DATA–BASED FUZZY MODELING

T. Slawinski and H. Kiendl
Department of Control Engineering, University of Dortmund, Germany

Keywords: Modeling, approximation, classification, fuzzy logic, control, data–based model, rule based model, Mamdani type fuzzy models, Takagi–Sugeno-Kang type fuzzy models, clustering, neural networks, fuzzy–neuro, decision trees, evolutionary algorithms

Contents

1. Introduction
2. Process of Data-Based Modeling
3. Concepts for Fuzzy Modeling
4. Established Methods
4.1. Cluster-Based Fuzzy Modeling
4.2. Neuro–Fuzzy Models
4.3. Tree-Oriented Approach
4.4. Rule-Oriented Approach
4.5. Fuzzy Modeling based on Evolutionary Algorithms
4.5.1. Pittsburgh Style
4.5.2. Michigan Style
5. Conclusion and Perspectives.
Glossary
Bibliography
Biographical Sketches

Summary

The use of fuzzy logic for the modeling of processes (such as a technical process or the behavior of a human operator in a process) is often motivated by the interpretability of the resulting fuzzy system/model. In fuzzy models, the dependencies of the variables of the considered process are described by qualitative (linguistic) *if–then* rules, which correspond to the way in which human knowledge is usually presented. The main advantages of interpretability are a higher acceptance by the users of the fuzzy model, and the increased possibility for tuning and adapting the fuzzy model by hand.

Most of the early fuzzy models were designed by the extraction of knowledge from experts via interviews. However, for more complex processes, or if no prior knowledge is available, such a knowledge-based modeling approach may be very time consuming or even impossible. In these cases, a data-based approach that automatically generates a fuzzy model by the evaluation of recorded process data is more promising. Currently, many methods for data-based fuzzy modeling are available, and their applicability and use has been established in numerous benchmark problems and real-world applications.

This chapter aims at giving a survey of the *state-of-the-art* fuzzy modeling techniques. At the beginning the process of data-based modeling is described in general (Section 2).

Based on this, different established concepts for data-based fuzzy modeling are discussed (Section 3). Then, some chosen illustrative methods are presented (Section 4). Finally, a conclusion is drawn and different perspectives are discussed.

1. Introduction

Models are extremely useful for understanding and purposefully controlling any 'process' such as a technical process or the behavior of a human process operator. Examples of the broad application spectrum of models are the analysis, optimization, simulation, control, and design of processes. Accordingly, modeling is a key objective, especially in the fields of natural science and engineering.

An increasing complexity and the linking of technical processes have led to a growing requirement for sophisticated modeling methods. Robust modeling methods that allow the processing of qualitative knowledge, or that deal with incomplete/uncertain information and the model adaptation to time variable environments, are especially wanted.

Furthermore, it is required that the modeling method be capable of being applied within an acceptable computing time, even for high-dimensional problems. From a consideration of these modeling objectives, the methods of *Computational Intelligence* (CI), such as *Fuzzy Control* (FC), artificial *Neural Networks* (NN) and of *Evolutionary Algorithms* (EA), provide promising modeling approaches. These techniques are inspired by nature, in particular by common-sense reasoning, by the neural structure of the brain, and by the evolution of species.

Regardless of the applied method, the modeling process can be divided into two main parts, the *identification* of the *structure* and the *optimization* of the free *parameters*. The modeling process must consider the modeling objective as well as the available information. An important distinguishing point is whether the objective is a *qualitative* or a *quantitative* model.

In qualitative models, the dependencies of the process variables — such as input, output, and internal variables — are expressed in an understandable manner, so that the model is interpretable. While interpretable equations are used to describe the dependencies in mathematical/physical modeling, fuzzy modeling is based on qualitative *if–then* rules. In these types of modeling, the identification of the structure corresponds to setting up equations and rules, respectively. Such a qualitative model can be converted to a quantitative model by identifying the free parameters — the physical parameters and the linguistic values, respectively — so that the interpretability usually is preserved.

The main aim of quantitative modeling is to determine the output values of the considered process for given input values. If nothing other than a quantitative model is desired, this can be obtained without first setting up a qualitative model. For example, for *approximation* purposes, the free parameters of general functions can be identified by means of regression algorithms. As well, NNs are often used for the approximation of a given input–output behavior. An essential advantage of these approaches is that they require no prior process knowledge. Rather, a suitable type of approximation function of the NN must be chosen as the structure for the model. Then, established methods can be applied for the identification of the free parameters (such as the parameters of the function, or the weights of the NN). However, the resulting quantitative model is hard to interpret and allows no insight into the process being considered. This impairs the acceptance of such pure qualitative models, especially in industrial practice.

For these reasons, it is important that there are data-based fuzzy modeling approaches which allow the generation of interpretable quantitative models, where the dependencies of the process variables are expressed by intuitively understandable *if–then* rules. Historically, the first fuzzy models were designed by knowledge acquisition from experts. However, interviewing an expert is often time-consuming and sometimes of questionable value, because access to the expert's knowledge can only be gained if the expert is conscious of the knowledge and is willing to surrender it. This applies especially for complex processes, and highlights the importance of data-based approaches for the generation of fuzzy models.

In the following, state-of-the-art methods are described, in particular established methods for the data-based generation of fuzzy models. The recent research activities in this field are mainly aimed at improving the accuracy of the resulting model and providing efficient generation methods that allow us to cope with high-dimensional tasks. Moreover, increasing attention is being paid to finding a good balance between the accuracy and the interpretability of the resulting fuzzy model. Furthermore, the present lack of systematic strategies for the choice and application of data-based fuzzy modeling approaches is briefly discussed.

2. Process of Data-Based Modeling

In this section, data-based modeling is used for the generation of interpretable quantitative fuzzy models. Quantitative modeling means to set up a model that has the same input–output behavior as the process to be modeled. Let $x_1,\ldots,x_n$ be the input variables of the process and y its output variables (for simplicity, we consider only one output variable; the case of more than one output variable can be handled analogously). We want to set up a model that has the same input variables as the process, and an output variable $\hat{y}$ so that $\hat{y}(x_1,\ldots,x_n) \approx y(x_1,\ldots,x_n)$ is valid for all possible values of the input variables (Fig. 1). The accuracy of the obtained model is said to higher because $\hat{y}$ corresponds more closely to y.

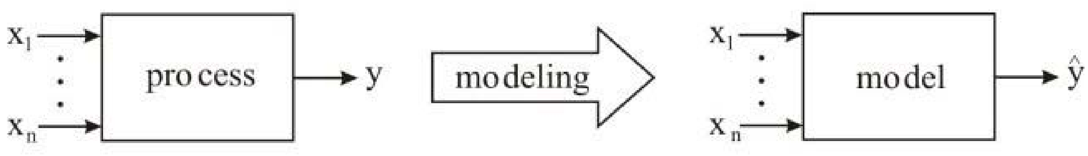

Figure 1. Task of quantitative modeling

As shown in numerous publications, data-based modeling has been successfully applied to a wide spectrum of problems. These can be classified as *approximation/regression* and *classification* problems. Approximation problems deal with a real valued process output variable y, e.g., the temperature of a chemical reactor, which may take any real value a within a certain range $a_{min} \leq a \leq a_{max}$. For the evaluation of the accuracy of the resulting model the mean absolute/square error is widely used.

Classification problems deal with process output variables y, which have a discrete value range $(a_1, a_2, \ldots, a_r)$. Examples of such output values — called classes — are integers and types of pathogenic agents. The modeling objective may be to obtain one model output value $\hat{a}$ that coincides for all possible inputs with the process output (hard classification). Alternatively, the aim may be to produce a model that supplies for each of the possible process output values $a_1, a_2, \ldots, a_r$ corresponding membership degrees $\mu_1, \mu_2, \ldots, \mu_r$ with $0 \leq \mu_r \leq 1$. These express how well each of the possible output values a_i is supported by the considered input values (soft classification). Fuzzy modeling can meet both objectives in a similar manner by using crisp or overlapping membership functions. Furthermore, some of the fuzzy modeling approaches can also deal with mixed types of input variables, e.g., continuous variables and hard classes. This capability is required for the solution of various real-world tasks, as for example, the characterization of client profiles by continuous variables such as age, and hard (discrete) classes such as sex. The accuracy of a model that is designed for classification purposes — also called classifier — is often evaluated by the relative number of classification errors.

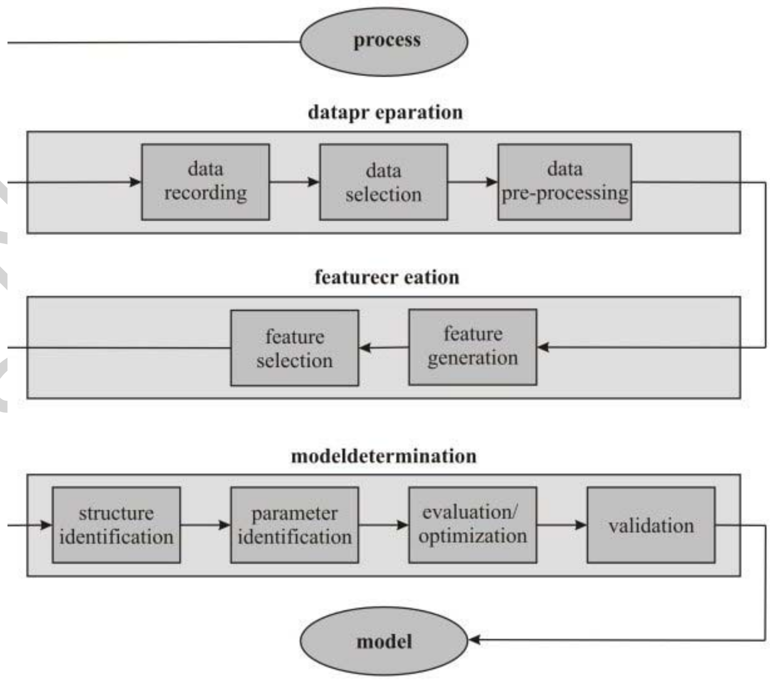

Figure 2. Procedure of data-based modeling

Independent of the type of modeling problem at hand, the process of data-based modeling can be divided into three main steps: data preparation, feature creation, and model determination. In the following, these steps are briefly described by means of Fig. 2.

Data Preparation: The data sets used for the modeling process are usually the result of a data recording. Depending on the available database it may be necessary to perform (additional) measurements or observations as the process is modeled. A subsequent data selection may be useful for the elimination of erroneous, inconsistent, or uncertain data points. Furthermore, an unbalanced data distribution can be avoided by using data selection. The final step of the data preparation is data pre-processing, where (standard) algorithms for data processing, such as normalization, smoothing, or numerical integration/derivation are applied.

Feature Creation: Depending on the modeling task at hand, the data points resulting from the data preparation can be used either directly for the subsequent modeling, or for creating aggregated features used as inputs for the subsequent modeling. The latter is, for example, often necessary for time series, such as stock exchange data, where features, such as trends, are generated based on the raw data points. Often it is not clear a priori which features are favorable for the intended modeling objective. Therefore, an automatic data-based feature generation is usually performed in two steps. In the first step, a number of potentially relevant features are generated. In a second step, from the set of all generated features, a smaller set of relevant and non-redundant features is selected.

Widely applied approaches for testing the features are to investigate, by measures such as correlation or information entropy, whether there is strong dependence between the feature and the output value. Furthermore, analysis with the same measures of whether two features are redundant is possible. In more advanced approaches, individual features are no longer tested and selected step by step. Instead of this, all (or all promising) subsets of the generated features are evaluated, each as a whole. Favorable feature selection usually considerably reduces the expenditure for the subsequent model determination, because the corresponding search space is smaller if fewer input variables have to be considered. Furthermore, the resulting model quality is often improved if features of little relevance are omitted.

Model Determination: The first step in the model determination is to choose/identify an adequate structure, e.g., a type of general function, a type of NN, or a type of fuzzy system for the model. Based on this chosen structure, the remaining free parameters, e.g., the constants of the functions, the weights of the NN, or the rules of the fuzzy system are identified based on the prepared data and created features. Then the model obtained is evaluated, and, if necessary, optimized. In order to asses the generalizing capability of the model and to avoid over-fitting, respectively, in a final validation step the determined model is tested on new data, i.e. data which were not used in the modeling process. In practice, these three main steps are strongly connected. Therefore, a successive work flow in the modeling process will not always lead to the best results in a single sequence. For complex problems in particular, an iterative approach is more promising. The following sections focus on the third step, the determination of the (fuzzy) model, with emphasis on the identification/generation of the fuzzy rules.

-
-
-

Bibliography

A. Krone, and H. Taeger. (2001). Data–based fuzzy rule test for fuzzy modelling. *Fuzzy Sets and Systems*, 126: pages 343–358. [In this article statistical tests are extended and applied to data–based fuzzy rules.]

A.K. Jain, and R.C. Dubes. (1988). *Algorithms for Clustering Data*. Prentice Hall, Englewood Cliffs, New Jersey, USA. [A detailed introduction in clustering techniques is given in this book.]

H. P. Schwefel, I. Wegener, and K. Weinert, editors. (2002). *Advances in Computational Intelligence*. Springer Verlag, Berlin. [This book presents state of the art and applications of computational intelligence techniques.]

H.-P. Schwefel. (1995). *Evolution and Optimum Seeking*. Sixth–Generation Computer Technology Series, New York, USA. [In this book evolutionary algorithms are introduced and applied to optimization tasks.]

J. Kacprzyk, and L. Zadeh, editors. (1999). *Computing with Words in Information/Intelligent Systems*. Physica-Verlag, Heidelberg. [This book describes recent developments and advances in fuzzy systems.]

J. R. Quinlan. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufman, San Mateo. [A detailed introduction in decision trees is given in this book.]

J.C. Bezdek, and S.K. Paul, editors. (1992). *Fuzzy Models for Pattern Recognition*. IEEE Press, Piscataway, NJ. [In this collection many popular fuzzy modeling approaches are presented, applied and compared.]

J.C. Bezdek. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, USA. [This book describes the application of fuzzy techniques to pattern recognition tasks.]

M. Last, A. Kandel, and H. Bunke, editors. (2001). *Data Mining and Computational Intelligence*. Physica-Verlag, Heidelberg. [This book gives a survey of state of the art data-mining techniques based on computational intelligence.]

T. Takagi, and M. Sugeno. (1985).Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1): pages 116–132. [A detailed introduction in TSK-fuzzy models is provided in this article.]

Biographical Sketches

Timo Slawinski was born in 1970 in Hannover. He studied physics and computer science at the University of Hannover, Germany. In 1997 he was awarded Dipl.-Phys. for his diploma thesis: “Experimental Design and Numerical Simulation for the Storage of Atomic Ensembles in Dipole Traps “in the domain of experimental physics in the special subject quantum optics.

He then turned to the field of electrical engineering and information technology, working as scientific assistant at the Chair of Control Engineering at the University of Dortmund, Germany, from 1997 to 2001. Thereby his research activities are focused on data-mining, data-based fuzzy modeling, data-preprocessing, CI-Techniques (fuzzy logic, evolutionary algorithms, and neural networks), hybrid/adaptive approaches and data-based design of process control/monitoring.

In 2001 he was awarded doctorate degree for his thesis: “Analysis and efficient Generation of Relevant Fuzzy Rules in High-Dimensional Search Spaces“. Thereafter he started working in industry at the Bayer AG, Germany. Since 2002 he is project manager for manufacturing execution systems for the integration of the enterprise resource planning system (SAP R/3) with the subsidiary process control systems in the Process Management Technology Department of the Bayer Technology Services.

Harro Kiendl was born in 1936 in Hamburg. He studied physics and mathematics at the University of Hamburg, and received doctorate degree in 1966 for his thesis that explored the domain of experimental solid state physics. From 1966 to 1967 he was leader of the project "cybernetic models and system analysis" at the Pedagogical Center of Berlin. He then turned to the field of control sciences, working at the Technical University of Berlin from 1967 to 1969, and at the Ruhr-University of Bochum from 1969 to 1973 where he qualified as a university lecturer in 1971. In 1973 he was appointed to the Chair of Control Sciences in the Faculty of Electrical Engineering at the University of Dortmund. His research interests are method-oriented and focus on the fields of classical control (nonlinear control, robustness and stability analysis, and modeling) and computational intelligence (fuzzy control, data-based modeling, and evolutionary optimization). Many of the methods he developed are used in industrial applications and commercial software tools. He obtained 21 patents, such as for hyperinference, which processes both positive and negative fuzzy rules, as well as for the inference filter, for torque defuzzification, and for implicit modeling.