View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Biblioteca Digital do IPB

Decomposition of a Greenhouse TS-Fuzzy Model by Clustering Process

Paulo Salgado* and Getúlio Igrejas**

* Departamento de Engenharias
CETAV – Universidade de Trás-os-Montes e Alto Douro
Quinta de Prados – Vila Real
Telf: +351 259 350 359; fax: +351 259 350 480; e-mail: psal@utad.pt

** Departamento de Electrotecnia
ESTiG – Instituto Politécnico de Bragança
Campus de Sta. Apolónia – Bragança
Telf: +351 273 303 134; fax: +351 273 313 051; e-mail: igrejas@ipb.pt

Abstract **— This paper presents a fuzzy c-means clustering method for decompose a T-S fuzzy system. This technique is used to organize the fuzzy greenhouse climate model into a new structure more interpretable, as in the case of the physical model. This new methodology was tested to split the inside greenhouse air temperature and humidity flat fuzzy models into fuzzy sub-models. These fuzzy sub-models are compared with its counterpart's physical sub-models. This algorithm is applied to the T-S fuzzy rules. The results are several clusters of rules where each cluster is a new fuzzy sub-system. This is a generalized Probabilistic Fuzzy C-Means (PFCM) algorithm applied to TS-Fuzzy System clustering. This allows automatic organization of one fuzzy system into a multimodel Hierarchical Structure.**

1. Introduction

The greenhouse climate model describes the dependence of temperature, air humidity and CO_2 concentration inside the greenhouse on the outside weather conditions and on the control equipment using a set of nonlinear differential equations of first order. These equations are formulated as result of a balance of energy and mass of many physical and biological processes.

However, many processes related to the climate of the greenhouse are difficult to describe mathematically, especially when the structure of the system and the relationship between variables is unknown or too complex, giving rise to intractable mathematical expressions. Other strategies called intelligent could also be valuable alternatives. This is the case of Fuzzy modelling. Combining the well-established learning techniques with the proposed methodology enables the ability to learn from real world observation and describes naturally the behaviour as a series of understandable linguistically human rules. Despite this effort, it is not possible to guarantee a direct relation between the linguistic information and the physical processes.

Clustering methods seeks to organize a set of items into clusters such that items within a given cluster have a high degree of similarity, whereas items belonging to different clusters have a high degree of dissimilarity [1] [2]. These methods have been widely applied in various areas such as taxonomy, image processing, information retrieval, data mining, etc [3][8]. Clustering techniques may be divided into hierarchical and partitioning methods: hierarchical methods yield complete hierarchy, i.e., a nested sequence of partitions of the input data, whereas partitioning methods seek to obtain a single partition of the input data in a fixed number of clusters, usually by optimizing an objective function. However, the generalization of these techniques to clustering imprecisely or uncertainly data or objects is not yet explored. Recently, fuzzy set theory is more and more frequently used in intelligent systems, because of its simplicity and similarity to human reasoning.

This work addresses this fundamental goal of fuzzy modelling by using an algorithm that implements Fuzzy Clustering of Fuzzy Rules (FCFRA) applied to the climate greenhouse model [4]. The proposed algorithm allows the decomposition of the fuzzy relation into sub-relations, through a process of unfolding the fuzzy rules. The obtained sub-rules are then grouped into v subgroups (clusters), by similarity association. As a result, the original fuzzy relation is layered through the v levels of the hierarchical fuzzy system [6]. In other words, the application of the FCFRA to the clustering of a flat fuzzy system leads to the distribution of the information carried by the system among various layers of a hierarchical collaborative structure (HCS) [5].

2. The Probabilistic Clustering Algorithm of T-S Fuzzy System

A fuzzy rule-based model suitable for describing a large class of nonlinear systems was introduced by Takagi and Sugeno [9][10] as follows:

$$R_l : IF\ x_1 \text{ is } A_1^l \text{ and } \cdots \text{ and } x_n \text{ is } A_n^l\ THEN\ y^l = f_l(\boldsymbol{x}, \boldsymbol{a}) \quad (1)$$

where l=1, 2, . . . , M, R^l denotes the l^{th} IF-THEN rule and M is the number of rules in the rule base. x_i, i = 1, ... , n, are individual input variables, and A_i^l are the associated individual antecedent fuzzy sets of each input variable. $y^l \in R$ is the output of each rule and a the vector of parameters of the nonlinear function f. As a special case of f, we have a polynomial function, where a's are the polynomials coefficients.

For any input vector, $\boldsymbol{x}=[x_1,\cdots,x_n]^T$, if the singleton fuzzifier, the product fuzzy inference and the centre average defuzzifier are applied, the output of the fuzzy model $\hat{y}$ is inferred as follows [11]:

$$\hat{y}=\sum_{l=1}^{M}\mu^l(x)y^l \Big/ \sum_{l=1}^{M}\mu^l(x) \quad (2)$$

where $\mu^l(x)=\prod_{i=1}^{n} A_i^l(x_i)$.

The objective of fuzzy clustering partition is to separate a set of fuzzy rules $\Im$={R_1, R_2,..., R_M} in c clusters in the antecedent space and e clusters in the consequent space, according to a "similarity" criterion. This process allows finding the optimal clusters centers V in the input space, the polynomial prototype Z at output space, the partition matrix, U, of combined input-output partition and the matrix W of scalars values. Each value u_{ijk} represents the membership degree of the k^{th} rule, R_k, belonging to the i^{th} cluster of the input space and j^{th} cluster of the output space. w_{jk} is a value that express the translation of the consequent of the k^{th} rule fuzzy sets in direction of the center of j^{th} the output center of cluster. So, the projection of y^l in the cluster j is the function y_j^l, with $y_j^l = w_{jl} y^l$ and is expectable that:

$$\sum_{j=1}^{e} w_{jl} = 1 \,, \quad l = 1,\cdots,M \quad (3)$$

with $w_{jl} \in \mathbb{R}$.

Let $x_k \in S$ be a point covered by one or more fuzzy rules. Naturally, the membership degree of point x_k belonging to $(ij)^{th}$ cluster is the sum of products between the relevance of the rules l in x_k point and the membership degree of the rule l belonging to cluster ij, u_{ijl}, for all rules, *i.e.*:

$$\sum_{i=1}^{c}\sum_{j=1}^{e}\sum_{l=1}^{M} u_{ijl} \cdot \Re_l(\vec{x}_k) = 1 \,, \quad \forall x_k \in S \quad (4)$$

where $\Re_l(\boldsymbol{x})$ represents the relevance function of the l^{th} fuzzy subsystem covering the point $\boldsymbol{x}$ of the Universe of Discourse.

The rule decomposition into $c \times e$ sub-relations will lead to an output fuzzy set decomposition as well.

For the Fuzzy Clustering of Fuzzy Rules Algorithm (FCFRA) the objective is to find U=[u_{ijl}], $\boldsymbol{V}=[v_1,\cdots,v_c]\in R^{n\times c}$ and $\boldsymbol{Z}=[z_1,\cdots,z_e]\in R^e$ where:

$$J=\sum_{k=1}^{n}\sum_{i=1}^{c}\sum_{j=1}^{e}\sum_{l=1}^{M}\left(u_{ijl}\Re_l(\boldsymbol{x}_k)\right)^m\left(\|\boldsymbol{x}_k-\boldsymbol{v}_i\|^2+\|f_l(\boldsymbol{x}_k)\cdot w_{jl}-\boldsymbol{z}_j\|^2\right) \quad (5)$$

is minimized, with a weighting constant m > 1, with equation (3) and (4) as a constraint. z_j is the prototype function of the j^{th} cluster in output space, here considered to be of polynomial type of order one:

$$z_j(\boldsymbol{x}_k)=[1\ \boldsymbol{x}_k]^T\boldsymbol{z}_j=\boldsymbol{x}_k^T\boldsymbol{z}_j \quad (6)$$

It can be shown that the following algorithm may lead the triplet (U*,V*,W*) to a minimum. The results can be expressed by the following algorithm:

Probabilistic Fuzzy Clustering Algorithm of Fuzzy Rules (FCAFR):

Step 1– For a set of points X={x_1,..., x_n}, with $x_i \in S$, and a set of rules $\Im$={R_1, R_2,..., R_M}, with relevance $\Re_l(\boldsymbol{x}_k)$, k = 1, … , M, keep c, 2 ≤ c < np, and initialize $U(0) \in M_{fcm}$.

Step 2– On the r^{th} iteration, with r = 0, 1, 2, ... , compute the c mean vectors.

$$v_i^{(r)}=\frac{\sum_{k=1}^{n}\left(\sum_{l=1}^{M}U_{il}^m\cdot\Re_l^m(x_k)\cdot x_k\right)}{\sum_{k=1}^{n}\left(\sum_{l=1}^{M}U_{il}^m\cdot\Re_l^m(x_k)\right)} \quad (7)$$

where $U_{il}^m=\sum_{j=1}^{e}u_{ijl}^m$, i =1, 2, ... , e and.

Step 3– Compute the new partition matrix $U(r+1)$ using the expression:

$$u_{ijl}^{(r+1)} = 1 \Bigg/ \sum_{r=1}^{c}\sum_{s=1}^{e}\left(\frac{\sum_{k=1}^{n}\Re_l^m(\boldsymbol{x}_k)\cdot D_{ijlk}}{\sum_{k=1}^{n}\Re_l^m(\boldsymbol{x}_k)\cdot D_{rslk}}\right)^{\frac{1}{m-1}} \quad (8)$$

where $D_{ijlk}=\|\boldsymbol{x}_k-\boldsymbol{v}_i\|^2+\|f_l(\boldsymbol{x}_k)\cdot w_{jl}-\boldsymbol{z}_j\|^2$, with 1 ≤ i ≤ c, 1 ≤ l ≤ M.

Step 4 – Compute the new partition matrix $W(r+1)$ with the expression:

$$w_{jl}^{(r+1)}=\left(1-\hat{\theta}_l^T\sum_{r=1}^{e}V_r\right)\Bigg/\sum_{r=1}^{e}\left(\frac{\overline{U}_{jl}}{\overline{U}_{rl}}\right)^m+\hat{\theta}_l V_j \quad (9)$$

with $\bar{U}_{jl}^{m} = \sum_{i=1}^{c} u_{ijl}^{m}$, $\hat{\theta}_l = \theta_l \Big/ \left(\theta_l^T \theta_l\right)$ and $\theta_l = \sum_{k=1}^{n} \Re_l^m(\boldsymbol{x}_k) f_l(\boldsymbol{x}_k)$.

Step 5 – Compute z_j with:

$$\boldsymbol{z}_j^{(r+1)} = \frac{\sum_{l=1}^{M}\left[U_{jl}^{m} \cdot w_{jl}\left(\sum_{k=1}^{n}\Re_l^m(\boldsymbol{x}_k)\cdot f_l(\boldsymbol{x}_k)\cdot\boldsymbol{x}_k\right)\right]}{\sum_{l=1}^{M}\left[U_{jl}^{m}\cdot\left(\sum_{k=1}^{n}\Re_l^m(\boldsymbol{x}_k)\cdot\boldsymbol{x}_k^T\cdot\boldsymbol{x}_k\right)\right]} \tag{10}$$

where $\bar{\Re}_l^m = \sum_{k=1}^{n} \Re_l^m(x_k)$.

Step 6– If || $U(r+1)$-U(r)|| < ε then the process ends. Otherwise let $r = r + 1$ and go to step 2.

More details about this method can be found in [7].

3. Experimental Results

In this section the results achieved with the proposed greenhouse climate model strategies are presented: physical sub-models and the clustered fuzzy models. The identification of both models was done by using input–output data previously collected. The physical model was built by adding the contributions from the different heat mechanisms involved in the greenhouse climate behaviour. This model is used as a comparison to the clustered fuzzy models.
The greenhouse climate model describes the dynamic behaviour of the stated variables using differential equations for the air temperature, humidity and CO_2 concentrations that results from a combination of the various physical processes involving heat and mass transfer taking place in the greenhouse and from the greenhouse to the outside air. In this paper, only the temperature variables presented in Figure 1 are considered.
The presented physical model has as the main objective of validating the proposed fuzzy model and the performance of the FCFRA algorithm. The model considers a greenhouse divided into the control system devices (heating and ventilation), the covering material, the air inside, the outside space and the soil beneath the greenhouse. These components are characterized by the heating pipe water temperature, T_{heat}, for the heating system; inside air temperature, T_{in}; outside air temperature, T_{out} and the soil temperature, T_{soil}. Other fundamental measured variables of the process are the solar radiation, *Rad* [W/m^{-2}], and two main control signals (control variables): heating ($0 \le u_{heat} \le 1$) and ventilation ($0 \le u_{vent} \le 1$), corresponding to a range from 0 to 100% of the actuator nominal power.
The energy balance in the greenhouse air is affected by the energy supplied by the heating system, $Q_{T,heat}$ [Wm^{-2}], the energy losses to the outside air due to the transmission through the greenhouse cover and the forced ventilation exchange, $Q_{T,out}$ [Wm^{-2}], the energy exchange with the soil $Q_{T,soil}$ [Wm^{-2}] and by the heat supplied by Sun's radiation, $Q_{T,rad}$ [Wm^{-2}]. These fluxes obeyed to the following relations:

- Sunlight radiant energy:

$$Q_{T,Rad} = c_{rad} \cdot Rad \tag{11}$$

where the coefficient c_{rad} denotes the vulgar fraction of the solar radiation transformed into the heat.

- Energy losses to the outside air:

$$Q_{T,out} = \left(c_{vent}\phi_{vent} + c_{leak,roof}\right)\cdot\left(T_{in} - T_{out}\right) \tag{12}$$

where $c_{leak,roof}$ is the heat transfer coefficient through greenhouse cover.
The ventilation flux, ϕ_{vent}, is the sum of natural ventilation ($\phi_{leak,vent}$) with the forced ventilation ($c_{force,vent} \cdot u_{vent}$), i.e., $\phi_{vent} = c_{force,vent} \cdot u_{vent} + \phi_{leak,vent}$ is c_{vent} is the heat capacity per unit volume of air.

- Energy transferred from soil:

$$Q_{T,soil} = c_{soil} \cdot \left(T_{soil} - T_{in}\right) \tag{13}$$

where c_{soil} is an heat transfer coefficient.

In the inside volume, a heat balance equation can be written and solved taking in account the different heat fluxes present here. From this analysis results the following deferential equation:

$$\frac{dT_{in}}{dt} = \frac{1}{C_{cap,T}}\left(Q_{T,heat} - Q_{T,out} + Q_{T,soil} + Q_{T,rad}\right) \tag{14}$$

where, $C_{cap,T}$ [J.m^{-2}.ºC^{-1}] is the heat capacity of the greenhouse air.

A more detailed description of the greenhouse installation, physical process and models can be found in [5].
The first task is to identify a fuzzy system that can match all the *N* pairs of collected data to a given level of accuracy. The training data is a collection of data (of inside and outside physical variables) recorded from a long period (3 months) with a sampling interval of 1 minute. The identification process was performed using triangular membership functions with 4×4×4×4 fuzzy sets. A set of 158 useful fuzzy rules were generated for the fuzzy Temperature model. The simulated and measured air temperatures are compared during a test period. These curves show a good agreement between measured data and the models output, with negligible error.
In Figure 1, the heat fluxes computed with the physical and the clustered fuzzy sub-models are plotted.
The natural heat leakage responses, $Q_{out,leak} = \left(c_{vent}\phi_{leak,vent} + c_{leak,roof}\right)\left(T_{in} - T_{out}\right)$ (without forced

ventilation) is plotted in (a); the heat input due to the sunlight radiation, $Q_{T,rad}$ in (b); the heat supplied by the heating system, $Q_{T,heat}$ in (c) and the heat exchanges between the inside air and the soil, $Q_{T,soil}$, in (d). These results show a good agreement between the clustered fuzzy model and its physical sub-models counterparts.

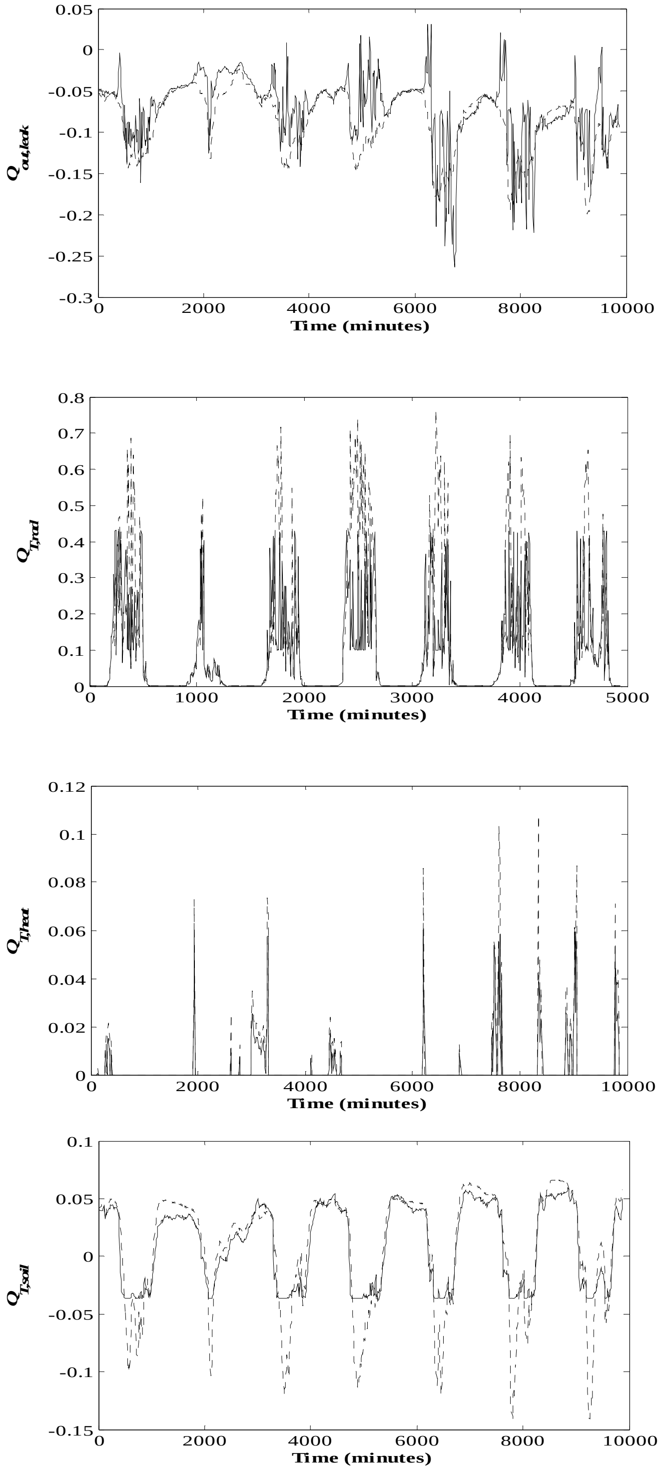

Fig .1. Heat fluxes computed with physical model (dotted line) and HCS fuzzy sub-model (full line).

4. Conclusion

This work has shown that the clustering strategy is capable to organise the fuzzy models in way to decompose its original structure into a hierarchical (collaborative) way. The result is a set of TSK- fuzzy sub-models that can reflect some individual relationships. In this article an application to the climate temperature fuzzy modelling of an agricultural greenhouse has been shown. The fuzzy clustering has been tested to split the inside greenhouse air temperature flat fuzzy model into fuzzy TSK sub-models. These sub-models have a similar counterpart on the physical model, representing the contributions of the process mechanisms involved in the global system dynamics.

Acknowledgement

This work was supported by Fundação para a Ciência e Tecnologia (FCT) under grant POSI/SRI/41975/2001 and by CITAB - Centro de Investigação e Tecnologias Agro-Abientais e Biológicas.

References

[1] Jain *et al.*, 1999 A.K. Jain, M.N. Murty and P.J. Flynn, “Data clustering: A review”, *ACM Comput. Surveys* 31 (3), pp. 264–323.

[2] Dunn, J.C., (1974). “A fuzzy relative of the isodata process and its use in detecting compact, well separated clusters”, *J. Cybernet.* 3, 95–104.

[3] Bezdek, J.C.(1981). Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, NY.

[4] Salgado, Paulo, (2005a). “Clustering and hierarchization of fuzzy systems”, *Soft Computing Journal*, Vol. 9, nº 10, pp. 715-731, October, 2005, Springer Verlag.

[5] Salgado, P. & Boaventura, J. (2005b). “Greenhouse climate hierarchical fuzzy modelling”, *Control Eng. Practice*, 13 pp. 613-628.

[6] Salgado, P. (2007). “Rule generation for hierarchical collaborative fuzzy system”, *Applied Mathematical Modelling*, (In Press).

[7] Salgado, P. (2007), “Hierarchical decomposition of the fuzzy systems by clustering process”, *accepted to publish*.

[8] Höppner, F., Klawonn, F., Kruse, R. & Runkler, T. (1999). Fuzzy Cluster Analysis—Methods for Image Recognition, Classification, and Data Analysis. New York: Wiley.

[9] Takagi, T. & Sugeno, M. (1985). "Fuzzy Identification of Systems and Its Applications to Modelling and Control." IEEE Transactions on Systems, Man, and Cybernetics 5(3): 116-132.

[10] Rovatti, R., “Takagi-sugeno models as approximators in Sobolev norms: the SISO case”, *Proc. 5th IEEE Conf. on Fuzzy Systems*, New Orleans, 1996, pp. 1060–1066.

[11] Wang, L.X., Adaptive fuzzy systems and control: design and stability analysis’ (Prentice-Hall, NJ, 1994).