

Fuzzy Modelling Using a New Compact Fuzzy System: A Special Application to the Prediction of the Mechanical Properties of Alloy Steels

Qian Zhang and Mahdi Mahfouf

Department of Automatic Control and Systems Engineering
The University of Sheffield
Sheffield, UK
qian.zhang@sheffield.ac.uk, m.mahfouf@sheffield.ac.uk

Abstract—In high-dimensional modelling cases, a fuzzy modelling approach based on the grid-partitioning of fuzzy sets always meets great challenges, as it cannot avoid the problem of introducing a huge number of fuzzy rules. To tackle this issue, a new grid-partitioning based fuzzy modelling paradigm is proposed in this paper to construct a compact fuzzy system by including ‘short fuzzy rules’, in which only a few but strategic premises are used. In the proposed approach, the generation of fuzzy rules is data-orientated, a consideration which can greatly reduce the computational complexity. A new framework for fuzzy reasoning and defuzzification is also devised, which employs some archived reference data to help choose the most suitable fuzzy rules. In material engineering, describing the behaviour of mechanical properties of alloys is often a high-dimensional modelling problem, which involves the complexity of materials’ chemical composites and their underlying physical processing mechanisms. In this paper, the proposed approach was successfully applied to generate models of ultimate tensile strength of alloy steel. Compared with the standard grid-partitioning based fuzzy modelling paradigms, the new method shows an improvement in both complexity and interpretability. Compared with the clustering-based fuzzy modelling approaches, the proposed method can achieve the same accuracy level and is more transparent.

Keywords—fuzzy; high-dimensional modelling; grid partitioning; interpretability; short fuzzy rule; alloy steel; mechanical property

I. INTRODUCTION

There are many complex real-life modelling problems for which precise physical models do not currently exist, while abundant data that represent the input-output relationships can be collected. In this situation, some forms of intelligent systems associated with data-driven techniques, such as fuzzy rule-based systems and artificial neural networks (ANNs), appear to be suitable for model construction.

The fundamental concepts related to fuzzy rule-based systems were first introduced by Zadeh [1], [2]. They are known as robust universal approximators and allow a system to be represented using an explicit knowledge representation (linguistic ‘IF-THEN’ rules), which can easily be understood and explained by humans to allow them to gain a deeper insight

into uncertain, complex, and ill-defined systems [3]. Compared with artificial neural networks, fuzzy systems are more transparent to humans to interpret. This is the most attractive property that distinguishes fuzzy systems from other black-box modelling approaches.

Normally, data-driven fuzzy modelling can be realised via two different routes: the grid-partitioning based method or the data-clustering based method. In the first method, the grid-partitioning initially defines a number of well-distributed fuzzy sets for each variable. The corresponding fuzzy rules are then elicited based on the fixed partitioning of fuzzy sets. In this approach, the fuzzy sets are also shared by all the fuzzy rules. In contrast, the second method employs data clustering (grouping) [4] information to define both the fuzzy sets and the fuzzy rules simultaneously. In this method, one cluster is associated to one fuzzy rule; the fuzzy sets are not shared by all the rules, but each set is only mapped into one particular fuzzy rule.

Compared with the grid-partitioning based method, the data-clustering based method needs less fuzzy rules to achieve a certain accuracy level; nevertheless its fuzzy sets are not evenly and well distributed and are not consistent with human knowledge and understanding. The grid-partitioning based method produces more interpretable systems, but it cannot avoid the problem of introducing a huge number of fuzzy rules for high-dimensional modelling problems. Another problem lies in high-dimensional modelling cases where a fuzzy rule may include too many antecedent conditions, a fact that defeats the interpretability property.

In this paper, a modelling paradigm using ‘short fuzzy rules’ is proposed to solve the above problems, where only a few but strategic antecedent conditions (premises) are included in a ‘short fuzzy rule’. This remarkably decreases the number of fuzzy rules, as one short fuzzy rule possesses more powerful generalisation ability and can cover a wider input space. On the other hand, the short expression of fuzzy rules is much easier to be understood by a human. Table 1 illustrates the comparison among different fuzzy modelling paradigms.

TABLE I. COMPARISON OF THE PROPOSED APPROACH OF ‘SHORT FUZZY RULES’ WITH OTHER FUZZY MODELLING PARADIGMS

		Clustering based methods	Grid-partitioning based methods	The proposed method
Interpretability	Number of fuzzy rules	Good	Bad	Medium
	The length of fuzzy rules	Bad	Bad	Good
	The number of fuzzy sets	Medium	Good	Good
	The distribution of fuzzy sets	Bad	Good	Good
	General performance	Medium	Medium	Good
Accuracy		Good	Medium	Medium

The remaining parts of this paper are organised as follows. Section II introduces the fuzzy systems with short fuzzy rules. Section III describes the proposed method to generate the compact fuzzy system. In Sections IV and V, a modified fuzzy reasoning and defuzzification mechanisms and a new framework for the proposed fuzzy system are proposed. Section VI presents an experimental study relating to the modelling of the mechanical property of the alloy steel. Finally, concluding remarks are drawn in Section VII.

II. FUZZY SYSTEMS WITH SHORT FUZZY RULES

Generally, a fuzzy rule-based system consists of four fundamental components: a fuzzy rule-base, a fuzzy inference engine, a fuzzifier, and a defuzzifier. Fig. 1 shows the diagram of a fuzzy system. The central part of a fuzzy system is the knowledge-base (rule-base) consisting of fuzzy rules. A fuzzy rule is an IF-THEN statement in which some words are characterised by continuous membership functions. All other components are used to implement these rules in a reasonable and in an efficient manner. The fuzzifier represents a mapping from a real-valued point $x \in U \subset R^i$ to a fuzzy set A in U , which is defined by a membership function $\mu_A(x): A \rightarrow [0, 1]$. In a fuzzy inference engine, fuzzy logic principles direct how to employ the fuzzy rules into a mapping from an input fuzzy set A to an output fuzzy set B . The defuzzifier is a mapping from the output fuzzy set B in $V \subset R$ to a real-valued point $y \in V$. Conceptually, the purpose of the defuzzifier is to specify a point in V that best represents B [5].

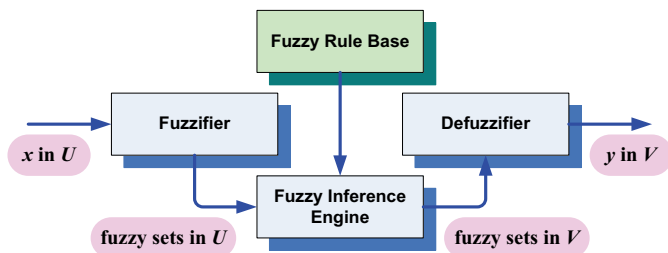


Figure 1. The basic configuration of fuzzy systems.

For an i -input 1-output system, a fuzzy rule-base could comprise the following Mamdani fuzzy rules:

Rule_j: IF x_1 is A_{1j} AND x_2 is A_{2j} AND ... AND x_i is A_{ij} , THEN y is B_j .

where $x = (x_1, x_2, \dots, x_i)^T \in U$ and $y \in V$ are the input and output (linguistic) variables of the fuzzy system; A_{ij} and B_j are fuzzy sets in $U_i \subset R$ and $V \subset R$, respectively.

In some literatures [6], ‘don’t care’ is employed as an additional antecedent fuzzy set, which is fully compatible with any input value. Its membership function is defined as follows:

$$\forall x: \mu_{\text{don't care}}(x) = 1. \quad (1)$$

Since ‘don’t care’ conditions are usually omitted from the antecedent part, fuzzy rules with many ‘don’t care’ conditions are short and interpretable. The following shows an example of a fuzzy rule with ‘don’t care’ conditions:

IF x_1 is *don't care* AND x_2 is *medium* AND x_3 is *don't care*, THEN y is *large*.

If two ‘don’t care’ conditions are omitted, the above fuzzy rule can be shortened as follows:

IF x_2 is *medium*, THEN y is *large*.

In this paper, the number of antecedent conditions (premises) of each fuzzy rule is referred to as the rule length. In the above case, the rule length is reduced from 3 to 1 by using the ‘don’t care’ label.

III. GENERATION OF THE ‘COMPACT FUZZY SYSTEM’

Generally, when constructing a fuzzy system, the fuzzy inference engine and the defuzzifiers are predefined. In the grid-partitioning based method, the uniformly-distributed fuzzy sets are also predefined for each variable. Thus, the primary work becomes to generating appropriate fuzzy rule-base.

In this work, the following steps are executed to elicit short fuzzy rules for the ‘compact fuzzy system’:

A. Find candidate short fuzzy rules from all possible fuzzy rules

The possible fuzzy rules include every combination of linguistic terms, although they are too many to be comprehensively evaluated. Suppose an i -input 1-output system has k linguistic terms (excluding ‘don’t care’) for every input variable and l linguistic terms for the output variable, the number of possible fuzzy rules could be $(k+1)^i * l$. The number of possible short fuzzy rules with rule length being 1 is $i * k * l$ and the number of possible short fuzzy rules with rule length being 2 is $\frac{i * (i-1)}{2} * k^2 * l$. In this paper, only the fuzzy rules no longer than 2 are considered.

As stated above, for a high-dimensional modelling problem, the possible short fuzzy rules may even be too many to be evaluated. In order to further decrease the number of candidate rules, it is required that only the rule that covers at least one training data point is considered as a candidate fuzzy rule.

Let C be a set of m training data c_n , where $c_n = [x_n, y_n]^T = [x_{1n}, x_{2n}, \dots, x_{in}, y_n]^T$ and $n = 1, 2, \dots, m$. $Rule_j$ is regarded as a candidate short fuzzy rule, if

$$\exists n: \min(\mu_{A_{1j}}(x_{1n}), \mu_{A_{2j}}(x_{2n}), \dots, \mu_{A_{ij}}(x_{in}), \mu_{B_j}(y_n)) > Th_1, \quad (2)$$

where ' Th_1 ' is a threshold value and it is set to 0.5 in this work without any loss of generality.

B. Calculate the 'confidence' value for each candidate short fuzzy rule

To estimate whether a fuzzy rule is redundant or not, an evaluation measure 'confidence' is applied [7], [8]. In the area of data mining, it has been used for evaluating association rules in the pattern classification problem. The *confidence* of rule $A \rightarrow B$ is defined as follows:

$$conf(A \rightarrow B) = \frac{|C(A) \cap C(B)|}{|C(A)|} \quad (3)$$

where $|C(A)|$ is the number of training patterns that are compatible with the antecedent A , and $|C(A) \cap C(B)|$ is the number of training patterns that are compatible with both the antecedent A and the consequent B . This *confidence* measure indicates that $conf \times 100\%$ of the training patterns that are compatible with the antecedent A are also compatible with the consequent B . In [7, 8], the *confidence* measure is extended to the fuzzy case and can be rewritten as follows:

$$conf(A \rightarrow B) = \frac{|C(A) \cap C(B)|}{|C(A)|} = \frac{\sum_{n=1}^m (N_A(x_n) \times \mu_B(y_n))}{\sum_{n=1}^m N_A(x_n)} \quad (4)$$

where $N_A(x_n)$ is the compatibility grade of the input vector x_n with the antecedent part $A = [A_1, A_2, \dots, A_i]^T$ of the fuzzy rule, and $\mu_B(y_n)$ is the compatibility grade of the output value y_n with the consequent part B of the rule. $N_A(x_n)$ is usually defined by the minimum operator or the product operator. Such as:

$$N_A(x_n) = \min(\mu_{A_{1j}}(x_{1n}), \mu_{A_{2j}}(x_{2n}), \dots, \mu_{A_{ij}}(x_{in})) \text{ or} \quad (5)$$

$$N_A(x_n) = \mu_{A_{1j}}(x_{1n}) \times \mu_{A_{2j}}(x_{2n}) \times \dots \times \mu_{A_{ij}}(x_{in}). \quad (6)$$

In this work, these *confidence* values are also saved as the rule weights and will be further used in the defuzzification stage, as will be seen in Section IV.

C. Determine the final short fuzzy rule-base

A subset of candidate short fuzzy rules is selected as the active rule-base by taking into account the obtained *confidence* values. It is designed as the minimum subset of candidate rules which have the largest *confidence* values and can also cover all the training data. Specifically, if RS_p represent the selected rules and RU_q represent those unselected rules,

$$\min_p conf(RS_p) \geq \max_q conf(RU_q), \quad (7)$$

$$\forall n, \exists p: \min(\mu_{A_{1p}}(x_{1n}), \mu_{A_{2p}}(x_{2n}), \dots, \mu_{A_{ip}}(x_{in}), \mu_{B_p}(y_n)) > Th_2 \quad (8)$$

where $conf(RS_p)$ is the *confidence* value of the p th short fuzzy rule RS_p ; Th_2 is a threshold and a value 0.5 is used in this work without any loss of generality.

IV. FUZZY REASONING AND DEFUZZIFICATION

Because of the different rule structures of the compact fuzzy system and standard fuzzy systems, it is not appropriate to implement the traditional fuzzy reasoning and defuzzification methods directly. For a compact fuzzy system with short fuzzy rules, the following two issues should be addressed. First, a shorter rule describes a more general circumstance and a longer rule reflects a more specific situation. In a system including various rule lengths, longer rules should have higher priority to take effect. Second, among the rules with the same rule length, each makes a different level of impact on the system. This can be observed from a variety of *confidence* values for different fuzzy rules. In this work, both of the above issues are taken into account in the proposed fuzzy reasoning and defuzzification mechanism.

Suppose the product inference engine and the centre average defuzzifier [5] are employed, the system output $y(x)$ is calculated as follows:

$$y(x) = \frac{\sum_p N_{A_p}(x) \times b_p}{\sum_p N_{A_p}(x)} \quad (9)$$

where $N_{A_p}(x)$ is the compatibility grade of the input vector x with the antecedent part $A_p = [A_{1p}, A_{2p}, \dots, A_{ip}]^T$ of the fuzzy rule RS_p ; b_p is a representative real number of the consequent fuzzy set B_p (the centre of its membership function). $N_{A_p}(x)$ is usually defined by the minimum operator or the product operator. Such as:

$$N_{A_p}(x) = \min(\mu_{A_{1p}}(x_1), \mu_{A_{2p}}(x_2), \dots, \mu_{A_{ip}}(x_i)) \text{ or} \quad (10)$$

$$N_{A_p}(x) = \mu_{A_{1p}}(x_1) \times \mu_{A_{2p}}(x_2) \times \dots \times \mu_{A_{ip}}(x_i). \quad (11)$$

By considering both the rule length and the rule weight, the above equation can be amended as follows:

$$y(x) = \frac{\sum_p (l_{RS_p})^{conf_{RS_p}} \times N_{A_p}(x) \times b_p}{\sum_p (l_{RS_p})^{conf_{RS_p}} \times N_{A_p}(x)} \quad (12)$$

where l_{RS_p} is the rule length of the p th short fuzzy rule RS_p , and $conf_{RS_p}$ is the rule weight (*confidence* value) of RS_p .

V. A NEW FRAMEWORK FOR THE COMPACT FUZZY SYSTEM

After some initial experimental tests, it was found that the above compact fuzzy system always tends to yield a moderate prediction. The reason behind this may lie in the existing parallel fuzzy rules, which have similar or the same antecedent conditions but dissimilar consequent parts. To distinguish which rules should be applied in a certain situation, a new framework for the compact fuzzy system is proposed, as shown in Fig. 2.

The modified system will not use all the fuzzy rules including in the rule-based in the fuzzy inference engine. Only a set of relevant fuzzy rules, which is carefully selected by comparing their compatibility with the relevant reference data, will be employed into the fuzzy reasoning.

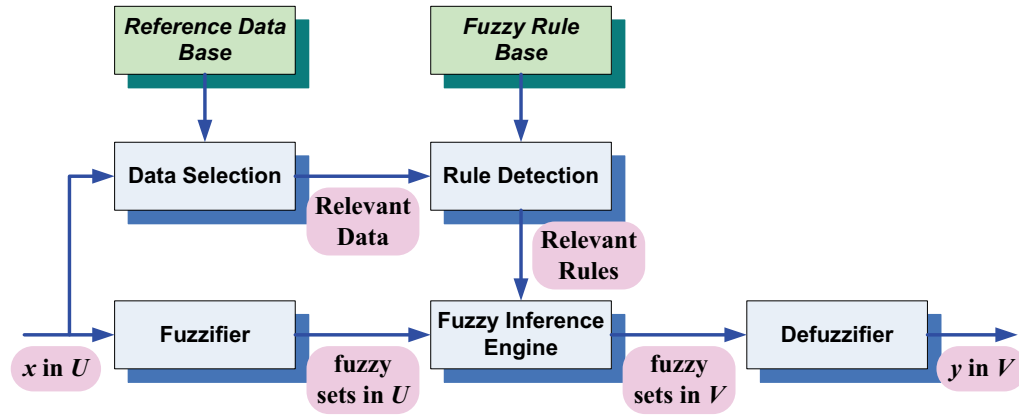


Figure 2. The diagram of the proposed fuzzy system.

Specifically, when a set of inputs $\mathbf{x} = [x_1, x_2, \dots, x_i]^T$ goes into the system, it will be compared with the archived reference data set $\mathbf{C}_R = [\mathbf{c}_{R1}, \mathbf{c}_{R2}, \dots, \mathbf{c}_{Rr}]$, where $\mathbf{c}_{Rr} = [x_{1Rr}, x_{2Rr}, \dots, x_{iRr}, y_{Rr}]^T$ and r is the number of all the reference data, and the most similar data $\mathbf{C}_S = [\mathbf{c}_{S1}, \mathbf{c}_{S2}, \dots, \mathbf{c}_{Ss}]$ will be selected, where $\mathbf{c}_{Ss} = [x_{1Ss}, x_{2Ss}, \dots, x_{iSs}, y_{Ss}]^T$ and s is the number of the selected reference data. To estimate the similarity between $\mathbf{x}$ and a reference data, the Euclidean distance in the input space is calculated. If it is less than a threshold ' Th_3 ', the reference data are regarded as close to $\mathbf{x}$ and they will be selected to be used in the further fuzzy rule detection. This can be described as follows:

$$\mathbf{c}_R \subset \mathbf{C}_S, \text{ if } \sqrt{\sum_{t=1}^i (x_t - x_{tR})^2} < Th_3. \quad (13)$$

In general, ' Th_3 ' is recommended to be set as follows, without any loss of generality:

$$Th_3 = 0.15 \times \sqrt{\sum_{t=1}^i (x_{t\max} - x_{t\min})^2} \quad (14)$$

where $x_{t\max}$ and $x_{t\min}$ represent the maximum value and the minimum value of all the reference data in their t th dimension, respectively.

In the subsequent rule detection, the obtained relevant reference data are employed to identify the most relevant fuzzy rules $\mathbf{R}_S$. If a fuzzy rule R is compatible with at least one reference data to some extent, the fuzzy rule is reckoned as relevant and will be passed through fuzzy reasoning phase. This can be described as follows:

$$R \subset \mathbf{R}_S, \text{ if } \exists n \in [1, 2, \dots, s]: \min(\mu_{A_1}(x_{1n}), \mu_{A_2}(x_{2n}), \dots, \mu_{A_i}(x_{in}), \mu_B(y_n)) > Th_4, \quad (15)$$

where ' Th_4 ' is a threshold value and it is set to be 0.5 in this work without any loss of generality.

VI. PREDICTION OF MECHANICAL PROPERTIES

In material engineering, it is crucial to establish transparent and reliable mechanical property prediction models for materials design and development [9]-[11]. But it may be 'tricky' to precisely describe the behaviour of mechanical properties using mathematical models alone due to the complexity of materials' chemical composites and their underlying physical processing mechanisms, such as hot rolling

and heat treatment. Thus, developing a fast, efficient and transparent data-driven modelling framework for material property prediction is still a priority.

In the past, some prediction models relating to mechanical properties have been developed, which were mainly based on linear regression methods, artificial neural networks [12], and fuzzy systems [13]-[15]. Such linear models were only designed for specific classes of steels and specific processing routes, and are not sophisticated enough to account for more complex interactions, while neural networks are black-box techniques and the knowledge behind this kind of models cannot be understood fully. The fuzzy modelling techniques included in [13]-[15] employed data clustering information, which lead to a fuzzy rule-base which was not interpretable enough.

Particularly in this work, one of the studied mechanical properties of heat-treated alloy steels, is the Ultimate Tensile Strength (UTS), which represents a measure of the maximum load that a material can withstand. This problem is a high-dimension modelling problem and includes fifteen (15) input variables, which include the weight percentages for various chemical composites as well as the processing parameters of the heat treatment, consisting of the weight percentages for the chemical composites, namely carbon (C), silica (Si), manganese (Mn), sulphur (S), chromium (Cr), molybdenum (Mo), nickel (Ni), aluminium (Al) and vanadium (V), the test depth, the bar size, the treatment site, the quenching medium, as well as the hardening and tempering temperatures. Moreover, this problem is associated with a large amount of industrial data (3760 UTS data), which are not well-distributed and may be quantitatively redundant for training purposes. All the data are provided by Tata Steel Europe. In the following experiment, 60% of the data were used for training, and the remaining 40% were used for final testing. This experiment was repeated 20 times. Only one set of typical results out of the 20 runs is chosen and shown in the following.

For configurations, the number of fuzzy sets for each input variable was set to be 7 and the number of fuzzy sets for the output variable was set to be 15. In such a grid partitioning of input space, the number of all the possible fuzzy rules is 5.28×10^{14} , the number of possible short fuzzy rules with rule length being 1 is 1575, and the number of possible short fuzzy

rules with rule length being 2 is 77175. By employing the proposed approach, only 214 ‘short fuzzy rules’ were elicited.

Fig. 3 shows the prediction performance (predicted outputs against measured outputs) of the compact fuzzy model using standard fuzzy reasoning and defuzzification. It can be seen that the predictions tend to be of moderate values. Fig. 4 shows the prediction performance of the same model with the new framework introduced in Section IV and the modified reasoning and defuzzification methods proposed in Section V. For this case, the Root Mean Square Error (RMSE) based on the training data and the testing data are 40.05 MPa and 47.92 MPa, respectively. This accuracy is as good as those reported in [14], [15], which established fuzzy models via data-clustering based methods with rules including many antecedents unnecessarily.

For more details about the developed UTS model, Fig. 5 compares the fuzzy rules between a full-length one generated using data clustering techniques [14], [15] and the short ones obtained in this work. It can be seen that the full-length fuzzy rule is too complicated to be interpreted. One can also observe another disadvantage of the fuzzy system elicited from data clustering that the fuzzy sets are not evenly and well distributed. For these fuzzy rules, they can be rewritten as the following approximate linguistic rules using the linguistic hedges approach [16]:

R_1 : IF Test Depth is *small* AND Size is *small medium* AND Site Number is *large* AND C is *large medium* AND Si is *medium* AND Mn is *medium small* AND S is *small* AND Cr is *medium small* AND Mo is *medium* AND Ni is *large medium* AND Al is *very small* AND V is *very small* AND Hardening Temperature is *small* AND Cooling Medium Number is *more or less large* AND Tempering Temperature is *medium large*, THEN UTS is *large medium*.

R_2 : IF Mn is *large medium*, THEN UTS is *small medium*.

R_3 : IF V is *small medium*, THEN UTS is *medium small*.

R_4 : IF Ni is *medium* AND Hardening Temperature is *small*, THEN UTS is *small medium*.

It is clear that the linguistic fuzzy rules as R_2 , R_3 , and R_4 , allow for a better insight into the heat-treated alloy steels process.

To verify the physical interpretation of the obtained models, Fig. 6 shows the three-dimensional response surfaces of the compact model. These surfaces are achieved by plotting two varying input variables against the output while keeping other input variables constant. The constant variables are set to the average values of the dominant steel grade, which is the 1%CrMo steel grade [17]. These plots in Fig. 6 are consistent with those variable effect plots in [17], which have been verified to follow the expected behaviour as predicted by theory and knowledge experts.

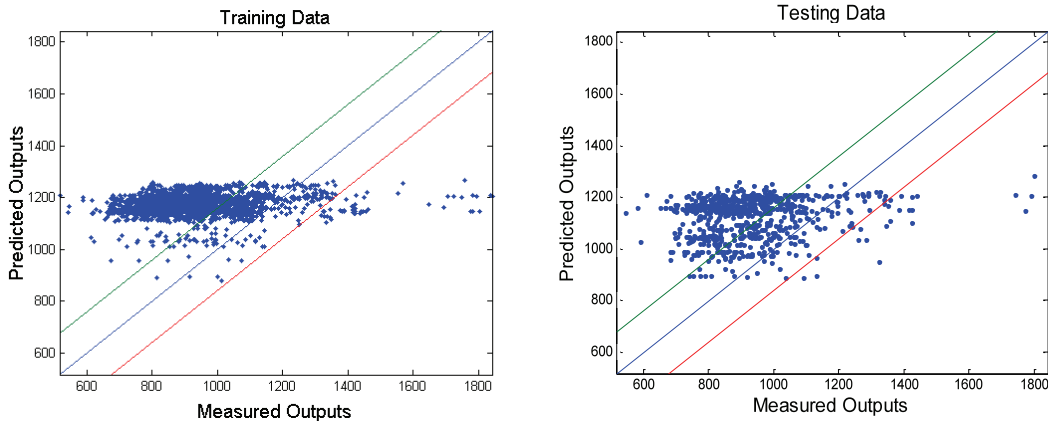


Figure 3. The UTS model’s predicted outputs versus the measured outputs with the standard fuzzy reasoning and defuzzification methods.

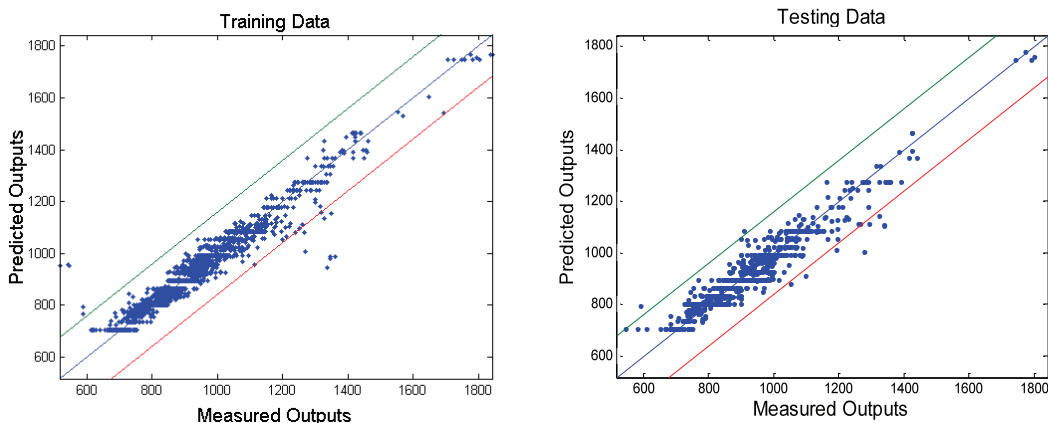


Figure 4. The UTS model’s predicted outputs versus the measured outputs with the proposed system structure.

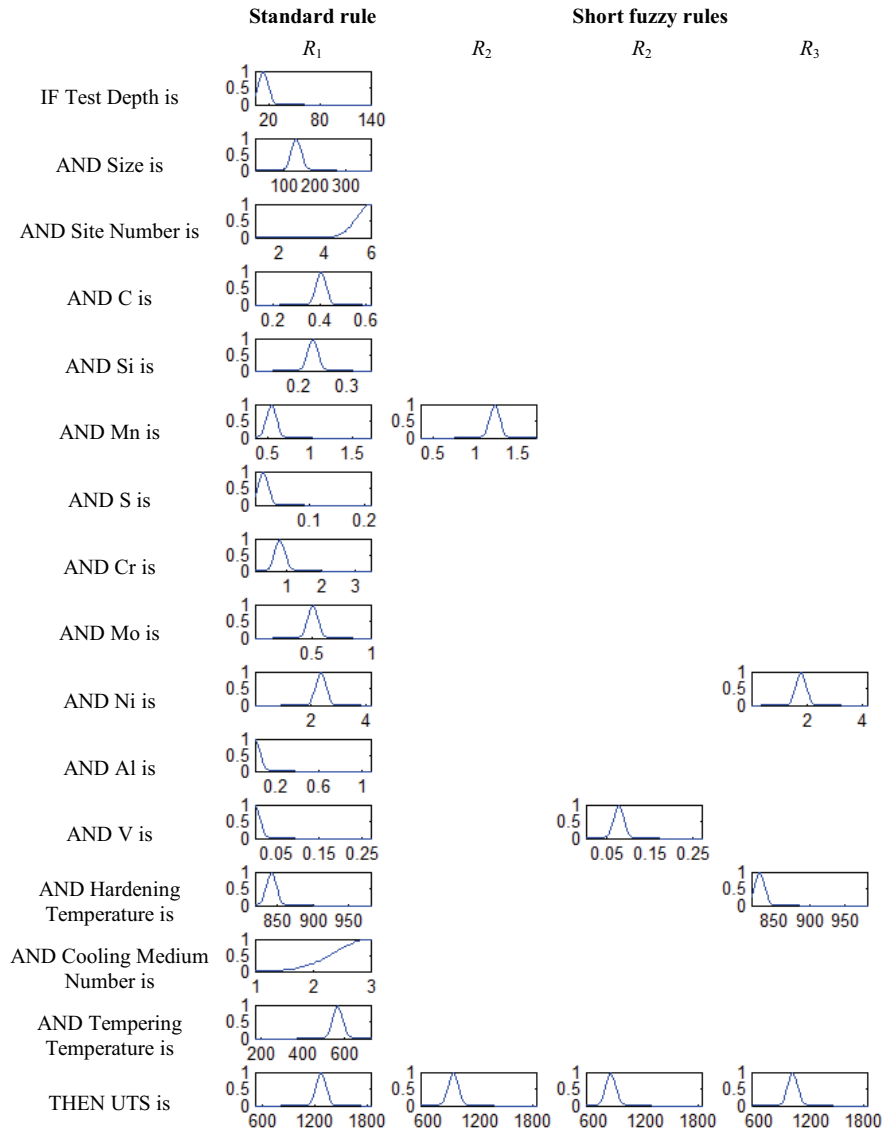


Figure 5. A comparison between a full-length and short fuzzy rules.

VII. CONCLUSION

For modelling problems, the most attractive advantage of fuzzy rule-based systems lies in their interpretability, which means that users should be able to understand a fuzzy system's behaviour by inspecting its associated rule-base. Normally, the interpretability issue involves the following widely accepted aspects: First, the number of fuzzy rules should be small enough to be comprehensible. Second, the rule premises should be simple in structure and contain only a few features (input variables). Third, the linguistic terms should be intuitively comprehensible. The form and parameters of the membership functions should correspond to the understanding of the linguistic expressions.

The grid-partitioning based fuzzy modelling approaches are considered to be more interpretable than data-clustering based ones since they include a set of well-distribution fuzzy sets. In the traditional grid-partitioning based methods, all the possible fuzzy rules in every partition are evaluated, which is time-

consuming and cannot be achieved in high-dimensional modelling cases. In this work, the generation of fuzzy rules is data-orientated using 'short fuzzy rules', which can greatly reduce the number of fuzzy rules as well as the computational complexity. Among the fuzzy rules with various rule length, the shorter ones represent more general status and the longer ones deal with more specific situation. This has been fully taken into account in the reasoning and defuzzification. To identify the most appropriate fuzzy rules in a certain circumstance, a new mechanism of reference rules detection has also been developed.

The proposed approach was then successfully applied to generate models of the mechanical properties (in our case one) of alloy steels. Compared with the standard grid-partitioning based fuzzy modelling paradigms, the new method reveals an improvement in both the computational complexity and the interpretability. Compared with the clustering-based fuzzy modelling approaches, the proposed method can achieve the same accuracy level with a more transparent rule-base.

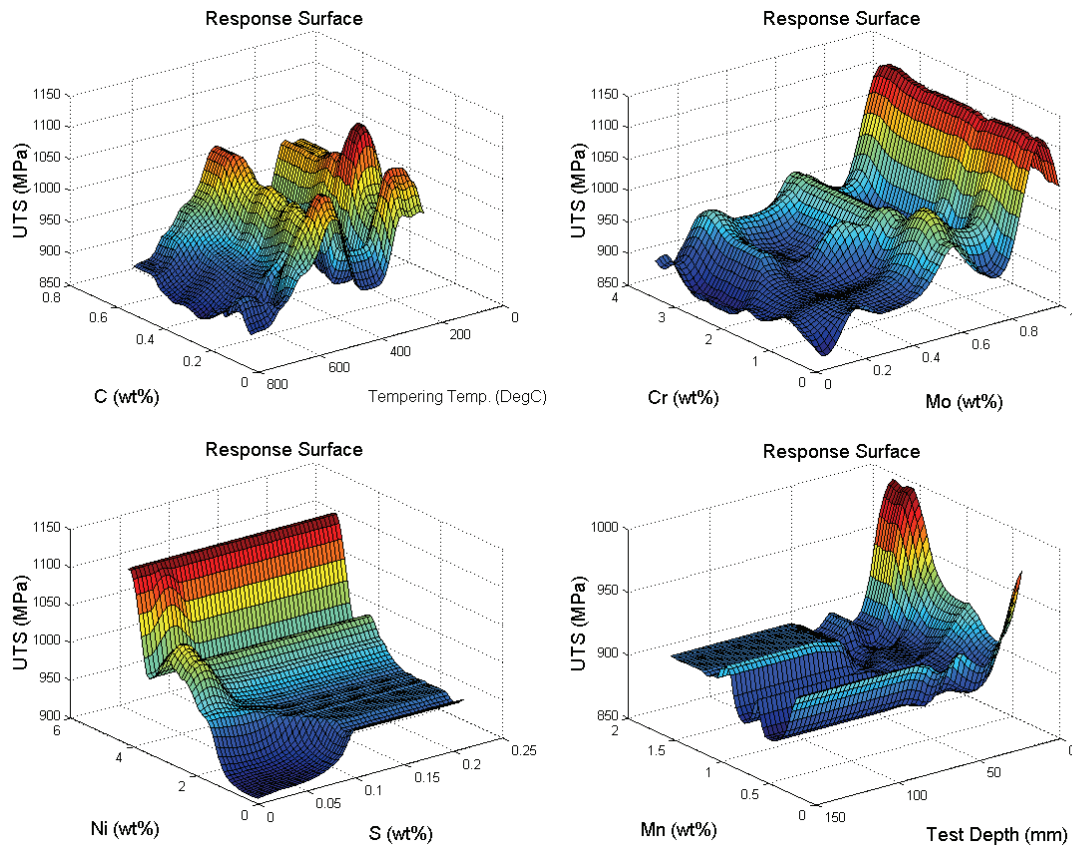


Figure 6. Response surfaces of the obtained UTS model.

In this paper, only the Mamdani-type fuzzy systems [18], which use linguistic expressions in both the antecedent and consequent parts, are considered. In the near future, the design could be extended to TSK-type fuzzy systems [19], which would replace the linguistic consequent parts of the Mamdani-type fuzzy system with deterministic forms. The authors will, also explore the possibility of extending the idea of employing short fuzzy rules within the framework of data-clustering based fuzzy modelling. It is also worth noting that the proposed framework does not include any optimisation strategy. In future, learning via effective optimisation could be embedded into the system to further improve overall accuracy while maintaining the same level of interpretability.

REFERENCES

- [1] L.A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [2] L.A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 28–44, 1973.
- [3] Q. Zhang, *Nature-Inspired Multi-Objective Optimisation and Transparent Knowledge Discovery via Hierarchical Fuzzy Modelling*. PhD Thesis, The Department of Automatic Control and Systems Engineering, The University of Sheffield, UK, 2008.
- [4] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping Multidimensional Data: Recent Advances in Clustering*, J. Kogan, C. Nicholas, and M. Teboulle, Eds. Berlin: Heidelberg: Springer, 2006, pp. 25–71.
- [5] L.-X. Wang, *A course in fuzzy systems and control*. Englewood Cliffs, NJ: Prentice-Hall, 1997.
- [6] H. Ishibuchi and T. Yamamoto, "Interpretability issues in fuzzy genetics-based machine learning for linguistic modelling," in *Modelling with Words*, J. Lawry, J. Shanahan, and A. Ralescu, Eds. Berlin: Heidelberg: Springer, 2003, pp. 209–228.
- [7] T.-P. Hong, C.-S. Kuo, and S.-C. Chi, "Trade-off between computation time and number of rules for fuzzy mining from quantitative data," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, pp. 587–604, 2001.
- [8] H. Ishibuchi, T. Yamamoto, and T. Nakashima, "Fuzzy data mining: effect of fuzzy discretization," in *Proceedings of the First IEEE International Conference on Data Mining*, pp. 241–248, 2001.
- [9] Q. Zhang and M. Mahfouf, "A Modified PSO with a Dynamically Varying Population and Its Application to the Multi-Objective Optimal Design of Alloy Steels," in *Proceedings of the 2009 IEEE Congress on Evolutionary Computation*, pp. 3241–3248, 2009.
- [10] Q. Zhang and M. Mahfouf, "A nature-inspired multi-objective optimisation strategy based on a new reduced space searching algorithm - special application to the design of alloy steels," *Engineering applications of artificial intelligence*, vol. 23, no. 5, pp. 660–675, 2010.
- [11] Q. Zhang, M. Mahfouf, J.R. Yates, C. Pinna, G. Panoutsos, S. Boumaiza, R.J. Greene, and L. de Leon, "Modeling and Optimal Design of Machining-Induced Residual Stresses in Aluminium Alloys Using a Fast Hierarchical Multiobjective Optimization Algorithm," *Materials and Manufacturing Processes*, vol. 26, no. 3, pp. 508–520, 2011.
- [12] C.L.P. Chen, C. Yang, and S.R. LeClair, "Materials structure-property prediction using a self-architecting neural network," *Journal of Alloys Compounds*, vol. 279, no. 1, pp. 30–38, 1998.
- [13] M.-Y. Chen and D.A. Linkens, "A systematic neuro-fuzzy modeling framework with application to material property prediction," *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, vol. 31, no. 5, pp. 781–790, 2001.
- [14] Q. Zhang, and M. Mahfouf, "Mamdani-type fuzzy modelling via hierarchical clustering and multi-objective particle swarm optimisation

- (FM-HCPSO)," International Journal of Computational Intelligence Research, vol. 4, no. 4, pp. 314–328, 2008.
- [15] Q. Zhang, and M. Mahfouf, "A hierarchical Mamdani-type fuzzy modelling approach with new training data selection and multi-objective optimisation mechanisms: A special application for the prediction of mechanical properties of alloy steels," Applied Soft Computing, vol. 11, no. 2, pp. 2419–2443, 2011.
 - [16] L.A. Zadeh, "A fuzzy-set-theoretic interpretation of linguistic hedges," Journal of Cybernetics, vol. 2, pp. 4–34, 1972.
 - [17] J. Tenner, Optimisation of the Heat Treatment of Steel Using Neural Networks, PhD Thesis, The Department of Automatic Control and Systems Engineering, The University of Sheffield, UK, 1999.
 - [18] E.H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," International Journal Man-Machine Studies, vol. 7, pp. 1–13, 1975.
 - [19] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," IEEE Transaction Systems, Man and Cybernetics, vol. 15, no. 1, pp. 116–132, 1985.