Incorporating Prior Knowledge in Fuzzy Model Identification

J. Abonyi[†], R. Babuška, H.B. Verbruggen, F. Szeifert[*]

Delft University of Technology, Department of Information Technology and Systems

Control Engineering Laboratory, P.O.Box 5031 2600 GA Delft, The Netherlands

[*]University of Veszprem Department of Chemical Engineering Cybernetics

P.O. Box 158, H-8201, Hungary

Abstract

This paper presents an algorithm for incorporating a priori knowledge into data-driven identification of dynamic fuzzy models of the Takagi-Sugeno type. Knowledge about the modelled process such as its stability, minimal or maximal static gain, or the settling time of its step response can be translated into inequality constraints on the consequent parameters. By using input-output data, optimal parameter values are then found by means of quadratic programming. The proposed approach has been applied to the identification of a laboratory liquid level process. The obtained fuzzy model has been used in model-based predictive control. Real-time control results show that when the proposed identification algorithm is applied, not only physically justified models are obtained, but also the performance of the model-based controller improves with regard to the case where no prior knowledge is involved.

1 Introduction

Recent years have witnessed a rapid growth in the use of fuzzy controllers for complex and poorly defined processes. Most fuzzy controllers developed until now are of the rule-based type, where the rules in the controller model the operator's response in particular process situations.

An alternative approach to the design of fuzzy controllers is the use of more advanced model-based design methods, including the system modeling and identification steps. When the process under control is nonlinear and cannot be described by first principles with sufficient accuracy, it is advantageous to use

[†]On leave from the University of Veszprem, Department of Chemical Engineering Cybernetics, P.O. Box 158, H-8201, Hungary, sponsored by the Hungarian Ministry of Culture and Education (MKM) Eötvos Foundation.

fuzzy modeling as a way of combining first-principle knowledge, linguistic rules describing the system, and process data.

Fuzzy identification is an effective tool for the approximation of uncertain nonlinear systems on the basis of measured data (Hellendoorn and Driankov, 1997). Data-driven identification techniques alone, however, sometimes yield unrealistic models in terms of steady-state characteristics, local linear behavior or physically impossible parameter values. This is typically due to insufficient information content of the identification data and due to over-parameterization of the models.

The Takagi-Sugeno (TS) fuzzy model is often used to represent nonlinear dynamic systems, by interpolating between local linear, time-invariant (LTI) ARX models. The TS fuzzy model is over-parameterized and when data-driven identification is used, the model can exhibit regimes which are not found in the original system (Babuška, 1998). It is demonstrated in this paper that this problem can be remedied by incorporating prior knowledge into the identification method.

Recently, combinations of a priori knowledge with black-box modeling techniques have been gaining considerable interest. Two different approaches can be distinguished: gray-box modeling and semi-mechanistic modeling. In gray-box modeling, a priori information enters a black-box model, for instance, as constraints on the model parameters or variables, smoothness of the system behavior, or open-loop stability (Tulleken, 1993; Johansen, 1996). One can also start with deriving a model based on first principles and include black-box elements as parts of the white-box model. This approach is usually denoted as hybrid-modeling or semi-physical modeling (Schubert, 1994; Thompson and Kramer, 1994; Psichogios and Ungar, 1992; van Can, et al., 1997).

The main contribution of this article is the development of a gray-box modeling approach for data-driven identification of dynamic Takagi–Sugeno (TS) fuzzy models. The main idea is to constrain the candidate model parameters of the rules in the TS fuzzy model. Knowledge about the process such stability, minimal or maximal gain, or the settling time are translated into inequality constraints on the parameters. The fuzzy model then can be identified from input-output data by quadratic programming.

The proposed approach is applied to a laboratory liquid level process. A fuzzy model is first obtained from input-output measurements by using the proposed identification technique. The model is then used in model-based predictive control. Real-time control results show that when the gray-box identification algorithm is used, not only physically justified model is obtained, but also the performance of the model-based controller is improved with regard to the case where no prior knowledge is used.

The paper is organized as follows. In Section 2, the applied TS fuzzy model. Section 3 describes how prior knowledge can be implemented as constraints in the data-driven generation of a fuzzy model. In Section 4, the identification technique is detailed and Section 5 presents the application example. Conclusions are given in Section 6.

2 The Takagi-Sugeno fuzzy model

This paper addresses the identification of fuzzy models with the structure proposed by Takagi and Sugeno (1985). This fuzzy model consists of a set of rules of the following form:

$$R_{i_1,\ldots,i_n}: \textbf{ If } z_1 \text{ is } A_{1,i_1} \textbf{ and } \ldots \textbf{ and } z_n \text{ is } A_{n,i_n} \textbf{ then } y = f_{i_1,\ldots,i_n}(z_1,\ldots,z_n), \tag{1}$$

where n is the number of inputs, $\mathbf{z} = [z_1, \ldots, z_n]$ is a vector containing all the inputs of the fuzzy model and $A_{j,i_j}(z_j)$ is the i_jth antecedent fuzzy set for the jth input. The same symbol is used to denote a fuzzy set and its membership function. M_j is the number of the fuzzy sets on the jth input domain.

$f_{i_1,\ldots,i_n}(\mathbf{z})$ is a (crisp) consequent function. For a given input, $\mathbf{z}$, the output of the fuzzy model, y, is inferred by computing the weighted average:

$$y = \frac{\sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \beta_{i_1,\ldots,i_n} f_{i_1,\ldots,i_n}(z_1,\ldots,z_n)}{\sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \beta_{i_1,\ldots,i_n}}, \tag{2}$$

where the weight, $\beta_{i_1,\ldots,i_n} > 0$, is the overall truth value of the $i_1,\ldots,i_n$th rule calculated as:

$$\beta_{i_1,\ldots,i_n} = \prod_{j=1}^{n} A_{j,i_j}(z_j)\,. \tag{3}$$

Triangular membership functions are used in this article to define the antecedent fuzzy sets as shown in Figure 1, where a_{j,i_j} denotes the cores of fuzzy sets defined by:

$$a_{j,i_j} = \text{core}(A_{j,i_j}(z_j)) = \left\{ z_j | A_{j,i_j}(z_j) = 1 \right\}\,. \tag{4}$$

The support of a set is determined by the cores of the adjacent fuzzy sets:

$$\begin{aligned} A_{j,i_j}(z_j) &= \frac{z_j - a_{j,i_j-1}}{a_{j,i_j} - a_{j,i_j-1}},\ a_{j,i_j-1} \leq z_j < a_{j,i_j} \\ A_{j,i_j}(z_j) &= \frac{a_{j,i_j+1} - z_j}{a_{j,i_j+1} - a_{j,i_j}},\ a_{j,i_j} \leq z_j < a_{j,i_{j+1}} \end{aligned} \tag{5}$$

This ensures that the sum of the membership functions is equal to one. Constraints such as this one help to obtain an interpretable, grid-type partitioning rule-bases. The consequent-estimation method presented in this paper is, however, independent of the membership functions used.

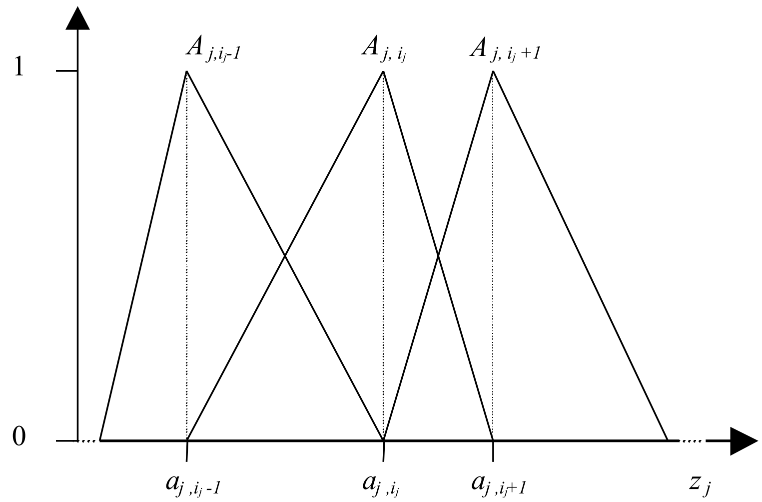

Figure 1. *The membership functions used.*

As the product operator (3) is used for the 'and' connective, the overall truth values of the rules fulfill:

$$\sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \beta_{i_1,\ldots,i_n} = 1\,. \tag{6}$$

Therefore, (2) can be simplified to:

$$y = \sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \left[\left(\prod_{j=1}^{n} A_{j,i_j}(z_j) \right) f_{i_1,\ldots,i_n}(z_1,\ldots,z_n) \right]\,. \tag{7}$$

3 Fuzzy identification of nonlinear dynamic systems using a priori knowledge

The Nonlinear AutoRegressive with eXogenous input (NARX) model is frequently used within many nonlinear identification methods, such as neural network models and fuzzy models. This model establishes a nonlinear relation between the past inputs and outputs and the predicted output:

$$y(k+1) = F\left(y(k),\ldots,y(k-n_y+1),u(k-n_d),\ldots,u(k-n_u-n_d+1)\right)\,. \tag{8}$$

Here, n_y and n_u are the maximum lags considered for the output, and input terms, respectively, n_d is the discrete dead time, and F represents the mapping of the fuzzy model.

The TS fuzzy model of the NARX type interpolates between local linear, time-invariant (LTI) ARX models as follows:

$$\begin{aligned} R_{i_1,\ldots,i_n} &: \ \textbf{If } z_1(k) \text{ is } A_{1,i_1} \textbf{ and } \ldots \textbf{ and } z_n(k) \text{ is } A_{n,i_n} \textbf{ then} \\ y^{i_1,\ldots,i_n}(k+1) &= \sum_{i=1}^{n_y} a_i^{i_1,\ldots,i_n} y(k-i+1) + \sum_{i=1}^{n_u} b_i^{i_1,\ldots,i_n} u(k-i-n_d+1) + c^{i_1,\ldots,i_n}\,. \end{aligned} \tag{9}$$

Here, $\mathbf{z}(\mathbf{k})$ is usually a subset of $\{y(k), \ldots, y(k-n_y+1), u(k-n_d), \ldots, u(k-n_u-n_d+1)\}$.

This fuzzy model can be regarded as a linear parameter varying (LPV) system:

$$\sum_{i=0}^{n_y} a_i y(k-i+1) = \sum_{i=1}^{n_u} b_i u(k-i-n_d+1) + c \tag{10}$$

with

$$a_0 = 1 \tag{11}$$

$$a_i = -\sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \left(\prod_{j=1}^{n} A_{j,i_j}(z_j) \right) a_i^{i_1,\ldots,i_n}, \quad i = 1, \ldots, n_y \tag{12}$$

$$b_i = \sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \left(\prod_{j=1}^{n} A_{j,i_j}(z_j) \right) b_i^{i_1,\ldots,i_n}, \quad i = 1, \ldots, n_u \tag{13}$$

$$c = \sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \left(\prod_{j=1}^{n} A_{j,i_j}(z_j) \right) c^{i_1,\ldots,i_n}. \tag{14}$$

The parameters a_i, b_i and c_i are always bounded within the convex sets (polytopes) delimited by the individual rules' parameters. This follows from equations (11) to (14) and the fact that the membership degrees add up to one, see (6).

It has been shown, that certain types of a priori knowledge about an LTI model can be expressed in the form of linear inequality constraints (Tulleken, 1993; Karny, et al., 1995; Timmons, et al., 1997):

$$\Lambda_{LTI} \boldsymbol{\theta}_{LTI} \leq \omega_{LTI}, \tag{15}$$

where $\boldsymbol{\theta}_{LTI} = \left[a_1, a_2, \ldots, a_{n_y}, b_1, \ldots, b_{n_u}, c\right]$ denotes the parameters of the LTI model. These constraints on the LTI model parameters define a convex parameter set Ω:

$$\Omega = \{\boldsymbol{\theta}_{LTI} | \Lambda_{LTI} \theta_{\mathbf{LTI}} \leq \omega_{\mathbf{LTI}}\}\ . \tag{16}$$

The aim of this paper to incorporate important prior knowledge into the fuzzy model (9). The parameter set, $\boldsymbol{\theta}_{LTI}$, of the LTI model (10) has to be a subset of Ω. Because of the convexity of Ω and the convexity of the applied fuzzy inference method (11)–(14), it is sufficient to check the constraints for the rule consequents. This means that the constraints can easily be adapted to the TS fuzzy model:

$$\Lambda^* \boldsymbol{\theta}_{i_1,\ldots,i_n} \leq \omega^*, \tag{17}$$

where $\boldsymbol{\theta}_{i_1,\ldots,i_n} = \left[a_1^{i_1,\ldots,i_n}, a_2^{i_1,\ldots,i_n}, \ldots, a_{n_y}^{i_1,\ldots,i_n}, b_1^{i_1,\ldots,i_n}, \ldots, b_{n_u}^{i_1,\ldots,i_n}, c^{i_1,\ldots,i_n}\right]$ denotes the parameters of the $i_1, \ldots, i_n$th local model, $\boldsymbol{\Lambda}^*$ and $\boldsymbol{\omega}^*$ represent global constraints on the fuzzy model. In the following, it will be shown how prior knowledge about the stability, stationary gains and the settling time of the process can be transformed into the linear inequalities (17).

3.1 Prior knowledge about process stability

It is well known that the poles of a stable discrete-time model that emerge from a correctly sampled, continuous-time system cannot be situated in $\mathbb{C}^-$, the left half of the complex plane. This knowledge on sampling can be translated into inequality constraints on the parameters on the model parameters:

$$(-1)^i \, a_i \geq 0, \quad 1 \leq i \leq n_y \, . \tag{18}$$

Additional constraints can be derived from stability considerations. Let $C(m, R)$ denote the set of complex numbers within or at a circle with a real-valued center m and radius R,

$$C(m, R) = \{z \in \mathbb{C} \mid \|z - m\| \leq R; \; m \in \mathbb{R}; \; R \in \mathbb{R}^+\} \, . \tag{19}$$

Tulleken (1993) has shown that for the poles of a linear discrete-time system to be in $C(m, R)$, the following linear constraints on the model parameters must be satisfied:

$$\Re \cdot \left[1, a_1, a_2, \ldots, a_{n_y}\right]^T \;\; \geq \;\; 0, \tag{20}$$

$$\Re \cdot \Im \cdot \left[1, a_1, a_2, \ldots, a_{n_y}\right]^T \;\; \geq \;\; 0, \tag{21}$$

where the non-zero elements of the right and left triangular matrices $\Re$ and $\Im$ are defined by

$$[\Re]_{ij} = (-1)^i \binom{j}{i}, \quad i = 0, 1, \ldots, n_y, \; j = 0, 1, \ldots, n_y, \tag{22}$$

$$[\Im]_{ij} = (m - R)^{i-j} \binom{n_y - j}{i - j} (2R)^{n_y - i} \, . \tag{23}$$

If $m = 0$ and $R = 1$, the previous equation constitutes the smallest convex hull of the admissible parameter region corresponding to the stable system model having all poles in the unit circle, $C(0, 1)$, in the complex plane. Therefore, these are necessary conditions for asymptotic stability. More details about the derivation of the above constraints can be found in (Tulleken, 1993).

Example 1. Let us consider a second-order system with $n_y = 2$ and $n_u = 1$: $y(k+1) = a_1 y(k) + a_2 y(k-1) + b_1 u(k) + c$. In this case $\Re \cdot \Im$ becomes:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ -4 & 2 & 0 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix} . \tag{24}$$

It can be seen from the coefficients of the above matrices and equation (20) that for the given second-order system the following holds:

$$\begin{aligned} 1 + a_1 + a_2 &\geq 0, \\ -a_1 - 2a_2 &\geq 0, \\ a_2 &\geq 0\,. \end{aligned}$$

From (21) and (24), the following inequality constraints are obtained:

$$\begin{aligned} 1 + a_1 + a_2 &\geq 0, \\ 2 - 2a_2 &\geq 0, \\ 1 - a_1 + a_2 &\geq 0\,. \end{aligned}$$

3.2 Knowledge about process gain

Often, not only the stability of the plant is known beforehand, but also the admissible intervals or at least the signs of the static gains are known. Since the steady-state gain is only defined for stable systems, open-loop stability must also be imposed on the estimates (see the Section 3.1). Given the minimal and maximal gains $K_{\min}$ and $K_{\max}$, for the steady-state gain of a linear system, the following holds:

$$K_{\min} \leq \frac{\sum_{i=1}^{n_b} b_i}{1 + \sum_{i=1}^{n_a} a_i} \leq K_{\max}\,. \tag{25}$$

This knowledge can be represented by the following linear inequality constrains:

$$K_{\min}\left(1 + \sum_{i=1}^{n_a} a_i\right) - \sum_{i=1}^{n_b} b_i \;\leq\; 0, \tag{26}$$

$$-K_{\max}\left(1 + \sum_{i=1}^{n_a} a_i\right) + \sum_{i=1}^{n_b} b_i \;\leq\; 0\,. \tag{27}$$

3.3 Knowledge about the settling time

The approximate settling time (the time required for a step response to settle within a band around its final value) is usually known. For second-order discrete-time systems, a known maximum settling time τ_s approximately translates to all poles lying within a circle centered at the origin in the z-plane with the radius $r = \exp\left(-4.6\frac{T_s}{\tau_s}\right)$, where T_s denotes the sampling time. Based on this consideration, the following inequality constraints can be developed, based on the method presented by Timmons, et al. (1997):

$$r_{\max}^2 a_1 + (r_{\min} - 2r_{\max})(-a_2) \;\leq\; -r_{\max}^2 r_{\min}, \tag{28}$$

$$
\begin{aligned}
-r_{\max}a_1 - a_2 &\leq r_{\max}^2, \\
-a_2 &\leq 0 .
\end{aligned}
$$

3.4 Knowledge about the nonlinearity

For nonlinear processes, besides the bounds on the gain and/or the settling time, the tendency of the change of these parameters may be known as well.

Example 2. Let us consider a nonlinear liquid level process (Posthlethwaite, 1994):

$$
A\frac{dy}{dy} = u - \alpha\sqrt{y}, \tag{29}
$$

where A is the cross-sectional area of the tank, y is the liquid level in the tank, u is the inlet flow-rate, and α is the outflow coefficient. The gain and the time constant of the process can be obtained by linearizing the process model: $K = \frac{2\sqrt{y}}{\alpha}$, $\tau = \frac{2A\sqrt{y}}{\alpha}$. The gain and the time constant thus increase when the level of the liquid in the tank increases.

This kind of a priori knowledge can be included in the TS fuzzy model identification by using local constraints instead of the global ones given in equation (17):

$$
\Lambda_{i_1,\ldots,i_n}\boldsymbol{\theta}_{i_1,\ldots,i_n} \leq \omega_{i_1,\ldots,i_n} . \tag{30}
$$

Here, $\Lambda_{i_1,\ldots,i_n}$ and $\omega_{i_1,\ldots,i_n}$ are the constraints of the $i_1,\ldots,i_n$th rule parameter vector $\boldsymbol{\theta}_{i_1,\ldots,i_n}$. By using local constraints, the prior information on the nonlinear behaviour around a given operating point can be used in the identification procedure, while the global constraints (17) represent knowledge that is independent of the operating point.

4 Constrained identification of TS fuzzy models

The output of the Takagi-Sugeno fuzzy model is linear in the consequent parameters. In the unconstrained case (no prior knowledge used), these parameters can be estimated by linear least-squares techniques. A global or a local approach can be followed. With the global approach the parameters of all rule consequents are estimated within one least-squares problem, yielding an optimal predictor. Due to its cubic complexity, global parameter estimation becomes computationally expensive for fuzzy systems with many rules (Fischer and Nelles, 1998). The local parameter estimation approach does not estimate all parameters simultaneously. It rather divides this task into $\prod_{j=1}^{n} M_j$ weighted least-squares problems. This method forces the local linear models to fit the data locally, but it does not give an optimal TS model in terms of a minimal global prediction error. As the identified dynamic fuzzy model will be used in model-based predictive control, it is important that it gives the least possible prediction error. Therefore, the global estimation approach is pursued in this paper.

The N data pairs and the truth values of the fuzzy rules are arranged in the following matrices.

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z^1} \\ \mathbf{z^2} \\ \vdots \\ \mathbf{z^N} \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^N \end{bmatrix}, \ \mathcal{B}_{i_1,\ldots,i_n} = \begin{bmatrix} \beta^1_{i_1,\ldots,i_n} & 0 & \cdots & 0 \\ 0 & \beta^2_{i_1,\ldots,i_n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta^N_{i_1,\ldots,i_n} \end{bmatrix}. \tag{31}$$

Appending a unitary column to $\mathbf{Z}$ gives the extended regression matrix $\mathbf{Z_e} = [\mathbf{Z}\ \mathbf{1}]$. Denote $\mathbf{Z}'$ the matrix composed of matrices $\mathcal{B}_{i_1,\ldots,i_n}$ and $\mathbf{Z_e}$ as follows:

$$\mathbf{Z}' = [\mathcal{B}_{1,\ldots,1,1}\mathbf{Z_e}, \mathcal{B}_{1,\ldots,1,2}\mathbf{Z_e}, \ldots, \mathcal{B}_{M_1,\ldots,M_n}\mathbf{Z_e}] \tag{32}$$

Denote $\boldsymbol{\theta}'$ the vector given by $\boldsymbol{\theta}' = [\boldsymbol{\theta}_{1,\ldots,1,1}, \boldsymbol{\theta}_{1,\ldots,1,2}, \ldots, \boldsymbol{\theta}_{M_1,\ldots,M_n}]$. The least-squares method can be applied to solve the underlying regression problem $\mathbf{y} = \mathbf{Z_e}\boldsymbol{\theta}' + \varepsilon$:

$$\boldsymbol{\theta}' = \left[\left(\mathbf{Z}'\right)^T \mathbf{Z}'\right]^{-1} \left(\mathbf{Z}'\right)^T \mathbf{y}\,. \tag{33}$$

In order to be able to use the constraints derived in the previous section, quadratic programming (QP) has to be used instead of the least-squares method.

The constrained optimization problem can be formulated as follows:

$$\min_{\boldsymbol{\theta}'} \left\{ \frac{1}{2} \left(\boldsymbol{\theta}'\right)^T \mathbf{H}\theta' + \mathbf{c^T}\boldsymbol{\theta}' \right\} \tag{34}$$

with $\mathbf{H} = \mathbf{2}\left(\mathbf{Z}'\right)^{\mathbf{T}} \mathbf{Z}'$, $\mathbf{c} = -\mathbf{2}\left(\mathbf{Z}'\right)^{\mathbf{T}} \mathbf{y}$ and the constraints on $\boldsymbol{\theta}'$:

$$\boldsymbol{\Lambda}'\boldsymbol{\theta}' \leq \boldsymbol{\omega}' \tag{35}$$

where $\boldsymbol{\Lambda}' = [\boldsymbol{\Lambda}_{1,\ldots,1,1}, \boldsymbol{\Lambda}_{1,\ldots,1,2}, \ldots, \boldsymbol{\Lambda}_{M_1,\ldots,M_n}]$ and $\boldsymbol{\omega}' = [\boldsymbol{\omega}_{1,\ldots,1,1}, \boldsymbol{\omega}_{1,\ldots,1,2}, \ldots, \boldsymbol{\omega}_{M_1,\ldots,M_n}]$

5 Example: identification for model predictive control of a liquid level process

This real-time control example is used to illustrate the advantages of the proposed identification method. A laboratory process consisting of two cascaded tanks is considered. Fuzzy models are constructed from input-output data and from different pieces of prior knowledge. The obtained models are used in a predictive control scheme. It is shown that by using prior knowledge a more reliable control model is obtained.

5.1 Process description

The laboratory process consists of two cascaded tanks depicted in Figure 2. Water is supplied into the upper tank through a controlled peristaltic pump. A pressure transmitter attached to the bottom of the lower tank measures the level of the liquid in this tank. The process is connected to a personal computer through a data acquisition board. The identification and control software is written under MATLAB, using the Real-Time Toolbox to collect data and to control the process. The sampling time is 2.5 s.

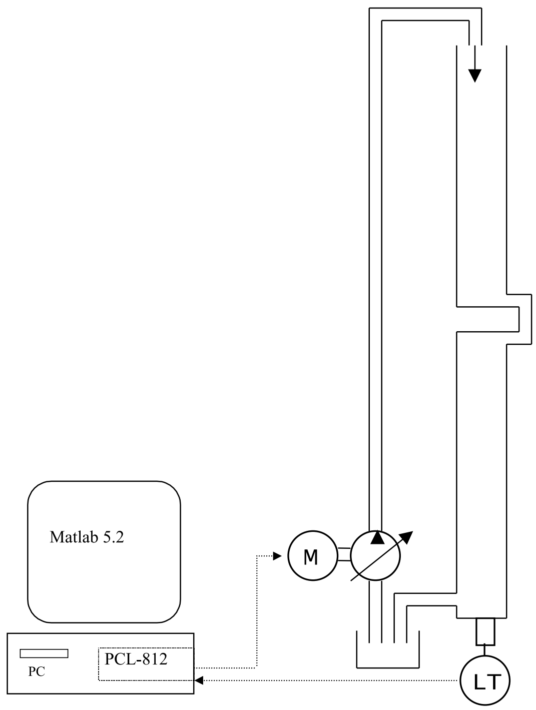

Figure 2. *The cascaded-tanks setup.*

5.2 Fuzzy model identification

The TS model was constructed from process measurements. The input variables of the fuzzy model were selected on the basis of prior knowledge about of the process. Since the system can be modelled approximately as a second-order system with a time delay, and the source of the nonlinearity is the level-dependent outflow from the tank, the antecedent variable of the fuzzy model is chosen to be $\mathbf{z}(\mathbf{k}) = \mathbf{y}(\mathbf{k})$. This results in the following TS fuzzy model structure:

$$R_i: \ \textbf{If } y(k) \text{ is } A_{1,i} \textbf{ then } y^i(k+1) = a_1^i y(k) + a_2^i y(k-1) + b_1^i u(k-2) + c^i \,. \qquad (36)$$

Six triangular fuzzy sets were used on the antecedent universe $y(k)$. Based on the range of the liquid level and after some manual optimization, the cores of the fuzzy sets were selected to be $\{0.0781\ 0.2000\ 0.4000\ 0.5000\ 0.6000\ 0.8081\}$. The identification data set contains $N = 1400$ samples, using an input signal shown in Figure 3.[1] The signal was designed to contain the important frequencies in the *expected* range of the process dynamics. Later on, we will see that the frequency content of this signal is not ideal, which is typical for real-life experiments.

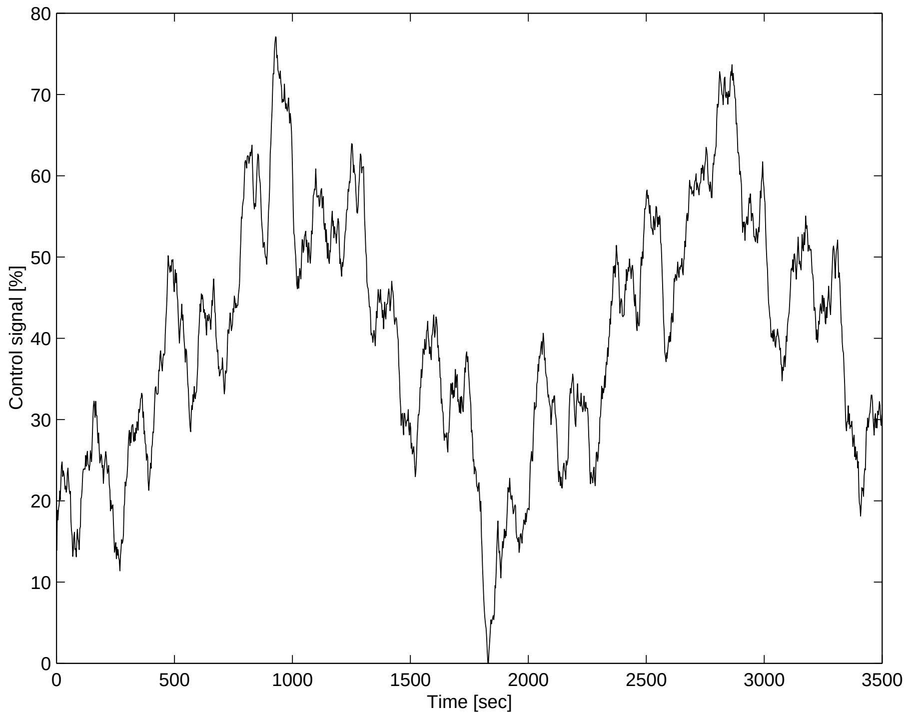

Figure 3. *The training control sequence.*

Three different models were identified:

Model 1 No prior knowledge was used during the identification.

Model 2 Prior knowledge on process stability was used and the minimal and maximal steady-state gain were assumed: $K_{\min} = 0$ and $K_{\max} = 2.5$.

Model 3 The second case was extended by the following constraints on the settling time of each submodel (Table 1).

[1] In this example, the process input and output signals are plotted in percentage of their full range.

Table 1. *Settling times of the local models.*

	1	2	3	4	5	6
$\tau_{\min}$	15	40	60	70	80	80
$\tau_{\max}$	30	60	100	120	140	140

The developed fuzzy models were validated by means of (recurrent) simulation on a separate validation data set. The simulated and the measured liquid levels are shown in Figure 4. The two curves can hardly be distinguished from each other.

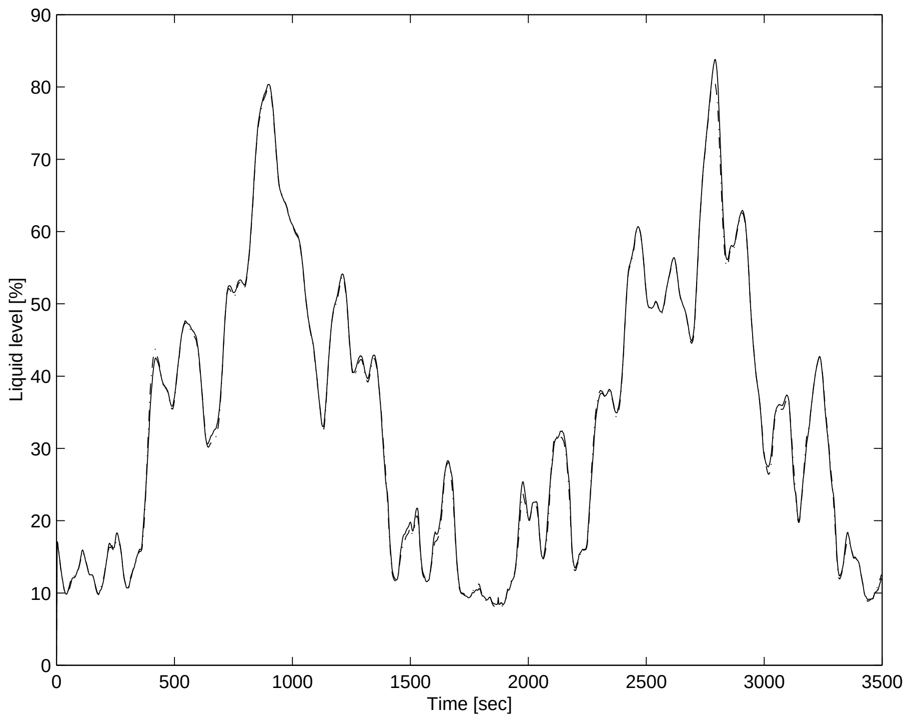

Figure 4. *The measured (–) and simulated (- -) process output.*

The performance of the obtained models was measured by the variance accounted for (VAF) index given by:

$$\text{VAF} = 100\% \left[1 - \frac{var\,(y - y_m)}{var\,(y)}\right], \tag{37}$$

where y is the true process output and y_m is the model output obtained by simulation of the fuzzy model. The VAF of two equal signals is 100 %, if the signals differ, the VAF value is lower.

Models 1 and 2 both achieve the performance VAF =99.89 and Model 3 has VAF=99.9. Clearly, there is no significant difference between the performance of the fuzzy models identified with and without the use of a priori knowledge. It is interesting to note that even though the modeling performance is almost unchanged, the consequent parameters of the three models differ considerably (Table 2).

Table 2. *Parameters of the local models obtained by different identification method.*

Model 1

	1	2	3	4	5	6
a_1^i	-0.2419	-0.0398	-0.0217	-0.0238	0.0375	0
a_2^i	-0.4630	-0.6574	-0.6499	-0.6470	-0.7032	-0.6680
b_1^i	0.0217	0.0691	0.0350	0.0392	0.0325	0.0384
c^i	0.1315	0.3112	0.6487	0.8094	0.9753	1.3141

Model 2

	1	2	3	4	5	6
a_1^i	1.1271	1.3291	1.3472	1.3452	1.4064	1.3689
a_2^i	-0.4630	-0.6574	-0.6499	-0.6470	-0.7032	-0.6680
b_1^i	0.0217	0.0691	0.0350	0.0392	0.0325	0.0384
c^i	0.0246	0.0374	0.1011	0.1250	0.1539	0.2079

Model 3

	1	2	3	4	5	6
a_1^i	1.3632	1.5495	1.6560	1.6134	1.6827	1.6422
a_2^i	-0.4646	-0.6047	-0.6815	-0.6407	-0.7049	-0.6668
b_1^i	0.0297	0.0718	0.0333	0.0394	0.0324	0.0384
c^i	0.0055	-0.0182	-0.0088	-0.0124	-0.0107	-0.0139

This effect is caused by the fact that the unconstrained global least-squares estimation method biases the estimate of the local model parameters, that may hamper the local interpretation of the TS fuzzy model, while the prior-knowledge-based constraints regularize the model parameters. If the incorporated constraints are adequate, this regularization does not result in worse modeling performance. Moreover, if the identification data do not contain all information about the process (which is the usual case), the constraints can even improve the modeling performance.

Figure 5 shows three measured step responses of the process and the step responses of the fuzzy model at these operating points.

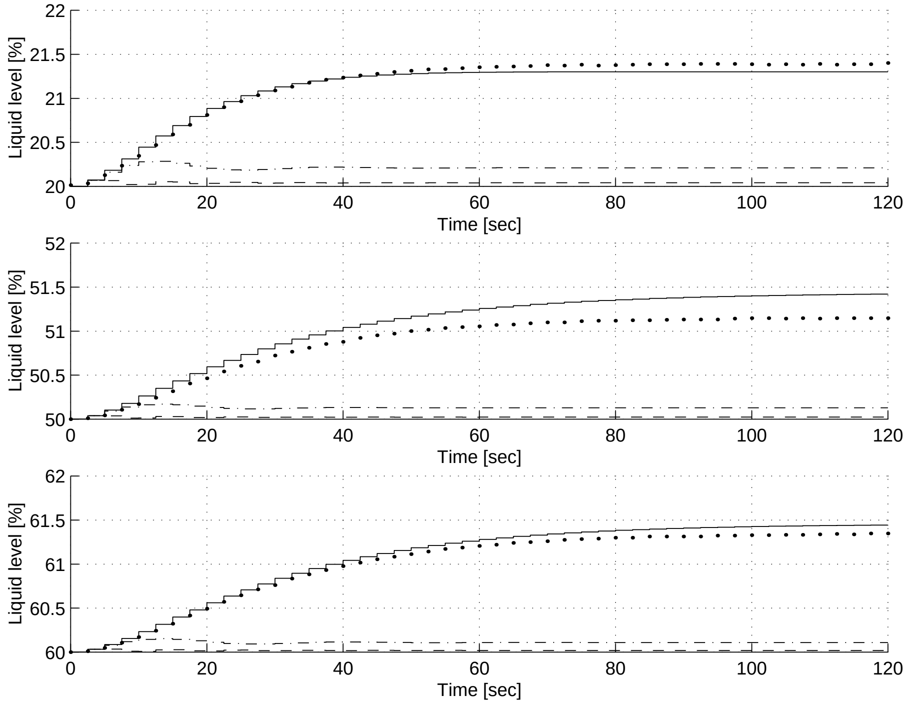

Figure 5. *Step responses of the fuzzy models (- - model 1, - · model 2, — model 3) and the measured step response at the corresponding operating points (dotted curve).*

The generated step responses clearly show the above-mentioned regularization effect of the constraints. One can see that by using more prior knowledge, more accurate local models can be identified from the global process data. The large deviation in the step response and the small difference in the modeling performance is mainly caused by the offset parameters, c^i, of the rules. By using the proposed constrained identification method, the offset parameters influence the steady-state behavior of the plant only (as one would normally expect). The reason why the constraints on the settling time give the most significant improvement is that the identification data do not contain enough information on low-frequency and steady-state behavior of the system. This is a typical real-life situation, as it is often difficult to design identification experiments that yield data with an appropriate balance between low and high frequencies. An advantage of the proposed method is that the information on low-frequency behavior can originate from other sources or separate experiments (step responses) and it can easily be integrated with the information coming from the dynamic data.

5.3 Fuzzy model-based predictive controller

A generalized predictive controller (GPC) has been designed for the considered process, using the fuzzy models obtained with the proposed identification method. The GPC algorithm computes the control sequence $\{u(k+j)\}$ such that the following quadratic cost function (Clarke, et al., 1989) is minimized:

$$J(H_{p1}, H_{p2}, H_c, \lambda) = \sum_{j=H_{p1}}^{H_{p2}} \left(w\left(k+j\right) - \hat{y}\left(k+j\right)\right)^2 + \lambda \sum_{j=1}^{H_c} \Delta u^2(k+j-1)\,. \tag{38}$$

Here, $\hat{y}(k+j)$ denotes the predicted process output, H_{p1} is the minimum costing horizon, H_{p2} is the maximum costing or prediction horizon, H_c is the control horizon, and λ is a weighting coefficient. The predicted outputs are calculated by using the following approximate step-response model:

$$\hat{y}(k+j) = \sum_{i=1}^{j} g_i \Delta u(k+j-1) + p_j, \tag{39}$$

where g_i represent the plant step-response coefficients, obtained from the fuzzy model at the current operating point, and p_j is the response of the fuzzy model at the $k+j$th step assuming that the future control signal remains constant. By combining all predicted outputs into the vector $\hat{\mathbf{y}} = [\hat{y}(k+H_{p1}), \ldots, \hat{y}(k+H_{p2})]$, the key equation of the GPC algorithm can be formulated:

$$\hat{\mathbf{y}} = \mathbf{G}\boldsymbol{\Delta}\overline{\mathbf{u}} + \mathbf{p} \tag{40}$$

Here, $\boldsymbol{\Delta}\overline{\mathbf{u}} = [\Delta u(k), \ldots, \Delta u(k+H_c)]$, and $\mathbf{p} = [p_1, p_2, \ldots, p_{H_{p_2}}]$, where the elements of $\mathbf{p}$ are obtained by using the fuzzy model as follows:

$$\begin{aligned} p_1 &= f(\hat{y}(k), \hat{y}(k-1), u(k-2)), \\ p_2 &= f(p_1, y(k), u(k-1)), \\ p_j &= f(p_{j-1}, p_{j-2}, u(k-1)), \quad \text{for } j > 2\,. \end{aligned} \tag{41}$$

$\mathbf{G}$ is a $(H_{p2} - H_{p1} + 1) \times H_c$ matrix with zero entries g_{ij} for $j - i > H_{p1}$:

$$\mathbf{G} = \begin{bmatrix} g_{H_{p1}} & g_{H_{p1}-1} & \cdots & 0 \\ g_{H_{p1}+1} & g_{H_{p1}} & g_{H_{p1}-1} & \\ \vdots & & \ddots & \\ g_{H_{p2}} & g_{H_{p2}-1} & \cdots & g_{H_{p2}-H_c} \end{bmatrix}. \tag{42}$$

The elements of this matrix are the step-response coefficients obtained from the fuzzy model at the current operating point:

$$g_j \;=\; 0,\ \forall j \le n_d, \tag{43}$$

$$g_j = -\sum_{i=1}^{j} a_i g_{j-i} + \sum_{i=1}^{j} b_i, \quad \text{for} j > n_d \,.$$

When constraints are considered, the minimum of (38) can be found by quadratic optimization with linear constraints:

$$\min_{\overline{\mathbf{u}}} \left\{ (\mathbf{G}\boldsymbol{\Delta}\overline{\mathbf{u}} + \mathbf{p} - \mathbf{w})^T (\mathbf{G}\boldsymbol{\Delta}\overline{\mathbf{u}} + \mathbf{p} - \mathbf{w}) + \lambda \boldsymbol{\Delta}\overline{\mathbf{u}}^{\mathbf{T}} \boldsymbol{\Delta}\overline{\mathbf{u}} \right\} = \min_{\overline{\mathbf{u}}} \left\{ \frac{1}{2} \boldsymbol{\Delta}\overline{\mathbf{u}}^{\mathbf{T}} \mathbf{H} \boldsymbol{\Delta}\overline{\mathbf{u}} + \mathbf{d}\boldsymbol{\Delta}\overline{\mathbf{u}} \right\} \quad (44)$$

with $\mathbf{H} = \mathbf{2}\left(\mathbf{G}^{\mathbf{T}}\mathbf{G} + \lambda\mathbf{I}\right)$, $\mathbf{d} = -\mathbf{2}\left(\mathbf{G}^{\mathbf{T}}(\mathbf{w} - \mathbf{p})\right)$, where $\mathbf{I}$ is an $(H_c \times H_c)$ unity matrix.

The constraints on u and Δu can be formulated with the following inequality:

$$\begin{pmatrix} \mathbf{I}_{\boldsymbol{\Delta}\overline{\mathbf{u}}} \\ -\mathbf{I}_{\boldsymbol{\Delta}\overline{\mathbf{u}}} \\ \mathbf{I}_{\mathbf{H_c}} \\ -\mathbf{I}_{\mathbf{H_c}} \end{pmatrix} \boldsymbol{\Delta}\overline{\mathbf{u}} \leq \begin{pmatrix} \mathbf{u}_{\max} - \mathbf{I}_{\overline{\mathbf{u}}}\mathbf{u}(\mathbf{k}-\mathbf{1}) \\ -\mathbf{u}_{\min} + \mathbf{I}_{\overline{\mathbf{u}}}\mathbf{u}(\mathbf{k}-\mathbf{1}) \\ \boldsymbol{\Delta}\mathbf{u}_{\max} \\ -\boldsymbol{\Delta}\mathbf{u}_{\min} \end{pmatrix} \quad (45)$$

where $\mathbf{I}_{\mathbf{H_c}}$ and $\mathbf{I}_{\overline{\mathbf{u}}}$ is an $H_c \times H_c$ unity matrix, $\mathbf{I}_{\boldsymbol{\Delta}\overline{\mathbf{u}}}$ is an $H_c \times H_c$ lower triangular matrix with all elements equal to one, and $\boldsymbol{\Delta}\mathbf{u}_{\min}$, $\boldsymbol{\Delta}\mathbf{u}_{\max}$, $\mathbf{u}_{\min}$, $\mathbf{u}_{\max}$ are H_c-vectors, with the constraints $\Delta u_{\min}$, $\Delta u_{\max}$, $u_{\min}$, $u_{\max}$ respectively.

5.4 Control results

In order to compare the control performance of the different fuzzy models, the following quadratic performance index is used:

$$\text{PI} = \sum_{k=1}^{T} \left\{ (w(k) - y(k))^2 + \lambda \Delta u^2(k) \right\}, \quad (46)$$

where T is the number of time steps in a control experiment.

The prediction and the control horizons are selected to be $H_{p1} = 3,\ H_{p2} = 10$, and $H_c = 2$. The move suppression coefficient, $\lambda = 0.15$. The constraints are set to $\Delta u_{\max} = -\Delta u_{\min} = 0.25,\ u_{\max} = 1,\ u_{\min} = 0$.

In order to cope with the model-plant mismatch and unmeasured disturbances, the GPC was implemented within an internal model control (IMC) scheme (Garcia and Morari, 1982), see Figure 6.

The inputs of the controller are the reference w, the predicted liquid level, $\hat{y}$, and the filtered modeling error signal e_{mf}. The following first-order low-pass Butterworth filter was used:

$$e_{mf}(k) = b_{f1} e_m(k) + b_{f2} e_m(k-1) - a_{f2} e_{mf}(k-1) \quad (47)$$

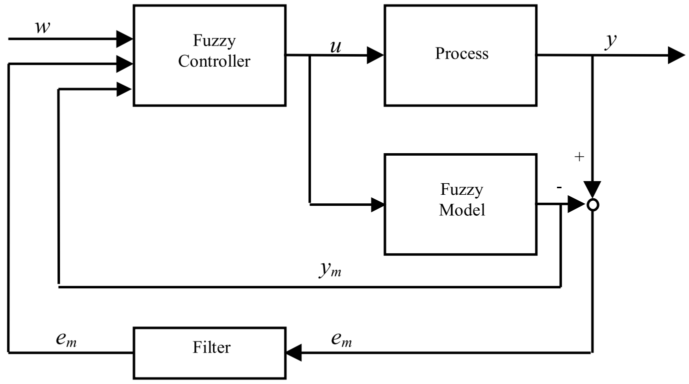

Figure 6. *Implementation of the GPC control in the IMC structure.*

with $b_{f1} = 0.2929$, $b_{f2} = 0.2929$, and $a_{f2} = -0.4142$. The filter parameters were designed based on the desired cut-off frequency in order to reliably filter out the measurement noise and to provide a fast response.

The control performance of the predictive controller based on model 1 (no prior knowledge involved) is shown in Figure 7. The corresponding performance is $PI = 1.2$. Note the oscillations of the control signal, caused by the fact that an inaccurate step-response model is obtained from the fuzzy model (44).

By using prior knowledge, the control performance improves considerably. For model 2, which was identified by using the knowledge about stability and the minimal and maximal steady-state gains, the performance index is $PI = 0.98$. If the knowledge about the approximately settling time is also used, the performance improves even further, yielding $PI = 0.73$. The reason is that by using prior knowledge in the identification, the local linear ARX models in the fuzzy model become more realistic. Therefore, the performance of the model-based controller is much better. The corresponding control result is shown in Figure 8. Note that the control signal becomes less oscillatory. The remaining small fluctuations at low liquid levels are caused by the measurement noise of the level sensor.

6 Conclusions

A new approach to data-driven identification of Takagi–Sugeno fuzzy models has been presented. It allows to translate prior knowledge about the process (including stability, minimal or maximal static gain and settling time) into constraints on the model parameters. This procedure allows the development of TS models also in cases where little experimental data are available. Experimental results have been obtained for a laboratory setup consisting of two cascaded tanks. It has been shown that fuzzy models

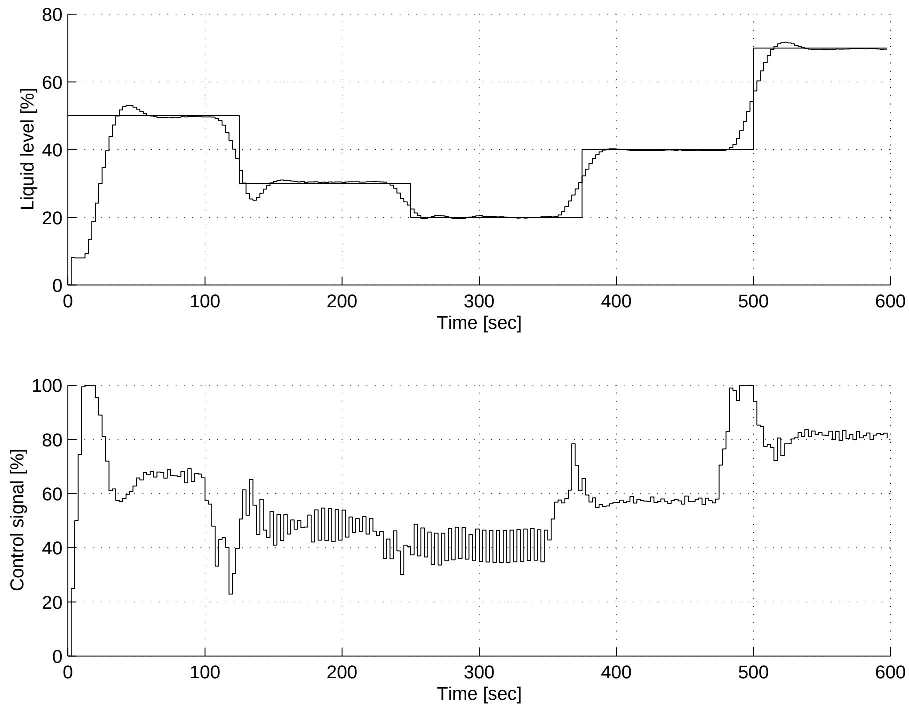

Figure 7. *Performance of the fuzzy-model-based controller when the fuzzy model is identified without using prior knowledge.*

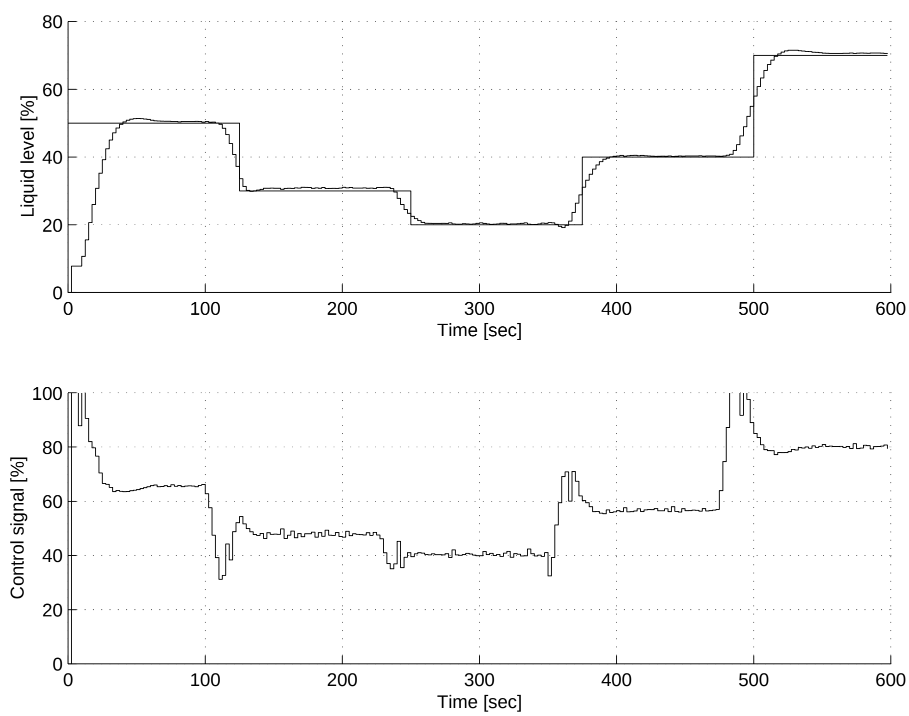

Figure 8. *Performance of the fuzzy-model-based controller when the fuzzy model is identified with the use of prior knowledge.*

built on the basis of data combined with prior knowledge perform better in control than models obtained from data only. This is because the prior-knowledge-based constraints result in better local modeling of the plant.

References

Babuška, R. (1998). *Fuzzy Modeling for Control*. Boston, USA: Kluwer Academic Publishers.

Can, van, H.J.L., H.A.B. te Braake, C. Hellinga, K.Ch.A.M. Luyben and J.J. Heijnen (1997). An efficient model development strategy for bioprocesses based on neural networks in macroscopic balances. *Biotechnology and Bioengineering 54*, 549–566.

Clarke, D. W., C. Mothadi and P.S.C. Tuffs (1989). Generalized predictive control — part I. the basic algorithm. *Automatica 23*, 137–148.

Fischer, M. and O. Nelles (1998). Predictive control based on local linear fuzzy models. *International Journal of Systems Science 29*, 679–697.

Garcia, C.E. and M. Morari (1982). Internal model control: 1. A unifying review and some new results. *Ind. Eng. Chem. Process Res. Dev. 21*, 308–323.

Hellendoorn, H. and D. Driankov (Eds.) (1997). *Fuzzy Model Identification: Selected Approaches*. Berlin, Germany: Springer.

Johansen, T.A. (1996). Identification of non-linear systems using empirical data and a priori knowledge – an optimisation approach. *Automatica 32*, 337–356.

Karny, M., A. Haluskova and P. Nedoma (1995). Recursive approximation by ARX model: A tool for grey-box modelling. *Int. J. Adaptive Control and Signal Processing 9*, 525–546.

Posthlethwaite, B. (1994). A model-based fuzzy controller. *Chem. Eng. Res. Des. (Trans IChemE) 72*, 39–48.

Psichogios, D.C. and L.H. Ungar (1992). A hybrid neural network – first principles approach to process modeling. *AIChE J. 38*, 1499–1511.

Schubert, J. (1994). Bioprocess optimization and control: Application of hybrid modelling. *Journal of Biotechnology 35*, 51–68.

Takagi, T. and M. Sugeno (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Systems, Man and Cybernetics 15*(1), 116–132.

Thompson, M.L. and M.A. Kramer (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE J. 40*, 1328–1340.

Timmons, W.D., H.J. Chizeck and P.G. Katona (1997). Parameter-constrained adaptive control. *Ind. Eng. Chem. Res. 36*, 4894–4905.

Tulleken, H.J.A.F (1993). Gray-box modelling and identification using pysical knowledge and Bayesian techniques. *Automatica 29*, 285–308.