SYSTEM IDENTIFICATION USING FUZZY MODELS

Robert Babuška
Delft University of Technology, Faculty of Information Technology and Systems, The Netherlands

Keywords: Nonlinear system identification, fuzzy logic, neuro-fuzzy systems, discrete-time systems, fuzzy clustering, parameter estimation

Contents

Summary

Modern processes in industry are characterized by nonlinear and time varying behavior, a large number of sensors, many loops, and interactions among them. This makes the design of control systems for these processes difficult and time consuming. Nonlinear system identification is becoming an important tool which can lead to improved control along with considerable time savings and cost reductions. Among the different nonlinear identification techniques, relatively new methods based on fuzzy logic models are gradually becoming established not only in the academia but also in industrial applications.

Fuzzy modeling can be regarded as a gray-box technique on the boundary between nonlinear black-box and qualitative models or expert systems. The tools for building fuzzy models from data are based on a variety of algorithms form the fields of fuzzy logic, approximate reasoning, neural networks, pattern recognition, and regression analysis.

This chapter is an introductory treatment of system identification techniques based on

fuzzy models. It starts with a brief discussion of the position of fuzzy modeling within the general nonlinear identification setting. It presents some of the most commonly used fuzzy models: the Mamdani model and the Takagi-Sugeno model. The basic techniques for the construction of these models are then given along with an illustrative example. The paper concludes with an overview of applications and some suggestions for further reading.

1. Introduction

Modern control systems are confronted with a large number of requirements posed by increasing competition, environmental requirements, energy and material costs and the demand for high quality customer-tailored products. Moreover, the trend in the current plant-wide information and control systems towards the integration of control, scheduling, supervision, fault detection and diagnosis requires methods that can handle qualitative and quantitative information with varying precision and complexity. These considerations impose extra demands on the effectiveness of process modeling techniques and tools. Many systems are not amenable to conventional modeling approaches due to the lack of precise, formal knowledge about the system, due to strongly nonlinear behavior, high degree of uncertainty, or time-varying characteristics.

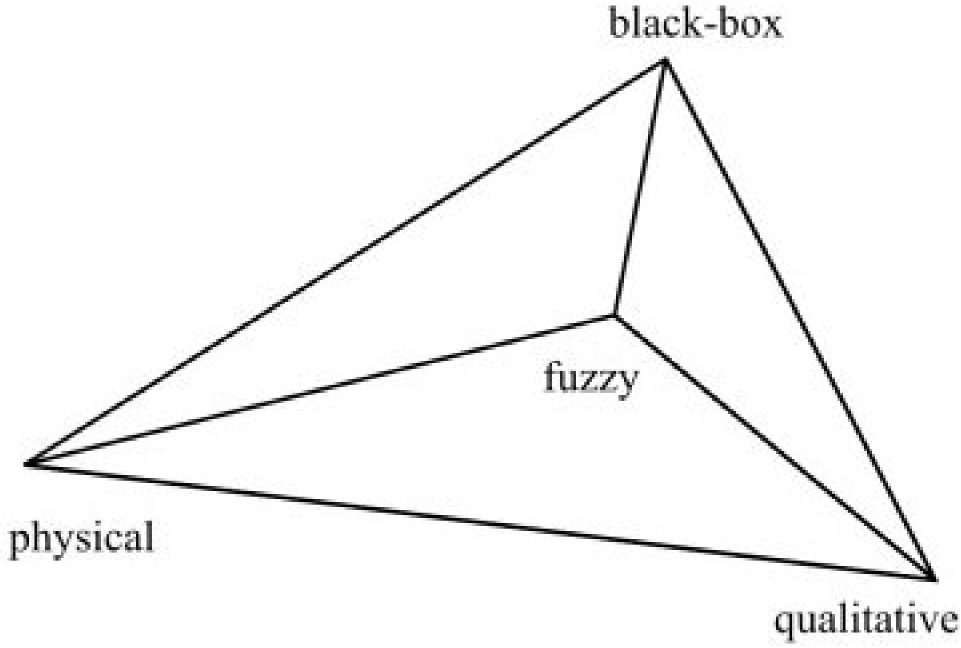

Figure 1: Fuzzy modeling as a flexible framework in between physical, black-box and qualitative modeling.

Fuzzy modeling, along with other related techniques such as neural networks, has been recognized as a powerful tool which can facilitate the effective development of models by combing information from different sources, such as physical laws, empirical models, heuristics and data. Fuzzy models describe systems by relating the relevant variables by means of if-then rules, such as "if x is small then y is large". The meaning of the qualitative terms like 'small' and 'large' is defined by using fuzzy sets, which serve as an interface between the qualitative information in the rule base and numerical data at the inputs and outputs of the model. Thanks to this structure, fuzzy models are to a certain degree transparent to interpretation and analysis, i.e., can be better used to explain solutions to users than completely black-box models such as neural networks. In this sense, we can regard fuzzy modeling as a technique belonging some-where between the standard physical modeling, qualitative methods and nonlinear black-box identification, see Figure1.

2. Nonlinear Dynamic Models for System Identification

A wide class of nonlinear dynamic systems with an input u and an output y can be described in discrete time by the NARX (nonlinear autoregressive with exogenous input) input-output model:

$$\hat{y}(k+1) = f(\mathbf{x}(k)) \quad \text{with} \quad \mathbf{x}(k) = \left[y(k) \ldots y(k-n+1)\ u(k) \ldots u(k-m+1) \right]^T \quad (1)$$

where $\hat{y}(k+1)$ denotes the output predicted at the future time instant $k+1$ and $\mathbf{x}(k)$ is the regression vector, consisting of a finite number of past inputs and outputs. The dynamic order of the system is represented by the number of lags m and n. Depending on the nature of noise disturbances (e.g., sensor noise, process noise, etc.), more complicated model structures can be chosen. Some common examples are the nonlinear output error (NOE) model, which involves the past model predictions instead of the data:

$$\mathbf{x}(k) = \left[\hat{y}(k) \ldots y(k-n+1) u(k) \ldots u(k-m+1) \right]^T, \qquad (2)$$

or the 'innovations' form offering additional freedom in describing the effect of disturbances. In the NARMAX model, for instance, the prediction error $e(k) = y(k) - \hat{y}(k)$ and its past values are included in the regression vector as well:

$$\mathbf{x}(k) = \left[y(k) \ldots y(k-n+1)\ u(k) \ldots u(k-m+1)\ e(k) \ldots e(k-n_e) \right]^T. \qquad (3)$$

For the sake of brevity, the discussion of fuzzy models in the subsequent sections is restricted to the NARX structure (1). In the sequel, we also drop the hat denoting predictions (the model output thus becomes y).

Although for simplicity stated with a scalar input and output, the NARX model can also be used for multivariable systems. In that case, however, the number of components of the regression vector usually becomes large and one may prefer the nonlinear state-space description:

$$\begin{aligned} \xi(k+1) &= g(\xi(k), \mathbf{u}(k)) \\ \mathbf{y}(k) &= h(\xi(k)) \end{aligned} \qquad (4)$$

where the state ξ is assumed to be measured or reconstructed. The problem of nonlinear system identification is to infer the unknown function f (or the functions g and h in the state-space model) from the available data sequences $\{(u(k),\ y(k))\,|k=1,2,\ldots,N\}$. As in most cases only input-output data are available, in the sequel we focus on the regression model (1).

The choice of f, or more specifically its parameterization $f(\mathbf{x};\theta)$ is of outmost importance. To set up the structure for the function f, one can follow two basic ways. The first one is the derivation of a model by using first principles (physical, chemical, biological and other laws).This method relies on a deep understanding of the underlying mechanisms and therefore tends to be expensive, time consuming and usually involves many unknown parameters that must be determined. Another possibility is to use a black-box model based on some approximation tools such as interpolated look-up tables, splines, neural networks, fuzzy systems, etc. If the aim of modeling is only to obtain an accurate predictor for y, there is not much difference between these models, as they all can approximate smooth nonlinear systems arbitrarily well. Often, however, besides accurate predictions one wants to obtain a model that can be used to learn something about the underlying system and analyze its properties. From this point of view, fuzzy systems are more transparent than most black-box techniques.

3. Fuzzy Models

A mathematical model which in some way uses fuzzy sets is called a fuzzy model. In system identification, rule-based fuzzy models are usually applied. In these models, the relationships between variables are represented by means of if-then rules with imprecise (ambiguous) predicates, such as:

If heating intensity is high then temperature will increase fast.

This rule defines in a rather qualitative way the relationship between the heating and the temperature in a room, for instance. To make such a model operational, the meaning of the terms ‘high’ and ‘fast’ must be defined more precisely. This is done by using fuzzy sets, i.e., sets where the membership changes gradually rather than in an abrupt way. Fuzzy sets are defined through their membership functions which map the elements of the considered universe to the real unit interval [0, 1]. The extreme values 0 and 1 denote complete membership and non-membership, respectively, while a degree between 0 and 1 means partial membership in the fuzzy set. Depending on the structure of the if-then rules, two main types of fuzzy models can be distinguished: the Mamdani (or linguistic) model and the Takagi-Sugeno model. For the ease of notation, multiple-input, single-output static models are first considered. The representation of dynamics in fuzzy models is discussed afterwards.

3.1. Mamdani Model

In this model, the antecedent (if-part of the rule) and the consequent (then-part of the rule) are fuzzy propositions:

$$\mathcal{R}_i : \text{if } x \text{ is } A_i \text{ then } y \text{ is } B_i, \quad i = 1, 2, \ldots, K. \tag{5}$$

Here A_i and B_i are linguistic terms (such as ‘small’, ‘large’, etc.), represented by fuzzy sets, and K is the number of rules in the model. The linguistic fuzzy model is useful for representing qualitative knowledge such as in the following illustrative example.

Example 3.1 Consider a qualitative description of the relationship between the oxygen supply to a gas burner (x) and its heating power (y):

$\mathcal{R}_1$: If O_2 flow rate is Low then heating power is Low.

$\mathcal{R}_2$: If O_2 flow rate is OK then heating power is High.

$\mathcal{R}_3$: If O_2 flow rate is High then heating power is Low.

The meaning of the antecedent linguistic terms {Low, OK, High} and consequent linguistic terms {Low, High} is defined by membership functions such as the ones depicted in Figure 2. Membership functions can be defined by the model developer (expert) based on prior knowledge or by using data. In this example, the membership functions and their domains are selected quite arbitrarily.

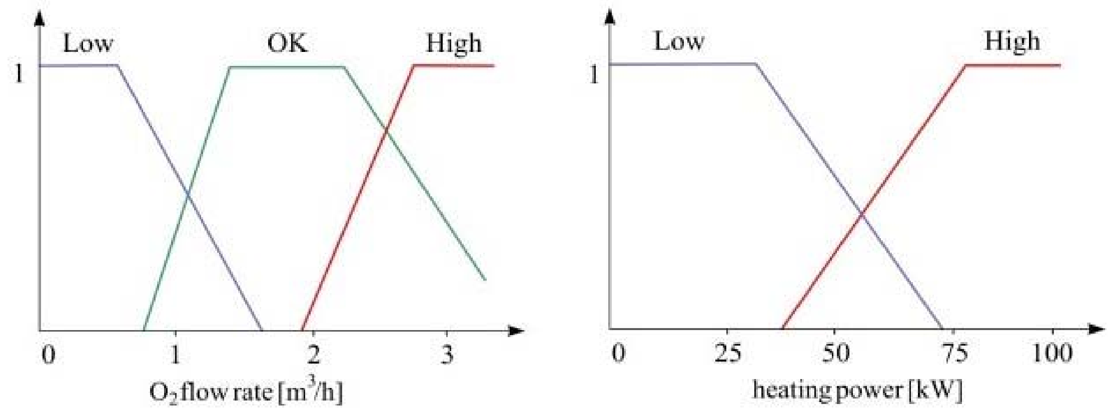

Figure 2: Membership functions for the Mamdani model in Example 3.1

The meaning of the linguistic terms is, of course, not universally given. In the above example, the definition of the fuzzy set OK, for instance, may depend on the flow-rate of the fuel gas, the type of burner, etc. When input-output data of the system under study are available, methods for constructing or adapting the membership functions from data can be applied, as discussed later on in Section 4.3. Note, however, that the qualitative relationship given by the rules is usually expected to be valid for a wide range of conditions. □

While the Mamdani model is typically used in knowledge-based (expert) systems, in data-driven identification, the following model due to Takagi and Sugeno has become popular.

3.2. Takagi-Sugeno model

The antecedent is defined in the same way as in the Mamdani model, while the consequent is a function of the input variables:

$$\mathcal{R}_i: \text{ If } \mathbf{x} \text{ is } A_i \text{ then } \mathbf{y}_i = f_i(\mathbf{x}), \quad i = 1, 2, \ldots, K. \quad (6)$$

This model combines the linguistic description with standard functional regression: the antecedents describe fuzzy regions in the input space in which the consequent functions are valid. The functions f_i are typically of the same structure, only their parameters are different in each rule. A simple and practically useful parameterization is the affine (linear in the parameters) form, yielding the rules:

$$\mathcal{R}_i: \text{ If } \mathbf{x} \text{ is } A_i \text{ then } y_i = \mathbf{a}_i^T \mathbf{x} + b_i \quad i = 1, 2, \ldots, K. \tag{7}$$

where $\mathbf{a}_i$ is the consequent parameter vector and b_i is a scalar offset. In the sequel, we will focus on this type of model which be called the affine Takagi-Sugeno (TS) model. The output y is computed by taking the weighted average of the individual rules' contributions:

$$y = \frac{\sum_{i=1}^{K} \beta_i(\mathbf{x})\, y_i}{\sum_{i=1}^{K} \beta_i(\mathbf{x})} = \frac{\sum_{i=1}^{K} \beta_i(\mathbf{x})\, (\mathbf{a}_i^T \mathbf{x} + b_i)}{\sum_{i=1}^{K} \beta_i(\mathbf{x})} \tag{8}$$

where $\beta_i(\mathbf{x})$ is the degree of fulfillment of the *i*-th rule. For the rule (7), it is equal to the membership degree $\mu_{A_i}(\mathbf{x})$, but it can also be more complicated expression, as shown later on. The antecedent fuzzy sets are usually defined to describe distinct partly overlapping regions in the input space. The parameters $\mathbf{a}_i$ are then (approximate) local linear models of the considered nonlinear system. The TS model can thus be regarded as a smooth piece-wise linear approximation of a nonlinear function or a parameter-scheduling model.

$\mathcal{R}_1$: If u is Negative then $y_1 = a_1\, u - b_1$

$\mathcal{R}_2$: If u is Zero then $y_2 = a_2\, u - b_2$

$\mathcal{R}_3$: If u is Positive then $y_3 = a_3\, u - b_3$

Example 3.2 Consider a static characteristic of an actuator with a dead-zone and a nonsymmetrical response for positive and negative inputs. Such a system can conveniently be represented by a TS model with three rules each covering a subset of the operating domain that can be approximated by a local linear model, see Figure 3. The corresponding rules are given in the right part of this figure.

$$y = \frac{\mu_{Neg}(u)y_1 + \mu_{Zero}(u)y_2 + \mu_{Pos}(u)y_3}{\mu_{Neg}(u) + \mu_{Zero}(u) + \mu_{Pos}(u)}$$

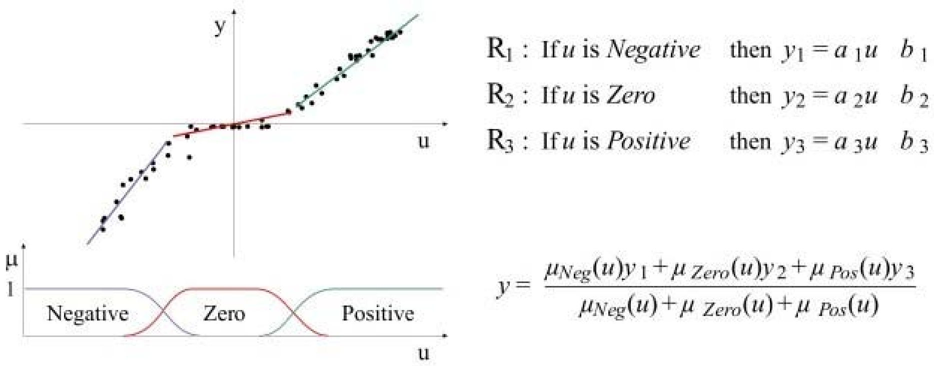

Figure 3: The Takagi-Sugeno fuzzy model as a piece-wise linear approximation of a nonlinear system.

3.3. Fuzzy Logic Operators

In (5) and (6), the antecedent fuzzy sets are defined on vector domains by their multivariate membership functions. It is usually more convenient to represent the antecedent as a combination of terms with univariate membership functions, using logic operators such as ‘and’ (conjunction), ‘or’ (disjunction) and ‘not’ (complement). In fuzzy set theory, several families of operators have been introduced for these logical connectives. Table 1 shows the two most common ones; the Zadeh min and max operators and the probabilistic operators.

	A and B	A or B	not A
Zadeh	$\min(\mu_A, \mu_B)$	$\max(\mu_A, \mu_B)$	$1-\mu_A$
probabilistic	$\mu_A \cdot \mu_B$	$\mu_A + \mu_B - \mu_A \cdot \mu_B$	$1-\mu_A$

Table 1: Commonly used functions for fuzzy logic operators.

Consider first the conjunctive form of the antecedent, which is given by:

$$\mathcal{R}_i: \text{ If } \mathbf{x} \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \text{ and } \ldots \text{and } x_p \text{ is } A_{ip} \text{ then } y_i = \mathbf{a}_i^T \mathbf{x} + b_i \qquad (9)$$

with the degree of fulfillment

$$\begin{aligned} \beta_i(\mathbf{x}) &= \min\left(\mu_{A_{i1}}(x_1),\, \mu_{A_{i2}}(x_2), \ldots,\, \mu_{A_{ip}}(x_p)\right) \text{ or} \\ \beta_i(\mathbf{x}) &= \mu_{A_{i1}}(x_1) \cdot \mu_{A_{i2}}(x_2) \cdots \mu_{A_{ip}}(x_p) \end{aligned} \qquad (10)$$

for the minimum and product conjunction operators, respectively. With this structure, the complete set of rules (9) divide the input domain into a lattice of overlapping axisparallel hyperboxes. Each of these hyperboxes is a Cartesian product intersection of the corresponding univariate fuzzy sets, see Figure 4a. In this way, multivariate membership functions are in fact created.

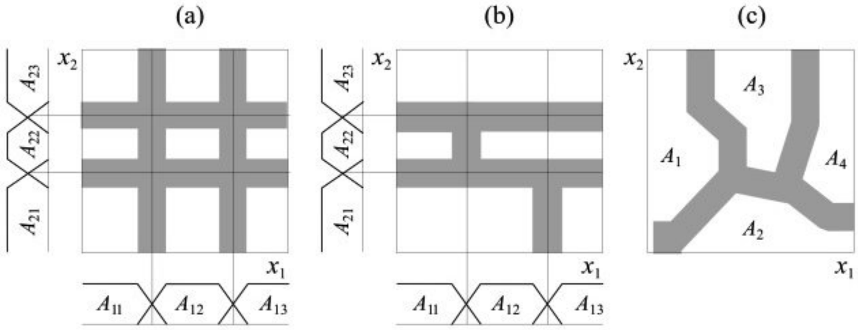

Figure 4: Different partitions of the antecedent space. The gray areas denote the overlapping regions of the fuzzy sets.

In the conjunctive form, the number of rules K, needed to cover the entire domain, is growing in an exponential manner:

$$K = \prod_{i=1}^{p} N_i, \tag{11}$$

where p is the dimension of the input space and N_i is the number of linguistic terms for the i-th antecedent variable. By combining conjunctions, disjunctions and complements, the number of rules can be reduced as illustrated in Figure 4b. Here, the rule antecedent covering, for instance, the lower left corner of the antecedent space is: If x_1 is not A_{13} and x_2 is A_{21} then … and the corresponding degree of fulfillment is computed by:

$$\beta(x_1,\ x_2) = \min(1-\mu_{A_{13}}(x_1),\ \mu_{A_{21}}(x_2)).$$

However, the boundaries between the regions still remain restricted to the rectangular grid defined by the fuzzy sets of the individual variables. The antecedent form with multivariate membership functions (6) is the most general one, as there is no restriction on the shape of the fuzzy regions.

The boundaries between these regions can be arbitrarily curved and oblique to the axes, as depicted in Figure 4c. Also the number of fuzzy sets needed to cover the antecedent space may be much smaller than in the previous cases.

Hence, for complex multivariable systems this partition may provide the most effective representation. Note that the fuzzy sets A_1 to A_4 in Figure 4c still can be projected onto x_1 and x_2 to obtain an approximate linguistic interpretation of the regions described. Fuzzy clustering techniques are often used to obtain fuzzy sets of this type from data (see Section 4.3).

-
-
-

TO ACCESS ALL THE **26 PAGES** OF THIS CHAPTER,
Click here

Bibliography

Abonyi J., Babuška R., Verbruggen H., Szeifert F. (2000). Incorporating prior knowledge in fuzzy model identification. *International Journal of Systems Science* **31**(5), 657-667. [This is a research article on constrained identification of fuzzy models. The constraints are derived from prior knowledge about the plant characteristics.].

Babuška R. (1998). *Fuzzy Modeling for Control*. Boston, USA: Kluwer Academic Publishers. [This is a research monograph on fuzzy modeling and identification, mainly covering clustering-based approaches. Model-based fuzzy control design is addressed as well.].

Bezdek J. (1981). *Pattern Recognition with Fuzzy Objective Function*. Plenum Press, New York.[This is a classical research monograph on fuzzy clustering, explaining the basic concepts as well more advanced issues and various clustering algorithms.].

Bezdek J., Pal S., eds. (1992). *Fuzzy Models for Pattern Recognition*. New York: IEEE Press. [This edited book is a collection of articles on fuzzy modeling and clustering technqiues applied to pattern recognition.].

Brown M., Harris C. (1994). *Neurofuzzy Adaptive Modelling and Control*. New York: Prentice Hall. [This book presents the basic principles and algorithms for neuro-fuzzy modeling.].

Duda R., Hart P. (1973). Pattern Classification and Scene Analysis. New York: John Wiley & Sons. [This is a classical book on pattern recognition and clustering.].

Gustafson D., Kessel W. (1979). Fuzzy clustering with a fuzzy covariance matrix. In Proc. IEEE CDC, pp. 761-766, San Diego, CA, USA. [In this paper, the so-called 'Gustafson-Kessel' algorithm was introduced.].

Hellendoorn H., Driankov D., eds. (1997). *Fuzzy Model Identification: Selected Approaches*. Berlin, Germany: Springer. [This is a collection of articles in fuzzy identification. A large part of the book is devoted to clustering-based approaches.].

Jain A., Dubes R. (1988). *Algorithms for Clustering Data*. Englewood Cliffs: Prentice Hall. [This is a classical book on clustering techniques.].

Jang J.S. (1993). ANFIS: Adaptive-network-based fuzzy inference systems. IEEE Transactions on Systems, Man & Cybernetics **23**(3), 665-685. [One of the first articles on neuro-fuzzy systems and learning.].

Klir G., Yuan B. (1995). *Fuzzy sets and fuzzy logic; theory and applications*. Prentice Hall. [A standard, detailed textbook on fuzzy set theory and fuzzy logic.].

Kortmann P., Unbehauen H. (1997). Structure identification of functional-type fuzzy models with application to modeling nonlinear dynamic plants. In Proceedings 5th Fuzzy Days. [A paper on structure identification in fuzzy models.].

Ljung L. (1987). *System Identification, Theory for the User*. New Jersey: Prentice-Hall. [This is a classical monograph on linear systems identification.].

Mamdani E. (1977). Application of fuzzy logic to approximate reasoning using linguistic systems. *Fuzzy Sets and Systems* **26**, 1182-1191. [In this article, the so-called 'linguistic' or 'Mamdani' fuzzy model was first introduced.].

Murray-Smith R., Johansen T.A., eds. (1997). *Multiple Model Approaches to Nonlinear Modeling and Control*. London, UK: Taylor & Francis. [A collection of articles on various multiple-model modeling and control methods.].

Nakamori Y., Ryoke M. (1994). Identification of fuzzy prediction models through hyperellipsoidal clustering. *IEEE Trans. Systems, Man and Cybernetics* **24**(8), 1153-73. [One of the first articles introducing the use of fuzzy clustering techniques in the identification of fuzzy systems.].

Nelles O. (2000). *Nonlinear System Identification*. New York: Springer. [This book contains an extensive treatment of various nonlinear system identification techniques, including fuzzy models and neural networks.].

Pedrycz W., ed. (1996). *Fuzzy Modelling: Paradigms and Practice*. Boston: Kluwer Academic Publishers. [This is a standard research monograh on fuzzy modeling.].

Sousa J., Babuška R., Verbruggen H. (1997). Fuzzy predictive control applied to an air-conditioning system. *Control Engineering Practice* **5**(10), 1395-1406. [This article describes a real-world application of fuzzy identification.].

Sugeno M., Kang G. (1986). Fuzzy modelling and control of multilayer incinerator. *Fuzzy Sets and Systems* **18**, 329. [This article describes a practical application of fuzzy identification.].

Sugeno M., Kang G. (1988). Structure identification of fuzzy model. *Fuzzy Sets and Systems* **28**, 15-33. [One of the first articles on structure identification in fuzzy systems.].

Sugeno M., Tanaka K. (1991). Successive identification of a fuzzy model and its application to prediction of a complex system. *Fuzzy Sets and Systems* **42**, 315-334. [An identification method based on successive partitioning of the antecedent space in a fuzzy model is proposed in this article. Application examples are given.].

Sugeno M., Yasukawa T. (1993). A fuzzy-logic-based approach to qualitative modeling. *IEEE Trans.Fuzzy Systems* **1**, 7-31. [An extensive article on Mamdani-type fuzzy modeling with several application examples.].

Takagi T., Sugeno M. (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics* **15**(1), 116-132. [In this article, the so-called 'Takagi-Sugeno' fuzzy model was first introduced. Application examples of fuzzy modeling are given.].

Zadeh L. (1965). *Fuzzy sets. Information and Control* **8**, 338-353. [This is the seminal paper on fuzzy sets.].

Zimmermann H.J. (1996). *Fuzzy Set Theory and its Application*. Boston: Kluwer Academic Publishers, third edn. [This is a standard book on fuzzy set theory.].

Biographical Sketch

Robert Babuška received the M.Sc. degree in control engineering from the Czech Technical University in Prague, in 1990, and the Ph.D. degree from the Delft University of Technology, the Netherlands, in 1997. Currently, he is a Professor at the Department of Control Systems Engineering, Faculty of Information Technology and Systems, Delft University of Technology. He is serving as an associate editor of the IEEE Transactions on Fuzzy Systems, Engineering Applications of Artificial Intelligence, and as an area editor of Fuzzy Sets and Systems. Dr. Babuška has co-authored more than 30 journal papers and chapters in books and has published a research monograph *Fuzzy Modeling for Control* (Kluwer Academic Publishers, Boston, 1998). His research interests include the use of fuzzy set techniques and neural networks in nonlinear system identification and control with applications in aerospace systems and process industry.