CONFINEMENT REGIME IDENTIFICATION IN NUCLEAR FUSION VIA AN INTERPRETABLE FUZZY LOGIC CLASSIFIER

Guido Vagliasindi, Paolo Arena
Dipartimento di Ingegneria Elettrica Elettronica e dei Sistemi
Università degli Studi di Catania
Italy
gvaglia@diees.unict.it

Andrea Murari
Consorzio RFX
Associazione EURATOM ENEA per la Fusione
Italy
murari@igi.pd.cnr.it

and JET-EFDA Contributors[1]

JET-EFDA, Culham Science Centre, OX14 3DB, Abingdon, UK

Abstract
In this paper a data driven methodology to automatically derive an interpretable Fuzzy Logic Classifier (FLC) has been applied to the problem of confinement regime identification in the Joint European Torus. The approach has been developed explicitly to handle with the complexities of the inference process in Magnetic Confinement Nuclear Fusion (MCNF). The first step of the method consists of a supervised, exploratory analysis performed with the approach of Classification and Regression Trees (CART), to extract the variables in the database which are the most critical for the problem under study. Then, a fully automated algorithm determines the membership functions and the most appropriate rules to reproduce the classification tree obtained with CART. The resulting FLI on the one hand attains very good performance in terms of generalization and classification, on the other hand provides a series of rules which can be easily interpreted and contributing to a very good first, intuitive understanding of the physics involved.

Key words
Tokamak, Confinement Regime, Knowledge from Data

1 Introduction

In magnetically confinement nuclear fusion, the data analysis process presents some unique challenges which have no direct counterpart in any other field of ”Big Physics” research. Indeed, whereas disciplines like astrophysics or high energy physics are faced with the task of extracting a small number of interesting events from an overwhelming sea of irrelevant background, the measurements performed on high temperature plasma are all relevant and there is no reason to discard any of them. On the contrary, in principle, a consistent model of a plasma discharge should be able to make sense of all the acquired signals by inserting them in a coherent picture. Till nowadays, no general model of the plasma evolution is available, due to the complexity and non linear character of the phenomena involved. As a consequence data analysis constitutes a precious tool to infer measurement driven knowledge in fusion, which thus is more similar to the health science, where a holistic view of the patient is also desirable, in order to optimize the treatment and reduce negative side effects.

The task of attaining a global and coherent view of the plasma state is a real challenge for a series of aspects of the involved physics. Since fusion plasmas are open systems far from equilibrium, it is very difficult to formulate satisfactory models starting from basic physics principles. Moreover the nonlinear interactions among many variables not only increase the difficulty of the interpretation but also pose some limits to the degree of control of the experiments which can be performed.

All these challenges are complicated by the amount of data which are produced by fusion diagnostics. A discharge of the Joint European Torus (JET) can produce more than 10 Gbytes of data. Since several tens of discharges of this sort can be performed in a day of operation, the shear amount of data renders the analysis prohibitively difficult. In order to alleviate the problems just mentioned, in the last years some efforts have been devoted to the development of statistical and soft computing methods to support the analysis and interpretation of the experiments [Rizzo and Xibilia, 2002]. On

[1] See the Appendix of F. Romanelli et al., Proceedings of the 22nd IAEA Fusion Energy Conference 2008, Geneva, Switzerland

the other hand, as summarized in [Murari *et al.*, 2008] and [Vega *et al.*, 2008], till nowadays the various developed automatic analysis techniques have been limited to exploratory aspects.

In this paper, a methodology which, starting from raw data, proceeds up to the point of providing an intuitive interpretation of physics involved in a purely automatic way is applied to a specific nuclear fusion problem, i.e. magnetic confinement regime classification. The first step consists of an exploratory phase, performed with the approach of the Classification and Regression Trees (CART) [Breiman *et al.*, 1984]. The CART algorithm allows making a data driven decision about the most important variables for the problem at hand and provides a classification tree. It is worth pointing out that the CART method is fully non-linear and unbiased. It does not require any form of preliminary signal processing (not even normalization) and therefore is considered particularly suited to the exploratory phase of identifying the data with the highest information content. A fully automated procedure has been developed to derive a fuzzy logic classifier (FLC) from the classification tree obtained in the previous step explicitly oriented to the production on an interpretable system. The following criteria, proposed in [Guillame, 2001] and [Zhou and Gan, 2008], were followed to achieve an interpretable fuzzy system:

1. the fuzzy partition should be readable, i.e. it should be possible to interpret the fuzzy set as linguistic labels, the fuzzy sets must be distinguishable and they should be in moderate number;
2. the set of rules should be as small as possible, preserving the performance;
3. the rules should be incomplete, i.e. the rules premises should not exceed the limit of 72 conditions.

The FLC so produced provides a series of explicit rules which turn out to be very useful for the interpretation of the physics involved. With the proposed method, the transparency of the fuzzy logic is combined with the maintenance ease typical of other black box methods like neural networks. To illustrate the potentiality of the method, in this paper a specific FLC is applied to the problem of confinement regime identification in JET. The results of fuzzy logic inference systems are quite positive. First of all the derived rules are very sound and quite intuitive; they formalize a body of common sense knowledge about the L-H transition that can contribute to acquire knowledge and deeply inspect the underlying physics. The validity of these rules is proved by the classification capabilities of the system, which can exceed 90 % over the whole discharge.

2 Confinement Regime At Jet

Tokamak configuration can be operated in different confinement regimes, which can be significantly different not only in terms of performance but also of physics interpretation and control requirements. The High Confinement regime, the so called H mode [Wagner *et al.*, 1982], is a particularly relevant example. Its performance can increase of more than a factor of two compared to the L-mode (Low confinement), but it is affected by edge instabilities. These require particular measures to avoid disruptions and could be very dangerous for the integrity of the entire device in the next generation of machines like ITER. Moreover the H-mode of confinement is a self-organized state of the plasma, which develops spontaneously when certain conditions are met. The transition from one mode of confinement to the other presents some characteristics of phase transitions already studied in many other physical systems. On the other hand, the details of the plasma evolution from the L-mode to the H-mode state have not been fully understood yet, neither from the dynamics aspects, nor from power requirements to trigger the transition. Even the optimal control parameters remain unclear.

In addition to physical interpretation, the H-mode of confinement presents some important challenges form the control point of view. The higher internal energy of the plasma, together with the increased plasma shaping used to improve performance, make the H-mode significantly more unstable and more prone to disruptions. On the other hand, up to now no reliable classifier of the confinement regime is available for real time operation and therefore normally tokamak devices are controlled in feed-forward, predicting a priori in which type of regime the plasma will be during the discharge. In case of unexpected transitions from state to another the control systems can therefore adopt a non optimal strategy which can not only limit the performance but also compromise the equilibrium and contribute to trigger disruptions [Franzen *et al.*, 1998]. In the perspective of ITER, in which accessing the H-mode is essential but disruptions can have very harmful consequences, it is becoming urgent to develop models to better interpret the H-mode physics and to identify the regimes in real time.

3 Database And Features Selection

The database used for the analysis reported in this paper covers the shot range between 55211 (21/03/2002) and 62723 (28/01/2004) and therefore refers mainly to JET divertor configuration with the Septum. More details about these discharges can be found in [Meakins, 2008]. The database is composed of 55 pulses and 29 signals which comprises both unprocessed quantities and processed ones. Among the last ones, some of the parameters taken into account, such as the electron temperature (T_e) and the axial toroidal magnetic field at 80% of the flux (B_{T80}), are not calculated routinely at JET but were calculated specifically for this work.

The objective of this work is to produce an interpretable fuzzy system. A limited number of input variables is, therefore, mandatory to have readable fuzzy sets and rule base (see [Guillame, 2001] and [Zhou and Gan, 2008]) and it can be achieved trough a feature selection step to select among the variables in the database the most relevant for the description of the problem. The instrument selected to perform this step is Classification and Regression Trees [Breiman *et al.*, 1984]. It is a non-parametric statistical method, which uses a decision tree to solve classification and regression problems using both categorical and continuous variables.

One of CART features is the evaluation of the importance of the different explanatory variables, i.e. the variables provided as input during the building of the tree, to describe the output through the so-called "variable ranking method". The importance value produced through the variable ranking allows sorting the different input signals from highest to lowest or zero important one. Since the aim of the work is, also, to derive an intuitive understanding of the confinement transition in tokamak plasmas, the time slices of the various discharges have been divided in three subsets and analyzed separately. The three subsets include data around the transition from L to H, around the transition back from H to L, and away from the transitions in steady state L and H mode phases. In more detail, the first two datasets include data acquired 300 ms before and after the L→H and the H→L transition. The last one is obtained using intervals of 500 ms in L mode, between 1200 ms and 700 ms before the L→H transition and intervals of 500 ms in H mode, between 1200 ms and 700 ms before the H→L transition.

The three subsets have been provided to CART in order to build three different trees. According to the variable ranking provided by CART the signals reported in Table 1 have been evaluated as the most relevant. In addition to these physical quantities, it has been considered important to test also the influence of geometrical parameters that can account for the position/shape of the plasma inside the vacuum vessel. Several tests have been performed using various quantities and the ones which have produced the best results are the radial and vertical position of the X point. These parameters have been appended to the ones reported in Table 1 in all the subsets used to obtain the results described in the following.

Table 1. The Four Most Relevant Variables For The Three Different Subsets. The Signals Are Sorted In Descendent Order Of Importance

L→H	
Symbol	Description
W_{mhd}	Magnetohydrodynamic energy
β_N	Beta normalized over diamagnetic energy
T_e	Electron temperature
B_{T80}	Axial toroidal magnetic field at ψ=0.8
H→L	
Symbol	Description
β_N	Beta normalized over diamagnetic energy
B_T	Toroidal magnetic field
FDWDT	Time derivative of diamagnetic energy
q_{95}	Safety factor at ψ=0.95
Steady State	
Symbol	Description
β_N	Beta normalized over diamagnetic energy
B_T	Toroidal magnetic field
Lid4	Outer interferometry channel
FDWDT	Time derivative of diamagnetic energy

4 From Cart Rules To Fuzzy Rule

Once the most relevant variables have been selected, a new tree is produced providing only the selected variables as predictors. The output from the CART represents the input for the automatic FLC construction which comprises three steps:

1. Extraction of the crisp rules from the classification tree.
2. Determination of the membership functions from the set of crisp rules obtained in the previous step.
3. Formulation of the fuzzy rules on the basis of the classification tree crisp rules and the membership functions.

4.1 Extraction Of The Crisp Rules From The Classification Tree

The tree is parsed automatically to determine a rule for each terminal node. The definition of these rules is performed in the following way. For each terminal node, the corresponding branch is scanned up to the root and a specific rule is devised for each intermediate node, on the basis of the inequality used by CART at each node to perform the split.

4.2 Determination Of The Membership Functions

In order to satisfy the first criteria for interpretability, i.e. the readability of fuzzy partitions, the number of fuzzy membership functions was limited to three trapezoidal function since they provides enough flexibility to cover each variable domain by dividing it in three different regions. The selection of the parameters identifying the trapezoidal function is performed evaluating the cumulative discriminating power [Vagliasindi *et al.*, 2009] of the various splitting values selected when the variable is used as a splitter. In other word, each crisp rule provided by the tree is ranked according to the number of sample it is able to discriminate. Since each rule is the combination of many conditions involving different variables, the splitters, and different values, the splitting value, a discriminating power is associated to each splitting value depending on the rule in which it compares. For each variable, the splitting values having the two highest cumulative discriminating powers are selected and used to build the trapezoidal membership functions, as in equation 1 where α and β, with $\alpha > \beta$, are the two splitting variables selected and δ is a parameter defining the slope of the shoulders.

$$\left.\begin{array}{l} \mu_{F_j}^1 = [MIN \quad MIN \quad \beta-\delta \quad \beta+\delta] \\ \mu_{F_j}^2 = [\beta-\delta \quad \beta+\delta \quad \alpha-\delta \quad \alpha+\delta] \\ \mu_{F_j}^3 = [\alpha-\delta \quad \alpha+\delta \quad MAX \quad MAX] \end{array}\right\} \qquad (1)$$

4.3 Formulation Of The Fuzzy Rules

Each rule provided by the tree has already a form similar to a fuzzy rule apart from the fact that the antecedent is composed of inequalities using crisp values. We have then to translate the inequalities of the tree rules for the terminal nodes into inequalities based on the membership functions defined in the previous section. This is achieved selecting the fuzzy membership function which best approximates the crisp inequalities produced during the tree construction. A detailed description of the above mentioned steps is available in [Vagliasindi *et al.*, 2009].

5 Evaluation Of The Performance And Discussion

Since the total number of rules is affected by the number of terminal nodes of the related tree, the complexity of the produced FIS also depends on nodes retained in the CART trees. On the other hand some of the terminal nodes discriminate a very limited number of samples; therefore the corresponding rules may introduce an excessive increase in complexity compared to the additional discrimination capability they provide. Therefore, an investigation of the number of terminal nodes, and consequently the number of rules, which give the best results in terms of correct classification of samples has been performed. At the same time, an investigation of the optimal threshold value of the output, which maximizes the classification performance, has also been carried out.

Figure 1 reports the results of the above mentioned investigations for a test set of 17 pulses and a total of 234429 samples. The final performance of the various threes is represented as a surface plot versus the numbers of nodes and the threshold values. From these surface plots, it can be noticed that with a high number of terminal nodes and, consequently, a high number of rules in the corresponding FIS, the performance tends to be low and the best threshold value is 0.5. This suggests that too many rules may not necessarily increase the discrimination capability of the system. On the contrary, being some of them very specific and related to a small subset of the data, they may cause overfitting and move the output of the network to a small neighborhood of 0.5. When reducing the number of nodes, the performance raises and tends to flatten both in the nodes and threshold direction. The flattening in the nodes direction may be caused by the data provided as test samples. Indeed, being the test data represented

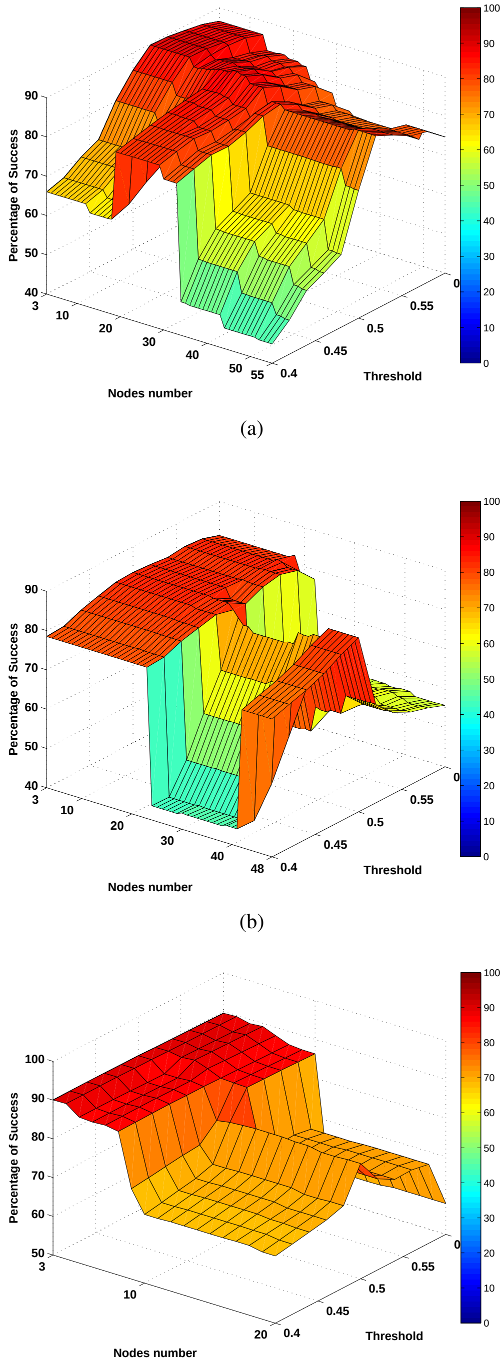
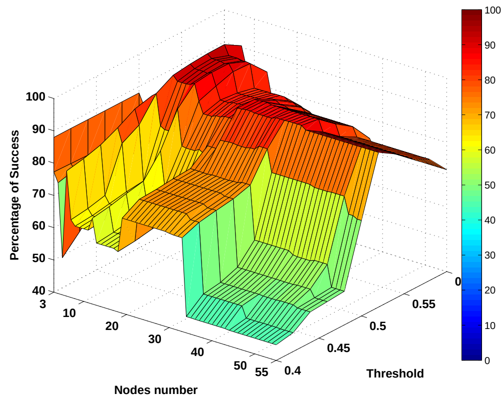
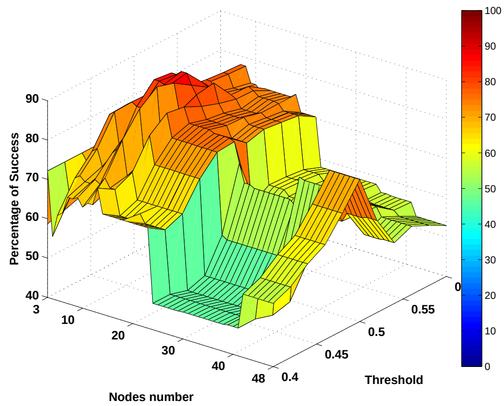


Figure 1. Performance of the FIS generated from the CART data on full set of data as function of the number of terminal nodes taken into account and the threshold chosen to discriminate between the L and H mode. a), b) and c) are the results when the FIS is trained using data near the L⟶H transition, near the H⟶L transition and in steady-state condition respectively.

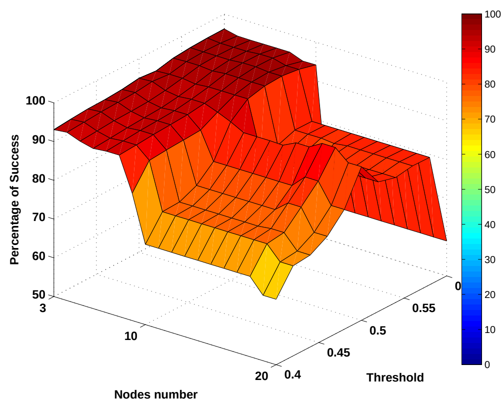
by whole pulses data, the majority of them are samples distant from a transition. Therefore, L and H samples present very different characteristics in terms of values in the space of parameters so that only a small number of rules are activated even if a large number is present. The same reason can explain the flattening



(a)



(b)



(c)

Figure 2. Performance of the FIS generated from the CART data on specific test intervals as a function of the number of terminal nodes taken into account and the threshold chosen to discriminate between the L and H mode. a), b) and c) are the results when the FIS is trained and tested using data near the L⟶H transition, near the H⟶L transition and in steady-state condition respectively.

in the threshold direction. Being the data mainly represented by samples distant from a transition, they are well discriminated by the systems for a wide range of thresholds.

To confirm these hypotheses, the three FIS have been tested on a subset of the full set of data available for the test. In particular they have been applied to data taken from the specific interval used for training them. So the fuzzy inference system developed starting from the classification tree built using data around the L→H transition (from now on referred to as LH-FIS) has been tested on data around the L→H transition, the FIS developed using data in the neighborhood of H→L transition (HL-FIS) with data around the H→L transition and the FIS devised starting from data far from the transition (referred to as SS (steady-state)-FIS) with data far from the transition.

The results of these tests are reported in Figure 2(a),2(b) and 2(c) respectively. Whereas the results achieved by the SS-FIS (Figure 2(c)) are comparable with Figure 1(c) although with an higher percentage of success, being the data similar but without the samples in the more uncertain region of L→H and H→L transition, the performance surface of LH-FIS and HL-FIS are slightly different. It is possible to observe (Figure 2(a)-2(b)) that the flattening of the surface at small node numbers is no more present and the maximum performance are achieved with 7 and 11 nodes respectively indicating that more information is required to distinguish between L and H mode samples near a transition. Table III summaries the above mentioned results in a numerical way, showing the maximum percentage of success for the various developed FIS, together with the number of nodes taken into account to build the FIS and the threshold value which best discriminate between L and H mode samples. With regard to the maximum percentages of success achievable by the various FIS, it can be noticed that while SS and LH are comparable, the classifier tuned on the HL exhibits significantly lower performance. This can be due to a greater uncertainty in the H→L transition times contained in the database and, therefore, in a more uncertain classification of the samples in the neighborhood of the H→L transition. This confirms the long suspected fact that the H→L is less defined and more difficult to pin point with the measurements available. The consequent uncertainties can lead to both erroneous learning during the training phase and a wrong estimation of the results during the testing phase.

The automatic procedure, in addition to classifying the plasma confinement regime with a very high rate of success, can also provide an intuitive interpretation of the plasma behavior at the transition. Qualitative indications about the plasma dynamics can indeed be obtained by the rules devised in an automatic way by the proposed methodology. An example of the rules de-

Table 2. Performance On Full Test Set And Specific Test Sets

	Full Test			Specific Test		
	(%)	Thr	Rules	(%)	Thr	Rules
L → H	89.70	0.52	22	93.58	0.58	7
H → L	84.31	0.5	42	87.45	0.48	11
SteadyState	90.13	0.44	4	96.14	0.56	3

Table 3. The Rules Of The Fis System For The L To H Transition With The Highest Performance

♯	**Linguistic Term**
1	IF T_e is not high and W_{MHD} is low and $XPzl$ is not low then Output is L
2	IF B_{T80} is not low and W_{MHD} is high then Output is H
3	IF β_N is high and B_{T80} is not high and T_E is high and W_{MHD} is not high then Output is L
4	IF β_N is not low and B_{T80} is not low and T_E is high and W_{MHD} is not high then Output is H
5	IF B_{T80} is not low and W_{MHD} is medium and $XPzl$ is not high then Output is H
6	IF B_{T80} is not high and $XPrl$ is low and T_E is not low and W_{MHD} is not high and $XPzl$ is not low then Output is H
7	IF B_{T80} is low and W_{MHD} is medium and $XPzl$ is not high then Output is L

rived by the FIS for the L to H transition is reported in Table 3.

These rules are considered quite realistic and summaries properly a significant amount of expert knowledge in the field. Therefore these rules confirm the good quality of the final FIS and the soundness of the automated method used to derive them. Moreover, the intuitive and fully transparent character of the final rules is a very positive outcome of the approach. Contrary to other methods, like Artificial Neural Networks, which provide good performance but are practically black boxes very difficult to interpret, the proposed automatic FIS gives a clear set of rules which can result particularly useful to get a first grasp of completely new or poorly understood phenomena.

6 Conclusion

An automatic learning system for fuzzy inference system based on crisp classification trees has been applied to a nuclear fusion problem. The method achieves a good compromise between accuracy and interpretability. It is, indeed, able to reach about 90 % of accuracy on the whole data set and up to the 96 % when tested on specific regions of the data set. The derived FIS, on the other hand, presents a high grade of interpretability. According to [Guillame, 2001], in order for a set of rules to be interpretable it should satisfy the following requirements: the fuzzy partition should be readable, i.e. it should be possible to interpret the fuzzy set as linguistic labels, the set of rules should be as small as possible and the rules should be incomplete, i.e. the rules premises should not involve all the input variables. All of these requirements are satisfied by the proposed approach. Indeed, the fuzzy partition is simple since the number of fuzzy set is limited to three and readable since each fuzzy set is associated to a linguistic label. According to Table 2, the best performance is usually achieved when the number of rules is small and, even when the best performing network has an high number of rules, the reduction in accuracy, when a lower number of rules is used, is limited to just a few %. Finally, as it is possible to observe in Table 3, where an example of the rules produced by the algorithm is reported, most of the rules involve just a few of the input variables leading to an incomplete rule set.

Acknowledgements

The work was partially supported by the project "Realtime visual feature extraction from plasma experiments for real time control," funded by ENEA-EURATOM, 2008.

References

Rizzo, A. and Xibilia G. (2002), An Innovative Intelligent System for Sensor Validation in Tokamak Machines, *IEEE Trans. On Control Systems Technology*, Vol. 10, No. 3.

Murari, A. *et al.* (2008). How to Extract Information and Knowledge from Fusion 1 Massive Databases. In *Burning Plasma Diagnostics: An International Conference*, vol. 988, pp. 457-470.

Vega, J. *et al.* (2008). Data mining technique for fast retrieval of similar waveforms in Fusion Massive databases. In *Fusion Eng. and Design*, vol. 83, no. 1, pp. 132-139.

Breiman, L. *et al.* (1984). *Classification and Regression Trees*. Chapman & Hall (Wadsworth, Inc.). New York

Wagner, F. *et al.* (1982). Regime of improved confinement and high beta in neutral beam heated divertor discharges of the ASDEX tokamak. In *Physical Review Letters* 49, 1408

Franzen, P. *et al.* (1998). On-line confinement regime identification for the discharge control system at ASDEX Upgrade. In *Fusion Technol.* vol. 33.

Meakins, A.J. (2008). *A Study of the L-H Transition in Tokamak Fusion Experiments*. PhD Thesis. Imperial College London.

Guillame, S. (2001). Designing Fuzzy Inference Systems from Data: An Interpretability-Oriented Review. In *IEEE Transactions on Fuzzy Systems* vol. 9, no. 3, pp. 426-443.

Zhou, S.M. and Gan, J.Q. (2008). Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modeling. In *Fuzzy Sets and Systems*, vol. 159, pp 3091-3131.

Vagliasindi, G. *et al.* (2009). CART data analysis to attain interpretability in a Fuzzy Logic Classifier. In *International Joint Conference on Neural Networks (IJCNN09)*, Atlanta, U.S.A., June 14-19, pp. 3164-3171.

Zadeh, L.A. (1965). Fuzzy sets. In *Infor. Control*, vol. 8, pp.338-353.