The choice of parameters in the noise matrix will affect the estimation results. This has an effect on the range of values that the d coefficients might take. As such, good care must be taken to assure that the chosen noise model fits the process reasonably well. As this point is extremely important, it is mandatory to examine it more closely.

5.3.3 The Choice of a Noise Model

The singular-pencil-based FDI strategy proposed in this and previous chapters requires that the matrix C_* elements be fixed before identification takes place. Ideally, such selection must reflect the *a priori* knowledge of the process noise. As the choice of C_* has an influence on the results achieved by the estimation, intuition suggests that the more closely it resembles the "true" noise in the process, the better the identification results should be. The results in the pilot plant tests confirmed this assumption. As seen in Figure 5.42, depending on the choice of elements of C_*, the estimation of the accepts losses coefficient d_1 renders different values for the ten points included in the same

Motor Current (M = 59.18); (SD = 0.4685)   Accept Pressure (M = 217.9); (SD = 4.095)
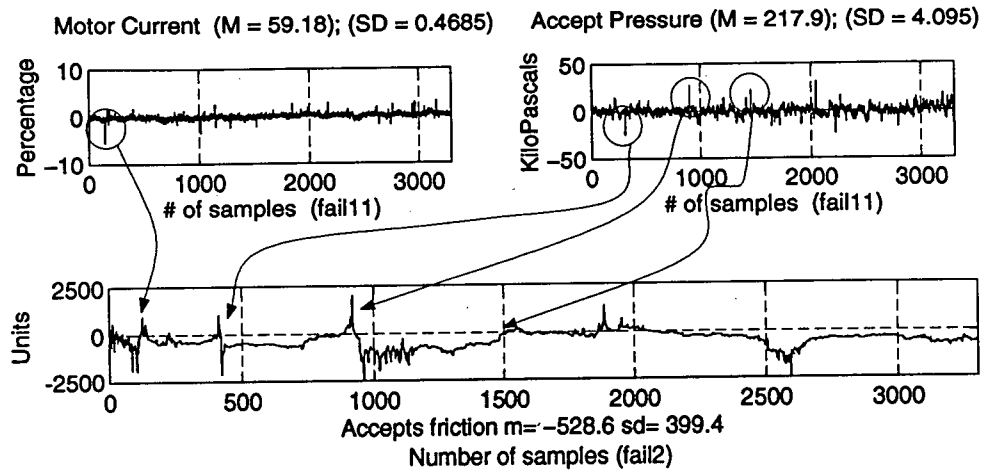
Figure 5.41: Accept friction coefficient and Accept Flow

trial. Although it is clear that blinding of the screen is identified by all the noise models (C_*1, C_*2, and C_*3), the one with the greatest difference between the pre-failure and the 'healthy' points values (C_*2) is the better suited for FDI. The reasons are simple: it is less prone to false alarms and it allows for the most time to take preventive actions. The meaning here is clear: proper matrix selection plays an important role in the overall success of the FDI strategy.

Rather than elaborating on how to make an *a priori* selection of C_* for identification (which will be dealt with in chapter 6), the purpose of this section is to determine whether a good enough choice of the matrix has been made, FDI-wise, when C_* is given.

At this point in the model validation procedure, the best tool to assess the appropriateness of the noise matrix C_* seems to be an evaluation of the statistical lack of fit of the different choices for it. To check the goodness of fit of the coefficients used in the model, one can look at the "whiteness" of its residuals by using auto-correlation functions or some statistical criterion such as Box-Pierce's, Durbin-Watson's or Akaike's [4]. In any case, the residuals must be calculated first using the estimation results.

81

Accept Friction Coefficient d1 for different C* choices
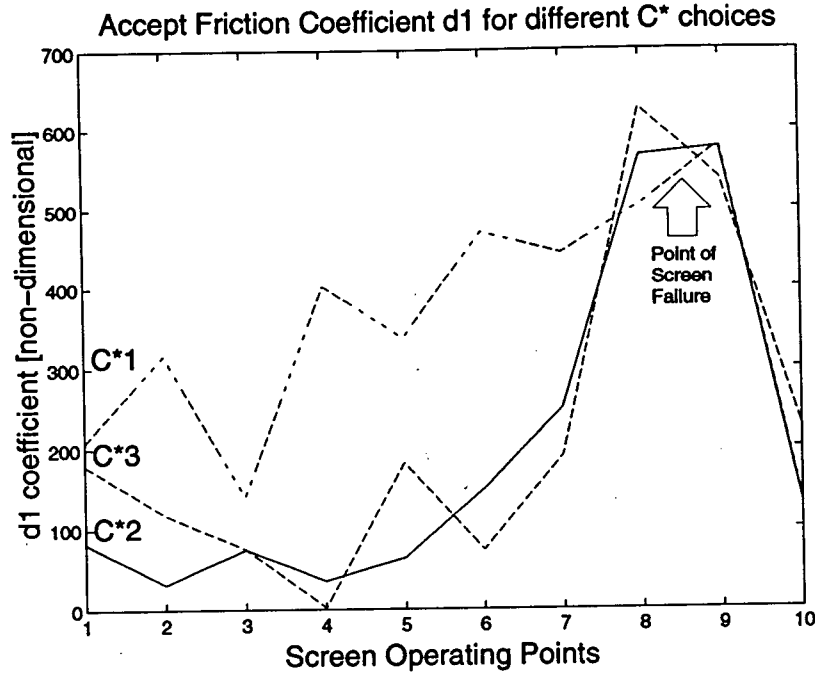
Figure 5.42: Accept friction coefficient and Accept Flow

Since our pencil matrix representation has the form

$$P(z) \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = \begin{bmatrix} E_* - zI & -A_* & B_* & C_* \\ E_0 & -A_0 & B_0 & I \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = 0 \tag{5.102}$$

it is easy to see that the residuals can be obtained as

$$e_k = A_0 y_k - E_0 x_k - B_0 u_k \tag{5.103}$$

If the choice of the noise coefficients is appropriate, $\{e_k\}$ should be a stochastic sequence with a Gaussian distribution, zero mean, and covariance Λ ($\Lambda > 0$).

Getting back to the three choices for matrix C_* that were shown above (C_*1, C_*2, and C_*3), when looking at the auto-correlations of the residuals during screen normalcy, i.e., at operating points 1, 2 or 3, the typical values look as seen in Figures 5.43, 5.45, and 5.44. It is obvious that choice C_*2 it is the best suited, as speculated, as it satisfies the 2σ condition which gives 95% statistical confidence.
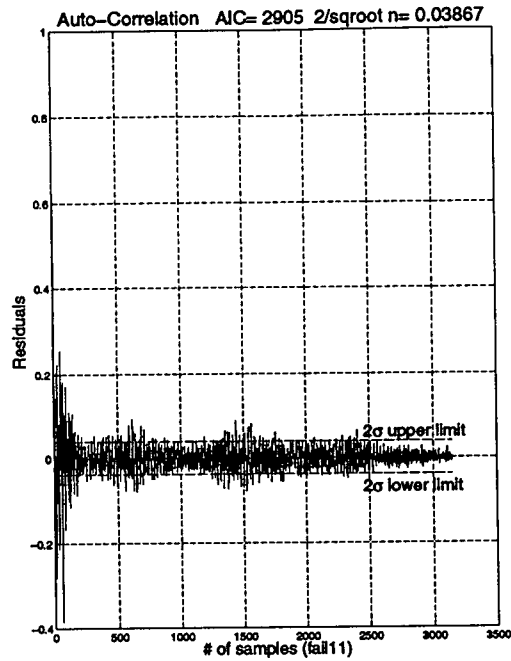
Figure 5.43: Residuals during normalcy. Model C*1
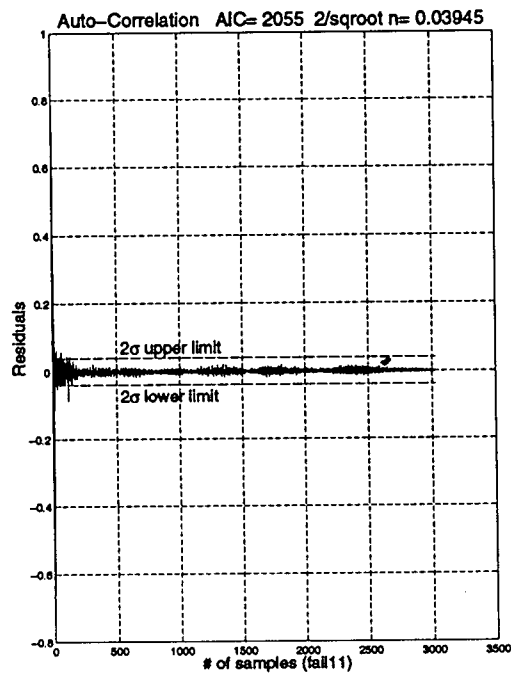
Figure 5.44: Residuals during normalcy. Model C*2
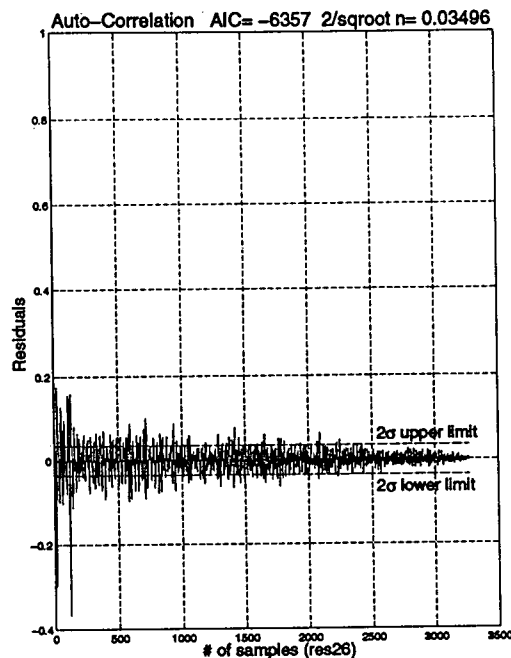
Figure 5.45: Residuals during normalcy. Model C*3

The circumstances change drastically when, instead of looking at normalcy, one looks at the points in which blinding is on the verge of making the screen plate fail. As seen in Figures 5.46, 5.47, and 5.48, now the auto-correlations tell us that no model passes the statistical confidence test. There is a clear mismatch between the model and physical reality but that's precisely what any process-coefficient-based FDI scheme would be looking for: indications of deviation in the behavior of a system. As the screen model is an attempt to reflect the physical reality during normal operating conditions, one would expect to see it become less and less suited when the screen deviates from those conditions. But the problem remains: how do we know if the choice of C_* is a good-enough one? Obviously, to decide on that, one is confined to checking the normalcy regions, as the faulty regions are by design expected to produce model mismatch. In laymen words: to know if the choice of C_* is good enough for FDI, one should test its statistical properties when the screen is known to be operating normally. Good fit in that region would seem to guarantee the best possible FDI range.

When using simulations to test the generality of the conclusions made upon the pilot plant trials, the results confirmed the initial assumptions. The choice of C_* which guarantees statistical fit during
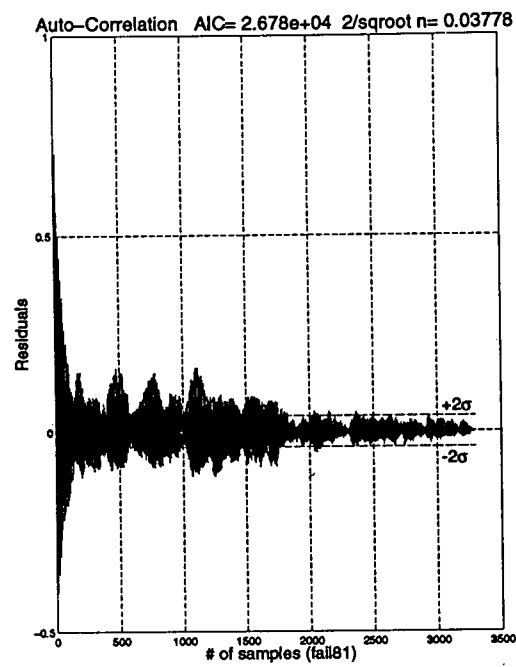
Figure 5.46: Residuals during failure. Model C*1
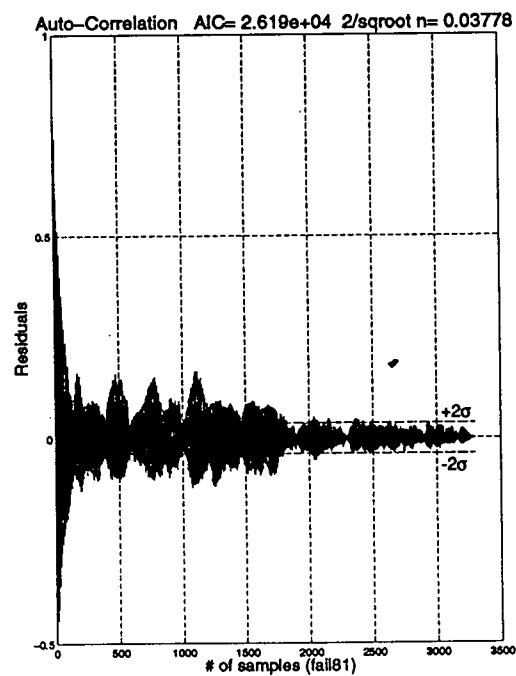
Figure 5.47: Residuals during failure. Model C*2

85
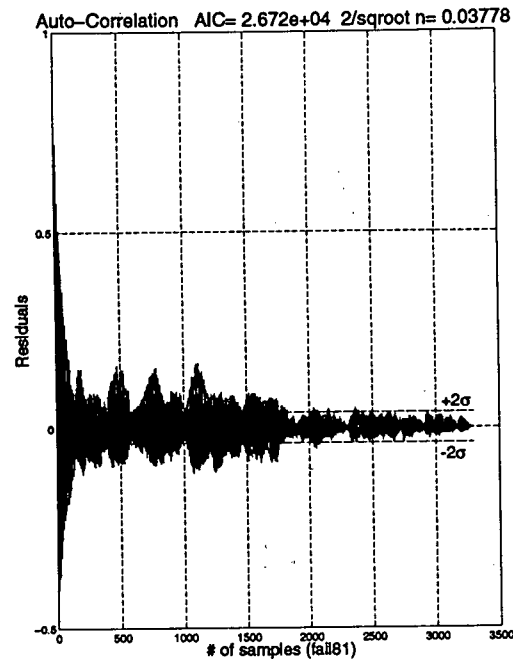
Figure 5.48: Residuals during failure. Model C*3

normalcy (Figure 5.49), becomes marginally fit under a mild fault (Figure 5.50), and is no longer statistically adequate under a fast developing fault (Figure 5.51), but since that is the characteristic FDI is interested in, to test for model adequacy one should confine the investigation to the normal regions of operation.

5.4 Model Reduction

In the process of establishing the appropriateness of a model, an issue which must be addressed is its complexity. If the original model can be simplified without much change on its input-output properties, this reduced version is much preferable and the first one is deemed as too complex. One procedure for achieving this simplification is reducing the model order [45].

In the case of the pressure screens, as the originating differential equations are first order, it is not possible to change the order of the model without totally losing its dynamic characteristics, but we can reduce the number of coefficients in the forming equations to achieve some simplifications. One such reduction which seemed suitable concerned the D'Arcy-Weisbach-based relationship.
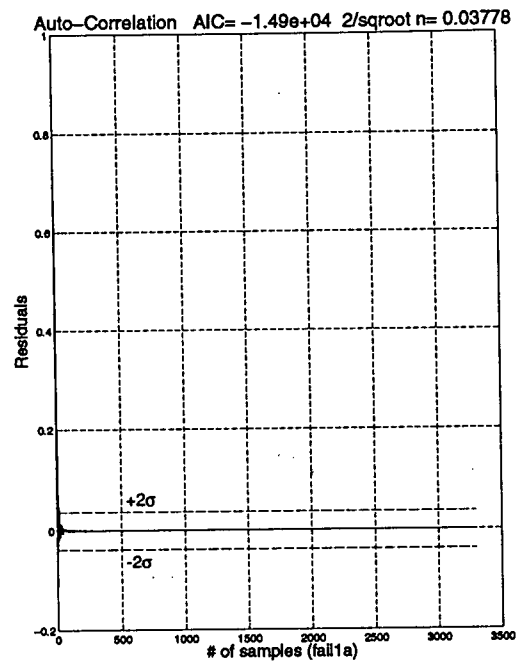
Figure 5.49: Simulation residuals. Normalcy
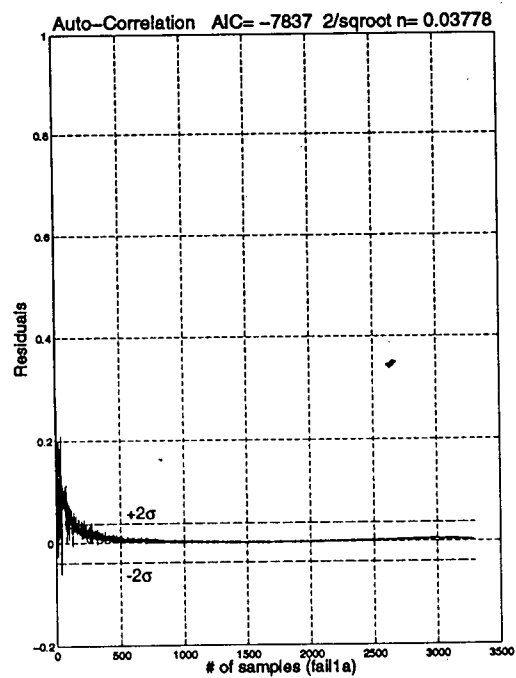
Figure 5.50: Simulation residuals. Incipient mild fault
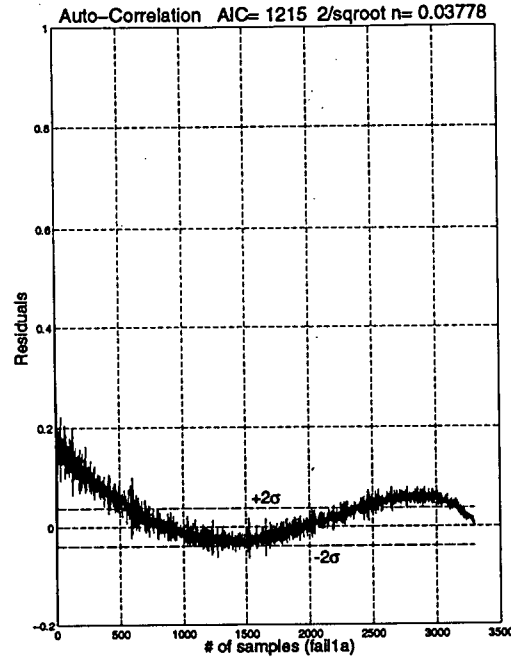
Figure 5.51: Simulation residuals. Fast fault

Instead of the original expression

$$\dot{k}_A = d_1 k_A + d_2 v_A^2 \tag{5.104}$$

which included two coefficients for every dynamic equation used in modelling, a simpler one

$$\dot{k}_A = k_A + d_{R1} v_A^2 \tag{5.105}$$

with only one coefficient was proposed. After linearization, that formulation leads to a continuous state-space representation of the form

$$\dot{\underline{x}} = A\underline{x} + B\underline{u}$$
$$\underline{y} = C\underline{x} + D\underline{u} \tag{5.106}$$

where

$$A = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

$$B = \begin{vmatrix} d_{R1} & -d_{R1} & d_{R1} & 0 & 0 & 0 & 0 & 0 \\ 0 & d_{R2} & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix} \tag{5.107}$$

$$C = \begin{vmatrix} h_8 & h_8 \end{vmatrix}$$

$$D = \begin{vmatrix} h_1 & -h_3 & h_2 & h_4 & -h_6 & -h_7 & h_5 & h_8 \end{vmatrix}$$

88

and where $\underline{x}$, $\underline{y}$, and $\underline{u}$ are the state, output and input vectors, and the d_R's, and h's are defined as for the original model. Unfortunately, when using such simplified model one problem immediately arises: although the reduced model is controllable (provided that $d_{R2} \neq 0$), it is not observable, as the rank of the matrix

$$
Q_b = \begin{vmatrix} C^T \\ C^T A \\ \vdots \\ C^T A^{m-1} \end{vmatrix}
\tag{5.108}
$$

is less than the order of matrix A. In order to achieve uniqueness in identification, a minimal realization of the system has to be obtained. One might decide to obtain such realization at this stage, using the proper state space transformations, or when rewriting the state-space model into a singular pencil model. The latter approach will be taken here for illustration purposes.

As explained in Chapter 4, when starting from a state-space representation

$$
\begin{aligned}
\underline{\dot{x}} &= A\underline{x} + B\underline{u} \\
\underline{y} &= C\underline{x} + D\underline{u}
\end{aligned}
\tag{5.109}
$$

the departing implicit equation is obtained by inspection as

$$
\begin{bmatrix} I \\ 0 \end{bmatrix} \lambda x_k = \begin{bmatrix} A \\ C \end{bmatrix} x_k + \begin{bmatrix} 0 & B \\ -I & D \end{bmatrix} \begin{bmatrix} y_k \\ u_k \end{bmatrix}
\tag{5.110}
$$

or, making

$$
\begin{aligned}
\begin{bmatrix} I \\ 0 \end{bmatrix} &= F \\
\begin{bmatrix} A \\ C \end{bmatrix} &= E \\
\begin{bmatrix} 0 & B \\ -I & D \end{bmatrix} &= G
\end{aligned}
\tag{5.111}
$$

as

$$
P \begin{bmatrix} x_k \\ w_k \end{bmatrix} = \begin{bmatrix} E - \lambda F & G \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ u_k \end{bmatrix} = 0
\tag{5.112}
$$

Since this representation is non minimal, one has to use elementary matrix operations equivalent to row and column compression to obtain a minimal one[14]. Intuitively, this is equivalent to removing all dependent variables in the state space vector x which do not affect the external behavior of the system. The elementary operations are row operations on $[E, F, G]$ or simultaneous column operations on E and G. Those required to achieve a minimal controllable and observable system can be summarized in the following algorithm [5]:

1. Make t = number of columns of E and F.

2. Make q = number of rows of $[E, F, G]$.

3. While t > 0 and q > 0

    Put rows 1,...,q in columns 1,...,t of F into upper compressed form of resulting rank k.

    a. If k = q

       stop.

    b. Else

       put rows k+1,...,q in columns 1,...,t of E into right compressed form of resulting rank r

       make q = k

       make t = t - r

       if r = 0 stop.

4. End.

Once the minimal representation is achieved and after applying transformation T_1, the new model has the by-now-well-known form

$$P(z) \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = \begin{bmatrix} E_* - zI & -A_* & B_* & C_* \\ E_0 & -A_0 & B_0 & I \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = 0 \qquad (5.113)$$

---

[14]   Upper row compression corresponds to pre-multiplication of matrix A by non-singular matrix W such that $WA = \begin{bmatrix} A_r \\ 0 \end{bmatrix}$ where A_r has full row rank. Lower row compression and left or right column compression are defined by analogy.

with the following values

$$E_* = |0|$$

$$E_0 = |1| \tag{5.114}$$

$$A_* = |0|$$

$$A_0 = |1|$$

$$B_* = |j_1(1 - h_1) + h_8 j_2 \quad (h_3 - 1)j_1 + h_8(j_3 - j_2) \quad j_1(1 - h_2) + h_8 j_2$$

$$-h_4 j_1 \quad -h_5 j_1 \quad h_6 j_1 \quad h_7 j_1 \quad -h_8 j_1| \tag{5.115}$$

$$B_0 = |h_1 \quad -h_3 \quad h_2 \quad h_4 \quad h_5 \quad -h_6 \quad -h_7 \quad h_8|$$

and C_* to be chosen from previous knowledge of the process noise. From here, the manipulations to perform the system identification described in Chapter 4 are straightforward, and one can proceed to estimate the minimal state and the physical coefficients in the screen model, namely, d_{R1} and d_{R2}.
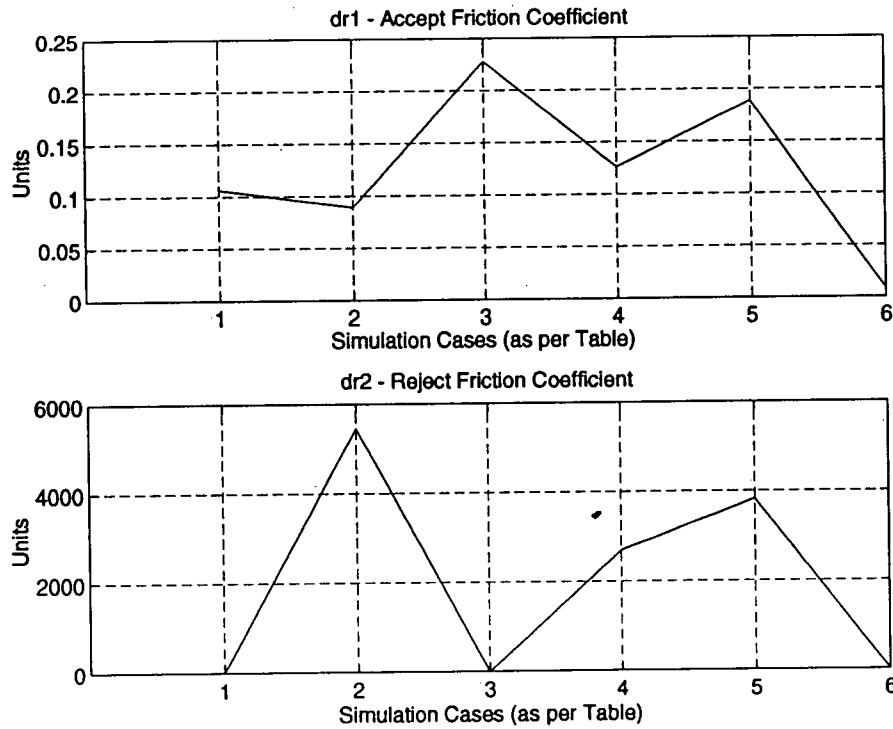
Figure 5.52: Reduced Model. Friction Coefficients in Simulation Trials

Unfortunately, when using the new reduced model in simulations it did not show the same usefulness as the original one, as seen in Figure 5.52. The original model seemed capable of displaying

the characteristics of the dynamics affecting the screen: faster accept flow dynamics would produce a higher value of the accept friction coefficient. This feature can easily be associated with the screen faults, particularly with blinding, which is the most important of them from an industrial perspective. As the graph purports, the reduced model has been stripped of that ability.

When the reduced model was applied to the pilot plant test trials, its usefulness was even worse, as seen in Figure 5.53. In this case, the variations in the magnitude of the accept friction coefficient could not be clearly related to the presence of faults.
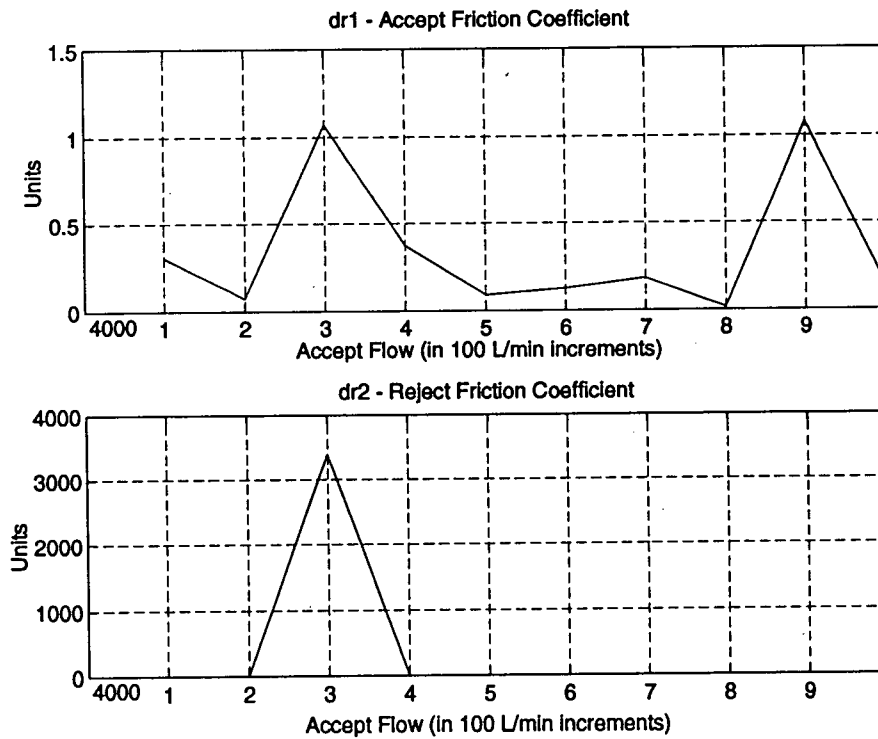
Figure 5.53: Reduced Model. Friction Coefficients in Pilot Plant Trials

An explanation for the inability of the new model to give a clear picture of the situation inside the screen can be obtained when looking at the original equation representing the screen losses

$$\dot{k}_A = k_A + d_{R1} v_A^2 \tag{5.116}$$

That relation gives a weighting function of the form

$$k_A(t) = k_0 e^t \tag{5.117}$$

and it seems clear that for the reduced model, the d_R coefficients have lost their relationship with the energy losses k and are only associated to the flow variations inside the screen. As such, any change in those flows will affect their dimension, regardless of what might be the cause for such change. Given the fact that the model proposed originally for the pressure screens is of first order, the evidence seems to indicate that any simplified version of it will be stripped of the desired FDI properties present in the original.

In all areas of model-based control, finding a model which describes the "true" system under study is an ideal. Unfortunately, it is philosophically impossible to discern whether a particular model is the "true" one or not, and one must settle for a good enough model for the purposes at hand. In this case, judging by the results, the original model proposed for the screens should be preferred.

5.5 Summary

In this chapter a 3–step methodology for testing the proposed model-based FDI approach on pressure screens has been presented and the main results obtained with this approach have been shown. After disclosing the repeatability results of healthy and faulty screens simulations, experimental results on the screen pilot plant have been presented. Those findings imply that the d coefficients included in the proposed screen model are good indicators of the presence of faults. In particular, when dealing with blinding, the coefficient d_1 has proven to be useful. The problem of choosing the noise matrix C_* included in the model has been explained. Finally, a reduced screen model has been proposed and tested and the results have been displayed. The shortcomings found when using this reduced model for FDI underline the benefits of utilizing the original (non-reduced) model.

Chapter 6

Additional Topics on Estimation

6.1 Noise Modelling

6.1.1 The Problem

In past sections it has been explained how the Singular Pencil approach can be used for FDI purposes on pulp screens. Some results have also been presented but, all along, when applying the system identification scheme, it has always been assumed that the coefficients present in the process noise model are known from the beginning. In practice, such knowledge seldom exists. Therefore, the natural question is: how can one obtain the coefficients of the noise? Let us take a look at that problem which, needless to say, is not a simple one.

Trial and Error

When facing the need to obtain *a priori* values of C_* for fault detection purposes, the first option for the user is the simplest: trial and error. The procedure would go along the following lines: take any starting value for c_{111}, c_{112}, and Λ in the matrix C_*, provided that they comply with the realization/causality conditions[15], and start the estimation during operating conditions which are known to be very safe, or well below in the screen "healthy" region. Look at the residuals and, if not "white" enough, try with different values. When some coefficients make the residuals look "white", stop the search and move the screen onto the desired (and more "risky") regions of operation. The results achieved in Chapter 5 show that this approach is feasible and that the chosen model should be FDI adequate. Unfortunately, feasibility does not equal practicality.

Trial and error is simple and does not involve many calculations but, from an applications point of view, is not recommendable. Although with enough iterations the operator can develop a certain "feel" for what coefficients might constitute a good choice, the process can be lengthy and time consuming, and there's no guarantee that a suitable-enough model for C_* will be found.

---

[15] A system described by a transfer function is realizable if it satisfies the causality principle, i.e., the output variables do not depend on future values of the inputs [32].

Non Parametric Estimation

A second option would be to obtain the *a priori* knowledge on the process noise by non-parametric methods such as the application of auto, cross-correlations and spectral density functions, and to use that information to obtain the coefficients in an ARMA transfer function which would produce colored noise of similar characteristics. That transfer function could then be transformed back into the required Singular Pencil representation using the reverse similarity transformations T described in Chapter 4.

Before further elaborating into the mathematical steps required for the complete procedure, one warning sign grabs the attention: as shown in Appendix B, all the T transformations (and their inverses) involved in the Singular Pencil approach require knowledge of the parameters in the augmented state vector s_k. Which is, precisely, the vector we want to estimate. In laymen words: we need to know the results of the estimation in order to select the noise coefficients that will help us estimate! Although some kind of recursive procedure can be thought of, by means of which the user could start with an arbitrary value and through several iterations try to obtain convergence towards "optimal" values, the approach seems rather elaborate and without guarantees of being successful. This fact suggests that time would be better invested looking first at some other simpler approaches for determining the "good" noise coefficients.

Dual Estimation

Given the difficulties found in trying to determine *a priori* the coefficients in the noise matrix C_* by using independent means, one might think of estimating simultaneously those coefficients and the system parameters. For that purpose, starting with the estimation equations shown in Chapter 4

$$x_{k+1} = E_* x_k + \tilde{G}_*(w_k)r + C_* e_k$$
$$0 = E_0 x_k + \tilde{G}_0(w_k)r - J w_k + e_k \tag{6.118}$$

letting

$$\tilde{C}_*(e_k) = diag\left(e_{k,}^T e_{k,}^T \ldots, e_{k,}^T\right) \tag{6.119}$$

be a n x $(n$ x $p)$ block diagonal matrix, then

$$C_* e_k = \tilde{C}_*(e_k)\eta_k \tag{6.120}$$

with η_k a column vector containing the parameters c_{ijk} in C_*. Using equation 6.120, and since the system is assumed time-invariant, the departing expression can be transformed into

$$s_{k+1} = \tilde{F}_k s_k$$

$$y_k = \tilde{H}_k s_k + e_k \tag{6.121}$$

with

$$s_k = \left[x_k^T, r_k^T, \eta_k^T\right]^T$$

$$\tilde{H}_k = \left[E_0 \tilde{G}_0(w_k)\ 0\right]$$

$$F_k = \begin{bmatrix} E_* & \tilde{G}_*(w_k) & \tilde{C}_*(e_k) \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \tag{6.122}$$

Now, the estimation of the augmented state s_k includes the state system parameters and the noise parameters. Unfortunately, as e_k is not directly measurable, it has to be estimated from equation 6.121, namely

$$\hat{e}_k = y_k - \tilde{H}_k \hat{s}_k = y_k - E_0 \hat{x}_k - \tilde{G}_0(w_k)\hat{r}_k \tag{6.123}$$

where $\hat{x}_k$ and $\hat{r}_k$ are the estimates of x_k and r_k at time k. Now it should be evident that the estimation problem has become non-linear, because when using $\hat{e}_k$ the matrix $\tilde{F}_k$ which is a function of e_k, has become a function of s_k [8].

One of the reasons for selecting a Singular Pencil approach in this research was the desire to avoid the problems encountered when trying to apply the Extended Kalman Filter for identification purposes in an industrial environment. By getting back to a non-linear problem for which the EKF seems the most suitable solution, we have defeated the benefits of the Singular Pencil and returned to the departing point. Perhaps a totally different strategy for the problem of the *a priori* knowledge of the coefficients in the noise matrix C_* is in order.

6.1.2 An Alternative

A detailed look at the results of Chapter 5 regarding the appropriateness of the coefficients in the noise matrix C_* seems to bring an alternative to the problem of their *a priori* selection. In Figure 5.42 it is evident that all the choices for C_* which were evaluated share a common characteristic: they

make the estimated coefficient $\mathbf{d_1}$ reach its maximum dimension when the dynamics are suspected to be fastest, i.e., very close to or at the point of screen failure. Instead of placing the attention on how to obtain the best possible estimate for the noise coefficients, perhaps an approach suitable for any "good enough" estimate would be preferable. Rather than dealing with intricate algorithms that will give us the "smoothest" trajectory of $\mathbf{d_1}$ values for the different operating points, it would be better to implement a simple indicator that can be applied to the calculated values of $\mathbf{d_1}$, even if the choice of C_* for obtaining them was not the best.

When looking at the shape of the variations in the dimension of $\mathbf{d_1}$ for the different screen operating points, the idea of using some sort of cumulative function of this variable immediately arises. It is a well known fact in the FDI literature that cumulative sum charts can "damp out" noise and "amplify" true changes in the process [28]. They also can add some sense of "threshold" for the results obtained. Intuitively, several alternatives appear to have enough potential for detecting changes in the nature of $\mathbf{d_1}$:

1. Successive averages.

2. Successive differences.

3. Range of two successive pairs of values from the expected value.

In the first case, the cumulative sum would be given by the expression

$$\frac{1}{n}\sum_{i=1}^{n} d_{1i} \tag{6.124}$$

where i represents the operating point at which the estimation of $\mathbf{d_1}$ has been done, and n the total number of operating points evaluated. In the second case, the function would be given by the expression

$$f(i) = d_{1i} - d_{1(i-1)}; \quad i = 1,\ldots,n \tag{6.125}$$

with i and n being the same as for the first case. In the last case, the function is given by

$$f(i) = [d_{1i} - E\{d_1\}]; \quad i = 1,\ldots,n \tag{6.126}$$

with i and n as explained above.

97

Successive Averages of d1 Coefficient for Different C* Matrices
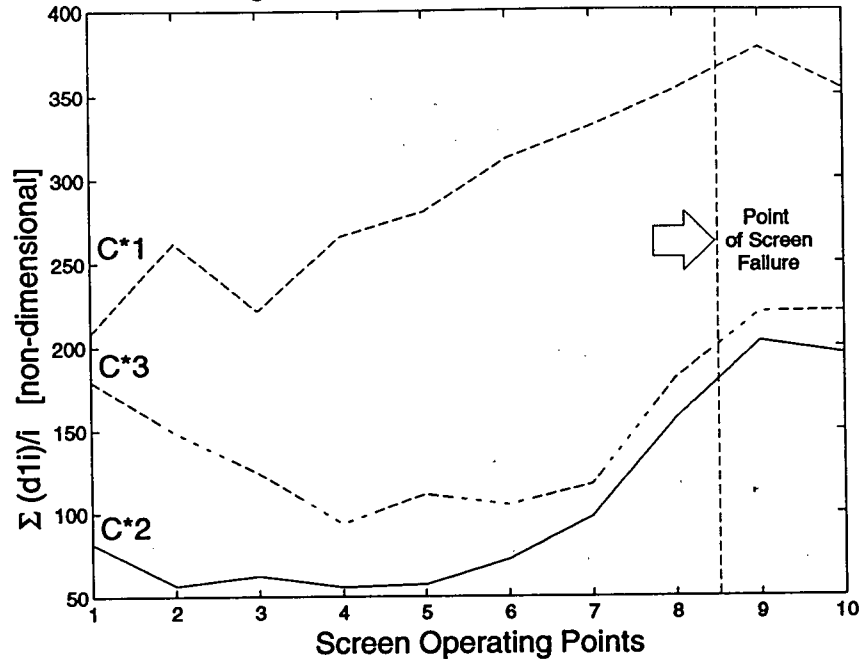
Figure 6.54: Coefficient d1 Cumulative Averages for Different Noise Matrix Choices

The usefulness of all three alternatives was tested against the d_1 estimation results obtained in the pilot plant trials. If we apply the first criterion, the results obtained are as shown in Figure 6.54. Regardless of the noise model chosen, an increasingly growing magnitude of d_1 seems to be a good indication of blinding. Two consecutive increases could be used as a practical limit on the achievable screen throughput in the mill. From the curves it is evident that the better suited the noise model, the larger the throughput allowed. In other words, when using averages, the better suited the noise model used in the estimation is, the bolder our decision making process (regarding screen operating points) can be. The best model (C_{*2}) would allow for a throughput around 4,700 Liters per Minute, whereas the poorest one (C_{*1}) would call for stopping further increases when reaching the 4,500 L/Min mark. All models will allow for avoidance of plugging, i.e., screen failure.

When we apply the second criterion, the results are displayed in Figures 6.55, 6.56, and 6.57. In this case, regardless of the noise model chosen, a good indication of imminent failure could be the starting dimensional value of d_1. Whenever a difference reaches this level, blinding can be considered as present and the throughput should be adjusted accordingly. Once again, the best suitable noise

98
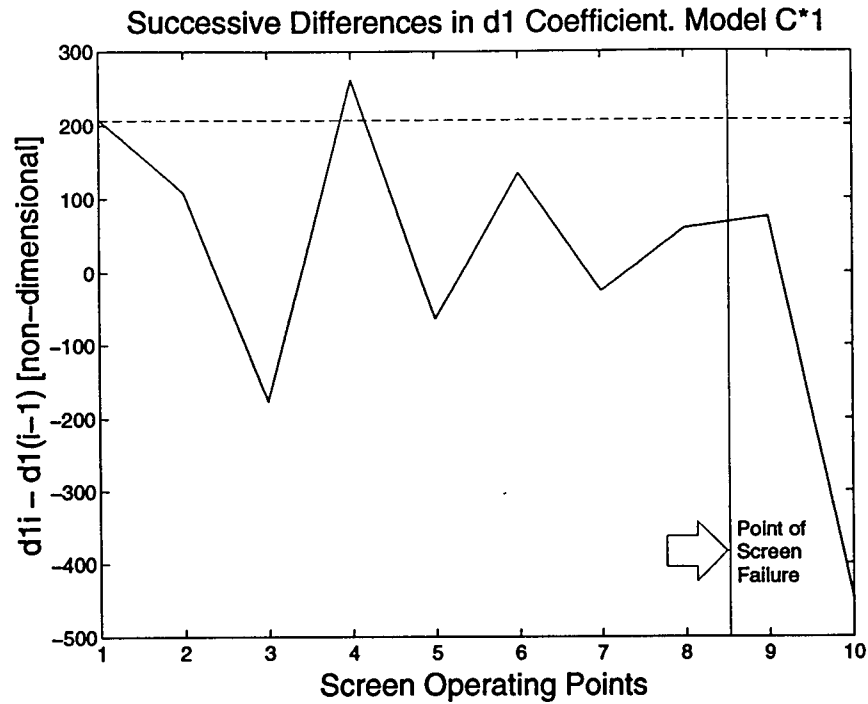
Figure 6.55: Successive Differences in d1 Coefficient. Model C*1
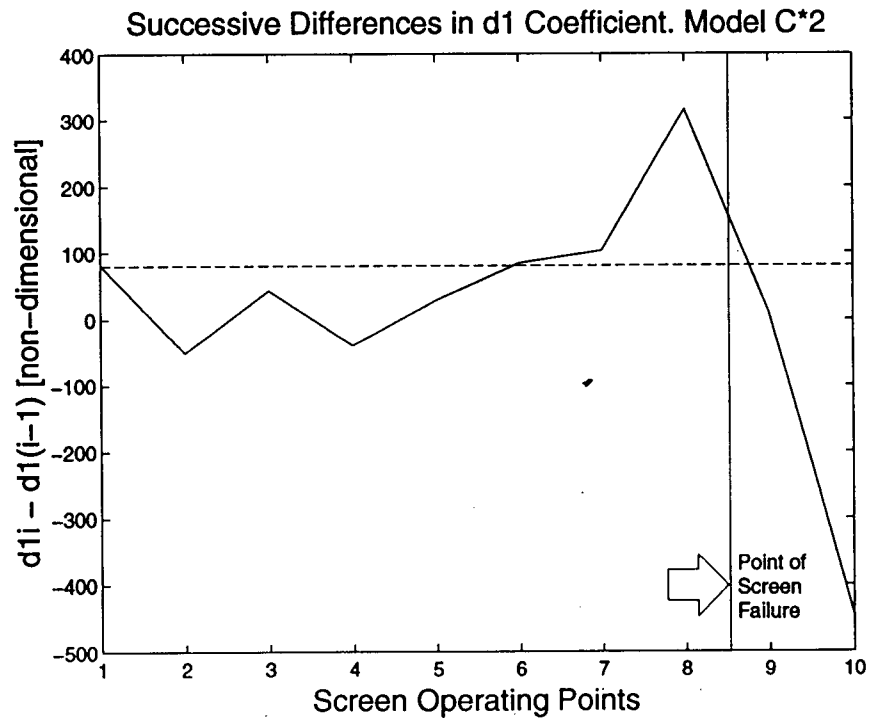
Figure 6.56: Successive Differences in d1 Coefficient. Model C*2
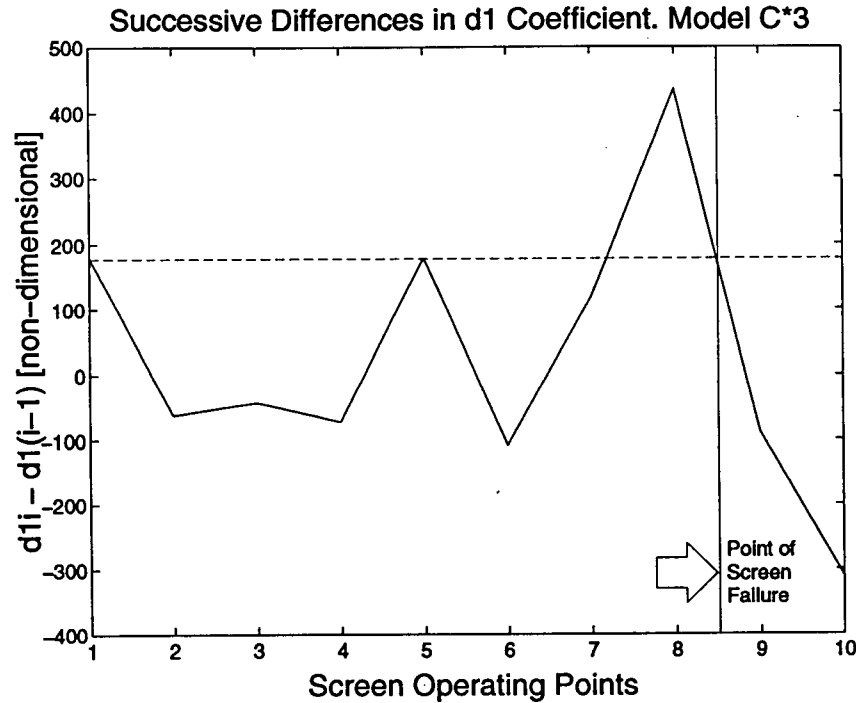
Figure 6.57: Successive Differences in d1 Coefficient. Model C*3

model (C_{*2}) seem to allow for bigger throughput margins in the decision making process than the poorest one (C_{*1}).

The results obtained when applying the last criterion can be seen in Figure 6.58. In this case, the guidelines to be observed are not as clear cut as in the previous cases. Nonetheless, a combination of rules could be applied to spell the imminence of failure: number of times the starting zero point is exceeded and/or number of consecutively increasing values of d_1. Again, the best adjusted noise coefficients render the best results.

This discussion on cumulative functions could be extended. There are, of course, many other cumulative sum charts suitable for application. It is not the intention to cover them all here. However, what must be underlined is the rationale for using this type of "statistical" criterion to infer information on screen fault detection:

a. All methods are simple to implement. They are very well suited for the typical mill environment.
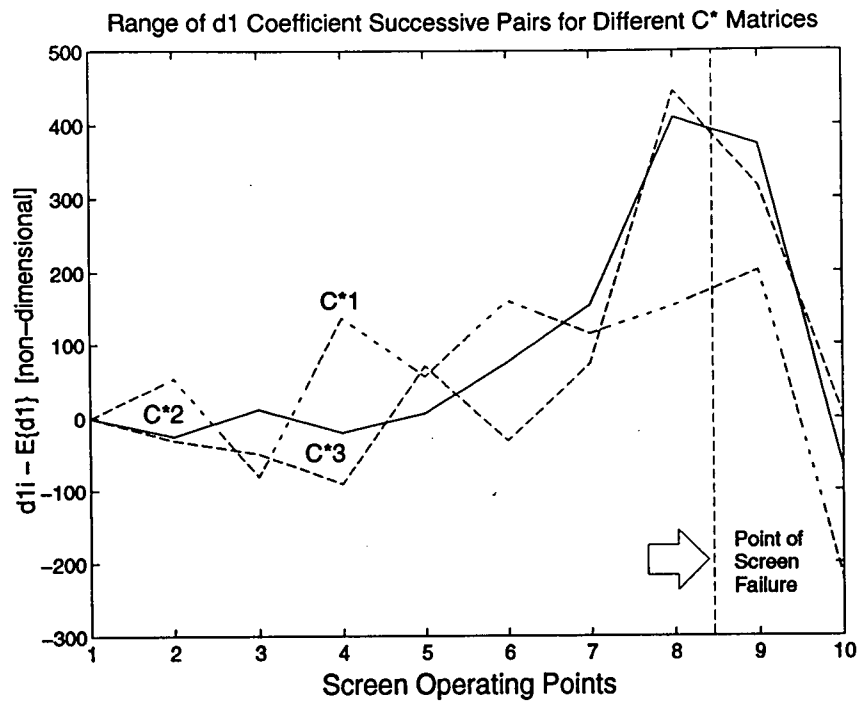
Figure 6.58: Cumulative Range of Coefficient d1 for Different Noise Matrix Choices

b. Using proper engineering judgement, it is possible to overcome the complicated problem of estimating the noise matrix elements, without affecting the ability of the singular pencil approach to give us valuable information on the screen faults.

c. The decision making process becomes independent of the individual values that the *d* coefficients may display, adding some sense of "threshold" that clearly displays whether the screen is operating normally or is under a fault.

In conclusion, it seems that a suitable FDI strategy for industrial pressure screens should incorporate the use of some "good-enough" approximated model for the noise matrix C_* (which may be found by means of a limited amount of trial and error testing) and the application of cumulative functions criteria. Successive averages and successive differences are two of those functions which seem particularly promising for such task.

6.2 Other Identification Techniques

6.2.1 Bootstrapping

During the course of this research the need for combined estimation of the parameters and the states included in our dynamic screen model arose. As mentioned in previous chapters, an strategy based on the Extended Kalman Filter was not considered advisable for FDI purposes, due to the well known fact that the estimates obtained with such method often diverge. Instead, we decided to use the SPM approach which has been described so far. However, there were some other methods at hand when this decision was made. One of the most promising at the time seemed to be "Bootstrapping". The main problem with "Bootstrapping" was that there has been an on going discussion in the literature on whether it gives biased estimates or not, and some articles have appeared which proved that, depending on the means used to estimate the parameters and the states, it indeed produces biased estimates [2], [62]. Under those circumstances, the Singular Pencil Matrix alternative looked as a more advisable choice.

The main idea behind the "Bootstrapping" approach is to split the nonlinear problem of the product of unknowns: model parameters and states, into two linear problems solved recursively. The procedure is carried out using the certainty equivalence principle. First, the estimation of the states takes place, assuming the parameters as known. Second, the parameters are estimated using the values obtained during the estimation of the states. Third, the states are estimated again, using the new estimates of the parameters, and so on and so forth, until both parameters and states achieve converging estimated values [3].

If new developments in the literature find a way for the problem of bias to be worked out, given the fact that this technique does not require any transformation of the estimates, its use would imply that some computational speed advantages could be achieved. Another advantage is that little (or none at all) *a priori* knowledge of the system would be required. Perhaps this is a good topic to focus future research on, and a *vis a vis* comparison with the dual estimation of parameters and noise using SPM plus the Extended Kalman Filter would be of interest.

6.2.2 New Developments

From the time the choice of identification technique was done for FDI on pressure screens, some work has been brought to light on how to deal with the simultaneous estimation of parameters and states. Besides documents dealing with applications of the Extended Kalman Filter, some other approaches have shown up in the literature on the subject. Some of these publications seem promising and well worth investigating rigorously in forthcoming research, although, due to time constraints, only a cursory overview is going to be offered here.

Kamas and Sanders have proposed the use of an adaptive observer for the estimation of both quantities. However, the convergence properties of their method are subject to strict conditions on the system [38]. It remains to be seen if those conditions can be applied to pressure screens. Du and Brdys introduced an Extended Luenberger Observer for estimation on Induction Motor Drives. The approach is deterministic and uses pre-filtering of the noise [13]. Given the noisy environment found when operating the screens, it is not simple to speculate on the benefits of such approach. Instead of concerning themselves with the dual estimation problem itself, Liu et al. have decided to focus the attention on the bias of the parameters and a way to detect its presence [44]. This is an intriguing approach which could be further extended for FDI. Oshman has presented the use of a Maximum Likelihood algorithm based on a square root filter and its derivatives for the estimation of the states and their parameter-based sensitivity functions [52]. By his own admission the method involves very expensive computations which may require a parallel processor. This obviously imposes a burden on its industrial application. Sproesser and Gissinger have used MIMO transfer functions and an improved version of the Recursive Least Squares which inclûdes a filter for the calculation of states [63], to perform FDI on sensors. The usefulness of their approach is limited, though, as the states are simply the derivatives of the measurable signals in the model. As it has been said, in the case of the pulp screens, the states are non-measurable. Finally, Feyo de Azevedo et al. have dealt with simulation of deterministic non-linear MIMO systems which are of limited applicability [10].

6.3 Summary

In this chapter some topics related to the estimation of the states and parameters in the screen

model have been presented. First, the difficulties involved in using analytic techniques for modelling the noise matrix present in the SPM model have been explored. Then, simpler alternatives based on cumulative functions have been explained and its usefulness has been shown. Finally, a cursory overview on several identification methods and its importance as future topics of research has been outlined.

<div align="center">

Chapter 7

Summary and Future Research

</div>

7.1 Summary

The preceding pages have presented the findings on the application of Fault Detection and Isolation (FDI) techniques to pulp pressure screens. The motivation of this work, all along, has been to find a technique suitable for implementation in an industrial mill with the potential to offer the operator advanced warnings of deviations in the.desired performance of the screens. For that purpose, the following methodology was used:

☐ First, the operation of pulp screens in industry was studied. The current techniques for detection of faults were presented and analyzed. It was established that no real FDI is performed in industry and that screen faults are not dealt with until they cause the device to become inoperative.

☐ Second, the FDI methods present in the literature were investigated. Their features were evaluated and their potential for application to pressure screens was screened.

☐ Third, an experimental setting to test the different FDI approaches was chosen. It involved the use of a fully instrumented pilot plant, capable of simulating industrial operation and flexible enough to allow for the creation of faulty conditions at will. The simpler non-model based approaches which showed promise were implemented and evaluated. The results gave clear indications that such techniques were not capable of giving advanced warning of faults.

☐ Fourth, a model-based approach was decided upon. It demanded the creation of a dynamic model of the typical pressure screen and the application of identification techniques to estimate the parameters and variables in the model. As the estimation turned out to be non-linear, a technique called the Singular Pencil Matrix (SPM) was implemented to deal with this problem. Its benefits and shortcomings were explored.

☐ Fifth, the model-based strategy was put to test. Computer simulations were conducted and the results analyzed. Pilot plant tests followed and their results confirmed the previous simulations. It was established that the physical parameters included in the screen model show strong potential

<div align="center">

105

</div>

for their use as an FDI technique in the mill. A simplified version of the model was then tested to explore further gains of the model-based approach in computing expense. In this case, the experimental outcome did not match the results of the full model.

☐ Finally, the role of noise modelling in the application of the Singular Pencil Matrix approach was explored and some conclusions were established. A cursory overview of some recent alternatives appearing in the literature for the non-linear estimation was attempted, and techniques with potential for further research were identified.

Every one of the steps outlined above bred findings and conclusions which have been explained in all the previous chapters of this thesis. For the benefit of the reader the most important ones will be repeated here.

- Fault Detection and Isolation (FDI) involves the early diagnosis of deviations, also called faults, in the expected behavior of a system.

- A typical pressure screen can suffer several possible faults. The most damaging and, therefore, the most worthy candidate for early detection is blinding. Blinding is caused by the accumulation of fibres in the screen plate orifices.

- Two non-model based approaches for FDI were attempted during this research. They involved monitoring the screen motor load and calculating the slope of the curve given by the squared accept flow versus the differential pressure. Although these methods are simpler to use, the results obtained when using them were not satisfactory.

- It was decided that an approach based on a model was indicated and, for that, a model has been proposed. The pressure screen model designed for FDI is based on mechanistic principles and an expression analogous to the D'Arcy-Weisbach principle to account for screen dynamics.

- By monitoring the d coefficients in the model, which represent the frictional changes in the screen, the experimental results showed that faults can be detected. In particular, blinding can be associated with changes in the *accept friction* coefficient d_1.

- When using the SPM technique, modelling the screen noise is an important issue and must be dealt with judiciously. Although the analytic procedures for this task are not simple, a

combination of selection of C_* through limited trial and error and use of cumulative functions can be used advantageously.

The major contribution of this work is, to the best of my knowledge, to be the very first work which has ever been written on the topic of Fault Detection and Isolation for pressure screens, and one of a few (in the profuse literature published on the subject) dealing with industrial applications of FDI. Having said that, some other contributions are:

- The introduction of faults in pulp pressure screen as a separate subject of study.
- The derivation of the first dynamic model in the literature to characterize pressure screen behavior.
- The development and experimental implementation of the first model-based approach for fault detection on pressure screens.
- The use of the Singular Pencil Matrix technique for dealing with non-linear estimation problems in a true practical application.
- The enhancement of knowledge on pressure screen operation and faults, in general, and the gain of valuable physical insight for better screen control.

7.2 Future Work

It is very difficult to address all the issues that arise while doing research. Besides the obvious need for focus, which forces one to put aside some unanswered questions that may appear to be of marginal importance for the achievement of the main goal or goals, time limitations are always a component to be reckoned with. This thesis is no exception and, therefore there are some topics which hopefully will be dealt with in future research, namely:

- <u>Pipe plugging</u>. Due to the difficulties in duplicating pulp-induced plugging in piping, most results presented in this thesis concentrated on the most important cause of screen failure: blinding. Nevertheless, it would be of interest to observe the behavior of the coefficients in the model under the presence of reject pipe plugging. Industrial simulation of such plugging could be achieved by means of altering the mechanical configuration of the piping used in the pilot plant.

107

- <u>Valve dynamics</u>. The screen model proposed in the research includes the device dynamics only. As such, it is intended to sit atop a screen control strategy, to be determined by the user. For such strategy to be successful, the control valve dynamics used in conjunction with the screen must be incorporated. The problem of incorporating the accept and reject valves dynamics into the screen model has not been addressed here.

- <u>Dual estimation of noise and physical coefficients</u>. Given the new developments in the literature regarding the Extended Kalman Filter, it would be of interest to study the performance of simultaneous dual estimation of the noise and physical coefficients present in the screen model which has been developed here. A detailed evaluation of its convergence and computing expense have appeal for its examination.

- <u>Use of new estimation techniques</u>. The application of other techniques different from the EKF to the non-linear estimation problem is also of interest. A comparison *vis a vis* the dual estimation using the EKF seems another interesting avenue to be explored.

References

[1] *Instrument Engineers Handbook.* B. G. Liptak, editor-in-chief, and K. Venczel, associate editor. Chilton Book Company, Revised edition, 1985.

[2] M.S. Ahmed. On bootstrap estimation of system parameters and states. *IEEE Transactions on Automatic Control, vol. AC–28, No. 7, July,* 1983.

[3] M.S. Ahmed and N. Sait. State-space adaptive control through a modified bootstrap algorithm for parameter and state estimation. *IEE Proceedings, vol. 136, Pt. D., No. 5, September,* 1989.

[4] M. Aoki. *State Space Modelling of Time Series.* Springer-Verlag, Berlin, 2nd. edition, 1990.

[5] J. D. Aplevich. *Implicit Linear Systems.* Springer-Verlag, 1991.

[6] R. G. Brown and P. Y. C. Hwang. *Introduction to Random Signals and Applied Kalman Filtering.* John Wiley and Sons, Inc., 2nd. edition, 1992.

[7] R. E. Carvill. High-consistency pressurized pulp screening. *Tappi Journal, vol. 51, No. 9, pp. 104A-107A,* 1968.

[8] Y.C. Chen, J.D. Aplevich, and W.J. Wilson. Simultaneous estimation of state variables and parameters for multivariable linear systems with singular pencil models. *IEE Proceedings, Vol. 133, Pt. D, No. 2, March,* 1986.

[9] R. L. Daugherty and J. B. Franzini. *Fluid Mechanics, Sixth Edition.* McGraw-Hill Book Company, 1965.

[10] S. Feyo de Azevedo, P. Pimenta, F. Oliveira, and E. Ferreira. Studies on on-line state and parameter estimation through a real-time process simulator. In *Proceedings of the IFAC Conference on Artificial Intelligence in Real-Time Control (Delft, The Netherlands),* 1992.

[11] M. M. Denn. *Process Modeling.* Pitman Publishing Inc., 1986.

[12] R. H. Dieck. *Measurement Uncertainty. Methods and Applications.* Instrument Society of America, 2nd. printing, 1995.

[13] T. Du and M. A. Brdys. Implementation of extended Luenberger observers for joint state and parameter estimation of PWM induction motor drive. In *Proceedings 5th Conference on Power Electronics and Applications (Brighton, U.K.)*, 1993.

[14] D. P. Dumdie. A systems approach to pressure screen control in the pulp mill. In *Tappi Journal, v. 74, No. 11, pp. 97–102*, 1991.

[15] L. Estévez-Reyes and G. A. Dumont. A new approach towards prevention of blinding in pulp screens. In *Advances in Instrumentation and Control, v. 50, Pt. 1, pp. 31–36*, 1995.

[16] R. M. Felder and R. W. Rousseau. *Elementary Principles of Chemical Processes*. John Wiley and Sons, Inc., 1986.

[17] G. Ferrara, U. Preti, and G.D. Schena. Phenomenological model of the hydrocyclone: Model development and verification for single-phase flow. *International Journal of Mineral Processing, vol. 22, No. 1–4, April*, 1988.

[18] P. M. Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – a survey and some new results. *Automatica, Vol. 26, No. 3, pp. 459–474*, 1990.

[19] J. Gertler. Analytical redundancy methods in fault detection and isolation. survey and synthesis. *Proceedings of the IFAC Fault Detection, Supervision and Safety for Technical Processes, pp. 9–21*, 1991.

[20] R. W. Gooding. The passage of fibres through slots in pulp screening. Master's thesis, The University of British Columbia, 1986.

[21] R. W. Gooding and D. F. Craig. The effect of slot spacing on pulp screen capacity. Technical report, PPR 830. PAPRICAN, 1990.

[22] R. W. Gooding and D. F. Craig. The effect of slot spacing on pulp screen capacity. *Tappi J., v. 75, No. 2, pp. 71–75*, 1992.

[23] R.W. Gooding and R.J. Kerekes. Derivation of performance equations for solid-solid screens. *Canadian Journal of Chemical Engineering, vol. 67, No. 5, October, pp. 801–805*, 1989.

[24] C. W. Helstrom. *Probability and Stochastic Processes for Engineers*. Macmillan Publishing Co., 1991.

[25] J. Hill, H. Höglund, and E. Johnsson. Evaluations of screens by optical measurements. *Tappi Journal, v. 58, No. 10, pp. 120–124*, 1975.

[26] J. Hill, T. Pettersson, and S. Rydefalk. The stfi long-fibre content meter and its process applications. In *International Mechanical Pulping Conference, vol. IV, session 39*, 1977.

[27] D. M. Himmelblau. *Basic Principles and Calculations in Chemical Engineering*. Prentice-Hall, Inc., 1974.

[28] D. M. Himmelblau. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*. Elsevier Scientific Publishing Co., 1978.

[29] A. W. Hooper. Screening of mechanical pulp — effect of control parameters and screen operation on pulp quality. In *Pulp and Paper Manufacture, Vol. 2, pp. 181*, 1987.

[30] A. W. Hooper. The screening of chemical pulp. In *Pulp and Paper Manufacture, Vol. 5*, 1989.

[31] R. Isermann. Process fault detection based on modeling and estimation methods – a survey. *Automatica, Vol. 20, No. 4, pp. 387–404*, 1984.

[32] R. Isermann. *Digital Control Systems, Vol. I*. Springer-Verlag, Berlin, 1989.

[33] R. Isermann. Fault diagnosis of machines via parameter estimation and knowledge processing. *Proceedings of the IFAC Fault Detection, Supervision and Safety for Technical Processes, pp. 43–55*, 1991.

[34] R. Isermann and B. Freyermuth. Process fault diagnosis based on process model knowledge. *Proceedings of the 1989 ASME Intl. Computers in Eng. Conference, pp. 631–642*, 1989.

[35] Y. Ishida. An application of qualitative reasoning to process diagnosis: Automatic rule generation by qualitative simulation. *Proceedings of the IEEE Fourth Annual Conference on Artificial Intelligence Applications, pp. 124–129*, 1988.

[36] D. Janigan. Operating experience of an innovative secondary fiber system. In *76th Annual Meeting, Technical Section Canadian Pulp and Paper Association, Preprints 'B'*, 1990.

[37] R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, New Jersey, 2nd. Ed., 1988.

[38] L. A. Kamas and S. R. Sanders. Parameter and state estimation in power electronic circuits. *IEEE Transactions on Circuits and Systems, vol. 40, No. 12, December*, 1993.

[39] A. Kärnä, H. Liimatainen, and H. Rusanen. Latest development trends in controlling the mechanical groundwood process. *Tappi J., v. 71, No. 7, pp. 87–92*, 1988.

[40] A. Kaukonen, H. Koivo, P. Virtanen, and P. Sumanen. Modelling and simulation of a paper machine's wet end process. In *Modelling and Simulation in Engineering, B. Walstrom and K. Leiviska Editors*, 1986.

[41] M. Kramer and F. E. Finch. Development and classification of expert systems for chemical process fault diagnosis. *Robotics & Computer-Integrated Manufacturing, Vol. 4, No. 3/4, pp. 437–446*, 1988.

[42] M.A. Kramer and J.A. Leonard. Diagnosis using backpropagation neural networks. Analysis and criticism. *Computers & Chemical Engineering, Vol. 14, No. 12, pp. 1323–1338*, 1990.

[43] D. F. Lehman. Pulp screening. In *Handbook of Pulp and Paper Technology, pp. 209*, 1970.

[44] W. E. Liu, F. F. Wu, and S. Lun. Estimation of parameters errors from measurement residuals in state estimation. *IEEE Transactions on Power Systems, vol. 7, No. 1, February*, 1992.

[45] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Inc, 1987.

[46] X. Lou, A.S. Willsky, and G.C. Verghese. Optimally robust redundancy relations for failure detection in uncertain systems. *Automatica, Vol. 22, No. 3, pp. 333–344*, 1986.

[47] A. C. Martin. Screen operation and maintenance. In *Pulp and Paper Manufacture, Vol. 2*, 1987.

[48] R. J. McGill. *Measurement and Control in Papermaking, pp. 380*. Adam Hilger Ltd., 1980.

[49] R. B. Michaelson. A summary of thermomechanical pulping plant advanced control applications. *Instrumentation in the Pulp and Paper Industry, I.S.A., v. 17*, 1978.

[50] G. L. Nelson. The screening quotient: a better index for screening performance. *Tappi Journal, vol. 64, No. 5, pp. 133–134, May*, 1981.

[51] L. W. Nelson and E. Stear. The simultaneous on-line estimation of parameters and states in linear systems. *IEEE Transactions on Automatic Control, vol. AC–21, No. 2, February*, 1976.

[52] Y. Oshman. Maximum likelihood state and parameter estimation via derivatives of the v-lambda filter. *Journal of Guidance, Control, and Dynamics, vol. 15, No. 3, May-June*, 1992.

[53] R. Patton, P. Frank, and R. Clark. *Fault Diagnosis in Dynamic Systems*. Prentice-Hall International, 1989.

[54] L.F. Pau. *Failure Diagnosis and Performance Monitoring*. Marcel Derrer, 1981.

[55] P. L. Poole and N. C. S. Chari. Experiences with a distributed control system in pulp mill operations. In *Tappi Engineering Conferences*, 1983.

[56] K. Årzén. *Realization of Expert System Based Feedback Control*. PhD thesis, Dept. of Automatic Control, Lund Institute of Technology, 1987.

[57] J. A. Roberson and C. T. Crowe. *Engineering Fluid Mechanics, Third Edition*. Houghton Miffling Company, 1985.

[58] P. M. Schaffrat. Pulp screening and cleaning. In *Pulp and Paper Science and Technology, Vol. 1, pp. 394*, 1962.

[59] W. A. Shewhart. *Statistical Method from the viewpoint of quality control*. Dover Publications, Inc., New York, 1986.

[60] G. A. Smook. *Handbook for Pulp and Paper Technologists*. Angus Wilde Publication, Vancouver, B.C., Canada, 1990.

[61] G. A. Smook. *Handbook of Pulp and Paper Terminology*. Angus Wilde Publication, Vancouver, B.C., Canada, 1990.

[62] D. Sprevak and M. M. Newton. Comments on "Bootstrap estimation of parameters and states of linear multivariable systems". *IEEE Transactions on Automatic Control, vol. AC–25, No. 1, February*, 1980.

[63] T. Sproesser and G. L. Gissinger. A method for fault detection using parameter and state estimation. In *Proceedings of the IFAC Conference on Artificial Intelligence in Real-Time Control (Delft, The Netherlands)*, 1992.

[64] C. S. Travers. Mill operating experience with a Beloit S Screen. In *74th Annual Meeting, Technical Section Canadian Pulp and Paper Association, Preprints 'B'*, 1988.

[65] S. Tzafestas. Second generation diagnostic expert systems: Requirements, architectures and prospects. *Proceedings of the IFAC Fault Detection, Supervision and Safety for Technical Processes, pp. 327–332*, 1991.

[66] S. Tzafestas, editor. *Applied Control. Current trends and modern methodologies*. Marcel Dekker, Inc., 1993.

[67] X. G. Wang. *Model Structure Selection for On-line System Identification Using Overlapping Singular Pencil Models*. PhD thesis, University of Waterloo, 1991.

[68] L. E. Wells. Advanced distributed multiple loop multiple stage stock screening control. In *1993 Tappi/ISA PUPID Process Control Conference, Nashville, Tenn.*, March 21–25, 1993.

[69] A.S. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica, Vol. 12, No. 6, pp. 601–611*, 1976.

[70] B.M. Wise, N.L. Ricker, D.F. Veltkamp, and B.R. Kowalski. Theoretical basis for the use of a principal component models for monitoring multivariate processes. *Process Control and Quality, Vol. 1, No. 1, pp. 41–51*, 1990.

[71] P. Young. Parameter estimation for continuous-time models—a survey. *Automatica, Vol. 17, No. 1, pp. 23–39*, 1981.

Appendix A Screen Modeling

A.1 Mass Balances

The fundamental relationship in a mass balance is expressed as

$$\text{Rate of mass accumulation} = \text{R. of mass in} - \text{R. of mass out} \tag{A.127}$$

In our case, three equations can be written. The first two are for the water in the slurry and the fibers, namely

$$\frac{d(m_{H_2O})}{dt} = \dot{M}_{F_{H_2O}} + \dot{M}_{\text{dilution}} - \dot{M}_{A_{H_2O}} - \dot{M}_{R_{H_2O}}$$
$$\frac{d(m_{sol})}{dt} = \dot{M}_{F_{sol}} - \dot{M}_{A_{sol}} - \dot{M}_{R_{sol}} \tag{A.128}$$

and the third one, representing the total mass, will come from adding these two. The total mass in the screen will be the sum of the water and the fibers inside it

$$M_{\text{Total}} = M_{H_2O} + M_{\text{solids}} \tag{A.129}$$

Using the definitions of *consistency* (C)

$$C = \frac{M_{\text{Solids}}}{M_{\text{Solids}} + M_{\text{Water}}} \tag{A.130}$$

and density

$$\rho = \frac{M}{V} \tag{A.131}$$

we arrive to

$$M_{\text{solids}} = \rho C V_{\text{Total}} \tag{A.132}$$

and

$$M_{H_2O} = \rho(1 - C)V_{\text{Total}} \tag{A.133}$$

115

Assuming that neither the consistencies nor the densities change too fast with time, the balances will finally become

$$\frac{d(V_{H_2O})}{dt} = \frac{1}{\rho_{H_2O}}[a_1 F_F - a_2 F_A - a_3 F_R] + F_{\text{dilution}}$$

$$\frac{d(V_{sol})}{dt} = \frac{1}{\rho_{sol}}[a_4 F_F - a_5 F_A - a_6 F_R]$$ (A.134)

$$\frac{d(V_{\text{Total}})}{dt} = F_F - F_A - F_R + F_{\text{dilution}} = 0$$

with all consistencies approximately equal and

$$a_1 = \rho_F(1 - C_F)$$

$$a_2 = \rho_A(1 - C_A)$$

$$a_3 = \rho_R(1 - C_R)$$ (A.135)

$$a_4 = \rho_F C_F$$

$$a_5 = \rho_A C_A$$

$$a_6 = \rho_R C_R$$

A.2 Pulp Screen Energy Balances

The fundamental relationship in an energy balance can be expressed as

Rate of energy accumulation = R. energy in − R. energy out

+ Heat + Work (A.136)

and

Energy = Kinetic + Potential + Internal (A.137)

As ours is an open system where no phase changes or reactions occur, and since our process is adiabatic, i.e. $Q = 0$, the starting point for an energy balance is the expression

$$\frac{d[U + KE + PE]}{dt} = \sum_{\text{input streams}} E_j - $$

$$\sum_{\text{output streams}} E_j + W$$ (A.138)

which can be written as

$$\frac{d[U + KE + PE]}{dt} = \sum_{\text{input streams}} \dot{M_i}\left[\frac{v_i^2}{2g_c} + \frac{g}{g_c}h_i + \underline{U_i}\right] -$$

$$\sum_{\text{output streams}} \dot{M_i}\left[\frac{v_o^2}{2g_c} + \frac{g}{g_c}h_o + \underline{U_o}\right] + W \qquad \text{(A.139)}$$

The work done on the system can be divided into hydraulic ($W_h = PV$) and shaft work (W_s) and the equation becomes

$$\frac{d[\text{Energy}]}{dt} = \sum_{\text{input streams}} \dot{M_i}\left[\frac{v_i^2}{2g_c} + \frac{g}{g_c}h_i + \underline{U_i} + \frac{P_i}{\rho}\right] -$$

$$\sum_{\text{output streams}} \dot{M_o}\left[\frac{v_o^2}{2g_c} + \frac{g}{g_c}h_o + \underline{U_o} + \frac{P_o}{\rho}\right] + W_s \qquad \text{(A.140)}$$

Our sources of incoming energy are the feed flow and the dilution flow. Our sources of outgoing energy are the accepts and rejects flows. The changes in the internal energy of this system can be accounted for in terms of the temperature of the system materials and, since the process is adiabatic, all thermal variations are due to friction. Establishing our rejects as our reference in height, i.e. $h_R = 0$, and using the fact that

$$v = \frac{F}{A} \qquad \text{(A.141)}$$

the above expression becomes

$$\frac{d[\text{Energy}]}{dt} = \sum_{\text{input streams}} \left[\rho_i \frac{F_i^3}{2g_cA_i^2} + \rho_iF_i\frac{g}{g_c}h_i + F_iP_i\right] -$$

$$\sum_{\text{output streams}} \left[\rho_o\frac{F_o^3}{2g_cA_o^2} + \rho_oF_o\frac{g}{g_c}h_o + F_oP_o\right] + \text{Frictional losses} + W_s \qquad \text{(A.142)}$$

which turns out to be

$$\frac{d[\text{Energy}]}{dt} = e_1\dot{V}_F^3 + e_2\dot{V}_F + \dot{V}_FP_F + e_3\dot{V}_{\text{Dil}}^3 + e_4\dot{V}_{\text{Dil}} + \dot{V}_{\text{Dil}}P_{\text{Dil}}$$

$$-e_5\dot{V}_A^3 - e_6\dot{V}_A - \dot{V}_AP_A - e_7\dot{V}_R^3 - \dot{V}_RP_R + W_s + k_A + k_R \qquad \text{(A.143)}$$

where k refers to the frictional losses inside the screen.

The left hand of the above equation is

$$\frac{d}{dt}[U + KE + PE] = \frac{d}{dt}\left(M\hat{U}\right) + \frac{d(KE)}{dt} + \frac{d(PE)}{dt} \qquad \text{(A.144)}$$

117

and the influence of the last two derivatives can be considered negligible, as the changes inside the screen are minimal. We can also safely assume that:

1. Temperature and composition of the system content do not vary with position within the system.

2. No phase changes or chemical reactions occur within the system.

3. Heat capacities do not vary with time.

4. No heat is added to the system by direct means.

therefore,

$$\frac{d[\text{Energy}]}{dt} = \frac{dU}{dt} = \frac{d\left(M\hat{U}\right)}{dt} = \frac{d\left(\rho V \hat{U}\right)}{dt} \tag{A.145}$$

since neither the density nor the Volume of the screen change, the above equation becomes

$$\frac{dU}{dt} = \rho V \frac{d\hat{U}}{dt} = \rho V C_v \frac{dT}{dt} \tag{A.146}$$

leading to the expression

$$\frac{dT}{dt} = f_1 F_F^3 + f_2 F_F + f_3 F_F P_F + f_4 F_{\text{Dil}}^3 + f_5 F_{\text{Dil}} + f_3 F_{\text{Dil}} P_{\text{Dil}}$$

$$-f_6 F_A^3 - f_7 F_A - f_3 F_A P_A - f_8 F_R^3 - f_3 F_R P_R + f_3 W_s + f_3 k_A + f_3 k_R \tag{A.147}$$

which, since $\frac{dT}{dt}$ varies very slowly when compared with the flows and pressures, becomes

$$0 = f_1 F_F^3 + f_2 F_F + f_3 F_F P_F + f_4 F_{\text{Dil}}^3 + f_5 F_{\text{Dil}} + f_3 F_{\text{Dil}} P_{\text{Dil}}$$

$$-f_6 F_A^3 - f_7 F_A - f_3 F_A P_A - f_8 F_R^3 - f_3 F_R P_R + f_3 W_s + f_3 k_A + f_3 k_R \tag{A.148}$$

with

$$f_1 = \frac{1}{2A_F^2 g_c V C_v}$$

$$f_2 = \frac{g(h_F - h_R)}{g_c V C_v}$$

$$f_3 = \frac{1}{g_c V C_v} \tag{A.149}$$

$$f_4 = \frac{1}{2A_{\text{Dil}}^2 g_c \rho V C_v}$$

$$f_5 = \frac{g(h_{\text{Dil}} - h_R)}{g_c \rho V C_v}$$

$$f_6 = \frac{1}{2A_A^2 g_c V C_v}$$

$$f_7 = \frac{g(h_A - h_R)}{g_c V C_v} \tag{A.150}$$

$$f_8 = \frac{1}{2A_R^2 g_c V C_v}$$

118

A.3 Mechanical Losses

The pipe-friction equation for flows of incompressible fluids in pipes is as follows

$$k_L = f\frac{L}{D}\frac{v^2}{2g} \tag{A.151}$$

with

$$
\begin{aligned}
&L : \text{length}\\
&D : \text{diameter}\\
&v : \text{velocity}\\
&g : \text{gravity} .\\
&f : \text{friction factor}\\
&k_L : \text{head loss}
\end{aligned}
\tag{A.152}
$$

it is obvious that the above is a static relationship. When dealing with a piece of pipe able to change its head losses dynamically, that equation is not good.

If we lump all terms referred to the piping into the term f_p and differentiate the D'Arcy-Weisbach's equation, namely,

$$\frac{dk}{dt} = \frac{d\left(f_p v^2\right)}{dt} \tag{A.153}$$

we obtain the following expression

$$\frac{dk}{dt} = v^2\frac{df_p}{dt} + f_p\frac{d\left(v^2\right)}{dt} \tag{A.154}$$

which, when using the fact that $k = f_p v^2$ gives us the result

$$\frac{dk}{dt} = v^2\frac{df_p}{dt} + \frac{k}{v^2}\frac{d\left(v^2\right)}{dt} \tag{A.155}$$

This shows that a dynamic expression for the losses inside the screen must include a term dependent on the square of the fluid velocity and another term dependent on the losses themselves. Following this line of reasoning, a dynamic relationship to account for the behavior of such device is proposed as

$$\frac{dk_L}{dt} = c_1 k_L + c_2 v^2 \tag{A.156}$$

119

The constant

$$c_1 = \frac{1}{v^2} \frac{d(v^2)}{dt} \qquad (A.157)$$

should be fairly sensitive to variations in the speed and, by implication, to flow changes. As the flow becomes smaller one should expect to see it grow in magnitude, fairly rapidly. The constant c_2 would reflect the rate of change of the parameter f_p which accounts for the internal characteristics of the piping. Under steady-state conditions this relationship reduces to

$$-c_1 k_L = c_2 v^2 \qquad (A.158)$$

or

$$k_L = -\frac{c_2}{c_1} v^2 \qquad (A.159)$$

Making

$$-\frac{c_2}{c_1} = f \frac{L}{D2g} \qquad (A.160)$$

our equation becomes the pipe friction equation.

What pipe fitting would change its head losses dynamically? Observations and empirical relations say that a pressure screen could be thought of as such a device. Since the screen has two diverting outgoing flows, it could be seen as some sort of "T" fitting with different losses: one for the rejects and another for the accepts line. Therefore, two equations would describe the losses in the screen

$$\frac{dk_A}{dt} = c_1 k_A + c_2 v_A^2$$
$$\frac{dk_R}{dt} = c_3 k_R + c_4 v_R^2 \qquad (A.161)$$

Blinding would be responsible for the dynamical behavior of the head losses.

As velocity is difficult to measure and seldom measured, we can substitute it by its equivalent expression

$$v = \frac{F}{A} \qquad (A.162)$$

and our equations become

$$\frac{dh_A}{dt} = c_1 h_A + c_5 F_A^2$$
$$\frac{dh_R}{dt} = c_3 h_R + c_6 F_R^2 \qquad (A.163)$$

with

$$F : \text{volumetric flow}$$

$$A : \text{Area of pipe}$$

$$c_5 = \frac{c_2}{A_A^2}$$

$$c_6 = \frac{c_4}{A_R^2}$$

(A.164)

A.4 Linearization

Using a Taylor series expansion and truncating with first order terms, we'll get

$$f(x,y) \approx f(\bar{x},\bar{y}) + \left.\frac{\partial f}{\partial y}\right|_{\cdot \bar{x},\bar{y}}(y - \bar{y}) + \left.\frac{\partial f}{\partial x}\right|_{\cdot \bar{x},\bar{y}}(x - \bar{x})$$

(A.165)

for our energy equation, the derivatives are

$$\left.\frac{\partial f}{\partial F_F}\right| = 3f_1 \bar{F}_F^2 + f_2 + f_3 \bar{P}_F$$

$$\left.\frac{\partial f}{\partial F_{\text{Dil}}}\right| = 3f_4 \bar{F}_{\text{Dil}}^2 + f_5 + f_3 \bar{P}_{\text{Dil}}$$

$$\left.\frac{\partial f}{\partial F_A}\right| = -3f_6 \bar{F}_A^2 - f_7 - f_3 \bar{P}_A$$

$$\left.\frac{\partial f}{\partial F_R}\right| = -3f_8 \bar{F}_R^2 - f_3 \bar{P}_R$$

(A.166)

$$\left.\frac{\partial f}{\partial P_F}\right| = f_3 \bar{F}_F$$

$$\left.\frac{\partial f}{\partial P_{\text{Dil}}}\right| = f_3 \bar{F}_{\text{Dil}}$$

$$\left.\frac{\partial f}{\partial P_A}\right| = -f_3 \bar{F}_A$$

$$\left.\frac{\partial f}{\partial P_R}\right| = -f_3 \bar{F}_R$$

(A.167)

$$\left.\frac{\partial f}{\partial W_s}\right| = f_3$$

$$\left.\frac{\partial f}{\partial k_A}\right| = f_3$$

$$\left.\frac{\partial f}{\partial k_R}\right| = f_3$$

(A.168)

and using deviation variables, i.e., $f' = f - \bar{f}$ the equation becomes

$$0 = g_1 F_F' + g_2 P_F' + g_3 F_{\text{Dil}}' + g_4 P_{\text{Dil}}'$$

$$-g_5 F_A' - g_6 P_A' - g_7 F_R' - g_8 P_R' + f_3 W_s' + f_3 k_A' + f_3 k_R'$$

(A.169)

121

with

$$g_1 = \frac{3\rho \bar{F}_F^2 + 2A_F^2 \rho g(h_F - h_R) + 2A_F^2 g_c \bar{P}_F}{2A_F^2 \rho g_c V C_v}$$

$$g_2 = \frac{3\bar{F}_{\text{Dil}}^2 + 2A_{\text{Dil}}^2 g(h_{\text{Dil}} - h_R) + 2A_{\text{Dil}}^2 g_c \bar{P}_{\text{Dil}}}{2A_{\text{Dil}}^2 \rho g_c V C_v}$$

$$g_3 = \frac{3\rho \bar{F}_A^2 + 2A_A^2 \rho g(h_A - h_R) + 2A_A^2 g_c \bar{P}_A}{2A_A^2 \rho g_c V C_v} \qquad \text{(A.170)}$$

$$g_4 = \frac{3\rho \bar{F}_R^2 + 2A_R^2 g_c \bar{P}_R}{2A_R^2 \rho g_c V C_v}$$

$$g_5 = \frac{\bar{F}_F}{\rho V C_v}$$

$$g_6 = \frac{\bar{F}_{\text{Dil}}}{\rho V C_v}$$

$$g_7 = \frac{\bar{F}_A}{\rho V C_v} \qquad \text{(A.171)}$$

$$g_8 = \frac{\bar{F}_R}{\rho V C_v}$$

$$f_3 = \frac{1}{\rho V C_v}$$

and the P''s and F''s are deviation variables from the set points $\bar{P}'$ and $\bar{F}'$.

Using the same approach for the remaining frictional losses equations, we'll get

$$\frac{dk_A'}{dt} = d_1 k_A' + d_2 F_F' + d_2 F_{\text{Dil}}' - d_2 F_R' \qquad \text{(A.172)}$$

$$\frac{dk_R'}{dt} = d_3 k_R' + d_4 F_R$$

with

$$d_1 = c_1$$

$$d_2 = \frac{2c_2 \bar{F}_A}{A_A^2}$$

$$d_3 = c_3 \qquad \text{(A.173)}$$

$$d_4 = \frac{2c_4 \bar{F}_R}{A_R^2}$$

A.5 Degrees of Freedom

From our chosen model we have three equations that represent the energy balance and the frictional losses. We also have six parameters: ρ, C_v, d_1, d_2, d_3, d_4; and eleven variables: flows, pressures, shaft work and frictional losses. From the formula for the degrees of freedom

$N_F = N_V - N_E$ it is found that we must identify eight variables as inputs and the remaining as outputs or states.

A.6 State Space Representation

As two of the three potential outputs in our model can not be measured directly (k_A and k_R), an input-output representation seems flawed for identification purposes. Instead, a state space representation approach in which the states aren't measurable is proposed. Making

$$x_1 = k'_A$$
$$x_2 = k'_R$$

$\qquad$ (A.174)

our system model becomes

$$\underline{\dot{x}} = A\underline{x} + B\underline{u}$$
$$y = C\underline{x} + D\underline{u}$$

$\qquad$ (A.175)

with

$$A = \begin{bmatrix} d_1 & 0 \\ 0 & d_3 \end{bmatrix}$$

$$B = \begin{bmatrix} d_2 & -d_2 & d_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & d_4 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\qquad$ (A.176)

$$C = \begin{bmatrix} h_8 & h_8 \end{bmatrix}$$

$$D = \begin{bmatrix} h_1 & -h_3 & h_2 & h_4 & -h_6 & -h_7 & h_5 & h_8 \end{bmatrix}$$

$\qquad$ (A.177)

and

$$\underline{x}^T = \begin{bmatrix} h_A & h_R \end{bmatrix}$$

$$\underline{u}^T = \begin{bmatrix} F'_F & F'_R & F'_{\text{Dil}} & P'_F & P'_A & P'_R & P'_{\text{Dil}} & W'_s \end{bmatrix}$$

$\qquad$ (A.178)

$$y = F'_A$$

This system is <u>controllable</u>, provided $d_4 \neq 0$ (which seems to be the case), and <u>observable</u>, provided $d_1 \neq d_3$ (which also seems the case), which makes it suitable for use of Singular Pencil modelling and identification[16].

---

[16] Let us take a closer look at these two conditions:

1. $d_4 = 0$ can mean two things, zero reject flow (a condition which would not be present in normal screen operation) or $c_4 = 0$. But $c_4 = 0$ would make the rate of losses independent from the flow, which goes against the hydrodynamic theory.

2. Under the light of our empirical observations regarding the friction coefficients experimental values, the condition $d_1 = d_3$ is not likely. This is backed by the knowledge on friction factors found in the literature. In the pipe-friction equation, these coefficients depend on the internal geography of the piping. As such, the frictional coefficients on uneven branches of a "T" (due to size or internal contour) would hardly be equal.

Appendix B   Singular Pencil Representation

B.1 Discrete Model

Given a continuous time state-space model

$$\underline{\dot{x}} = A\underline{x} + B\underline{u}$$

$$y = C\underline{x} + D\underline{u}$$

(B.179)

the equivalent discrete time representation will be of the form

$$\underline{x}((k+1)T) = G(T)\underline{x}(kT) + H(T)\underline{u}(kT)$$

$$y(kT) = C\underline{x}(kT) + D\underline{u}(kT)$$

(B.180)

with

$$G(T) = e^{AT}$$

$$H(T) = \left(\int_0^T e^{A\lambda}d\lambda \right) B$$

(B.181)

and $\lambda = T - t$. Our discrete matrices then become

$$G(T) = \begin{bmatrix} j_1 & 0 \\ 0 & j_2 \end{bmatrix}$$

$$H(T) = \begin{bmatrix} j_3 & -j_3 & j_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & j_4 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(B.182)

$$C = \begin{bmatrix} h_8 & h_8 \end{bmatrix}$$

$$D = \begin{bmatrix} h_1 & -h_3 & h_2 & h_4 & -h_6 & -h_7 & h_5 & h_8 \end{bmatrix}$$

(B.183)

with

$$j_1 = e^{c_1 T}$$

$$j_2 = e^{c_3 T}$$

(B.184)

$$j_3 = \frac{d_2}{c_1}\left(e^{c_1 T} - 1\right)$$

$$j_4 = \frac{d_4}{c_3}\left(e^{c_3 T} - 1\right)$$

$$h_1 = \frac{g_1}{g_3} = \frac{3\rho\left(\bar{F}_F^2/A_F^2\right) + 2\rho g(h_F - h_R) + 2g_c\bar{P}_F}{3\rho\left(\bar{F}_A^2/A_A^2\right) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A}$$

$$h_2 = \frac{g_2}{g_3} = \frac{3\left(\bar{F}_{Dil}^2/A_{Dil}^2\right) + 2g(h_{Dil} - h_R) + 2g_c\bar{P}_{Dil}}{3\rho\left(\bar{F}_A^2/A_A^2\right) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A}$$

$$h_3 = \frac{g_4}{g_3} = \frac{3\rho\left(\bar{F}_R^2/A_R^2\right) + 2g_c\bar{P}_R}{3\rho\left(\bar{F}_A^2/A_A^2\right) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A}$$

$$h_4 = \frac{g_5}{g_3} = \frac{2g_c\bar{F}_F}{3\rho\left(\bar{F}_A^2/A_A^2\right) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A}$$

(B.185)

124

$$h_5 = \frac{g_6}{g_3} = \frac{2g_c\bar{F}_{Dil}}{3\rho(\bar{F}_A^2/A_A^2) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A}$$

$$h_6 = \frac{g_7}{g_3} = \frac{2g_c\bar{F}_A}{3\rho(\bar{F}_A^2/A_A^2) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A}$$

$$h_7 = \frac{g_8}{g_3} = \frac{2g_c\bar{F}_R}{3\rho(\bar{F}_A^2/A_A^2) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A} \tag{B.186}$$

$$h_8 = \frac{f_3}{g_3} = \frac{2g_c}{3\rho(\bar{F}_A^2/A_A^2) + 2\rho g(h_A - h_R) + 2g_c\bar{P}_A}$$

B.2 Singular Pencil Matrix

As explained before, if we have the input-output representation of a system, it can be brought as a singular pencil model as follows

$$P(z)\begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = \begin{bmatrix} E_* - zI & -A_* & B_* & C_* \\ E_0 & -A_0 & B_0 & I \end{bmatrix}\begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = 0 \tag{B.187}$$

In our case, since we already have a state space representation, the matrix $P(z)$ will look as

$$P(z) = \begin{bmatrix} A - zI & 0 & B \\ C & -I & D \end{bmatrix} \tag{B.188}$$

with A, B, C, and D as defined in the previous section, i.e.,

$$P(z) = \begin{bmatrix} j_1 - z & 0 & 0 & j_3 & -j_3 & j_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & j_2 - z & 0 & 0 & j_4 & 0 & 0 & 0 & 0 & 0 & 0 \\ h_8 & h_8 & 1 & h_1 & -h_3 & h_2 & h_4 & -h_6 & -h_7 & h_5 & h_8 \end{bmatrix} \tag{B.189}$$

from here, it takes only a few linear transformations on the matrix (row operations) to bring it to a canonical representation that will allow the separation of the parameters and the states for identification purposes.

Our departing equation can also be expressed as

$$P(z)\begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} E_* - \lambda I_n & G_* \\ E_0 & G_0 \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix} = 0 \tag{B.190}$$

125

with

$$E_* = A$$

$$E_0 = C$$

$$G_* = [0 \quad B]$$

(B.191)

$$G_0 = [-I_1 \quad D]$$

and the operator z replaced by the algebraic indeterminate λ. First, we use a transformation $T \Rightarrow x = T^{-1}x'$ and, consequently,

$$\begin{bmatrix} TE_*T^{-1} - \lambda I_n & TG_* \\ E_0T^{-1} & G_0 \end{bmatrix} \begin{bmatrix} x' \\ w \end{bmatrix} = \begin{bmatrix} E'_* - \lambda I_n & G'_* \\ E'_0 & G_0 \end{bmatrix} \begin{bmatrix} x' \\ w \end{bmatrix} = 0$$

(B.192)

To get T we select a chained basis from the matrix

$$M = \begin{bmatrix} E_0(E_*)^n \\ \vdots \\ E_0E_* \\ E_0 \end{bmatrix}$$

(B.193)

which, in our case, will be

$$M = \begin{bmatrix} h_8j_1^2 & h_8j_2^2 \\ h_8j1 & h_8j_2 \\ h_8 & h_8 \end{bmatrix}$$

(B.194)

a basis is obviously $\{(h_8 \ h_8), (h_8j_1 \ h_8j_2)\}$ and, therefore, the transformation becomes

$$T = \begin{bmatrix} h_8j_1 & h_8j_2 \\ h_8 & h_8 \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} \frac{1}{h_8(j_1-j_2)} & \frac{-j_2}{h_8(j_1-j_2)} \\ \frac{-1}{h_8(j_1-j_2)} & \frac{j_1}{h_8(j_1-j_2)} \end{bmatrix}$$

(B.195)

which will give the new matrices

$$E'_* = \begin{bmatrix} j_1 + j_2 & -j_1j_2 \\ 1 & 0 \end{bmatrix}$$

(B.196)

$$E'_0 = [0 \quad 1]$$

$$G'_* = \begin{bmatrix} 0 & h_8j_1j_3 & h_8(j_2j_4 - j_1j_3) & h_8j_1j_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_8j_3 & h_8(j_4 - j_3) & h_8j_3 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(B.197)

$$G'_0 = [-1 \quad h_1 \quad -h_3 \quad h_2 \quad h_4 \quad h_5 \quad -h_6 \quad -h_7 \quad h_8]$$

which constitute the basic canonical form.

126

A second transformation is required, using a similarity transformation $\widehat{T}$, where

$$\widehat{T} = I_n + \begin{bmatrix} f\left(E_*^{11}\right) & \cdots & f\left(E_*^{1p}\right) \\ \vdots & & \vdots \\ f\left(E_*^{p1}\right) & \cdots & f(E_*^{pp}) \end{bmatrix} \tag{B.198}$$

and $f()$ is a matrix function that produces the following result

$$f(A) = I_n + \begin{bmatrix} 0 & -a_{11} & -a_{12} & \cdots & -a_{1(n-1)} \\ 0 & 0 & -a_{11} & \cdots & -a_{1(n-2)} \\ 0 & 0 & 0 & & \\ \vdots & \cdots & & & \vdots \end{bmatrix} \tag{B.199}$$

where $f(A)$ is m x n. To apply transformation $\widehat{T}$, we need to determine the observability indices of the system (n_i), as well as its positive integers (ξ_k). The observability indices (n_i with $i = 1, \ldots, p$ and p: number of rows in E_0) give the smallest integer j for which $E_{0i}E_*^j$ is linearly dependent on the inferior rows in M. In our case, there is only one observability index: $n_1 = 2$, and that satisfies the proposition $\sum_{i=1}^{p} n_i = n$ with n: number of rows in E_*. As for the positive integers ξ_k (with $k = 1, \ldots, p$), they are the column indices of the left most linearly-independent columns in G_0. In our case, $\{\xi_k\} = \{1\}$. As we must apply $f()$ on E_*^{ii}, it is important to notice that the latter is a n_i-square companion matrix, sub-block of the block matrix

$$\begin{bmatrix} (J'T)E_*(J'T)^{-1} \\ E_0(J'T)^{-1} \end{bmatrix} = \begin{bmatrix} E_*^{11} & \cdots & E_*^{1p} \\ \vdots & & \vdots \\ E_*^{p1} & \cdots & E_*^{pp} \\ E_0^{11} & \cdots & E_0^{1p} \\ \vdots & \bullet & \vdots \\ E_0^{p1} & \cdots & E_0^{pp} \end{bmatrix} \tag{B.200}$$

with

$$J'T = \begin{bmatrix} E_{01}E_*^{n_1-1} \\ \vdots \\ E_{01} \\ E_{02}E_*^{n_2-1} \\ \vdots \\ E_{0p} \end{bmatrix} \tag{B.201}$$

127

In our case, it is obvious that $J'T = T$ and then $E_*^{ii} = E_*'$. Therefore,

$$\widehat{T} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -(j_1 + j_2) \\ 0 & 0 \end{bmatrix} \tag{B.202}$$

or equivalently,

$$\widehat{T} = \begin{bmatrix} 1 & -(j_1 + j_2) \\ 0 & 1 \end{bmatrix} \tag{B.203}$$

and the new transformed matrices are:

$$E_*'' = \begin{bmatrix} 0 & -j_1 j_2 \\ 1 & j_1 + j_2 \end{bmatrix} \tag{B.204}$$

$$E_0'' = \begin{bmatrix} 0 & 1 \end{bmatrix},$$

$$G_*'' = \begin{bmatrix} 0 & -h_8 j_2 j_3 & h_8(j_2 j_3 - j_1 j_4) & -h_8 j_2 j_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_8 j_3 & h_8(j_4 - j_3) & h_8 j_3 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{B.205}$$

$$G_0'' = \begin{bmatrix} -1 & h_1 & -h_3 & h_2 & h_4 & h_5 & -h_6 & -h_7 & h_8 \end{bmatrix}$$

Finally, a third transformation using matrices L and R such that

$$\begin{bmatrix} I & L \\ 0 & R \end{bmatrix} P(\lambda) = \begin{bmatrix} \widetilde{E_*} - zI & \widetilde{G_*} \\ \widetilde{E_0} & \widetilde{G_0} \end{bmatrix} \tag{B.206}$$

with the $\tilde{\ }$ matrices being the system canonical representation suitable for identification and R a non-singular matrix. In our case,

$$L = \begin{bmatrix} j_1 j_2 \\ -(j_1 + j_2) \end{bmatrix} \tag{B.207}$$

and $R = 1$ which gives the transformation matrix

$$T_3 = \begin{bmatrix} 1 & 0 & j_1 j_2 \\ 0 & 1 & -(j_1 + j_2) \\ 0 & 0 & 1 \end{bmatrix} \tag{B.208}$$

and the resulting matrices for the model

$$\widetilde{E_*} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{B.209}$$

$$\widetilde{E_0} = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

$$\widetilde{G_*} = \begin{bmatrix} -j_1 j_2 & j_2(h_1 j_1 - h_8 j_3) & h_8(j_2 j_3 - j_1 j_4) - h_3 j_1 j_2 & j_2(h_2 j_1 - h_8 j_3) & 0 & \dots & 0 \\ j_1 + j_2 & h_8 j_3 - h_1(j_1 + j_2) & h_8(j_4 - j_3) + h_3(j_1 + j_2) & h_8 j_3 - h_2(j_1 + j_2) & 0 & \dots & 0 \end{bmatrix}$$

$$\widetilde{G_0} = \begin{bmatrix} -1 & h_1 & -h_3 & h_2 & h_4 & h_5 & -h_6 & -h_7 & h_8 \end{bmatrix} \tag{B.210}$$

B.3 System Identification

The system can now be represented as

$$P(z)\begin{bmatrix} x_k \\ y_k \\ u_k \\ e_k \end{bmatrix} = \begin{bmatrix} \widetilde{E_*} - zI & -\widetilde{A_*} & \widetilde{B_*} & C_* \\ \widetilde{E_0} & -\widetilde{A_0} & \widetilde{B_0} & I \end{bmatrix}\begin{bmatrix} x'_k \\ y'_k \\ u'_k \\ e_k \end{bmatrix} = 0 \tag{B.211}$$

where

$$\begin{bmatrix} \widetilde{G_*} \end{bmatrix} = [-\widetilde{A_*} \quad \widetilde{B_*}]$$
$$\begin{bmatrix} \widetilde{G_0} \end{bmatrix} = [-\widetilde{A_0} \quad \widetilde{B_0}] \tag{B.212}$$

and e_k a zero mean white Gaussian random noise sequence with covariance Λ ($\Lambda > 0$).

The measurements y_k and u_k are known. The state vector is known and so are the parameters in $\widetilde{G_*}$ and $\widetilde{G_0}$. Making

$$w_k = [y_k^T, u_k^T]^T \tag{B.213}$$

and r a column vector containing the non-identically zero, non-pivot parameters a_{ijk} and b_{ijk} in $\widetilde{A_*}$, $\widetilde{A_0}$, $\widetilde{B_*}$, $\widetilde{B_0}$ enumerated in some specific order, the equation above can be put as

$$x_{k+1} = E_* x_k + \widehat{G_*}(w_k)r + C_* e_k$$
$$0 = E_0 x_k + \widehat{G_0}(w_k)r - Jw_k + e_k \tag{B.214}$$

where

$$\widehat{G_*}(w_k)r = [-\widetilde{A_*} \quad \widetilde{B_*}]w_k$$
$$\widehat{G_0}(w_k)r = [-\widetilde{A_0} \quad \widetilde{B_0}]w_k + Jw_k \tag{B.215}$$

and

$$J = [I_p \quad 0_{p \times m}] \tag{B.216}$$

a matrix which includes the pivot[17]. $\widehat{G_*}(w_k)$ and $\widehat{G_0}(w_k)$ are not unique and are constructed to isolate the unknown parameters a_{ijk} and b_{ijk} into the column vector r. As such

$$\widehat{G_*}(w_k) = \begin{bmatrix} F'_A & F'_F & F'_R & F'_{Dil} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & F'_A & F'_F & F'_R & F'_{Dil} & 0 & \cdots & 0 \end{bmatrix} \tag{B.217}$$

$$\widehat{G_0}(w_k) = [0 \quad \cdots \quad 0 \quad F'_F \quad F'_R \quad F'_{Dil} \quad P'_F \quad P'_A \quad P'_R \quad P'_{Dil} \quad W'_s]$$

---

[17] $y \in \Re^p$, $x \in \Re^n$, and $u \in \Re^m$

and

$$r = \begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} & g_{21} & \cdots & g_{24} & h_1 & \cdots & h_8 \end{bmatrix} \tag{B.218}$$

with the g_{ij} indicating the elements in row i and column j in $\widetilde{G_*}$. As the system is considered time-invariant,

$$r = r_k = r_{k+1} \tag{B.219}$$

and then, if

$$s_k = [x_k^T, r_k^T]^T \tag{B.220}$$

our working equations become

$$s_{k+1} = F_k s_k + D_* e_k$$

$$y_k = H_k s_k + e_k \tag{B.221}$$

with

$$H_k = [\widetilde{E_0} \quad \widehat{G_0}(w_k)]$$

$$F_k = \begin{bmatrix} \widetilde{E_*} & \widehat{G_*}(w_k) \\ 0 & I_l \end{bmatrix} \tag{B.222}$$

$$D_* = \begin{bmatrix} C_* \\ 0_{l \times n} \end{bmatrix}$$

and l is the dimension of the vector r. If the matrix C_* and the covariance Λ of $e_k \in \Re^p$ are assumed known, the Kalman filter algorithm can be applied to give the optimal linear estimate of the augmented state vector s_k. Knowing s_k, we will know the system state vector x_k and the parameter vector r_k. The recursive equations become

$$\hat{s}_{k+1} = F_k \hat{s}_k + K_k (y_k - H_k \hat{s}_k)$$

$$K_k = \left(F_k P_k H_k^T + S \right) \left(F_k P_k H_k^T + \Lambda \right)^{-1} \tag{B.223}$$

$$P_{k+1} = F_k P_k F_k^T + Q - K_k \left(F_k P_k H_k^T + \Lambda \right) K_k^T$$

where

$$S = \begin{bmatrix} C_* \Lambda \\ 0_{l \times p} \end{bmatrix}$$

$$Q = \begin{bmatrix} C_* \Lambda C_*^T & 0_{n \times l} \\ 0_{l \times n} & 0_{l \times l} \end{bmatrix} \geq 0 \tag{B.224}$$

with initial conditions

$$\hat{s}_0 = E[s_0]$$

$$P_0 = E\left[(s_0 - E[s_0])(s_0 - E[s_0])^T \right] \geq 0 \tag{B.225}$$

and $E[(.)]$ is the expected value of $(.)$