APPROXIMATION OF NON–LINEAR SYSTEMS WITH IDENTIFIED HYBRID MODELS

Silvio Simani * **Cesare fantuzzi** **

* *Dipartimento di Ingegneria, Università di Ferrara*
Via Saragat 1, 44100 Ferrara - ITALY
Tel: +39 0532 97 4844, fax: +39 0532 97 4870
e-mail: ssimani@ing.unife.it
** *Dipipartimento di Scienze e Metodi per l'Ingegneria*
Università di Modena e Reggio Emilia, ITALY

Abstract: This paper addresses the identification of non–linear dynamic systems. A wide class of these systems can be described using non–linear time-invariant regression models, that can be approximated by means of piecewise affine prototypes with an arbitrary degree of accuracy. This work concerns the identification of piecewise affine model structure through input–output data acquired from a dynamic process. In order to show the effectiveness of the developed technique, the results obtained in the identification of both a simple simulated system and a real dynamic process are reported.

Keywords: Hybrid models, dynamic system identification, parameter estimation, data processing, non–linear systems.

1. INTRODUCTION

The key step towards system analysis and control design is to find a suitable mathematical description of the process under investigation. In some cases system modelling based on insight on the physical laws which govern the real process behaviour might be cumbersome and practically infeasible. On the other hand, input–output process measurements can be successfully used to infer an analytical description of the system in the framework of a parametric structure which possess approximation properties with respect to the complex, non–linear, unknown analytical functions that are amenable as candidate to describe the real behaviour of the observed process (Juditsky *et al.*, 1995). The problem of identification of hybrid dynamical systems, by using hyperplanes and Wiener piecewise affine autoregressive exogenous models was also addressed in (Roll *et al.*, 2004). The approach suggested in this work refers to non–linear processes that operate at different regimes, as it usually assumed in several engineering fields, in which distinct models are associated to each admissible operating condition. A switching function governs the transition among different models or interpolations of models. Such mathematical descriptions are referred in current literature as piecewise models or hybrid models, as suggested by several authors (Bemporad and Morari, 1999; Sontag, 1981). The interpolation properties of such models with respect to non–linear discrete–time regression functions are exploited. Particular emphasis has been given to the piecewise affine model identification when data are acquired form a dynamic process. The presented approach differs from the cited bibliography as it uses Error–in–Variables Models (EIV) with the assumption suggested by Frisch (Frisch,

1934) and revisited by Kalman (Kalman, 1982) regarding the stochastic characterisation of the noises corrupting the data. In this context, it is possible to generalise some basic results on single model proposed in (Beghelli *et al.*, 1990), which regard model structure and parameter identification with measurement noise rejection.

2. MODEL DESCRIPTION

The main idea underlying the mathematical description of non–linear dynamic systems is based on the interpretation of single input–single output, non–linear, time–invariant regression models such as:

$$y(t+n) = F\big(y(t+n-1),\cdots,y(t), u(t+n-1),\cdots,u(t)\big) \quad (1)$$

for $t = 0, 1, \ldots$. $u(\cdot)$ and $y(\cdot)$ belong respectively to the bounded input $\mathcal{U}$ and output $\mathcal{Y}$ sets, n is the finite system memory (*i.e.* the model order) and $F(\cdot)$ is a continuous non–linear function defining a hypersurface from a $\mathcal{A}_n$ to $\mathcal{Y}$, being $\mathcal{A}_n$ the Cartesian product $\mathcal{U}^n \times \mathcal{Y}^n$. The identification of the non–linear system can be translated to the approximation of its mathematical model given by Eq. (1) using a parametric structure that exhibits arbitrary accuracy interpolation properties. A piecewise model defined through the composition of simple models having local validity is the natural candidate to perform this task, as it combines function interpolation properties with mathematical tractability. In the following the proposed piecewise structure is defined and its properties in terms of interpolation characteristics of arbitrary non–linear functions are briefly discussed. The piecewise model is formed by a collection of parametric submodels of the type

$$y(t+n) = \sum_{j=0}^{n-1} \alpha_j^{(i)} y(t+j) + \sum_{j=0}^{n-1} \beta_j^{(i)} u(t+j) + b^{(i)} \quad (2)$$

in which the system operating point is described by the input and output samples $y(t+n-1),\cdots,y(t)$ and $u(t+n-1),\cdots,u(t)$, that can be collected with a vector $\mathbf{x}_n(t) = [y(t),\cdots,y(t+n-1), u(t),\cdots,u(t+n-1)]^T$, for $t = 0, 1, \ldots$. The *switching* function $\chi_i\big(\mathbf{x}_n(t)\big), i = 1,\ldots,M$ is:

$$\chi_i\big(\mathbf{x}_n(t)\big) = \begin{cases} \chi_i\big(\mathbf{x}_n(t)\big) = 1 \text{ if } \mathbf{x}_n(t) \in \mathcal{A}_n^{(i)} \\ \chi_i\big(\mathbf{x}_n(t)\big) = 0 \text{ otherwise} \end{cases} \quad (3)$$

where $\{\mathcal{A}_n^{(1)},\ldots,\mathcal{A}_n^{(M)}\}$ is a partition of $\mathcal{A}_n$, whose structure will be characterised in the following. Thus, the output $y(t+n)$ of the non–linear dynamic system described by Eq. (1) can be approximated by the *piecewise affine model* $f(\cdot)$ in the form

$$y(t+n) = f\big(\mathbf{x}_n(t)\big) = \sum_{i=1}^{M} \chi_i\big(\mathbf{x}_n(t)\big)[\mathbf{x}_n(t),1]^T \mathbf{a}_n^{(i)} \quad (4)$$

where the model parameters are collected in the vector $\mathbf{a}_n^{(i)} = [\alpha_0^{(i)},\ldots,\alpha_{n-1}^{(i)}, \beta_0^{(i)},\ldots,\beta_{n-1}^{(i)}, b^{(i)}]^T$.

It is worthwhile noting that the model is affine in each $\mathcal{A}_n^{(i)}$, $\mathbf{a}_n^{(i)}$ being the affine submodel parameters.

It is worth noting that this piecewise affine function aims to interpolate an arbitrary non–linear model, conceptually enabling us to capture the behaviour of any physical process. Thus, it is important to underline the approximation capabilities of this piecewise affine model with respect to the target function. On the other hand, since the model in the form of Eq. (1) is supposed continuous, $f(\cdot)$ is forced to be continuous over the whole $\mathcal{A}_n$. In such a case the parameter vectors are constrained to satisfy the following relation:

$$\lim_{\mathbf{x}_n(t)\to\bar{\mathbf{x}}_n, \mathbf{x}_n(t)\in\mathcal{A}_n^{(i')}} f\big(\mathbf{x}_n(t)\big) = \lim_{\mathbf{x}_n(t)\to\bar{\mathbf{x}}_n, \mathbf{x}_n(t)\in\mathcal{A}_n^{(i'')}} f\big(\mathbf{x}_n(t)\big) \quad (5)$$

$\bar{\mathbf{x}}_n$ being an accumulation point for both $\mathcal{A}_n^{(i')}$ and $\mathcal{A}_n^{(i'')}$, *i.e.* if

$$[\bar{\mathbf{x}}_n(t),1]^T\ \mathbf{a}_n^{(i')} = [\bar{\mathbf{x}}_n(t),1]^T\ \mathbf{a}_n^{(i'')} \quad (6)$$

The straightforward application of Eq. (6) to all the accumulation points common to neighbouring regions leads to an infinite number of constraints. However, the adoption of regions with straight borders guarantees that only a finite number of them is linearly independent. This results suggests that regions whose boundaries are convex polyhedra should be considered. In this case, in fact, continuity can be ensured simply by setting the value of the local models only on the vertices of the boundaries. In this case, the continuity constraints (one for each polyhedral vertex) can be collected in a finite matrix C_n such that:

$$C_n A_n = \mathbf{0} \text{ being } A_n = \Big[\ \mathbf{a}_n^{(1)^T}\ \ldots\ \mathbf{a}_n^{(M)^T}\ \Big]^T. \quad (7)$$

In particular, it is undoubtedly convenient to "triangulate" the domain $\mathcal{A}_n$, *i.e.* to partition it into $2n$–dimensional simplexes. Moreover, we will assume that the triangulation is such that two simplexes are either disjoint, or have in common a whole k–dimensional boundary, with $k = 0, 1, \ldots, 2n-1$. In this way, the local affine model described by Eq. (4) can be forced to assume given values at most in $2n+1$ vertices of each simplex, which are affinely independent points.

3. LOCAL IDENTIFICATION

Let us assume that the input–output data $u(t)$ and $y(t)$, $(t = 0, 1, , \ldots, L_i)$ generated by a system of the type of Eq. (2) are available. Restricting our investigation to find order n and parameters $\mathbf{a}_n^{(i)}$ for local model in the form of Eq. (2) in region $\mathcal{A}_n^{(i)}$, the following matrix should be defined:

$$X_k^{(i)} = \begin{bmatrix} y(k) & \mathbf{x}_k^T(0) & 1 \\ y(k+1) & \mathbf{x}_k^T(1) & 1 \\ \vdots & \vdots & \\ y(k+N_i-1) & \mathbf{x}_k^T(N_i-1) & 1 \end{bmatrix} \tag{8}$$

$$\Sigma_k^{(i)} = \left(X_k^{(i)}\right)^T X_k^{(i)}$$

with $k + N_i - 1 \leq L_i$ and N_i is chosen so that $k+N_i-1$ is large enough to avoid unwanted linear dependence relationships due to limitations in the dimension of the vector spaces involved.

To determine the model order n in region $\mathcal{A}_n^{(i)}$, it is possible to consider the sequence of increasing–dimension positive definite or positive semidefinite $\big((2k+2)\times(2k+2)\big)$ symmetric matrices

$$\Sigma_2^{(i)}, \;\; \Sigma_3^{(i)}, \;\; \ldots \;\; \Sigma_k^{(i)}, \;\; \ldots \tag{9}$$

testing their singularity as k increases. As soon as a singular matrix $\Sigma_k^{(i)}$ is found then $n = k$, and the parameters $\mathbf{a}_n^{(i)}$ describe the dependence relationship of the first vector of $\Sigma_n^{(i)}$ on the remaining ones as

$$\Sigma_n^{(i)} \left[-1 \;\; \mathbf{a}_n^{(i)\,T} \right]^T = 0 \tag{10}$$

It is worth noting that the vectors $\mathbf{x}_n(0)$, $\mathbf{x}_n(1)$, …, $\mathbf{x}_n(N_i - 1)$ in equation (8) must belong to the region $\mathcal{A}_n^{(i)}$ according to the partition defined in equation (3). Note also that in the presence of noise the above procedure described to determine order and model parameters would obviously be useless since matrices Σ_k would always be non–singular (positive definite) (Beghelli *et al.*, 1990).

In order to solve the problem in a mathematical framework, it is necessary to characterise the noise affecting the input-output data. Following common assumptions (Frisch, 1934; Kalman, 1982; Beghelli *et al.*, 1990), the noises $\tilde{u}(t)$ and $\tilde{y}(t)$ are assumed additive on input-output data $u^*(t)$ and $y^*(t)$ and region independent, so that

$$\begin{cases} u(t) \;=\; u^*(t) + \tilde{u}(t) \\ y(t) \;=\; y^*(t) + \tilde{y}(t). \end{cases} \tag{11}$$

Obviously, only $u(t)$ and $y(t)$ are available for the identification procedure, and moreover every noise term $\tilde{u}(t)$ and $\tilde{y}(t)$ is modelled with a zero–mean white process and is supposed to be independent of every other term. These structures are also commonly known as Error–In–Variables (EIV) models. Under these assumptions, and $\bar{\tilde{\sigma}}_u$ and $\bar{\tilde{\sigma}}_y$ being the input and output noise variances respectively, the generic positive definite matrix $\Sigma_k^{(i)}$ associated with the input-output noise-corrupted sequences can always be expressed as the sum of two terms $\Sigma_k^{(i)} = \Sigma_k^{*(i)} + \bar{\tilde{\Sigma}}_k$ where

$$\bar{\tilde{\Sigma}}_k = \mathrm{diag}[\bar{\tilde{\sigma}}_y I_{k+1}, \bar{\tilde{\sigma}}_u I_k, 0] \geq 0. \tag{12}$$

Thus, it is again possible to determine the order and parameters of the model in region $\mathcal{A}_n^{(i)}$ from the analysis of the sequence of increasing–dimension $\big((2k+2)\times(2k+2)\big)$ symmetric positive definite matrices

$$\Sigma_2^{(i)}, \;\; \Sigma_3^{(i)}, \;\; \ldots \;\; \Sigma_k^{(i)}, \;\; \ldots \tag{13}$$

The solution of the above identification problem requires the computation of the unknown noise covariances $\tilde{\bar{\sigma}}_u$ and $\tilde{\bar{\sigma}}_y$, that can be achieved solving the following relation:

$$\Sigma_k^{*(i)} = \Sigma_k^{(i)} - \tilde{\Sigma}_k \geq 0. \tag{14}$$

in the variables $\tilde{\sigma}_u, \tilde{\sigma}_y$, where $\tilde{\Sigma}_k = \mathrm{diag}\big[\, \tilde{\sigma}_y I_{k+1}, \tilde{\sigma}_u I_k \;, 0\big]$. It is worth noting that the set of values of variables $\tilde{\sigma}_u, \tilde{\sigma}_y$ which make matrix $\Sigma_k^{*(i)}$ positive semidefinite forms a curve. Unfortunately the relation (14) admits for any k an infinite solution set describing a curve $\Gamma_k^{(i)}(\tilde{\sigma}_y, \tilde{\sigma}_u) = 0$ in the first orthant of the noise plane whose concavity faces the origin. In (Beghelli *et al.*, 1990) a constructive methodology to numerically compute this curve is given. 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. As proven in (Beghelli *et al.*, 1990), the solution sets of condition (14) for different values of k are non–crossing curves in the noise plane $(\tilde{\sigma}_y, \tilde{\sigma}_u)$.

It is also important to observe that, since we assume that a system described by Eq. (2) has generated the noiseless data, for $k \geq n$ all the curves in the form of Eq. (14) have necessarily at least one common point, *i.e.* point $(\bar{\tilde{\sigma}}_u, \bar{\tilde{\sigma}}_y)$ corresponding to the true variances of the noise affecting the input and the output data. The search for a solution for the identification problem can thus start from the determination of this point in the noise space. This task can be achieved on the basis of the following properties: With reference to the diagonal non–negative definite matrices $\bar{\tilde{\Sigma}}_k$, the following properties hold: *(i)* If $k < n$ the matrices $\Sigma_k^{*(i)}$ are positive definite. *(ii)* If $k > n$ the dimension of the null space of $\Sigma_k^{*(i)}$ and consequently, the number of eigenvalues equal to zero is $(k-n+1)$. *(iii)* For $k = n$, matrix $\Sigma_k^{*(i)}$ is characterised by a linear dependence relation among its $2k+2$ vectors, and the coefficients which link the first vector of $\Sigma_k^{*(i)}$ to the remaining ones are the parameters $\mathbf{a}_n^{(i)}$, of the system described by Eq. (2) which has generated the noiseless sequences. *(vi)* For $k \geq (n+1)$, all the $k - n + 1$ linear dependence relations among the vectors of the matrix Σ_k^* are characterised by the same $2n+2$ coefficients $\mathbf{a}_n^{(i)}$. Figure (1) shows the above properties for a system in the form of Eq. (2) with $n = 3$. The point marked by a circle corresponds to the input-output noise variances $\bar{\tilde{\sigma}}_y$ and $\bar{\tilde{\sigma}}_u$ affecting the measurements.

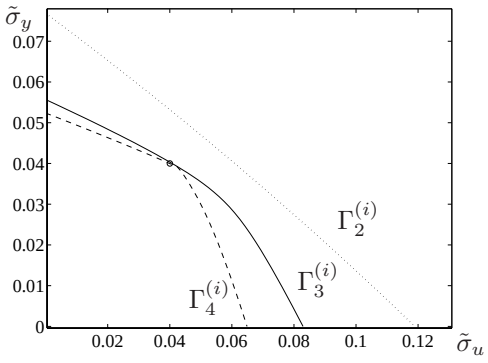

Fig. 1. Singularity curves in the noise space.

4. GLOBAL MODEL IDENTIFICATION

In the previous section we discussed a procedure for the identification of the noise variances $\bar{\tilde{\sigma}}_u$ and $\bar{\tilde{\sigma}}_y$ and of the system order n, with respect to a particular region $\mathcal{A}_n^{(i)}$.

If the noise characteristics are common to all the regions $\mathcal{A}_n^{(i)}$, since the physical nature of the process generating the noise is independent of the model structure and of the partition of $\mathcal{A}_n$, and all assumptions regarding the Frisch scheme are fulfilled, a common point $(\bar{\tilde{\sigma}}_y, \bar{\tilde{\sigma}}_u)$ in the noise plane exists for the singularity curves. When the order n has been determined, the parameters $\mathbf{a}_n^{(i)}, i = 1, \ldots, M$ can be identified solving the following equation:

$$(\Sigma_n^{(i)} - \bar{\tilde{\Sigma}}_n)\mathbf{a}_n^{(i)} = \mathbf{0} \quad \text{for } i = 1, \ldots, M. \tag{15}$$

The previous result can be fully applied when the assumptions behind the Frisch scheme are satisfied (independence between input–output sequences, additive noise, noise whiteness).

In real applications, we are forced to relax these assumptions, thus no common point can be determined among curves $\Gamma_n^{(i)} = 0$ in the noise plane and a unique solution to the identification problem can be obtained only by introducing a criterion to select a different noisy point for each region as best approximation of the ideal case. With reference to the identification of the system order n in the i–th region $\mathcal{A}_n^{(i)}$, it must be noted that the $\Gamma_{n+1}^{(i)} = 0$ curve has a single point in common with the $\Gamma_n^{(i)} = 0$ curve in ideal conditions, which corresponds to a double singularity of the matrix $\Sigma_{n+1}^{*(i)}$. In real cases, the order n can be computed finding the point $(\tilde{\sigma}_u, \tilde{\sigma}_y) \in \Gamma_{n+1}^{(i)} = 0$ that makes $\Sigma_{n+1}^{*(i)}$ closer to the double singular condition (*i.e.* minimal eigenvalue equal to zero and the second minimum eigenvalue near to zero). As n is unknown, increasing system orders k must be tested, and the value of k associated to the minimum of the second eigenvalue of the matrix $\Sigma_{k+1}^{*(i)}$ corresponds to the order n. This criterion is consistent as it leads to the common point of the curves when the assumptions of the Frisch scheme are not violated.

Note that since the order n of the piecewise model described by Eq. (4) is region independent, it can be determined by choosing k that fulfil the following inequality

$$\max_{i=1,\ldots,M_k} \lambda_k^{(i)} < \epsilon \tag{16}$$

when ϵ is an arbitrary positive constant and $\lambda_k^{(i)}$ is the minimal eigenvalue different from zero of matrix $\Sigma_{k+1}^{*(i)}$. This result led to derive the following algorithm for selection of the model order: (i) fix ϵ, k and M_k (k is the initial hypothesis on model order). (ii) Construct partition $\{\mathcal{A}_k^{(1)}, \ldots, \mathcal{A}_k^{(M_k)}\}$. (iii) Cluster data into partition. Compute matrices $\Sigma_{k+1}^{*(i)}$ from data clustered in region $\mathcal{A}_k^{(i)}$. (vi) Compute test (16): *If success*: $n = k$, *exit else* $k = k + 1$, *goto* (ii).

Once the model order n is selected, the parameters $\mathbf{a}_n^{(i)}, i = 1, \ldots, M$ cannot be computed from Eq. (15), as the curves $\Gamma_n^{(i)} = 0$ do not share the common point $(\bar{\tilde{\sigma}}_u, \bar{\tilde{\sigma}}_y)$. In this case, for each region a different noise $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$ must be considered and relation (14) should be rewritten as

$$\Sigma_n^{*(i)} = \Sigma_n^{(i)} - \tilde{\Sigma}_n^{(i)} \geq 0 \tag{17}$$

where $\tilde{\Sigma}_n^{(i)} = \mathrm{diag}[\bar{\tilde{\sigma}}_u^{(i)} I_{n+1}, \bar{\tilde{\sigma}}_y^{(i)} I_n, 0]$. The values $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$ can be computed by solving an optimisation problem which minimises both the distances between $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$ and $(\bar{\tilde{\sigma}}_u^{(j)}, \bar{\tilde{\sigma}}_y^{(j)})$ with $i \neq j$ and the continuity constraints described by Eq. (7)

$$\begin{aligned} &J\big((\bar{\tilde{\sigma}}_u^{(1)}, \bar{\tilde{\sigma}}_y^{(1)}), \ldots, (\bar{\tilde{\sigma}}_u^{(M)}, \bar{\tilde{\sigma}}_y^{(M)})\big) = \\ &d\big((\bar{\tilde{\sigma}}_u^{(1)}, \bar{\tilde{\sigma}}_y^{(1)}), \ldots, (\bar{\tilde{\sigma}}_u^{(M)}, \bar{\tilde{\sigma}}_y^{(M)})\big) + (C_n A_n)^T H C_n A_n \end{aligned} \tag{18}$$

H being a definite positive weighting matrix and $d(\cdot)$ a distance defined as

$$\begin{aligned} &d\left((\bar{\tilde{\sigma}}_u^{(1)}, \bar{\tilde{\sigma}}_y^{(1)}), \ldots, (\bar{\tilde{\sigma}}_u^{(M)}, \bar{\tilde{\sigma}}_y^{(M)})\right) = \\ &= \sum_{i=1}^{M} \sum_{j=i+1}^{M} \sqrt{(\bar{\tilde{\sigma}}_u^{(i)} - \bar{\tilde{\sigma}}_u^{(j)})^2 + (\bar{\tilde{\sigma}}_y^{(i)} - \bar{\tilde{\sigma}}_y^{(j)})^2}. \end{aligned} \tag{19}$$

It is worth observing that the matrix A_n collects the parameters $\mathbf{a}_n^{(i)}, i = 1, \ldots, M$ which depend on $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$.

Minimisation of cost function described by Eq. (18) can be computationally difficult, as it depends on $2M$ independent variables. Therefore, in order to decrease the complexity of the problem, a common parametrisation can be defined for all the curves $\Gamma_n^{(i)}(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)}) = 0$ by introducing polar coordinates

$$\begin{cases} \bar{\tilde{\sigma}}_u^{(i)} = \rho^{(i)} \cos \dfrac{\pi}{2} q \\ \bar{\tilde{\sigma}}_y^{(i)} = \rho^{(i)} \sin \dfrac{\pi}{2} q \end{cases} \tag{20}$$

where $\rho^{(i)}$ is determined so that $\Gamma_n^{(i)}\ \big(\rho^{(i)}\cos\frac{\pi}{2}q,\ \rho^{(i)}\sin\frac{\pi}{2}q\big)=0$ and $q\in[0,1]$. In such a way, the cost function has the form:

$$J(q)=d\Big((\bar{\tilde{\sigma}}_u^{(1)}(q),\bar{\tilde{\sigma}}_y^{(1)}(q)),\ldots,(\bar{\tilde{\sigma}}_u^{(M)}(q),\bar{\tilde{\sigma}}_y^{(M)}(q))\Big)+ \\ +(C_nA_n)^THC_nA_n. \tag{21}$$

The parametrisation chosen to simplify the minimisation problem leads to consistent results. In fact, when the data are generated by a continuous piecewise affine dynamic system, all assumptions regarding the Frisch scheme being fulfilled and noise being region-independent, the curves $\Gamma_n^{(i)}=0$ share a common point in the noise plane. In these conditions, cost function $J(q)=0$ and the variances $(\bar{\tilde{\sigma}}_u,\bar{\tilde{\sigma}}_y)$ are identified exactly.

Finally, one should note how once the parameter q minimising the cost function in the form of Eq. (21) is computed, the matrices $\tilde{\Sigma}_n^{(i)}$ can be built and the model parameter $\mathbf{a}_n^{(i)}, i=1,\ldots,M$ determined by means of relation:

$$(\Sigma_n^{(i)}-\tilde{\Sigma}_n^{(i)})\mathbf{a}_n^{(i)}=\mathbf{0}\quad \text{for } i=1,\ldots,M. \tag{22}$$

This completes the multiple model identification procedure.

5. APPLICATION EXAMPLE

To summarise the results which have been proposed in the previous sections, the identification of a piecewise affine model for an industrial dynamic process is considered. The process under consideration is a 40l laboratory fermenter which contains 25l of water. At the bottom of the fermenter, air is fed into the water at a specified flow rate which is kept at a desired value by a local mass–flow controller. The air pressure in the head space can be controlled by the position of an outlet valve at the top of the fermenter as shown in figure (2). This process has two inputs: the position of the outlet valve, denoted by $u(t)$, and the inlet air flow rate and one output, *i.e.* the pressure in the head space, denoted by $y(t)$. The inlet flow rate can be kept constant, in which case the process is a single–input, single–output system. Because of the underlying physical mechanisms, and because of the non–linear characteristic of the outlet valve, the process has a non–linear behaviour.

Two series of data ($N=380$) were acquired from the process. The first one has been used for model identification and the second one, has been exploited for the validation task.

According to the algorithm derived in Section 4 for the selection of the model order, the initial value of $k=1$ and $\epsilon=10^{-7}$ have been fixed.

Under these assumptions, the triangulation of the 2-dimensional domain $\mathcal{U}\times\mathcal{Y}$ into simplexes has been performed. The partition of the domain was obtained by exploiting the MATLAB© toolbox for piecewise linear approximation of multidimensional mappings. As shown previously, the partition of the domain for the real process with $k=1$ has been achieved by considering the Cartesian product of the intervals $\mathcal{I}_i^{\mathcal{U}}$ and $\mathcal{I}_i^{\mathcal{Y}}$.

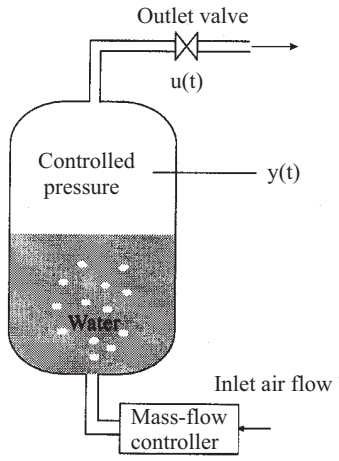

Fig. 2. The fermentation tank.

In can be also shown that $M_1=3$ regions only contain enough data N_i $(i=1,2,3)$ for applying Eqs. (8) and to perform the identification task. Three models have therefore been estimated. In this case, $\mathbf{x}_1(t)=[y(t),u(t)]^T$ and the data belonging to the domain $\mathcal{U}\times\mathcal{Y}$ have been clustered into the considered partition $\{\mathcal{A}_1^{(4)},\mathcal{A}_1^{(6)},\mathcal{A}_1^{(7)}\}$ $(k=1,\ M_1=3)$, the $\Sigma_2^{*(i)}$ matrices in Eq. (17) $(i=4,6,7)$ have been computed and the test of Eq. (16) performed. In such a case, $\max_{i=4,6,7}\lambda_k^{(i)}=2.4765\times10^{-9}$. This value is below the selected accuracy ϵ, so the model order can be estimated as $n=1$. In such a situation, with $\mathrm{H}=0.01\ \mathrm{I}_3$ the parameters for the identified piecewise models are $\mathbf{a}_1^{(4)}=[6.73\cdot10^{-1},9.33\cdot10^{-4},3.37\cdot10^{-1}]^T$, $\mathbf{a}_1^{(6)}=[8.17\cdot10^{-1},3.17\cdot10^{-3},2.51\cdot10^{-2}]^T$ and $\mathbf{a}_1^{(7)}=[9.13\cdot10^{-1},5.69\cdot10^{-1},-3.02\cdot10^{-1}]^T$.

The identified model with $n=1$ can be therefore represented by the following piecewise affine system, $y(t+1)=f\big(y(t),u(t)\big)$, defined as:

$$y(t+1)=\begin{cases}6.73\cdot10^{-1}\ y(t)+9.33\cdot10^{-4}\ u(t)+\\ +3.37\cdot10^{-1}\ \text{if } \mathbf{x}_1(t)\in\mathcal{A}_1^{(4)}\\ 8.17\cdot10^{-1}\ y(t)+3.17\cdot10^{-3}\ u(t)+\\ +2.51\cdot10^{-2}\ \text{if } \mathbf{x}_1(t)\in\mathcal{A}_1^{(6)}\\ 9.13\cdot10^{-1}\ y(y)+5.69\cdot10^{-1}\ u(t)+\\ -3.02\cdot10^{-1}\ \text{if } \mathbf{x}_1(t)\in\mathcal{A}_1^{(7)}\end{cases} \tag{23}$$

The standard deviation of the noise on the identification input data was estimated to be 14% with respect to the standard deviation of the measurement of the input sequence. For the output sequence, a value of 21% was identified. In such

case, with identification data, the mean square error between measured and simulated outputs is 0.0044. On the other hand, figure (4) shows the comparison between the output $y(t)$ acquired from the real process and the output $y_{\text{est}}(t)$ given by the first order estimated piecewise affine model given by Eq. (23) when they are both driven by the *validation* input $u(t)$. Using validation data, the mean square error between measured and simulated outputs is equal to 0.0066. It is worth-

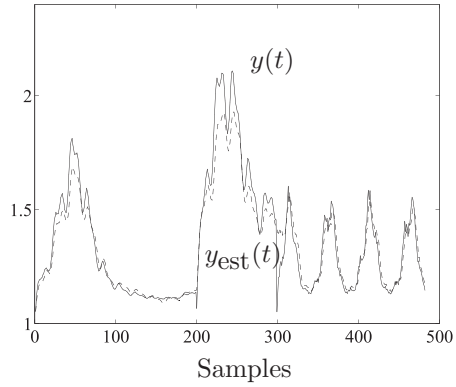

Fig. 3. Comparisons between the process and the estimated model outputs with identification data.

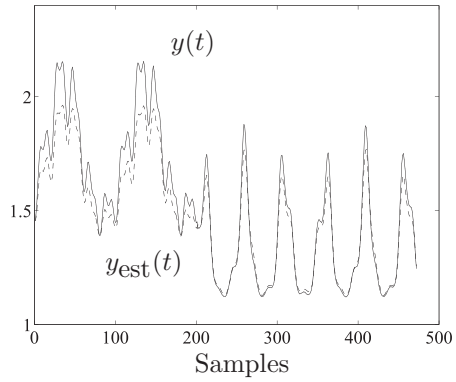

Fig. 4. Comparisons between the process and the estimated model outputs with validation data.

while noting how the validation was performed on a different data set than the one used for identification. The estimated output was obtained from the input in *full simulation*: $y_{\text{est}}(t+1) = f\big(y_{\text{est}}(t), u(t)\big)$, which imposes stronger requirements on its accuracy than the one–step–ahead prediction which uses the process output $y(t)$ instead of the predicted value $y_{\text{est}}(t)$. From figure (4) one can see that the identified model follows the measured process output with a reasonable accuracy.

Finally, more experiments were performed by increasing the order k of the considered model and the number of the domain partition M_k. The results are reported in table 1.

Table 1. Mean square errors when order and number of clusters vary.

Clusters/Order	$k=1$	$k=2$	$k=3$
$M_k=1$	0.0405	0.0456	0.0454
$M_k=2$	0.0078	0.2313	0.1037
$M_k=3$	**0.0066**	0.1748	0.1234
$M_k=4$	0.0089	0.2239	0.0768
$M_k=5$	0.0117	0.7869	0.9463

It is worth noting how the identified model described by Eq. (23) represents a trade–off between simulation accuracy (also dependent on the available data in each region) and structure complexity.

6. CONCLUSION

In this paper an off–line procedure was proposed for the approximation of a non–linear, discrete–time, continuous dynamic system from real data, using a multiple model approach. The multiple model consists of several local affine models each describing a different operating condition of the process. The identification approach exploited to estimate order and parameters of the local affine models is based on the Frisch scheme method for EIV models. Furthermore, an optimisation technique taking into account both continuity constraint fulfilment and EIV noise assumptions was considered.

REFERENCES

Beghelli, S., R. P. Guidorzi and U. Soverini (1990). The Frisch scheme in dynamic system identification. *Automatica* **26**(1), 171–176.

Bemporad, A. and M. Morari (1999). Control of systems integrating logic, dynamics, and constraints. *Automatica* **35**(3), 407–428.

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

Juditsky, A., H. Hjalmarsson, A. Beneviste, B. Delyon, L. Ljung, J. Sjöberg and Q. Zhang (1995). Nonlinear black–box modelling in system identification: a mathematical foundation. *Automatica* **31**(12), 1691–1724.

Kalman, R. E. (1982). System Identification from Noisy Data. In: *Dynamical System II* (A. R. Bednarek and L. Cesari, Eds.). pp. 135–164. Academic Press. New York.

Roll, Jacob, A. Bemporad and Lennart Ljung (2004). Identification of piecewise affine systems via mixed-integer programming. *Automatica* **40**(1), 37–50.

Sontag, E.D. (1981). Nonlinear regulation: The piecewise linear approach. *IEEE Trans. on Automatic Control* **26**, 346–358.