

Multiobjective Identification of Takagi–Sugeno Fuzzy Models

Tor A. Johansen and Robert Babuška

Abstract—The problem of identifying the parameters of the constituent local linear models of Takagi–Sugeno fuzzy models is considered. In order to address the tradeoff between global model accuracy and interpretability of the local models as linearizations of a nonlinear system, two multiobjective identification algorithms are studied. Particular attention is paid to the analysis of conflicts between objectives, and we show that such information can be easily computed from the solution of the multiobjective optimization. This information is useful to diagnose the model and tune the weighting/priorities of the multiobjective optimization. Moreover, the result of the conflict analysis can be used as a constructive tool to modify the fuzzy model structure (including membership functions) in order to meet the multiple objectives. Simple illustrative examples as well as experimental results show the usefulness of the method.

Index Terms—Interpretability, multiobjective optimization, nonlinear system identification.

I. INTRODUCTION

WE CONSIDER the problem of identifying the parameters of the constituent local linear models of Takagi–Sugeno (T–S) fuzzy models [23]. It is well known that several tradeoffs are involved in this problem.

- Models identified by minimizing the global prediction error need not have constituent local linear models which are interpretable as valid linearizations of the underlying nonlinear system [1], [13], [16], [21], [24].
- When identifying the local linear models by minimizing individual locally weighted prediction error criteria, the identified local linear models have locally valid interpretations as linearizations, under assumptions on identifiability and persistence of excitation [1], [13], [16], [21], [24]. On the other hand, the global prediction performance is typically inferior to what can be achieved with a global performance criterion.
- The use of constraints and regularization are useful to identify interpretable local models, especially when the assumptions on persistence of excitation or identifiability

are not strongly satisfied [2], [8], [9], [13], [17]. For dynamic systems, this is typically the case for local models associated with transient operating regimes [13], [21].

- The lack of local interpretability with global identification is also closely linked to poor identifiability and/or the choice of fuzzy membership functions. This is partly due to the interaction between the local models, or the “degree of orthogonality” among them, which is related to the degree of overlap between their associated membership functions and the degree of smoothness of the model [16]. In other words, with a global identification approach there is a tradeoff between local interpretability and smoothness. However, as shown in [16] it is possible to optimize the individual membership function parameters to address this tradeoff.

This paper is organized as follows. In Section II we review the T–S model and the common local and global least squares algorithms. With a simple example, we illustrate the tradeoff between local model interpretability and global model accuracy. Two multiobjective identification algorithms are described in Sections III and IV together with a conflict/sensitivity analysis. Using the same simple example as above, it is illustrated how these algorithms can be applied to address the tradeoff. A theoretical result on the relationship between the two algorithms is provided in Section V, and in Sections VI and VII we provide application examples before concluding in Section VIII.

II. PRELIMINARIES

A. T–S Fuzzy Model

The framework presented here is the identification of dynamic T–S fuzzy input–output models of the form

$$y(t) = \sum_{i=1}^N (-a_{1,i}y(t-1) - \dots - a_{n_y,i}y(t-n_y) + b_{0,i}u(t) + \dots + b_{n_u,i}u(t-n_u) + d_i)w_i(z(t)) + e(t) \quad (1)$$

where $u(t) \in R^r$ is the input, $y(t) \in R^m$ is the output, $e(t) \in R^m$ accounts for un-modeled phenomena. N is the number of rules/local models, n_u and n_y define the dynamic order, and $z(t) \in R^d$ is a vector of premise variables derived from the information vector

$$\tilde{\psi}(t) = (-y(t-1), \dots, -y(t-n_y), u(t), u(t-1), \dots, u(t-n_u))^T. \quad (2)$$

Manuscript received June 26, 2000; revised September 18, 2001, June 17, 2002, and April 25, 2003. This work was supported by a research fellowship granted to T. A. Johansen by the Delft University of Technology, Delft, The Netherlands.

T. A. Johansen is with the Department of Engineering Cybernetics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway (e-mail: Tor.Arne.Johansen@itk.ntnu.no).

R. Babuška is with the Systems and Control Engineering Group, Department of Electrical Engineering, Delft University of Technology, 2600 GA Delft, The Netherlands (e-mail: R.Babuska@its.tudelft.nl).

Digital Object Identifier 10.1109/TFUZZ.2003.819824

Fuzzy models of the form (1) result from fuzzy inference on a set of fuzzy rules

$$\begin{aligned} \text{IF } z(t) \in Z_i \text{ THEN} \\ y(t) = -a_{1,i}y(t-1) - \dots - a_{n_y,i}y(t-n_y) \\ + b_{0,i}u(t) + \dots + b_{n_u,i}u(t-n_u) + d_i \end{aligned} \quad (3)$$

where the premise is defined by a fuzzy set $Z_i \subset R^p$ and the consequent is a local linear dynamic model. The function $w_i: R^p \rightarrow [0, 1]$ is defined by the membership functions $\mu_i: R^p \rightarrow [0, 1]$ of Z_i

$$w_i(z) = \frac{\mu_i(z)}{\sum_{j=1}^N \mu_j(z)} \quad (4)$$

see [23] for details. The only assumption we make on the set of fuzzy rules is that it is complete in the sense that $\mu_j(z) > 0$ for some j for all z , such that (4) is well defined. Equation (1) can be reformulated into a form that is more convenient for system identification by introducing the definitions

$$\psi(t) = \begin{pmatrix} \tilde{\psi}(t) \\ 1 \end{pmatrix} \quad (5)$$

$$\theta_i(t) = (a_{1,i}, \dots, a_{n_y,i}, b_{0,i}, \dots, b_{n_u,i}, d_i)^T \quad (6)$$

where $\psi(t)$ is the information vector augmented with a constant element, and θ_i are the possibly unknown parameters associated with the local linear model of the i th rule. With these definitions

$$y(t) = \sum_{i=1}^N \psi^T(t)\theta_i w_i(z(t)) + e(t). \quad (7)$$

Furthermore, defining

$$\varphi(t) = \begin{pmatrix} \psi(t)w_1(z(t)) \\ \vdots \\ \psi(t)w_N(z(t)) \end{pmatrix} \quad \theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_N \end{pmatrix} \quad (8)$$

the linear regression form follows:

$$y(t) = \varphi^T(t)\theta + e(t). \quad (9)$$

For the purpose of system identification, assume that a data set $(y(1), y(2), \dots, y(n))$, $(u(1), u(2), \dots, u(n))$ is available. In the remainder of this section we briefly review the well known global and locally weighted identification methods and present a simple example that illustrates some of the problems associated with these simple methods; see [16] for further details. This serves as a background for the remaining parts of this paper, where alternative algorithms and tools are proposed.

B. Global Identification Algorithm

The objective of this algorithm is to identify the local model parameters $\theta_1, \dots, \theta_N$ that give a global model with the best

prediction performance [23]. Consider the global least squares prediction error criterion

$$V(\theta) = \frac{1}{n} \sum_{t=1}^n (y(t) - \varphi^T(t)\theta)^2 \quad (10)$$

subject to equality and inequality constraints on the parameters

$$H_i(\theta_i) = 0 \quad F_i(\theta_i) \leq 0, \quad i = 1, 2, \dots, N \quad (11)$$

where H_i and F_i are affine functions of θ_i . The usefulness of such constraints in order to improve the accuracy and robustness of the estimate is shown in, e.g., [2] and [13] and the references therein. Solving this problem is a convex quadratic program for which efficient standard methods and software exist [15].

C. Locally Weighted Identification Algorithm

The objective of this algorithm is to identify local model parameters $\theta_1, \dots, \theta_N$ that give local models which are close local approximations to the underlying nonlinear system [11]. Consider a locally weighted least squares prediction error criterion associated with each local model

$$V_i(\theta_i) = \frac{1}{n} \sum_{t=1}^n (y(t) - \psi^T(t)\theta_i)^2 w_i(z(t)) \quad (12)$$

subject to the constraints (11). The weighting factor $w_i(z(t))$ ensures that the parameters θ_i are influenced only by the data points within the fuzzy set Z_i that defines the region of validity of the i th local model. Again, solving this problem is a convex quadratic program.

D. Example

The purpose of this example is to illustrate the possible pitfalls of these basic algorithms; see also [1], [13], [16], [21], and [24]. The same example will be used in later sections to illustrate the properties of the suggested algorithms and tools. Consider the simple static nonlinear system

$$y(t) = u^3(t) \quad (13)$$

where $u(t) \in [-1, 1]$. Consider a T-S fuzzy model with affine local models of the form

$$y(t) = b_{0,i}u(t) + d_i \quad (14)$$

where $b_{0,i}$ and d_i are the local model parameters. The T-S fuzzy model consists of 5 local models, weighted according to the smooth weighting functions w_i (defined by Gaussian membership functions for the fuzzy sets Z_i) shown in Fig. 1.

For this system we generate a data set consisting of 201 equally spaced inputs in the interval $u(t) \in [-1, 1]$ and the corresponding noise-free outputs defined by (13). The global and local identification criteria give results as shown in Figs. 2 and 3. We observe that the constituent local models of the fuzzy model identified with the global criterion are far from being local linearizations of the nonlinear system (13).

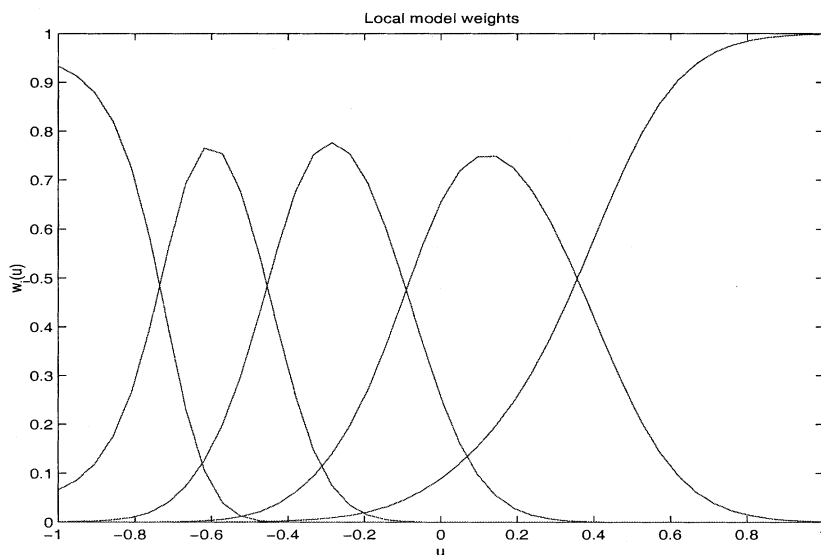


Fig. 1. Smooth weighting functions w_i of T-S fuzzy model.

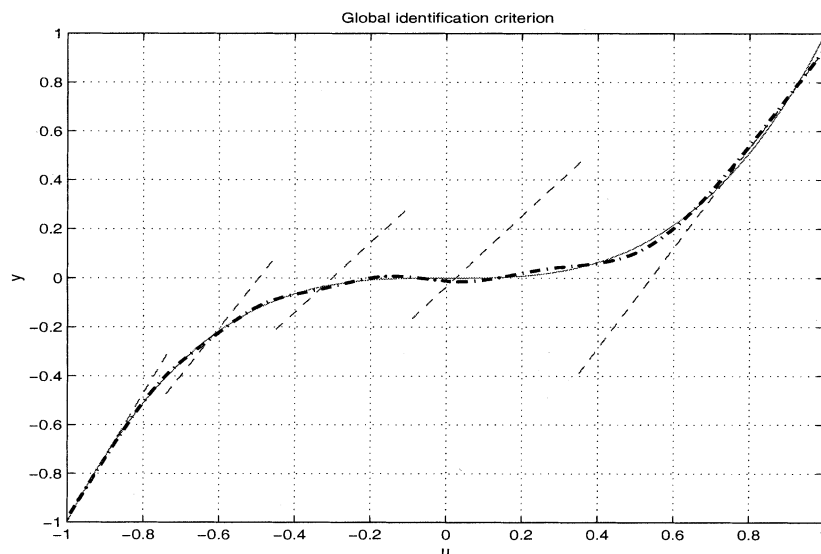


Fig. 2. Global identification results. Solid curve is the true system, dashed-dotted curve is the global model, while the dashed lines are the local models.

On the other hand, the globally identified model achieves a root-mean-squared prediction error of 0.013 659 while the global model identified with local criteria achieves a root-mean-squared prediction error of 0.054 824, which is considerably larger.

III. MULTIOBJECTIVE IDENTIFICATION ALGORITHM I

A. Algorithm

It was suggested in [24] to minimize the weighted sum of the global and local identification criteria (10) and (12). Here, we apply a slight extension, including the constraints (11) and individual weighting parameters for each of the individual local models. The algorithm solves the optimization problem

$$\min_{\theta} \left(V(\theta) + \sum_{i=1}^N \beta_i V_i(\theta_i) \right) \quad (15)$$

subject to (11). The solution can be computed by solving a convex quadratic program, which in the unconstrained case reduces to simple matrix manipulations. The weighting parameters $\beta_i \geq 0$ parameterize the set of Pareto-optimal solutions of the underlying multiobjective optimization problem, and essentially determine the tradeoff between the possibly conflicting objectives of global model accuracy and local model interpretability.

B. Conflict Analysis

The selection $\beta_i = 1$ in the multiobjective criterion (15) will in general give a fairly balanced tradeoff, due to the use of w_i (which forms a partition of unity) for weighting in (12). It is still of interest to study in detail how the choice of β influences the tradeoff. This problem is not discussed in [24]. In particular, it is of interest to analyze the degree of conflict between the different objectives in (15) for a given data sequence and different values of β . This will provide the user with information that can be

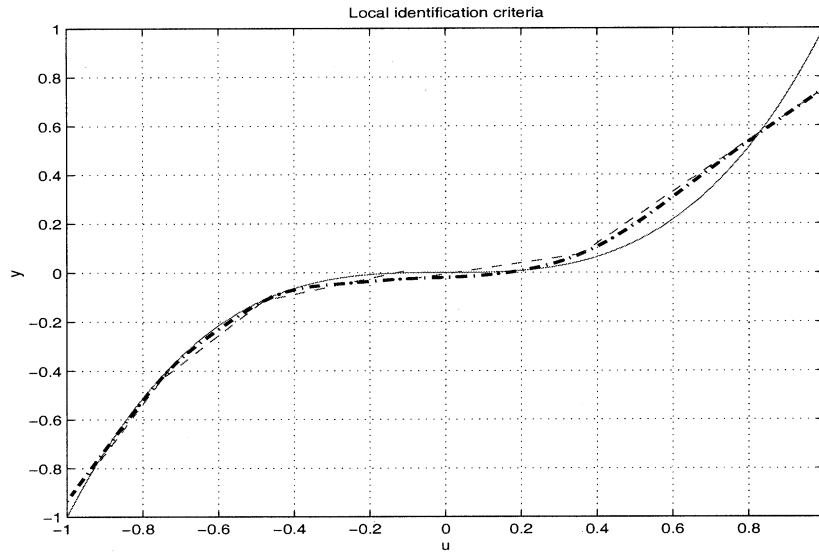


Fig. 3. Local identification results. Solid curve is the true system, dashed-dotted curve is the global model, while the dashed lines are the local models.

used to validate the model and data as well as improving the model, not only by selecting a better value of β but possibly also by modifying the model structure or membership functions, or adding/removing constraints.

In this section, let the minimum of (15) be denoted $\hat{\theta}(\beta)$ for a given data sequence and a vector of weights β . The minimum of (15) satisfies the Karush–Kuhn–Tucker (KKT) conditions

$$0 = \frac{\partial V}{\partial \theta}(\hat{\theta}(\beta)) + \sum_{i=1}^N \beta_i \frac{\partial V_i}{\partial \theta}(\hat{\theta}(\beta)) + \sum_{i=1}^N \hat{\mu}_i^T(\beta) \frac{\partial F_i}{\partial \theta}(\hat{\theta}(\beta)) + \sum_{i=1}^N \hat{\lambda}_i^T(\beta) \frac{\partial H_i}{\partial \theta}(\hat{\theta}(\beta)) \quad (16)$$

where $\hat{\mu}_i(\beta) \geq 0$ and $\hat{\lambda}_i(\beta)$ are the Lagrange multipliers (vectors) associated with the estimate $\hat{\theta}_i(\beta)$ of the local model parameter vector. If there are no conflicts among the objectives and constraints, i.e., $\hat{\theta}(\beta)$ minimizes all of the individual objectives simultaneously and none of the inequality constraints are active, then each of the terms in (16) will be zero. If there are conflicts, on the other hand, the directions and lengths of each of the (vector) terms of (16) will indicate the degree of conflict and which constraints and objectives are actually in conflict with each other.

For the unconstrained case, considering the parameter vector θ_i of the i th local models, (16) leads to

$$\frac{\partial V}{\partial \theta_i}(\hat{\theta}(\beta)) + \beta_i \frac{\partial V_i}{\partial \theta_i}(\hat{\theta}(\beta)) = 0. \quad (17)$$

Next, define the following sensitivity measures associated with each parameter $\theta_{i,j}$, which is the j th parameter in the i th local model:

$$\pi_{i,j}^g(\beta) = -\frac{\partial V}{\partial \theta_{i,j}}(\hat{\theta}(\beta)) \quad (18)$$

$$\pi_{i,j}^l(\beta) = -\frac{\partial V_i}{\partial \theta_{i,j}}(\hat{\theta}(\beta)) = \frac{1}{\beta_i} \frac{\partial V}{\partial \theta_{i,j}}(\hat{\theta}(\beta)) \quad (19)$$

which can easily be computed from the data/regressor matrices. The quantity $\pi_{i,j}^g(\beta)$ can be interpreted as the small decrease in the global identification criterion V that can be achieved by a small increase in $\hat{\theta}_{i,j}(\beta)$. Likewise, the quantity $\pi_{i,j}^l(\beta)$ can be interpreted as the small decrease in the local identification criterion V_i that can be achieved by a small increase in $\hat{\theta}_{i,j}(\beta)$. Hence, large values of $\pi_{i,j}^g(\beta)$ and $\pi_{i,j}^l(\beta)$ indicate conflicts between global and local performance. Notice that due to (17)

$$\pi_{i,j}^g(\beta) + \beta_i \pi_{i,j}^l(\beta) = 0. \quad (20)$$

An analysis of $\pi_{i,j}^g(\beta)$ and $\pi_{i,j}^l(\beta)$ can provide the user with significant information about the model and data, as illustrated in the example that follows.

C. Example (Continued)

Consider the system and data introduced in Section II-D. Assume $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta^*$ since there is no indication to discriminate between the different local models. Models were identified with the multiobjective identification algorithm for the four different cases $\beta^* \in \{0.01, 0.1, 1, 10\}$. The four models and their associated sensitivity measures $\pi_{i,j}^g(\beta)$ and $\pi_{i,j}^l(\beta)$ are illustrated in Fig. 4, and the values of the cost function $V(\hat{\theta}(\beta))$ are tabulated in Table I.

From the sensitivity measures and the cost function values for these four cases we can make the following observation (without any knowledge of the true system or the curves in the left column of Fig. 4): As β^* decreases, the global performance is improved at the cost of reduced local interpretability in the regions Z_3, Z_4 and Z_5 , i.e., the regions of validity of the local models with indexes 3–5. The reasoning behind this observation is as follows: It can be seen from Table I that the global performance indeed increases as β^* decreases. From Fig. 4 we observe at low values of β^* that the sensitivities $\pi_{i,j}^l(\beta)$ and $\pi_{i,j}^g(\beta)$ are significant only for $i \in \{3, 4, 5\}$. In other words, there is much stronger conflict between local and global performance in the regions Z_3, Z_4, Z_5

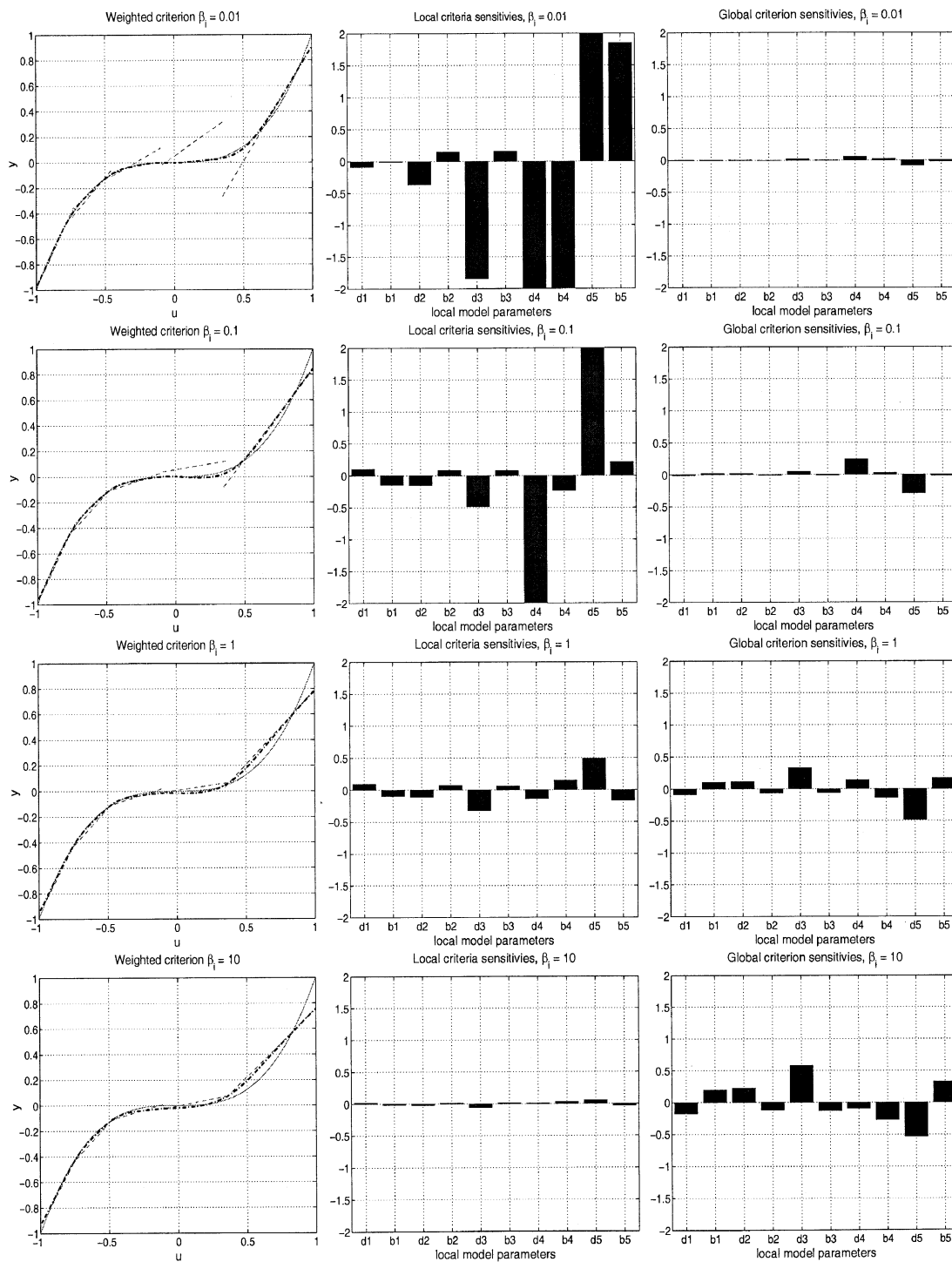


Fig. 4. Results with the multiobjective identification algorithm I, with $\beta^* = 0.01$ (first row), $\beta^* = 0.1$ (second row), $\beta^* = 1$ (third row), and $\beta^* = 10$ (fourth row). In the first column, the solid curve is the true system, dashed-dotted is the global model, and dashed are the local models. The bar-charts illustrate the local and global sensitivity measures associated with each model parameter. For example, b2 refers to the model parameter $b_{0,2}$ and d2 refers to the model parameter d_2 .

TABLE I
GLOBAL MODEL PERFORMANCE WITH THE MULTIOBJECTIVE IDENTIFICATION ALGORITHM I. NOTE THAT $\beta^* = 0$ AND $\beta^* \rightarrow \infty$ CORRESPONDS TO THE GLOBAL AND LOCAL IDENTIFICATION ALGORITHMS, RESPECTIVELY

	$\beta^* = 0$	$\beta^* = 0.01$	$\beta^* = 0.1$	$\beta^* = 1$	$\beta^* = 10$	$\beta^* \rightarrow \infty$
$V(\hat{\theta}(\beta))$	0.013659	0.017799	0.031397	0.045733	0.053372	0.054824

than in the regions Z_1, Z_2 . Indeed, this can be verified by examining the local and global models in the first column of Fig. 4,

where the local model with index 1 is valid for small u and the local model with index 5 is valid for large u , see also Fig. 1.

The observation that the conflicts between local and global performance are related to the local models with indexes 3–5 can be used in several ways to improve the model. For example, in order to improve the global performance without any decrease in local performance the membership functions might be tuned toward a finer partition of the area $Z_3 \cup Z_4 \cup Z_5$ and a correspondingly coarser partition in the area $Z_1 \cup Z_2$, or more local models should be introduced in this area. Thus, the conflict analysis points at specific regions and local models that need further consideration. Note that the selected four values of β^* are nothing more than first guesses that are not expected to lead to an optimal choice, and therefore need further refinement

IV. MULTIOBJECTIVE IDENTIFICATION ALGORITHM II

A. Algorithm

The idea is to first compute the locally weighted estimates $\tilde{\theta}_i$ as described in Section II-C. In the second step of the algorithm the global least squares prediction error is minimized subject to small deviations of the relevant parameters from the locally identified parameters, i.e.,

$$\min_{\theta} V(\theta) \quad (21)$$

subject to (11) and

$$\tilde{\theta}_i - \Delta\theta_i \leq \theta_i \leq \tilde{\theta}_i + \Delta\theta_i \quad (22)$$

for $i = 1, 2, \dots, N$. The allowed deviation $\Delta\theta_i$ from the locally identified parameters $\tilde{\theta}_i$ might be specified by the user and/or might be generated from some uncertainty estimate (e.g., standard deviation) of the locally weighted estimate $\tilde{\theta}_i$. Hence, it is guaranteed that the local model parameters retain their interpretability as local linearizations (within a user-specified tolerance) when they are tuned for global model performance in the second step of the algorithm. Notice that the interpretability of the local model parameters as linearizations may or may not involve the offset parameters d_i . Hence, depending on the application, the offset parameters d_i may in some cases be allowed to vary freely in the global optimization (21). As before, the optimization problems involved are convex quadratic programs, and the estimate $\hat{\theta}$ can be computed efficiently.

B. Conflict Analysis

Also, with this algorithm, a study of the KKT conditions contains information on conflicts between the objectives. In particular, it is useful to study the Lagrange multipliers associated with the deviation constraints (22). A zero Lagrange multiplier associated with some constraint $\theta_{i,j} \leq \tilde{\theta}_{i,j} + \Delta\theta_{i,j}$ (or $\tilde{\theta}_{i,j} - \Delta\theta_{i,j} \leq \theta_{i,j}$) means that the global prediction performance cannot be improved by increasing the allowed deviation $\Delta\theta_{i,j}$ alone since the Lagrange multipliers have the following sensitivity interpretation [15]:

$$\hat{\lambda}_{i,j} = -\frac{\partial V}{\partial \Delta\theta_{i,j}}(\hat{\theta}). \quad (23)$$

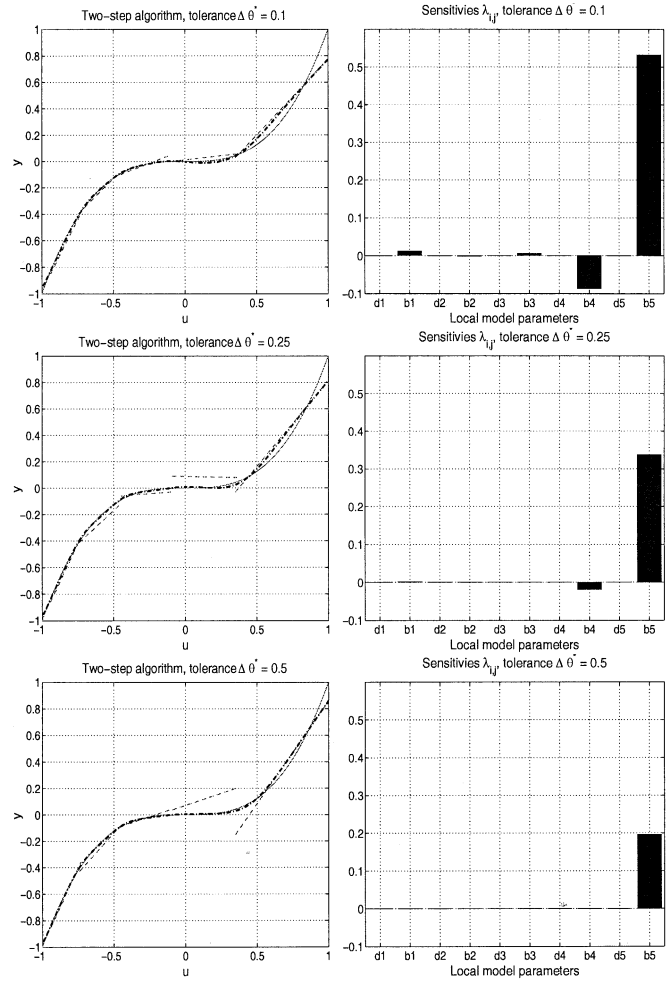


Fig. 5. Results with the multiobjective identification algorithm II, with $\Delta\theta^* = 0.1$ (first row), $\Delta\theta^* = 0.25$ (second row), $\Delta\theta^* = 0.5$ (third row). In the first column, the solid curve is the true system, dashed-dotted is the global model, and dashed are the local models. The bar-charts illustrate the sensitivity $\lambda_{i,j}$ associated with each model parameter.

On the other hand, the nonnegative value of the Lagrange multiplier $\hat{\lambda}_{i,j}$ tells us how much the global prediction error criterion $V(\theta)$ might be reduced by a small increase in the allowed deviation $\Delta\theta_{i,j}$. Thus, it is straightforward to determine those local model parameters that contribute to the conflict between objectives; they have nonzero Lagrange multipliers for some associated constraint. This information can be applied in several ways either to reformulate the model (for example by refining the model structure or membership functions) or reconsider the allowed deviation.

C. Example (Continued)

Consider the system and data introduced in Sections II-D. Assume for simplicity $\Delta\theta_{i,j} = \Delta\theta^*$ for all i, j . Models were identified with the multiobjective identification algorithm for the three different cases $\Delta\theta^* \in \{0.1, 0.25, 0.5\}$. The three models and their associated sensitivities $\lambda_{i,j}$ are shown in Fig. 5. We use the convention that

$$\lambda_{i,j} = \begin{cases} \hat{\lambda}_{i,j}, & \text{if the upper bound is active} \\ -\hat{\lambda}_{i,j}, & \text{if the lower bound is active} \\ 0, & \text{if no bound is active.} \end{cases} \quad (24)$$

TABLE II

GLOBAL MODEL PERFORMANCE WITH THE MULTIOBJECTIVE IDENTIFICATION ALGORITHM II. NOTE THAT $\Delta\theta^* = \infty$ AND $\Delta\theta^* = 0$ CORRESPONDS TO THE GLOBAL AND LOCAL IDENTIFICATION ALGORITHMS, RESPECTIVELY

	$\Delta\theta^* = \infty$	$\Delta\theta^* = 0.5$	$\Delta\theta^* = 0.25$	$\Delta\theta^* = 0.1$	$\Delta\theta^* = 0$
$V(\hat{\theta}(\Delta\theta))$	0.013659	0.025819	0.035974	0.044985	0.054824

Hence, $\lambda_{i,j} > 0$ indicate that the global performance can be improved by relaxing the upper bound, while $\lambda_{i,j} < 0$ indicate that the global performance can be improved by relaxing the lower bound. The values of the cost function $V(\hat{\theta}(\Delta\theta))$ are tabulated in Table II.

From the sensitivity measures and cost function values for these four cases we can make the following observation (without any knowledge of the true system or the leftmost curves in Fig. 5): *As the allowed tolerance $\Delta\theta^*$ on deviation from the locally identified parameters is increased, the global prediction performance improves and for sufficiently large $\Delta\theta^*$ the improvement in global prediction performance is limited by the constraints on the parameter $b_{0,5}$. Again, the conflict analysis points at a certain parameter in a certain local model, indicating that higher tolerance on deviation from locally interpretable parameter value or tuning of the membership functions associated with Z_5 is required. This is not unexpected, since Z_5 is the largest fuzzy set. Note that the three values of $\Delta\theta^*$ were selected by trial and error to give a large variety of solutions, and further refinement might be needed to get the best final model*

V. LIMITING SENSITIVITY MEASURES

The sensitivity measures associated with the two different multiojective identification algorithms have exactly the same interpretation in a limiting case.

Theorem 1: If there are no constraints (11), the sensitivity measures (18) and (23) satisfy

$$\lim_{\beta \rightarrow \infty} \pi_{i,j}^g(\hat{\theta}(\beta)) = \lim_{\Delta\theta \rightarrow 0} \hat{\lambda}_{i,j}. \quad (25)$$

Proof: In the limit $\beta \rightarrow \infty$ (i.e., all elements of the vector β go to infinity)

$$\hat{\theta}(\beta) \rightarrow \tilde{\theta} \quad (26)$$

where $\tilde{\theta}$ is the minimum of (12). It follows from (18) that

$$\lim_{\beta \rightarrow \infty} \pi_{i,j}^g(\hat{\theta}(\beta)) = -\frac{\partial V}{\partial \theta_{i,j}}(\tilde{\theta}). \quad (27)$$

From (23), there exists a diagonal matrix \mathcal{E} whose elements satisfy $|\varepsilon_{k,l}| \leq 1$ such that

$$\lim_{\Delta\theta \rightarrow 0} \hat{\lambda}_{i,j} = \lim_{\Delta\theta \rightarrow 0} \left(-\frac{\partial V}{\partial \Delta\theta_{i,j}}(\tilde{\theta} + \mathcal{E}\Delta\theta) \right). \quad (28)$$

Hence

$$\begin{aligned} & \lim_{\Delta\theta \rightarrow 0} \hat{\lambda}_{i,j} \\ &= -\lim_{\Delta\theta \rightarrow 0} \lim_{\delta \rightarrow 0} \frac{V(\tilde{\theta} + \mathcal{E}(\Delta\theta + \delta\mathcal{I}_{i,j})) - V(\tilde{\theta} + \mathcal{E}\Delta\theta)}{\delta} \end{aligned} \quad (29)$$

where $\mathcal{I}_{i,j}$ is a unit-vector of same dimension as θ , and all elements are zero, except the element corresponding to $\theta_{i,j}$ which is one. Consider the case $\hat{\lambda}_{i,j} > 0$ and interchange the limits (notice that V is a continuous function of θ):

$$\begin{aligned} \lim_{\Delta\theta \rightarrow 0} \hat{\lambda}_{i,j} &= -\lim_{\delta \rightarrow 0} \lim_{\Delta\theta \rightarrow 0} \\ & \frac{V(\tilde{\theta} + \mathcal{E}(\Delta\theta + \delta\mathcal{I}_{i,j})) - V(\tilde{\theta} + \mathcal{E}\Delta\theta)}{\delta} \end{aligned} \quad (30)$$

$$= -\lim_{\delta \rightarrow 0} \frac{V(\tilde{\theta} + \delta\mathcal{E}\mathcal{I}_{i,j}) - V(\tilde{\theta})}{\delta} \quad (31)$$

$$= -\lim_{\delta \rightarrow 0} \frac{V(\tilde{\theta} + \delta\varepsilon_{i,j}) - V(\tilde{\theta})}{\delta}. \quad (32)$$

Since $\hat{\lambda}_{i,j} > 0$ it is clear that $\varepsilon_{i,j} = 1$ and

$$\lim_{\Delta\theta \rightarrow 0} \hat{\lambda}_{i,j} = -\frac{\partial V}{\partial \theta_{i,j}}(\tilde{\theta}). \quad (33)$$

Equation (25) follows from (27) and (33). The cases $\hat{\lambda}_{i,j} < 0$ and $\hat{\lambda}_{i,j} = 0$ are similar. \square

This result shows that the sensitivity measures associated with the two algorithms have a common interpretation in the limiting case. Consequently, the sensitivities might be compared directly without any scaling or normalization.

VI. EXPERIMENTAL RESULTS: ESTIMATION OF LUNGS RESPIRATION DYNAMICS

In order to further illustrate the suggested methods, we consider the problem of estimation of respiration dynamics parameters described in [4]. This is important for monitoring respiratory mechanics in patients on ventilatory support, for example to assess patients' pulmonary conditions (for which the interpretability of the local model parameters is of vital importance) and to automatically control or optimize the ventilator settings (for which the global model accuracy may be more important).

A. Model Structure

Consider the following dynamic relationship [4], [14], [19]:

$$P = EV + R\frac{dV}{dt} + P_0 \quad (34)$$

where $V(l)$ is the lungs volume, $\dot{V}(l/s)$ is the flow rate through the ventilation tube, and $P(hPa)$ is the pressure. We consider identification of the three parameters of this equation, namely the respiratory elastance $E(hPa/l)$, the resistance $R((hPa \cdot s)/l)$ and the elastic recoil pressure $P_0(hPa)$. This problem is viewed as a regression problem where T-S fuzzy models consisting of multiple linear models of the form

$$P = E_i V + R_i \frac{dV}{dt} + P_{0,i} \quad (35)$$

are identified. In other words, the models predict P as a function of V and \dot{V} .

Consider three different T-S fuzzy model structures, each with four local models of the form (35), and Gaussian membership functions as shown in Fig. 6.

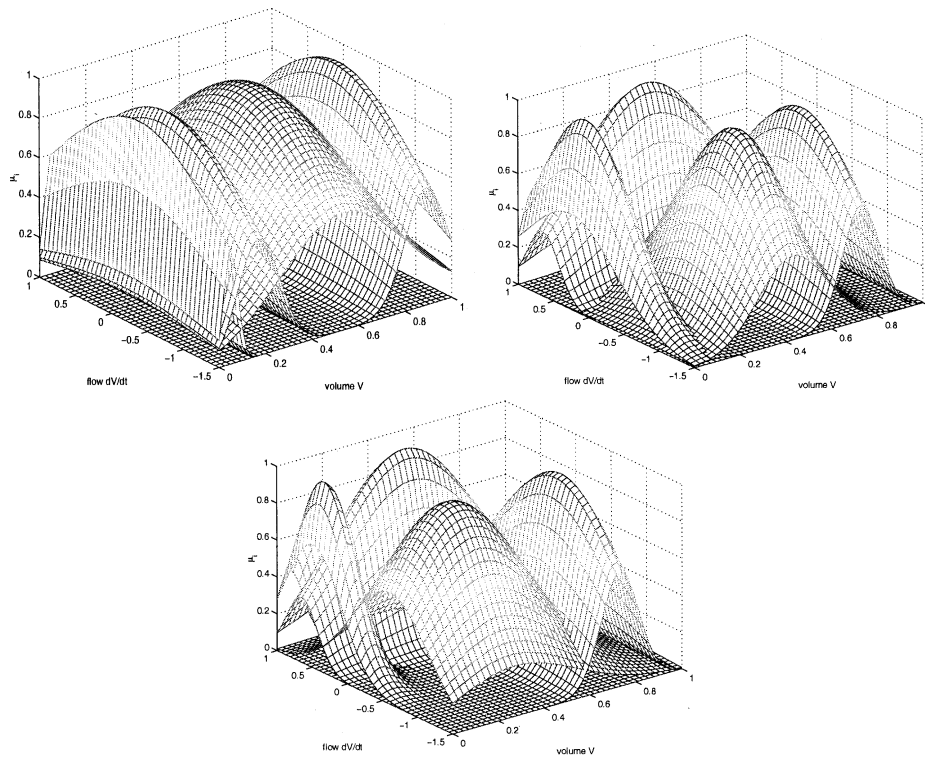


Fig. 6. Membership functions μ_i for the T-S lungs models: Model structure A (top left), B (top right), and C (bottom).

- **Model Structure A:** A fuzzy model where the membership functions are identified using the algorithm described in [12] (this algorithm is similar to [22]).
- **Model Structure B:** A fuzzy model where the membership function are selected “manually” to be consistent with the experiments using fuzzy clustering reported in [4].
- **Model Structure C:** Similar to Model Structure B, with with somewhat different membership functions.

In the continuation we assume the membership function of these three model structures are fixed, and consider identification of the consequent parameters.

B. Data and Identification

A single cycle from a typical data sequence is shown in Fig. 7. The respiration cycle consists of three phases; i) inspiration phase (forced inlet flow), ii) respiration pause (no flow), and iii) expiration phase (free outlet flow). Fig. 8 illustrates the fuzzy partition of the three model structures under consideration, with the identification data sequence projected onto the plane described by (V, \dot{V}) . The identification data sequence consists of measurements from ten respiration cycles for a single patient. Notice that patients may have different respiration dynamics, so it is generally desirable to identify a model for each patient [4].

C. Results

For each of the model structures A, B, and C we identified the local model parameters using i) locally weighted least squares, ii) least squares, iii) multiobjective algorithm I with $\beta^* = 1$, and iv) multiobjective algorithm II with $\Delta\theta^* = 1$. The results

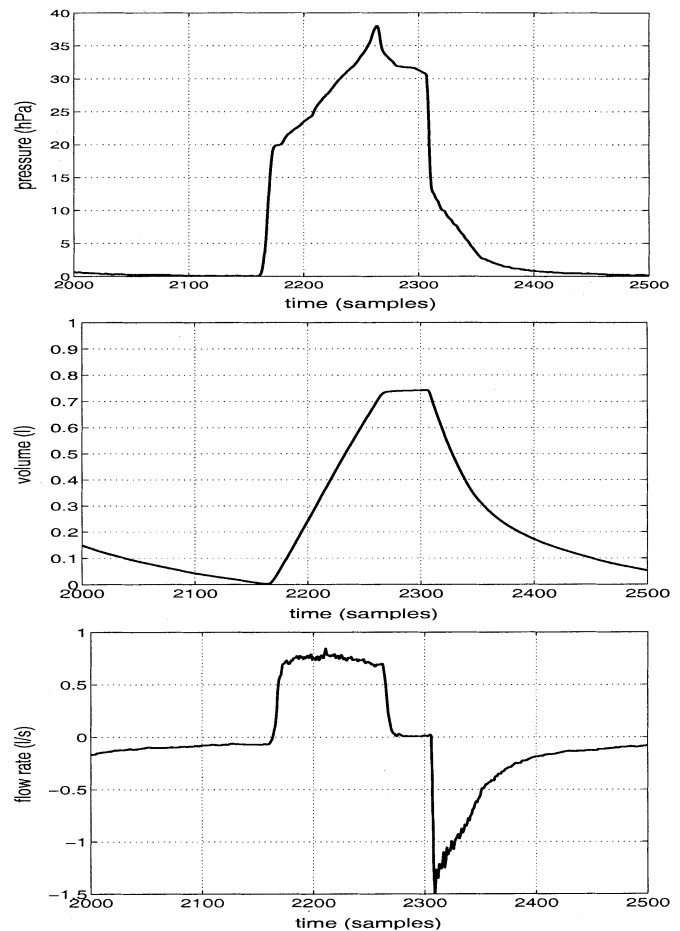


Fig. 7. Subset of a typical data sequence with respiratory data.

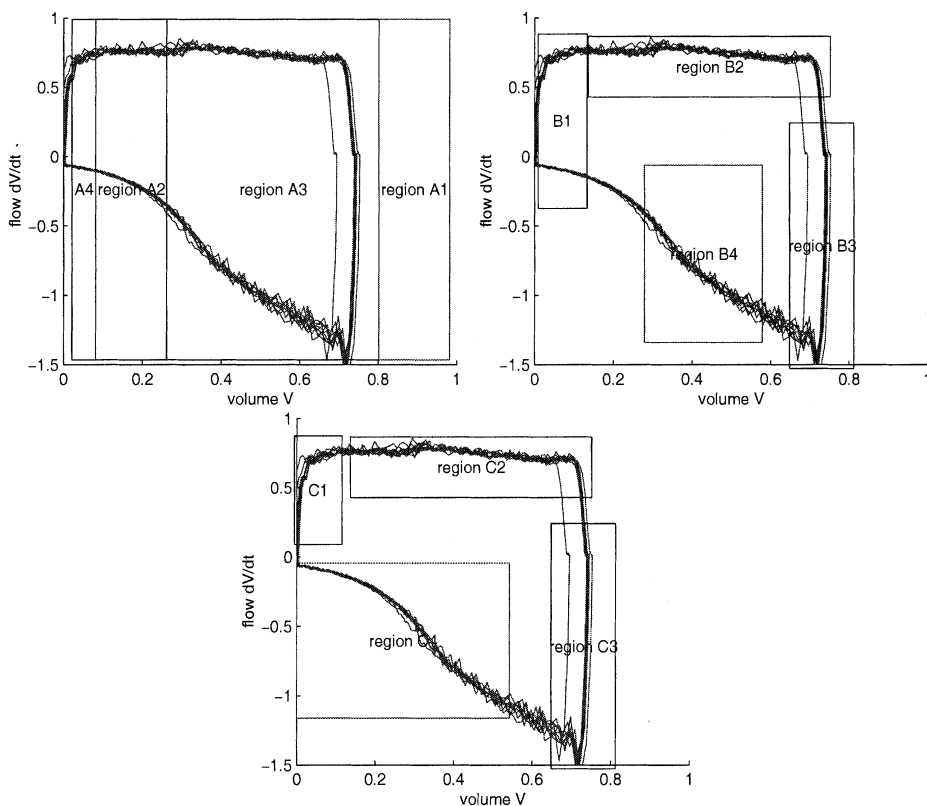


Fig. 8. (Top left) Partitioning of input domain with projected data for model structure A. (Top right) Partitioning of input domain with projected data for model structure B. (Bottom) Partitioning of input domain with projected data for model structure C. The rectangles provide a simplified illustration of the fuzzy sets.

TABLE III
LOCAL PARAMETERS OF MODEL STRUCTURE A, USING DIFFERENT IDENTIFICATION METHODS. ALGORITHM I IS WITH $\beta^* = 1$ AND ALGORITHM II IS WITH $\Delta\theta^* = 1$

Region	Parameter	LWLS	LS	Alg. I	Alg. II
Region A1	E	38.93	-696.5	33.93	39.88
	R	9.329	-12.84	6.469	8.329
	P_0	3.038	565.5	7.092	4.038
Region A2	E	28.36	-1.364	23.57	27.36
	R	20.93	26.67	22.80	21.93
	P_0	0.9699	1.429	1.371	-0.031
Region A3	E	42.47	21.20	42.14	41.47
	R	13.47	12.66	13.05	12.63
	P_0	-0.1901	10.42	0.075	0.665
Region A4	E	8.346	25.53	-2.554	7.346
	R	23.21	24.33	24.39	24.21
	P_0	1.822	-1.38	2.265	2.389

TABLE IV
LOCAL PARAMETERS OF MODEL STRUCTURE B, USING DIFFERENT IDENTIFICATION METHODS. ALGORITHM I IS WITH $\beta^* = 1$ AND ALGORITHM II IS WITH $\Delta\theta^* = 1$

Region	Parameter	LWLS	LS	Alg. I	Alg. II
Region B1	E	12.39	8.578	10.71	11.39
	R	22.89	23.18	23.35	23.89
	P_0	2.080	2.273	2.241	2.318
Region B2	E	29.86	30.18	29.76	30.48
	R	11.89	6.202	10.83	10.89
	P_0	7.325	12.63	8.100	7.852
Region B3	E	53.69	60.06	54.74	54.69
	R	10.41	9.989	10.12	9.704
	P_0	-7.809	-10.95	-8.579	-8.582
Region B4	E	50.85	19.10	51.31	51.36
	R	12.83	-1.049	12.42	11.83
	P_0	-5.029	-2.620	-5.485	-6.029

are given in Tables III–V. The average root-mean-square residuals for these cases are given in Table VI, and the sensitivity measures associated with the multiobjective identification algorithms are shown in Fig. 9.

D. Discussion of the Results

There are large differences between the locally and globally identified local model parameters, especially with model structure A and C. In many cases, the globally identified local model

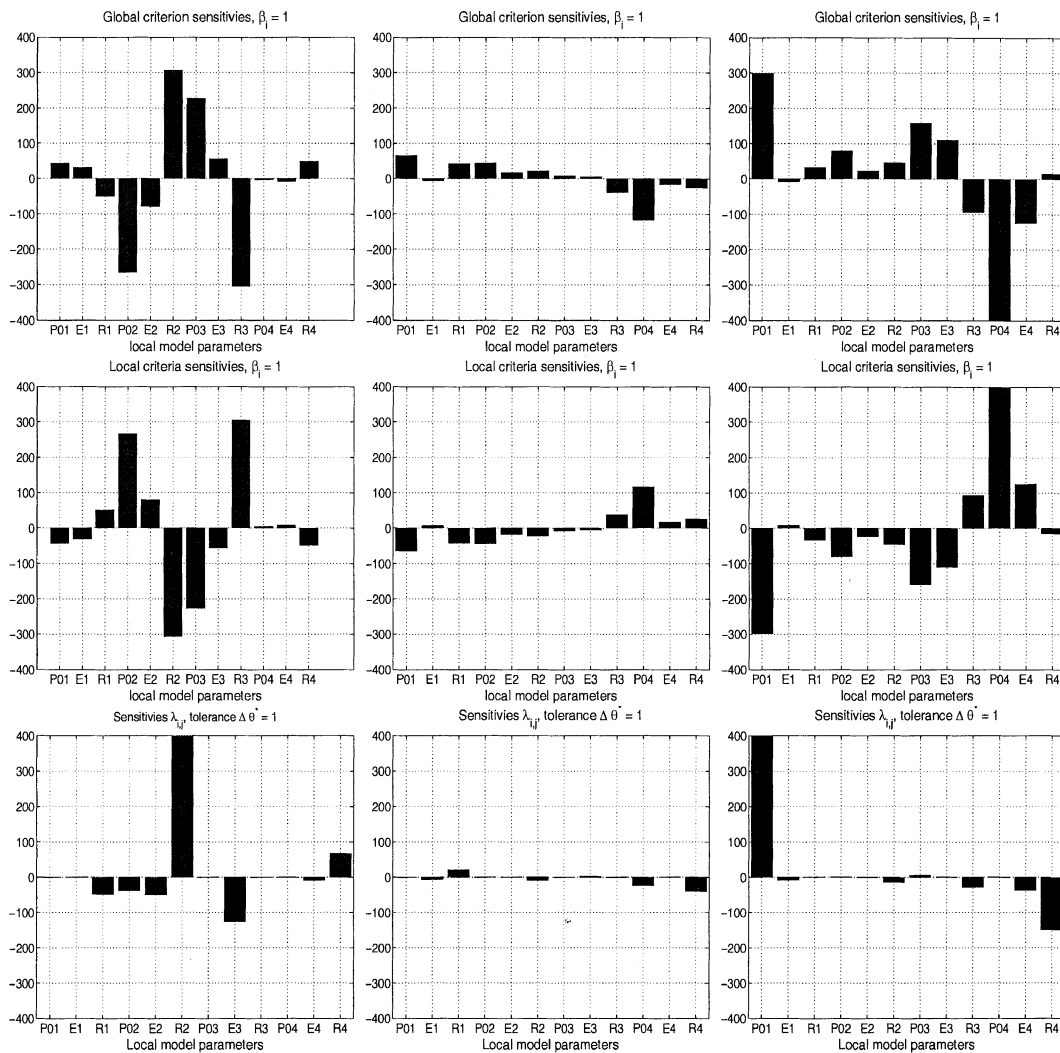


Fig. 9. Left column: Sensitivities with model structure A. Middle column: Sensitivities with model structure B. Right column: Sensitivities with model structure C.

TABLE V
LOCAL PARAMETERS OF MODEL STRUCTURE C, USING DIFFERENT IDENTIFICATION METHODS. ALGORITHM I IS WITH $\beta^* = 1$ AND ALGORITHM II IS WITH $\Delta\theta^* = 1$

Region	Parameter	LWLS	LS	Alg. I	Alg. II
	E	8.669	-19.49	-6.867	7.669
Region C1	R	23.15	28.94	23.75	23.50
	P_0	2.116	-0.381	3.295	3.116
	E	29.68	21.18	29.21	29.40
Region C2	R	11.30	4.342	9.673	10.30
	P_0	7.736	18.49	9.214	8.736
	E	54.37	51.76	53.26	53.92
Region C3	R	10.47	14.08	10.03	9.468
	P_0	-8.296	1.486	-7.229	-7.296
	E	42.07	-28.53	40.68	41.07
Region C4	R	13.63	-18.51	12.64	12.63
	P_0	-1.300	3.004	-1.547	-2.186

TABLE VI
ROOT-MEAN-SQUARE RESIDUALS

Model	LWLS	LS	Alg. I	Alg. II
Model structure A	1.9130	1.1067	1.6854	1.7368
Model structure B	1.4631	1.3758	1.4354	1.4209
Model structure C	1.9977	1.3859	1.8582	1.8565

parameter estimates of E and R are in conflict with their physical interpretation (e.g., when they are negative). Both multi-objective identification algorithms can be used to address this tradeoff. The selection of the tuning parameters (β and $\Delta\theta$) should take into account the actual application of the model. A discussion of this is beyond the scope of the present paper, so we just make the observation that these parameters offers useful degrees of freedom to address conflicts and tradeoffs.

When comparing the sensitivities of the conflict analysis for model structures A and B we observe that the sensitivities with model structure A are up to one order of magnitude larger than the sensitivities with model structure B. This indicates that the use of model structure A leads to a significant conflict

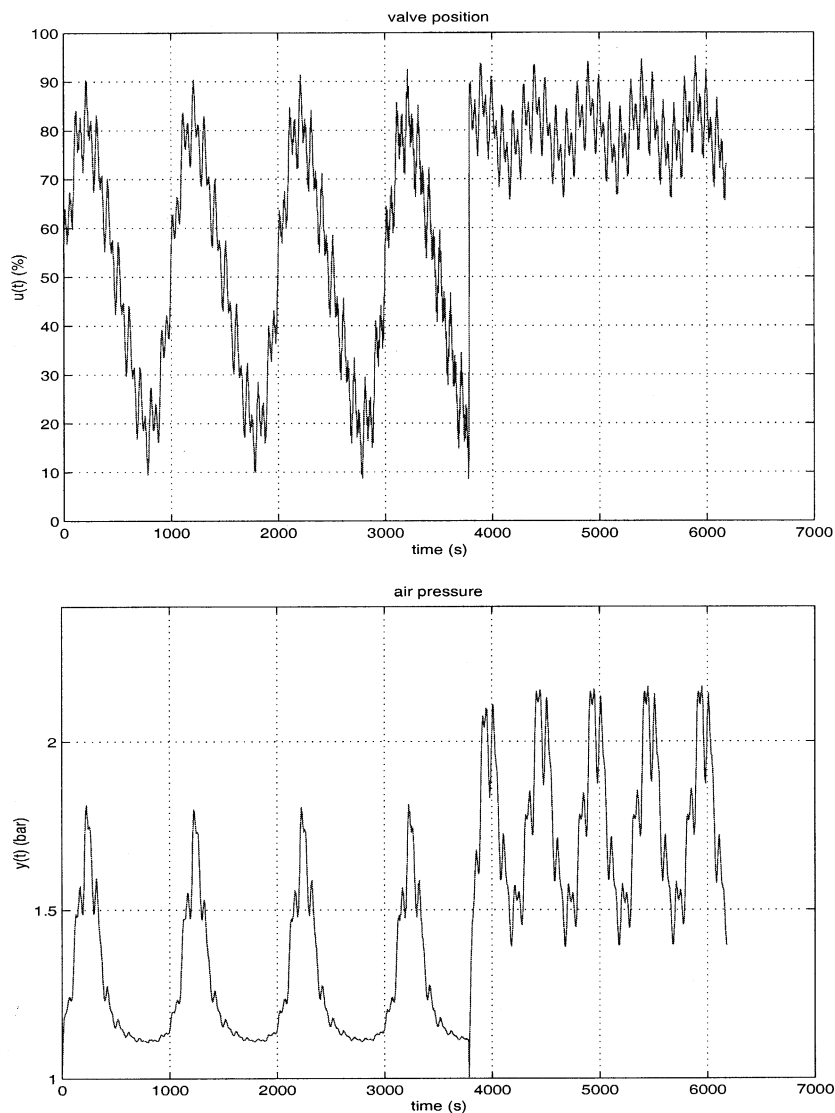


Fig. 10. Fermenter pressure data used for identification.

between local model interpretability and global prediction accuracy, compared to model structure B. Certainly, model structure A leads to a smaller global prediction error than model structure B when the models are identified using global least squares. However, model structure B admits a more intuitive and appealing interpretation of its membership functions as the partitioning of the input domain is closely related to the different phases of the respiration cycle; region B1 corresponds to the first part of the inspiration phase, region B2 corresponds to the final part of the inspiration phase, region B3 includes the respiration pause, while region B4 contains the expiration phase. This is to be expected, since model structure B was derived using a combination of experimental data and prior knowledge about the application, while model structure A is the result of a structural identification algorithm.

Perhaps more interestingly, we observe that the sensitivities with model structure C are also up to one order of magnitude larger than the sensitivities with model structure B. This large difference points out a major conflict in model structure C, especially since Figs. 6 and 8 show that the structure (member-

ship functions) are only slightly different. The only difference is that the fuzzy sets B1 and B4 (in model structure B) are slightly shifted and resized versions of C1 and C4 (in model structure C). The other two fuzzy sets are exactly the same in both model structures. Hence, the conflict analysis for model structure C points out a structural problem, namely that there will be a conflict between global prediction performance and local interpretability due to an unfortunate interaction between the membership functions for the fuzzy sets C1 and C4. As “proved” by model structure B, this conflict can be resolved by a small modification of these fuzzy sets.

As expected, a high level of sensitivity (or conflict) seems to be correlated with a high dependence of the residuals on which parameter identification criterion is being used; cf. Fig. 6. Especially for model structures A and C, there is a large difference between locally weighted least squares and the global least squares algorithms, while this difference is small for model structure B.

In summary, it is clear that model structure B is better than A and C, even though this is not evident from only inspecting

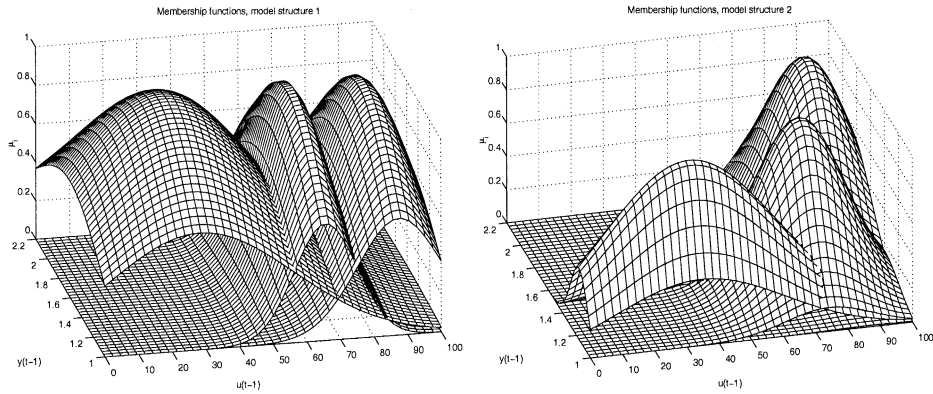


Fig. 11. Membership functions μ_i for the T-S pressure dynamics models. Model structure 1 (left) and 2 (right).

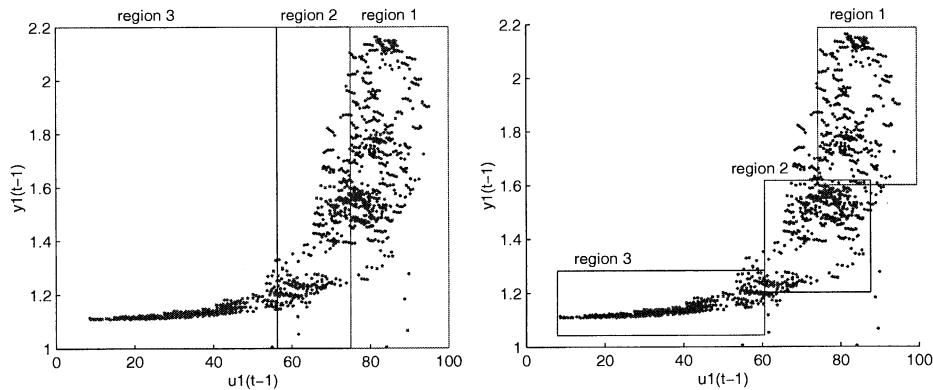


Fig. 12. Projected identification data and fuzzy regions (the rectangles provide a simplified representation of the fuzzy sets). Model structure 1 (left) and 2 (right).

the residuals in Table VI. The conflict analysis can be used as a constructive tool to assist modification of the model structure (including membership functions).

VII. EXPERIMENTAL RESULTS: FERMENTER PRESSURE DYNAMICS

This section describes identification of the highly nonlinear pressure dynamics in a laboratory fermenter [3]. Both global accuracy and local interpretability may be of interest if the model is to be used in a model-based controller.

The fermenter under consideration consists of a 40 l tank containing 25 l of water. At the bottom of the tank, air is fed into the water at a specified constant flow rate. The air pressure above the water level, $y(t)$ (bar), is controlled by an outlet valve $u(t)$ (% closed). Nonlinearities are both due to the valve characteristics and the air compression curve. The two concatenated data sequences used for identification are shown in Fig. 10.

We consider two model structures, each with three local models of the affine form

$$y(t) = -a_{1,i}y(t-1) + b_{1,i}u(t-1) + d_i. \quad (36)$$

- **Model structure 1.** A fuzzy model where the membership functions are identified using the algorithm described in [12].
- **Model structure 2:** A fuzzy model where the membership functions were originally found by clustering [3] and then modified to have diagonal covariance matrices.

The membership functions of the fuzzy sets defining these two T-S fuzzy models are shown in Fig. 11 and a projection of the data into the regions are shown in Fig. 12.

The local model parameters of both of these model structures were identified using i) locally weighted least squares, ii) global least squares, iii) the multiobjective algorithm I with $\beta^* = 0.1$, and iv) the multiobjective algorithm II with $\Delta\theta^* = 0.02$. The root mean squared residuals are shown in Table VII, and the sensitivity measures are shown in Fig. 13. From Table VIII, it is clear that model structure 1 with global least squares achieves the smallest average value of the residuals. However, the conflict analysis based on the sensitivity measures shows that this is achieved at the cost of lack of local interpretability, since model structure 2 leads to less conflicts than model structure 1. This can be verified by examining the actual parameter estimates where it can be seen that there are unreasonable variations between regions with model structure 1. Again, the multiobjective algorithms can be applied to address the tradeoffs. In particular, it can be observed from the sensitivity measures that the conflicts between global accuracy and local interpretability are mainly associated with the high-pressure regions, where it is known that the nonlinearities are strongest. This indicates that these regions should have their membership functions tuned for less interaction, or that more fuzzy rules should be introduced.

The conclusion that model structure 2 is in general better than model structure 1 is confirmed by validation on a separate data set consisting of 281 samples. The root mean squared prediction errors are tabulated in Table VIII. The application of the

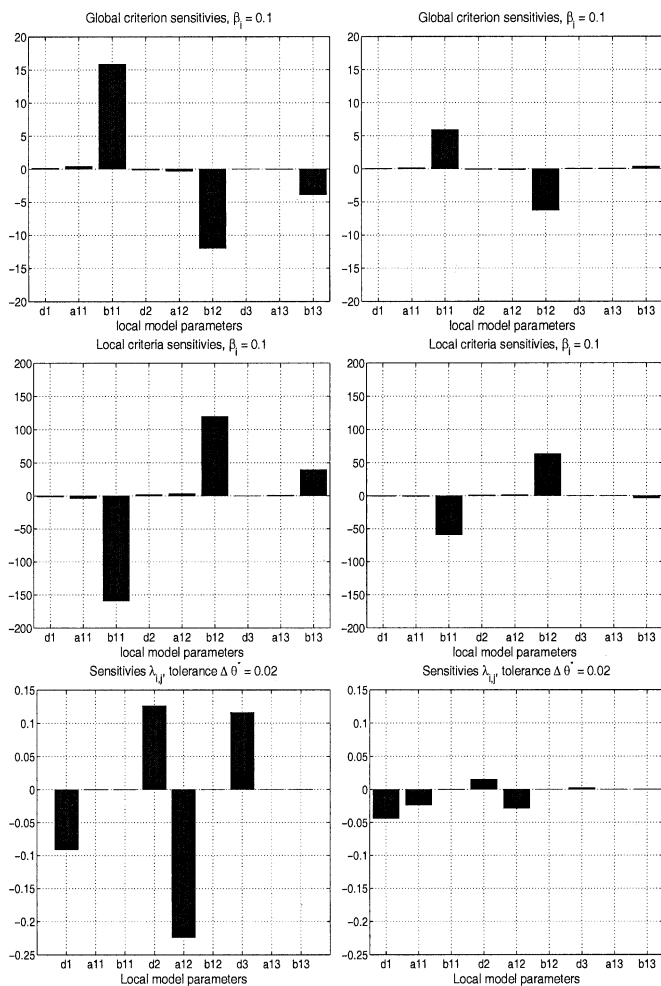


Fig. 13. Left column: Sensitivities with model structure 1. Right column: Sensitivities with model structure 2.

TABLE VII
ROOT-MEAN-SQUARE RESIDUALS (ONE-STEP-AHEAD PREDICTION USING THE SAME IDENTIFICATION DATA)

Model	LWLS	LS	Alg. I	Alg. II
Model structure 1	0.0106	0.0037	0.0049	0.0075
Model structure 2	0.0058	0.0042	0.0043	0.0045

TABLE VIII
ROOT-MEAN-SQUARE ONE-STEP-AHEAD PREDICTION ERROR USING A SEPARATE DATA SET

Model	LWLS	LS	Alg. I	Alg. II
Model structure 1	0.0175	0.0065	0.0132	0.0157
Model structure 2	0.0069	0.0051	0.0055	0.0062

proposed algorithms have correctly pointed out conflicts between interpretability and accuracy and suggested how to improve model structure 1.

VIII. CONCLUSION

The T-S model is a generic nonlinear model representation that has been studied extensively over the last years. One of the reason for its success might be that this is a model that admits a useful (local) interpretation. Hence, it is not surprising that the (not well studied) tradeoff between model interpretability and accuracy in system identification is particularly relevant in the identification of T-S fuzzy models.

A multiobjective optimization formulation of the identification problem arises naturally due to the two conflicting objectives. In this paper, we have studied two multiobjective formulations (one of them was proposed in [24]) and suggested algorithms for their solution and tools for analysis of the solution in terms of conflicts and sensitivity. As shown by the examples, this conflict/sensitivity analysis provides useful information not only about the local model parameter estimates, but also about the adequateness of the model structure and membership functions. Hence, we believe multiobjective optimization is a useful tool for T-S fuzzy identification. Further improvements can be made by combining the suggested methods with other tools for selecting the weighting coefficients, such as [5] where an interactive procedure for tuning the weighting coefficients is provided through a linguistic fuzzy rule-base.

Another application example is given in [25].

REFERENCES

- [1] J. Abonyi and R. Babuška, "Local and global identification and interpretation of parameters in Takagi-Sugeno fuzzy models," in *Proc. IEEE Conf. Fuzzy Systems*, San Antonio, TX, 2000, pp. 835–840.
- [2] J. Abonyi, R. Babuška, H. B. Verbruggen, and F. Szeifert, "Incorporating prior knowledge in fuzzy model identification," *Int. J. Syst. Sci.*, vol. 31, pp. 657–667, 2000.
- [3] R. Babuška, *Fuzzy Modeling for Control*. Norwell, MA: Kluwer, 1998.
- [4] R. Babuška, L. Alic, M. S. Lourens, A. F. M. Verbraak, and J. Bogaard, "Estimation of respiratory parameters via fuzzy clustering," *Art. Intell. Med.*, vol. 21, pp. 91–105, 2001.
- [5] L. F. B. Baptistella and A. Ollero, "Fuzzy methodologies for interactive multicriteria optimization," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-10, pp. 355–365, Apr. 1980.
- [6] M. Fischer and O. Nelles, "Predictive control based on local linear fuzzy models," *Int. J. Syst. Sci.*, vol. 29, pp. 679–697, 1998.
- [7] K. J. Hunt and T. A. Johansen, "Design and analysis of gain-scheduled local controller networks," *Int. J. Control*, vol. 66, pp. 619–651, 1997.
- [8] Y. Jin, "Fuzzy modeling of high-dimensional systems: Complexity reduction," *IEEE Trans. Fuzzy Syst.*, vol. 8, pp. 212–221, Apr. 2000.
- [9] T. A. Johansen, "Robust identification of Takagi-Sugeno-Kkang fuzzy models using regularization," in *Proc. IEEE Conf. Fuzzy Systems*, New Orleans, LA, 1996, pp. 180–186.
- [10] —, "Multi-objective identification of FIR models," in *Preprints IFAC Symp. System Identification*, Santa Barbara, CA, 2000, pp. 917–922.
- [11] T. A. Johansen and B. A. Foss, "Constructing NARMAX models using ARMAX models," *Int. J. Control*, vol. 58, pp. 1125–1153, 1993.
- [12] —, "Identification of nonlinear system structure and parameters using regime decomposition," *Automatica*, vol. 31, pp. 321–326, 1995.
- [13] T. A. Johansen, R. Shorten, and R. Murray-Smith, "On the interpretation and identification of Takagi-Sugeno fuzzy models," *IEEE Trans. Fuzzy Syst.*, vol. 8, pp. 297–313, Apr. 2000.
- [14] A. M. Lauzon and J. H. T. Bates, "Estimation of time-varying respiratory mechanical parameters by recursive least squares," *J. Appl. Physiol.*, vol. 71, pp. 1159–1165, 1991.
- [15] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1984.
- [16] R. Murray-Smith and T. A. Johansen, "Local learning in local model networks," in *Multiple Model Approaches to Modeling and Control*, R. Murray-Smith and T. A. Johansen, Eds. London, U.K.: Taylor and Francis, 1997.

- [17] R. Murray-Smith, T. A. Johansen, and R. Shorten, "On transient dynamics, off-equilibrium behavior and identification in blended multiple model structures," in *Proceedings of the European Control Conference*, Karlsruhe, Germany, 1999, pp. F1020–F1027.
- [18] D. Passaquay, M. Bross, R. Babuška, S. Boverie, and A. Titli, "Performance evaluation of a fuzzy rule base for control purpose," in *Proc. IEEE Conf. Fuzzy Syst.*, San Antonio, TX, 2000, pp. 423–428.
- [19] R. Peslin, F. da Silva, C. Duvivier, and F. Chabot, "Respiration dynamics studied by multiple linear regression in unsedated ventilated patients," *Eur. Respiration J.*, vol. 5, pp. 871–878, 1992.
- [20] H. Roubos, S. Mollov, R. Babuška, and H. B. Verbruggen, "Fuzzy model based predictive control by using Takagi–Sugeno fuzzy models," *Int. J. Approx. Reason.*, vol. 22, pp. 3–30, 1999.
- [21] R. Shorten, R. Murray-Smith, R. Bjørgan, and H. Gollee, "On the interpretation of local models in blended multiple model structures," *Int. J. Control*, vol. 72, pp. 620–628, 1999.
- [22] M. Sugeno and G. T. Kang, "Structure identification of fuzzy model," *Fuzzy Sets Syst.*, vol. 26, pp. 15–33, 1988.
- [23] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its application to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, pp. 116–132, Feb. 1985.
- [24] J. Yen, L. Wang, and C. W. Gillespie, "Improving the interpretability of TSK fuzzy models by combining global learning and local learning," *IEEE Trans. Fuzzy Syst.*, vol. 6, pp. 530–537, Aug. 1998.
- [25] K. Maertens, T. A. Johansen, and R. babuska, "Engine load precision in off-road vehicles using multi-objective nonlinear identification," *Control Eng. Practice*, 2004, to be published.

Tor A. Johansen received the Dr.Eng. (Ph.D.) degree in electrical and computer engineering from the Norwegian University of Science and Technology, Trondheim, Norway, in 1994.

He is currently a Professor in the Department of Engineering Cybernetics at the Norwegian University of Science and Technology. From 1995 to 1997, he was a Research Engineer with SINTEF Electronics and Cybernetics, where he is currently a Scientific Advisor. He has been Research Visitor at the University of Southern California, Los Angeles, Technical University in Delft, Delft, The Netherlands, and the University of California at San Diego. His research interests include optimization-based control, nonlinear systems, multiple model methods, and industrial applications of systems engineering and real-time control.

Dr. Johansen serves as an Associate Editor of *Automatica*. He is a former Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS, a Member of the IFAC Technical Committee on Neural and Fuzzy Systems, and a former Member of the IEEE Technical Committee on Fuzzy Systems.

Robert Babuška received the M.Sc. degree in control engineering from the Czech Technical University, Prague, Czech Republic, in 1990 and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 1997.

Currently, he is a Professor in the Department of Control Systems Engineering, Faculty of Information Technology and Systems, Delft University of Technology. He has coauthored more than 50 journal papers and chapters in books and has published the research monograph *Fuzzy Modeling for Control* (Norwell, MA: Kluwer, 1998). His research interests include the use of fuzzy set techniques and neural networks in nonlinear system identification and control.

Dr. Babuška serves as an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS and *Engineering Applications of Artificial Intelligence*, and as an Area Editor of *Fuzzy Sets and Systems*. He is the Chairman of the IFAC Technical Committee on Cognition and Control.