OBJECT-BASED DIAGNOSTIC NETWORK BASED ON STATISTICAL LEARNING

Sang Ok Song, En Sup Yoon

School of Chemical Engineering, Seoul National University

Abstract: This study proposes an effective diagnostic method using data readily available from the process. The system is decomposed into its local diagnostic models based on the direct and local causalities of process variables and the statistical learning model for each local relation is developed using data available from the process. The decomposed local models and the underlying fault assumptions compose of an object-based diagnostic network to perform on-line fault diagnosis. The diagnostic performance of the proposed method has been successfully illustrated in CSTR process.

Keywords: fault diagnosis, fault detection and isolation, statistical learning, residual generation and evaluation.

1. INTRODUCTION

As increasing the importance of early and accurate fault detection and diagnosis for modern industrial plants, various process monitoring techniques have been developed for the past decade. While techniques based on first-principles models have been around for more than two decades, their contribution to industrial practice has not been pervasive and data-driven approaches (process history based methods) have been widely used for process monitoring due to the high cost and time required to develop sufficiently detailed physical models.

Most of conventional data-driven monitoring methods use the process data collected during both normal operating conditions and specific faults to develop the measures for diagnosing faults based on supervised classification. However, the process data during abnormal situations are hardly available in real world industries and then conventional data-based fault diagnosis has a major disadvantage for their application to real processes (Lee, 2003)

2. PROBLEM DEFINITION

The fault diagnosis procedure consists of two separate stages: residual generation and residual evaluation. In other words, the automatic fault diagnosis can be viewed as a sequential process involving the symptom extraction and the actual diagnostic task. As usual, the residual generation process is based on a comparison between the measured and predicted system outputs. As a result, the difference of the so-called residual is expected to be near zero under normal operation, but on the occurrence of fault a deviation from zero should appear. In the meantime, the residual evaluation module is dedicated to the analysis of the residual signal in order to determine whether a fault has occurred and to isolate the fault in a particular system device.

One of the most known approaches to residual generation is a model-based concept. In the general case, this concept can be realized using different kinds of models: analytical, knowledge-based and data-based ones. Unfortunately, the analytical model-based approach is usually restricted to simpler systems described by linear models. When there are no mathematical models of the diagnosed system or the complexity of a system increases and the task of modelling is very hard to achieve, analytical models cannot be applied or cannot give satisfactory results (Patan, 2000). In these cases data-based models can be considered.

3. CONVENTIONAL METHODS AND THEIR LIMITATIONS

Technological plants are often complex dynamic systems described by nonlinear high-order differential equations. For their quantitative

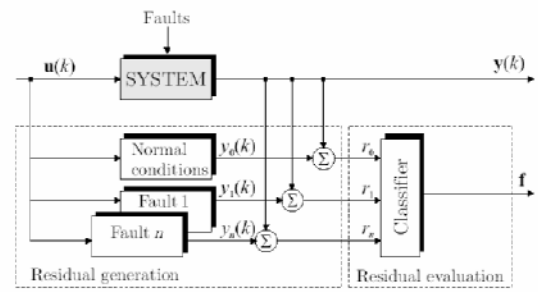

Fig.I. General scheme of model-based fault diagnosis

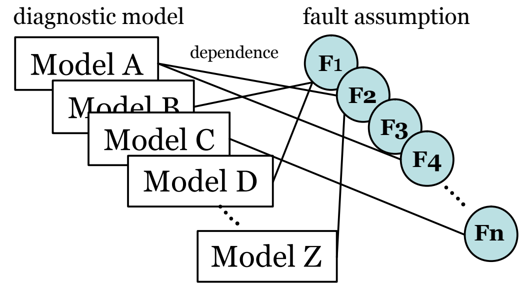

Fig. II. Object-based diagnostic network

modelling for residual generation, simplifications are inevitable. This usually concerns both the reduction of dynamics order and linearization. Another problem arises from unknown or time variant process parameters. Due to all these difficulties, conventional analytical models often turn out to be not accurate enough for an effective residual generation. In this case knowledge-based models are only alternatives. Replacing unknown parameters by qualitative knowledge-based approach enhances the robustness of the model versus unknown or time-dependent parameters. As one of the data-based approaches, the neural network replaces the analytical model that describes the process under normal operation. Fig. I illustrates how the fault diagnostic system can be designed using neural network techniques

However, these approaches have some inherent limitations on their ability to effectively identify and diagnose faults. As mentioned above, this approach requires data in every faulty situation and those data are not generally available at the real process. Hence methods that require information about faulty situations cannot be used to diagnose faults that have not previously occurred. This is a major advantage of conventional data-based approaches. Even if you have a good simulation model for learning data, it is obvious and natural that analytical model-based approaches are more favourable than data-based model approaches. To take advantage of data-based model approaches and overcome their limitations, a practical diagnostic method that uses data available from the real process only should be developed.

4. THE PROPOSED DIAGNOSTIC METHOD

In order to overcome the aforementioned limitation of data-driven model based diagnostic method, an object-based diagnostic network has been developed (Fig. II). The diagnostic network consists of a series of diagnostic data-based models with qualitative knowledge base. The qualitative knowledge base is represented as a set of assumptions about underlying faults. These fault assumptions represent the assumptions that the associated model is based upon (Petti, 1992). Each diagnostic model is trained using normal operation data. The data-based models generate residuals based on a comparison between the measured and predicted system outputs. If all the assumptions of a model hold, the satisfaction of the model is guaranteed, i.e., the residual is close to zero. By examining the direction and extent to which each model is violated against its fault assumptions, the most likely failed assumption (fault) can be deduced. An assumption that is common to many violated models is strongly suspect, whereas satisfaction of models provides evidence that the associated assumptions are valid. How the models can be formulated and how the fault assumptions can be associated with the models is explained in the next sub-sections.

4.1 Design of local diagnostic models

The formulation of diagnostic models is based on system decomposition by using local and direct causalities between process variables. The system decomposition approach narrows the diagnostic focus to a particular decomposed subsystem and performs diagnosis on this subsystem (Finch, 1988). System decomposition has the advantages of providing flexible diagnosis throughout operating condition changes, reducing the size of the knowledge base, and simplifying the understanding of complex process to process interactions.

The local subsystem models can be designed based on the information about local causalities between process variables. An example of subsystem models for a simple gravity flow tank is shown in Fig. III. The variables having a direct effect on L (target variable) are F_{in} and F_{out}. And F_{out} (target variable) is first affected by R_s and L. From these direct and local causalities, two models for L and F_{out} can be designed as shown in Fig. III. F_{in} and R_s are considered exogenous to a system because their values are determined outside of the system and controlled externally.

The availability of local subsystem models highly depends on the measurability of the process variables. Notice that unmeasured variables cannot participate in formulating diagnostic models due to their unavailability during training models. The outflow model contains an unmeasured variable R_s (the pipe resistance) and can not be learned from history data. However, the outflow model may be valid for diagnostic task in which case the unmeasured variable R_s can be considered as a constant.

4.2 Fault assumptions

Associated with each diagnostic model is a set of assumptions which if satisfied, guarantee the satisfaction of the model. The fault assumptions represent the faults that the model can detect. Some

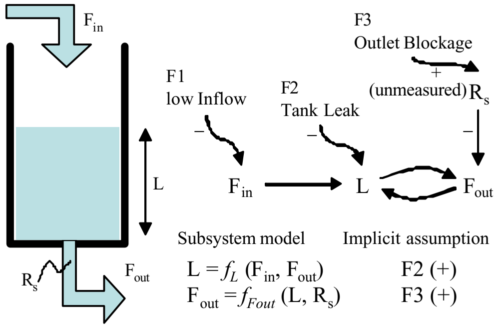

Fig. III. A gravity flow tank with subsystem models

of the assumptions are explicit in the model, such as correct sensor readings, and some are implicit such as the fact that there are no tank leaks. Consider the tank in Fig. III once again. The model L has following assumptions.

- Inflow sensor is OK (explicit).
- Outflow sensor is OK (explicit).
- No tank leaks (implicit).

If all the assumptions above hold, the satisfaction of the model L is guaranteed and the residual is close to zero.

Every explicit fault assumption is sensor fault and it is included in the fault assumption set naturally. On the other hand, implicit fault assumptions are associated with each diagnostic model according to the following rules.

For each fault (F_k) there must exist a primary deviation variable (x_{Fk}), i.e., a variable included in the system which is first affected by the root cause.

Case 1: When x_{Fk} is measured,

- if there exists the diagnostic local model for x_{Fk}, x_{Fk} is associated with model x_{Fk}.
- if there does not exist the diagnostic local model for x_{Fk}, this fault is not detectable except that x_{Fk} is in steady state condition.

Case 2: When x_{Fk} is unmeasured, any measured variable which is first affected by x_{Fk} is considered as the primary deviation variable and then apply to *case 1*.

In the tank example, the tank level (L) is the primary deviation variable for tank leak (F2) and there exists the local model for the tank level (L). Therefore F2 is associated with the level model and can be considered as the implicit fault assumption of the level model. Outflow (F_{out}) is considered as the primary deviation variable for the outlet blockage (F3) because the pipe resistance R_s which is actually first affected by F3 is unmeasured. In this case, F3 is the implicit fault assumption of the outflow model.

In case of low inflow disturbance (F1), we do not have the local model for the primary deviation variable F_{in} because F_{in} is an exogenous variable. If F_{in} is steady state, F1 can be detected by checking some tolerance limit.

The fault assumptions associated with their local diagnostic models have their qualitative information. The qualitative information represents the direction to which the associated local model will be violated when a fault assumption is not satisfied. This symptom acts as the basis to diagnosis faults. If a fault assumption is connected with its associated local model positively (negatively), it implies that the predicted output of the local model will deviate from its measured value in the positive (negative) direction when the fault occurs

4.3 Training local models

For residual generation, each local diagnostic model designed at previous step has to be trained. Learning data can be collected directly from the process if possible or from a simulation model that is as realistic as possible. Because a simulation model is assumed to be unavailable, data available from normal operating condition only should be used for the training. In the proposed method, each local model can be learned from the data set representing the local relations between the input variables and the output variable. The required data sets can be easily obtained during normal operation. Sometimes, the training data may be usually bounded on a small region under steady state condition and the fault assumptions may not be valid for the outside of the training data region. Therefore, the available data which can be obtained in the presence of intentional step changes or external disturbances that occur frequently in the process can be used for the training in order to guarantee the fault assumptions in as large region as possible. In this case, it must be noted that the data set obtained in the presence of a particular fault should not be used to train the local model(s) with same fault assumption.

In the proposed method, the statistical learning techniques, specially focusing on regression are used for training local models. The value of each target variable can be estimated by using statistical learning model (learning machine) for each local model. Various statistical methods for regression such as splines, neural networks and Support Vector Machine can be a tool for building learning machine. The current prevailing view in the statistical and neural network community is that there is no single best method for all regression problems (Friedman, 1994) and the comparative work of various learning techniques is not the point of this research. You can find a comparative study for various regression methods in (Cherkassky, 1998). In our work, Support Vector Regression method was selected as a learning technique because it leads to a more tractable formulation of the optimization problem and the notion of model complexity can be separated from dimensionality. In addition, SVM can handle nonlinearity of the model in the easy manner by selecting appropriate kernel function. Other advantages of SVM can be found in many references (Cherkassky, 1998; Vapnik, 1995).

4.4 Residual generation and evaluation

After finishing the training via an appropriate statistical learning technique, the trained local models are ready for on-line residual generation. The residuals, between predicted values from local models and measured values are monitored and the CUSUM (cumulative summation) control chart is used to detect the changes from normal operating state of a process. The CUSUM test first introduced by Page (Page, 1954) has been shown to be efficient in detecting changes in the mean of a process . As the name implies, the CUSUM chart cumulates deviations of the sample readings (x_i) from the target or desired value (μ). Once these cumulative summations reach either a high or low limit, an out-of-control signal is given. The parameters in the CUSUM chart are defined as follows:

k: the threshold for cumulative summation, which can be defined as the minimum difference between sample average and target that will cause the cumulative summation to begin.

SH_i and SL_i: the cumulative summation terms which are calculated as follows.

$$\begin{aligned} SH_i &= \max\left[0, SH_i^t + SH_{i-1}\right] \\ SL_i &= \max\left[0, SL_i^t + SL_{i-1}\right] \\ &\text{where} \\ SH_i^t &= \text{highsidecumulativeterm} \\ &= (x_i - \mu) - k \\ SL_i^t &= \text{lowsidecumulativeterm} \\ &= -(x_i - \mu) - k = (\mu - x_i) - k \end{aligned} \tag{2}$$

h: the control limit

Since the residuals are not uniform in magnitude, they are transformed into a metric between −1 and 1 which indicates the degree to which the model equation is satisfied: 0 for perfectly satisfied, 1 for severely violated high, and −1 for severely violated low. These values constitute the satisfaction vector, sf, which is calculated using the CUSUM results, SH_i and SL_i.

For the jth diagnostic model,

$$sf_j = \frac{sign\left(SH_j - SL_j\right)\left(\left(SH_j - SL_j\right)/h_j\right)^n}{1 + \left(\left(SH_j - SL_j\right)/h_j\right)^n} \tag{3}$$

The value of sf_j is given a positive value for a positive residual and a negative value for a negative residual. The curve is a general sigmoidal function with the steepness determined by the constant n.

A matrix of sensitivity values, S which describes the relationship between each local model and fault assumption is assigned to weight the sf values as evidence. The ijth element of S, represents the sensitivity of the jth model to the ith assumption. Unless experience suggests otherwise, these values are usually given −1, 0, or 1; 0 for the local models independent of an assumption, −1 for the local models with negative effect by an assumption, and 1 for positive effect. In some cases, the effect of a fault assumption is uncertain if negative or positive. For these cases, additional sensitivity matrix S' is created, whose values are given 1 or 0; 1 for the local model dependent of an assumption, 0 for the local model independent of an assumption.

In addition, a matrix of non-sensitivity values, NS which also describes the relationship between each local model and fault assumption is assigned to give some penalty to non-sensitivity symptoms. These values are given 1 or 0; 1 for the local model independent of an assumption, 0 for the others

Conclusion about the satisfaction of each assumption (fault) is made by combining the evidence from the local model, sf with consideration to the sensitivity matrix S, S', and the compensation matrix NS. This is done through the calculation of a vector of failure likelihoods, FL such that

$$FL_i = \frac{\sum_{j=1}^{N}\left(S_{ij} \cdot sf_j\right)}{\sum_{j=1}^{N}\left|S_{ij}\right|} + \frac{\sum_{j=1}^{N}\left(\left|S'_{ij} \cdot sf_j\right|\right)}{\sum_{j=1}^{N} S'_{ij}} - \frac{\sum_{j=1}^{N}\left(\left|NS_{ij} \cdot sf_j\right|\right)}{\sum_{j=1}^{N} NS_{ij}} \tag{4}$$

where N is the number of local model. It is evident that this method of combination allows the sf values of those local models which are most sensitive to ith fault assumption to be weighted the most heavily in the calculation of FL_i. The failure likelihood is interpreted as indicating a likely condition of ith fault assumption failing high as the value of FL_i approaches 1, while an FL_i tending toward −1 indicates a likely failure low.

5. ILLUSTRATIVE EXAMPLE

The example process is simple but displays the most common characteristics of industrial processes. This process is simulated by the model of Sorsa (Sorsa, 1991). The sampling interval is 5 s and faults are introduced at 100 s. Some selected faults of the CSTR process is described in Table 1.

5.1 Design of local diagnostic models

Based on the knowledge of the CSTR process, we obtained a direct causal relation between variables. The resulting local causality relation of the CSTR process is illustrated in Fig V. The whole system could be decomposed into several subsystems from these direct and local causalities. For example, the variables having a direct effect on T_R are F_R, T, and F_W. From this causality, a local model for recycle flow rate (T_R) can be designed such as $T_R = f(F_W, F_R, T)$.

5.2 Fault assumptions

Each fault should be associated with corresponding local models before on-line diagnostic task. Explicit fault assumptions such as sensor faults are naturally associated with the local models that contain the corresponding sensor variables. For example, *FSBL* (flow sensor biased low) represents a failure for the recycle flow (F_R) sensor. If the sensor for F_R is not normal, the local model for F_R deviates from its normal structure. In addition, the local models for T and T_R that are directly affected by the real value of recycle flow make wrong estimates because they have incorrect information about the recycle flow rate.

Table 1 Selected faults of the CSTR process

NO.	name	description
1	*FSBL*	flow sensor biased low
2	*FVBL*	flow control valve biased low
3	*LEAK*	reactor leaking
4	*LSBL*	level sensor biased low
5	*LVBH*	level control valve biased high
6	*TSBH*	temperature sensor biased high
7	*TVBH*	temperature control valve biased high

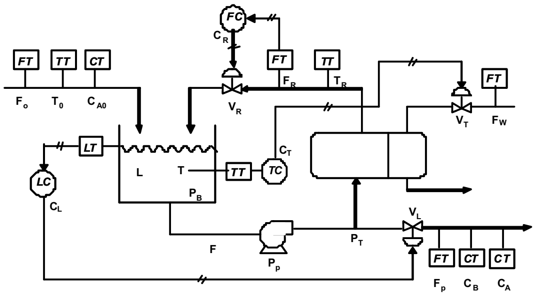

Fig. IV. Process flow diagram of the CSTR process

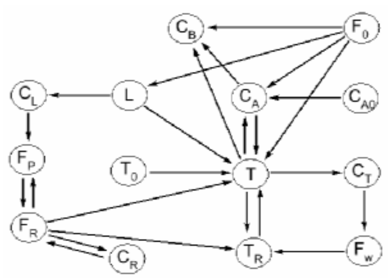

Fig. V. The direct causalities of the CSTR process

On the other hand, implicit fault assumptions are associated with their corresponding local model according to the adding rule explained in section 4.2. For example, *LVBH* (level control valve biased high) represents a fault for valve V_L (Fig IV.). This fault affects the valve output first. Because the valve output is not measured, the output flow rate (F_P) that is first affected by the valve output (V_L) becomes the primary deviation variable for the fault *LVBH* and the local model for F_P is associated with the fault , *LVBH*. In this manner, all faults in Table 1 are associated with decomposed local models and the object-based diagnostic network for the CSTR process is constructed as shown in Fig VI. The solid and dotted lines indicate the direction in which the associated model will deviate when the fault occurs. For example, *FSBL* is connected with the local model for F_R positively (solid line). It implies that the predicted output of the model F_R will deviate from its measured value in the positive direction when the fault *FSBL* occurs.

5.3 Training local models

Each diagnostic local model has been trained via support vector regression technique. Learning data for training the local models designed at the previous steps were collected under the condition where some external disturbances occurred or the set points of the controllers were changed. These kinds of data can be easily obtained in the process industries.

External disturbance data
Cooling water temperature high
Feed composition change low
Feed flow rate change high
Feed temperature change low

Controller set point change data
Level controller set point high
Temperature controller set point high
Flow controller set point high

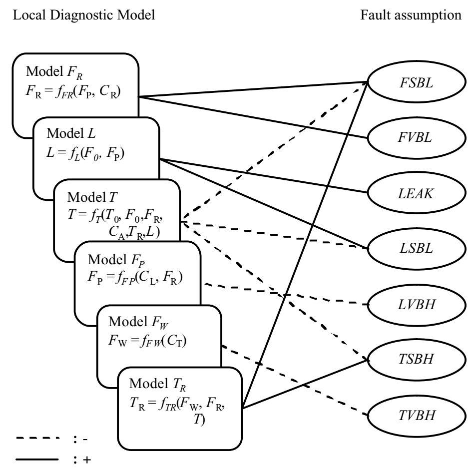

Fig. VI. The object-based diagnostic network for the CSTR process

The input matrix of each local model includes a part of its past observations to handle the time-dependency of the variables. The past values for all input variables always do not contribute to predict the output of a local model. For some input variables, both their past values and current values affect the current output of a local model. But in other cases, only the current values of the input mainly affect the output of a model. Whether or not the past values of the variables should be included in the input matrix is determined on the basis of the knowledge of the system and the comparison of the prediction performance of a local model via cross validation.
Polynomial kernel was considered when applying Support Vector Machine regression technique. The determination of appropriate kernel parameters (e.g., polynomial degree of the polynomial kernel) was performed by cross validation. Table 2 summarizes the input-output structure of each local model and model parameters selected. Table 2 summarizes the input-output structure of each local model and model parameters selected by cross validation.

5.4 On-line residual generation and evaluation

This study used 2σ of the residual distribution as the threshold for cumulative summation (k) and 6σ of the CUSUM distribution as the control limit (h).

Table 2 The input-output structure of diagnostic local models and selected parameters

Local Model	Inputs	Output	Kernel function in SVM
F_R	F_R(t-1), F_P(t-1) C_R(t), F_P(t)	F_R(t)	Polynomial (n=1)
L	L(t-1) F_0(t), F_P(t)	L(t)	Polynomial (n=1)
T	T(t-1) T_0(t), F_0(t), C_A(t) L(t), T_R(t), F_R(t)	L(t)	Polynomial (n=1)
F_P	F_P(t-1) C_L(t), F_R(t)	F_P(t)	Polynomial (n=1)
F_W	C_T(t)	F_W(t)	Polynomial (n=1)
T_R	T_R(t-1), F_W(t-1) T(t-1), F_R(t-1) F_W(t), T(t)	T_R(t)	Polynomial (n=2)

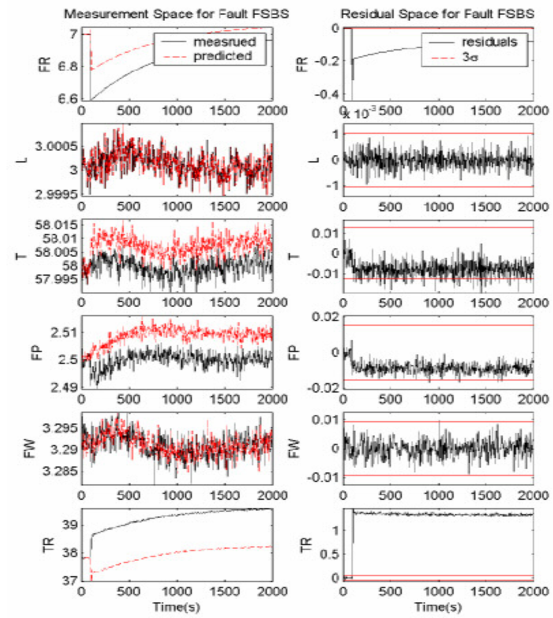

Fig. VII. The measured/predicted values and residuals for *FSBL*

Fig. VII and VIII show the results for *FSBS (flow sensor biased low)*. The flow sensor for recycle flow is biased low (-0.7 kg/s). This fault affects the measurement F_R directly and the biased measurements break the internal relationships of the corresponding local models. The diagnostic network for the CSTR process (Fig. VI) shows that the fault *FSBL* is associated with the local models for F_R, T_R and T. The local model for F_R detected an abnormal behaviour at 100s and the local model for T_R gave a high signal at 105s and the local model for T gave a low signal at 110s (Fig. VII). Fig. VIII shows the fault likelihoods of considered faults for 10 sampling period. The fault likelihood of *FSBL* was largest after 105 sec and could be successfully diagnosed.
The diagnosis results of selected faults are summarized in Table 3.

5. CONCLUSION

This study proposes an object-based diagnostic network which can diagnosis process faults effectively without faulty data and an illustrative example showed the proficiency of the proposed method. The proposed diagnostic method is very practical in that data readily available from the process are used only and the modelling effort required for the development of diagnostic system is very cheap in comparison to analytical modelling.

Table 1 Selected faults of the CSTR process

Fault	Isolation Time (s)	Mismatch rate
FSBL	105	0.0026
FVBL	100	0
LEAK	110	0.2128
LSBL	105	0.2868
LVBH	100	0
TSBH	105	0.0026
TVBH	100	0

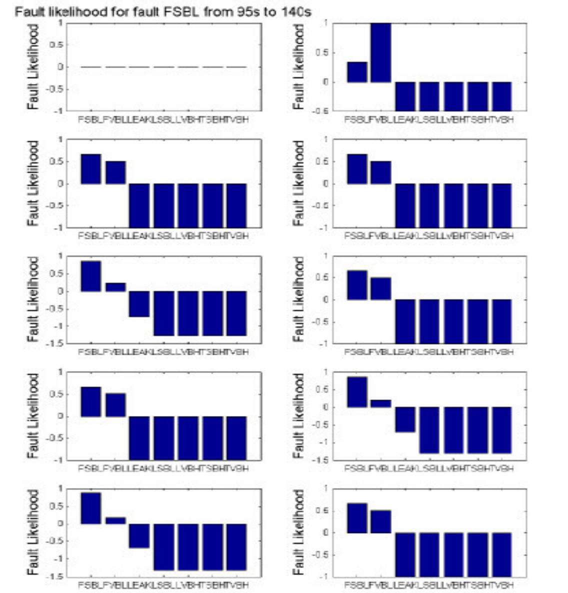

Fig. VIII. The failure likelihoods for *FSBL*

This approach provides a natural way of integrating qualitative information of the process into statistical learning structure

ACKNOWLEDGEMENT

We acknowledge the financial aid for this research provided by the Brain Korea 21 Program supported by the Ministry of Education. In addition, we would like to thank the Automation & Systems Research Institute and Research Institute of Engineering Science of Seoul National University.

REFERENCES

Lee, G and S.-O. Song (2003). Multiple-fault diagnosis based on system decomposition and dynamic PLS. *Ind. Eng. Chem. Res.*, **42**, 6145-6154.

Patan, K (2000). *Artificial dynamic neural networks and their applications in modelling of industrial processes*. Chapter 5, Ph. D. Dissertation, Warsaw University of Technology, Poland.

Petti, T. F. (1992). *Using mathematical models in knowledge-based control systems*. Ph. D. Thesis, University of Delaware, USA.

Finch, F. E. and M. A. Kramer (1988). Narrowing diagnostic focus using functional decomposition. *AIChE Journal*, **34**, 25-36.

Friedman, J. H. (1994). An overview of predictive learning and function approximation. In: *From Statistics to Neural Networks*, NATO ASI Series F 136, Springer Verlag, New York.

Cherkassky, V. and F. Muler (1998). *Learning from data: Concepts, Theory, and Methods*, John Wiley & Sons, New York.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*, Springer Verlag, New York.

Page, E. S. (1954). Continuous inspection scheme, *Biometrica*, **41**, 100-115.

Sorsa, T. and H. N. Koivo, Neural Networks in Process Fault Diagnosis, *Trans. Systems, Man and Cybernetics*, **21**, 815-825.