On Constructing Parsimonious Type-2 Fuzzy Logic Systems via Influential Rule Selection

Shang-Ming Zhou, *Member, IEEE,* Jonathan M. Garibaldi, Robert I. John, *Member, IEEE,* and Francisco Chiclana

***Abstract*—Type-2 fuzzy systems are increasing in popularity and there are many examples of successful applications. While many techniques have been proposed for creating parsimonious type-1 fuzzy systems, there is a lack of such techniques for type-2 systems. The essential problem is to reduce the number of rules, whilst maintaining the system's approximation performance. In this paper, four novel indices for ranking the relative contribution of type-2 fuzzy rules are proposed, termed R-values, c-values, ω_1-values and ω_2-values. The R-values of type-2 fuzzy rules are obtained by applying a QR decomposition pivoting algorithm to the firing strength matrices of the trained fuzzy model. The c-values rank rules based on the effects of rule consequents, whilst the ω_1-values and ω_2-values consider both the rule base structure (via firing strength matrices) and the output contribution of fuzzy rule consequents. Two procedures for utilising these indices in fuzzy rule selection (termed 'forward selection' and 'backward elimination') are described. Experiments are presented which demonstrate that, by using the proposed methodology, the most influential type-2 fuzzy rules can be effectively retained in order to construct parsimonious type-2 fuzzy models.**

***Index Terms*—Type-2 fuzzy sets, parsimony, rule ranking, rule selection, index, QR, SVD-QR**

I. Introduction

TYPE-2 fuzzy sets were initially proposed by Zadeh in 1975 [1]. Unlike type-1 fuzzy sets whose membership values are precise numbers in [0, 1], membership grades of a type-2 fuzzy set are themselves type-1 fuzzy sets, so type-2 fuzzy sets offer an opportunity to model higher level uncertainty in the human decision making process than type-1 fuzzy sets [2]–[5]. In a type-2 fuzzy inference system (T2FIS), some fuzzy sets used in the antecedent and/or consequent parts and each rule inference output are type-2 fuzzy sets. T2FISs have been used in many successful applications in various areas where uncertainties occur, such as in decision making [6]–[8], diagnostic medicine [9], [10], signal processing [11], [12], traffic forecasting [13], mobile robot control [14], pattern recognition [15]–[17], intelligent control [18], [19], and ambient intelligent environments [20].

However, one challenge in type-1 fuzzy systems also remains in T2FISs, that is, the *curse of dimensionality*: the number of fuzzy rules required increases exponentially with the dimensionality of the input space. An additional challenge in T2FIS modelling is that it involves higher computational overhead than type-1 fuzzy inference system modelling [2]. Importantly, in data-driven type-2 fuzzy modelling, the type-2 fuzzy rule-base generated by the back-propagation method [2], [21] will almost certainly suffer from rule redundancy [26]. This is because, as an accuracy-oriented method, the back-propagation training process is concerned only with producing an accurate system model, without any consideration of parsimony, consistency, and transparency of generated rule-base. Hence, accurate and parsimonious modelling techniques are urgently needed for T2FIS modelling. As a matter of fact, even in type-1 FIS modelling, the problem of developing parsimonious fuzzy modelling technique with as few fuzzy rules as possible is a very important research topic [22]–[25]. Interestingly, Liang and Mendel have suggested using the *SVD-QR* method [27], [28] to perform rule reduction for the sake of designing parsimonious interval T2FIS (IT2FIS) [26], in which *QR decomposition with column pivoting* is further applied to the unitary matrix V after the *singular-value decomposition* (SVD) of the firing strength matrix. However, one issue arising in applying their method is the estimation of an effective rank. Some research on type-1 fuzzy models indicates that rule reduction using the *SVD-QR with column pivoting* algorithm is very sensitive to the chosen rank, in that different estimates of the rank often produce dramatically different rule reduction results [29], [30].

In order to avoid the estimation of the rank for the SVD-QR with column pivoting algorithm, this paper applies the pivoted QR decomposition algorithm to type-2 fuzzy rule reduction. The absolute values of diagonal elements of matrix R in QR decomposition (termed the R-values) of fuzzy rules tend to track the singular values and can be used for rule ranking. However, both the pivoted QR decomposition algorithm and the SVD-QR with column pivoting algorithm only consider the rule-base structure, without paying attention to the effect of rule consequents during rule selection. In other words, both algorithms only employ the information from the premises of the fuzzy rules when carrying out rule reduction, but ignore the information from the rule consequents. In type-1 fuzzy system modelling, one way of considering the effects of fuzzy rule consequents on rule selection is via the orthogonal-least squares (OLS) technique [31]. Wang and Mendel first applied the OLS method to fuzzy rule selection [32], in which it was used to select the most important fuzzy basis functions, each of which was associated with a fuzzy rule. Later, Wang and Langari employed the OLS method to remove less important consequent terms for a first-order Tagaki-Sugeno fuzzy model [33]. Mastorocostas *et al.* proposed a constrained OLS ap-

Manuscript received October 10, 2007; revised February 18, 2008.

S.-M. Zhou, R. I. John, and F. Chiclana are with the Centre for Computational Intelligence, School of Computing, De Montfort University, Leicester, LE1 9BH, UK (email: smzhou@ieee.org, rij@dmu.ac.uk, chiclana@dmu.ac.uk).

J. M. Garibaldi is with ASAP, School of Computer Science and Information Technology, University of Nottingham, Nottingham, NG8 1BB, UK (email:jmg@cs.nott.ac.uk).

proach as an improvement to the basic OLS approach for producing compact fuzzy rule-bases [34]. The OLS based methods select the most important type-1 fuzzy rules based on their contributions of variance to the variance of the output. Another way of considering the output contribution of the fuzzy rules has been proposed by Zhou and Gan in fuzzy modelling based on support vector machines (SVMs) [25]. The Lagrangian multipliers of an SVM (termed the α-values of fuzzy rules) are used to rank the contributions of associated rule consequents. In this paper, we propose a novel method to consider the effects of type-2 rule consequents in order to select the most influential rules, which we term c-values. It can be seen that the existing methods of selecting influential fuzzy rules by considering the contributions of rule consequents [25], [32]–[34] are type-1 fuzzy system-oriented, which can not be applied in type-2 fuzzy models. Another advantage of the proposed c-values of fuzzy rules is that these values can be derived directly from given type-2 fuzzy rules: no additional computational techniques like OLS or SVM are needed. In type-1 fuzzy system modeling, other techniques of constructing parsimonious rules such as the one in [35] focus on merging similar fuzzy sets, do not take into account the scheme of selecting the influential rules in terms of rule ranking results.

Ideally, a rule ranking method should take into account both the structure of the rule-base and the effects of rule consequents in order to generate a parsimonious fuzzy model with good generalisation performance. However, currently this kind of scheme is very rare, even in type-1 fuzzy system modelling. In this paper, we propose ω_1-values and ω_2-values in order to take into account both the type-2 rule base structure and the contributions of type-2 rule consequents.

The organization of this paper is as follows. Section 2 reviews T2FIS with emphasis on IT2FIS whose secondary membership grades are unity. In Section 3, we propose some new rule ranking indices forT2FISS. Two procedures for utilising these indices in type-2 fuzzy rule selection are described in Section 4. Section 5 includes the experimental results of evaluating the proposed methods in three examples. Finally, discussion of the various indices is presented in Section 6, together with conclusions.

II. Overview of Type-2 Fuzzy Logic Systems

A T2FIS is a fuzzy logic system in which at least one of the fuzzy sets used in the antecedent and/or consequent parts and each rule inference output is a type-2 fuzzy set. Consider a type-2 Mamdani FIS having n inputs $x_1 \in X_1, \cdots, x_n \in X_n$ and one output $y \in Y$. The rule base contains K type-2 fuzzy rules expressed in the following form:

$$R^k: \ if\ x_1\ is\ \tilde{F}_1^k\ and\ \cdots and\ x_n\ is\ \tilde{F}_n^k,\ then\ y\ is\ \tilde{G}^k \quad (1)$$

where $k = 1, \cdots, K$, $\tilde{F}_i^k$ and $\tilde{G}^k$ are type-2 fuzzy sets. These rules represents fuzzy relations between the multiple dimensional input space $X \triangleq X_1 \times \cdots \times X_n$ and the output space Y. However, the computing load involved in deriving the system output from a general T2FIS model is high in practice, and this has become a major factor curtailing applications of general T2FIS. Advances have recently been made in computing general T2FIS inference by using geometrical approximations [36], [37]. For an IT2FIS in which the fuzzy sets $\tilde{F}_i^k$ and $\tilde{G}^k$ are interval fuzzy sets, the computing ofT2FISS can be greatly simplified. The membership grades of interval fuzzy sets can be fully characterised by their lower and upper membership grades of the footprint of uncertainty (FOU) separately [2]. Without loss of generality, let $\mu_{\tilde{F}_i^k}(x) = \left[\underline{\mu}_{\tilde{F}_i^k}(x),\ \overline{\mu}_{\tilde{F}_i^k}(x)\right]$ and $\mu_{\tilde{G}^k}(y) = \left[\underline{\mu}_{\tilde{G}^k}(y),\ \overline{\mu}_{\tilde{G}^k}(y)\right]$ for each sample $(x,\ y)$. The firing strength of an IT2FIS $\mu_{\tilde{F}^k}(x) = \sqcap_{i=1}^n \mu_{\tilde{F}_i^k}(x)$ is an interval [2], i.e.,

$$\mu_{\tilde{F}^k}(x) = \left[\underline{f}^k(x),\ \overline{f}^k(x)\right], \quad (2)$$

where

$$\underline{f}^k(x) = \underline{\mu}_{\tilde{F}_1^k}(x) * \cdots * \underline{\mu}_{\tilde{F}_n^k}(x) \quad (3)$$

and

$$\overline{f}^k(x) = \overline{\mu}_{\tilde{F}_1^k}(x) * \cdots * \overline{\mu}_{\tilde{F}_n^k}(x), \quad (4)$$

where * is a t-norm operator. In this paper, the singleton fuzzifier is used in the type-2 fuzzy inference process. The centroid of the type-2 interval consequent set $\tilde{G}^k$ is an interval calculated as follows,

$$C_{\tilde{G}^k} \triangleq [\underline{y}^k, \overline{y}^k] = \left\{ \frac{\sum_{i=1}^N y_i\theta_i}{\sum_{i=1}^N \theta_i} \ \middle|\ \theta_1 \in \left[\underline{\mu}_{\tilde{G}^k}(y_1),\ \overline{\mu}_{\tilde{G}^k}(y_1)\right], \cdots, \theta_N \in \left[\underline{\mu}_{\tilde{G}^k}(y_N),\ \overline{\mu}_{\tilde{G}^k}(y_N)\right] \right\} \quad (5)$$

for the discretised y-domain $\{y_1, \cdots, y_N\}$. The IT2FIS output set via type-reduction, $Y_c(x)$, is also an interval set having the following structure,

$$Y_c \triangleq [y_l, y_r] = \left\{ \frac{\sum_{i=1}^K y^i f^i}{\sum_{i=1}^K f^i} \ \middle|\ y^1 \in [\underline{y}^1, \overline{y}^1], \cdots, y^K \in [\underline{y}^K, \overline{y}^K], f^1 \in [\underline{f}^1, \overline{f}^1], \cdots, f^K \in [\underline{f}^K, \overline{f}^K] \right\} \quad (6)$$

Then the defuzzified output of the IT2FIS is

$$y = \frac{y_l + y_r}{2} \quad (7)$$

However, special attention should be paid to the calculations of the end points of $Y_c(x)$, y_l and y_r. From (6), we see that y_l and y_r can be expressed separately as follows [2]:

$$y_l = \frac{\sum_{i=1}^K \underline{y}^i f_l^i}{\sum_{i=1}^K f_l^i} = \sum_{i=1}^K \underline{y}^i p_l^i \quad (8)$$

$$y_r = \frac{\sum_{i=1}^K \overline{y}^i f_r^i}{\sum_{i=1}^K f_r^i} = \sum_{i=1}^K \overline{y}^i p_r^i \quad (9)$$

where $f_l^i = \underline{f}^i$ or $\overline{f}^i$ contributing to y_l, $f_r^i = \underline{f}^i$ or $\overline{f}^i$ contributing to y_r, $p_l^i = f_l^i / \sum_{i=1}^K f_l^i$ and $p_r^i = f_r^i / \sum_{i=1}^K f_r^i$. Hence, it is necessary to determine which of $\left\{\underline{f}^i\right\}_{i=1}^K$ and $\left\{\overline{f}^i\right\}_{i=1}^K$ contribute to y_l and which of $\left\{\underline{f}^i\right\}_{i=1}^K$ and $\left\{\overline{f}^i\right\}_{i=1}^K$ contribute to y_r. This can be done by the Karnik-Mendel procedure developed in [2].

Given the data set $\left\{x^{(i)}, y^{(i)}\right\}_{i=1}^S$ for designing an IT2FIS,

the back-propagation algorithm can be used to train an IT2FIS such that the following mean-square error (MSE) is minimised:

$$e = \frac{1}{2}\sum_{t=1}^{S}\left(y(x^{(t)}) - y^{(t)}\right)^2 \tag{10}$$

This differentiable MSE measure is a common error minimization heuristic used in system modelling. The validity of using the MSE as an objective function to minimize error relies on the assumption that system outputs are offset by inherent Gaussian noise. For regression tasks, in which the objective is to approximate the function of an arbitrary signal, this presumption often holds. However, this assumption may be invalid for some classification tasks, where other error metrics such as cross-entropy [38] or maximal class margin [39] may be more suited. Our paper focuses on regression-type problems.

In IT2FIS modelling, currently the primary membership functions are usually defined by correspondingly blurring type-1 fuzzy membership functions, such as triangular functions, trapezoid functions and Gaussian functions, in different ways. Among these, Gaussian primary membership functions are the most widely used ones [2]. This is because it is necessary to compute the derivatives in the back-propagation algorithm for tuning IT2FIS parameters. Computing such derivatives is much more challenging than in a type-1 fuzzy logic system. Obviously, the continuity and differentiability of Gaussian primary membership functions make the computation involved much more accessible to fuzzy logic system designers [21], [40]. However, the proposed methods in the next sections of this paper are applicable to any kind of type-2 membership functions, because the proposed rule ranking and rule selection in this paper are performed without any need to specify the nature of the type-2 membership functions.

For the sake of constructing an initial IT2FIS model in this paper, we use Gaussians with uncertain means as the primary membership functions of the type-2 fuzzy sets $\tilde{F}_i^k$ and $\tilde{G}^k$ [2], i.e.,

$$\mu_{\tilde{F}_i^k}(x_i) = \exp\left(-\frac{1}{2}\left(\frac{x_i - m_i^k}{\sigma_i^k}\right)^2\right) \quad m_i^k \in [m_{i1}^k, m_{i2}^k] \tag{11}$$

and

$$\mu_{\tilde{G}^k}(y) = \exp\left(-\frac{1}{2}\left(\frac{y - m^k}{\sigma^k}\right)^2\right) \quad m^k \in [m_1^k, m_2^k] \tag{12}$$

The back-propagation method can then be used to tune the antecedent and consequent parameters in (11) and (12) so as to minimise the mean-square error (10). Details of this training method can be found in [2] [21].

One issue arising is that the type-2 fuzzy rule-base generated by the back-propagation method will almost certainly suffer from rule redundancy, so we need to select the most important fuzzy rules and remove the redundant ones from the generated rule-base. Liang and Mendel applied the SVD-QR with column pivoting algorithm to an IT2FIS for generating a compact type-2 rule-base by removing the redundant rules [26], but it is necessary to estimate an efficient rank r_n in this algorithm. In our experiments using the SVD-QR algorithm for rule ranking, it was found that different r_n produced different rule ranking results. Hence, in order to avoid the estimation of rank r_n, we propose to apply the pivoted QR decomposition method to rule ranking, further suggest the indices of c-values, ω_1-values and ω_2-values of fuzzy rules for rule ranking from different aspects of IT2FIS with the goal of constructing a parsimonious model.

III. New Rule Ranking Indices

A. The R-values of fuzzy rules

The R-values of fuzzy rules are obtained by applying the pivoted QR decomposition method to the firing strength matrices. The idea behind this method is to assign a rule significance index to each fuzzy rule, then rank and select the influential fuzzy rules in terms of this index.

However, one issue arising in designing an IT2FIS is that the fuzzy rule sequence in the rule-base is changed after calculating y_r and y_l. This is because in the Karnik-Mendel procedure [2], it is necessary to arrange the $\{\overline{y}^i\}_{i=1}^K$ in ascending order for calculating y_r (and, similarly, to arrange the $\{\underline{y}^i\}_{i=1}^K$ in ascending order for calculating y_l). Hence, in applying the pivoted QR decomposition method, it is necessary to restore the fuzzy rule sequence when calculating the firing strength matrices [2]. Specifically, for y_r, let the original rule order be $I = [1,\ 2, \cdots, K]^T$. After re-ordering $\{\overline{y}^i\}_{i=1}^K$ in ascending order, the rule order becomes $I' = \Phi I$ (where Φ is a permutation matrix). Next, the rule order in I' is re-numbered as $1,\ 2, \cdots, K$ for calculating y_r by the Karnik-Mendel procedure. Then the firing strength matrices in the original order I are calculated as follows.

First, the number $\hat{K}$ determined in the Karnik-Mendel procedure is very important, because for $i \le \hat{K}$, $f_r^i = \underline{f}^i$, and for $i > \hat{K}$, $f_r^i = \overline{f}^i$. Thus,

$$y_r = \frac{\sum_{i=1}^{\hat{K}} \overline{y}^i \underline{f}^i + \sum_{i=\hat{K}+1}^{K} \overline{y}^i \overline{f}^i}{\sum_{i=1}^{\hat{K}} \underline{f}^i + \sum_{i=\hat{K}+1}^{K} \overline{f}^i} \tag{13}$$

So,

$$p_r^i = \begin{cases} \underline{f}^i \Big/ \left(\sum_{i=1}^{\hat{K}} \underline{f}^i + \sum_{i=\hat{K}+1}^{K} \overline{f}^i\right) & i \le \hat{K} \\ \bar{f}^i \Big/ \left(\sum_{i=1}^{\hat{K}} \underline{f}^i + \sum_{i=\hat{K}+1}^{K} \overline{f}^i\right) & i > \hat{K} \end{cases} \tag{14}$$

Then a firing strength vector given an input x is obtained by restoring the original rule order,

$$p(x) = \Phi^{-1}\left[p_r^1, \cdots, p_r^{\hat{K}}, p_r^{\hat{K}+1}, \cdots, p_r^K\right]^T \tag{15}$$

So the S training samples $\{x^{(i)}, y^{(i)}\}_{i=1}^S$ lead to S firing strength vectors composing a firing strength matrix P_r of size $K \times S$

$$P_r = \left[p\big(x^{(1)}\big), \cdots, p\big(x^{(S)}\big)\right]^T \tag{16}$$

Finally, the QR with column pivoting algorithm addressed in the following steps is applied to P_r, in which each rule is assigned an R-value as its significance index value.

Step 1. Calculate the QR decomposition of P_r and get the permutation matrix Π via $P_r\Pi = QR$, where Q is a unitary matrix and R is an upper triangular matrix. The absolute values of the diagonal elements of R, denoted by $|R_{ii}|$, decrease as i increases and are termed R-values.

Step 2. Rank the fuzzy rules and find their positions in the rule-base in terms of the R-values and the permutation matrix Π. Each column of Π has one element taking the value one and all the other elements taking the value zero. Each column of Π corresponds to a fuzzy rule. The numbering of the jth most important rule in the original rule-base is indicated by the numbering of the row where the element of the jth column containing one is located. For example, if the value one in the 1^{st} column is in the 4^{th} row, then the 4^{th} rule is the most important one and its importance is measured as $|R_{11}|$. The rule corresponding to the first column is the most important, and in descending order the rule corresponding to the last column is the least important.

Because the R-values of P_r tend to approach to singular values of P_r, they can be used as a rule ranking index in designing a T2FIS with a compact rule-base. In practice, we usually use the normalised R-values defined as follows to perform the rule ranking:

$$R_n^i = \frac{|R_{ii}|}{\max\limits_i |R_{ii}|} \tag{17}$$

In terms of the normalised R-values obtained by applying QR with column pivoting to P_r, the available fuzzy rules in rule-base are ranked. Let Ω_r denote the set with these ranked rules.

Similarly, the fuzzy rules are ranked in terms of the normalised R-values obtained by applying QR pivoted decomposition to the firing strength matrix P_l for y_l. Let Ω_l denote the set with these ranked rules. Hence, two sets of ranked fuzzy rules, Ω_r and Ω_l, are obtained.

B. The c-values of fuzzy rules

Both the SVD-QR with column pivoting method and pivoted QR method only take into account the rule-base structure, focusing on the rule antecedent parts when applied to rule reduction of IT2FIS. In the following, we further propose an approach to ranking type-2 fuzzy rules based on the effects of rule consequents, $\tilde{G}^i$.

As a matter of fact, it can be seen from the procedure for designing IT2FIS described in the above section that for each type-2 fuzzy rule, the magnitude of the left and right end-points of the centroid of the consequent set $\tilde{G}^i$, $\left|\underline{y}^i\right|$ and $\left|\overline{y}^i\right|$, separately determine the strength of the effects of the rule consequent on the output end-points y_l and y_r. Hence, $\left|\underline{y}^i\right|$ and $\left|\overline{y}^i\right|$ are very useful indices for measuring the output contributions of type-2 fuzzy rules. These $\left|\underline{y}^i\right|$ and $\left|\overline{y}^i\right|$ are called the c-values of type-2 fuzzy rules in this paper.

In practice, these c-values of fuzzy rules, $C^{(i)} = \left|\overline{y}^i\right|$ or $\left|\underline{y}^i\right|$, are usually normalised as well for rule ranking:

$$C_n^i = \frac{C^{(i)}}{\max\limits_i C^{(i)}} \tag{18}$$

In the rule ranking process, firstly, the normalised c-values $\left|\overline{y}^i\right|$ are used as rule ranking indices for calculating y_r. Let Ω_r denote the set with ranked rules in terms of $\left|\overline{y}^i\right|$. Secondly, the fuzzy rules are ranked in terms of the normalised c-values $\left|\underline{y}^i\right|$ for calculating y_l, and let Ω_l denote the set with ranked rules in terms of $\left|\underline{y}^i\right|$.

C. The ω_1-values and ω_2-values of fuzzy rules

Although type-2 fuzzy rule ranking by c-values takes into account the output contribution of the rule consequents, it ignores the rule-base structure. In order to consider both the rule-base structure and the consequent contribution of fuzzy rules for rule ranking, another two new types of rule ranking indices, termed the ω_1-values and ω_2-values, are separately suggested as follows. Firstly,

$$\omega_1^i = C_n^i \cdot R_n^i \tag{19}$$

where $C^{(i)} = \left|\underline{y}^i\right|$ (or $\left|\overline{y}^i\right|$) for y_l (or y_r), and $|R_{ii}|$ are the R-values of P_l (or P_r). Secondly,

$$\omega_2^i = \min\left(C_n^i, R_n^i\right) \tag{20}$$

where $C^{(i)} = \left|\underline{y}^i\right|$ (or $\left|\overline{y}^i\right|$) for y_l (or y_r), and $|R_{ii}|$ are the R-values of P_l (or P_r).

The choice behind the definition of the ω_1 in (19) lies in that a fuzzy rule ranking result considering both rule-base structure and contributions of rule consequents is expected to be a monotonically increasing function of ranking results from individual aspects (i.e., from rule-base structure or contributions of rule consequents), a higher (lower) ranking result from one aspect leads to a higher (lower) ω_1-value. However, a higher ranking result in terms of ω_2-value defined in (20) is obtained only if both its two operands are higher.

To obtain the rule ranking, firstly the ω_1- (or ω_2-) values are calculated by choosing $C^{(i)}$ as $\left|\overline{y}^i\right|$ for y_r and $|R_{ii}|$ of P_r in (19) (or (20)). Let Ω_r denote the set of rules ranked in terms of these calculated ω_1- (or ω_2-) values. Secondly, the ω_1- (or ω_2-) values are calculated by choosing $C^{(i)}$ as $\left|\underline{y}^i\right|$ for y_l and $|R_{ii}|$ of P_l in (19) (or (20)). Let Ω_l denote the set of rules ranked in terms of these calculated ω_1- (or ω_2-) values.

IV. Type-2 Fuzzy Rule Selection and Implementation

A. Type-2 Fuzzy Rule Selection Procedures

Let D_T be the test dataset and D_V be the validation dataset. Given a type-2 fuzzy model *T2FIS** constructed from data by the back-propagation algorithm, the generalisation performance of the *T2FIS** is measured in terms of the $Err_t^{(*)}$ obtained by applying the model to the testing samples in the data set D_T, whilst the $Err_v^{(*)}$ denotes the validation performance obtained by applying to the validation samples in the data set D_V. The R-values, c-values, ω_1-values and ω_2-values can be used to identify the most influential type-2 fuzzy rules in the *T2FIS**. Assume Ω_l and Ω_r are the rule ranking results obtained for calculating y_l and y_r separately in terms of the chosen rule ranking index:

$$\Omega_l = \left\{\Omega_l(1), \Omega_l(2), \cdots, \Omega_l(K)\right\}$$

$$\Omega_r = \left\{ \Omega_r(1), \Omega_r(2), \cdots, \Omega_r(K) \right\}$$

where K denotes the number of fuzzy rules in rule base of the initial fuzzy model *T2FIS**. The rule importance denoted by $\Omega_l(i)$ or $\Omega_r(i)$ decreases as $i = 1, 2, \cdots, K$; correspondingly, the rule redundancy denoted by $\Omega_l(i)$ or $\Omega_r(i)$ decreases as $i = K, K-1, \cdots, 1$. In this paper, two type-2 fuzzy rule selection procedures, termed forward selection (FS) and backward elimination (BE), are described, based on the proposed rule ranking indices. Given an error tolerance threshold e_h, the two procedures are used to determine the smallest possible model that explains the available data well from the rule base of the model *T2FIS**. The threshold e_h lies between $Err_v^{(*)}$ and $\widetilde{Err}_v$, i.e., $e_h \in [Err_v^{(*)}, \widetilde{Err}_v]$, where $\widetilde{Err}_v$ denotes the maximal validation RMSE of type-2 fuzzy models with one rule in rule-bases. The choice of $e_h > Err_v^{(*)}$ implies that one is ready to sacrifice some system approximation ability for the sake of obtaining a more compact rule-base for a fuzzy model. The FS procedure acts to directly select the influential fuzzy rules from Ω_l and Ω_r, whilst the BE procedure acts to remove the redundant rules from Ω_l and Ω_r (correspondingly retaining the important fuzzy rules). Let Σ_i be the rule subset selected recursively.

1) The FS Procedure: The FS procedure is a heuristics for rule selection which starts with an empty set of type-2 fuzzy rules (i.e. $\Sigma_0 = \emptyset$). One at a time, the most important type-2 fuzzy rules from Ω_l and Ω_r are added (separately) to Σ_i (the set of selected rules), the validation root-mean-square-error (RMSE) $Err_v^{(i)}$ of the fuzzy model constructed by Σ_i tends to be smaller. This rule selection process continues until the chosen criterion (the validation RMSE of the fuzzy model) is below a model error tolerance threshold, e_h. As indicated in Figure 1, the FS procedure consists of the following steps:

Step 1. Set $\Sigma_0 = \emptyset$, $i = 1$, and assign a model error tolerance threshold e_h.

Step 2. Select the most important type-2 fuzzy rules from Ω_l and Ω_r as follows:

$$\Sigma_i = \Sigma_{i-1} \cup \left\{ \Omega_l(i) \right\} \cup \left\{ \Omega_r(i) \right\}$$

where $\Omega_l(i)$ and $\Omega_r(i)$ are the i^{th} most important rules in Ω_l and Ω_r respectively.

Step 3. Construct a type-2 fuzzy model *T2FIS*$_i$ by using the rules in Σ_i;

Step 4. Apply *T2FIS*$_i$ to the validation dataset D_V and the test dataset D_T to obtain new RMSEs: $Err_v^{(i)}$ and $Err_t^{(i)}$;

Step 5. If $Err_v^{(i)} \leq e_h$, stop the selection process and use *T2FIS*$_i$ as the final compact fuzzy model with selected rule set Σ_i, and with $Err_t^{(i)}$ as the measure of generalization performance for *T2FIS*$_i$; otherwise, increase i by 1, and go to *Step 2.*

It should be noted that, because $e_h \geq Err_v^{(*)}$, the termination of the FS procedure is guaranteed, in the sense that the FS procedure always finds a non-empty set of influential rules Σ_i. At least, the initial model *T2FIS** satisfies the termination condition.

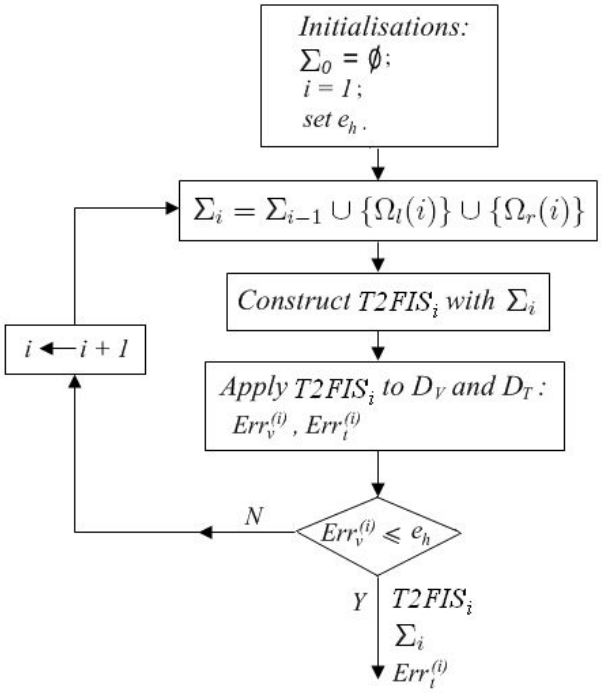

Fig. 1. Forward selection procedure for selecting important fuzzy rules.

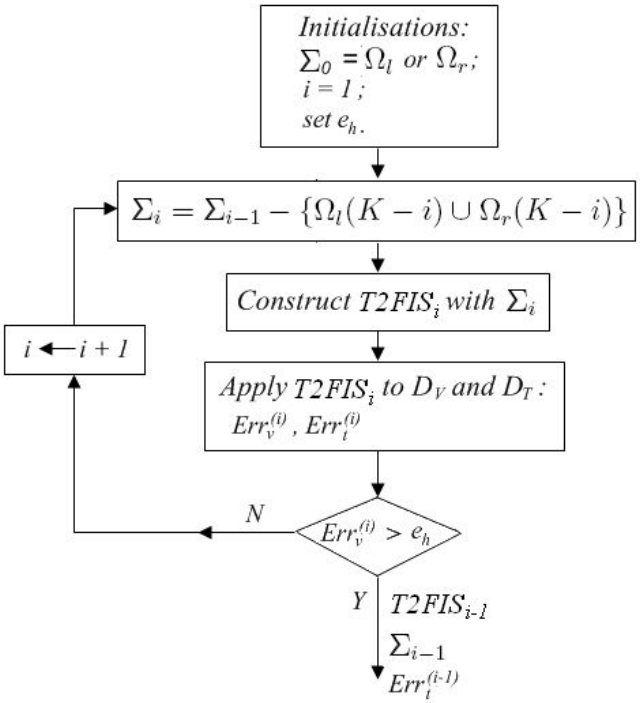

Fig. 2. Backward elimination procedure for removing redundant fuzzy rules.

2) The BE Procedure: The BE procedure is a heuristic for removing redundant rules while retaining influential rules. It works by starting from the full set of ranked rules, i.e. $\Sigma_0 = \Omega_l$ (or Ω_r), for the initial fuzzy model *T2FIS**. One at a time, the most redundant type-2 fuzzy rules from Ω_l and Ω_r are deleted (separately) from Σ_i (the set of selected rules), the validation

RMSE $Err_v^{(i)}$ of the fuzzy model constructed by the selected rules in Σ_i tends to be greater. This rule reduction process continues until the chosen criterion (the validation RMSE of the fuzzy model) is above a model error tolerance threshold, e_h. As indicated in Figure 2, the BE procedure consists of the following steps:

Step 1. Set $\Sigma_0 = \Omega_l$ (or Ω_r), $i = 1$, and assign a model error tolerance threshold e_h.

Step 2. Remove the most redundant type-2 fuzzy rules by performing a difference of rule subsets as follows:

$$\Sigma_i = \Sigma_{i-1} - \left\{ \{ \Omega_l(K+1-i) \} \cup \{\Omega_r(K+1-i)\} \right\}$$

Step 3. Construct a type-2 fuzzy model ***T2FIS***$_i$ by using the rules in Σ_i;

Step 4. Apply ***T2FIS***$_i$ to the validation dataset D_V and the test dataset D_T to obtain new RMSEs: $Err_v^{(i)}$ and $Err_t^{(i)}$;

Step 5. If $Err_v^{(i)} > e_h$ or $i = K$ (the full number of rules), stop the reduction process and use ***T2FIS***$_{i-1}$ as the final compact fuzzy model with selected rule set Σ_i, and with $Err_t^{(i-1)}$ as the measure of generalization performance for ***T2FIS***$_{i-1}$; otherwise, increase i by 1, and go to *Step 2.*

Because $Err_v^{(*)} \leq e_h \leq \widetilde{Err}_v$, the termination of the BE criteria is enforced when the rule set has been reduced to just $\{\Omega_l(1) \cup \Omega_r(1)\}$. In practice, if this occurred, then the BE procedure could be repeated with a lower e_h, while still ensuring $e_h \geq Err_v^{(*)}$.

As a further comment, it should be noted that an idiosyncrasy of type-2 rule ranking is that the two ranked sets Ω_l and Ω_r, obtained by calculating y_l and y_r separately, may be different. Consequently, the final fuzzy rule selection results obtained through the FS and BE procedures described above may be different.

B. Implementation of The Proposed Procedures

Given a training dataset $\left\{x^{(i)}, y^{(i)}\right\}_{i=1}^{S}$, the process of applying these proposed procedures is as follows.

Step 1. Set the initial mean and width parameters for the Gaussian primary functions with uncertain means in an initial input space partition.

Step 2. Train the IT2FIS model by applying the back-propagation algorithm.

Step 3. Rank the fuzzy rules of the trained IT2FIS model in terms of the selected rule ranking index, such as one of the proposed indices introduced in Section III.

Step 4. Conduct one or both of the FS and BE procedures described in Subsection IV-A to obtain the reduced type-2 fuzzy model(s).

It should be noted that not only the proposed indices in Section III, but also the rule ranking index obtained by the SVD-QR method [26] can be used in the above procedures. Indeed, the rule selection process used in the SVD-QR method is the same as the FS procedure using the rule ranking results obtained by SVD-QR decomposition.

C. Applications to Type-1 Mamdani Fuzzy Systems

It is known that an interval type-2 Mamdani fuzzy logic system is a generalisation of a type-1 Mamdani system (T1MFIS). If the lower and upper membership grades of the FOU in a IT2FIS are equal, then the IT2FIS reduces to the T1MFIS. Hence, the proposed methods can be applied to T1MFIS as well by treating the T1MFIS as a special case of the IT2FIS.

Specifically speaking, for a T1MFIS the $K \times S$ firing strength matrices P_l and P_r are the same (i.e. $P_l = P_r$), so only one group of R-values used as rule ranking index is obtained for constructing the T1MFIS with a compact rule base. Similarly, the left end-points of the centroid of the consequent set $\tilde{G}^i$, $\underline{y}^i$, will be equal to the right end-points, $\overline{y}^i$ (i.e. $\underline{y}^i = \overline{y}^i$). Consequently, they both reduce to the traditional centroid of the consequent set obtained by defuzzification of the T1MFIS. Hence the traditional centroids of the consequent sets in a T1MFIS are the c-values of type-1 fuzzy rules for measuring the strength of the effects of the rule consequents on the output. Accordingly, there is only one group of ω_1-values (ω_2-values) for considering the output contribution of rule consequents as well as the rule-base structure. As the sets of ranked rules Ω_l and Ω_r are the same, the final rule selection results obtained through the FS and BE procedures are the same. That is to say, either the FS or BE procedures can be used in identifying the influential type-1 fuzzy rules for the T1MFIS, with the same results.

V. Experiments

In this section, we use three examples to evaluate the proposed rule reduction methods for constructing parsimonious IT2FISs, and compare with the established *SVD-QR with column pivoting* method for rule ranking. Furthermore, type-1 fuzzy logic models are compared with the constructed IT2FIS models in order to examine the benefits of the type-2 approach. The first example is to recover an original signal from data highly contaminated by noise. Although the modelled system in this example seems simple, it is known in the signal processing community that it is rather challenging to recover the original signal from data highly contaminated by noise, without prior knowledge. The second example is a real world problem in which we wish to predict automobile fuel consumption in MPG (miles per gallon) based on several attributes of an automobile's profile. Automobile MPG prediction is a typical nonlinear regression problem. The third example considers a liquid-saturated steam heat exchanger [45]. The main motivation for the choice of the heat exchange process is that this plant is a significant benchmark for nonlinear control design purposes, because it is characterised by a non-minimum phase behaviour which makes the design of suitable controllers particularly challenging even in a linear design context [45]. Hence, it is highly suitable as a context in which an IT2FIS approach is used to predict the system behaviours.

A. Signal Recovery Problem

In the experiments for this example, the noisy signal is generated by

$$y(t) = \tilde{v}(t) + \tilde{\theta}(t) \tag{21}$$

where $\tilde{v}$ is the original signal and $\tilde{\theta}$ is an interference signal. The original signal is generated by

$$\tilde{v}(t) = \sin\left(40/(x(t) + 0.03)\right) + x(t-1)/10. \tag{22}$$

The interference signal is generated from another Gaussian noise source $\tilde{n}$, with a mean of zero and a standard deviation of one, via an unknown nonlinear process

$$\tilde{\theta}(t) = 4\sin\left(\tilde{n}(t)\right)\tilde{n}(t-1)/\left(1+\tilde{n}(t-1)^2\right) \qquad (23)$$

The measured signal y is the sum of the original information signal $\tilde{v}$ and the interference $\tilde{\theta}$. However, we do not know the interference signal $\tilde{\theta}$. The only signal available to us is the measured signal y. The task is to learn the characteristics of the original information signal $\tilde{v}$ from the measured signal y, then recover the original signal. A T2FIS may be suitable for such a signal processing problem due to its strong capability of characterising higher uncertainty exhibited within such noisy data [12].

1) Initial type-2 fuzzy model: In the following, we constructed an initial IT2FIS model with two inputs $x(t), x(t-1)$ selected in terms of regularity criteria [43] and one output $y(t)$. The antecedent and consequent parameters in (11) and (12) were optimised by the back-propagation algorithm [2] [21]. In order to train the interval type-2 fuzzy model to represent the nonlinearity and higher uncertainty of the system, the data generation process (21), (22) and (23) was run 10 times. In each of the runs, 100 samples $\left\{x^{(h)}, y_k^{(h)}\right\}_{h=1}^{S}$ (S=100, $k = 1, \cdots, 10$) were generated with $x^{(h)} \in [2,\ 5]$ and $y_k^{(h)}$ obtained by (21). Then the data set $\left\{x^{(h)}, \min\limits_k y_k^{(h)}\right\}_{h=1}^{100}$ was used to generate the antecedent means $\{m_{i1}\}$ in (11) by the fuzzy c-means (FCM) unsupervised clustering algorithm [41], whilst the data set $\left\{x^{(h)}, \max\limits_k y_k^{(h)}\right\}_{h=1}^{100}$ was used to generate the means $\{m_{i2}\}$. The consequent means $\{m_1\}$ and $\{m_2\}$ in (12) were randomly selected from the output data samples. The width parameters in (11) and (12) were determined using the nearest neighbour heuristic suggested by Moody and Daken [42], based on the corresponding data sets. Hence, all the initial antecedent and consequent parameters were determined from given data sets, rather than manually derived.

Four initial type-2 fuzzy sets were generated for each input variable, which led to 16 rules in the initial interval type-2 fuzzy model. After the training process, this interval type-2 fuzzy model had the ability to recover the original signal well with RMSE 0.2739, as illustrated in Figure 3. Further evaluation of the generalisation performance of this trained model was undertaken in order to examine whether, given system inputs which were different from the training samples, the outputs of the trained model could follow the characteristics of the original signal rather than act as the noisy signal. A test dataset with 40 testing inputs that were different from the above available training inputs was generated. The generalisation performance of the trained fuzzy model was measured by the RMSE of the original signal values calculated by (22) and the trained model outputs given these testing inputs.

In our experiments, this type-2 fuzzy model achieved a generalisation performance with the RMSE 0.1178 when being applied to the testing samples.

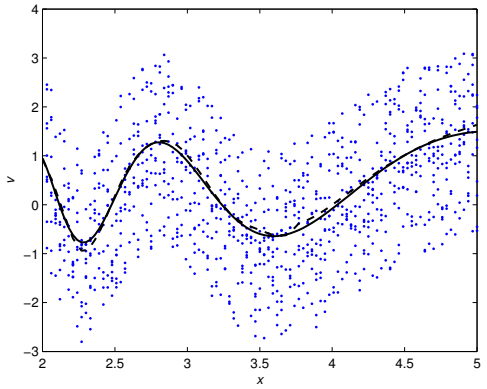

Fig. 3. Signal recovering by IT2FIS model on training samples: the dots represent the measured signal, solid line represents the original signal $\tilde{v}$ and dashed line represents the recovered signal.

2) Type-2 fuzzy rule ranking: Next, we applied the proposed rule reduction methods to the trained interval type-2 fuzzy model, and evaluated them in comparison with the SVD-QR rule reduction method. First, the QR with column pivoting algorithm was applied to the firing strength matrices P_l and P_r, in which the R-values of fuzzy rules were generated for selecting influential rules in calculating y_l and y_r separately. Figure 4 depicts the corresponding R-values and singular values of fuzzy rules in descending order obtained from the firing strength matrices P_l and P_r individually. These indicate that the R-values track the singular values well, so the R-values of P_l and P_r can be used to rank the fuzzy rules. Figure 5 illustrates the ω_1-values, normalised R-values and c-values of fuzzy rules with the rule orders in rule base, whilst Figure 6 depicts the ω_2-values, normalised R-values and c-values of fuzzy rules with the rule orders in rule base. These two figures show that there exist certain correlations among the proposed indices. However, it can be seen that each rule has different index values, and that a rule with higher R-value does not necessarily mean it has higher c-value, ω_1-value, or ω_2-value, etc. Hence, these indices evaluate the importance of fuzzy rules in their own ways. Table I illustrates the rule ranking results obtained in terms of the proposed indices: the *R-values, c-values, ω_1-values* and *ω_2-values* of fuzzy rules obtained when applied to the firing strength matrices P_l and P_r. As a comparison, the rule ranking results obtained by the SVD-QR with column pivoting algorithm are summarised in Table II. It can be seen that the rule ranking obtained by SVD-QR with column pivoting algorithm indeed depends on the rank r_n.

3) Type-2 fuzzy rule selection by the FS procedure: In this signal recovery problem, because the original signal remained unknown and the only available data was the measured signal, the noisy data was used as the validation dataset during the rule selection process. The testing dataset was the above dataset with 40 testing samples. The model performance was measured by the RMSEs of a constructed fuzzy model on the two datasets, Err_v and Err_t. The model parsimony was evaluated in terms of the number of fuzzy rules in the rule-

TABLE I
THE RULE RANKING RESULTS BY THE PROPOSED RULE INDICES IN THE SIGNAL RECOVERING PROBLEM

Indices	*Matrices*	*Rule Ranking Results*
R-values	P_r	16 11 15 1 7 10 6 12 5 9 2 14 3 13 8 4
	P_l	11 16 15 1 6 10 2 12 7 5 3 14 9 8 13 4
c-values	P_r	9 1 16 11 12 7 10 6 15 5 14 2 8 13 4 3
	P_l	16 1 7 11 12 3 6 2 8 9 5 13 10 14 15 4
ω_1-*values*	P_r	16 11 1 7 15 10 9 6 12 5 2 14 8 13 3 4
	P_l	16 11 1 6 2 7 12 15 10 3 5 9 14 8 13 4
ω_2-*values*	P_r	16 11 1 7 15 10 9 6 12 5 2 14 8 13 3 4
	P_l	16 11 1 6 2 7 12 15 10 3 5 9 14 8 13 4

TABLE II
RULE RANKING RESULTS BY SVD-QR WITH COLUMN PIVOTING IN THE SIGNAL RECOVERING PROBLEM

Algorithms	*Matrices*	*Rule Ranking Results*
SVD-QR($r_n = 4$)	P_r	1 16 12 8 5 6 7 4 9 10 11 3 13 14 15 2
	P_l	1 16 14 13 5 6 7 8 9 10 11 12 4 3 15 2
SVD-QR ($r_n = 5$)	P_r	1 16 12 8 14 6 7 4 9 10 11 3 13 5 15 2
	P_l	1 16 13 12 15 6 7 8 9 10 11 4 3 14 5 2
SVD-QR ($r_n = 6$)	P_r	1 12 16 13 14 7 6 8 9 10 11 2 4 5 15 3
	P_l	1 13 16 12 14 15 7 8 9 10 11 4 2 5 6 3

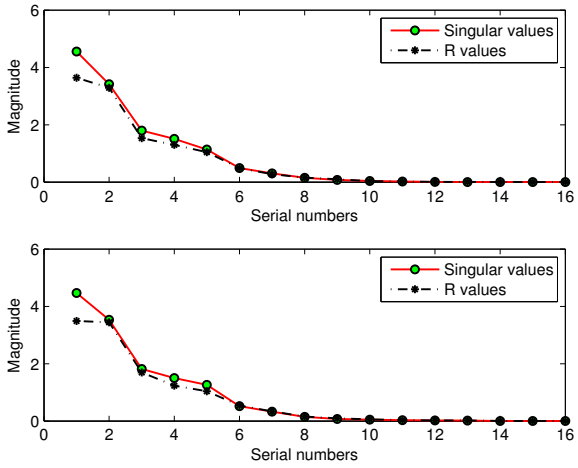

Fig. 4. The R-values and singular values of firing strength matrices P_l (bottom) and P_r (top) in descending order in the signal recovering problem

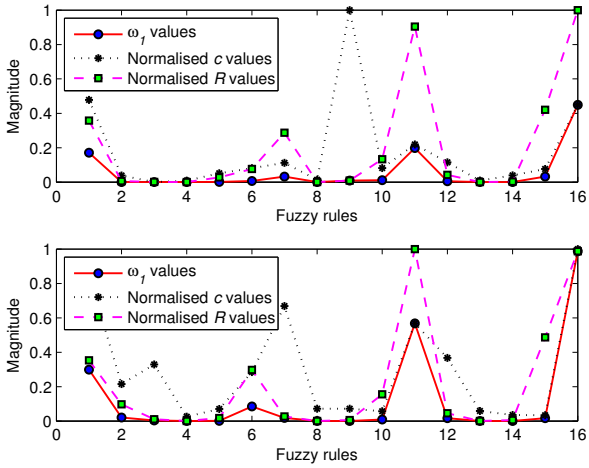

Fig. 5. ω_1-values, normalised c-values and R-values of type-2 fuzzy rules with the rules orders in rule base: ω_1-values with normalised $\left|\overline{y}^i\right|$ and R-values of P_r (top); ω_1-values with normalised $\left|\underline{y}^i\right|$ and R-values of P_l (bottom) in signal recovering problem

base of the constructed model. The RMSE tolerance threshold e_h was set to be 0.3. Table III summarises the rule selection results obtained by applying the FS procedure using each of the proposed indices (R-values, c-values, ω_1-values and ω_2-values). As a comparison, Table IV summarises the rule selection results obtained by applying the FS procedure using the SVD-QR with column pivoting algorithm with $r_n = 4$, 5 and 6. In this signal recovery problem, with the consideration of the consequent effects of trained fuzzy rules on the overall system output, the normalised c-values $\left|\overline{y}^i\right|$ and $\left|\underline{y}^i\right|$ of fuzzy rules depicted in Figure 5 or Figure 6 led to a parsimonious interval type-2 fuzzy model featuring 11 rules, selected from 16 rules. For the normalised R-values shown in Figure 5, ω_1-values shown in Figure 5 and ω_2-values shown in Figure 6, 12 important rules were identified for the final model. As a comparison, 15 fuzzy rules were chosen by the SVD-QR pivoting algorithm with $r_n = 4$ and 5, whilst for $r_n = 6$ the SVD-QR algorithm did not indicate any rules to be removed.

4) Type-2 fuzzy rule selection by the BE procedure: The proposed rule ranking indices were then used in the BE procedure to select the most important fuzzy rules, and the results were compared to those obtained by SVD-QR pivoting algorithm. Table V summarises the rule selection results obtained by applying the BE procedure using each of the proposed indices (R-values, c-values, ω_1-values and ω_2-values). As a comparison, Table VI summaries the rule selection results obtained by applying the BE procedure using the SVD-QR with column pivoting algorithm with $r_n = 4$, 5

TABLE III
RULE SELECTION RESULTS BY FS PROCEDURE WITH THE PROPOSED RULE RANKING INDICES IN THE SIGNAL RECOVERING PROBLEM

R-values			***c-values***			ω_1-***values***			ω_2-***values***		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
8	1.507	1.478	8	1.200	1.185	8	1.507	1.478	8	1.507	1.478
9	1.506	1.477	9	1.148	1.145	9	0.684	0.529	9	1.506	1.477
10	1.206	1.124	10	1.144	1.145	10	0.682	0.529	10	1.206	1.124
11	0.538	0.372	**11**	**0.288**	**0.124**	11	0.538	0.372	11	0.538	0.372
12	**0.275**	**0.128**	12	0.287	0.122	**12**	**0.275**	**0.128**	**12**	**0.275**	**0.128**
13	0.274	0.119	13	0.274	0.126	13	0.274	0.119	13	0.274	0.119

TABLE IV
RULE SELECTION RESULTS VIA FS PROCEDURE WITH THE RULE RANKING BY SVD-QR WITH COLUMN PIVOTING METHOD IN THE SIGNAL RECOVERING PROBLEM

SVD-QR ($r_n = 4$)			**SVD-QR** ($r_n = 5$)			**SVD-QR** ($r_n = 6$)		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
12	1.797	1.817	12	0.947	0.848	12	0.735	0.562
13	1.245	1.201	13	0.735	0.561	13	0.680	0.525
14	1.140	1.140	14	0.301	0.115	14	0.680	0.524
15	**0.276**	**0.100**	**15**	**0.276**	**0.100**	15	0.536	0.367
16	0.274	0.118	16	0.274	0.118	**16**	**0.274**	**0.118**

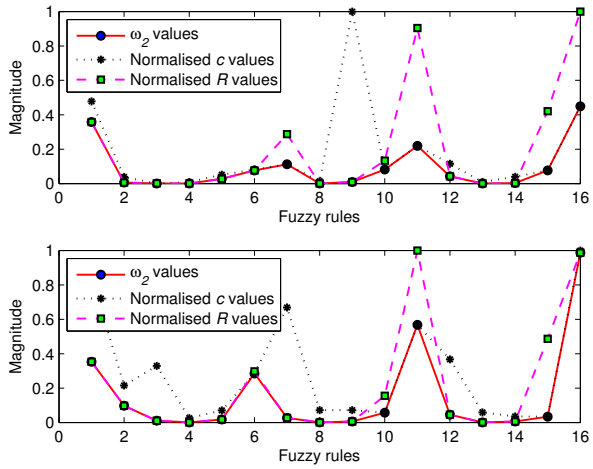

Fig. 6. ω_2-values, normalised c-values and R-values of type-2 fuzzy rules with the rules orders in rule base: ω_2-values with normalised $\left|\overline{y}^i\right|$ and R-values of P_r (top); ω_2-values with normalised $\left|\underline{y}^i\right|$ and R-values of P_l (bottom) in signal recovering problem

and 6. It can be seen that for the R-values and ω_2-values, 13 important rules were identified, whilst the BE procedure with c-values and ω_1-values identified 14 rules. As a comparison, in this procedure 15 fuzzy rules were selected by the SVD-QR pivoting algorithm with $r_n = 4$ and 5. Again, the SVD-QR algorithm with $r_n = 6$ did not indicate any rules to remove.

5) Comparison with a type-1 Mamdani approach: As a further comparison, we constructed a type-1 Mamdani fuzzy system model for this signal recovery problem, in which the initial antecedent means of Gaussian membership functions were the averages of the $\{m_{i1}\}$ and $\{m_{i2}\}$ used for constructing above initial IT2FIS model, the consequent means were randomly selected from the output samples, and the width parameters were determined using the nearest neighbour heuristic [42] based on the given data sets. The trained type-1 Mamdani fuzzy system model recovered the original signal with RMSE 0.3790 on the training samples and produced an RMSE of 0.2551 for outputs on the 40 testing inputs with the corresponding original signal values. By setting the RMSE tolerance threshold e_h for rule selection to be 0.4, the best result of rule selection for constructing a parsimonious type-1 Mamdani fuzzy model was obtained in terms of the c-values of fuzzy rules. In this best model, 11 influential type-1 fuzzy rules were selected from the 16 rules, producing an RMSE of 0.3951 when recovering the signal from the training inputs and an RMSE of 0.2474 for testing samples.

6) Effects of noise changes: In this signal recovery problem, it is natural to want to know what happens if the noise characteristics change, after a type-2 fuzzy model has been trained according the available measured data. As stated above, given a measured signal y, for example generated by (21), the task is to construct a type-2 fuzzy model to learn the characteristics of the original information signal $\tilde{v}$ from the measured signal y, and then recover the original signal. After the training process, the trained type-2 fuzzy model has learned the characteristics of the original signal (22). That is to say, this model possesses the ability to approximate the original signal, given the inputs x. The model output is the recovered signal. In rule selection, rule reduction or other such procedures, no matter how the noise level $\tilde{\theta}$ in (23) is changed again, the signal recovery performance achieved by the trained model will not be changed. This is because the original signal source (22) is not impacted (i.e., the characteristics of the original signal are not changed, only the measured signal (21)

TABLE V
RULE REDUCTION RESULTS BY BE PROCEDURE WITH THE PROPOSED RULE RANKING INDICES IN THE SIGNAL RECOVERING PROBLEM

***R*-values**			***c*-values**			ω_1-***values***			ω_2-***values***		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
15	0.274	0.118	15	0.274	0.118	15	0.274	0.118	15	0.274	0.118
14	0.274	0.118	**14**	**0.288**	**0.224**	**14**	**0.274**	**0.118**	14	0.274	0.118
13	**0.274**	**0.119**	13	0.537	0.367	13	0.537	0.368	**13**	**0.274**	**0.119**
12	0.609	0.573	12	1.230	1.194	12	0.537	0.368	12	0.609	0.573

TABLE VI
RULE REDUCTION RESULTS VIA BE PROCEDURE WITH THE RULE RANKING BY SVD-QR WITH COLUMN PIVOTING METHOD IN THE SIGNAL RECOVERING PROBLEM

SVD-QR ($r_n = 4$)			**SVD-QR** ($r_n = 5$)			**SVD-QR** ($r_n = 6$)		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
16	0.274	0.118	16	0.274	0.118	**16**	**0.274**	**0.118**
15	**0.276**	**0.100**	**15**	**0.276**	**0.100**	15	0.536	0.367
14	1.140	1.140	14	0.301	0.115	14	0.615	0.440

is influenced by the changes of the noise $\tilde{\theta}$). Hence, there is no need to re-train the model. However, in an application of the trained model, the RMSE between the model output and the measured signal given an input will vary along with the change of the noise $\tilde{\theta}$. In our experiments we did not use the RMSE between the model output and the original signal given an input as the training performance measure for the trained model, because the training data hails from the measured signal (21). However, if the original signal source (22) is changed, no matter whether or not the noise level $\tilde{\theta}$ in (23) is changed, the type-2 fuzzy model must be re-trained according to the new measured data. This is because the characteristics of the fundamental signal in the measured data have been changed.

B. MPG Prediction Problem

In this example of automobile MPG prediction, the data collected from automobiles of various makers and models is available in the UCI (Univ. of California at Irvine) Machine Learning Repository. In the available data set with 392 samples, there are six input attributes (*number of cylinders*, *displacement*, *horsepower*, *weight*, *acceleration*, and *model year*) and one output attribute (*the fuel consumption in MPG*). However, only three input variables (*weight*, *acceleration*, and *model year*) were considered here, based on the regularity criterion [43]. The 392 samples were randomly partitioned into a training set (196 samples), a testing set (120 samples), and a validation set (76 samples) for building and evaluating an initial IT2FIS model. The validation set was used in the rule selection and reduction process, whilst the performance of the IT2FIS models constructed was evaluated in terms of the RMSEs on the testing samples.

1) Initial type-2 fuzzy model: Given the input-output data samples, the FCM algorithm generated 3 clusters $\{m_x\}$ according to a partition entropy measure [35]. Then, the means of three initial antecedent interval type-2 fuzzy sets for each input variable were produced in a manner similar to that suggested by Mendel [2]: $[m_x-0.5\sigma_x-5.5\sigma_n,\ \ m_x-0.5\sigma_x+5.5\sigma_n]$ for the weight attribute and $[m_x-0.5\sigma_x-0.25\sigma_n,\ \ m_x-0.5\sigma_x+0.25\sigma_n]$ for the acceleration and year attributes, where σ_x and σ_n are the standard deviation of the training samples and additive noise respectively. The means of consequent interval type-2 fuzzy sets were randomly selected from the output samples, whilst the width parameters in (11) and (12) were chosen as $0.5\sigma_x$. The above initial input-output partition led to 27 rules in an initial interval type-2 fuzzy model. After training by the back-propagation algorithm with training RMSE 2.37, the interval type-2 fuzzy model predicted the fuel consumption of testing samples reasonably well with testing RMSE 2.46.

2) Type-2 fuzzy rule ranking: The proposed rule reduction methods were then applied to the trained interval type-2 fuzzy model. Table VII shows the rule ranking results obtained by the proposed indices. For comparison, the rule ranking results obtained by the SVD-QR algorithm with $r_n = 4$, 5 and 6 are illustrated in Table VIII. Once again, it can be seen that the rule ranking results vary with the choice of r_n.

3) Type-2 fuzzy rule selection by the FS procedure: The most influential type-2 fuzzy rules were selected based on the above rule ranking results via the FS procedure to construct parsimonious type-2 fuzzy models, and the results were compared with those obtained by the SVD-QR pivoting algorithm. The RMSE tolerance threshold e_h was set to 3.0. Table IX summarises the rule selection results obtained using the FS procedure with the proposed indices. Table X summarises the rule selection results using the FS procedure with the SVD-QR pivoting algorithm for $r_n = 4, 5$ and 6. In this real world problem, the R-values and ω_1-values led to a parsimonious interval type-2 fuzzy model constructed with only 6 rules selected from the original 27 rules. For the c-values and ω_2-values, 20 rules were retained. As a comparison, 23, 24 and 25 fuzzy rules were separately identified by the SVD-QR algorithm with $r_n = 4, 5$ and 6, respectively.

TABLE VII
THE RULE RANKING RESULTS BY THE PROPOSED RULE INDICES IN THE MPG PREDICTION PROBLEM

Indices	*Matrices*	*Rule Ranking Results*
R-values	P_r	14 23 15 13 5 6 22 17 18 10 19 24 20 4 8 11 26 12 7 9 16 25 3 21 27 2 1
	P_l	14 23 15 5 13 6 22 17 10 18 19 8 24 4 20 11 26 12 7 9 16 25 3 21 2 27 1
c-values	P_r	15 9 4 14 5 16 8 12 21 3 17 7 25 24 6 26 22 23 19 11 1 13 10 18 20 27 2
	P_l	4 7 6 18 9 3 12 13 8 10 27 21 20 14 5 15 26 25 1 17 23 19 11 2 22 16 24
ω_1-*values*	P_r	14 15 5 23 17 4 6 8 22 24 12 13 19 9 26 7 16 11 10 18 20 25 3 21 27 1 2
	P_l	4 6 14 13 18 15 5 10 23 7 8 20 12 9 17 19 26 11 22 3 25 24 21 16 27 1 2
ω_2-*values*	P_r	14 15 5 6 22 17 23 19 24 4 8 13 11 26 12 10 18 20 7 9 16 25 3 21 27 2 1
	P_l	6 18 13 10 8 4 20 14 5 15 26 12 17 23 7 9 19 11 22 25 16 3 24 21 2 27 1

TABLE VIII
RULE RANKING RESULTS BY SVD-QR WITH COLUMN PIVOTING IN THE MPG PREDICTION PROBLEM

Algorithms	*Matrices*	*Rule Ranking Results*
SVD-QR ($r_n = 4$)	P_r	9 21 17 27 5 6 7 8 1 10 11 12 13 14 15 16 3 18 19 20 2 22 23 24 25 26 4
	P_l	9 21 17 27 5 6 7 8 1 10 11 12 13 14 15 16 3 18 19 20 2 22 23 24 25 26 4
SVD-QR ($r_n = 5$)	P_r	9 21 27 17 6 5 7 8 1 10 11 12 13 14 15 16 4 18 19 20 2 22 23 24 25 26 3
	P_l	9 21 27 17 6 5 7 8 1 10 11 12 13 14 15 16 4 18 19 20 2 22 23 24 25 26 3
SVD-QR ($r_n = 6$)	P_r	9 21 27 26 17 6 7 8 1 10 11 12 13 14 15 16 5 18 19 20 2 22 23 24 25 4 3
	P_l	14 9 27 21 17 6 7 8 2 10 11 12 13 1 15 16 5 18 19 20 4 22 23 24 25 26 3

TABLE IX
RULE SELECTION RESULTS BY FS PROCEDURE WITH THE PROPOSED RULE RANKING INDICES IN THE MPG PREDICTION PROBLEM

R-values			***c-values***			ω_1-***values***			ω_2-***values***		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
4	6.234	5.382	17	3.744	3.938	4	4.455	5.194	17	3.118	3.221
5	4.850	3.719	18	3.744	3.918	5	3.361	4.561	18	3.106	3.098
6	**2.699**	**3.003**	19	3.110	3.060	**6**	**2.699**	**3.233**	19	3.106	2.842
7	2.699	2.879	**20**	**2.700**	**3.059**	7	2.699	2.967	**20**	**2.551**	**2.479**
			21	2.700	3.059				21	2.551	2.479

TABLE X
RULE REDUCTION RESULTS VIA FS PROCEDURE WITH THE RULE RANKING BY SVD-QR WITH COLUMN PIVOTING METHOD IN THE MPG PREDICTION PROBLEM

SVD-QR ($r_n = 4$)			**SVD-QR** ($r_n = 5$)			**SVD-QR** ($r_n = 6$)		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
21	3.045	2.880	22	3.312	2.569	23	3.310	2.568
22	3.044	2.654	23	3.019	2.460	24	3.016	2.460
23	**2.571**	**2.549**	**24**	**2.981**	**2.459**	**25**	**2.978**	**2.460**
24	2.514	2.548	25	2.975	2.459	26	2.970	2.459

4) Type-2 fuzzy rule selection by the BE procedure: Finally, the BE procedure was applied and the results compared with those obtained by the SVD-QR algorithm. The RMSE tolerance threshold e_h was again set as 3.0. Table XI summarises the results obtained using the BE procedure with the proposed indices. Table XII summarises the results obtained using the BE procedure with the SVD-QR algorithm for $r_n = 4$, 5 and 6. In this problem, the ω_1-values resulted in 18 important rules being selected, whilst the R-values, c-values and ω_2-values produced 20, 21 and 23 rules, respectively. As a comparison, 23, 24 and 25 rules were selected by the SVD-QR pivoting algorithm with $r_n = 4$, 5 and 6, respectively. It can be seen from these results that the proposed rule ranking indices outperform the existing SVD-QR with column pivoting algorithm, and are very effective in identifying the significant rules and removing redundant ones.

5) Comparison with a type-1 Mamdani approach and other approaches: For further comparison, we used the same training samples, testing samples and validation samples to construct a parsimonious type-1 Mamdani fuzzy system model in this MPG prediction problem. The initial antecedent means of Gaussian membership functions were the averages of the means used for constructing above initial IT2FIS model, the consequent means were randomly selected from the output

TABLE XI
RULE SELECTION RESULTS BY BE PROCEDURE WITH THE PROPOSED RULE RANKING INDICES IN THE MPG PREDICTION PROBLEM

R-values			***c-values***			ω_1-***values***			ω_2-***values***		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
23	2.562	2.455	23	2.619	2.484	20	2.608	2.480	25	2.971	2.455
21	2.551	2.455	22	2.619	2.583	19	2.608	2.480	24	2.972	2.455
20	**2.551**	**2.479**	**21**	**2.617**	**2.589**	**18**	**2.607**	**2.484**	**23**	**2.972**	**2.455**
19	3.106	2.842	20	3.237	2.901	17	3.228	2.807	22	3.010	2.455
18	3.106	3.098	19	3.249	3.065	16	3.228	2.898	21	3.010	2.455

TABLE XII
RULE REDUCTION RESULTS VIA BE PROCEDURE WITH THE RULE RANKING BY SVD-QR WITH COLUMN PIVOTING METHOD IN THE MPG PREDICTION PROBLEM

SVD-QR ($r_n = 4$)			**SVD-QR** ($r_n = 5$)			**SVD-QR** ($r_n = 6$)		
No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t	No. of Rules	Err_v	Err_t
24	2.514	2.548	25	2.975	2.459	26	2.970	2.459
23	**2.571**	**2.549**	**24**	**2.981**	**2.459**	**25**	**2.975**	**2.459**
22	3.044	2.654	23	3.019	2.460	24	3.031	2.548
21	3.045	2.880	22	3.312	2.569	23	3.036	2.548

samples, and the width parameters were chosen as $0.5\sigma_x$. The trained type-1 Mamdani fuzzy system model predicted the fuel consumptions with an RMSE of 2.4762 on training samples and an RMSE of 2.6348 on testing samples. By setting the RMSE tolerance threshold e_h for rule selection as 3.0, the best result of rule selection for constructing a parsimonious type-1 Mamdani fuzzy model was obtained in terms of the c-values of fuzzy rules, in which 22 influential type-1 fuzzy rules were selected from the 27 rules. This type-1 Mamdani model with the 22 selected rules predicted the testing samples with an RMSE of 3.5954. Many other researchers have also built system modelling algorithm to perform MPG prediction for this problem. Competitive results include those of Kilic *et al* in which they proposed a modelling method with a generalisation performance (RMSE) of 2.61 [44]. It can be seen that the results obtained in our parsimonious type-2 approach are superior to a type-1 approach and comparable, if not a little better than other approaches (although we stress that this cannot be stated as a definitive conclusion).

C. Steam Heat Exchanger

In the last example, water is heated by pressurized saturated steam through a copper tube. Saturated steam is used to provide primary heat to a process fluid in a heat exchanger. The process plant is illustrated in Figure 7. The plant output is the outlet liquid temperature, and the inputs are the liquid flow rate, the steam temperature, and the inlet liquid temperature. In this experiment, the steam temperature and the inlet liquid temperature are kept constant to their nominal values, so we only considered the liquid flow rate as the plant input variable.

In our experiment, 500 heat exchanging samples were used to construct an IT2FIS model with 4 inputs selected in terms of the regularity criterion [43]. That is, $v_t = f(v_{t-1}, v_{t-2}, v_{t-4}, u_t)$, where v_t is the outlet liquid temperature, and u_t the liquid flow rate, at time t. These 500

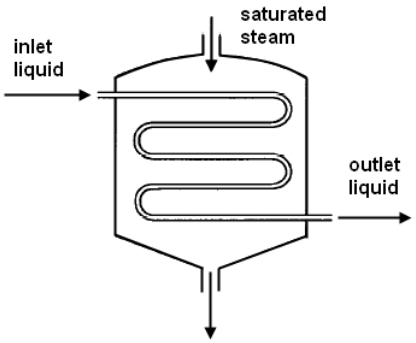

Fig. 7. The structure of the steam heat exchanger

samples were randomly partitioned into 300 training samples, 100 testing samples and 100 validation samples. The training samples were used to build an initial IT2FIS model, the model performance was evaluated in terms of the RMSEs on the testing samples, while the validation samples were used for rule selection and reduction.

1) Initial type-2 fuzzy model: Given the training samples, the FCM algorithm generated two clusters $\{m_x\}$ based on the given input-output samples.

The means of three initial antecedent interval type-2 fuzzy sets for each input variable were again produced in a manner similar to that suggested by Mendel [2], as $[m_x - 0.3\sigma_x - 0.1\sigma_n,\ \ m_x - 0.3\sigma_x + 0.1\sigma_n]$ for the liquid flow rate attribute and $[m_x - 0.3\sigma_x - 0.25\sigma_n,\ \ m_x - 0.3\sigma_x + 0.25\sigma_n]$ for other attributes, where σ_x and σ_n are the standard deviation of the training samples and additive noise, respectively. The means of consequent interval type-2 fuzzy sets were randomly selected from the output samples, whilst the width parameters in (11) and (12) were chosen as $0.3\sigma_x$. The above initial input-output partition resulted in 16 rules in the initial interval type-2 fuzzy

model. After training by the back-propagation algorithm with a training RMSE of 0.2642, the initial type-2 fuzzy model predicted the outlet liquid temperature well for the testing samples with an RMSE of 0.2436.

2) Type-2 fuzzy rule ranking and rule selection: The proposed rule ranking indices were then calculated for the trained interval type-2 fuzzy model. The most influential type-2 fuzzy rules were selected based on the rule ranking results. The best result achieved for constructing a parsimonious type-2 fuzzy rule base was obtained by using the BE procedure with c-value index and $e_h = 1.0$. In this case, eleven rules were selected with a validation RMSE of 0.6276 and a testing RMSE of 0.5698.

3) Comparison with type-1 Mamdani approach: For further comparison, we used the same training, testing and validation samples to construct a parsimonious type-1 Mamdani fuzzy system model for this problem. The initial antecedent means of Gaussian membership functions were the averages of the means used for constructing the initial IT2FIS model described above, the consequent means were randomly selected from the output samples, and the width parameters were chosen as $0.3\sigma_x$. The trained type-1 Mamdani fuzzy system model predicted the outlet liquid temperature with an RMSE of 0.2784 on training samples and an RMSE of 0.2847 on testing samples. By setting the RMSE tolerance threshold e_h for rule selection as 1.0, the best result of rule selection for constructing a parsimonious type-1 Mamdani fuzzy model was obtained with c-value index. In this case, twelve type-1 fuzzy rules were selected with an RMSE of 0.8052 on the testing samples. Again, it can be seen that the type-1 approach was inferior to the type-2 approach.

VI. Discussion and Conclusions

We have proposed some novel indices of type-2 fuzzy rules which focus on different aspects of type-2 fuzzy logic systems in order to determine the relative importance of the various rules. These indices are termed R-values, c-values, ω_1-values and ω_2-values. The R-values of type-2 fuzzy rules obtained by QR decomposition pay attention to the rule-base structure, the c-values focus on contributions of rule consequents, whilst the ω_1-values and ω_2-values take into account both the rule-base structure and contributions of rule consequents. Moreover, two procedures, the FS procedure and BE procedure, have been described for utilising these indices in determining a parsimonious rule-base. The experimental results have demonstrated that parsimonious type-2 fuzzy system models can be effectively constructed in terms of the fuzzy rules selected by these proposed indices, and that the proposed methods outperform the existing *SVD-QR with column pivoting* method for rule reduction. Indeed, in the second example (MPG prediction), the proposed methods selected only six rules to be used, while maintaining a reasonable RMSE of 2.699. In contrast, the best SVD-QR performance resulted in retaining 23 rules with an RMSE of 2.571.

One possible issue arising in the proposed methodology which has not been addressed in this paper is which fuzzy rule index is best among the proposed four indices. In practice, one can simply pick any one of these rule ranking indices to construct a parsimonious type-2 fuzzy model; alternatively, one can run both the FS and BE procedures on each of these indices in turn, then pick the best. This issue certainly merits further research.

Acknowledgment

The author would like to thank the anonymous reviewers for their excellent suggestions that have been incorporated into this paper. This work has been supported by the EPSRC Research Grants EP/C542207/1 and EP/C542215/1.

References

[1] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-1," Information Sciences, vol. 8, pp. 199-249, 1975.

[2] J. M. Mendel, *Uncertainty Rule-based Fuzzy Logic Systems: Introduction and New Directions*, Prentice Hall, Inc., 2001.

[3] J. M. Mendel, and R. I. John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy Systems*, vol.10, no.2, pp.117-127, 2002.

[4] J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Trans. on Fuzzy Systems*, vol.14, no.6, pp.808-821, 2006.

[5] R. John, and S. Coupland, "Type-2 fuzzy logic: a historical view," *IEEE Computational Intelligence Magazine*, vol.2, no.1, pp.57-62, 2007.

[6] J. L. Chaneau, M. Gunaratne, and A. G. Altschaeffl, "An application of type-2 sets to decision making in engineering," in J. C. Bezdek, Ed., *Analysis of Fuzzy Information-Vol. II: Artificial Intelligence and Decision Systems*, Boca Raton, FL: CRC, 1987.

[7] R. R. Yager, "Fuzzy subsets of type II in decisions," *Journal of Cybernetics*, vol.10, pp. 137-159, 1980.

[8] J. M. Garibaldi and T. Ozen, "Uncertain fuzzy reasoning: a case study in modelling expert decision making," *IEEE Transactions on Fuzzy Systems*, vol.15, no.1, pp.16-30, 2007

[9] R. I. John, and P. R. Innocent, "Modeling uncertainty in clinical diagnosis using fuzzy logic," *IEEE Trans. on Systems, Man and Cybernetics-part B*, vol.35, no.6, pp.1340-1350, 2005.

[10] L. Di Lascio, A. Gisolfi, and A. Nappi, "Medical differential diagnosis through Type-2 Fuzzy Sets," *Proc. of FUZZ-IEEE*, pp.371-376, Reno, NV, USA, May 2005.

[11] N. N. Karnik, J. M. Mendel, and Q. Liang, "Type-2 fuzzy logic systems," *IEEE Trans. on Fuzzy Systems*, vol.7, no.6, pp.643-658, 1999.

[12] J. M. Mendel, "Uncertainty, fuzzy logic, and signal processing," *Signal Processing*, vol.80, no.6, pp. 913-933, 2000.

[13] L. Li, W.-H. Lin and H. Liu, "Type-2 fuzzy logic approach for short-term traffic forecasting," *IEE Proc. Intelligent Transportation System*, vol.153, no.1, pp.33-40, 2006.

[14] H. A. Hagras, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Trans. on Fuzzy Systems*, vol.12, no.4, pp.524-539, 2004.

[15] H. B. Mitchell, "Pattern recognition using type-II fuzzy sets," *Information Sciences*, vol.170, pp.409-418, 2005.

[16] J. Zeng and Z. Q. Liu, "Type-2 fuzzy hidden Markov models and their application to speech recognition," *IEEE Trans. on Fuzzy Systems*, vol. 14, no.3, pp. 454-467, June 2006.

[17] H. Wu and J. M. Mendel, "Classification of battlefield ground vehicles using acoustic features and fuzzy logic rule-based classifiers," *IEEE Trans. on Fuzzy Systems*, vol. 15, no.1, pp.56-72, Feb. 2007.

[18] O. Castillo, P. Melin, O. Montiel, A. Rodriguez-Diaz and R. Sepulveda, "Handling uncertainty in controllers using type-2 fuzzy logic," *Journal of Intelligent Systems*, vol. 14, no. 3, pp.237-262, August 2005.

[19] R. Sepulveda, O. Castillo, P. Melin, A. Rodriguez-Diaz and O. Montiel, "Experimental study of intelligent controllers under uncertainty using type-1 and type-2 fuzzy logic," *Information Sciences*, vol. 177, no. 10, pp.2023-2048, 2007.

[20] H. Hagras, F. Doctor, A. Lopez, and V. Callaghan, "An incremental adaptive life long learning approach for type-2 fuzzy embedded agents in ambient intelligent environments," *IEEE Transactions on Fuzzy Systems*, vol. 15, no.1, pp. 41-55, 2007.

[21] H. Hagras, "Comments on "Dynamical optimal training for interval type-2 fuzzy neural network (T2FNN)"," *IEEE Trans. on Systems, Man, and Cybernetics,* vol.36, no.5, pp.1206-1209, 2006.

[22] K. M. Bossley, D. J. Mills, M. Brown, and C. J. Harris, "Construction and design of parsimonious neurofuzzy systems," in K. J. Hunt, G. R. Irwin, and K. Warwick (Eds), *Proceedings of Neural Network Engineering in Control Systems*, pp.153-177, Springer-Verlag, 1995.

[23] M. J. Er, and S. Wu, "A fast learning algorithm for parsimonious fuzzy neural systems," *Fuzzy Sets and Systems*, vol.126, no.3, pp.337-351, 2002.

[24] H. Roubos nd M. Setnes, "Compact fuzzy models through complexity reduction and evolutionary optimization," *Proc. of FUZZ-IEEE*, San Antonio, USA, May, pp.762-767, 2000.

[25] S.-M. Zhou, and J. Q. Gan, "Constructing parsimonious fuzzy classifiers based on L2-SVM in high-dimensional space with automatic model selection and fuzzy rule ranking," *IEEE Trans. on Fuzzy Systems*, vol.15, no.3, pp.398 409, 2007.

[26] Q. Liang, and J. M. Mendel, "Designing interval type-2 fuzzy logic systems using an SVD-QR method: rule reduction," *International Journal of Intelligent Systems*, vol.15, pp.939-957, 2000.

[27] G. C. Mouzouris, and J. M. Mendel, "A singular-value-QR decomposition based method for training fuzzy logic systems in uncertain environments," *Journal of Intelligent and Fuzzy Systems*, vol. 5, pp.367-374, 1997.

[28] G. C. Mouzouris, and J. M. Mendel, "Designing fuzzy logic systems for uncertain environments using a singular-value-QR decomposition method," *Proc of FUZZ-IEEE*, New Orleans, LA, USA, pp.295-301, 1996.

[29] M. Setnes and R. Babuska, "Rule base reduction: some comments on the use of orthogonal transforms," *IEEE Trans. on Systems Man and Cybernetics-Part C*, Vol.31, No.2, pp.199-206, 2001.

[30] M. Setnes and H. Hellendoorn, "Orthogonal transforms for ordering and reduction of fuzzy rules," *Proc. of the 9^{th} IEEE int. Conference on Fuzzy Systems*, San Antonio, TX, USA, pp.700-705. 2000.

[31] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. on Neural Networks,* vol.2, no.2, pp.302–309, 1991.

[32] L. X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least squares learning," *IEEE Trans. on Neural Networks,* vol.3, no.5, pp.807–814, 1992.

[33] L. Wang and R. Langari, "Building Sugeno-type models using fuzzy discretization and orthogonal parameter estimation techniques," *IEEE Trans. on Fuzzy Systems,* vol.3, no.4, pp.454–458, 1995.

[34] P. A. Mastorocostas, J. B. Theocharis, and V. S. Petridis, "A constrained orthogonal least squares method for generating TSK fuzzy models: application to short-term load forecasting," *Fuzzy Sets and Systems,* vol.118, no.2, pp.215-233, 2001.

[35] S.-M. Zhou, and J. Q. Gan, "Constructing accurate and parsimonious fuzzy models with distinguishable fuzzy sets based on an entropy measure," *Fuzzy Sets and Systems*, vol. 157, no.8, pp. 1057-1074, 2006.

[36] S. Coupland, and R. I. John, "Geometric type-1 and type-2 fuzzy logic systems," *IEEE Transactions on Fuzzy Systems*, vol.15, no.1, pp.3-15, 2007.

[37] S. Coupland, and R. John, "A fast geometric method for defuzzification of type-2 fuzzy sets," *IEEE Trans. on Fuzzy Systems*(Accepted) 2008.

[38] M. Rimer, and T. Martinez, "Classification-based objective functions," *Machine Learning*, vol. 63, no. 2, pp.183-205, 2006.

[39] V. Vapnik, *The Nature of Statistical Learning*, New York: Springer Verlag, 1995.

[40] J. M. Mendel, "Computing derivatives in interval type-2 fuzzy logic systems," *IEEE Trans. on Fuzzy Systems*, vol.12, no.1, pp.84-98, 2004.

[41] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[42] M. J. Moody and C. J. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Computation*, vol.1, no.2, pp.281-294, 1989.

[43] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modelling," *IEEE Trans. on Fuzzy Systems*, vol.1, no.1, pp.7-31, 1993.

[44] K. Kilic, B. A. Sproulec, I. B. Turksen and C. A. Naranjo, "A fuzzy system modeling algorithm for data analysis and approximate reasoning," *Robotics and Autonomous Systems*, vol.49, no.3-4, pp.173-180, 2004.

[45] S. Bittanti and L. Piroddi, "Nonlinear identification and control of a heat exchanger: a neural network approach," *Journal of the Franklin Institute*, vol.334, no.1, pp.135-153, 1997.

Shang-Ming Zhou (M'01) received the BSc. degree in mathematics from Liaocheng University, Shandong Province, China, the MSc. degree in applied mathematics from Beijing Normal University, Beijing, China, and the PhD degree in computer science from the University of Essex, UK in 1985, 1992, and 2006 respectively. Currently he is with the Centre for Computational Intelligence, School of Computing, De Montfort University in the UK. His research interests include fuzzy logic systems (type-1 and type-2) and applications; decision support systems under uncertainty environments; model interpretability and transparency in data-driven neurofuzzy systems; kernel machine learning (inc support vector machines) with fuzzy computing; artificial neural networks and pattern recognition; intelligent signal and image processing. He has published extensively on these topics.

Jonathan M. Garibaldi received the B.Sc (Hons) degree in Physics from Bristol University, U.K., and the M.Sc. degree in Intelligent Systems and the Ph.D. degree in uncertainty handling in immediate neonatal assessment from the University of Plymouth, U.K., in 1984, 1990, and 1997, respectively.

He is currently an Associate Professor and Reader within the Intelligent Modelling and Analysis (IMA) Research Group in the School of Computer Science at the University of Nottingham, U.K. The IMA research group undertakes research into intelligent modelling, utilising data analysis and transformation techniques to enable deeper and clearer understanding of complex problems. Dr Garibaldi has published over 40 papers on fuzzy expert systems and fuzzy modelling, including three book chapters, and has edited two books. His main research interests are modelling uncertainty in human reasoning and especially in modelling the variation in normal human decision making, particularly in medical domains. He has created and implemented fuzzy expert systems, and developed methods for fuzzy model optimisation.

Robert I. John received the B.Sc. (Hons.) degree in mathematics from Leicester Polytechnic, Leicester, U.K., the M.Sc. degree in statistics from UMIST, Manchester, U.K., and the Ph.D. degree in type-2 fuzzy logic De Montfort University, Leicester, U.K., in 1979, 1981, and 2000, respectively.

Currently, he is a Professor and the Director of the Centre for Computational Intelligence at De Montfort University. He is a Member of the Editorial Board of International Journal of Cognitive Neurodynamics, International Journal of Computational Intelligence, International Journal for Computational Intelligence and Information and Systems Sciences. Professor John is a Vice-chairman of the Fuzzy Technical Committee (FTC) of the IEEE Neural Networks, and a Member of EPSRC College in the UK. He was a co-general chair of the IEEE International Conference on Fuzzy Systems in London in 24-26 July, 2007. He has published extensively in the area of type-2 fuzzy logic and his research interests are in the general field of modelling human decision making using type-2 fuzzy logic.

Francisco Chiclana received the B.Sc. and Ph.D degrees in Mathematics, both from the University of Granada, Spain, in 1989 and 2000, respectively.

In August 2003, he joined the Centre for Computational Intelligence (CCI) Research Group at De Montfort University, Leicester, UK. He is currently a Reader in Computational Intelligence. His research interests include fuzzy logic based systems, fuzzy preference modelling, decision making problems with heterogeneous fuzzy information, decision support systems, the consensus reaching process, aggregation and consistency of information, and modelling situations with missing information.