Second Annual Symposium, Mary Kay O'Connor Process Safety Center
"Beyond Regulatory Compliance: Making Safety Second Nature"
Reed Arena, Texas A&M University, College Station, Texas
October 30-31, 2001

# Issues in Fault Diagnosis and Isolation

**Rajaraman Srinivasan,**
W. Rogers, and M.S. Mannan
Mary Kay O'Connor Process Safety Center
Chemical Engineering Department
Texas A&M University
College Station TX 77843-3122

Phone: 979/862-3985
Email: srini@neo.tamu.edu

**ABSTRACT**

In the chemical and other related industries, there has been an impetus to produce higher quality products, to reduce rejection rates, limit down time, and to satisfy increasingly stringent safety and environment regulations. To meet higher standards, chemical processes depend on a large number of variables operating under closed loop conditions. Standard process controllers (PID controllers, model predictive controllers, etc.) are designed for steady state operating condition by compensating for small deviations in process variables. However, there occur some deviations that are unpermitted from usual or acceptable conditions. Hence, according to IFAC Technical committee: SAFEPROCESS (Fault Detection, Supervision and Safety for Technical Processes) has defined a fault as "An unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable/usual/standard condition."

This paper discusses generic approaches taken in fault detection and diagnosis in engineering systems, viz. robust model based techniques (residual generation, eigen-structure assignment, parameter estimation, and parity relation), stochastic methods (univariate analysis, partial least square technique, Fisher discriminant analysis, principal component analysis), knowledge based methods for decision making using neuro-fuzzy hybrid models, and various other approaches taken by research scientists, process engineers, and people in academia.

# Issues in Fault Diagnosis and Isolation

Rajaraman Srinivasan and M. Sam Mannan
Mary Kay O'Connor Process Safety Center

Chemical Engineering Department
Texas A&M University
College Station TX 77843-3122

**Abstract**

In the chemical and other related industries, there has been an impetus to produce higher quality products, to reduce rejection rates, limit down time, and to satisfy increasingly stringent safety and environment regulations. To meet higher standards, chemical processes depend on a large number of variables operating under closed loop conditions. Standard process controllers (PID controllers, model predictive controllers, etc.) are designed for steady state operating conditions by compensating for small deviations in process variables. However, there occur some deviations that are unpermitted from usual or acceptable conditions. Hence, according to IFAC Technical committee: SAFEPROCESS (Fault Detection, Supervision and Safety for Technical Processes) has defined a fault as "An unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable/usual/standard condition."

This paper discusses a model-based approach for fault detection and isolation (FDI) along with its application on a lab scale of a dynamic chemical process. In the end we enunciate the pros and cons of observer based state estimations of the process variables to detect abnormalities and suggest required diagnostic procedures.

## 1.0 Introduction

Modern chemical plants are large scale, particularly complex, and operate with a large number of variables under closed loop. Early and accurate fault detection and diagnosis for these plants can reduce downtime, increase the safety of plant operations, and reduce manufacturing costs. With chemical plants becoming highly instrument ridden and being cascaded in nature drives the need for accurate estimations of variables and interpret its variation under the influence of faults.

Any dynamic linear system can be modeled in a unique form called state space representation [Luenberger, 1979]. In the event the system is nonlinear, the former can be linearized around a favorable operating condition and the same modeling can be applied. Many sophisticated analytical procedures for control design are based on the assumption that the full state vector is available for measurement. These specify the current input value, as a function of the current value of the state vector, i.e., the control is a static function of the state. However in the actual system the entire state vector is not available for measurement [Luenberger, 1979]. In chemical processes, e.g., concentration of the desired component is inaccessible to measurement, and proper estimation of the variables becomes a key issue in model-based fault detection and isolation (FDI) [Fogler, 1992].

## 2.0 Background

### 2.1 Importance of fault diagnosis

Modern control systems are becoming more complex and control algorithms more sophisticated. Consequently, the issues of availability, cost efficiency, and reliability, operating safety, and environment protection are of major importance. For safety critical systems, the consequences of faults can be extremely serious in terms of human mortality, environmental impact, and economic loss. Therefore, there is a growing need for on-line supervision and fault diagnosis to increase the reliability of such safety-critical systems. Early indications concerning which faults are developing can help avoid system breakdown, mission abortion, and catastrophes. For systems that are not safety-critical, on-line fault diagnosis techniques can be used to improve plant efficiency, maintainability, availability, and reliability.

## 2.2 Fault diagnosis terminology

A " fault" is an unexpected change of system function [Isermann, 1984] although it may not represent physical failure or breakdown. Such a fault or malfunction hampers or disturbs the normal operation of an automatic system, thus causing unacceptable deterioration of system performance or even leading to dangerous situations. The term "failure" suggests a complete breakdown of a system component or a function, while the term fault may be used to indicate that a malfunction may be tolerable at its present stage.

A monitoring system, which is used to detect faults and diagnose their location and significance in a system, is called a " fault diagnosis system" [Jie Chen and Ron Patton, 1999]. The following entail the process fault diagnosis and isolation:

- **Fault detection**: to make a binary decision- either that something has gone wrong or that everything is fine.

- **Fault isolation**: to determine the location of the fault, e.g. which sensor or actuator has become faulty.

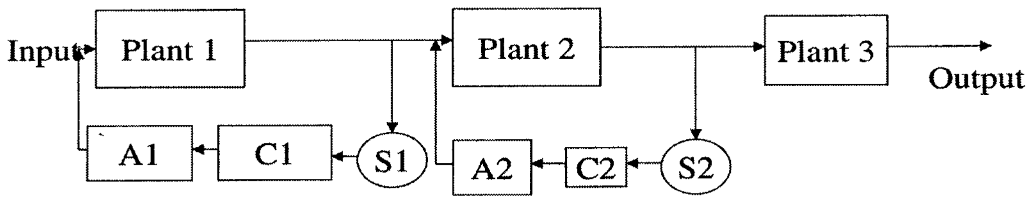
- **Fault identification**: to estimate the size and type or nature of the fault.

The relative importance of the three tasks is type dependent and situational, however detection is an absolute must for any practical system and isolation is almost equally important.

### 3 Model-based fault diagnosis

In practice, the most frequently used diagnosis method is to monitor the level (or trend) of a particular signal, and to take action when the signal reaches a given threshold. This method of limit checking although very simple, has serious drawbacks. The first drawback is the possibility of false alarms in the event of noise, input variations, and changes in operating point. The second and genuinely technical problem is that a single fault can cause multiple signals to exceed their limit, and hence fault isolation becomes very difficult, as shown in the Figure 1.

## Plant model




Legends:
S1: Sensor 1; C1: Control scheme 1; A1: Actuator 1
S2: Sensor 2 ;C2: Control scheme 2; A2: Actuator 2

Figure 1: General scheme of a cascaded plant

The model based technique or analytical redundancy as it is called uses the knowledge of the dynamic process to generate a signal for consistency checking of variables known as *residual signal*, as shown in Figure 2. In principle, the residual should be zero-valued when the system is normal, and should diverge from zero when the fault in the system occurs.

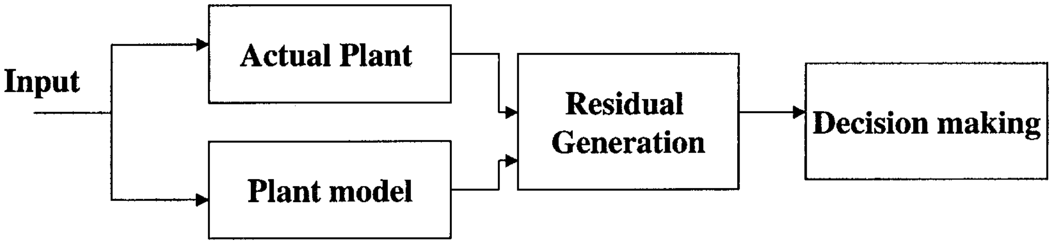



Figure 2: Residual generation scheme

Consistency checking through analytical redundancy is normally achieved through comparison of a measured signal with its estimation. The estimation is generated by a mathematical model of the system being considered. Hence, model-based fault diagnosis can be defined as the "*determination of faults of a system from the comparison of available system measurements with a priori information represented by the system's mathematical model, through generation of residual quantities and their analysis*". [Jie Chen and Ron Patton, 1999]. The major advantage of the model-based approach is that no additional hardware components are needed to realize an FDI algorithm, and hence it can be implemented in the software of the process control computer.

## 3.1 Model-based fault diagnosis methods

As discussed earlier model-based fault diagnosis can be defined as sequence of detection, isolation, and characterization of faults in components of a system by inferring the residual signal.

Faults are detected by setting a (fixed or variable) threshold on a residual quantity generated from the difference between the real measurements and estimates of these measurements using a mathematical model. There are mainly two stages in fault diagnosis procedure. These two main stages are described as follows:

(1) **Residual generation:** Its purpose is to generate a fault-indicating signal, which is generated with only the input and output of the monitored process. The residual

signal should be normally zero or close to zero when no fault is present, but is distinguishably different when a fault occurs. This means that the residual is characteristically independent of the system inputs and outputs, in ideal conditions. The algorithm used to generate residuals is called a *residual generator*.

(2) **Decision-making:** The residuals are examined for the likelihood of faults, and a decision rule is then applied to determine if any faults have occurred. A decision process may consist of a simple threshold test on the instantaneous values or moving averages of the residuals, or it may consist of methods of statistical decision theory, e.g., generalized likelihood ratio (GLR) testing or sequential probability ratio testing (SPRT), as shown in Figure 3.

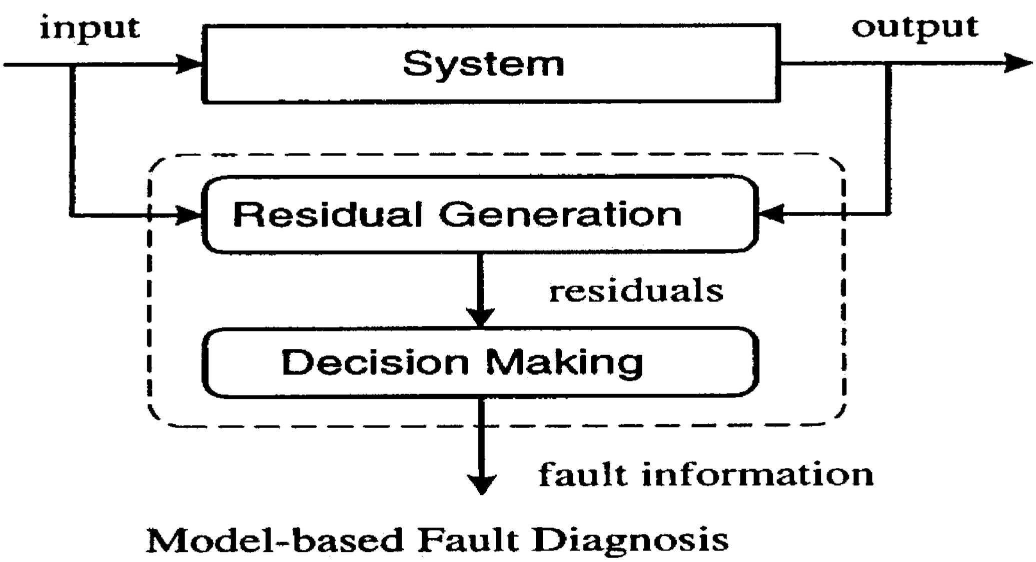



Figure 3: Model –based residual generation scheme

3.2 On-line fault diagnosis

The primary advantage of model-based fault diagnosis is its application in on-line fault diagnosis, which is carried during system operation. This is because the system input and output information required by the model-based FDI is available only when the system is in operation.

The information used for FDI is the measured output from sensors and the input to the actuators. The measured output is normally needed in the feedback control, whereas the input to the actuators is the required control action generated by the controller, which is normally implemented in the microprocessor.

The convenience of the model-based fault diagnosis is that the system model required in model-based FDI is the open-loop system model although we consider that the system is in the control loop, as shown in Figure 4. This is because the input and output information required in model-based FDI is related to the open-loop system. By the separation principle (O' Reilly, 1983) in modern control theory it is not necessary to consider the controller in the design of the fault diagnosis scheme.

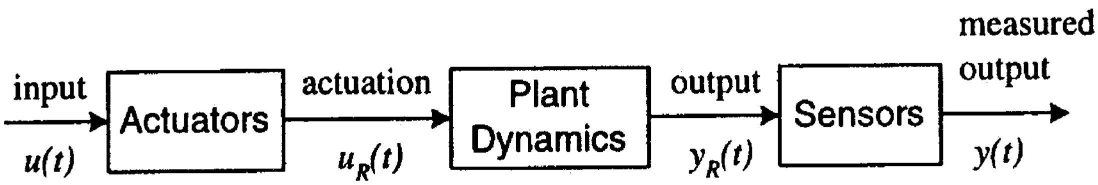



Figure 4: Open loop system

**4.0 <u>Modeling of faulty systems</u>**

The system dynamics, as shown in Figure 4, can be described by the state space model as:

$$x(t) = Ax(t) + Bu_R(t)$$
$$y_R = Cx(t) + Du_R(t)$$

Where $x \in \Re^n$ is the state vector, $u_R \in \Re^r$ is the input vector to the actuator, and $y_R \in \Re^m$ is the real system output vector: A, B, C, and D are known system matrices with appropriate dimensions. When the system has all possible sensor, component, and actuator faults (this is the most common situation to be considered), the system model is described as:

$$
\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) + Bf_a(t) + f_c(t) \\
y(t) &= Cx(t) + Du(t) + Df_a(t) + f_s(t)
\end{aligned}
$$

Considering the general case, a system with all possible faults can be described by the state space model as:

$$
\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) + R_1 f(t) \\
y(t) &= Cx(t) + Du(t) + R_2 f(t)
\end{aligned}
$$

Where $f(t)\ \varepsilon\ \Re^g$ is a fault vector, each element $f_i$ (t) (I = 1,2…g) corresponds to a specific fault.

### 4.1 Observer based residual generation

The main concern in FDI is the estimation of outputs using an observer, while the estimation of the state vector is unnecessary. It is desired to estimate a linear function of the state, i.e., $Lx(t)$, using a functional Luenberger observer (Luenberger, 1979) with the following structure.

$$
\begin{aligned}
\dot{z}(t) &= Fz(t) + Ky(t) + Ju(t) \\
w(t) &= Gz(t) + Ry(t) + Su(t)
\end{aligned}
$$

Where $z(t)\ \varepsilon\ \Re^q$ is the state vector of this observer, F, K, J, R, G, and S are matrices with appropriate dimensions. The output $w(t)$ of this observer is said to be an estimate of $Lx(t)$, for the system in an asymptotic sense if in the absence of faults: $w(t) \rightarrow Lx(t)$ as $t \rightarrow \infty$.

Now, the residual generator based on a generalized Luenberger observer to estimate a linear function of the state can be defined by the following equation:

$$z(t) = Fz(t) + Ky(t) + Ju(t)$$
$$r(t) = L_1 z(t) + L_2 y(t) + L_3 u(t)$$

When we apply the residual generator described above to the former system, the residual will be:

$$\dot{e}(t) = Fe(t) - TR_1(t) + KR_2 f(t)$$
$$r(t) = L_1 e(t) + L_2 R_2 f(t)$$

Where $e(t) = z(t) - Tx(t)$. It can be seen that the residual depends solely on the faults, and that in the absence of faults, $e(t)$ tends to zero as time proceeds thereby making $r(t)$ tend to zero. The simplest method in observer-based residual generation is to use a full order observer.

$$T = I \qquad L_1 = QC$$
$$F = A - KC \qquad L_2 = -Q$$
$$J = B - KD \qquad L_3 = QD$$

## 5.0 Illustrative example

To illustrate the main aspects of the proposed full order based observer for residual generation, a representative chemical reactor example is considered in this section. In particular, we consider an ideal continuous stirred tank reactor (CSTR) in non-isothermal operation, where the following exothermic irreversible reaction between sodium thiosulphate and hydrogen peroxide is taking place (Vejtasa and Schmitz, 1970; Fogler, 1992):

$$2Na_2S_2O_3 + 4H_2O_2 \rightarrow Na_2S_3O_6 + Na_2SO_4 + 4H_2O$$

By using the capital letters A, B, C, D, and E we denote the chemical compounds. $Na_2S_2O_3$, $H_2O_2$, $Na_2S_3O_6$, $Na_2SO_4$, $4H_2O$, respectively. The reaction kinetic law is reported in the literature to be (Vejtasa and Schmitz, 1970):

$$-rA = k(T)cAcB = ko\exp(-\frac{E}{RT})cAcB.$$

where $k(T)$ is the reaction rate constant, $k_o$ is the reaction frequency factor, $E$ is the activation energy, $R$ is the gas constant, $T$ is the temperature, and $c_A, c_B$ are the concentrations of the species A and B respectively. The overall mass and energy balance of the species A and B give the following non-linear dynamics.

$$\frac{dcA}{dt} = \frac{F}{V}(cA,in - cA) - 2k(T)cA^2$$

$$\frac{dT}{dt} = \frac{F}{V}(Tin - T) + 2\frac{(-\Delta H)R}{cP\rho}k(T)cA^2$$

The kinetic parameter values can be obtained from the pertinent literature (Vejtasa and Schmitz, 1970). Moreover, experimental investigations in the same paper show that there exists a multiplicity of steady states for a specific reaction system, as shown in Figure 5. The objective is to detect faults in the system discussed above at the middle unstable steady state with state feed back control law.

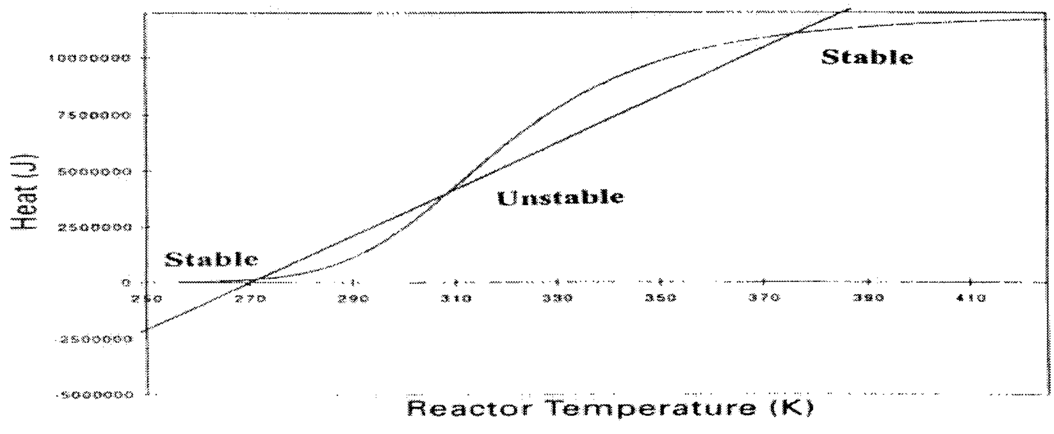



Figure 5: Multiplicity of steady states

The following Figures 6-10 show the result of the simulation on the above CSTR.

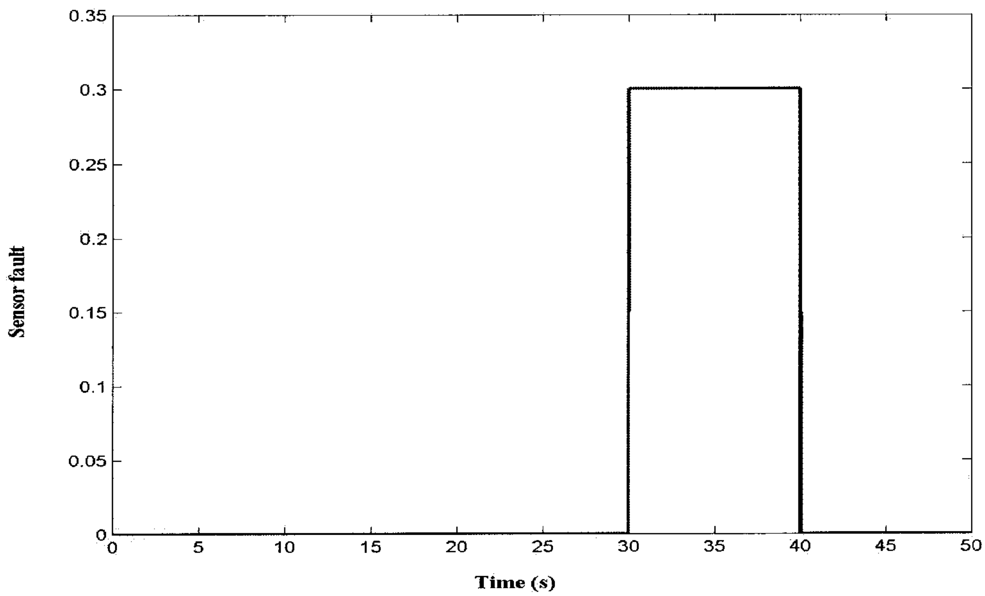



Figure 6: Fault signal pattern

Figure 6 shows the simulated fault pattern entering the system through the sensor.

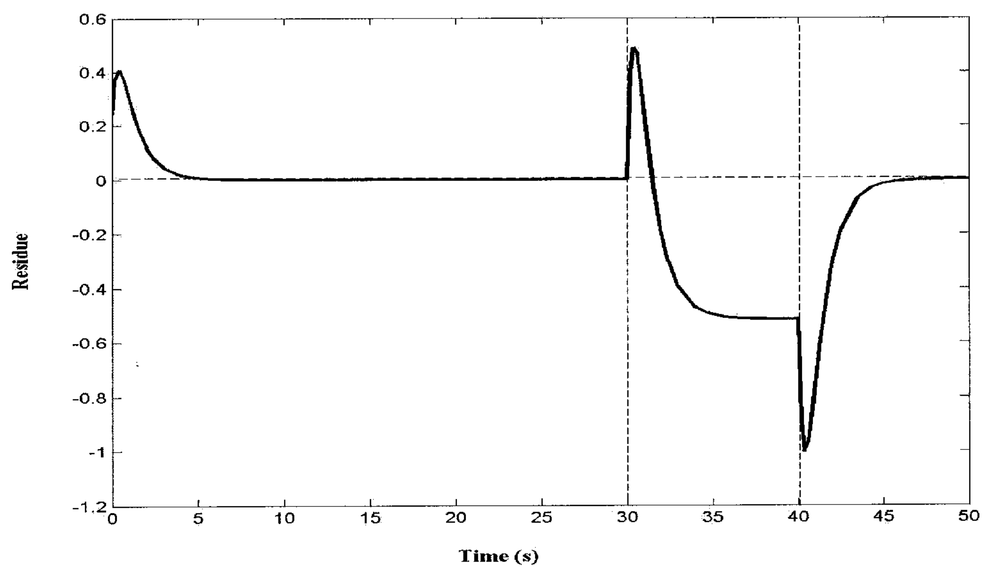



Figure 7: Fault residue signal

Figure 7 shows the effect of the sensor fault on the residue. The above figure clearly shows that there was a fault in the system from 30 to 40 seconds, whose effect lasted until about 45 seconds.

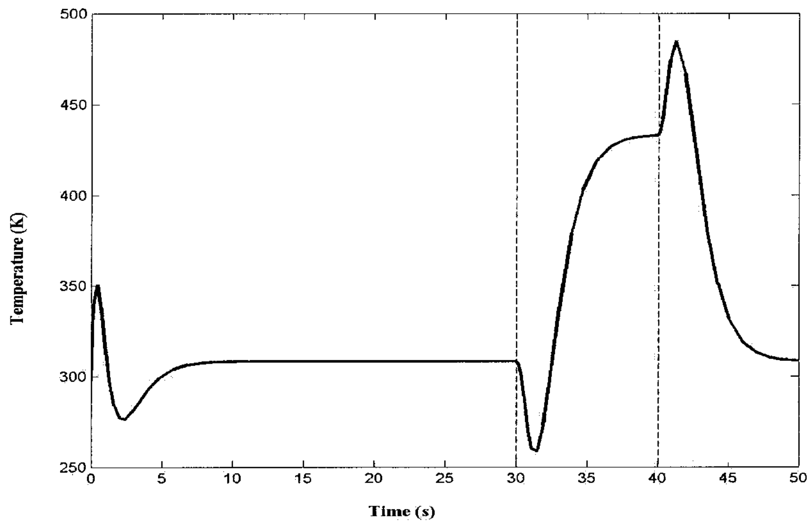



Figure 8: Temperature signal

Figure 8 shows the effect of the sensor fault on the measurable output of the system, namely the temperature.

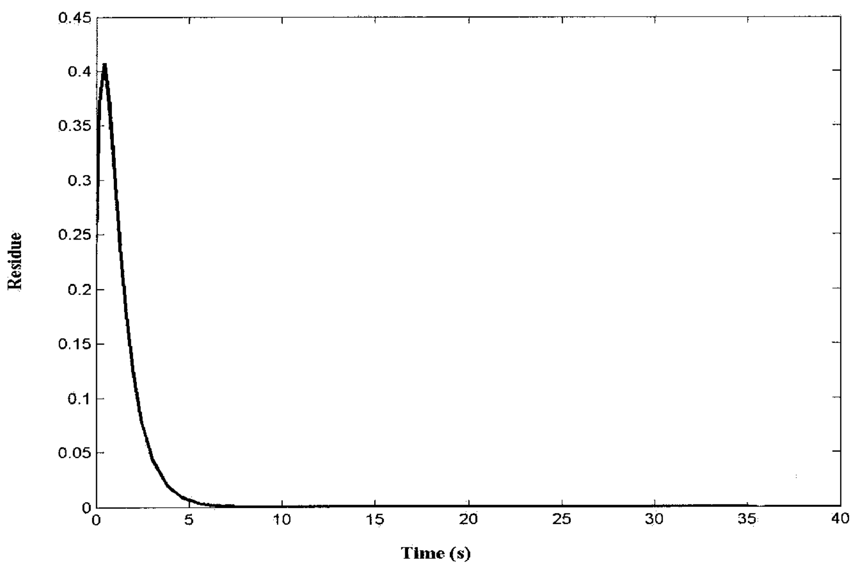



Figure 9: No Fault residue signal

Figure 9 shows the residue signal in the absence of fault. This figure clearly shows that the residue approaches zero in the absence of a fault and hence is independent of the system input and output. However, the speed of approach to zero depends on the speed of the observation [Luenberger, 1979].

A look at Figure 8 does not provide the idea of fault, but Figure 7 indicates clearly that there was a fault from time 30 to 40 s whose effect lasted approximately 45 s after which the controller was able to drive the process back to the operating conditions. Clearly under the absence of faults the state feed back regulator is able to bring the process back to the operating conditions in spite of disturbance, as shown in Figure 10.

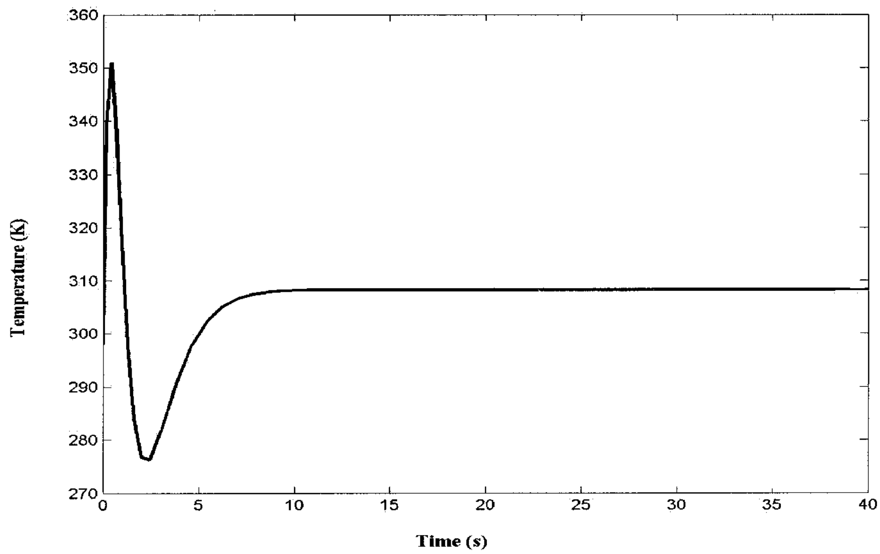



Figure 10: Faultless temperature signal

## 6.0 Conclusions

The above methodology for residual generation in fault diagnosis uses a linear estimation of the states of the process to be monitored. The residual is independent of the inputs and the outputs of the process but sensitive only to the faults and hence is appropriate for detection and isolation of process faults. The method proposed can be achieved through powerful computation, which is feasible and is on-line in nature and does not permit opening the control loop. However, this treatment would be suitable only for linear systems with the variables in the vicinity of the operating point. Also the issue of model uncertainty hence the issue of robustness is not dealt here. Our future work will explore other complications due to modeling of the system and will work with stochastic systems.

## 7.0 References


1. Luenberger, D.G (1979) *Introduction to Dynamic systems, Theory, Models and applications,* John Wiley & Sons, Inc.

2. Fogler, H.S (1992). *Elements of chemical reaction engineering*. Englewood Cliffs, NJ: Prentice Hall.

3. Isermann, R. (1984). Process fault detection based on modeling and estimation methods: A survey, *Automatica* 20(4): 387-404.

4. Jie Chen and Ron Patton (1999). *Robust model based fault diagnosis and isolation*, Kluwer Academic Publishers.

5. O'Reilly, J. (1983). *Observers for Linear systems*, Academic Press.

6. Vejtasa, S.A., & Schmitz, R.A. (1970). An experimental study of steady state multiplicity and stability in an adiabatic stirred reactor. *A.I.C.H.E. Journal*, 3, 410.