

On Data-driven Takagi-Sugeno Modeling of Heterogeneous Systems with Multidimensional Membership Functions

A. Kroll*. S. Soldan**

*University of Kassel, Faculty of Mechanical Engineering, Department of
Measurement and Control, Mönchebergstraße 7, 34125 Kassel, Germany
(e-mail: andreas.kroll@mrt.uni-kassel.de)

** same Department (e-mail: samuel.soldan@mrt.uni-kassel.de)

Abstract: In case of a structural mismatch between a Takagi-Sugeno model and a nonlinear system that is to be modeled, identification algorithms tend to compensate this by a finer granular partitioning. This is in conflict with the objective of parsimonious models. A novel TS model type with multi-dimensional fuzzy sets and locally adjustable/heterogeneous fuzziness is presented with a method for the adjustment. This new approach enhances the structural flexibility of the model while the number of parameters increases negligibly. Two case studies, including compressor modeling, illustrate the performance of the proposed method.

Keywords: system identification, nonlinear modeling, multidim. membership functions, compressor

1. INTRODUCTION

The availability of models is the prerequisite for model-based analysis and design methods. Rigorous modeling has many advantages but is costly and sometimes the required knowledge is not available. An alternative is data driven modeling. High approximation capability and ease of use have made Takagi-Sugeno (TS) and piecewise-affine (PWA) models popular for modeling nonlinear systems. Both result from a weighted superposition of local models that can typically be interpreted as linearised systems. This permits the transfer of methods from linear systems theory, making analysis, interpretation and control design significantly easier.

Local models in PWA models have hard boundaries and no interpolation takes place. Various identification concepts have been proposed (e.g. (Paoletti, et al., 2007; Roll, et al., 2004; Ferrari-Trecante, et al., 2003)). PWA models fit well applied for switching and PWA target systems. However, in case of systems with smooth nonlinearities, approximation quality is gained by a finer granular partitioning in the respective regions which inflates the number of model parameters. On the contrary, TS models are typically used to model systems that feature smooth nonlinearities (see e.g. (Babuska, 1998; Nelles, 2001; Ruano, 2005) and references therein). However, systems may feature a heterogeneously shaped nonlinearity: Fig. 1 shows a compressor map with sharp edges and different soft curvatures. Such characteristics are difficult to model using PWA models.

In case of data-driven identification, a parsimonious model representation is advisable: The variance of a parameter estimate increases with the relation of the number of parameters to identification data sets. In practical cases, the number of identification data sets is limited such that the parametric uncertainty of the model increases with the number of model parameters. Therefore it is important that the model structure

fits well to the modeling problem. This article builds on TS models with multidimensional membership functions (MF) that are identified by fuzzy clustering and least squares parameter estimation (Kroll, 1996) due to their structural flexibility. In order to archive high prediction quality with parsimonious parameterization in case of modeling systems with heterogeneous nonlinearities, the common uniformly chosen partition fuzziness (fuzziness parameter weighting exponent m) is replaced by a locally adjustable one.

The idea is that the increased flexibility in shaping transitions permits to structurally better approach approximation problems and therefore reduces the required number of partitions.

The next section briefly reviews TS modeling and identification schemes. Section 3 introduces TS models with locally adapted fuzziness and a method to identify them. The performance is demonstrated by approximating a compressor map in section 4. Section 5 summarizes the results and provides an outlook on further research.

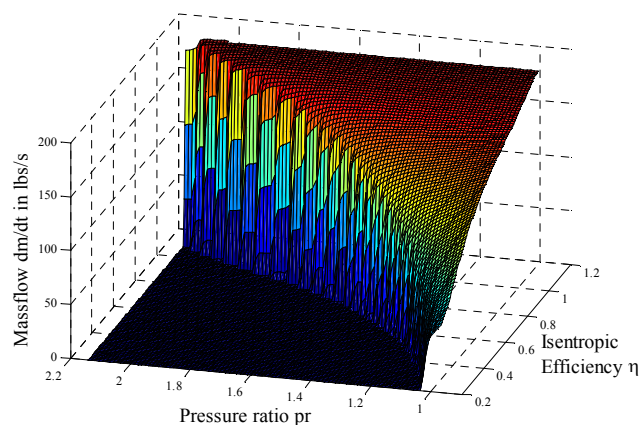


Fig. 1. Characteristic map of a single-stage axial compressor Nasa CR-72694 (based on data from (Flynn, et al., 1970)).

2. TAKAGI-SUGENO MODEL DESCRIPTION AND IDENTIFICATION

2.1 Model description

Takagi-Sugeno fuzzy models (Takagi & Sugeno, 1985) constitute a rule-based system description. A TS model is composed of c fuzzy rules. The antecedent of a rule defines a fuzzy partition in which the local model in the rule's conclusion is valid. The i -th multi-input multi-output rule of a TS model is given as:

IF $\mathbf{X}$ IS $\mathbf{A}_i$ THEN $\mathbf{y}_i = \mathbf{f}_i(\mathbf{x}), i = 1, \dots, c$.

with $\mathbf{X}$ the linguistic and $\mathbf{x} \in \mathcal{R}^n$ the crisp input variable, $\mathbf{A}_i$ the antecedent linguistic term, $\mathbf{f}_i$ a crisp vector-valued function that specifies the local model, and $\mathbf{y}_i \in \mathcal{R}^p$ the crisp output variable. In the following, a scalar function f_i will be used for the ease of notation, i.e. MISO models will be considered. The fuzzy set $\mathbf{A}_i$ is defined by the corresponding membership function (MF) $\mu_i(\mathbf{x}) := \mu_{\mathbf{A}_i}(\mathbf{x}) : \mathcal{R}^n \rightarrow [0;1]$. The model transfer characteristic results from a weighted superposition of all c local models

$$\mathbf{y}(\mathbf{x}) = \frac{\sum_{i=1}^c \mu_i(\mathbf{x}) \cdot \mathbf{y}_i(\mathbf{x})}{\sum_{i=1}^c \mu_i(\mathbf{x})} =: \sum_{i=1}^c \phi_i(\mathbf{x}) \cdot \mathbf{y}_i(\mathbf{x}) \quad (1)$$

where ϕ_i is the fuzzy basis function (FBF). For orthogonal MF, $\phi_i \equiv \mu_i$ holds.

Often, scalar (e.g. trapezoidal or Gaussian) MFs are used. Then the antecedent is written in conjunctive form which is in general implemented by the product operator

$$\mu_i(\mathbf{x}) = \prod_{j=1}^n \mu_{i,j}(x_j). \quad (2)$$

Using scalar MF provides for axis-parallel partitions that ease interpretation. On the contrary, they inflate the number of local models if nonlinearities are not strictly directed along the coordinate axes. Therefore, multi-dimensional MFs

$$\mu_i(\mathbf{x}) = \left[\sum_{j=1}^c \left(\frac{\|\mathbf{x} - \mathbf{v}_i\|_{\mathbf{A}_i}^2}{\|\mathbf{x} - \mathbf{v}_j\|_{\mathbf{A}_j}^2} \right)^{\frac{1}{m-1}} \right]^{-1} \quad (3)$$

with prototypic points (partition centers) $\mathbf{v}$ as used in (probabilistic) fuzzy clustering with $\mathbf{x}, \mathbf{v}_i, \mathbf{v}_j \in \mathcal{R}^n, c \in \mathbb{N}_+$ and the

fuzziness parameter $m \in \mathcal{R}^{>1}$ will be preferred in the following. In contrast to multi-dimensional Gaussian MFs or MFs from possibilistic clustering, these are orthogonal. This avoids unintuitive effects due to the normalization in (1). $\|\cdot\|$ is a vector norm such as an inner product norm

$$\|\mathbf{x} - \mathbf{v}_r\|_{\mathbf{A}_r} = \sqrt{(\mathbf{x} - \mathbf{v}_r)^T \cdot \mathbf{A}_r \cdot (\mathbf{x} - \mathbf{v}_r)} \quad (4)$$

with pos. definite matrix $\mathbf{A}_r = [a_{h,l}^r] \in \mathcal{R}^{n \times n}$ or an L_p -norm.

The f_i are often chosen as affine functions

$$\mathbf{y}_i(\mathbf{x}) = [a_{1,i}; \dots; a_{n,i}; b_i] [\mathbf{x}^T; 1]^T = \boldsymbol{\Theta}_i^T \tilde{\mathbf{x}}, \quad (5)$$

as linear functions ($b_i = 0$) severely limit the approximation capability: each linear model has to pass through the origin of the coordinate system.

2.2 Model identification

Model building includes partition and conclusion identification. These tasks can be carried out successively, simultaneously or alternating (iteratively) with the first approach being most common in TS modeling, see e.g. (Babuska, 1998; Nelles, 2001; Ruano, 2005) and references therein. In the related area of PWA model identification, simultaneous identification received interest after suggesting to use MILP/MIQP solution methods, see e.g. (Roll, et al., 2005). However, several local minima exist as the resulting optimization problem is highly nonconvex. In this paper, a multi-stage algorithm is used, which determines an initial model by clustering and least-squares estimation of the conclusion parameters. This is followed by an optimization stage.

Partition identification can be achieved using fuzzy clustering, e.g. the Fuzzy-c-Means (FCM), Gustafson-Kessel (FGK) or Gath-Geva algorithm (Bezdek, 1981; Babuska, 1998; Höppner, 1999). These minimize the cost function

$$J_{CL} = \sum_{k=1}^N \sum_{i=1}^c (\mu_{i,k})^m \|\mathbf{x}_k - \mathbf{v}_i\|_{\mathbf{A}_i}^2 \quad (6)$$

with N being the number of data sets. The cluster algorithms are typically applied to the input (X) or product data space ($X \times Y$). MFs (3) as provided by FCM or FGK clustering can directly be used in the rule antecedents, with product space clustering requiring reduction to the input space. Alternatively, MFs (3) can be projected to the coordinate axis and be approximated by scalar MFs. Once the structure of f_i is selected, parameters $\boldsymbol{\Theta}_i$ of the local models are typically determined from minimizing a quadratic cost function. The $\boldsymbol{\Theta}_i, i = 1, \dots, c$, of the local models can be identified independently from each other by minimizing the deviation of the local output (5) from the reference $\hat{\boldsymbol{\Theta}}_i : \arg \min_{\boldsymbol{\Theta}_i}$

$\sum_{k=1}^N w_i(k) \cdot (y(k) - \hat{y}_i(k, \boldsymbol{\Theta}_i))^2$ ('local estimation', LE) with a weighting factor $w_i(k)$. Alternatively, all c vectors $\boldsymbol{\Theta}_i$ can be computed simultaneously by optimizing the TS model output $\hat{\boldsymbol{\Theta}} : \arg \min_{\boldsymbol{\Theta}} \sum_{k=1}^N (y(k) - \hat{y}(k, \boldsymbol{\Theta}))^2; \hat{\boldsymbol{\Theta}} := [\hat{\boldsymbol{\Theta}}_1; \dots; \hat{\boldsymbol{\Theta}}_c]$ ('global estimation', GE). LE provides for local models that can be interpreted as local linearization of the target system. Due to the isolated estimation of local models, LE does not assess the transitional behavior. GE optimizes the overall model prediction quality including the transitions between models. Therefore, the overall model prediction quality is in general higher than for LE (if no counter-effects due to overfitting occur). On the contrary, GE can produce local models that cannot be interpreted as local linearizations. See (Zimmerschied & Isermann, 2008) for further discussions.

The ordinary weighted least squares (WLS) estimator can be used if the problem is linear in the parameters. E.g. for LE the parameters are computed as

$$\hat{\Theta}_i = [\Phi^T \mathbf{W}_i \Phi]^{-1} \Phi^T \mathbf{W}_i \mathbf{Y} \quad (7)$$

with $\Phi = [\tilde{\mathbf{x}}_k^T]$ the regression matrix, $\mathbf{W}_i = \text{diag}[w_{i,k}]$ the diagonal weighting matrix and $\mathbf{Y} = [y_k]$ the vector of the output observations. The memberships assigned by fuzzy clustering can be used as weights in (7). Alternatively weighting strategies include determining α -level sets of the MF and selecting the set members (crisply or fuzzily). Partition and conclusion function parameters can also be derived simultaneously from the result of a FGK clustering in the product space (Babuska, 1998).

Different criteria can be used to assess a model (Ljung, 1999; Nelles, 2001). These include averaging ones as the sum of the root mean square error (RMSE)

$$J_{\text{RMSE}} := \frac{1}{N} \sqrt{\sum_{k=1}^N (y(k) - \hat{y}(k))^2} \quad (8)$$

and local ones such as the maximum norm of the residuals.

3. MODELS WITH HETEROGENOUS FUZZINESS

3.1 Methods to locally adjust the fuzziness parameter m

The fuzziness parameter m is the key parameter to adjust the shape of the transitions between local models of TS fuzzy systems given by (1), (3), (5), see Fig. 2. A local adjustment of m can be achieved by assigning a local fuzziness parameter $m_i \in \mathbb{R}^{>1}$ to each partition yielding MFs

$$\mu_i(\mathbf{x}) = \left[\sum_{j=1}^c \left(\frac{\|\mathbf{x} - \mathbf{v}_i\|_{A_i}^2}{\|\mathbf{x} - \mathbf{v}_j\|_{A_j}^2} \right)^{\frac{1}{m_i-1}} \right]^{-1}. \quad (9)$$

This yields non-orthogonal MF. The transformation to FBF in (1) may provide for unintuitive results: a crisper FBF may result from a fuzzier MF and vice versa. Secondly, this approach does not permit the assignment of a degree of fuzziness to the transitional regimes, but only to the partitions.

Alternatively, the fuzziness parameter in (3) can be adjusted by the same global function $m(\mathbf{x})$ for all μ_i :

$$\mu_i(\mathbf{x}) = \left[\sum_{j=1}^c \left(\frac{\|\mathbf{x} - \mathbf{v}_i\|_{A_i}^2}{\|\mathbf{x} - \mathbf{v}_j\|_{A_j}^2} \right)^{\frac{1}{m(\mathbf{x})-1}} \right]^{-1}. \quad (10)$$

This provides for advantageous orthogonal MFs. $m(\mathbf{x})$ can also depend just on a subset of the argument $\mathbf{x}$.

In scalar problems, m can e.g. be switched when passing the centre of a MF. As the distance from the centre is zero when the switching occurs, this scheme does not cause discontinuities in the course of μ_i that would occur if m would be changed outside the centre. Fig. 3 illustrates that this concept permits to adjust the shape of the transitions in wide ranges: Soft transitions are implemented between the first two and the last two partition centers and a sharp between partition centers 2 and 3. The peripheral areas use a small m in this example to avoid significant reactivation effects.

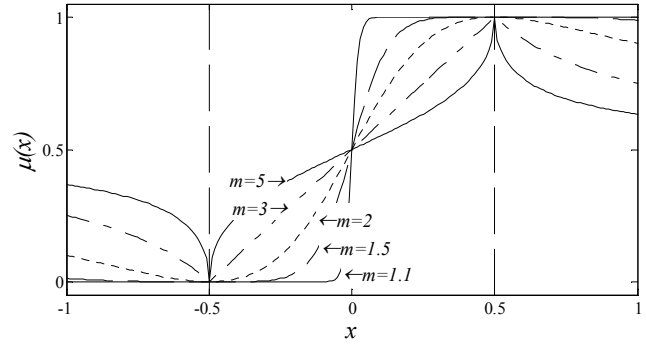


Fig. 2. Shape of MF (3) for different choices of m (example with prototypes $\mathbf{v}_i$ in -0.5 and 0.5 and Euclidean norm).

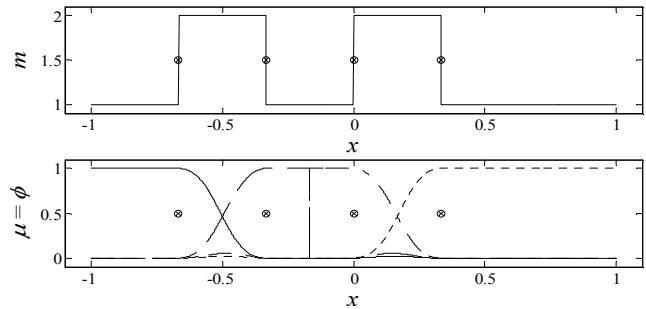


Fig. 3. 1-dim. example for interval-wise assignment of m with changes occurring in partition centers $\mathbf{v}_i$ (circles): assigned m (top) and resulting MF (bottom).

In contrast, multi-dim. MFs do not have just two tails and require a different approach. A possible method is to copy the concept behind MFs (3): (fuzziness-) m -class centers $\mathbf{v}_{i,j}$ are located in the middle of the connecting line of neighboring partition centers $\mathbf{v}_{i,j} = (\mathbf{v}_i + \mathbf{v}_j)/2$ in order to adjust m primarily in the transitional regions. A target value $m_{i,j}$ is assigned to each m -class center $\mathbf{v}_{i,j}$. The local fuzziness is computed by

$$m(\mathbf{x}) = \sum_{i,j:i \neq j} \mu_{i,j}(\mathbf{x}) \cdot m_{i,j} \quad \text{with} \quad \mu_{i,j}(\mathbf{x}) = \left[\sum_{r,s} \left(\frac{\|\mathbf{x} - \mathbf{v}_{i,j}\|_{A_i}^2}{\|\mathbf{x} - \mathbf{v}_{r,s}\|_{A_j}^2} \right)^{\frac{1}{\hat{m}-1}} \right]^{-1}. \quad (11)$$

$\hat{m}$ should be chosen large enough not to provide sharp changes and small enough to provide for extended areas of the target fuzziness for the transitional regime. Good experience was gained with $\hat{m} = 1.3$. The fuzziness values $m_{i,j}$ can e.g. be initialized using a priori knowledge or from a preceding optimization of a model with uniformly chosen fuzziness m . Alternatively, the $m_{i,j}$ can be chosen as random numbers (e.g. $m_{i,j} \in (1;3)$). Succeeding, they are numerically optimized. This method only introduces the $m_{i,j}$ as additional model parameters.

In case of more than two partitions, a m -class centre must not be placed between any two partition centers. Instead, they should only be positioned on the connecting lines between next neighboring centers. The selection of next neighbors including connection lines can be achieved by Delaunay triangulation. Different algorithms are available (Klein, 2005), while the Matlab implementation of the Quickhull algorithm was chosen for simplicity. Fig. 4 shows the result for a 2D-example with 10 partitions. A further reduction of the number of m -class centers is achieved by requiring them to have a minimal distance in between. Instead of merging centers (which would introduce additional model parameters), the connection lines that belong to a triangle featuring an interior angle below a threshold (e.g. 10°) are eliminated. Fig. 4 illustrates the operation of this reduction method.

3.2 Academic example

The proposed approach is illustrated by an artificial example with 3 partitions. The three centers are $\mathbf{v}_1^T = [-0.6; -0.75]$, $\mathbf{v}_2^T = [-0.6; 0.75]$ and $\mathbf{v}_3^T = [0.7; 0]$. The local models are $y_1 = [0; 0; 0][x_1; x_2; 1]^T$, $y_2 = [0; 0; -1][x_1; x_2; 1]^T$ and $y_3 = [0; 0; 1][x_1; x_2; 1]^T$. Two data sets with 500 points each (uniformly distributed random numbers in $x_1 \in [-1; 1]$ and $x_2 \in [-1; 1]$) were generated, one for training and another one for validation. The transition between local models 1 & 2 and 1 & 3 is crisp ($m_{1,2} = m_{1,3} = 1_+$) and soft between 2 & 3 ($m_{2,3} = 2$). The resulting m -class centers are $\mathbf{v}_{1,2}^T = [-0.6; 0]$, $\mathbf{v}_{1,3}^T = [0.05; -0.375]$ and $\mathbf{v}_{2,3}^T = [0.05; 0.375]$. The graphs of $y(\mathbf{x})$, $m(\mathbf{x})$ and the resulting MFs in Fig. 5 illustrate the heterogeneity of the modeling problem.

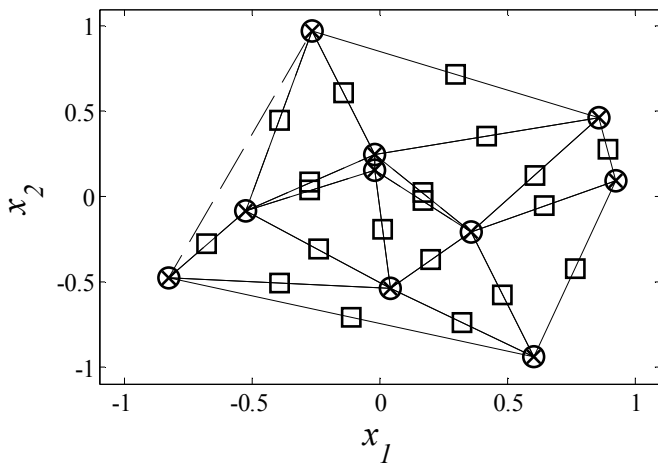


Fig. 4. Delaunay triangulation (all lines) of example with 10 partition centers (circles), eliminated connection lines (dashed), and resulting m -class centers (boxes).

Using FCM with Euclidean, $c = 3$, $m = 1.2$ and GE provides for a model with $J_{\text{RMSE}} = 0.194$ (Fig. 6). This modeling step was repeated 20 times to avoid convergence of the FCM to an unfavorable local minimum. Comparing Fig. 5 top and Fig. 6 indicates that the resulting model is moderately good.

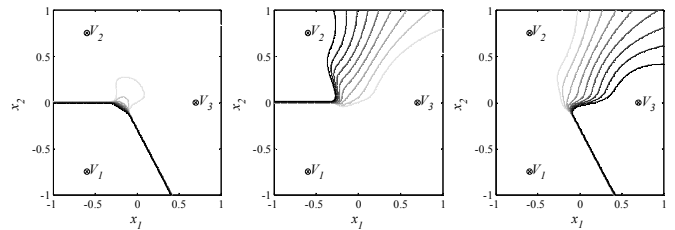
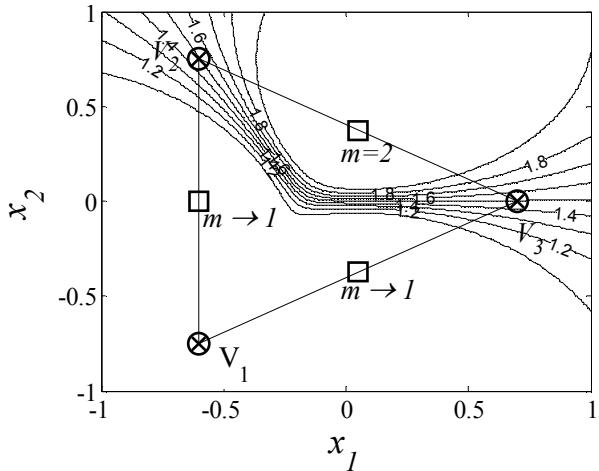
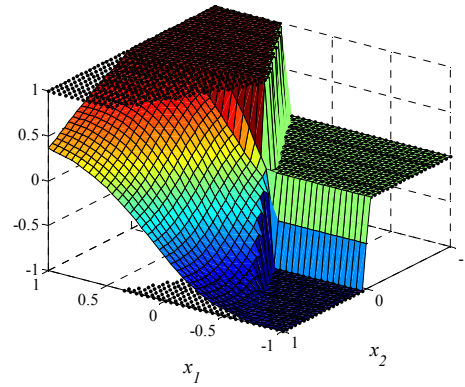


Fig. 5. Graph of y (top), level curves of m (middle) and level curves of μ_1, μ_2, μ_3 (bottom) for the academic example.

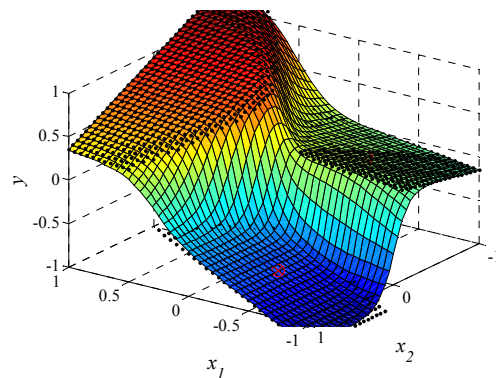


Fig. 6. Graph of y after base step (FCM and GE).

Optimization of the location of the partition centers and of the (global) fuzziness parameter yields $m = 1.29$ and provides for negligible qualitative and quantitative improvement ($J_{\text{RMSE}} = 0.156$). Permitting inhomogeneous fuzziness and optimizing the partition centers as well as the local uncertainty parameters leads to a good reproduction of the partitioning

(Fig. 7) and overall model transfer characteristics (Fig. 8) with $J_{\text{RMSE}} = 0.119$. The local reference uncertainties were identified as 1_+ , 1.03 and 1.95 which are close to the true values $(1_+; 1_+; 2)$. While the partition centers have only been identified moderately well, the resulting partitioning fits very well (Fig. 8).

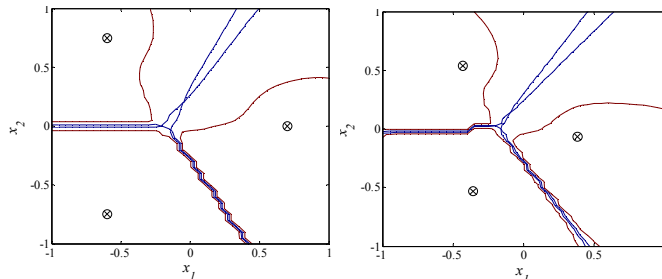


Fig. 7. Level curves of MF for original system (left) and after full identification procedure (right).

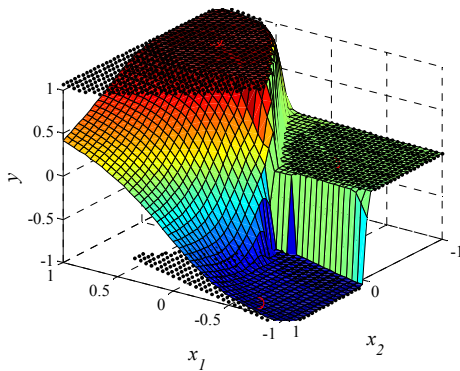


Fig. 8. Overall model characteristic after full procedure.

4. COMPRESSOR CASE STUDY

This section describes the application of the proposed methods to model the characteristic map of an axial compressor (Fig. 1). Training and test data were generated by a tool specialized for compressor modeling (Kurzke, 2009). The map belongs to a single-stage axial compressor (NASA CR-72694). After data export and pre-processing, 56,248 data points were available. From these, two data sets with 500 points each were selected randomly for training and testing purposes, after the data was normalized (mean 0 and scaled to feature an absolute maximum of 1).

The FCM with Euclidean distance norm was used for partitioning. Product space clustering and conclusion parameter estimation were repeated 20 times for any parameterization ($c = 3; 4; \dots; 9$, $m = [1; 2.5]$) in order to prevent termination in a disadvantageous cost function minimum. The best result for each parameterization was considered. In consequence, $c = 6$ and $m = 1.2$ were chosen as base values and Fig. 9 shows the partitioning. Obviously the contour of the significant change is not well met. Parameters were determined using LE with weighting in the product space or GE with MFs in the input space, respectively. Fig. 10 shows an example of the resulting model when using GE.

In the next step, the global fuzziness parameter m and the position of the partition centers $\mathbf{v}_i$ were optimized using the simplex method as provided by the Matlab routine “fmin-search”. LE and GE give similar results in this case, as the optimized values are close to 1 ($m|_{\text{opt}} \approx 1.12$). Fig. 11 and 12 show the result after this step. The “step change” in the contour is now well captured.

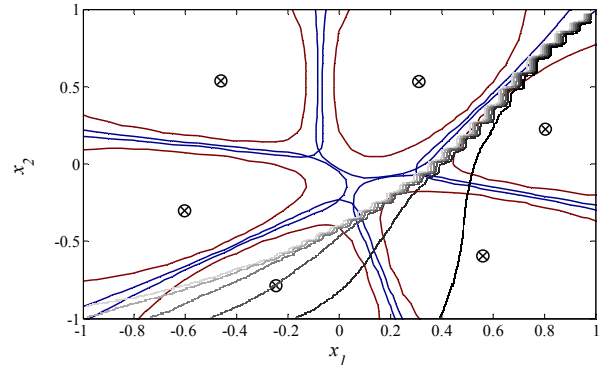


Fig. 9. Partitioning after application of FCM with Euclidean distance norm, $c = 6$ and $m = 1.2$. The gray lines indicate the contour of the original data.

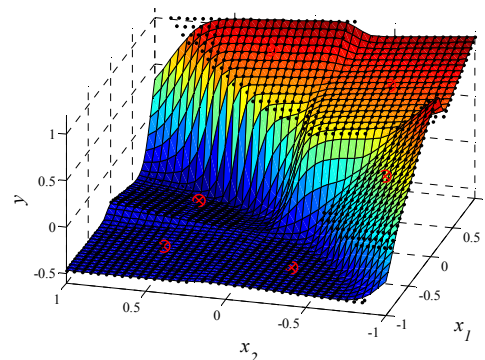


Fig. 10. Local models from GE for partitioning as of Fig. 9.

In a third step, the proposed method for inhomogeneously adjusting the fuzziness $m(\mathbf{x})$ was applied. The resulting values ranged within $m_{i,j} \in [1_+; 1.19]$. The progress of the modeling performance over the different stages is summarized in Table 1. The resulting level curves for m are shown in Fig. 13. Fig. 14 shows the model graph after this step, which well approximates the reference.

Table 1. RMSE after the sequential modeling steps

RMS Error	Training	Test
FCM and WLS	0.1158	0.1326
With optimized $\mathbf{v}$ and m	0.0369	0.0636
With inhomogeneous m	0.0330	0.0619

The adjustment of the partition centers had obviously the largest impact, as this permits to approximately quarter the training and to half the test error. The reason for the improvement is that clustering algorithms minimize a different cost functional than the one that is used for the model as-

assessment. This can cause a partitioning that does not fit the modeling task well. Optimizing the partition centers wrt. the model cost functional moves the partition centers to better fit the contour of the system nonlinearity. An inhomogeneous choice of m permitted to further cut the training error by approx. 10%, while the test error remained almost the same. Münz (2006) identified a PWA model (using Delaunay triangulation for partitioning) with 18 (instead of 6) partitions for a similar problem as shown in Fig 1. The TS model identified in this paper captures the nonlinear origin better while using 3-times less local models.

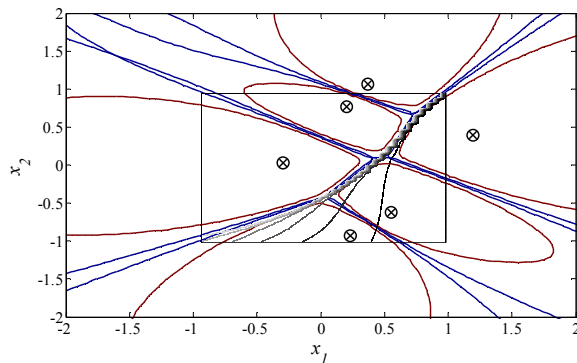


Fig. 11. Partitioning after optimization of all $\mathbf{v}_i$ and m .

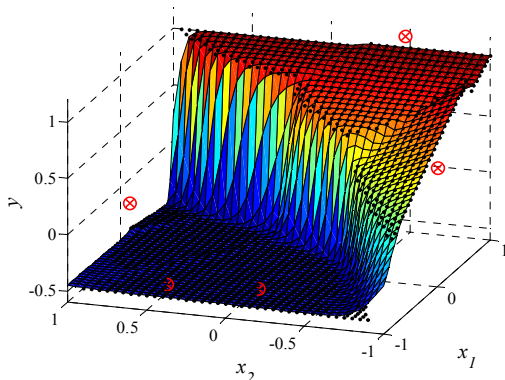


Fig. 12. Graphs of TS model and local models for partitioning as of Fig. 11.

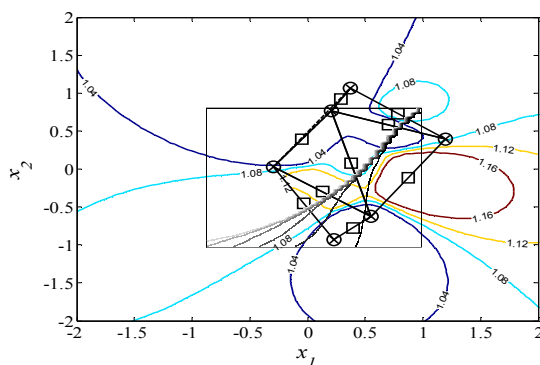


Fig. 13. Level curves of $m(\mathbf{x})$ for final model.

5. CONCLUSIONS AND OUTLOOK

In this article, a novel type of TS models with local adjustable fuzziness $m(\mathbf{x})$ and a corresponding identification me-

thod were presented. The improved performance for heterogeneous modeling problems was demonstrated in an academic example and for approximating a compressor map. It is planned to extend the method to permit Gustafson-Kessel type clustering and to model dynamic systems. Moreover, the identification algorithm has just been designed for prove of concept and can certainly be improved.

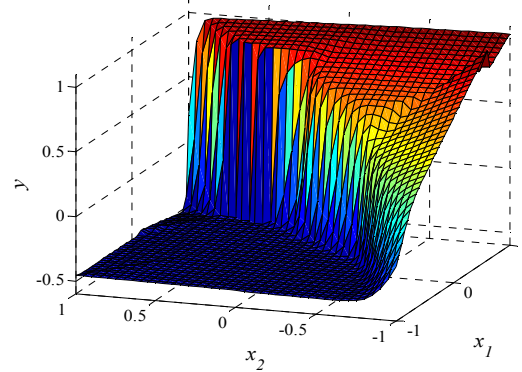


Fig. 14. Prediction of final model.

REFERENCES

- Babuska, R. (1998). *Fuzzy modelling for control*. Kluwer.
- Bezdek, J.C. (1981). *Pattern recognition with fuzzy objective function algorithms*. Kluwer Academic.
- Ferrari-Trecante, G., Muselli, M., Liberati, D., Morari, M. (2003). A clustering technique for the identification of piecewise affine systems. *Automatica*, 39, 205-217.
- Flynn, J.T., Keenan, M.J., Sulam, D.H. (1970). "Single-stage evaluation of highly-loaded high Mach-number compressor stages" NASA Technical Report.
- Höppner, F., Kruse, R., Klawonn, F., Runkler, T. (1999). *Fuzzy Cluster Analysis*, Wiley.
- Klein, R. (2005). *Algorithmische Geometrie: Grundlagen, Methoden, Anwendungen*, Springer.
- Kroll, A. (1996). Identification of functional fuzzy models using multidimensional reference fuzzy sets. *Fuzzy Sets and Systems*, 80, 149-158.
- Kurzke, J., (2009), "Smooth C 8.2: Preparing Compressor Maps for Gas Turbine Performance Modeling," User's Manual.
- Ljung, L. (1999). *System identification*. Prentice Hall.
- Münz, E. (2006). Identifikation und Diagnose hybrider dynamischer Systeme. PhD Dissertation, Karlsruhe.
- Nelles, O. (2001). *Nonlinear system identification*. Springer.
- Paoletti, S., Joloski, A.L., Ferrari-Trecante, G., Vidal, R. (2007). Identification of hybrid systems: a tutorial. *European J. of Control*, 13 (3), 242-260.
- Roll, J., Bemporad, A., Ljung, L. (2004). Identification of piecewise affine systems via mixed-integer programming, *Automatica*, 49, 37-50.
- Ruano, A.E. (2005) *Intelligent control systems using computational intelligence techniques*, IEEE Control Series 70.
- Takagi, T., Sugeno, M. (1985). Fuzzy identification of systems and its application to modelling and control. *IEEE Trans. Systems, Man, and Cybernetics*, 15 (1), 116-132.
- Zimmerschied, R., Isermann, R. (2008). Regularisierungsverfahren für die Identifikation mittels lokal-affiner Modelle. *at-Automatisierungstechnik*, 56 (7), 339-349.