Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

ELSEVIER International Journal of Approximate Reasoning 35 (2004) 275–289
www.elsevier.com/locate/ijar

An approach for fuzzy rule-base adaptation using on-line clustering

Plamen Angelov *

Department of Communication Systems, Faculty of Applied Sciences, Lancaster University, Bailrigg, Lancaster, LA1 4YR, UK

Received 1 February 2003; accepted 1 August 2003

Abstract

A recursive approach for adaptation of fuzzy rule-based model structure has been developed and tested. It uses *on-line* clustering of the input–output data with a recursively calculated spatial proximity measure. Centres of these clusters are then used as prototypes of the centres of the fuzzy rules (as their focal points). The recursive nature of the algorithm makes possible to design an evolving fuzzy rule-base in *on-line* mode, which adapts to the variations of the data pattern. The proposed algorithm is instrumental for *on-line* identification of Takagi–Sugeno models, exploiting their dual nature and combined with the recursive modified weighted least squares estimation of the parameters of the consequent part of the model. The resulting evolving fuzzy rule-based models have high degree of transparency, compact form, and computational efficiency. This makes them strongly competitive candidates for *on-line* modelling, estimation and control in comparison with the neural networks, polynomial and regression models. The approach has been tested with data from a fermentation process of lactose oxidation.

Keywords: On-line clustering; Fuzzy rule-based models identification; Parameter estimation; Takagi–Sugeno fuzzy models

* Tel.: +44-1524-594-188; fax: +44-1524-592713.
E-mail address: p.angelov@lancaster.ac.uk (P. Angelov).

0888-613X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2003.08.006

1. Introduction

Recently, fuzzy rule-based models of the so-called Takagi–Sugeno (TSK) type [10] have been widely used in modelling and control applications of complex systems, because of their computational efficiency, transparency and flexibility [1,2]. The main reason for this is their dual, quasi-linear nature: they are non-linear in general and thus suitable for complex systems, but they can also be treated as partially linear in respect to sub-models of the consequent. In fact, TSK models are multi-model construct [14] with fuzzily defined regions of validity of the local models. They have close nature to other types of models, which use superposition of local models, such as Gabor–Kolmogorov's Gaussian polynomials [17], group method of data handling [18], radial-basis function neural networks [19], generalised regression neural networks [3], quantum states and von Neumann lattice [4] etc. The main advantage of the TSK models is that since the local regions are *fuzzily* defined, the resulting global model can be non-linear (of high order) while the local models can be as simple as necessary. While, in general, the local sub-models can be non-linear, usually linear (first-order) or singleton (zero-order) sub-models are considered [10,12]. The original formulation of TSK models [10] assumes linear sub-models and in this paper we will make the same assumption, without loss of generality.

The design problem for such models can, in general, be solved by two different approaches:

1. Directly solving the non-linear optimisation problem, which constitutes the problem of identification of parameters of the fuzzy membership functions and the structure of the fuzzy rule-based models (unless it is assumed that it is known or it is supplied by experts) [15,16].
2. Decomposing the identification problem into two separate tasks:
   (a) Identification of the regions of interest (the regions of validity of the local sub-models), which in some control applications could be viewed as 'regimes' or 'operating modes' [14];
   (b) Identification of the parameters of the output (consequent part) of the model.

Since the problem of the design of fuzzy rule-based models, which is a structure (rule-base) and parameter (of the membership functions) identification problem, is a non-linear one in its nature, the first approach can lead to a near-optimal solution (subject to the Kuhn–Taker conditions of optimality [15]). The second approach ensures a sub-optimal solution only, because of the decomposition. The first approach is computationally more involved and not readily suitable for on-line applications [1,5–8,12]. The second approach, in the opposite, have large potential for real-time and on-line applications, because of

the effective de-coupling of the identification of the antecedent and consequent parts of the fuzzy rule-based models. This approach will be considered in the present paper.

The regions of interest (task (a)) can be defined by input–output data space clustering [5–8]. The second task (b) can be solved by using recursive least squares estimation technique [5–8,11].

The problem of model identification has, normally, several stages, the last being model validation [11]. In reality, however, there is no guarantee that a model or controller based on such model will perform satisfactory even if the validation has been successful [5]. There could be different reasons, the main being possible changes in the object of modelling/control or the environment in which this object exists. Internal changes could be due to development of a fault, ageing or simply due to dynamic processes not represented in the training and validation data. External changes could have different nature, very often of seasonal character. Adaptation of parameters only has limited effect. More interesting is the case when the model structure itself can be adapted, is evolving [5–8]. If consider on-line mode and suppose an evolving model structure, then the task (a) is modified/complicated to on-line clustering with an option for evolution of the cluster centres (and parameters). The task (b) will not be considered in this paper. More details about it can be found in [6,8].

Existing clusters (used as prototypes of the fuzzy rules) can be replaced with the new ones based on ranking using the information measure based on the distance between *all* of the data samples. In this way, the rule-base structure can evolve by inheriting the bulk of the previous structure (all but one rules) and updating, when new informative data become available, rather then being retrained from the scratch [5,6].

The approach has been tested with data from a fermentation process of lactose oxidation, which is characterised as highly non-linear and non-stationary [5].

2. Off-line identification of the fuzzy rule-base

Considerations in the rest of the paper are exemplified by TSK fuzzy rule-based model without loss of generality. While the off-line clustering approach can be applied to identification of other fuzzy rule-based models, the full effect of the application of the on-line clustering approach, presented in this paper is evident in the on-line identification of TSK models [8]. TSK model is a rule-based model with fuzzy antecedents and functional (usually linear) consequence [10]:

$$\begin{aligned} \Re_i: \quad & \boldsymbol{IF}(x_1 \text{ is } \aleph_{i1})\boldsymbol{AND}\ldots\boldsymbol{AND}(x_n \text{ is } \aleph_{in}) \\ & \boldsymbol{THEN}(y_i = p_0 + p_{i1}x_1 + p_{i2}x_2 + \cdots + p_{in}x_n); \quad i = 1, 2, \ldots, R \end{aligned} \tag{1}$$

where $\Re_i$ denotes the ith fuzzy rule; R is the number of fuzzy rules; $\boldsymbol{x}$ is the input vector; $\boldsymbol{x} = [1, x_1, x_2, \ldots, x_n]^{\mathrm{T}}$; $\aleph_{ij}$ denotes the antecedent fuzzy sets, $j = 1, 2, \ldots, n$; y_i is the output of the ith linear subsystem; p_{il} are its parameters, $l = 0, 1, \ldots, n$.

Each fuzzy rule has a degree of firing, which is proportional to the level of contribution of the corresponding sub-model to the overall output of the TSK model:

$$\begin{aligned} &\tau_i = \mu_{i1}(x_1) \times \mu_{i2}(x_2) \times \cdots \times \mu_{in}(x_n) = \bigcap_{j=1}^{n} \mu_{ij}(x_j) \\ &\mu_{ij} = \mathrm{e}^{-\alpha \|x_j - x_i^*\|}; \quad i = 1, 2, \ldots, R; \; j = 1, 2, \ldots, n \end{aligned} \tag{2}$$

where $\alpha = 4/r^2$ and r is a positive constant, which defines the spread of the antecedent and the zone of influence of the ith model; x_i^* is the focal point (centre) of the ith fuzzy rule antecedent.

The TS model output is calculated by weighted averaging of individual rules' contributions [10,13]:

$$y = \sum_{i=1}^{R} \lambda_i y_i = \sum_{i=1}^{R} \lambda_i x^{\mathrm{T}} \pi_i \tag{3}$$

where $\lambda_i = \tau_i / \sum_{j=1}^{R} \tau_j$ is the normalized firing level of the ith rule; y_i represents the output of the ith linear sub-model; $\pi_i = [p_{i1} p_{i2} \ldots p_{in}]^{\mathrm{T}}$, $i = 1, 2, \ldots, R$, is the vector of parameters of the ith linear sub-model.

Generally, the problem of identification of TSK model can be divided into two sub-tasks [5–10,13]:

(a) Identification of the antecedent part of the model (1); this consists of determination of the centres of the clusters, which are prototypes of the fuzzy rules ($x_i^*; i = 1, 2, \ldots, R$) and spreads ($r$) of the respective membership functions.
(b) Estimation of the parameters of the sub-models ($p_{ij}; i = 1, 2, \ldots, R; j = 1, 2, \ldots, n$) of the consequent part.

First sub-task, the antecedent parameters *off-line* learning, when all the input–output data are available, can be solved by clustering the input–output data space ($z = [x^{\mathrm{T}}; y]^{\mathrm{T}}$), for example by *subtractive clustering* [9]. It uses the data points as candidate prototype cluster centres and ranks them by their potential to be a cluster centre. It is a measure of the spatial proximity between a particular point z_i and *all other* data points (the value of the potential is higher for a data point that is surrounded by a large number of close data points).

One can use different functions to express the potential, but they should be monotonic and inversely proportional to the distance between *all* of the data

points. The distance can be Euclidean (then the clusters are circular [5–9], Mahalonobis (then the clusters are ellipsoids) [4] etc. In this paper so-called Cauchy type function is used:

$$P_i = \frac{1}{1 + \frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{n+1}(d_{ij}^l)^2}; \quad i = 1, 2, \ldots, N \tag{4}$$

where $d_{ij}^l = z_i^l - z_j^l$, denotes projection of the distance between two data points (z_i^l and z_j^l) on the axis z^l; $l = 1, 2, \ldots n+1$.

This form of the potential is also monotonic and inversely proportional to the distance as the exponential Gaussian one, but, as it will be demonstrated in the next section, it makes possible *recursive* calculation of the potential.

The procedure of the *subtractive clustering* includes the following steps [9]:

(i) The first cluster centre is established at a point with the highest potential:

$$P_1^* := \max_{i=1}^{N} P_i \tag{5}$$

where P_1^* denotes the potential of the first centre.

(ii) The potential of *all* other data points is reduced by an amount proportional to the potential of the chosen point and inversely proportional to the distance to this centre:

$$P_i := P_i - P_k^* \mathrm{e}^{-\beta\|z_i - z_k^*\|^2}; \quad i = 1, 2, \ldots, N \tag{6}$$

where P_k^* denotes the potential of the kth centre; $k = 1, \ldots, N$.
$\beta = \frac{4}{r_b^2}$; where r_b is a positive constant, determining the radius of the neighborhood that will have measurable reductions in the potential because of the closeness to an existing centre; recommended value of r_b is $r_b = 1.5r$ [9].
The next centre is found also as the data point with the highest (after this subtraction) potential.

(iii) The procedure is repeated until the potential of all data points is reduced below certain threshold, determined as a function of the maximal potential called the 'reference' potential (P^{ref}): $\overline{\varepsilon}^* P^{\mathrm{ref}}$.

(iv) Similarly, a lower threshold is defined ($\underline{\varepsilon}^* P^{\mathrm{ref}}$). If the potential of a point lies between the two thresholds ($\overline{\varepsilon}^* P^{\mathrm{ref}}$ and $\underline{\varepsilon}^* P^{\mathrm{ref}}$), the shortest of the distances ($\delta_{\min}$) between the new candidate to be a cluster centre (z_k^*) and all previously found cluster centres is decisive [9]. The following inequality, express the trade-off between the potential value and the closeness to the previous centres:

$$\frac{\delta_{\min}}{r} + \frac{P_k^*}{P_1^*} \geqslant 1 \tag{7}$$

This approach has been used for initial estimation of the antecedent parameters in identification of the fuzzy rule-based TSK model and the resulting cluster centres define the centres of the rules of the model [9].

The second sub-task, estimation of the parameters of the consequent linear sub-models can be solved as least squared problem for fixed antecedent parameters [6,8].

3. On-line adaptation of the evolving fuzzy rule-base

The previous approach is suitable for a fixed set of data supplied in batch mode and under the assumption that the model structure (rule-base) remains unchanged. When the training data are collected continuously, some of them will reinforce and confirm the information contained in the previous data, but the other could bring new information [5]. This new information could concern a change in operating conditions, development of a fault or simply more significant change in the dynamic of the process. They may content enough new information *to form a new rule or to modify* an existing one. The value of the information they bring is closely related to the information, which is possessed in the data collected so far. In cases when the new data are more informative, than the data used as prototypes for forming the antecedent part of the rules an adaptation or *evolution* of the model structure (rule-base) is necessary. It is based on the *on-line* clustering of the input–output data space with gradually evolving (shifting) regions of interest. For development of *on-line* clustering procedure (applicable in *real-time*) the two main potentials have to be calculated *recursively*:

(a) Potential of the new data sample.
(b) Potential of the existing centres have to be updated recursively.

The potential of the new data sample can be calculated by (4). This formula requires, however, the distances, and hence the co-ordinates of *all* of the previous data points. It can be re-arranged expressing the distances as follows:

$$P_i = \frac{1}{1+\frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{n+1}(d_{ij}^l)^2} = \frac{1}{1+\frac{1}{N}\sum_{j=1}^{N}\sum_{l=1}^{n+1}\left((z_i^l)^2 - 2z_i^l z_j^l + (z_j^l)^2\right)}; \quad i = 1,2,\ldots,N \tag{8}$$

This expression for the potential of a data sample at the moment k using *all* of the previous data samples available at this moment can be represented as:

$$P_k(z_k) = \frac{1}{1+\frac{(k-1)b_k - 2h_k + g_k}{k-1}} = \frac{k-1}{(k-1)(1+b_k) - 2h_k + g_k}; \quad k = 2,3,\ldots \tag{9}$$

where $P_k(z_k)$ denotes the potential of the data point (z_k) calculated at the moment k;

$$b_k = \sum_{j=1}^{n+1}(z_k^j)^2, \quad g_k = \sum_{i=1}^{k-1}\sum_{j=1}^{n+1}(z_i^j)^2, \quad h_k = \sum_{j=1}^{n+1} z_k^j p_k^j, \quad p_k^j = \sum_{i=1}^{k-1} z_i^j$$

Parameters b_k and h_k are calculated from the current data point z_k, while p_k^j and g_k are recursively updated by $g_k = g_{k-1} + \sum_{j=1}^{n+1}(z_{k-1}^j)^2$; $p_k^j = p_{k-1}^j + z_{k-1}^j$.

The newly collected data influence the potentials of the established centres of the clusters (z_i^*, $i = 1, 2, \ldots, R$), because by definition the potential depends on the distance to *all* of the data points, including the new ones. The potential of a centre of a cluster (z_i^*) at the moment k can be calculated from (8) or (9) as:

$$P_k(z_i^*) = \frac{1}{1 + \frac{1}{k-1}\sum_{j=1}^{k-1}\sum_{l=1}^{n+1}(d_{ij}^l)^2}; \quad k = 2, 3, \ldots \tag{10}$$

where $P_k(z_i^*)$ is the potential of the at the moment k of the cluster, which is a prototype of the ith rule.

Similarly, for $k - 1$ we have:

$$P_{k-1}(z_i^*) = \frac{1}{1 + \frac{1}{k-2}\sum_{j=1}^{k-2}\sum_{l=1}^{n+1}(d_{ij}^l)^2}; \quad k = 3, 4, \ldots \tag{10a}$$

The potential of the cluster/rule centre can be expressed *recursively* from the value at the previous moment (before the new data sample is available) as:

$$P_k(z_i^*) = \frac{(k-1)P_{k-1}(z_i^*)}{k - 2 + P_{k-1}(z_i^*) + P_{k-1}(z_i^*)\sum_{j=1}^{n+1}(d_{i(k-1)}^j)^2} \tag{11}$$

The following procedure for *on-line* clustering is considered [6,8]:

1. We establish the first data point as the centre of the first cluster. Its co-ordinates are used to form the antecedent part of the fuzzy rule and its potential is equal to 1.
2. Starting from the next data point onwards the potential of each new data point is calculated *recursively* by (9).
3. The potentials of the existing centres are *recursively* updated by (11).
4. Potentials of the *new* data points are compared to the updated potential of the centres of the existing clusters.
   (a) If the potential of the *new* data point is *higher* than the potential of the *existing* centres then the *new* data point is added as a *new centre* and a *new rule* is formed ($R := R + 1$; $x_R^* = x_k$). The condition to have higher potential is a very strong one, which restricts form an excessively large rule-base being formed.
   (b) If in addition to the previous condition the new data point is close to an old centre then the new data point (z_k) replaces this centre ($z_j^* := z_k$).

This *on-line* clustering approach ensures an *evolving* rule-base by dynamically upgrading and modifying it while inheriting the bulk of the rules ($R - 1$ of the rules are preserved even when a modification or an upgrade take place).

Assumed that the rule-base is *gradually evolving* the number of rules as well as the parameters of the antecedent part will vary. Then weighted recursive least squares estimation [8] could be used to solve the second sub-task: to identify parameters of these loosely coupled linear sub-models [8].

The procedure for the *on-line* rule-base adaptation (Fig. 1) can be summarised as follows:

1. Initialisation of the rule-base structure (antecedent part of the rules).
2. Reading the *next* data sample at the next time step.
3. *Recursive* calculation of the potential of each new data sample by (9).
4. *Recursive* update of the potentials of the established centres by (11) because of the influence of the *new* data.
5. Possible *modification* or *innovation of* the rule-base *structure* based on the comparison between the potential of the *new* datasample and the potential of the *existing* rules' centres.

   The algorithm continues for the next time step from step 2.

4. Experimental results

On-line modelling the fermentation of *Kluyveromyces marxianus var. lactis MC 5* in lactose oxidation [5] is considered as an example of the application of the adaptive rule-base approach. Biotechnological processes (an important part of which is fermentation processes) are characterised by uncertainties, highly non-stationary dynamics, non-reproducibility (uniqueness), existing of non-quantified factors, such as smell, taste, colour, morpho-physiological specifics etc. The basic source of difficulties in modelling biotechnological processes is the lack of adequate knowledge and information about the complex intimate nature of these processes, involving complex metabolite pathways. Practically, each cell is a living organism, and it should be considered as a separate multi-component, self-organising system, being in a constant energetic and metabolite exchange with the environment. The relation between the product synthesis and the metabolism is not always clear and obvious. The level and precision of sensors, especially of cell mass, substrates and product concentrations is far from the level of 'classical' temperature or pH sensors. Each run of fermentation differs by the type of used species, tanks etc. Strongly speaking, if we change the strain, tank or inoculate, the process is yet a new one.

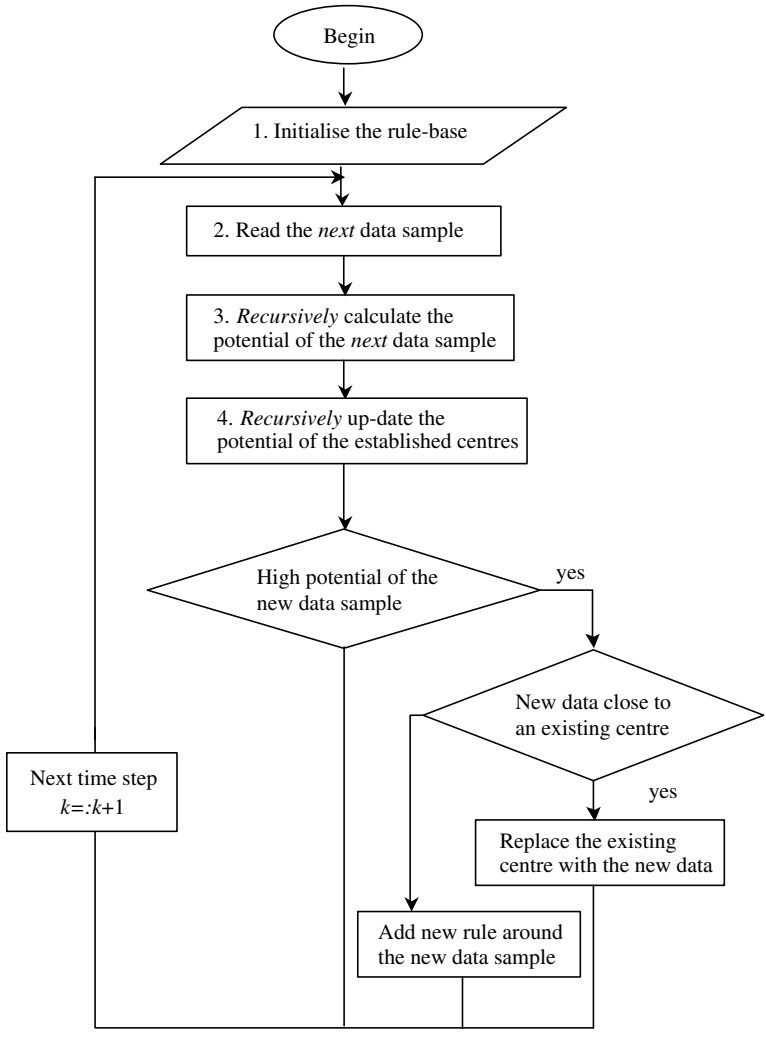

Fig. 1. Flow chart of the on-line clustering algorithm.

The model of this particular process includes the dependence between concentrations of the cell mass (X) and the basic energetic substrates: lactose (S) and dissolved oxygen concentration (DO). In very general terms, the following first principles model based on mass- and energy-balance can be formulated, but the complexity is hidden in this form in the specific rates, which are highly non-linear, non-stationary functions of many factors [5]:

$$\begin{aligned}\frac{\mathrm{d}X}{\mathrm{d}t} &= \mu_X(S, \mathrm{DO}, T, pH, n, \ldots)X \\ \frac{\mathrm{d}S}{\mathrm{d}t} &= -q_{\mathrm{S}}(S, \mathrm{DO}, T, pH, n, \ldots)X \\ \frac{\mathrm{d}\,\mathrm{DO}}{\mathrm{d}t} &= q_{\mathrm{CO}}(S, \mathrm{DO}, T, pH, n, \ldots)X\end{aligned} \tag{12}$$

where X denotes cell mass concentration, μ_X denotes specific growth rate, q_S denotes specific lactose consumption rate, q_{DO} denotes specific DO consumption rate, n denotes the rotation speed, T denotes the temperature.

There is no a general mathematical model of the microbial synthesis because of the extreme complexity and great variety of living activity of the microorganisms, although various models of different parts of the whey fermentation exist.

In our experiment (Fig. 2), six independent batch runs of fermentation where carried out in aerobic cultivation of *Kluyveromyces lactis* using a laboratory bioreactor ABR 02M with capacity 2 l [5].

The lactose concentration (S) is modelled *on-line* using data for the cell mass concentration (X) and the dissolved oxygen concentration (DO):

$$\begin{aligned}y_{k+1} &= S_{k+1}; \\ x_{k+1} &= \{X_{k+1}, \mathrm{DO}_{k+1}\}\end{aligned} \tag{13}$$

where y denotes the output vector; x denotes the inputs vector.

The design of the TSK model started with the first data point established as a centre for the first rule. Then, in on-line mode reading each new data sample the decision to upgrade (or modify) the existing rule-base is taken as described in the previous section. In this way the model structure (rule-base) has evolved, and finally it contains eight rules. The parameters of the consequents (local linear sub-models) has been updated recursively by Kalman filtration. Their evolution is illustrated in Figs. 5 and 6. The values of the parameters of the

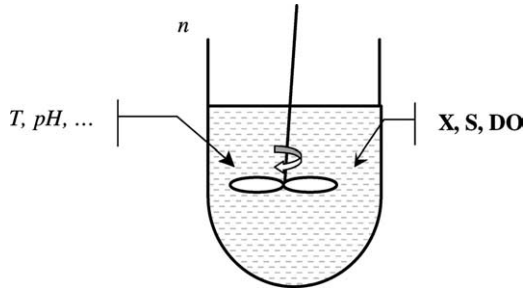

Fig. 2. Schematic representation of a fermentation process in a stirred tank reactor.

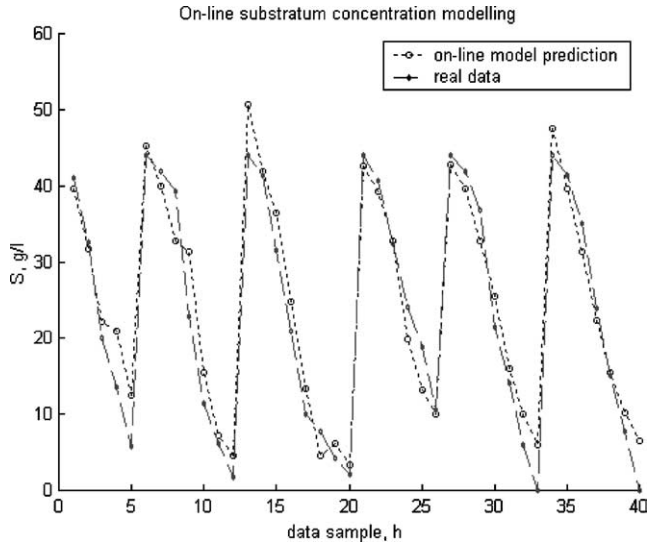

Fig. 3. On-line modelling lactose concentration (six fermentations).

antecedent part (rule centres) are tabulated in Table 1. The model finally looks as follows:

R_1: ***IF*** (X *is Extremely Low*) ***AND*** (DO *is Extremely High*)
THEN $S = a_0^1 + a_1^1 X + a_2^1 \mathrm{DO}$

R_2: ***IF*** (X *is Low*) ***AND*** (DO *is Rather High*)
THEN $S = a_0^2 + a_1^2 X + a_2^2 \mathrm{DO}$

R_3: ***IF*** (X *is Very Low*) ***AND*** (DO *is Very High*)
THEN $S = a_0^3 + a_1^3 X + a_2^3 \mathrm{DO}$

R_4: ***IF*** (X *is Medium*) ***AND*** (DO *is High*)
THEN $S = a_0^4 + a_1^4 X + a_2^4 \mathrm{DO}$

R_5: ***IF*** (X *is Medium*) ***AND*** (DO *is Very Low*)
THEN $S = a_0^5 + a_1^5 X + a_2^5 \mathrm{DO}$

R_6: ***IF*** (X *is Very High*) ***AND*** (DO *is Very Low*)
THEN $S = a_0^6 + a_1^6 X + a_2^6 \mathrm{DO}$

R_7: ***IF*** (X *is Extremely High*) ***AND*** (DO *is Low*)
THEN $S = a_0^7 + a_1^7 X + a_2^7 \mathrm{DO}$

R_8: ***IF*** (X *is High*) ***AND*** (DO *is Extremely Low*)
THEN $S = a_0^8 + a_1^8 X + a_2^8 \mathrm{DO}$

The RMS error in prediction of the lactose concentration is 3.85 g/l (Fig. 4). Having in mind that the range of values for this concentration is 0–50 g/l and

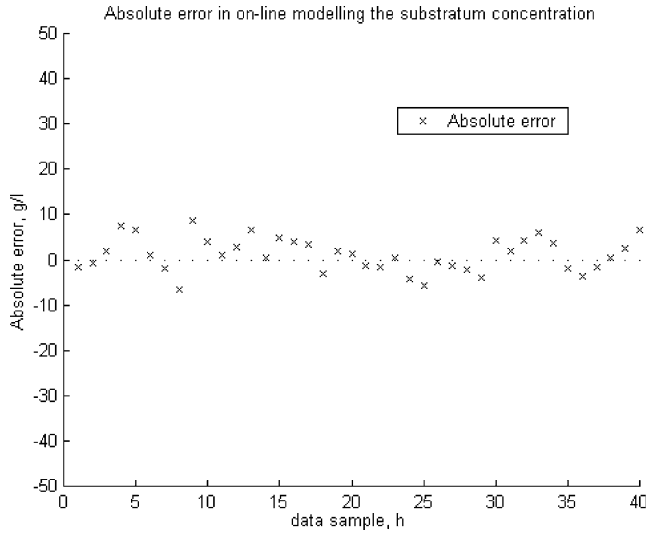

Fig. 4. Absolute error in prediction the cell mass concentration.

the practical difficulties with measuring cell mass concentrations, this level is satisfactory. Moreover, the results from all of the six different fermentations indicate that the model fits well the real experimental data (Fig. 3), which compares favourably with the off-line modelling [5], where for at least one or two fermentation runs the results are significantly poorer.

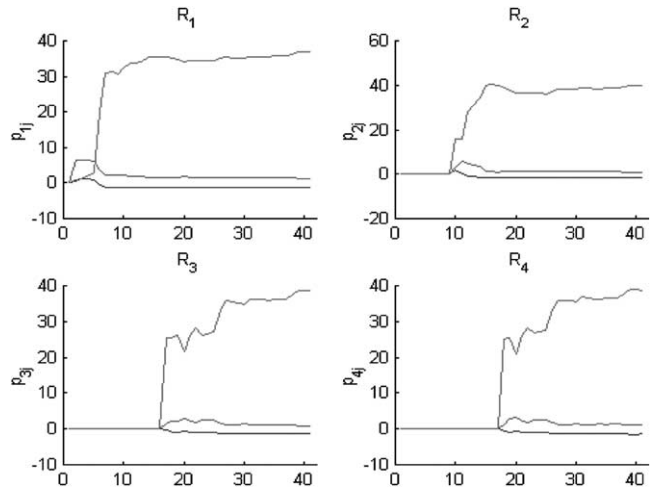

Fig. 5. Evolution of the parameters of the sub-models 1–4.

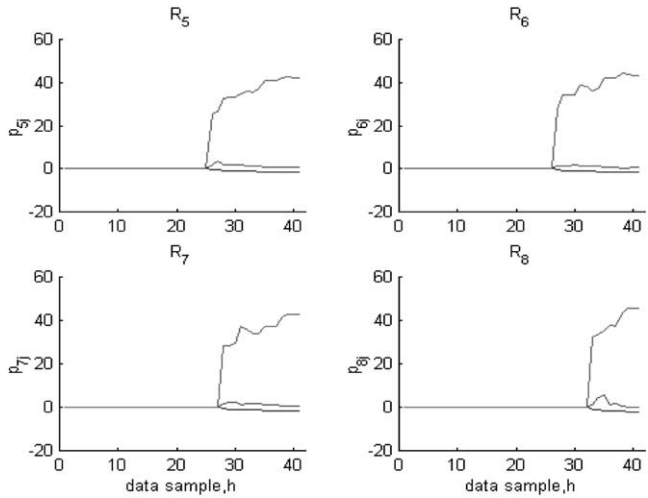

Fig. 6. Evolution of the parameters of the sub-models 5–8.

Table 1
Centres of the antecedent part of the rules

Centre	X_k	g/l	DO_k	–
	Linguistic label	Value	Linguistic label	Value
x_1^*	Extremely Low	3.2	Extremely High	5.59
x_2^*	Low	6.5	Rather High	2.54
x_3^*	Very Low	5.8	Very High	4.04
x_4^*	Medium	11.6	High	2.80
x_5^*	Medium	11.4	Very Low	0.13
x_6^*	Very High	18	Very Low	0.14
x_7^*	Extremely High	21.9	Low	0.18
x_8^*	High	15.8	Extremely Low	0.07

5. Conclusion

A recursive approach for adaptation of fuzzy rule-based model structure has been developed and tested. It uses *on-line* clustering of the input–output data with a recursively calculated spatial proximity measure. Centres of these clusters are then used as prototypes of the centres of the fuzzy rules. The recursive nature of the algorithm makes possible to design an evolving fuzzy rule-base in *on-line* mode, which adapts to the variations of the data pattern.

The proposed algorithm is instrumental for *on-line* identification of TSK models, exploiting their dual nature and combined with the recursive modified weighted least squares estimation of the parameters of the consequent part of the model. The resulting evolving fuzzy rule-based models have high degree of

transparency, compact form, and computational efficiency. This makes them strongly competitive candidates for *on-line* modelling, estimation and control in comparison with the neural networks, polynomial and regression models. The approach has been tested with data from a fermentation process of lactose oxidation.

Acknowledgement

The author acknowledges the helpful comments of Dr. D.P. Filev.

References

[1] W.S. Lin, C.-H. Tsai, Self-organizing fuzzy control of multi-variable systems using learning vector quantization network, Fuzzy Sets and Systems 124 (2001) 197–212.
[2] F.-J. Lin, C.-H. Lin, P.-H. Shen, Self-constructing fuzzy neural network speed controller for permanent-magnet synchronous motor drive, IEEE Transactions on Fuzzy Systems 9 (5) (2001) 751–759.
[3] D.F. Specht, A general regression neural network, IEEE Transactions on NN 2 (6) (1991) 568–576.
[4] L.K. Stergioulas, V.S. Vassiliadis, A. Vourdas, Construction of quantum states from an optimally truncated von Neumann lattice of coherent states, Journal of Physics A: Mathematical and General 32 (1999) 3169–3178.
[5] P.P. Angelov, Evolving rule-based models: a tool for design of flexible adaptive systems, Springer, Physica-Verlag, Heidelberg, 2002.
[6] P.P. Angelov, R.A. Buswell, Identification of evolving fuzzy rule-based models, IEEE Transactions on Fuzzy Systems 5 (10) (2002) 667–677.
[7] P.P. Angelov et al., Evolving rule-based control, in: EUNITE Symposium, Tenerife, Spain, 2001, pp. 36–41.
[8] P.P. Angelov, D.P. Filev, An approach to on-line identification of Takagi–Sugeno fuzzy models, IEEE Transactions on Systems Man and Cybernetics Part B 33 (13) (2003).
[9] S.L. Chiu, Fuzzy model identification based on cluster estimation, Journal of Intelligent and Fuzzy Systems 2 (1994) 267–278.
[10] T. Takagi, M. Sugeno, Fuzzy identification of systems and its application to modeling and control, IEEE Transactions on SMC 15 (1985) 116–132.
[11] K.J. Astroem, B. Wittenmark, Adaptive Control, Addison Wesley, MA, USA, 1989.
[12] N.K. Kasabov, Q. Song, DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction, IEEE Transactions on Fuzzy Systems 10 (2) (2002) 144–154.
[13] M. Sugeno, T. Yasukawa, A fuzzy logic based approach to qualitative modeling, IEEE Transactions on Fuzzy Systems 1 (1) (1993) 7–31.
[14] T.A. Johanson, R. Murray-Smith, Operating regime approach to non-linear modeling and control, in: R. Murray-Smith, T.A. Johanson (Eds.), Multiple Model Approaches to Modeling and Control, Taylor Francis, Hants, UK, 1997, pp. 3–72.
[15] P. Angelov, R. Buswell, Automatic generation of fuzzy rule-based models from data by genetic algorithms, Information Sciences 150 (1/2) (2003) 17–31.
[16] M.H. Lim, S. Rahardja, B.H. Gwee, A GA paradigm for learning fuzzy rules, Fuzzy Sets and Systems 82 (1996) 177–186.

[17] D. Gabor, W. Wildes, R. Woodcock, A universal nonlinear filter, predictor and simulator which optimizes itself by a learning process, Proceedings of the IEE 108B (1961) 422–438.
[18] A.G. Ivakhnenko, Polynomial theory of complex systems, IEEE Transactions on Systems Man and Cybernetics 1 (4) (1971) 364–378.
[19] J.S.R. Jang, ANFIS: adaptive network-based fuzzy inference systems, IEEE Transactions on Systems Man and Cybernetics 23 (3) (1993) 665–685.