

# Takagi-Sugeno Neuro-Fuzzy Modeling of a Multivariable Nonlinear Antenna System

E. A. Al-Gallaf\*

Department of Electrical and Electronics Engineering, College of Engineering  
University of Bahrain, P.O. Box 13184, Kingdom of Bahrain

Received 17 August 2003; accepted 7 April 2004

نموذج تاكاجي - سوكينو معتمد على الخلايا العصبية والمنطق المضرب لنظام هوائي غير خطي ومتعدد المتغيرات

**المستخلص:** يهدف هذا البحث الى استخدام وسائل التعليم الذكية (مثل الشبكات العصبية الالكترونية) لنظم التحكم وذلك من اجل اعداد نماذج للعمليات الديناميكية الغير خطية ذات المتغيرات العده والتي يصعب تمثيلها بوسائل التمثيل الاعتيادية. وفي هذا الاطار فقد تم توزيع استجابة النظام المراد تمثيلة باستخدام (Clustering) ومن ثم تم استخدام الشبكات العصبية الالكترونية المكونة من خمسة طبقات لتمثيل و بناء نظام (Fuzzy Logic) والذي سوف يستخدم لبناء نموذج لنظام غير خطي ديناميكي و هو لنظام اتصالات (Multi-variable Antenna System). واخيرا تم استخدام الطرق الاحصائية لاثبات ان النموذج المقترح قادر على تمثيل العمليات المعقدة ومن ثم بناء نظام التحكم.

**المفردات المفتاحية:** الأنظمة العصبية المضربة - التجمع العنقودي المضرب - أنظمة تاكاجي - سوكينو، الأنظمة اللاخطية.

**Abstract:** This article investigates the use of a clustered based neuro-fuzzy system to nonlinear dynamic system modeling. It is focused on the modeling via Takagi-Sugeno (T-S) modeling procedure and the employment of fuzzy clustering to generate suitable initial membership functions. The T-S fuzzy modeling has been applied to model a nonlinear antenna dynamic system with two coupled inputs and outputs. Compared to other well-known approximation techniques such as artificial neural networks, the employed neuro-fuzzy system has provided a more transparent representation of the nonlinear antenna system under study, mainly due to the possible linguistic interpretation in the form of rules. Created initial memberships are then employed to construct suitable T-S models. Furthermore, the T-S fuzzy models have been validated and checked through the use of some standard model validation techniques (like the correlation functions). This intelligent modeling scheme is very useful once making complicated systems linguistically transparent in terms of the fuzzy if-then rules.

**Keywords:** Neuro-fuzzy systems, Fuzzy clustering, Takagi-Sugeno modeling, Nonlinear systems

## 1. Introduction

### 1.1 Takagi-Sugeno Models

Developing mathematical models of real systems is a central topic in many disciplines of engineering and science. Models can be used for simulations, analysis of the system's behavior, better understanding of the underlying mechanisms in the system, design of new processes, or design of controllers. Takagi-Sugeno (T-S) modeling plays an essential role in deriving local linear models of the nonlinear dynamic system under concern (Lo and Chen, 1999; Takagi and Sugeno, 1985). Through the use of the heuristic rules inherent in the fuzzy systems, T-S fuzzy models, then make it possible to have a transparent like system which is governed by the fuzzy inference system and rules. Fuzzy modeling concerns the methods of describing the characteristics of a system using fuzzy inference rules. Fuzzy modeling methods have a distin-

guishing feature in that they can express complex nonlinear systems linguistically.

In a similar way fuzzy clustering has been utilized as well in classifying data-driven fuzzy modeling, since it draws a methodology for assigning label to similar data. Such assignment does give quantitative directions for shaping the fuzzy membership functions. Model validation and verification is also an important task in the modeling paradigm. This is due to the choice of the right model from a number of models that might present similar characteristics. Statistically validated models in addition to probabilistic validation are used sometimes to make the suitable choice of a system model. In general, fuzzy control systems can be classified as linguistic (Lo and Chen, 1999; Mamdani and Assilian, 1975). The linguistic type fuzzy control system is well recognized and received by the control society. The T-S type fuzzy system, which will be used in this article mainly focuses on the modeling aspect. It has been reported that a T-S fuzzy system can exactly model any nonlinear system Wang, *et al.* (2000). On the other hand there is a main drawback

\*Corresponding author E-mail: ebrgallaf@eng.uob.bh

of the linguistic model compared with the T-S model in that there is a difficulty in dealing with a multidimensional system since a large number of fuzzy rules have to be used.

Gorzalczany *et al.* (2000), has briefly presented and compared four neuro-fuzzy systems used for rule-based modeling of dynamic processes (*chaotic Mackey-Glass time series*). The following systems have been considered: NFMOD - the proposed system, the well-known ANFIS and NFIDENT systems, and an alternative neuro-fuzzy system already reported in literature. The main criterion of comparison of all systems is their performance (modeling accuracy) versus interpretability (the transparency and the ability to explain generated decisions; it also includes an analysis and pruning of obtained fuzzy-rule bases). On the other hand, Zhang and Knoll (1995), have proposed an approach for solving multivariate modeling problems with neuro-fuzzy systems. Instead of using selected input variables, statistical indices are extracted to feed a fuzzy controller. The original input space was transformed into an eigen-space. If a sequence of training data are sampled in a local context, a small number of eigenvectors which possess larger eigen-values provide a good summary of all the original variables. Fuzzy controllers can be trained for mapping the input projection in the eigen-space to the outputs. Implementations with the prediction of time series was used to validate the concept.

The article of Ikonen and Kortela (2000) is concerned with a process modeling using fuzzy neural networks. In Distributed Logic Processors (DLP) the rule base is parameterized. The DLP derivatives required by gradient-based training methods are given, and the recursive prediction error method is used to adjust the model parameters. The power of the approach is illustrated with a modeling example where NO<sub>x</sub>-emission data from a full-scale fluidized-bed combustion district heating plant are used. The method presented in their paper was general, and can be applied to other complex processes as well. Bologna (2001) has presented a new neuro-fuzzy model denoted as Fuzzy Discretized Interpretable Multi-Layer Perceptron (FDIMLP). Fuzzy rules were extracted in polynomial time with respect to the size of the problem and the size of the network. He applied our model to three classification problems of the public domain. It turned out that FDIMLP networks compared favorably with respect to EFUNN and ANFIS neuro-fuzzy systems.

Ning *et al.* (2001) presented a fuzzy satisfactory clustering algorithm in their paper. It started with two cluster centers and adds new center if necessary. A system data set was quickly divided into several satisfactory fuzzy clusters by the algorithm. A T-S type fuzzy model was then, identified. Chen and Linkens (1998) introduced a three-layered RBF (Radial Basis Function) network to implement a fuzzy model. Differing from existing clustering-based methods, in their approach the structure identification of the fuzzy model, including input selecting and partition validating, was implemented on the basis of

a class of sub-clusters created by a self-organizing network instead of raw data. The important input variables which independently and significantly influence the system output can be extracted by a fuzzy neural network. On the other hand, the optimal number of fuzzy rules can be determined separately via the fuzzy c-means clustering algorithm with a modified fuzzy entropy measure as the criterion of cluster validation.

Akkizidis and Roberts (2001) proposed an algorithmic methodology for identifying and modeling non-linear control strategies. The methodology presented was based on choices of different fuzzy clustering algorithms, projection of clusters and merging techniques. The best features of well-known clustering methods such as the Gustafson-Kessel and mountain method were combined. The latter was used to determine and define the number and the approximate positions of the cluster prototypes; whereas the former was used to define the shapes of the clusters according to the data distribution. The projection of the prototypes and variables of clusters was a recognized approach to extracting the information included in the data clusters into fuzzy sets. Merging these fuzzy sets, based on proposed guidelines, can minimize the number of rules and make the identifying control strategy more transparent. Bossley (1997) has looked into the problem of antenna modeling via neuro-fuzzy systems, however, getting an optimized five layers neural network was not easily achieved due to the large number of generated fuzzy rules.

## 1.2 Article Contribution

The system under study is typical of the type used for oceanary satellite communication systems and has a high nonlinear coupling among its two outputs. Hence, it is required to have transparent sub-models. This class of multivariable system has been modeled via a classical Neuro-fuzzy system as in Bossley (1997). However, it did result in a large number of rules, and large number of training patterns were required. In this respect, this research frame work is investigating the use of clustered fuzzy rules, that makes it easy for the training mechanism to be achieved in less time with fewer number of rules. Fuzzy sets in the antecedent of the rules are obtained from the partition matrix by projection onto certain antecedent variables. The obtained point-wise fuzzy sets are then approximated by some suitable parametric functions. The transparency of the antenna model obtained using the above approach may be hindered by the redundancy present in the form of many overlapping (*compatible*) membership functions. Certain similarity measures were used in order to assess the compatibility (*pair-wise similarity*) of fuzzy sets in the rule base, in order to detect sets that can be merged. Fuzzy sets estimated from antenna training data can also be similar to the universal set, thus adding no information to the model. Sets of such nature were removed from the antecedent of the rules, thus reducing the number of the fuzzy rules.

## 2. Intelligent Dynamic Systems Modeling

### 2.1 Intelligent Modeling

Fuzzy modeling and control are typical examples of techniques that make use of human knowledge and deductive processes. Various alternative approaches have been proposed, Fuzzy Logic and Set Theory being one of them. Artificial neural networks and fuzzy models belong to the most popular model structures used. From the input-output view, fuzzy systems are flexible mathematical functions, which can approximate other functions or just data measurements with a desired accuracy. Compared to well-known approximation techniques such as Neural Networks, fuzzy systems provide a more transparent representation of the system under study, which is mainly due to the possible linguistic interpretation in the form of rules. The logical structure of the rules facilitates the understanding and analysis of the model in a semi-qualitative manner, close to the way human reason about the real world.

Given the state of a system with a given input, the next state  $x(k+1)$  can be determined. In the sense of discrete-time setting, it can be written as:

$$x(k+1) = f(x(k), u(k)) \quad (1)$$

where  $x(k)$  and  $u(k)$  are the state and the input at time  $k$ , respectively, and  $f$  is a static function. Fuzzy models of different types can be used to approximate the state-transition function. As the state of a system is often not measured, input-output modeling is usually applied. The most common is the NARX (Nonlinear Auto-Regressive with Exogenous input) model, as defined by

$$\begin{aligned} y(k+1) = f(y(k), y(k-1), \dots, y(k-n_y+1), \\ u(k), u(k-1), \dots, u(k-n_u+1)) \end{aligned} \quad (2)$$

where  $y(k), \dots, y(k-n_y+1)$ , and  $u(k), \dots, u(k-n_u+1)$  denote the past model outputs and inputs respectively and  $n_y$  and  $n_u$  are integers related to the model order (usually selected by the designer). For instance in Eq. (3), a linguistic fuzzy model of a dynamic system may consist of rules of the following form:

$$\begin{aligned} R_i: \text{if } y(k) \text{ is } A_{i1} \text{ and } y(k-1) \text{ is } A_{i2} \text{ and, \dots, } y(k-n+1) \\ \text{is } A_{in} \text{ and } u(k) \text{ is } B_{i1} \text{ and } u(k-1) \text{ is } B_{i2} \\ \text{and, \dots, } u(k-m+1) \text{ is } B_{im} \text{ then } y(k+1) \text{ is } C_i \end{aligned} \quad (3)$$

In Eq. (3), the input dynamic filter is a simple generator of the lagged inputs and outputs, and no output filter is used. Since the fuzzy models can approximate any smooth function to any degree of accuracy, models of the type in Eq. (3) can approximate any observable and controllable modes of a large class of discrete-time nonlinear systems. To facilitate data-driven optimization of fuzzy models (learning), differentiable operators (product, sum)

are often preferred to the standard min and max operators. Once the structure is fixed, the performance of a fuzzy model can be fine-tuned by adjusting its parameters. Tunable parameters of linguistic models are the parameters of antecedent and consequent membership functions (determine their shape and position) and the rules (determine the mapping between the antecedent and consequent fuzzy regions).

It is assumed that a set of  $N$  input-output data pairs  $\{\mathbf{x}_i, y_i\}_{i=1,2,\dots,N}$  is available. Recalling that  $\mathbf{x}_i \in \mathfrak{R}^{F_{inp}}$  are input vectors and  $y_i$  are output scalars. Denote  $\mathbf{X} \in \mathfrak{R}^{N \times F_{inp}}$  a matrix having the vectors  $\mathbf{x}_k^T$  in its rows, and  $\mathbf{y} \in \mathfrak{R}^N$  a vector containing the outputs  $y_k$ :

$$\mathbf{x} = (x_1, \dots, x_N)^T \quad (4)$$

$$\mathbf{y} = (y_1, \dots, y_N)^T \quad (5)$$

## 3. Neuro-Fuzzy Modeling

Figure 1 shows typical five layers of a neuro-fuzzy system that can be employed to accomplish a rule network. Typically, such rules are

$$\text{if } x_1 \text{ is } A_{11} \text{ and } x_2 \text{ is } A_{21} \text{ then } y=b_1 \quad (6)$$

$$\text{if } x_1 \text{ is } A_{12} \text{ and } x_2 \text{ is } A_{22} \text{ then } y=b_2 \quad (7)$$

Nodes in the first layer compute the membership degree of the inputs in the antecedent fuzzy sets. The product nodes  $\Pi$  in the second layer represent the antecedent conjunction operator. The normalization node  $N$  and the summation node  $\Sigma$  realize the fuzzy-mean operator. Using smooth antecedent membership functions, such as a Gaussian function, as given below

$$\mu_{A_{ij}}(x_j, c_{ij}, \tau_{ij}) = \exp\left(-\left(\frac{x_j - c_{ij}}{2\tau_{ij}}\right)^2\right) \quad (8)$$

in which  $c_{ij}$  and  $\tau_{ij}$  parameters are adjusted by gradient-descent learning algorithms, such as back-propagation. This allows a fine-tuning of the fuzzy model to the available data in order to optimize its prediction accuracy. There may be a lot of structure/parameter combinations which make the fuzzy model behave in a satisfactory way. The problem can be formulated as that of finding the structure complexity which will give the best performance in generalization. In our approach we choose the number of rules as the measure of complexity to be properly tuned on the basis of available data. We adopt an incremental approach where different architectures having different complexity (*i.e.* number of rules) are first assessed in cross-validation and then compared in order to select the best one.

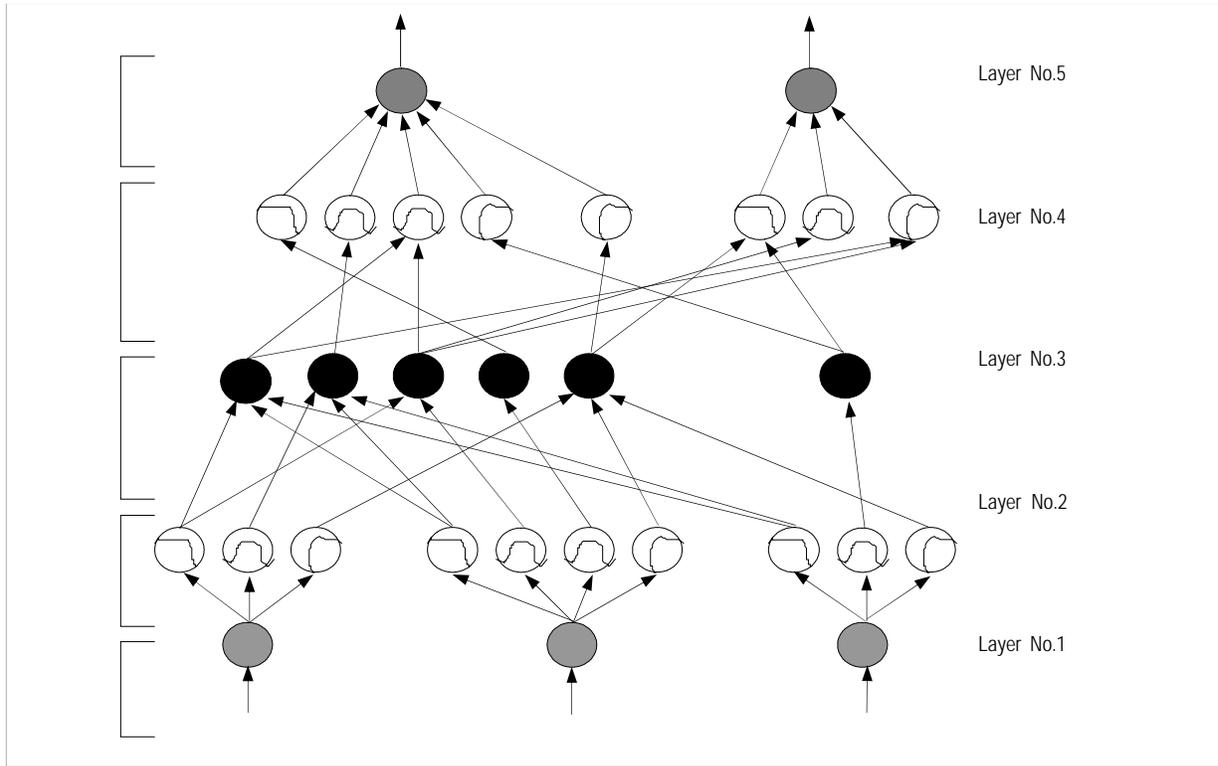


Figure 1. A five layer neurofuzzy network architecture

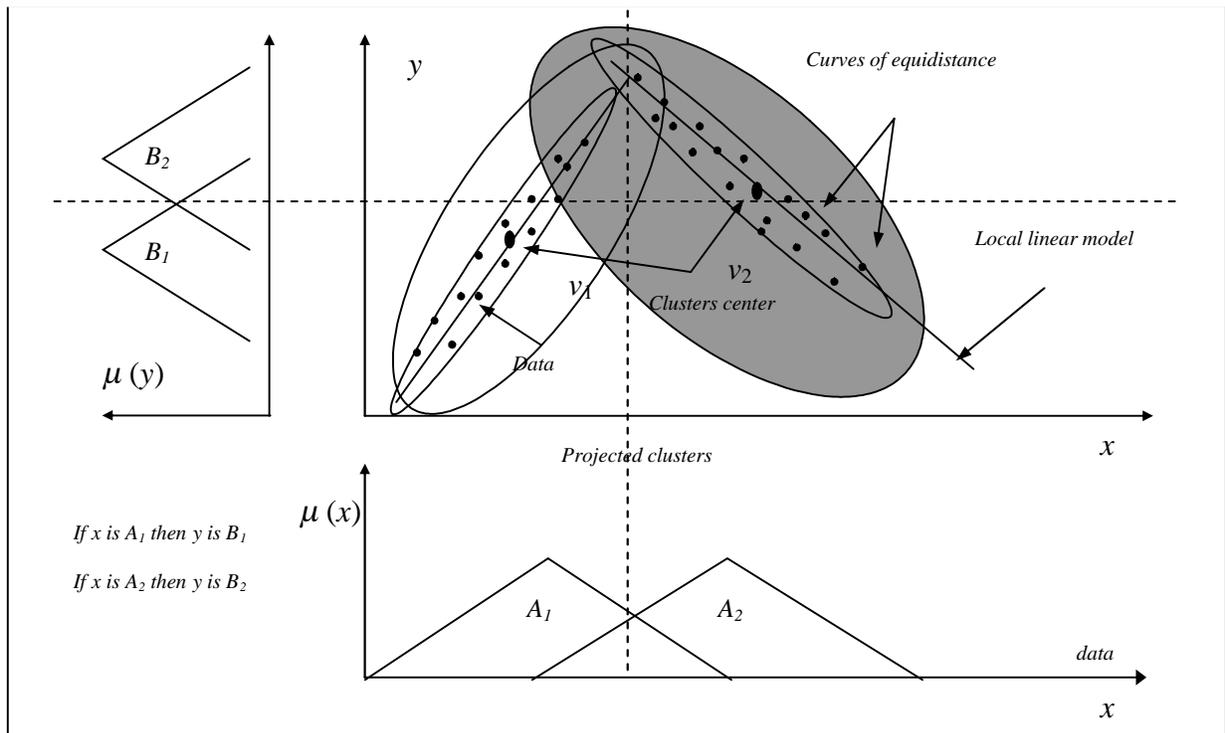


Figure 2. Hyper ellipsoidal fuzzy clusters

The initialization of the architecture is provided by a hyper-ellipsoid fuzzy clustering procedure inspired by Babuska and Verburggen (1995). This procedure is clustering the data in the input-output domain obtaining a set of hyper-ellipsoids which are a preliminary rough representation of the input/output mapping. Methods for initializing the parameters of a fuzzy inference system form the outcome of the fuzzy clustering procedure. Here we use the axes of the ellipsoids (eigenvectors of the scatter matrix) to initialize the parameters of the consequent functions. We project the cluster on the input domain to initialize the centers of the antecedents and we adopt the scatter matrix to compute the width of the membership functions. Once the initialization is done, the learning procedure begins. In the case of linear T-S models this minimization procedure can be decomposed in a least-squares problem to estimate the linear parameter of the consequent models and a nonlinear minimization to find the parameters of the membership functions. The structural identification loop (the outer one) searches for the best structure, in terms of optimal number of rules, by increasing gradually the number of local models.

## 4. Fuzzy Pattern Clustering

### 4.1 Fuzzy Clustering

Identification methods based on fuzzy clustering originate from data analysis and pattern recognition, where the concept of graded membership is employed to represent the degree to which a given object, represented as a vector of features, is similar to some prototypical object. Based on that similarity, feature vectors can be clustered such that vectors within a cluster are as similar as possible, and vectors from different clusters are as dissimilar as possible. This thought of fuzzy clustering is depicted in Fig. 2. Data is clustered into two groups with prototypes  $v_1$  and  $v_2$ , using the Euclidean distance measure. The partitioning of the data is expressed in the fuzzy partition matrix whose elements  $\mu_{ij}$  are degrees of membership of data points  $(x_i, y_i)$  in a fuzzy cluster with prototypes  $v_j$ . The concept of similarity of data to a given prototype leaves enough space for the choice of an appropriate distance measure and of the character of the prototype itself. Prototypes can be defined as linear subspaces, or the clusters can be ellipsoids with adaptively determined shape Akkizidis and Roberts, (2001). From these clusters, the antecedent membership functions and the consequent parameters of the T-S model can be extracted as follows, (Bossley, 1997).

$$\begin{aligned} \text{if } x \text{ is } A_1 \quad \text{then } y &= a_1x + b_1 \\ \text{if } x \text{ is } A_2 \quad \text{then } y &= a_2x + b_2 \end{aligned} \quad (9)$$

Each obtained cluster is represented by one rule in the T-S model. Membership functions for fuzzy sets  $A_1$  and  $A_2$  are generated by pointwise projection of the partition matrix

onto the antecedent variables. Such pointwise defined fuzzy sets are then approximated by a suitable parametric function.

### 4.2 Fuzzy Clustering Algorithm

Consider a finite set of elements  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$  as being elements of the  $F_{\text{inp}}$  dimensional Euclidean space  $\mathcal{R}^{F_{\text{inp}}}$ , that is,  $x_j \in \mathcal{R}^{F_{\text{inp}}}$ ,  $j=1,2,\dots,n$ . The issue is to perform a partition of such collection of elements into  $c$  fuzzy sets with respect to a given criterion, where  $c$  is a given number of clusters. The criterion is usually to optimize an objective function that acts as a performance index of clustering. The end result of fuzzy clustering can be expressed by a partition matrix  $U$  such that :

$$U = [u_{ij}]_i \quad i = 1, \dots, c, \quad j = 1, \dots, n \quad (10)$$

In Eq. (10),  $u_{ij}$  is a numerical value in  $[0,1]$  and expresses the degree to which an element  $x_j$  belongs to the  $i^{\text{th}}$  cluster. However, there are two additional constraints on the value of  $u_{ij}$ . First, a total membership of the element  $x_j \in X$  in all classes is equal to unity; that is :

$$\sum_{i=1}^c u_{ij} = 1 \quad \text{for all } j=1,2,\dots,n \quad (11)$$

Second, every constructed cluster is nonempty and different from the entire set, that is,

$$0 < \sum_{j=1}^n u_{ij} < n \quad \text{for all } i=1,2,\dots,c \quad (12)$$

The general form of the objective function used in fuzzy clustering is

$$J(u_{ij}, v_k) = \sum_{i=1}^c \sum_{j=1}^n \sum_{k=1}^c g[w(x_i), u_{ij}] d(x_j, v_k) \quad (13)$$

where  $w(x_i)$  is a prior weight for each  $x_i$  and  $d(x_j, v_k)$  is the degree of dissimilarity between the data  $x_i$  and the supplemental element  $v_k$ , which can be considered as the central vector of the  $k^{\text{th}}$  cluster. Degree of dissimilarity is defined as a measure that satisfies two assumptions given by

$$d(x_j, v_k) \geq 0, \quad (14)$$

$$d(x_j, v_k) = d(v_k, x_j) \quad (15)$$

Based on the above background, fuzzy clustering can be precisely formulated as an optimization problem:

Minimize

$$J(u_{ij}, v_k) = \sum_{i=1}^c \sum_{j=1}^n \sum_{k=1}^c g [w(x_i, u_{ij})] d(x_j, v_k) \quad (16)$$

$i, k = 1, 2, \dots, c; j = 1, 2, \dots, n$

subject to

$$\sum_{i=1}^c u_{ij} = 1 \quad \text{for all } j = 1, 2, \dots, n \quad \text{and}$$

$$0 < \sum_{j=1}^n u_{ij} < n \quad \text{for all } i = 1, 2, \dots, c$$

One of the widely employed clustering methods based on Eq. (16) is the Fuzzy C-Means (FCM) algorithm. The objective function of the FCM algorithm is expressed in the form of

$$J(u_{ij}, v_k) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|x_j - v_i\|^2 \quad m > 1 \quad (17)$$

where  $m$  is called exponential weight that influences the degree of fuzziness of the membership (partition) matrix. To solve this minimization problem, the objective function  $J(u_{ij}, v_k)$  in Eq. (17) is differentiated with respect to  $v_k$  (for fixed  $u_{ij}$ ,  $i=1, \dots, c$ ,  $j=1, \dots, n$ ) and to  $u_{ij}$  (for fixed  $v_k$ ,  $i=1, \dots, c$ ) and the conditions of Eq. (11), are applied obtaining

$$v_i = \frac{1}{\sum_{j=1}^n (u_{ij})^n} \sum_{j=1}^n (u_{ij})^n x_j, i = 1, 2, \dots, c \quad (18)$$

$$u_{ij} = \frac{\left( \frac{\|x_j - v_i\|^2}{\sum_{k=1}^c \|x_j - v_k\|^2} \right)^{\frac{1}{m-1}}}{\sum_{k=1}^c \left( \frac{\|x_j - v_k\|^2}{\sum_{l=1}^c \|x_j - v_l\|^2} \right)^{\frac{1}{m-1}}}, \quad (19)$$

$i = 1, 2, \dots, c; j = 1, 2, \dots, n$

The system described by the Eqs. (18) and (19) cannot be solved analytically. However, the FCM algorithm provides an iterative approach to approximating the minimum of the objective function starting from a given position.

## 5. Linear State Space Models Extraction

### 5.1 T-S Fuzzy Space Model

At each sample time  $k$ , given an operating point condition (for example  $u(k-1)$  and  $y(k-1)$ ), a local linear fuzzy state-space model can be constructed via calculating the degree of fulfillment  $\mu_i(x(k))$  of the antecedents, using product as the fuzzy logic AND operator. The inference of the entire structure (hierarchy) due to rule  $i$  results in a sub-model (1) which can be expressed as:

$$y_l(k+1) = \frac{\sum_{i=1}^r \mu_{li}[x_l(k)] y_{li}(k+1)}{\sum_{i=1}^r \mu_{li}[x_l(k)]} \quad (20)$$

$$y_{li}(k+1) = [\zeta_{li} y(k) + \eta_{li} u(k) + \theta_{li}] \quad (21)$$

Defining  $\zeta_l$ ,  $\eta_l$  and  $\theta_l$  as follows :

$$\zeta_l = \frac{\sum_{i=1}^r \mu_{li}[x_l(k)] \zeta_{li}}{\sum_{i=1}^r \mu_{li}[x_l(k)]} \quad (22)$$

$$\eta_l = \frac{\sum_{i=1}^r \mu_{li}[x_l(k)] \eta_{li}}{\sum_{i=1}^r \mu_{li}[x_l(k)]} \quad (23)$$

$$\theta_l = \frac{\sum_{i=1}^r \mu_{li}(x_l(k)) \theta_{li}}{\sum_{i=1}^r \mu_{li}(x_l(k))} \quad (24)$$

where  $x(k)$ ,  $u(k)$  and  $y(k)$  for the state-space description are defined as:

$$\mathbf{x}(k) = \begin{bmatrix} x_1(k) \dots x_1(k-n_{y1}) \dots x_{n_o}(k) \\ \dots x_{n_o}(k-n_{yn_o}) \dots u_1(k-1) \dots \\ u_{n_i}(k-n_d-n_u) \end{bmatrix}^T \quad (25)$$

$$\mathbf{u}(k) = [u_1(k) \quad u_2(k) \quad \dots \quad u_{n_i}(k)]^T \quad (26)$$

$$\mathbf{y}(k) = [x_1(k) \quad x_2(k) \quad \dots \quad x_{n_i}(k)]^T \quad (27)$$

In order to employ Quadratic Programming for systems which depend on current as well as on the previous inputs, it is necessary to construct a state-space representation, such that the state vector  $x(k)$  to accommodate not only the state variables, appearing in  $y(k)$ , but also the previous inputs and the offset as last element. This results in a system with only current inputs, but leads to a more complex A-matrix.

The latter contains also  $\eta$  s, corresponding to the previous inputs. If the maximal delay in the input  $i$ ,  $i=1, \dots, n_i$  is  $u_{i,dmax}$ , then the number of the additional columns is  $\sum_{i=1}^{n_i} \max(u_{i,dmax} - 1, 0)$ . In the last column of A are stored the offsets  $\theta$  s. The columns with  $\eta$  s correspond to the previous inputs, stored in the state vector; these columns are not included in B. The ones in A correspond to the delayed values of a certain variable. The local linear system matrices are derived as follows: A is  $\alpha \times \alpha$  square matrix,

where

$$\alpha = \alpha_1 + \sum_{i=1}^{n_i} \max(u_{i,\alpha_{\max}} - 1, 0) + 1,$$

$$\alpha_1 = \sum_{j=1}^{n_{o_j}} \max(n_{y_j} - 1)$$

B is an  $\alpha \times n_i$  and C is a  $n_o \times \alpha$  matrix:

$$A = \begin{bmatrix} \zeta_{1,1} & \zeta_{1,2} & \zeta_{1,3} & \dots & \dots & \dots & \zeta_{1,\alpha_1} & \eta_{1,i} & \dots & \eta_{1,j} & \theta_{1,1} \\ 1 & 0 & 0 & \dots & \dots & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \vdots & \ddots & \dots & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & \ddots & \ddots & \dots & \dots & \vdots & \vdots & \dots & \vdots & \vdots \\ \zeta_{2,1} & \zeta_{2,2} & \dots & \dots & \dots & \dots & \zeta_{2,\alpha_1} & \eta_{2,i} & \dots & \eta_{2,j} & \theta_{2,1} \\ 0 & \vdots & \ddots & \ddots & \dots & \dots & \vdots & \vdots & \dots & \vdots & \vdots \\ \zeta_{n_o,1} & \zeta_{n_o,2} & \dots & \dots & \dots & \dots & \zeta_{n_o,\alpha_1} & \eta_{n_o,i} & \dots & \eta_{n_o,j} & \theta_{n_o,\alpha_1} \\ 0 & \dots & 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \quad (28)$$

$$B = \begin{bmatrix} \eta_{1,1} & \eta_{1,2} & \dots & \eta_{1,n_i} \\ 0 & \dots & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & \dots & \dots & 0 \\ \eta_{2,1} & \eta_{2,2} & \dots & \eta_{2,n_i} \\ \vdots & \ddots & \ddots & \vdots \\ \eta_{n_o,1} & \eta_{n_o,2} & \dots & \eta_{n_o,n_i} \\ \vdots & \ddots & \ddots & \vdots \end{bmatrix} \quad (29)$$

$$C = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 1 & 0 & \dots & 0 \end{bmatrix} \quad (30)$$

The ones in C are positioned such that  $y_j(k) = x_j(k)$ . At any time index  $k$ , initially the control signal  $u(k-1)$  is used. However, after the optimization,  $u(k)$  is available and could be used in next iterations.

## 6. Takagi-Sugeno Fuzzy Model Validation

### 6.1 Correlation Tests

Traditionally more rigorous statistical validation tests are employed in which model residuals are examined, and if found to be sufficiently correlated with a function of the data then the model is inadequate. This is achieved by defining a matrix  $Z$ , where  $Z(x^t)$  is

$$Z(x^t) = [m(t), m(t-1), \dots, m(t-t_d)]^T \quad (31)$$

in which  $x^t$  is the observational vector of inputs, outputs and errors seen up to time step  $t$ , and  $m(t)$  represents the degree of dependency of the two training signals  $y(t)$  and  $u(t)$ , i.e.

$$x^t = (u^{t-1}, y^{t-1}, e^{t-1})^T \quad (32)$$

and  $m(t-1)$  is a monomial of the vector  $x^t$  given by

$$m(t) = y(t-1)u^2(t-2) \quad (33)$$

The following two hypotheses are defined :

$H_o$  :  $e(t)$  is uncorrelated with  $Z(x^t)$ ,  $E(e(t)Z(t)) = 0$  and

$H_1$  :  $e(t)$  is correlated with  $Z(x^t)$ ,  $E(e(t)Z(t)) \neq 0$

where the purpose of validation is to use the data to decide if  $H_o$  holds. Two different test statistics have been proposed in the literature, the most common being the standard sample correlation measure,  $\rho(k)$ , is defined as,

$$\rho(k) = \frac{1/N \sum_{t=1}^{N-K+1} [m(t-k+1)e(t)]}{\sqrt{\sum_{t=1}^{N-K+1} [m(t-k+1)m(t-k+1)] \sum_{t=1}^{N-K+1} [e(t)e(t)]}} \quad (34)$$

in which  $k = 1, \dots, t_d$  and  $\rho(k) \in [-1, +1]$ . If  $H_o$  holds, this statistic asymptotically approaches a normal distribution, and with 95% confidence limits,  $H_o$  is accepted if  $\rho(k) \in [-1.96/\sqrt{N}, 1.96/\sqrt{N}]$ . An alternative statistic is given by:

$$d = N \left[ \sum_{t=1}^{N-t_d} [e^2(t)] \right]^{-1} \left[ \sum_{t=1}^{N-t_d} [Z(t)e(t)] \right] \left[ \sum_{t=1}^{N-t_d} [Z(t)Z(t)] \right]^{-1} \left[ \sum_{t=1}^{N-t_d} [Z(t)e(t)] \right] \quad (35)$$

where  $H_o$  hold is asymptotically a  $X^2(s)$  distribution where  $s$  is the number of delays,  $t_d$ . For a given acceptance level (typically 95%) a critical point is found, and if elements of  $d$  are outside this acceptance region,  $H_o$  is rejected.

## 7. Modeling of a Nonlinear System : A Case Study

### 7.1 Antenna System (Input-Output Training Pattern)

To test these proposed neuro-fuzzy methodologies further, they are applied to model a realistic nonlinear dynamical system. The system considered is a nonlinear (MIMO) dynamics of an antenna system with two coupled inputs

and outputs. A data set containing 500 samples of training patterns were produced by applying random torques to the different channels, with suitable sampling rate and an amplitude drawn from uniform random distribution in the range  $(-1.5, +1.5)$  N/m.

### Antenna system

A coupled two degree of freedom satellite dish, typical of the type used for oceanary satellite communication systems, is presented.

The behavior of the antenna is described by the following nonlinear idealized time invariant State space equations:

$$\mathbf{z} = [\varphi \quad \dot{\varphi} \quad \psi \quad \dot{\psi}]^T \quad (36)$$

$$\dot{\mathbf{z}} = \begin{bmatrix} \frac{T_\varphi - b_\varphi \dot{\varphi} - \dot{\varphi} \dot{\psi} (I' - I) \sin(2\psi)}{I' \sin^2(\psi) + I \cos^2(\psi)} \\ \dot{\varphi} \\ \frac{T_\psi - b_\psi \dot{\psi} - \frac{1}{2} (I' - I) \dot{\varphi}^2 \sin(2\psi)}{I} \\ \dot{\psi} \end{bmatrix} \quad (37)$$

and

$$\mathbf{y} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{z} + \begin{bmatrix} e_\varphi(t) \\ e_\psi(t) \end{bmatrix} \quad (38)$$

where  $\varphi$  is the azimuth angle,  $\psi$  is the elevation angle,  $b_\varphi$  and  $b_\psi$  are the associated friction coefficients, and  $T_\varphi$  and  $T_\psi$  are the torques applied to the axes. To produce a more realistic simulation, the outputs are corrupted by additive Gaussian noise,  $[e_\varphi(t) \quad e_\psi(t)]^T$ , representing a crude approximation to measurement noise. The azimuth is permitted to turn through a complete revolution, while end stops restrict the elevation to the interval  $[0, \pi]$ . In this antenna there are essentially two sources of nonlinearity: that produced from the end stops on the elevation and the other as a results of the non-isotropic moment of inertia tensor. Indeed when isotropy is present the state-space equations (above) are linear. The strength of this non-linearity depends on the degree of anti-isotropy and the angular velocities of the antenna. These torques are chosen to emulate typical operating conditions. Such block diagram used to produce the identification data was simulated through SIMULINK/MATLAB, using a set of nonlinear differential equations that describe the antenna system. Half of the training pattern was used in the modeling of the dynamic system whereas the other half was used to validate the fuzzy models resulting from the modeling. For a typical sequence of training data, such responses of the antenna inputs-outputs is shown in Fig. 3.

### 7.2 Training Pattern and Clustering

As discussed in section (3), fuzzy modeling of any dynamical system could be achieved through clustering the training data. In this respect, Fig. 3 shows the employed Input-Output data training pattern. For this simulation example, clustering has been applied to the antenna training pattern. In Fig. 4 it is shown the training pattern following applying the clustering algorithm, where it illustrates clearly the clusters and their three associated centers is shown in Fig 4. For instant, the figure shows the training pattern which has been clustered into three. To reduce the fuzzy rules while preserving the model accuracy, the number of clusters were chosen to be three clusters. The fuzziness parameter  $m$  was kept at 2.2 with a termination criterion  $\epsilon = 0.01$ . The result of the clustering algorithm is the fuzzy partition matrix and the cluster centers matrix, which will be used to construct the fuzzy model for the antenna system.

### 7.3 Neuro-Fuzzy Modeling

Neuro-fuzzy modeling is applied to the problem of identifying a discrete model of the antenna. A fuzzy model can be constructed from data by using the output of the clustering algorithm and by constructing regressors to form inputs to the neuro-fuzzy network. Hence a conventional linear difference model with regressors is constructed containing previous inputs and outputs, *i.e.*

$$\varphi(k) = \begin{bmatrix} T_\varphi(k-1), T_\varphi(k-2), T_\psi(k-1) \\ \varphi(k-1), \varphi(k-2) \\ \psi(k-1), \psi(k-2), \Delta\varphi(k-1) \end{bmatrix}^T \quad (39)$$

$$\psi(k) = \begin{bmatrix} T_\varphi(k-1), T_\psi(k-1), T_\psi(k-2) \\ \varphi(k-1), \varphi(k-2) \\ \psi(k-1), \psi(k-2), \Delta\psi(k-1) \end{bmatrix}^T \quad (40)$$

Fuzzy IF-THEN rules can be extracted by projecting the clusters onto the axes and the membership functions of the fuzzy sets generated by pointwise projection of the partition matrix onto the antecedent variables. Then consequent parameters for each rule are obtained as least squares estimates. When an initial structure is obtained through clustering, the membership functions and the consequent parameters are tuned to satisfy certain cost function through the learning procedure of the neuro-fuzzy.

### 7.4 Membership Functions and Associated Fuzzy Rules

As a result, the membership functions of all the inputs (regressors) and outputs are shown in Fig. 5 for azimuth angle. The antenna system has six inputs (*in terms of fuzzy model*) and two outputs, hence two groups of seven sets of MFs are shown. Each universe of discourse (set) has three

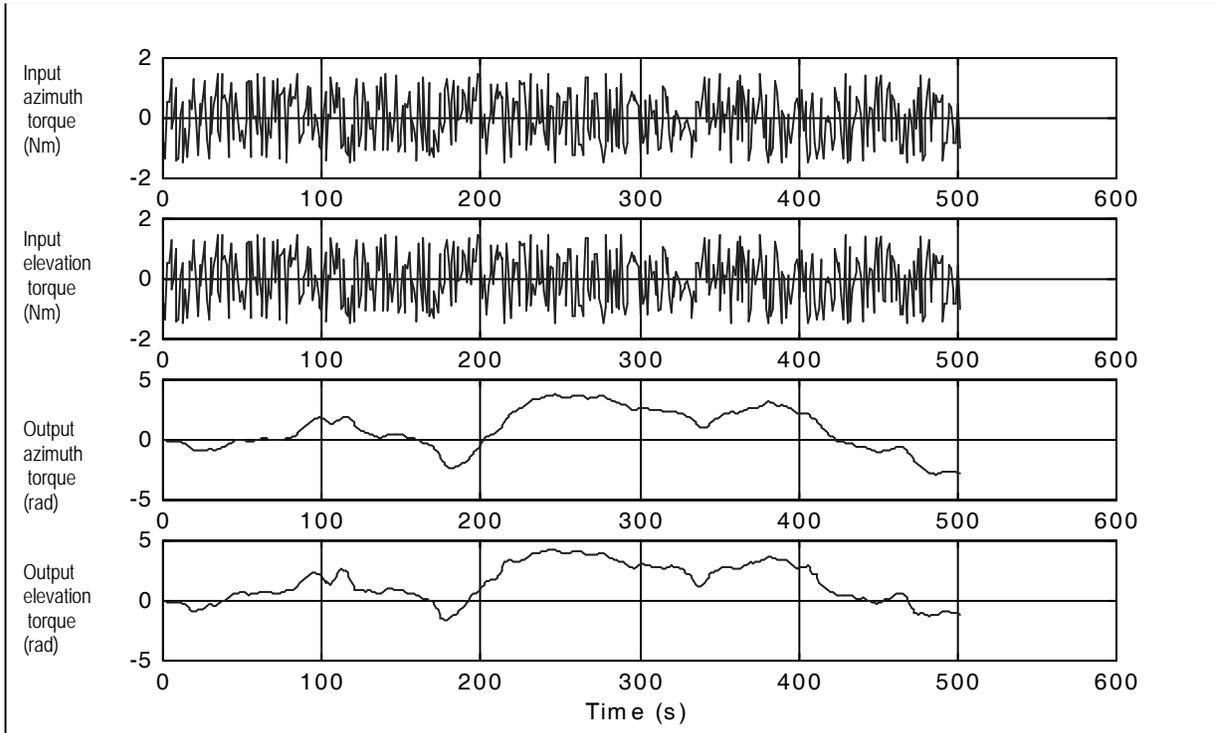


Figure 3. Input-Output data training pattern

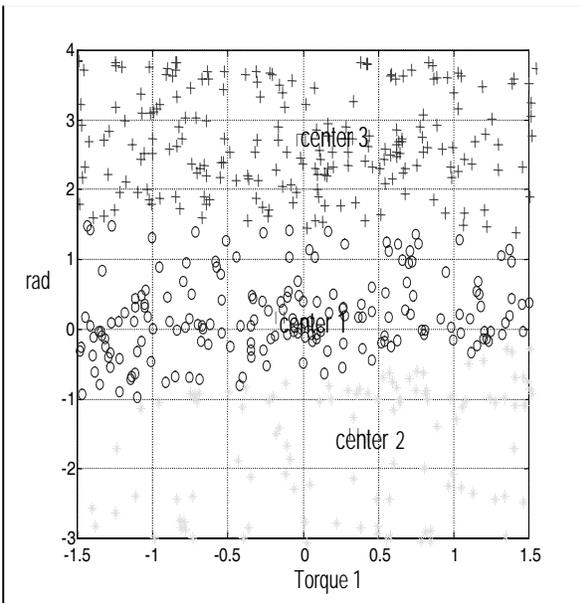


Figure 4. Extracted clusters functions associated azimuth angle

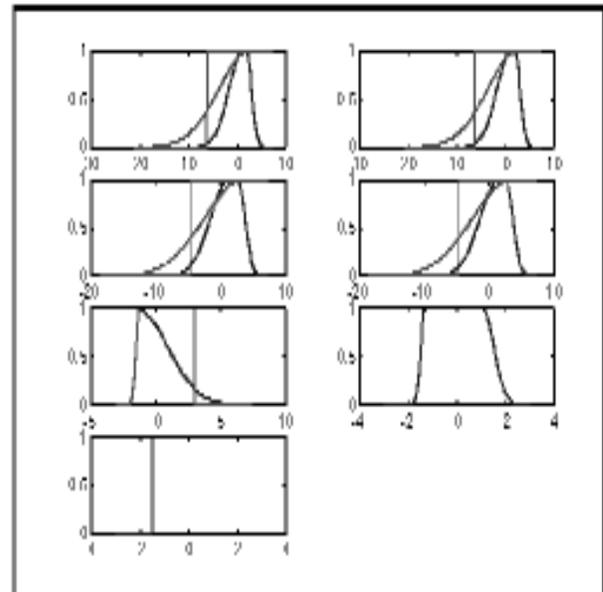


Figure 5. Extracted membership functions of all inputs

MFs representing the assigned three clusters. Such memberships are representing the range of the inputs.

### 7.5 Fuzzy Sub-models:

The T-S fuzzy model presented has been used to identify the nonlinear antenna system. As was mentioned before, the number of rules in the T-S fuzzy model equals the number of clusters in the product space. The consequent of each rule is a local model that approximates the output of the real function for the range of  $x$  for which the rule is applicable. As a result of the modeling development, the following rules are obtained for azimuth and elevation angles :

Rule 1: IF  $T_\varphi(k-1)$  is  $F_1^1$  and  $T_\varphi(k-2)$  is  $F_2^1$  and  $T_\psi(k-1)$  is  $F_3^1$  and  $\varphi(k-1)$  is  $F_4^1$  and  $\varphi(k-2)$  is  $F_5^1$   $\psi(k-1)$  is  $F_6^1$  and  $\psi(k-2)$  is  $F_7^1$  then  $\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{B}_1\mathbf{u}(t)$  and  $\mathbf{y}(t) = \mathbf{C}_1\mathbf{x}(t)$ .

Rule 2: IF  $T_\varphi(k-1)$  is  $F_1^2$  and  $T_\varphi(k-2)$  is  $F_2^2$  and  $T_\psi(k-1)$  is  $F_3^2$  and  $\varphi(k-1)$  is  $F_4^2$  and  $\varphi(k-2)$  is  $F_5^2$   $\psi(k-1)$  is  $F_6^2$  and  $\psi(k-2)$  is  $F_7^2$  then  $\dot{\mathbf{x}}(t) = \mathbf{A}_2\mathbf{x}(t) + \mathbf{B}_2\mathbf{u}(t)$  and  $\mathbf{y}(t) = \mathbf{C}_2\mathbf{x}(t)$ .

Rule 3: IF  $T_\varphi(k-1)$  is  $F_1^3$  and  $T_\varphi(k-2)$  is  $F_2^3$  and  $T_\psi(k-1)$  is  $F_3^3$  and  $\varphi(k-1)$  is  $F_4^3$  and  $\varphi(k-2)$  is  $F_5^3$   $\psi(k-1)$  is  $F_6^3$  and  $\psi(k-2)$  is  $F_7^3$  then  $\dot{\mathbf{x}}(t) = \mathbf{A}_3\mathbf{x}(t) + \mathbf{B}_3\mathbf{u}(t)$  and  $\mathbf{y}(t) = \mathbf{C}_3\mathbf{x}(t)$

where,

$$\mathbf{A}_1 = \begin{bmatrix} 0.09510 & -0.4265 & 0.0300 & -0.2998 & 0.0116 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0.0348 & -0.0273 & 0.8022 & -0.3315 & 0 & 0.0013 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{A}_2 = \begin{bmatrix} 0.9354 & -0.4667 & 0.0525 & -0.0475 & 0.0119 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0.0373 & -0.0384 & 0.0844 & -0.3715 & 0 & -0.0002 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{A}_3 = \begin{bmatrix} 0.9020 & -0.4099 & 0.0142 & -0.0157 & 0.0108 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0.1294 & -0.1098 & 0.7887 & -0.2987 & 0 & 0.0002 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{B}_1 = \begin{bmatrix} 0.0128 & 0 \\ 0 & 0 \\ 0 & 0.0240 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 0.0128 & 0 \\ 0 & 0 \\ 0 & 0.0137 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

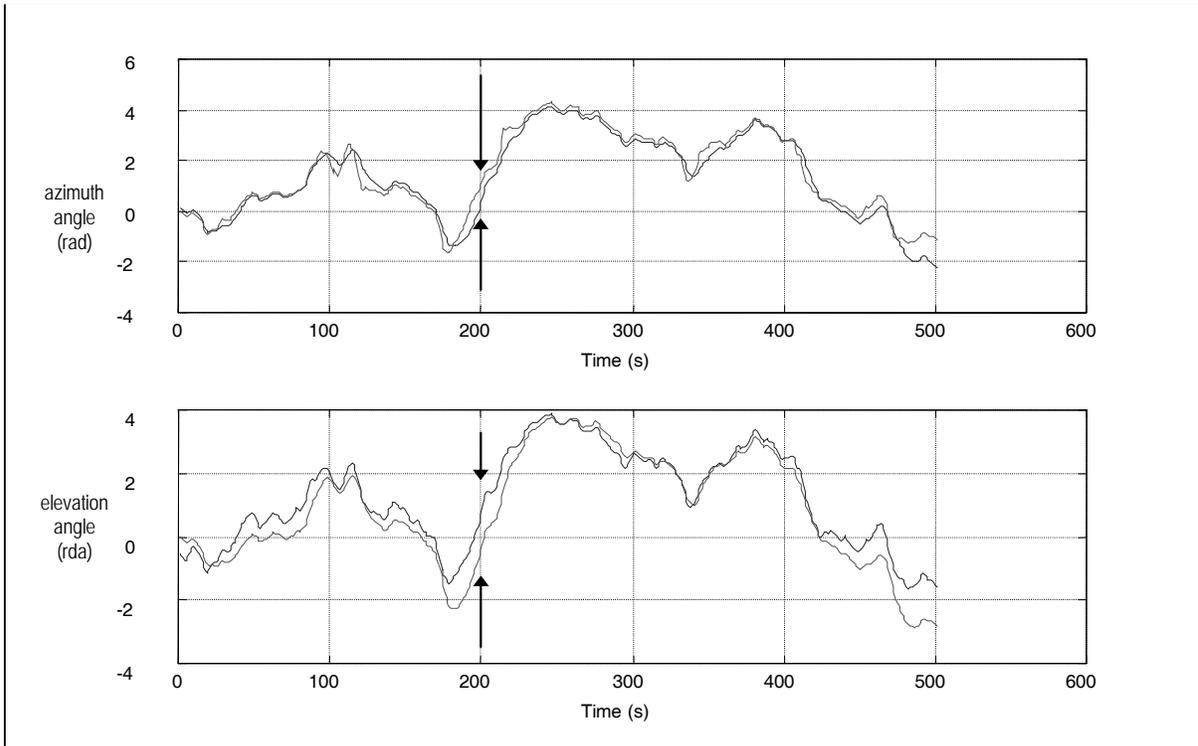
$$\mathbf{B}_3 = \begin{bmatrix} 0.0118 & 0 \\ 0 & 0 \\ 0 & 0.0028 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{D}_1 = \mathbf{D}_2 = \mathbf{D}_3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Here the C and D matrices are common for all of the three fuzzy sub-models, and the D matrix is equal to zero. Furthermore, the elevation angle dynamics is of the same above structure. The antenna simulation system incorporating the three models are shown in Fig. 6. Consequently, Fig. 6 shows the actual antenna output superimposed over the evaluated fuzzy model output. From the figure, it is apparent how the fuzzy model output resembles the actual system output.

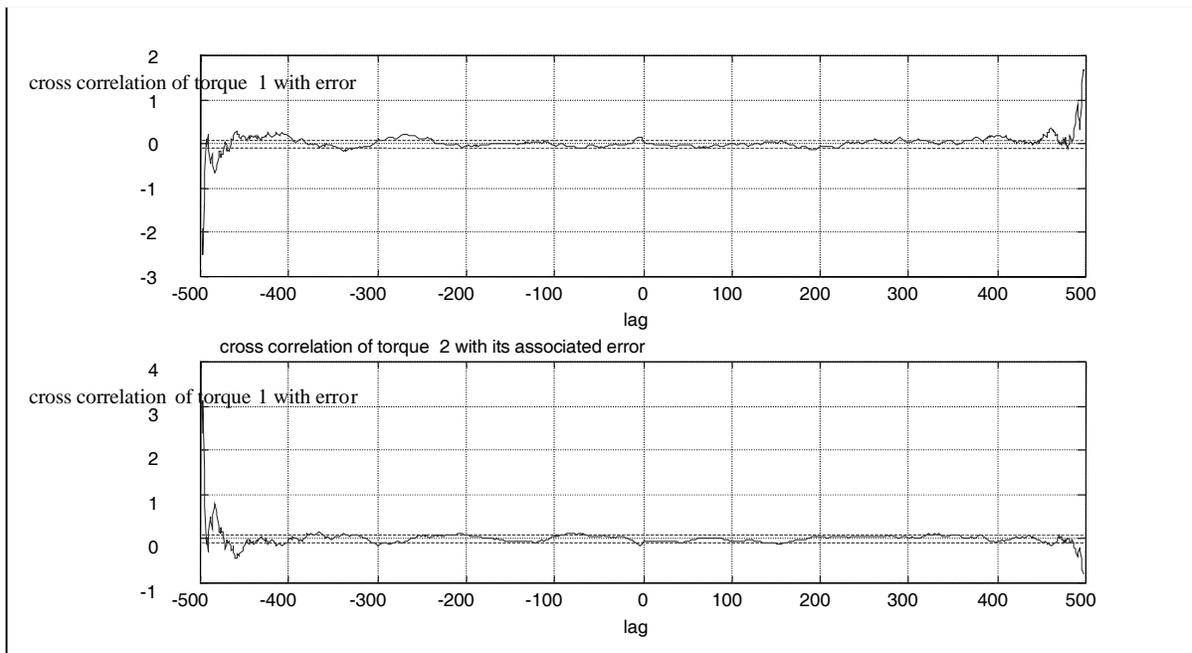
### 7.6 Fuzzy Sub-models Validation

Figure 7 displays the cross-correlation function of the error signal with the first input signal of the antenna. The correlation in the figure is within the confidence interval, which indicates that the two signals are not correlated. To further investigate the constructed local linear sub-models of the antenna, Figure 8 shows the attained linearized sub-models over the antenna time response. In terms of antenna nonlinear behavior, it is obvious that the entire operating region has been sub-divided into a number of local models which could be employed for further control synthesis. From the shown antenna response, fuzzy models are useful for describing the antenna dynamics where

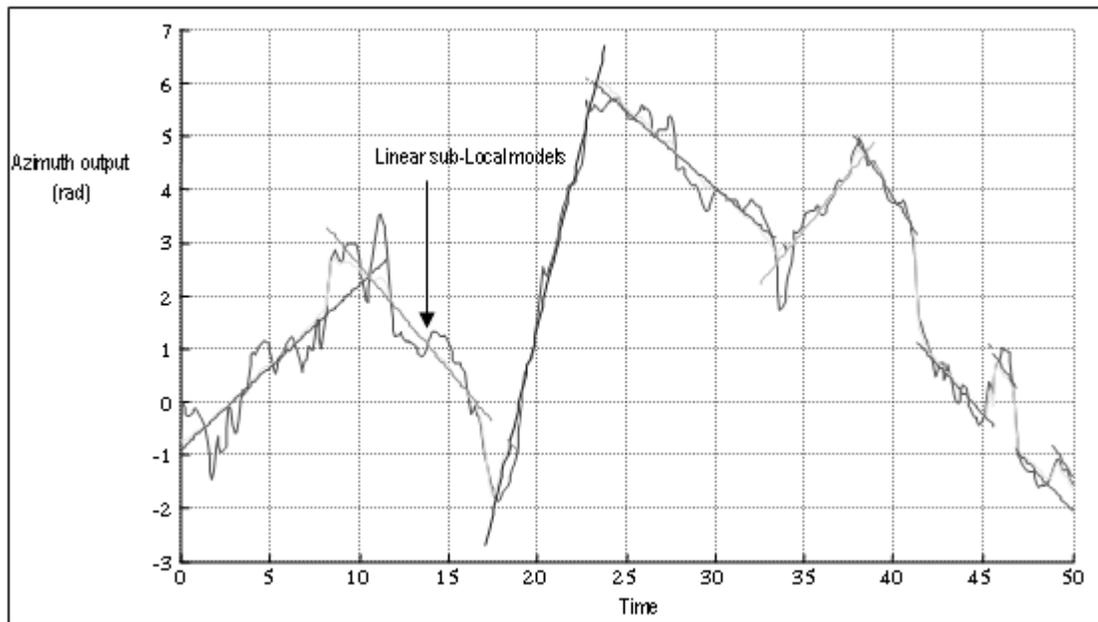
the underlying physical mechanisms are not completely known and the antenna behavior is understood in qualitative terms. Consequently, an important property of fuzzy models is their capability to represent nonlinear dynamic systems. Therefore, the obtained fuzzy sub-models can also be applied to systems that are well understood but due to the nonlinearities untraceable with standard linear methods. Rule-based structure of fuzzy models allows for integrating heuristic knowledge with information obtained



**Figure 6. Fuzzy model responses (azimuth and elevation angles) compared to the antenna outputs**



**Figure 7. Cross correlation of system first input with its associate error of the antenna system**



**Figure 8. Obtained linearized antenna sub-models**

from antenna measurements. The global operation of the antenna nonlinear process is divided into several local operating regions. Within each region  $R_i$ , a reduced order linear model in ARMAX form, is used to represent the local antenna behavior. That was not restrictive, and any appropriate model forms can also be used.

## 8. Conclusions

This article has concentrated on the modeling of nonlinear dynamic systems via the utilization of the well known fuzzy modeling paradigm, the Takagi-Sugeno (T-S) technique. T-S models depend heavily on some initial membership centers of the universe of discourse of used fuzzy variables, such centers have been obtained by employing clustering algorithm. Once such centers are computed, a fuzzy system can establish initial membership centers through which they are updated via a neural network learning mechanism. One of the advantages of T-S modeling is that systems can be modeled by few rules, and consequently fewer linear sub-models. This advantage has overcome the problem of the large number of rules in the fuzzy modeling. Fuzzy models have also been verified and validated through some standard validation techniques, where they have shown clearly the successful ability of T-S techniques to model nonlinear systems with a good degree of accuracy.

## References

- Akkizidis, S. and Roberts, N., 2001, "Fuzzy Clustering Methods for Identifying and Modeling of Non-Linear Control Strategies": Proceedings of the Institution of Mechanical Engineers. Part I: *Journal of Systems and Control Engineering*, Vol. 1(215), pp. 437-452.
- Babuska, R. and Verburggen, H., 1995, "Identification of Composite Linear Models via Fuzzy Clustering," Proceedings of the European Control Conference, Italy, pp. 1207-1212.
- Bologna, G., 2001, "A New Neuro-Fuzzy Model," *Proceedings of the International Joint Conference on Neural Networks (IJCNN'01)*, Vol. 2, Washington-DC (USA), pp. 1328-1333.
- Bossley, K., 1997, "Neuro-Fuzzy Modeling Approaches in System Identification," Ph.D. Thesis, University of Southampton, U.K.
- Gorzalczany, E., Marian, B. and Gluszek, A., 2000, "Neuro-Fuzzy Networks in Time Series Modeling," *Proceedings of the International Conference on Knowledge-Based Intelligent Electronic Systems*, Vol. 1, Brighton, (UK), pp. 450-453.
- Ikonen, E., Najim, K. and Kortela, U., 2000, "Neuro-Fuzzy Modelling of Power Plant Flue-Gas Emissions," *J. of Engineering Applications of Artificial Intelligence*, Vol. 3(6), pp. 705-717.
- Lo, J. and Chen, Y., 1999, "Stability Issues on Takagi-Sugeno Fuzzy Model, Parametric Approach", *IEEE Transaction of Fuzzy Systems*, Vol. 7(5), pp. 597-607.
- Mamdani, E. and Assilian, S., 1975, "An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller," *International J. of Man-Machine Studies*, Vol. 3(5), pp. 1-13.
- Min-You, Chen. and Linkens, A., 1998, "Fast Fuzzy Modeling Approach Using Clustering Neural Networks," *IEEE International Conference on Fuzzy Systems*, Vol. 1(2), IEEE World Congress on Computational Intelligence, pp. 1088-1093.

- Ning, L., Shaoyuan, L. and Yugeng, X., 2001, "Modeling pH Neutralization Processes Using Fuzzy Satisfactory Clustering," IEEE International Conference on Fuzzy Systems, Vol. 1(1), pp. 308-311.
- Takagi, T. and Sugeno, M., 1985, "Fuzzy Identification of Systems and its Application to Modeling and Control," IEEE Transaction Systems, Man and Cybernetics, Vol. 15(1), pp. 116-132.
- Wang, H., Li, J., Niemann, D. and Tanaka, K., 2000, "T-S Fuzzy Model with Linear Rule Consequence and PDC Controller: A Universal Framework for Nonlinear Control Systems," Proceedings of the 9th IEEE International Conference on Fuzzy Systems.
- Zhang, J. and Knoll, A., 1999, "Modelling Multivariate Data by Neuro-Fuzzy Systems", Proceedings of the 1999 IEEE/IAE Conference on Computational Intelligence for Financial Engineering (CIFEr), New York, (USA), pp. 267-270.





