Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005



# Fuzzy Model Identification of a Sugar Cane Crushing Process for Fault Diagnosis Application

Silvio Simani

*Abstract*— **This work proposes a method for input–output sensor fault detection and isolation of an industrial processes using fuzzy process models. The presented technique concerns the identification of a piecewise affine fuzzy system based on Takagi–Sugeno models. The process under investigation may, in fact, be represented as a composition of several Takagi-Sugeno models selected according to the process operating conditions. This work also addresses a method for the identification of the local Takagi-Sugeno models from a sequence of noisy measurements acquired from the real process. The fault detection scheme adopted to generate residuals uses the Takagi-Sugeno fuzzy model. The developed technique was applied to fault diagnosis of input-output sensors of a sugar cane crushing mill.**

## I. INTRODUCTION

The application of fuzzy logic and control to model-based fault diagnosis has gained increasing attention in both theory and application in recent years. The former method exploits fuzzy set theory to express cause-effect relations in expert systems. On the other hand, the key idea of the model-based approaches is the generation of signals, symptoms or residuals, obtained by using observers, parameter estimation or parity equations designed on the basis of mathematical models of the monitored system [1].

The majority of industrial processes cannot be modelled by using a single model for all operating conditions because they are non–linear and non-stationary [2].

Since a mathematical model is a description of system behaviour, accurate modelling for a complex non–linear system is very difficult to achieve in practice.

Sometime, for some non–linear systems, it can be impossible to describe them by analytical equations. Instead of exploiting complicated non–linear models obtained by modelling techniques, it is also possible to describe the plant by a collection of local affine models obtained by identification procedures.

Symptoms are signals representing inconsistencies between the model and the actual system being monitored. Any inconsistency will indicate a fault in the system. Residual must, therefore, be different from zero when a fault occurs and zero otherwise. However, the deviation between the model and the plant is influenced not only by the presence of the fault but also the modelling error.

Several techniques had been proposed for fault detection and isolation (FDI) in dynamic systems [1]. In particular, in this work, fuzzy modelling is combined with the model-based method to formulate a FDI technique exploiting the Takagi-Sugeno fuzzy model [3] for residual generation.

The Takagi-Sugeno fuzzy model for non–linear dynamic systems is described by a number of local affine models. Each submodel approximates the system locally around an operating point and a selection procedure determines which particular submodel has to be used. Such a multiple model structure is be called multiple model approach [4].

Under such a fuzzy scheme, a number of local affine models are designed and the estimate of outputs is given by a fuzzy fusion of local outputs. The diagnostic signal (symptom or residual) is the difference between the estimated and actual system output.

In this paper, the different operating points are self-selected with a fuzzy clustering method [5]. On the basis of knowledge of the operating-point regions, the identification of the structure and the parameters of each local Takagi-Sugeno model can be performed [6], [7], [8], [9].

The remainder of this paper is organised as follows. Section II presents the structure of the multiple model, while Section III briefly illustrates how to integrate a well-established identification method [10], [11] for the estimation of affine systems within a general procedure for non–linear fuzzy identification [6], [7], [8], [9]. Section IV shows the design of the diagnostic scheme for FDI of dynamic systems. The application of such a FDI approach to a real power plant is described in Section V. The example demonstrates the effectiveness of the technique proposed. Finally, some concluding remarks are included in Section VI.

## II. FUZZY PROCESS MODELS

This section deals with the decomposition of input-output data $u(t)$ and $y(t)$ $(t = 1, \ldots, N)$, acquired from the actual process, into fuzzy subsets which can be approximated by local affine input-output models. Each submodel represents the system behaviour around the operating point.

Fuzzy clustering algorithms can be used as a tool to obtain partitioning of data into subsets, which can be approximated by local affine models.

It is assumed that the dynamics of the system under observation can be described by the following equation error (EE) model [12]:

$$y(t) = f\big(x(t)\big) + \varepsilon(t) \tag{1}$$

where the vector $y(t)$ is the system output, $x(t)$ is a collection of a finite number of inputs and outputs, the vector $x^T(t) = [y(t-1), \ldots, y(t-n), u(t-1), \ldots, u(t-n)]$, $f(\cdot)$ describes the input-output link, while $\varepsilon(t)$ reflects the fact

Silvio Simani is with Dipartimento di Ingegneria, Università di Ferrara, Via Saragat 1, 44100 Ferrara, Italy ssimani@ing.unife.it

that $y(t)$ is not an exact function of $x(t)$. $n$ is an integer related to the system order.

The objective of fuzzy clustering algorithms is to partition the set of observed inputs and outputs $\{x(t)\}_t$ of an unknown system into a number $M$ of fuzzy subsets. Each subset, $R_i$, representing an operating condition of the dynamic system, can be approximated by an affine dynamic model.

Partition of the data set into fuzzy subset can be achieved, for instance, by using the well-established Gustafson-Kessel (GK) clustering algorithm in [13].

Each cluster $R_i$ $(i = 1, \cdots, M)$ obtained by fuzzy partitioning is regarded as a local approximation of the real process. The global EE model (1) can be conveniently represented using local affine Takagi-Sugeno rules [3] $y_i(t)$:

$$x(t) \ \in \ R_i \Rightarrow y_i(t) = \theta_i^T x(t) \tag{2}$$

where $\theta_i$ is the $i$-th parameter vector of the $i$-th submodel, with $i = 1, \cdots, M$.

The Takagi-Sugeno fuzzy model is a simple way to describe a non–linear dynamic system using local affine models. By means of Takagi-Sugeno models, any dynamic system can be linearised around a number of operating points. The global system behaviour is described by a fuzzy fusion of all linear model outputs:

$$\hat{y}(t) = \frac{\sum_{i=1}^{M} \mu_i\big(x(t)\big) y_i(t)}{\sum_{i=1}^{M} \mu_i\big(x(t)\big)}. \tag{3}$$

in which $\hat{y}(t)$ is the estimate of the output $y(t)$ at the instant $t$. The results of the clustering algorithms are $M$, the membership functions $\mu_i(\cdot)$ and the subsets of input-output data $\{x_i(t)\}_{i=1}^{M}$ with $x_i(t) \in R_i$ [5].

These subsets can be processed according the Frisch scheme identification procedure [11], in order to estimate the $theta_i$ and $n$ parameters for each submodel.

## III. LINEAR DYNAMIC SYSTEM IDENTIFICATION

This section shows how to integrate the fuzzy clustering technique, described in the previous section, and the Frisch scheme identification method.

Without loss of generality, a simple single-input single-output (SIS0) discrete system will be considered.

For the identification in each region $R_i$ of the local affine model, a finite sequence of the variables $x_i(t) \in R_i$ observed with a constant sampling interval is considered.

If dynamic linear relations exist among these variables, they can be described by models of the type:

$$y_i(t) = \theta_i^T x_i(t) \tag{4}$$

which represent a linear SISO discrete-time system whose order is $n$ and whose parameters are the entries of the vector $\theta_i$.

If the following vectors and matrices are defined:

$$\begin{array}{lcl} u^N(t+k) & = & \big[u(t+k) \ \ldots \ u(t+k+N-1)\big]^T \\ y^N(t+k) & = & \big[y(t+k) \ \ldots \ y(t+k+N-1)\big]^T \\ X_k(u) & = & \big[u^N(t) \ \ldots \ u^N(t+k-2)\big] \\ X_k(y) & = & \big[y^N(t) \ \ldots \ y^N(t+k-1)\big] \\ \hat{\Sigma}_k(uu) & = & X_k^T(u) X_k(u) \\ \hat{\Sigma}_k(yy) & = & X_k^T(y) X_k(y) \\ \hat{\Sigma}_k(yu) & = & X_k^T(y) X_k(u) = \hat{\Sigma}_k^T(uy) \end{array} \tag{5}$$

where $N$ is assumed large enough to solve the problem considered. Let us partition now the matrix $\hat{\Sigma}_k$ as follows:

$$\hat{\Sigma}_k = \left[ \begin{array}{cc} \hat{\Sigma}_k(yy) & \hat{\Sigma}_k(yu) \\ \hat{\Sigma}_k(uy) & \hat{\Sigma}_k(uu) \end{array} \right]. \tag{6}$$

To solve the realization problem (noise-free data) it is possible to consider the sequence of increasing–dimension matrices:

$$\hat{\Sigma}_2, \ \ \hat{\Sigma}_3, \ \ \ldots \ \ \hat{\Sigma}_k, \ \ \ldots \tag{7}$$

testing their singularity. As soon as a singular matrix $\hat{\Sigma}_k$ is found then $n = k-1$ and the $2n-1$ parameters $\theta_i$ describe the dependence relationship of the $n$-th vector of $\hat{\Sigma}_{n+1}$ on the remaining ones.

It has been assumed that $N$ is large enough to avoid unwanted linear dependence relationships due to limitations in the dimension of the involved vector spaces; this means $N \geq 2n+1$.

In the noisy case, the following identification method was proposed [11].

In this condition, the procedure described for the solution of the realization problem (7) would obviously be useless since matrices $\hat{\Sigma}_k$ would always be non-singular because of the presence of noise.

In the Frisch scheme it is normally assumed that:

$$\left\{ \begin{array}{lcl} u(t) & = & u^*(t) + \tilde{u}(t) \\ y(t) & = & y^*(t) + \tilde{y}(t) \end{array} \right. \tag{8}$$

where $u^*(t)$ and $y^*(t)$ are the noise-free data and noise terms $\tilde{u}(t)$ and $\tilde{y}(t)$ are independent of every other term and only $u(t)$ and $y(t)$ are known. Note that, in the realization problem, $u(t) = u^*(t)$ and $y(t) = y^*(t)$ since $\tilde{u}(t) = 0$ and $\tilde{y}(t) = 0$.

Consequently the generic positive definite matrix $\Sigma_k$ associated with the input-output noise-corrupted sequences may always be expressed as the sum of two terms $\Sigma_k = \hat{\Sigma}_k + \tilde{\Sigma}_k$ where

$$\tilde{\Sigma}_k = diag[\tilde{\sigma}_y I_k, \tilde{\sigma}_u I_{k-1}] \geq 0 \tag{9}$$

since no correlation has been assumed among the noise samples at different times. This condition is verified for additive white noise with variance $\tilde{\sigma}_y$ and $\tilde{\sigma}_u$ on the input-output sequences.

In the stochastic case, the following problem should be solved.

Given a sequence of increasing–dimension $\big((2k-1) \times (2k-1)\big)$ symmetric positive definite covariance matrices:

$$\Sigma_2, \ \ \Sigma_3, \ \ \ldots \ \ \Sigma_k, \ \ \ldots \tag{10}$$

find, for each $k$, all diagonal non-negative definite matrices $\tilde{\Sigma}_k = diag[\tilde{\sigma}_y I_k, \tilde{\sigma}_u I_{k-1}]$ such that

$$\hat{\Sigma}_k = \Sigma_k - diag[\tilde{\sigma}_y I_k, \tilde{\sigma}_u I_{k-1}] \geq 0 \ . \qquad (11)$$

It can be noted that for each $k$ the solution set of relation (11) describes, in the first orthant of the $(\tilde{\sigma}_y, \tilde{\sigma}_u)$ hyperplane, a hypersurface whose concavity faces the origin [11].

Previous results hold for every value of $k$. Since determination of the system order requires the increasing values of $k$ to be tested, it is relevant to analyse the behaviour of the associated curves when $k$ varies. This corresponds to a comparison of the admissible solution sets for different model orders. It can be shown that the solution sets of condition (11) for different values of $k$ are non–crossing curves [11].

It is also important to observe that, since we assume that a system (4) has generated the noiseless data, for $k > n$ all the hypersurfaces of type (11) have necessarily at least one common point, i.e. point $(\tilde{\sigma}_{y^*}, \tilde{\sigma}_{u^*})$ corresponding to the true variances $\tilde{\sigma}_{y^*}$ and $\tilde{\sigma}_{u^*}$ of the noise affecting the output and the input of the system. The search for a solution for the identification problem can thus start from the determination in the noise space of this point. The following considerations can now be stated.

With reference to the diagonal non-negative definite matrices $\tilde{\Sigma}_k = diag[\tilde{\sigma}_{y^*} I_k, \tilde{\sigma}_{u^*} I_{k-1}]$, the following properties hold:

1) If $k \leq n$ the matrices $\hat{\Sigma}_k$ are positive definite.
2) If $k > n$ the dimension of the null space of $\hat{\Sigma}_k$ and, consequently, the multiplicity of its least eigenvalue, is equal to $(k-n)$.
3) For $k = (n+1)$ matrix $\hat{\Sigma}_k$ is characterised by a linear dependence relation among its $2k-1$ vectors and the coefficients which link the $k$-th vector of $\hat{\Sigma}_k$ to the remaining ones are the parameters in $\theta_i$ of the system (4) which has generated the noiseless sequences.
4) For $k > (n+1)$ all linear dependence relations among the vectors of the matrix $\hat{\Sigma}_k$ are characterised by the same $2n-1$ coefficients in $\theta_i$.

If $m$ models of the type (4) are used to describe the mathematical behaviour of a multivariable dynamic system with $r$ inputs and $m$ outputs, the previous identification procedure must be repeated $m$ times. At every step the identification procedure must lead to the same values for the input noise variances $(\tilde{\sigma}_{y^*}, \tilde{\sigma}_{u_1^*}, \ldots, \tilde{\sigma}_{u_r^*})$

It is worthy to note how this approach cannot be applied immediately to the identification of real processes, since the hypotheses on the linearity, finite dimensionality and time independence of the system and on the additivity and whiteness of the noise are not usually verified, so that the hyper-surfaces (11) have no common point for $k > n$. The definition of a suitable criterion of model selection in such cases was suggested in [14].

## IV. MULTI MODEL–BASED FAULT DETECTION

The problem treated in this section regards the detection and isolation of the input-output sensors faults of a process on the basis of the knowledge of the measured noisy sequences $u(t)$ and $y(t)$.

In the following it is assumed that the monitored system, depicted in Figure (1), can be described by a model of the type (1). $y(t) \in \Re^m$ is the system output vector and $u(t) \in \Re^r$ the control input vector. The term $\varepsilon(t)$ takes into account the modelling error, which is due to process noises, parameter variations, etc. According to Eqs. (8), in

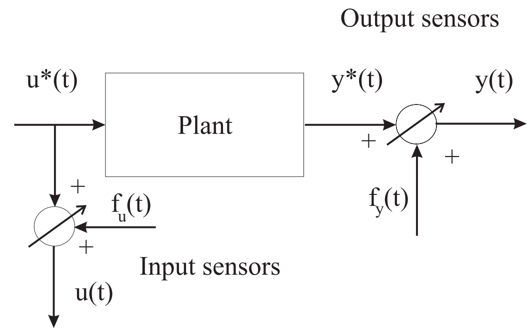


Fig. 1. The structure of the monitored system.

real applications variables $u^*(t)$ and $y^*(t)$ are measured by means of sensors whose outputs are affected by noise.

Neglecting sensor dynamics, faults on the measured input and output signals $u(t)$ and $y(t)$ are modelled as:

$$\begin{cases} u(t) &= u^*(t) + f_u(t) \\ y(t) &= y^*(t) + f_y(t) \end{cases} \qquad (12)$$

in which, the vectors $f_u(t) \in \Re^r$ and $f_y(t) \in \Re^m$ is composed of additive signals which assume values different from zero only in the presence of faults.

Usually these signals are described by step and ramp functions representing, respectively, abrupt and incipient faults (bias or drift).

There are different approaches to generate the diagnostic signals, residuals or symptoms, from which it will be possible to diagnose faults associated to sensors. In this work, a model-based approach is used to estimate the outputs of the system from the input-output measurements.

As depicted in Figure (2), residuals can be generated by the comparison of measured and estimated outputs

$$r(t) = \hat{y}(t) - y(t). \qquad (13)$$

The symptom evaluation refers to a logic device which processes the redundant signals generated by the first block in order to estimate when a fault occurs and to univocally identify the unreliable sensor. Faults can be detected by using a simple thresholding logic:

$$|r(t)| \begin{cases} \leq \text{Threshold} & \text{Fault-free conditions,} \\ > \text{Threshold} & \text{Faulty conditions.} \end{cases} \qquad (14)$$

## V. SUGAR CANE CRUSHING PROCESS

The proposed methodology was applied to the identification and fault diagnosis of a sugar cane crushing mill.

The plant is shown in Figure (3) where the inputs (turbine speed and chute flap) and the outputs (turbine torque and

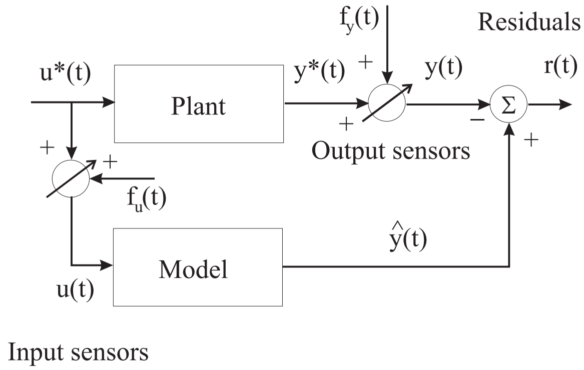


Fig. 2. The residual generation scheme.

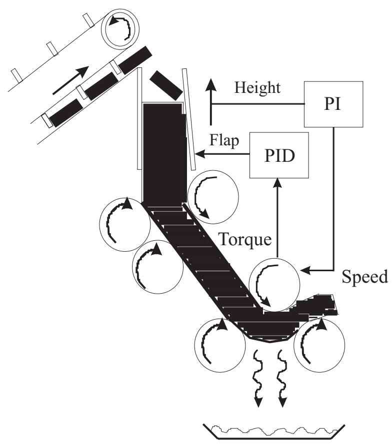


Fig. 3. The structure of the sugar cane crushing process.

chute height) are highlighted. The available data from the control inputs ($r = 2$) were 1400 samples from normal operating records of $u_1(t)$ (turbine speed) and $u_2(t)$ (chute flap). The data from the output sensors ($m = 2$) were the corresponding values of $y_1(t)$ (turbine torque) and $y_2(t)$ (chute height). The sampling time was of $0.1$ s.

Because of the underlying physical mechanisms and because of the modes of the control signals, the process has non–linear steady state as well as dynamic characteristics.

The GK clustering algorithm was used with $M = 3$ clusters (operating conditions) for each output and $n = 3$ the number of shifts of inputs and outputs.

After clustering, the system parameters $\theta_i$, with $i = 1, \cdots, M$ for each output, were estimated using the Frisch scheme identification method. The model was then validated on a separate data set.

The $i$-th output $y(t)$ of the plant ($i = 1, \cdots, m$ and $m = 2$) can be characterised as a TS fuzzy multiple–input single–output (MISO) model (3) with $r = 2$ inputs.

The mean square errors of the output estimation errors $r(t)$, under no-fault conditions, are $0.2549$ for the first output and $0.0125$ for the second one. The fuzzy multiple models approximate the real process very accurately.

Using these models, a model-based approach for fault diagnosis can be exploited and applied to the actual process.

Single faults in the sensors were generated by adding variations (step functions of different amplitudes) in the input and output sensor signals. It was decided to consider fault occurrences during a transient since, in this case, the residual error due to model approximation is maximum and therefore it represents the most critical case in fault detection [6], [7], [8], [9].

Single faults occurring on the input or on the output sensors cause alteration of the sensor signals $u(t)$ and $y(t)$ and therefore of the residuals $r(t)$ given by the predictive model (3) using $u(t)$ as inputs. Residuals indicate fault occurrence according to (14) whether their values are lower or higher than the thresholds fixed in fault-free conditions.

In order to determine the thresholds above which the faults are detectable, the simulation of different amplitude faults in the sensor signals was performed. The threshold value depends on the residual error amount due to the model approximation. These thresholds were settled on the basis of fault-free residuals. A margin of $10\%$ between the thresholds and the residual values was imposed.

To summarise the performance of the FDI technique, the minimal detectable faults on the various sensors, expressed as per cent of the mean values of the relative signals, are collected in Table (I), in case of step faults. The minimum values shown in Table (I) are relative to the case in which the fault must be detected as soon as it occurs.

TABLE I
MINIMAL DETECTABLE STEP FAULTS.

| Sensor | $u_1(t)$ | $u_2(t)$ | $y_1(t)$ | $y_2(t)$ |
|---|---|---|---|---|
| Value | $2.0\%$ | $2.5\%$ | $1.0\%$ | $7.0\%$ |

## VI. CONCLUSIONS

In this paper an off–line procedure was proposed for the identification and fault diagnosis of an industrial process using a fuzzy multiple model identified from noisy input-output measurements. A fuzzy multiple model consists of several local linear models each for different operating point of the process. The identification algorithm was based on fuzzy clustering in order to determine the regions in which measured data can be approximated by affine local dynamic models. Parameters and orders of submodels were estimated using a technique based on the rules of the Frisch scheme. This identification approach gives a reliable model of the plant under investigation which can be exploited to generate redundant residuals for fault diagnosis. The effectiveness of these procedures were tested on real data acquired from a sugar cane crushing mill.

## REFERENCES


[1] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publisher, January 1999.

[2] J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*. Basel, Switzerland: Marcel Dekker, 1998.

[3] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its application to modelling and control," in *IEEE Trans. Sys. Man. & Cyber.*, vol. 15, pp. 116–132, 1985.

[4] I. J. Leonaritis and S. A. Billings, "Input-output parametric models for non-linear systems," *International Journal of Control*, vol. 41, pp. 303–344, 1985.

[5] R. Babuska, J. Keizer, and M. Verhaegen, "Identification of nonlinear dynamic systems as a composition of local linear parametric or state space models," in *Proc. of SYSID'97*, (Fukuoka, Japan), 1997.

[6] R. Rovatti, C. Fantuzzi, and S. Simani, "High–speed DSP–based implementation of piecewise–affine and piecewise–quadratic fuzzy systems," *The Signal Processing Journal*, vol. 80, pp. 951–963, November 2000.

[7] S. Simani, C. Fantuzzi, R. Rovatti, and S. Beghelli, "Parameter identification for piecewise linear fuzzy models in noisy environment," *International Journal of Approximate Reasoning*, vol. 1–2, pp. 149–167, November 1999.

[8] C. Fantuzzi, S. Simani, S. Beghelli, and R. Rovatti, "Identification of piecewise affine models in noisy environment," *International Journal of Control*, vol. 75, no. 18, pp. 1472–1485, 2002.

[9] S. Simani, C. Fantuzzi, and R. J. Patton, *Model-based fault diagnosis in dynamic systems using identification techniques*. Advances in Industrial Control, London, UK: Springer–Verlag, first ed., November 2002. ISBN 1852336854.

[10] R. Frisch, *Statistical Confluenece Analysis by Means of Complete Regression Systems*. University of Oslo, Economic Institute, publication n. 5 ed., 1934.

[11] S. Beghelli, R. P. Guidorzi, and U. Soverini, "The Frisch scheme in dynamic system identification," *Automatica*, vol. 26, no. 1, pp. 171–176, 1990.

[12] T. Söderström and P. Stoica, *System Identification*. Englewood Cliffs, N.J.: Prentice Hall, 1987.

[13] D. E. Gustafson and W. C. Kessel, "Fuzzy Clustering with a Fuzzy Covariance Matrix," in *Proc. IEEE CDC*, (San Diego, CA, USA), pp. 161–166, 1979.

[14] S. Beghelli, P. Castaldi, R. P. Guidorzi, and U. Soverini, "A comparison between different model selection criteria in Frisch scheme identification," *Systems Science Journal*, vol. 20, pp. 77–84, Sept. 1994. Wroclaw, Polonia.