On the Efficiency of Fuzzy Logic for Stochastic Modeling

Ö. Ciftcioglu

Department of Building Technology
Delft University of Technology
Berlageweg 1, 2628 CR Delft, The Netherlands

o.ciftcioglu@tudelft.nl

I.S. Sariyildiz

Department of Building Technology
Delft University of Technology
Berlageweg 1, 2628 CR Delft, The Netherlands

i.s.sariyildiz@tudelft.nl

Abstract - **The properties of fuzzy modeling is investigated for statistical signals. The research makes explicit comparative investigations to position fuzzy modeling in the statistical signal processing domain next to nonlinear dynamic system modeling. It is found that, the nonlinear system dynamics with exogenic inputs can adequately be represented by fuzzy modeling, including the stochastic, as well as deterministic inputs. From the statistical pattern recognition viewpoint, it is interesting to note that, fuzzy modeling has conspicuous ability to represent such input/output relations adequately with relatively small number of fuzzy sets compared to other possibilities, like RBF networks. However, for pattern recognition case, the fuzzy sets determined by data driven methods and approximated by predetermined fuzzy set functions may result in inadequate representations. While such representations can still be used for interpretability reasons, the unshaped data driven projected fuzzy sets remain the essential representation of the fuzzy sets.**

I. INTRODUCTION

After the introduction of artificial intelligence concept *computing with words* is a fundamental contribution of fuzzy logic [1] to this concept. Computing with words became feasible via the utilization of linguistic variables where the words can be interpreted as semantic labels to the fuzzy sets which are the machinery of fuzzy logic. Consequently, human comprehensible computer representation of the domain issues can be created. In this respect symbolic logic associated with fuzzy logic conceivably can do quite a lot that it yielded strong anticipations on the progress of artificial intelligence through expert systems, in the year 80s. Today there remains much to do and human comprehensible computer representations are still in progress. To bring a pragmatic as well as engineering approach to this issue, the computational intelligence concept introduced in 90s became one of the key areas artificial intelligence. Here, the computation is not with words but numbers. However, the words are converted to numbers via fuzzy logic, and the computations are made with well-defined fuzzy computations via fuzzy sets.

By means of computational intelligence, in particular with fuzzy logic, the uncertainty, vagueness imprecision etc can be dealt with. Alternatively, this new domain is referred to as *soft computing*. On one side, dealing with such fuzzy qualities quantitatively is a significant step in the artificial intelligence domain. On the other side, due to the same fuzzy qualities, the interpretability issues appear [2]. While fuzzy logic contributes to science to deal with domain related fuzzy issues, it is natural to anticipate that fuzzy logic associated with the probability theory and statistics can better deal with fuzziness of the domain issues, spanning exact sciences and soft sciences.

Fuzzy modeling is an important means to deal with fuzziness. The type of fuzzy models can vary being dependent on the area it is used. In some applications, qualitative aspects may play essential role while in some other application the numerical properties are more prominent. Both the interpretability properties and numerical properties are important aspects of fuzzy modeling while the gravity is still dependent on the application. In engineering systems, generally fuzzy models are data driven models. In this research, to explore the properties of fuzzy logic in connection to probability theory and statistics, data driven fuzzy models are considered. In this way the statistical properties of fuzzy modeling are conveniently dealt with. Referring to probability theory and statistics, fuzzy logic is a different concept and therefore there is not much work on unifying probability and fuzzy logic in a single context. Therefore the statistical aspects of fuzzy modeling are relatively less received attention than the aspects of *computing with words* and *soft computing*. For dealing with the latter two aspects Mamdani type of fuzzy models are more convenient [3,4], addressing soft issues especially in soft domains. In contrast to this, Takagi-Sugeno (TS) type fuzzy models [5,6] are presumably more convenient in engineering systems to integrate the linguistic information or expert knowledge into engineering design as the consequents are local linear models rather than fuzzy sets subject to aggregation.

To understand the affinity of fuzzy logic with probability theory and statistics several strategies can be adopted. In this research especially stochastic signals with fuzzy modeling are considered since such signals are rich of probabilistic and statistical information to probe fuzzy logic in this context. The fuzzy model is considered as the representation of a general nonlinear dynamic system. The organization of the paper is as follows. Section 2 briefly describes the TS type data driven fuzzy modeling. Section 3 describes the performance of TS fuzzy modeling of a nonlinear dynamic system with stochastic inputs. Section 4 deals with the application of fuzzy modeling to pattern recognition problems where stochastic signals are essential in this particular concern. The paper ends with discussions and conclusions.

II. TS FUZZY MODELING

Takagi-Sugeno (TS) type fuzzy modeling [5,6] consists of set of fuzzy rules a local input-output relation in a linear form as

$$R_i : If \quad x_1 \;\; is \;\; A_{i1} \;\; and \ldots x_n \;\; is \;\; A_{in}$$

$$Then \;\; \hat{y}_i = a_i x + b_i, \qquad i = 1,2,......,K \tag{1}$$

where R_i is the ith rule, $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in X$ is the vector of input variables; $A_{i1}, A_{i2}, \ldots, A_{in}$ are fuzzy sets and y_i is the rule output; K is the number of rules. The output of the model is calculated through the weighted average of the rule consequents of the form

$$\hat{y} = \frac{\sum_{i=1}^{K} \beta_i(x) \hat{y}_i}{\sum_{i=1}^{K} \beta_i(x)} \tag{2}$$

n (2), $\beta_i(x)$ is the degree of activation of the ith rule

$$\beta_i(x) = \prod_{j=1}^{n} \mu_{A_{ij}}(x_j), \quad i = 1,2,....,K \tag{3}$$

where $\mu_{A_{ij}}(x_j)$ is the membership function of the fuzzy set A_{ij} at the input (antecedent) of R_i. To form the fuzzy system model from the data set with N data samples, given by

$$X = [\mathbf{x}_1, \mathbf{x}_2, \ldots\ldots, \mathbf{x}_N]^T, \;\; Y = [y_1, y_2, \ldots\ldots, y_N]^T \tag{4}$$

where each data sample has a dimension of n (N≫n). First the structure is determined and afterwards the parameters of the structure are identified. The number of rules characterizes the structure of a fuzzy system. The number of rules is determined by clustering methods. Fuzzy clustering in the Cartesian product-space X×Y is applied to partition the training data. The partitions correspond to the characteristic regions where the system's behaviour is approximated by local linear models in the multidimensional space.

Given the training data Γ and the number of clusters K, a suitable clustering algorithm [7] is applied. One of such clustering algorithms is known as Gustafson-Kessel (GK) [8]. As result of the clustering process a fuzzy partition matrix U is obtained. The fuzzy sets in the antecedent of the rules are identified by means of the partition matrix U which has dimensions [N×K]; N is the size of the data set. The *ik*-th element of $\mu_{ik} \in [0,1]$ is the membership degree of the i*th* data item in cluster k; that is, the i*th* row of U contains the point wise description of a multidimensional fuzzy set. One-dimensional fuzzy sets A_{ij} are obtained from the multidimensional fuzzy sets by projections onto the space of the input variables x_j. This is expressed by the point-wise projection operator of the form [9,10]

$$\mu_{A_{ij}}(x_{jk}) = proj_j(\mu_{ik}) \tag{5}$$

The point-wise defined fuzzy sets A_{ij} are then approximated by appropriate parametric functions. The consequent parameters for each rule are obtained by means of linear least square estimation. For this, consider the matrices $X = [\mathbf{x}_1, \ldots, \mathbf{x}_N]^T$, $X_e[X, \mathbf{1}]$ (extended matrix $[N \times (n+1)]$) ; Λ_i (diagonal matrix dimension of $[N \times N]$) and

$$X_E = [(\Lambda_1 X_e); (\Lambda_2 X_e); \ldots\ldots (\Lambda_K X_e)] \;\; ([N \times K(n+1)]) \tag{6}$$

where the diagonal matrix Λ_i consists of normalized membership degree as its k-th diagonal element

$$_n\beta_i(x_k) = \frac{\beta_i(x_k)}{\sum_{j=1}^{K} \beta_j(x_k)} \tag{7}$$

The parameter vector ϑ dimension of $[K \times (n+1)]$ is given by

$$\vartheta = [\vartheta_1^T \;\; \vartheta_2^T \;\; \ldots\ldots \;\; \vartheta_K^T]^T$$

where $\vartheta_i^T = [a_i^T \;\; b_i] \;\; (1 \leq i \leq K). \tag{8}$

Now, if we denote the input and output data sets as X_E and Y respectively then, the fuzzy system can be represented as a regression model with the random error *e* of the matrix form

$$Y = X_E \;\vartheta + e \tag{9}$$

For a model with single output (9) becomes a linear regression model of the form

$$y = X \;\vartheta + e \tag{10}$$

where $\vartheta^T = [a^T \;\; b]$ is the regression parameters matrix.

III. DYNAMIC SYSTEM FUZZY MODELING

For the investigation of fuzzy modeling with stochastic excitations, a nonlinear system

$$y(t) = 1 - e^{-x(t)/\tau} \tag{11}$$

is considered. Here *x(t)* is the system variable. For a data driven fuzzy modeling approach, the system representation is cast into a recursive form as

$$y(t) = a(t) \, y(t-1) + u(t) \tag{12}$$

where the time varying AR coefficient *a(t)* and the input *u(t)* are given by

$$a = e^{-[x_2(t) - x_1(t)]/\tau} \tag{13}$$

$$u(t) = 1 - e^{-[x_2(t) - x_1(t)]/\tau}$$

For fuzzy modeling, first the system variable $x(t)$ is considered as band limited white noise and the system response is obtained from (12) for 200 samples. Based on this data the TS fuzzy model of the system is established for three clusters, i.e. three local models. The membership functions and the system performance are shown in figure 1, for 100 samples. In the lower plot the true model output and the fuzzy model output are shown together. There is insignificant difference between these outputs and this is constructive for generalization capability of the model for unknown (test) inputs. In the model τ is taken as $\tau = 25$. The interesting point in this experiment is the ability of the fuzzy model capturing the system dynamics by means of relatively small number of membership functions at the antecedent space
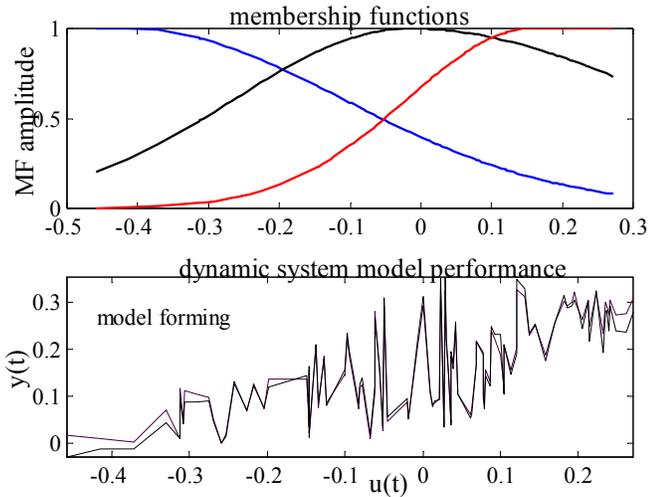
Fig. 1 Membership functions (upper) and the fuzzy model outputs (lower). True model outputs and the estimated counterparts are shown together.

Figure 2 represents the model performance for the test input $u(t)$. The true and the estimated model outputs are shown together in the uppermost plot. The difference between the real and the estimated outputs is in the range of the difference as to the training phase. This is a conspicuous aspect of fuzzy model as robust modeling. The model inputs $u(t)$ corresponding to both outputs of model forming and the test cases are shown in the same figure as middle and lower plots, respectively for 100 samples. From figures 1 and 2, it is to conclude that the fuzzy model has satisfactory performance for stochastic inputs. The system variable $x(t)$ in (11) used to form the model is from band-limited white noise. The system test inputs are from the perception research from a virtual agent reported elsewhere [11]. They are also stochastic and colored by the interaction of the random rays with environment which is subjected to its openness perception measurement. The non-linear system considered above maps visual perception to visual openness perception which is subject to measurement. Note that, in the recursive form the input $u(t)$, given in (12), to the system has wide frequency range. However, the nonlinear system behaves a nonlinear low-pass filter so that three local models give satisfactory

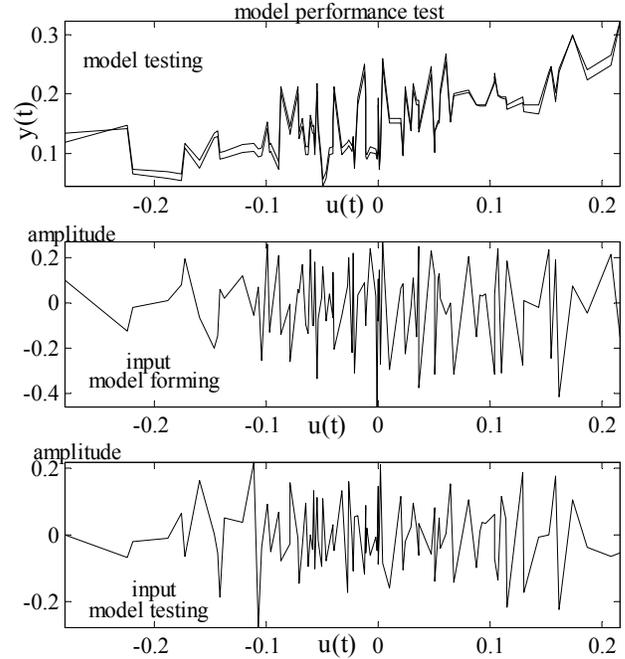estimated system outputs matching the true counterparts rather satisfactorily.

Fig. 2 True model output and its estimation by fuzzy modeling (uppermost plot); the input u(t) to nonlinear system used for fuzzy model formation (middle) and the input u(t) to the same system used for testing the fuzzy model performance (lowermost).

To compare the fuzzy model performance with neural net approach for stochastic inputs, an RBF structure is considered. The equivalence of an RBF structure to a fuzzy model is already well established [12-13]. The structure of the RBF network used for the performance comparison is shown in figure 3. Note that (12) refers to a NARX structure so that the RBF network takes the form of a recurrent network. This is shown in figure 3.
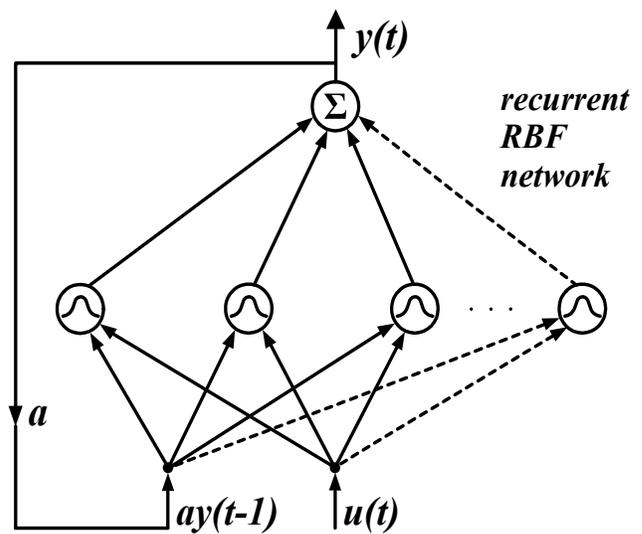

Fig. 3 The RBF network for the performance comparison studies with a fuzzy model.

For comparison, the same excitations as well as the same dynamic system modelling conditions are applied to the RBF network. For five RBF centres at the hidden layer, the results obtained are presented in figure 4.
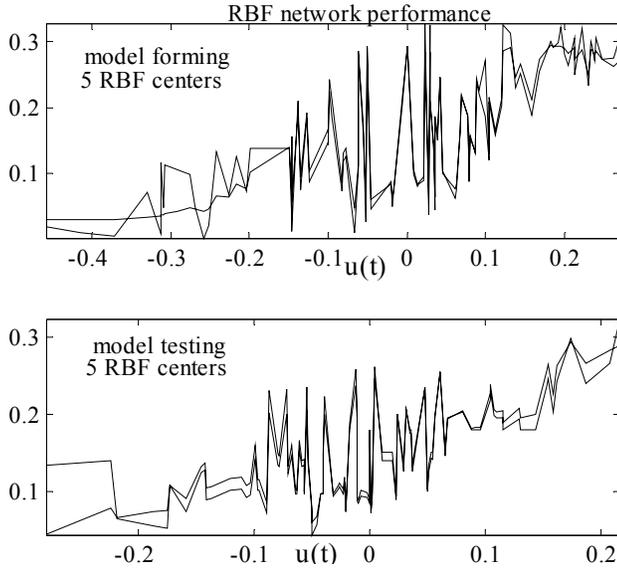

Fig. 4 The results from RBF model subject to comparison with the results from fuzzy model in figure 2.

In figure 4, both results of training and testing are presented. Referring to the differences between true and estimated outputs in both cases, the results are rather unsatisfactory in both cases. To obtain comparable results as to the results from the fuzzy model, the number of nodes of the RBF network is raised to 10. These results are presented in figure 5.
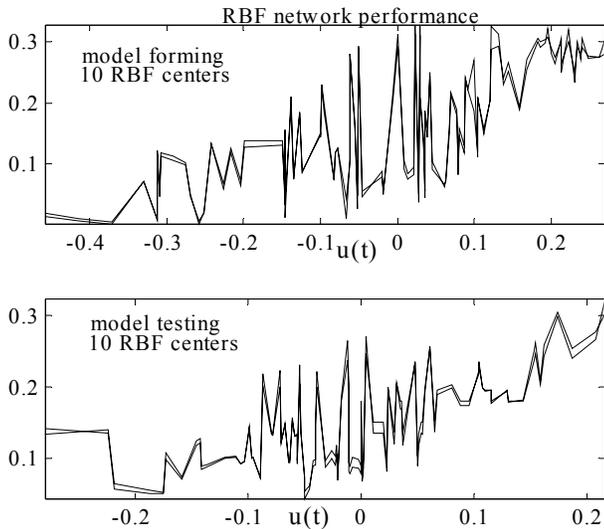

Fig.5 The results from RBF model subject to comparison with the results from fuzzy model in figure 2.

The results presented in figure 5 are comparable with those obtained from the fuzzy model and they are both satisfactory.

Note that the same performance is obtained by 10 nodes in RBF case as compared to three membership functions in fuzzy logic case.

Further studies are carried out to compare the robustness of the models. In this case a uniform white noise source $n(t)$ with mean zero is included to the input of both models so that (12) takes the form

$$y(t) = a(t)\, y(t-1) + u(t) + n(t) \qquad (21)$$

In this case the results from the model are presented in figure 6 for the model-forming case together with the fuzzy membership functions involved.
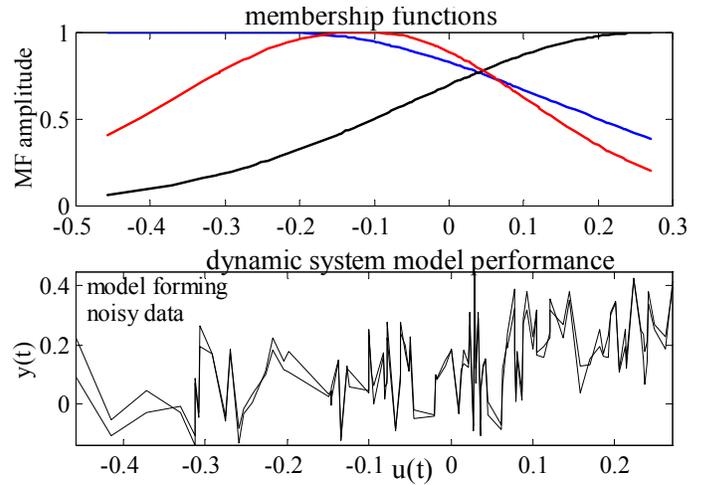

Fig. 6 Membership functions (upper) and the fuzzy model outputs (lower). True model outputs and the estimated counterparts are shown together.

The difference between the true and the estimated counterpart of the model outputs due to noise is noticeable although is not significant. This is an indication of the robustness of the fuzzy model, as this will be clear shortly afterwards. It is noteworthy to mention that, the significant difference only occurs at the location where the input $u(t)$ is at the lower end. To explain this phenomenon, first it is to remember that the model formation is made using white noise and the test inputs are also stochastic while they are colored by the interaction of the random rays with environment subject to its openness perception measurement. The probability density function of $u(t)$ in (12) is shown in figure 7 [14]. The fuzzy model developed in this research is a data-driven type meaning that the data at hand plays the role to determine the location and shape of the fuzzy membership function and consequently the model properties. From figure 7 it is clearly seen that, the gravity of data forming the model is at the positive values of $u(t)$ rather than the negative ones. This means the fuzzy model is well established where the data are favorably concentrates and conversely, model is relatively naïve where the data are unfavorably sparse. The former refers to higher end of $u(t)$ and the latter refers to lower end of $u(t)$. These observations having been made, it is obvious that the model performance is

expectedly favorable at the higher values of u(t) and the reverse for the lower values. This is observed in figure 6.
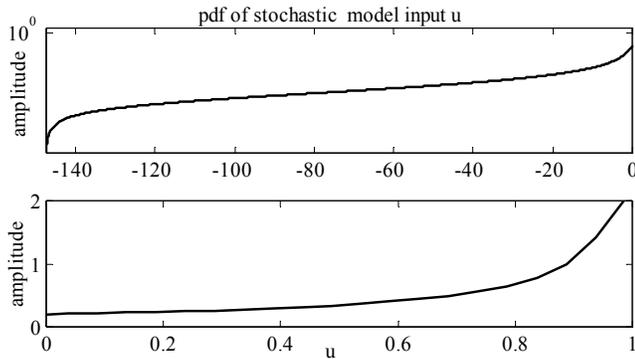

Fig.7 Probability density function of the stochastic model input u(t).

After the model-forming experiment described above having been made, the same experiment is repeated for test data using the same model formed by the noisy data as mathematically described by (13). The results are presented in figure 8. It is interesting to note that, in figure 8, the model testing experiment gave more favorable results than those obtained during model forming. This is a demonstration of a quite typical and well-known phenomenon in data-driven model formation. Namely, in general terms of neuro-fuzzy systems, to introduce some noise to training data improves the generalization capabilities of the model. This is what is happening in the present experiment.
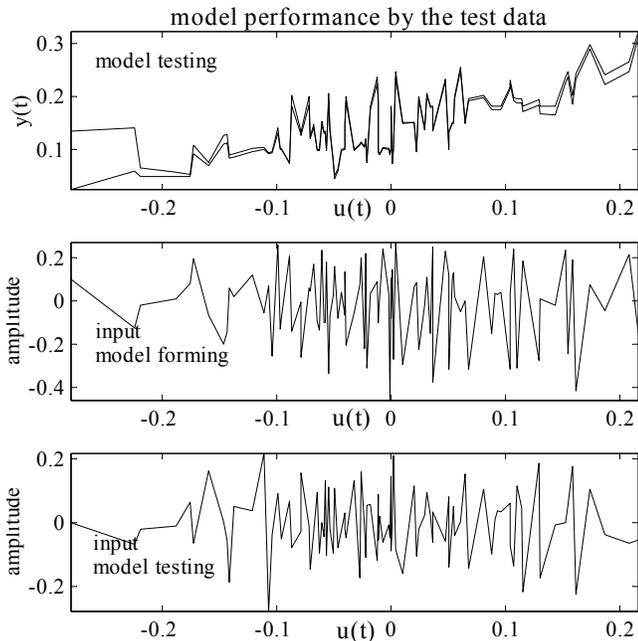

Fig. 8 True model output and its estimation by fuzzy modeling (uppermost plot); input u(t) to the system used for fuzzy model formation (middle) and the input u(t) to the same system used for testing the model performance (lowermost). In the model forming white noise is added to u(t)..

The same experiment with noise described above is repeated again using the RBF network in figure 3. It is already established that, we expect 10 RBF centers in the RBF net for comparable results. These results are presented in figure 9.
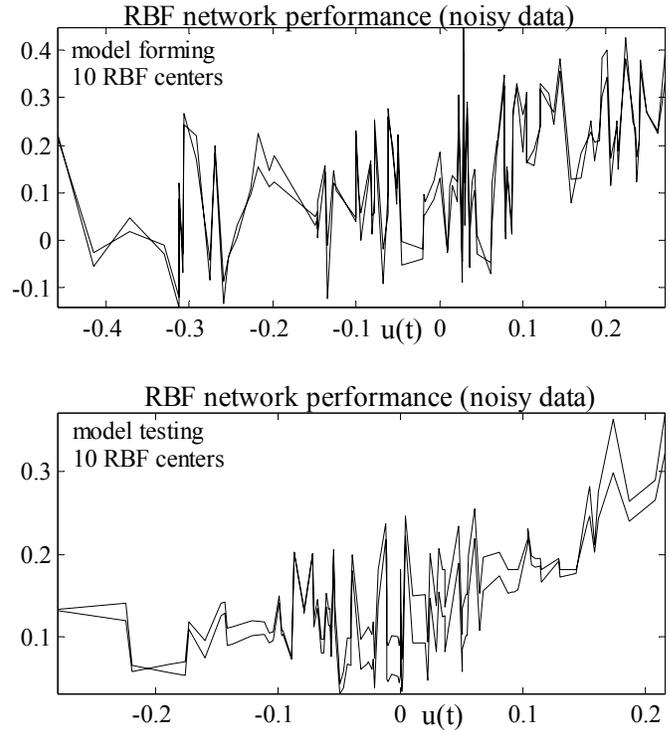

Fig.9 The results from RBF model subject to comparison with the results from fuzzy model in figure 8.

The afore mentioned effect of noise on fuzzy model forming is valid also for the RBF case and this perfectly observed in the upper plot in figure 8. On the contrary to fuzzy model, the test experiments made with the RBF network formed with noisy data, revealed the poor performance of the RBF network in this case. This indicates that the number of nodes in the RBF net should be increased for better performance while apparently ten centers are not enough to capture the non-linear system dynamics. By way of passing, it is worth to mention that, to demonstrate the above-mentioned data-driven model properties with stochastic inputs, the x-axes in the plots are reserved to u(t), in place of parameter t where t represents the time.

IV. FUZZY LOGIC AND PATTERN RECOGNITION

In order to investigate the pattern representation capabilities of fuzzy modeling a block of time-series signal and its wavelet transform is considered. The time-series signal is in particular band-limited white noise. The clustering process together with the local linear models is shown in figure 10 where the number of clusters is four. The corresponding membership functions and the fuzzy model representation of the wavelet coefficients are shown in figure 11.
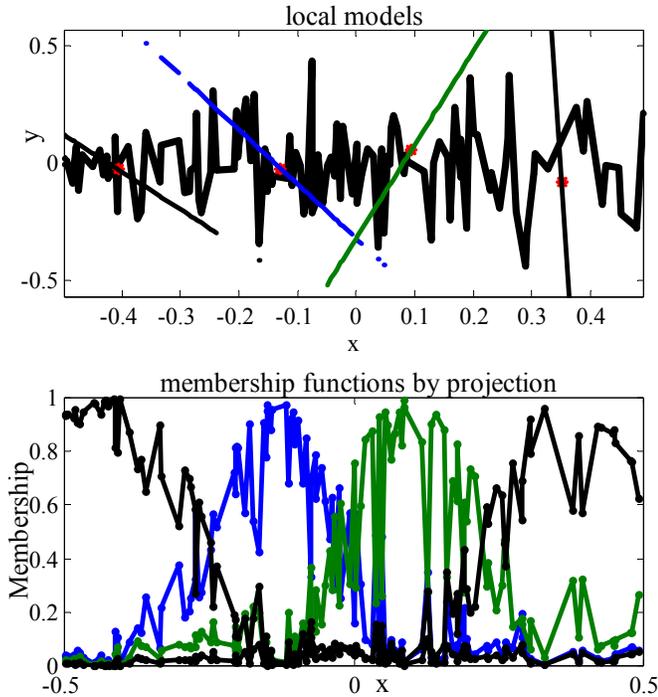
Fig.10 Four local linear models of TS fuzzy modeling (upper) and the projected fuzzy membership functions (lower) where the data are 128-sample wavelet transform coefficients of a time-series data.
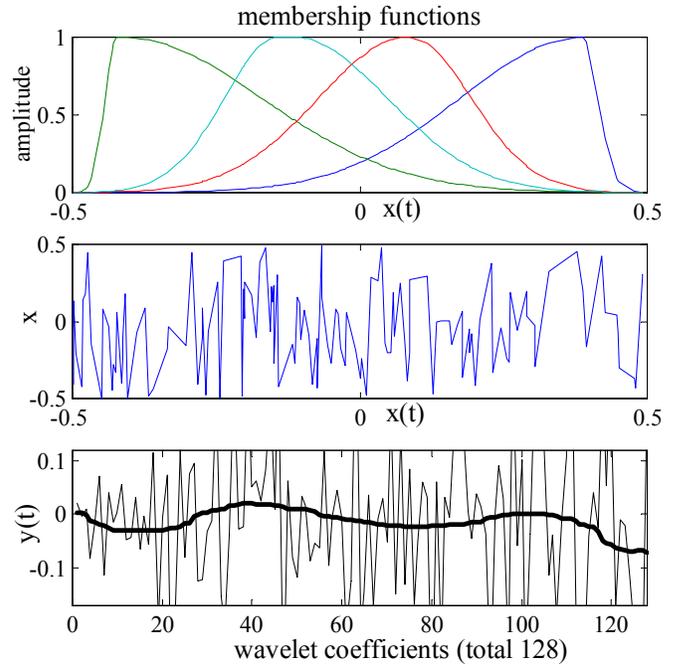

Fig.11 Four fuzzy membership functions Upper), white noise time-series data subjected to wavelet decomposition (middle) and the data formed by the wavelet transform coefficients (lower) where also the estimated counterpart is also shown. The estimation performance is extremely poor.

In figure 11, the membership functions are seen in the upper plot and the time-series data subjected to wavelet transform is shown in the middle plot. In particular, the model outputs as true outputs and their estimated counterparts are seen in the lower plot, in the figure. The model performance for estimation is extremely low.

The above reported computer experiments so far show that TS fuzzy modeling is effective, modeling nonlinear dynamic systems. In the nonlinear dynamic system representation, since the system is restricted to lower frequency region, the membership functions identified by clustering and shaped via appropriate modified gaussians seen in figures 1,6 and 11 are capable to represent the system dynamics adequately. However in the pattern representation, for instance the wavelet transform coefficients as is the case here, the frequency band is wide as the time-series data subjected to wavelet decomposition is band-limited white. Therefore, the shaped fuzzy membership functions cannot be used to represent the wavelet transform data given. This phenomenon is observed in the figure 11. With other words, the shaped gaussians are not narrow enough and their total number is not sufficient enough to represent the local variations The solution is, in place of shaped gaussian membership functions, to use the membership functions obtained directly from the clustering process. The results are shown in figure 12. By means of unshaped membership functions much improved estimation is obtained compared to that obtained in figure 11.

By means of unshaped membership functions much improved estimation is obtained compared to that obtained in figure 11. In the upper plot of this figure, it is interesting to note that, the wavelet coefficients are relatively small at the lower range, i.e., the coefficients from 1 to 16, due to the logarithmically descending resolution bandwidth peculiar to wavelet transform This phenomenon is anifested better at lower resolutions, that is, the resolution band becomes very narrow. Due to small coefficient values at the lower resolutions, which are close to support boundaries of the transform, the estimation results are relatively poor at this region. However, much improved estimation of the wavelet transform coefficients via the fuzzy model with higher membership functions is obtained as compared to the estimation seen in figure 12.

The same computer experiments as presented in figures 10 and 11, are repeated with the increased number of membership functions. For six membership functions the results are shown in figures 13 and 14 as counterparts of figures 10 and 11. As it is seen in figure 14, the shaped gaussians are not narrow enough to represent the local variations. The solution as before is, in place of shaped gaussian membership functions, to use the membership functions obtained directly from the clustering process. By doing so, the results obtained are shown in figure 15 where definitely much improved estimation is obtained compared to that given in figure 14 (lower plot) but also compared to that given in figure 11 (lower plot), as well. With other words, the

local variations can not be represented by restricted number of local linear models defined by the number
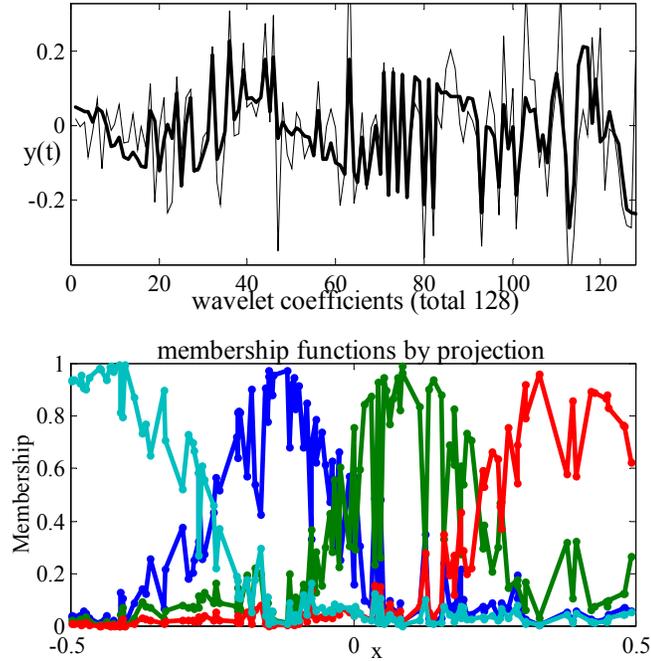

Fig. 12 The projected but unshaped membership functions (lower) and the estimated wavelet coefficients by fuzzy model (upper).
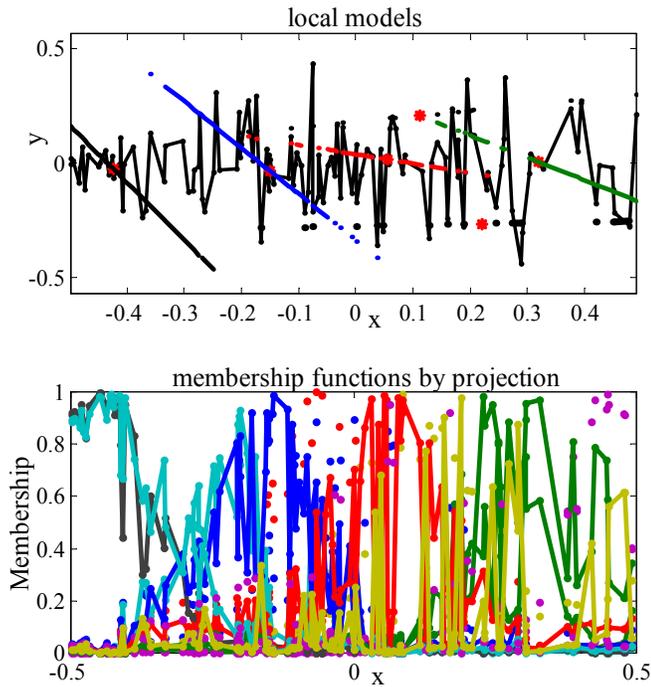

Fig. 13 Six local linear models of TS fuzzy modeling (upper) and the projected fuzzy membership functions (lower) where the data are 128-sample wavelet transform coefficients of a time-series data.
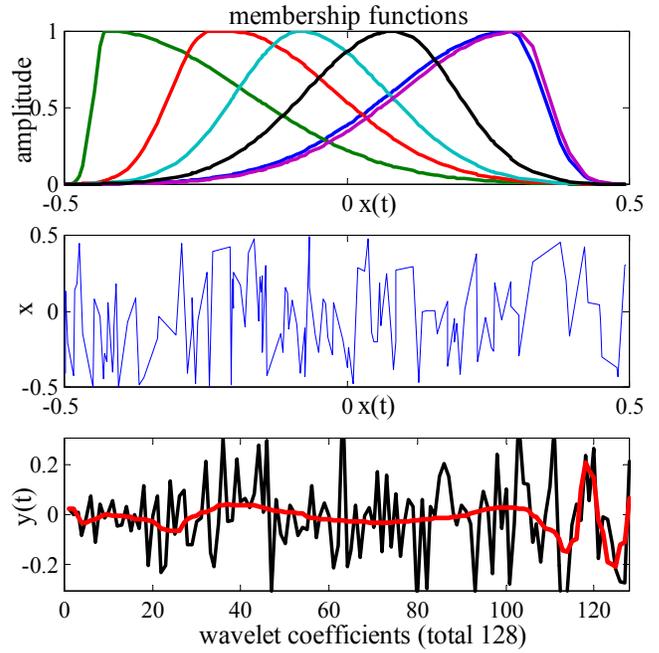

Fig. 14 Six fuzzy membership functions (upper), white noise time-series data subjected to wavelet decomposition (middle) and the data formed by the wavelet transform coefficients (lower) where also the estimated counterpart is also shown. The estimation performance is extremely poor.
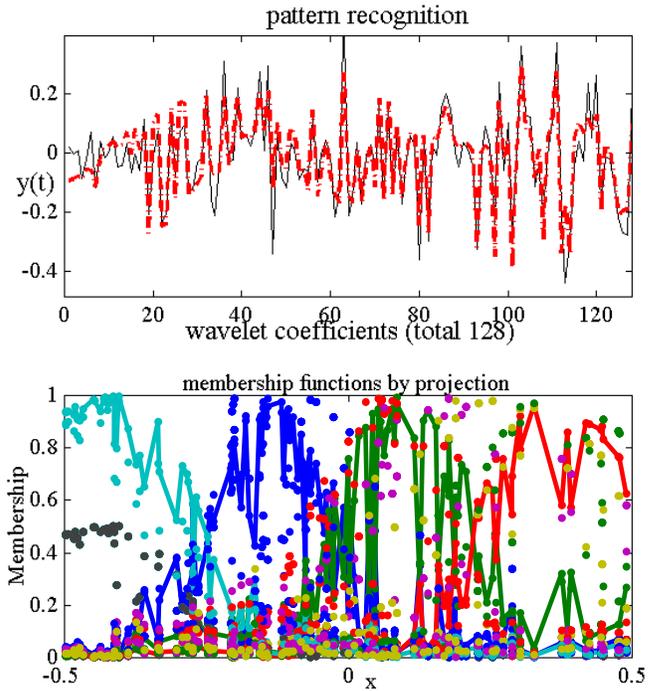

Fig.15 The projected but unshaped membership functions (lower) and the estimated wavelet coefficients by fuzzy model (upper).

of clusters. A similar situation where the membership functions directly from the clustering are employed, is the case of multivariable fuzzy modeling. In this case the cause is different but the consequence is the same. Namely, in the multivariable case, irrecoverable projection error due to projected and shaped membership functions from the clustered data prevents accurate representation of the dynamic model in multidimensional space [15].

By considering such basic features of fuzzy modeling stated above, the fuzzy logic can be conveniently associated with the probabilistic entities, as these are stochastic signals and patterns, in this work. This is especially the case in perception studies being carried out in the department at Delft University of Technology, The Netherlands, where probability plays the essential role to model the perception process [11]. In these studies associations are made from visual perceptions of an environment to perceptual qualities of it. Because of the softness of perception, association of probability with fuzzy logic is much demanded and the research progresses along that line. In perception studies, a snapshot of a real-time perception measurement via a virtual agent in a virtual reality environment is shown in figure 16.
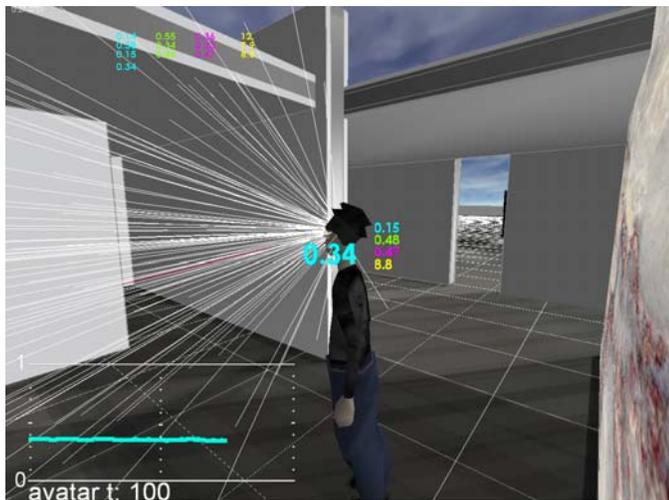

Fig. 16 Perception experiment by a virtual agent in virtual reality. The interacting rays, which simulate the vision process and ensuing visual openness measurement variations in real-time are also shown.

V. DISCUSSION AND CONCLUSIONS

TS fuzzy modeling is an essential means for representation of nonlinear dynamic systems for identification, control etc. Such system dynamics is represented relatively small number of fuzzy sets compared to other approaches, the RBF network for instance. This is especially due to singletons, i.e., normalized output weights as consequents in the RBF model. For nonlinear dynamic system identification, probability density investigations of stochastic model inputs and outputs can reveal important information about the unknown system subject to stochastic modeling. In this respect, the capabilities of fuzzy modeling and its behavior with stochastic excitations are demonstrated in this work. It is

shown that the pattern representation capability of fuzzy modeling is satisfactory even for stochastic time-series data. As result of these studies a clear message is conveyed that, effective associations of probabilistic data can be made with fuzzy logic and these associations can be exploited in variety of ways. The above reported computer experiments show that TS fuzzy modeling is effective modeling nonlinear dynamic systems and representation of patterns. In this research of nonlinear dynamic system modeling, since the system is restricted to lower frequency region, the membership functions identified by clustering and modified with gaussian shapes are capable to represent the system adequately. However in the pattern representation, the frequency band is wide as the time-series data are band-limited white. Therefore, in place of shaped gaussian membership functions, the membership functions obtained directly from the clustering process should be used. Otherwise, the shaped gaussians do not localize enough to represent the local variations. In RBF networks this localization is accomplished with increased number of nodes.

REFERENCES

[1] L. A. Zadeh, 'Fuzzy logic=computing with words', *IEEE Transactions on Fuzzy Systems*, 4(2),pp.103-111, May 1996.
[2] J. Casillas, O. Cordon, F. Herrara, L. Magdalena (Eds.), *InterpretabilityIssues in Fuzzy Modeling,* Springer, 2003.
[3] E.H. Mamdani, "Applications of fuzzy algorithms for control a simple dynamic plant", *Proceedings of the IEE*, 121, 12, pp.1585-1588, 1974.
[4] E.H. Mamdani and S. Asillian, "An experiment in linguistic synthesis with fuzzy logic controller", *International Journal of Man-Machine Studies*, 7, pp.1-13, 1975.
[5] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its application to modelling and control", *IEEE trans. On Systems, Man, and Cybernetics*, 15, pp.116-132, 1985.
[6] M. Sugeno and G.T. Kang, "Structure identification of fuzzy model", *Fuzzy Sets and Systems*, 29, pp.15-33, 1988.
[7] J. C. Bezdek, Pattern Recognition with fuzzy objective function algorithms, New York, Plenum, 1981.
[8] D. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix", in *Proc. IEEE CDC*, San Diego, CA, pp. 761-766, 1979.
[9] R. Kruse, J. Gebhardt and F. Klawonn, *Foundations of Fuzzy Systems*, Wiley, New York, 1994.
[10] K.H. Lee, *First Course on Fuzzy Theory and Applications*, Springer, Berlin, 2005.
[11] M. Bittermann and Ő. Ciftcioglu, "Real-time measurement of perceptual qualities in conceptual design", Proc. *International Symposium series on Tools and Methods of Competitive Engineering*, *TMCE 2006*, April 18–22, 2006, Ljubljana, Slovenia.
[12] Hunt K.J., Haas R. and Murray-Smith R., , "Extending the functional equivalence of radial basis function networks and fuzzy inference systems", *IEEE Trans. Neural Networks,* Vol.7, No.3, May 1998.
[13] J.S.R Jang and C.T. Sun, "Functional equivalence between radial basis function networks and fuzzy inference systems", *IEEE Trans. Neural Networks,* Vol.4, No.1, January, 1993.
[14] Ő. Ciftcioglu and I.S. Sariyildiz, "Fuzzy logic for stochastic modeling", Proc. *Soft Methods in Probability and Statistics, SMPS 2006*, University of Bristol – UK, September 5-7, 2006.
[15] Ő. Ciftcioglu and I.S. Sariyildiz, "On the efficiency of multivariable TS fuzzy modeling', *FUZZ-IEEE 2006,, Proc. World Congress on Computational Intelligence*, July 16-21, 2006, Vancouver, Canada. 2006.