

A. Murari, G.Vagliasindi, P. Arena, L. Fortuna, O.Barana, M. Johnson
and JET-EFDA Contributors

Prototype of an Adaptive Disruption Predictor for JET Based on Fuzzy Logic and Regression Trees

“This document is intended for publication in the open literature. It is made available on the understanding that it may not be further circulated and extracts or references may not be published prior to publication of the original when applicable, or without the consent of the Publications Officer, EFDA, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK.”

“Enquiries about Copyright and reproduction should be addressed to the Publications Officer, EFDA, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK.”

Prototype of an Adaptive Disruption Predictor for JET Based on Fuzzy Logic and Regression Trees

A. Murari¹, G.Vagliasindi², P. Arena², L. Fortuna², O.Barana¹, M. Johnson³
and JET-EFDA Contributors*

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

¹*Consorzio RFX-Associazione EURATOM ENEA per la Fusione, I-35127 Padova, Italy.*

²*Dipartimento di Ingegneria Elettrica Elettronica e dei Sistemi-Universit degli Studi di Catania, 95125 Catania, Italy*

³*EURATOM-UKAEA Fusion Association, Culham Science Centre, Abingdon, Oxon OX14 3DB, UK*

** See annex of M.L. Watkins et al, "Overview of JET Results",
(Proc. 21st IAEA Fusion Energy Conference, Chengdu, China (2006)).*

ABSTRACT

Disruptions remain one of the most hazardous events in the operation of a Tokamak device, since they can cause damage to the vacuum vessel and surrounding structures. Their potential danger increases with the plasma volume and energy content and therefore they will constitute an even more serious issue for the next generation of machines. For these reasons, in the last years a lot of attention has been devoted to devise predictors, capable of foreseeing the imminence of a disruption sufficiently in advance, to allow time for undertaking remedial actions. In this paper, the results of applying Fuzzy Logic and Classification and Regression Trees (CART) to the problem of predicting disruptions at JET are reported. The conceptual tools of Fuzzy Logic, in addition to being well suited to accommodate the opinion of experts even if not formulated in mathematical but linguistic terms, are also fully transparent, since their governing rules are human defined. They can therefore help not only in forecasting disruptions but also in studying their behaviour. The analysis leading to the rules of the Fuzzy Predictor has been complemented with a systematic investigation of the correlation between the various experimental signals and the imminence of a disruption. This has been performed with an exhaustive, non-linear and unbiased method based on Decision and Regression Trees (CART). This investigation has confirmed that the relative importance of various signals can change significantly depending on the position of the plasma in the parameter space. On the basis of the results provided by CART on the information content of the various quantities, the prototype of an adaptive Fuzzy Logic predictor was trained and tested on JET database. Its performance is significantly better than the previous static one, proving that more flexible prediction strategies, not uniform over the whole discharge but tuned to the operational region of the plasma at any given time, can be very competitive and should be investigated systematically.

1. INTRODUCTION AND RATIONALE

The Tokamak configuration is affected by uncontrolled and irreversible losses of confinement, called disruptions. Plasma disruptions are considered one of the most critical issues for next step devices, since they constitute a severe hazard for the integrity of the machine [1]. This is due to the fact that, on a typical time scale of a few milliseconds, the energy stored in the plasma is lost, producing a large heat flux to the plasma facing components. Moreover, during the fast evolution of the magnetic configuration, high electro-mechanical stresses are induced on the Tokamak structures; a significant amount of runaway electrons can also be generated, with additional potential detrimental effects on the vacuum vessel.

Even if disruptions can be potentially very dangerous, in present day machines the quest for better performance pushes the plasma close to its operational limits. In this perspective, reliable methods to predict the occurrence of a disruption are extremely important, since forecasting sufficiently in advance the imminence of disruption can give the opportunity to intervene and avoid or alleviate the hard conditions a disruption can lead to. This will become even more crucial in the next generation of experiments, where unforeseen disruptions could have serious consequences.

Since Tokamak fusion plasmas are complex, nonlinear systems, driven far from equilibrium by powerful additional heating systems, many causes can destabilise them and cause a disruption. A complete model, allowing to reliably predict the plasma configurations which are going to disrupt, does not exist. Therefore, due to the lack of algorithmic solutions, in the past many efforts have been devoted to predicting disruptions using neural networks. This approach of neural networks was used to predict the disruption boundary for the high- β_N disruption in DIII-D [2]. The network, a three layer feedforward perceptron, used as inputs 33 diagnostic signals and produced as output the value of the normalized beta $\beta_N = \beta_t(aB)/I_p$ at the time of disruption for each discharge. The ratio of the actual value of β_N and the predicted one was used as the detection parameter. A neural network for predicting the time to disruption (ttd) in ASDEX Upgrade tokamak [3] was developed using seven signals and some of their derivatives as inputs. In JET a two layer feedforward perceptron model [4] was developed and trained to distinguish between pre-disruptive and stable plasmas. The best performance (90% of disruptive pulses detected and 5% of false alarms) were obtained using seven input parameters: locked mode amplitude, density, input power, radiated power, the safety factor (q_{95}), the plasma internal inductance (l_i), the poloidal beta (β_p) and the derivative of the stored energy. Further developments in this direction are described in [5], which reports the results of estimating the probability of disruption using an artificial neural network (ANN). The ANN was trained using both disrupted and safe pulses with the following conditions: plasma current $I_{pla} > 1.5MA$, X-point configuration and flat-top plasma current profile. In [6] a fuzzy neural approach for plasma disruption prediction was proposed. A Fuzzy Inference System (FIS) was used in combination with a neural network to predict the time to disruption (ttd), achieving a Root Mean Squared Error (RMSE), averaged on the total number of shots, of 0.0191. An alternative fuzzy time series approach was also proposed in [7], again providing the ttd as output and predicting the onset of a disruption, in the range from 400 to 0.5 ms before the disruption, with a success rate of the order of 93% with very limited false alarms.

A Fuzzy Logic predictor was also developed by some of the authors as reported in [8]. The approach adopted in that case (see section 4 for more details) consisted of developing a series of Fuzzy Rules to determine the probability of disruption. Contrary to NNs, which are black boxes and can be used only as transfer functions, Fuzzy Logic is fully algorithmic and therefore allows drawing a direct link between any inference rule and its effects on the prediction accuracy. This renders much easier the interaction with the experts and, on the other hand, can contribute to the learning process, since the various theories can be confirmed or falsified by the output of the predictor. The Fuzzy predictor described in [8] provided as output the probability of disruption, which was considered a more consistent quantity to provide than the time to disruption, which is for example very difficult to interpret if, for any reason, the disruption does not really take place after the time of its occurrence having been predicted. Moreover, the probability of disruption, as determined by the position of the plasma in the operational space, is also a much better quantity to use in the phase of scenario development, when new configurations are designed and the associated risk has to be evaluated. Indeed this information on the probability can be combined with estimates of the potential damage in a consistent risk evaluation

assessment. On the other hand, as in all the previous works, a static predictor was devised, optimised on the basis of the data of the whole database and valid for the entire flat top phase of the discharge.

In this paper, an alternative approach to disruption prediction for JET is presented, based again on Fuzzy Logic but with an adaptive character. This was motivated by a feature selection analysis, performed on the Classification And Regression Tree (CART) [9] method, which was undertaken to assess the relative importance of the various input signals, the same used in [8]. This has shown that, depending on the plasma status, and therefore on the probability of disruption, the various signals have a clearly different relevance in detecting the probability of disruption (see section 5). Therefore it was decided to develop an adaptive fuzzy predictor (see section 6), which consists of different sets of Fuzzy Rules optimised for various plasma regimes, characterised by a different probability of disruption. Given the positive results obtained with this prototype, possible strategies for a real time predictor, based on this adaptive approach, have been identified, which would allow providing an estimate of the disruption probability at various future times on the basis of the present plasma state .

With regard to the structure of this paper, in the next session a brief review of the disruptions evolution and their causes are given to set the stage. In section three the basic elements of Fuzzy Logic are summarised to provide the reader, not familiar with this topic, with the background indispensable to follow the rest of the paper. The description of the adopted database, of the diagnostic signals used as inputs and a brief overview of the predictor originally developed in [8], are the subject of section four. In section five, the correlation analysis of the inputs and the nonlinear correlation between the inputs and the desired output, i.e. the probability of disruption are reported. In the following section six the main characteristics of the new adaptive Predictor and the obtained performance are described in detail. Section seven contains a summary of the results and an outline of the possible evolution of the work, both on JET and other devices.

2. DISRUPTION EVOLUTION AND CAUSES

In a disruption it is possible to identify four different phases [1], [10]:

- 1) the initiating event
- 2) the precursor phase
- 3) the thermal quench
- 4) the current quench

The initiating event can be external to the plasma, for instance a mechanical failure or a glitch in the control system, or internal, i.e. unforeseen plasma conditions, like unusually high radiated power, that can bring the plasma to a disruptive phase. During the precursor period, which depends strictly on the initiating event, MHD instabilities develop in the plasma, leading it to the thermal quench. During the thermal quench almost the entire thermal content of the plasma is lost. This phase is followed by the current quench, during which the magnetic energy is dissipated mainly by radiation [11].

In order to have a sufficient time for the intervention, a predictor should produce an alarm during

the “precursor phase”. Depending on the different initiating event, there are different precursor events that should be taken into account and these can be better identified by looking at different signals (see section 5). Given the nature of the precursor phase, for classification purposes it is customary to distinguish different disruption classes:

- a) Mode Lock (ML)
- b) Density Limit (DL)
- c) High radiated power (RP)
- d) H/L mode transition (HL)
- e) Internal Transport Barrier (IT)
- f) Vertical Displacement Event (VDE)

2.1. MODE LOCK

During plasma operation, MHD instabilities can occur due to the presence of an external perturbation of the magnetic field or to instable current or pressure gradients, leading to the creation of the so called magnetic islands. These perturbations typically rotate in the frame of reference of the laboratory, inducing eddy currents in the wall of the vacuum vessel. These currents produce a magnetic field that, due to the Lenz’s law, acts to stabilize the instability. They also produce a slowing down of the rotation velocity of the instability, leading to a decrease of the eddy currents themselves and to a reduction of their stabilizing effects. It can happen that the instability reaches a state in which it ceases to rotate and “locks to the wall”. In these conditions, the instability can grow much more rapidly and eventually reach dimension that can occlude a whole sector of the torus, so modifying the current profile to the point of inducing a disruption.

2.2. DENSITY LIMIT

Every tokamak has both a low and high density limit [12]. The upper density limit is the most relevant as the fusion reaction rate scales with the square of the density, so the aim of a fusion experiment would consist of working at the highest density achievable avoiding disruptions in order to maximize the energy output. In general, there is no fixed high density limit as it depends on the plasma configuration. Usually, increasing the density causes an increase in the radiated power, in particular due to line radiation from the edge of the plasma, where the local electron temperature decreases at high densities. When the radiated power exceeds the local heating power, the plasma detaches from the wall and the temperature and current profile contract. This situation is unstable and typically leads to a disruption.

2.3. HIGH RADIATED POWER

They are quite similar to the density limit disruptions apart from the fact that, in this case, the increment in the radiated power occurs slightly before the increment in the plasma density. It is due to an increase

in the presence of impurities, which cause an increment of the radiation. A particularly dangerous radiating instability is the MARFE (Multifaceted Asymmetric Radiation From the Edge) [13]. It is a poloidally localised, toroidally symmetric region of increased emission where the power lost by radiation becomes larger than the power flowing into this region by conduction. As the plasma cools, the radiating power increases, further enhancing the cooling process and leading to a disruption.

2.4. H/L TRANSITION

The H-mode, or High confinement, regime is a state of the plasma, reached when the plasma heating overcomes a characteristic power threshold, characterized by a better confinement. An H/L transition is a transition from the H-mode to the L-mode, or low confinement, regime. When this occurs at high density, usually after a premature reduction of the additional heating, it can lead to a disruption (typically a density limit type disruption).

2.5. INTERNAL TRANSPORT BARRIER

In some plasma configurations, usually called “advanced scenarios”, an Internal Transport Barrier (ITB) can appear. These barriers reduce considerably the transport of particles and energy from the inner to the outer side of the plasma. An ITB requires a steep pressure gradient that can lead to instabilities and disruptions. This kind of disruption is very difficult to predict as the time-scale of the precursor phase is really low. Usually the time from the occurrence of the precursor signal (a high variation in the Poloidal Beta signal) and the thermal and current quench phase is very short, of the order of few tens of ms.

2.6. VERTICAL DISPLACEMENT EVENT

Other things being equal, configurations with elongated shapes have better performance with respect to plasmas with circular cross-section. These configurations, on the other hand, are vertically unstable so they require the presence of a vertical position and velocity controller. If this controller fails, the plasma naturally moves towards the vacuum vessel producing a vertical displacement disruption. Usually this kind of disruption is very difficult to foresee as the dynamic of the event is very fast, but it is very unlikely for these disruptions to happen spontaneously. Usually they are due to failures of the control system.

The high number of possible disruption causes, interacting nonlinearly, and the unavailability of a dynamical model of the plasma behaviour, particularly at the boundary of the operational space where disruptions are more likely, are the main reasons which in the past motivated the adoption of black box approaches like ANNs for disruption prediction. On the other hand, the methods of Fuzzy Logic seem to be quite suited to this application, since they allow modelling complex non linear systems and handling operational spaces with non Boolean boundaries. Fuzzy Logic has also been applied successfully in the few last years to manage knowledge expressed in linguistic form. This is somehow also the case of disruption prediction in Tokamaks, where the experts possess a significant amount of

knowledge, which is difficult to express in the formal language of rigorous mathematics but can be extremely important in helping to predict the occurrence of disruptions efficiently. The basic methods of Fuzzy Logic to operate with fuzzy sets and nonlinear systems and their potential to handle linguistic variables, and therefore to mimic human reasoning, are briefly summarised in the next section.

3. BRIEF INTRODUCTION TO FUZZY LOGIC METHODS

The term ‘‘Fuzzy Logic’’ [14] has become of widespread use today to identify a coherent logical edifice, based not on crisp sets, like Boolean logic, but on fuzzy sets. A fuzzy set is an alternative to the traditional notion of set membership and logic that has its origin in ancient Greek philosophy and in the so-called ‘‘Law of the Excluded Middle’’ originally formulated by Aristotle. It states that an element X must either be in set A or in set not- A , so that a classical or ‘‘crisp’’ set is a container that completely includes or excludes any given element.

In many practical applications, it is often necessary to handle phenomena characterised by unsharp boundaries, for which it is not always easy to attribute drastically the membership of an element to a set or its complementary. To overcome this difficulty, Zadeh introduced the notion of fuzzy set, which allows defining the membership of an element to a set using the whole interval $[0,1]$ and not only the Boolean discrete values 0 and 1. In more detail, a fuzzy set F defined on U (called the universe of discourse) is given by [14]:

$$F = \{(u, \mu_F(u)) \mid u \in U\}$$

where $\mu_F(u)$ is the membership function, a curve that specifies how each element of U is mapped to the real interval $[0, 1]$, that is

$$\mu_F(u) : U \rightarrow [0, 1]$$

In other terms, for each $u \in U$, the membership function defines the degree of membership of u to the fuzzy set F . The degree varies continuously from zero (no membership) to one (full membership) according to the particular properties of the fuzzy set.

In analogy to classical set theory, for fuzzy sets it is possible to define a series of quantities and properties. One of the most basic is the *support* of a fuzzy set F , which is the crisp set $S(F)$ composed by the elements which have a nonzero degree of membership to the fuzzy set F : $SF = \{u \in U \mid \mu_F(u) > 0\}$. Another essential definition is that of the *complement* of a fuzzy set \bar{F} , which is the fuzzy set F defined by the membership function $\mu_{\bar{F}}(u) = 1 - \mu_F(u)$.

Since it has been demonstrated that fuzzy logic can be considered a superset of standard Boolean or crisp logic, it is possible to extend the logical operation AND and OR. Given two fuzzy sets A and B , the union of the two fuzzy sets A OR B is a fuzzy set given by the membership function $\mu_{A \text{ or } B} = \max(\mu_A(u), \mu_B(u))$. The intersection of two fuzzy sets, A AND B is the fuzzy set defined by the membership function $\mu_{A \text{ and } B} = \min(\mu_A(u), \mu_B(u))$.

On the basis of the previously described fuzzy sets, it is possible to define a fuzzy system, which

consists of a nonlinear set of fuzzy rules and an appropriate inference mechanism. The set of rules represents the knowledge base of the fuzzy system and the inference mechanism processes the knowledge base in a quantitative way to yield the desired results. Fuzzy rules are particularly suited to formulate conditional statements based on fuzzy notions and can therefore be used to process information in a way much more similar to human decision making [15] than classical logic. Usually a fuzzy rules is a if-then rule of the form:

if x is A then y is B

where A and B are two fuzzy sets defined in the universes of discourse X and Y , with $x \in X$ and $y \in Y$.

A typical fuzzy system is a logical procedure that maps a set of scalar inputs to one scalar output. In general, a typical fuzzy system with N inputs u_1, \dots, u_N , one output variable v_0 and M fuzzy rules can be written in the following form:

Rule 1. IF $(u_1, F_{1,1})$ AND $(u_2, F_{2,1})$ AND ... AND $(u_N, F_{N,1})$ THEN (v_0, G_1)

...

Rule j . IF $(u_1, F_{1,j})$ AND $(u_2, F_{2,j})$ AND ... AND $(u_N, F_{N,j})$ THEN (v_0, G_j)

...

Rule M . IF $(u_1, F_{1,M})$ AND $(u_2, F_{2,M})$ AND ... AND $(u_N, F_{N,M})$ THEN (v_0, G_M)

$F_{i,j}$ identifies the fuzzy set associated with the i_{th} input variable in the j_{th} rule and G_j is the consequent fuzzy set associated with the output variable in the same rule.

Evaluating an if-then rule implies two different steps: first evaluating the antecedent, i.e. fuzzyfying the input and applying the necessary fuzzy operators recalled in the antecedent clause, then applying the result to the consequent, the so-called *implication*.

Evaluating the antecedent implies the calculation of the degree of activation of the rule. Referring to the typical fuzzy set reported before, when the operation among the input in the antecedent clause is an AND, the degree of activation λ_j for the j_{th} rule is evaluated as follows:

$$\lambda_j = \min \{ \mu_{F_{i,j}}(u_i) \}$$

The resulting consequent can be considered a new fuzzy set G'_j defined by the membership function:

$$\mu_{G'_j}(u) = \lambda_j \mu_{G_j}(u)$$

Once evaluated the rules, that is the degree of activation of all the single fuzzy rules and the degree of membership of the output variable to the various fuzzy sets involved in the implication, it is necessary to aggregate the results of the various rules together in order to obtain a scalar value, that represents the output of the fuzzy system, v_0 . This process is called *inference* or *defuzzification*.

4. DATABASE, INPUT SIGNALS AND THE STATIC FUZZY PREDICTOR

4.1 DATABASE

At JET a disruption database has been created and is regularly maintained to monitor the disruptability (the ratio of disruptive and safe discharges) of the scenarios and to record the force each disruption produces on the machine. This database contains all disruptions since discharge 27000. A subset of this database is devoted to the EFDA period and records all disruptions since year 2000. The database used for the original Fuzzy detector described in [8] was extracted from this EFDA database, adopting the same criteria as applied in previous works done at JET [5], i.e. plasma current above 1.5 MA and flat-top X-point configurations. At the time of the work reported in [8], the aforementioned criteria left 814 disruptive pulses. These disruptive pulses of the database have been studied by experts at JET, to provide a classification of the disruption causes. The ones for which a solid classification was available have been used in the present work and form the so called “disruptive database”. To define and optimise the Fuzzy rules, other databases were created including temporal windows, which do not present disruptions and that can be considered safe. A database includes the same disruptive pulses but 2 seconds before the disruption takes place. Such an interval is considered adequate to guarantee the absence of precursors in the vast majority of the cases. This database is called “Pre-disruptive Safe”. Moreover, non disruptive pulses have been selected from the same campaigns. In this case the time window was chosen considering the time when the plasma reaches the X-point configuration and adding 7s, because this is the time phase when there is the highest probability of disruption in the disruptive database. This is called “Safe” database. Finally, two other sets of pulses were selected for testing the performance of the predictor, whose pulses were not used for developing the fuzzy inference system. They are called, respectively, “Disruptive Test” and “Safe test”. Altogether therefore five sets of pulses were used to devise the original static Fuzzy predictor described in [8]. They are summarised in Table 1 and fully described in [8].

4.2 DIAGNOSTIC SIGNALS

The diagnostic signals utilised have been chosen on the basis of previous experience on disruption prediction [3][4][5]. They present the following properties that make them suitable for being a disruption predictor input:

- a) none of the signals are post-processed and are therefore available in real time
- b) they refer to global measurements of the plasma parameters
- c) they can be normalized to obtain machine independent parameters, as they are calculated and measured almost in every tokamak

After careful observation of the signals from the various pulses, both disruptive and safe, it has been noticed that useful information can also be obtained from the trend that some of the diagnostic signals present during the time window. So, some of the diagnostic signals taken into account have been derived. These signals are the poloidal beta, the safety factor and the internal inductance.

Moreover, the radiated power is a useful indicator when it is close or overcomes the total input power.

So instead of using its absolute value, the difference $P_{\text{net}} = P_{\text{input}} - P_{\text{rad}}$ proves to be a better choice. In the end a total of 12 signals has been used as input to the static fuzzy predictor

4.3 THE STATIC FUZZY PREDICTOR

As reported in [8][16], a Fuzzy Inference System [16] was previously devised, to predict the proximity of a disruption at JET and to compare its performance with other methods. This predictor is called static because it was trained to work for whole discharges, without any allowance for different phases during the shot. It has 12 inputs, which are the ones described in the previous section, and one output, the probability of a disruption, in the range [0, 1], where 1 represent 100% probability of an incipient disruption. For every input to the Fuzzy Inference System 3, 4 or 5 membership functions were selected according to the experimental trend of the input signals. After a systematic analysis of the input signal behaviour for disruptive and non disruptive pulses in the database, 36 if-then rules were devised. In order to evaluate the performance of the designed predictor, a rule for discriminating between correct and wrong prediction was needed. Since the output of the fuzzy network can be interpreted as the probability of the specific configuration (in terms of input signals) to be disruptive, a shot can be considered disruptive if the output is above a certain threshold that has been chosen around 0.5 but that can be changed in order to reflect the “risk aversion” of the experimental team with regard to a certain plasma configuration. In order to prevent false detection in case of spurious transients in the plasma, that can be similar to pre-disruption phase, the condition of remaining above the threshold for at least 2 consecutives time steps or, in other words, at least 40ms was added.

All the details about the structure rules and performance of this Static Predictor are given in references [8].

5. CORRELATION ANALYSIS

In order to investigate the relative importance the signals given as input to the predictor, a feature selection analysis based on the Classification and Regression Trees has been performed. A brief description of the CART method is given in the following section 5.1; the details of its application to the problem subject of this work are provided in section 5.2.

5.1 CLASSIFICATION AND REGRESSION TREES ANALYSIS

Classification and Regression Trees [9] is a non-parametric statistical method, which uses a decision tree to solve classification and regression problems using both categorical and continuous variables. The method was introduced by Breiman et al. [9] to build a decision tree, which describes one output variable as a function of different explanatory variables. When the output is categorical, CART produces a classification tree, whereas if the output is continuous it will produce a regression tree. The analysis roughly consists of three steps.

During the first stage, an overgrown tree is produced using a recursive partitioning technique to select variables and split points using a splitting criterion. Several criteria are available for determining

the splits, including Gini, Twoing and Ordered Twoing. For a more detailed description of the mentioned criteria one can refer to Breiman et al. [9]. In the application described in this paper, the Gini criterion was adopted. According to this method, to find the best variable for splitting a node, the algorithm checks all possible splitting variables (called splitters), as well as all possible values of the variable to be used to split the node. To choose the best splitter, called also primary variable, the maximization of the average “purity” of the two child nodes is sought. For a binary target variable, for example being disruptive or not, the aim of the split is to group all the inputs into a group of disruptive discharges and another group without disruptive discharges. Since a complete separation is typically not achievable with one single variable, the procedure is repeated for the child nodes until pure terminal nodes are obtained (solutions exist to handle the cases in which the variables are not enough to reach a complete set of pure nodes). In addition to selecting the primary variables, surrogate variables, which are those variables that give in a specific node a similar split to the primary variable and provide the second best reduction in impurity in the child nodes, can also be identified and selected. They may be used in classifying observations having missing values for the primary variables.

The tree obtained at this stage is called maximal tree. It closely describes the training set and usually shows overfitting of the training data. So very often additional measures have to be taken to converge on a better a compromise between the tree complexity and its predictive power.

The second step in CART algorithm is, then, the pruning of the maximal tree, which results is the creation of smaller subtrees obtained by successively cutting terminal branches. The pruning relies on a cost-complexity method, in which both the tree accuracy and complexity are considered. This method relies on a complexity parameter, called α , which is gradually increased during the pruning process. Beginning from the terminal nodes, the child nodes are pruned away if the resulting change in the predicted misclassification cost (a measure of the accuracy) is less than α times the change in the tree complexity. Thus, α is a measure of how much accuracy a split must add to the entire tree to counterpart the additional complexity. As α is increased, a series of trees with decreasing complexity is obtained.

The last stage is the selection of the optimal tree among the various pruned trees derived during the previous step. This selection is based on the evaluation of the predictive error also called misclassification or relative cost. The goal consists of selecting the optimal tree maximizing the relative cost and minimizing the tree complexity (i.e. the number of nodes) so that the information in the learning data set is fit but not overfit. As the number of nodes increases, the relative cost decreases monotonically for the database used during the learning process. This corresponds to the fact that the maximal tree will always give the best fit to the learning data set. In contrast, the expected cost for an independent dataset reaches a minimum, and then increases as the complexity increases. This reflects the fact that an overfitted and overly complex tree will not perform well on a new set of data.

In general this step would need an independent set of data to be accessed, but this requirement can be avoided using the technique of Cross-Validation (CV). It consists of dividing the entire sample randomly into N (usually 10) sub-samples, stratified by the response variable. One sub-sample is then used as the test sample and the other $N - 1$ (e.g., nine) are used to construct a large tree. The entire

model-building procedure, i.e. tree growing and pruning, is repeated N times, with a different subset of the data reserved for use as the test dataset each time. Thus, N different models are produced, each one of which can be tested against an independent subset of the data. Trees within the sequences are matched up, based on their number of terminal nodes, to produce an estimate of the performance of the tree in predicting outcomes for a new independent dataset, as a function of the number of terminal nodes or complexity. Using this method, a minimum cost is obtained when the tree is complex enough to fit the information in the learning dataset, but not so complex that noise in the data is fit as well.

The final tree obtained has various nodes at which different variables are used for the splitting criteria. A variable that is not selected in the final tree could be considered as less important in describing the dataset than the variables that appears in the tree. However, it could happen that the variable is masked. For instance, a variable x_1 could be a surrogate of a variable x_2 , never occurring in this way as a primary splitter although is the best splitter after variable x_2 . CART allows the evaluation of the importance of the different explanatory variables to describe the output in the selected dataset through the so-called “variable ranking method”. Variable importance is the sum across all nodes in the tree of the improvement scores that the variable induces when it acts as a primary or surrogate splitter. In this way, variables that never appear in the tree being always a surrogate of another variable, are also considered in the final classification of the variables.

The importance values so produced allow ranking the different input signals from high to low importance. In this way, CART can be used for feature selection, being able to identify the most important variables to describe the output.

5.2 CART APPLICATION TO THE DISRUPTION PREDICTION PROBLEM

It is this last feature of CART that has been exploited to evaluate the relative importance of the different input signals in the disruption prediction process. In order to do so, the database used for the development and test of the static fuzzy predictor has been provided as input for the building of a classification tree. The samples from a disruptive pulse were labelled with a ‘1’ while the data from safe pulses as ‘0’. The data from the various databases were joined together and, subsequently, separated in time intervals of 100ms producing a total of three datasets (see later). From the full data, which comprises a total of 400ms sampled at 20 ms, the last three samples were discarded as considered too near to the disruption and, also, in order to produce a result that is easily comparable with the other predictors, whose performance are evaluated at 100ms from the disruption. Each dataset has been used to build a classification tree and to evaluate the variable importance in each time interval.

Table 3 reports the variable ranking for the time intervals $[t_d-440, t_d-100]$ (the full data set), $[t_d-440, t_d-340]$, $[t_d-320, t_d-220]$ and $[t_d-200, t_d-100]$ which represent the three sub-datasets.

It is possible to observe how the variable ranking changes depending on the distance from the disruption. For instance, the dW_{dia}/dt is very important in discriminating a disruptive from a safe pulse near to the disruption, whereas it plays a less important role earlier in the discharge. The CART method therefore indicates very clearly that the various input signals have a different importance as

indicators of a disruption, depending on the time interval considered. It is worth mentioning that the correlation technique presented, which performs an exhaustive search for the best univariate splits over all the variables of the data base, is fully non linear and unbiased. The relevance of the CART results on the optimisation of disruption predictors is considered in the next section.

6. THE PROTOTYPE OF AN ADAPTIVE FUZZY PREDICTOR

Following the results illustrated in the previous section, an additional predictor has been devised from the classification tree built for the input signals in the interval $[t_d-440, t_d-340]$. This new fuzzy predictor, called “earlier predictor” in the following, has been developed starting from the original one, which had been optimised for the whole discharge. The new optimization has been obtained by performing an iterative adaptation to the subset of the data pertaining to the interval $[t_d-440, t_d-340]$. First of all, some rules have been rescaled according to the variable ranking information from the CART. In particular, the rules involving only variables that have a relative importance below 50 have been multiplied by a factor 0.5. Some of them, as they do not have a significant impact after rescaling, have been deleted in order to reduce the complexity of the predictor. Then, to increase further its performance, additional rules have been developed, interpreting the classification tree and the threshold values derived from the primary splitters. This optimisation process has converged on a final fuzzy network with a total of 26 rules.

This “earlier predictor” has been tested on the data samples distant from the disruption and its output compared with the fuzzy network developed for the full discharge. Table 4 reports the comparison of the two predictors. A shot is considered disruptive if the output of the network remains above the threshold of 0.45 for more than two time steps [8]. The results of the table show very clearly that, with the new partial predictor, at 340 ms from the disruption the correct prediction rate has increased and, even if the number of false alarms grows, the overall performance is significantly improved.

The performance improvements of the “earlier predictor” suggest alternative strategies for the implementation of disruption predictors. Instead of developing static predictors, valid for the whole discharge and therefore for any value of disruption probability, it is conceivable to devise “adaptive predictors”, namely fuzzy nets optimised for different disruption probabilities. In this way, a predictor could be fine tuned according to the disruption probability and, therefore, the imminence of disruption. A possible strategy could consist of identifying different predictors optimised for different disruption probabilities and switch from one to the other depending on the plasma status. To illustrate the potential of such an approach, a combined use of the “earlier” and the static predictors has been tested. The strategy for the simultaneous operation of the two predictors has been organised starting always with the “earlier” predictor and switching to the static predictor when the output of the first network is above a primary threshold for two time steps. Once the switching occurred, the shot is considered disruptive if the output of the static predictor is above a secondary threshold just for one sample. The primary threshold has been

chosen smaller than the secondary one and equal to 0.4. The secondary thresholds, to compare the results with those obtained adopting only the static predictor, are 0.45 and 0.51.

Table 5 and Table 6 report the comparison between the new adaptive predictor here illustrated and the static one using the threshold of 0.51 and 0.45 respectively. The overall performance of the predictor is improved using the new approach as regards both the number of Missed Alarms (MA) and the number of false alarms (FA).

CONCLUSIONS

In this work, a critical revision of a previous static fuzzy predictor for JET disruptions has been reported. The correlation analysis of the inputs proves the limitations of a static approach. At different levels of disruption probability, and therefore at different times before a disruption, the same diagnostic signal can carry information of significantly different importance and therefore must be weighted accordingly. Consequently, a prototypical adaptive predictor has been developed, which consists of a set of different Fuzzy systems optimised for different disruption probabilities. The various Fuzzy systems can be used when the monitored signals enter the relative subspace of disruption probability. This new adaptive predictor provides a tool to define a more optimised and better performing prediction strategy than the previous static one. On the other hand, it maintains all the advantages of a white box approach and therefore can also be systematically used to complement the analysis of the experts in the study of disruption physics.

The present work must be interpreted as a first step because additional refinements could be easily implemented. First of all, more expert knowledge could be inserted in the rules, which were devised more on the basis of the database contents than mathematical modelling of the underlying physics. Moreover, the results obtained with the prototype of an adaptive Fuzzy predictor open the way to a more systematic investigation of adaptive strategies for disruption prediction at JET and elsewhere. A more sophisticated method, than the preliminary one tested so far, could utilise different fuzzy nets, each one optimised for a specific time interval of the training data set. A preliminary analysis seems to indicate that intervals of the order of 100 ms would be adequate far from the disruption, whereas shorter intervals could be beneficial in the last time period. These various nets could be run in parallel and would provide the probability of disruption for their respective time interval in the future (for example a predictor optimised for the interval 500-400ms before the disruption would give the probability of disruption for this interval in the future). A global algorithmic controller could then manage the predictions of these nets and issue an alarm when considered appropriate. A set of thresholds for the various predictors could be identified, to find the best trade-off between machine integrity and the scenario development needs. To conclude, it is worth mentioning that also other approaches, based on ANNs [5] or Support Vector Machines [17], could also be implemented in an adaptive way and their performance compared, to converge on the most suited technique for the next generation of devices.

REFERENCES

- [1] F.C. Schuller , Disruption in tokamaks, Plasma Phys. Control Fusion **37** (1995), A135-A162
- [2] D. Wroblewski et al., “Tokamak disruption alarm based on a neural network model of the high- β limit”, Nucl. Fusion, vol. **37**, no. 6, 1997, pp. 725-740.
- [3] G. Pautasso et al., “On-line prediction and mitigation of disruptions in ASDEX Upgrade”, Nucl. Fusion, vol. **42**, pp.100-108, 2002.
- [4] Milani F., “Disruption prediction in JET”, PhD Thesis, Univ. of Aston, Birmingham (1998)
- [5] B. Cannas et al., “Disruption forecasting at JET using neural networks”, Nucl. Fusion **44**, pp. 68-76, 2004.
- [6] F.C. Morabito and M. Versaci, “A fuzzy neural approach to plasma disruption prediction in tokamak reactors”, Int. Joint Conf. on Neural Network, Washington, D.C., 1999.
- [7] M. Versaci and F.C. Morabito, “Fuzzy time series approach for disruption prediction in tokamak reactors”, IEEE Trans. On Magnetism, vol. **39**, no. 3, pp. 1503-1506, May 2003.
- [8] G. Vagliasindi et al., “Fuzzy Logic Approach to Disruption Prediction at JET”, submitted to IEEE Trans. on Plasma Science.
- [9] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, “Classification and Regression Trees” Chapman & Hall (Wadsworth, Inc.): New York, 1984.
- [10] J.A. Wesson, “Tokamaks”, Clarendon Press, 2004.
- [11] J.A. Wesson et al., “Disruptions in JET”, Nucl. Fusion, vol. **29**, n. 4, pp.641-666, 1989.
- [12] R.G. Kleva, J.F. Drake, “Density limit disruptions in tokamaks”, Phys. Fluids B 3 (2), February 1991
- [13] B. Lipschultz, “MARFE: An edge plasma phenomenon,” Nucl. Fusion, vol. **24**, no. 8, pp. 977–89, 1984.
- [14] L.A. Zadeh, “Fuzzy sets”, Infor. Control, 8, pp.338-353,1965.
- [15] B. Kosko, Neural Network and Fuzzy Systems, Prentice-Hall, Englewood Cliffs, NJ(1992)
- [16] Fuzzy Logic Toolbox User’s Guide, Version 2, The Mathworks, Inc., 2002.
- [17] B. Cannas et al., “Support Vector Machines for Disruption Prediction and Novelty Detection at JET”, Proc. of 24th Symposium on Fusion Technology (SOFT ‘06), Warsaw, Poland, September 11-15, 2006

Databases	Number of pulses
Disruptive	164
Disruptive test	128
Safe	84
Pre-disruptive Safe	164
Safe test	137

Table 1: list of the databases used for the analysis

Signal name	Unit
Plasma current I_{pla}	[A]
Mode Lock Amplitude $Loca$	[T]
Plasma density $Dens$	[m ⁻³]
Total Input Power P_{inp}	[W]
Plasma Internal Inductance L_i	
Stored Diamag. Energy Derivative dW_{dia}/dt	[W]
Safety factor at 95% of minor radius q_{95}	
Poloidal beta β_p	
Safety factor derivative dq_{95}/dt	[s ⁻¹]
Plasma Internal Inductance derivative dL_i/dt	[s ⁻¹]
Poloidal Beta derivative $d\beta_p/dt$	[s ⁻¹]
Net power P_{net}	[W]

Table 2: list of the signals used as input of the fuzzy predictor.

[td-440, td-100]		[td-200, td-100]		[td-320, td-220]		[td-440, td-340]	
Variable	Importance	Variable	Importance	Variable	Importance	Variable	Importance
dW_{dia}/dt	100,00	dW_{dia}/dt	100,00	dl_i/dt	100,00	$Ipla$	100,00
dl_i/dt	82,88	dl_i/dt	50,46	dW_{dia}/dt	99,20	dl_i/dt	69,42
$Ipla$	70,88	$d\beta_p/dt$	37,01	$Ipla$	79,40	P_{net}	68,23
P_{net}	67,96	P_{net}	35,03	l_i	70,47	l_i	54,78
dq_{95}/dt	54,90	q_{95}	27,54	q_{95}	62,06	dq_{95}/dt	53,92
$Dens$	53,24	\leq_p	24,80	P_{net}	59,00	P_{inp}	49,58
$d\beta_p/dt$	52,96	l_i	24,37	β_p	57,05	dW_{dia}/dt	38,78
l_i	52,36	$Loca$	22,74	dq_{95}/dt	56,19	Q_{95}	37,97
$Loca$	46,75	P_{inp}	20,49	$d\leq_p/dt$	38,68	$Loca$	37,39
P_{inp}	45,26	$Ipla$	16,36	P_{inp}	38,41	$d\leq_p/dt$	36,57
β_p	43,47	$Dens$	13,34	$Dens$	37,91	$Dens$	35,55
q_{95}	33,55	dq_{95}/dt	9,87	$Loca$	30,13	β_p	28,77

Table 3: Ranking of the various signals as calculated by CART for the time intervals $[t_d-440, t_d-100]$ (the full data set), $[t_d-440, t_d-340]$, $[t_d-320, t_d-220]$ and $[t_d-200, t_d-100]$.

	Missed Alarms (MA)	False Alarms (FA)	Correct Prediction
“Earlier” predictor			
Shots	83 / 292 (28%)	52 / 221 (23.5%)	378 / 513 (74%)
Static predictor			
Shots	203 / 292 (70%)	7 / 221 (3.1%)	303 / 513 (59%)

Table 4: comparison between the “earlier predictor” and the “static predictor”. The first column reports the total number of missed alarms while evaluating disruptive pulses in the datasets; the second column reports the total number of false alarms in safe pulses; the last column contains the total number of correct predictions, i.e. pulses identified as disruptive in disruptive databases and pulses identified as non-disruptive in “safe” databases.

Adaptive predictor		Static predictor	
<i>Database set</i>	<i>MA/TOT</i>	<i>Database set</i>	<i>MA/TOT</i>
Disruptive	7 / 164 (4.3%)	Disruptive	16 / 164 (9.7%)
Disruptive test	19 / 128 (14.8%)	Disruptive test	18 / 128 (14.0%)
Total	26 / 292 (8.9%)	Total	34 / 292 (11.6%)
<i>Database set</i>	<i>FA/TOT</i>	<i>Database set</i>	<i>FA/TOT</i>
Safe	8 / 84 (9.5%)	Safe	9 / 84 (10.7%)
Pre-disruptive Safe	16 / 164 (9.7%)	Pre-disruptive Safe	15 / 164 (9.1%)
Safe test	3 / 137 (2.2%)	Safe test	11 / 137 (8.7%)
Total	27 / 385 (7.0%)	Total	31 / 385 (9.1%)

Table 5: comparison of the predictors' output using a secondary threshold of 0.51

Adaptive predictor		Static predictor	
<i>Database set</i>	<i>MA/TOT</i>	<i>Database set</i>	<i>MA/TOT</i>
Disruptive	3 / 164 (1.8%)	Disruptive	9 / 164 (5.5%)
Disruptive test	12 / 128 (9.4%)	Disruptive test	14 / 128 (10.9%)
Total	15 / 292 (5.1%)	Total	23 / 292 (7.9%)
<i>Database set</i>	<i>FA/TOT</i>	<i>Database set</i>	<i>FA/TOT</i>
Safe	11 / 84 (13.1%)	Safe	11 / 84 (13.1%)
Pre-disruptive Safe	29 / 164 (17.7%)	Pre-disruptive Safe	30 / 164 (18.3%)
Safe test	6 / 137 (4.3%)	Safe test	27 / 137 (19.7%)
Total	46 / 385 (11.9%)	Total	68 / 385 (17.7%)

Table 6: comparison of the predictors' output using as secondary threshold 0.45

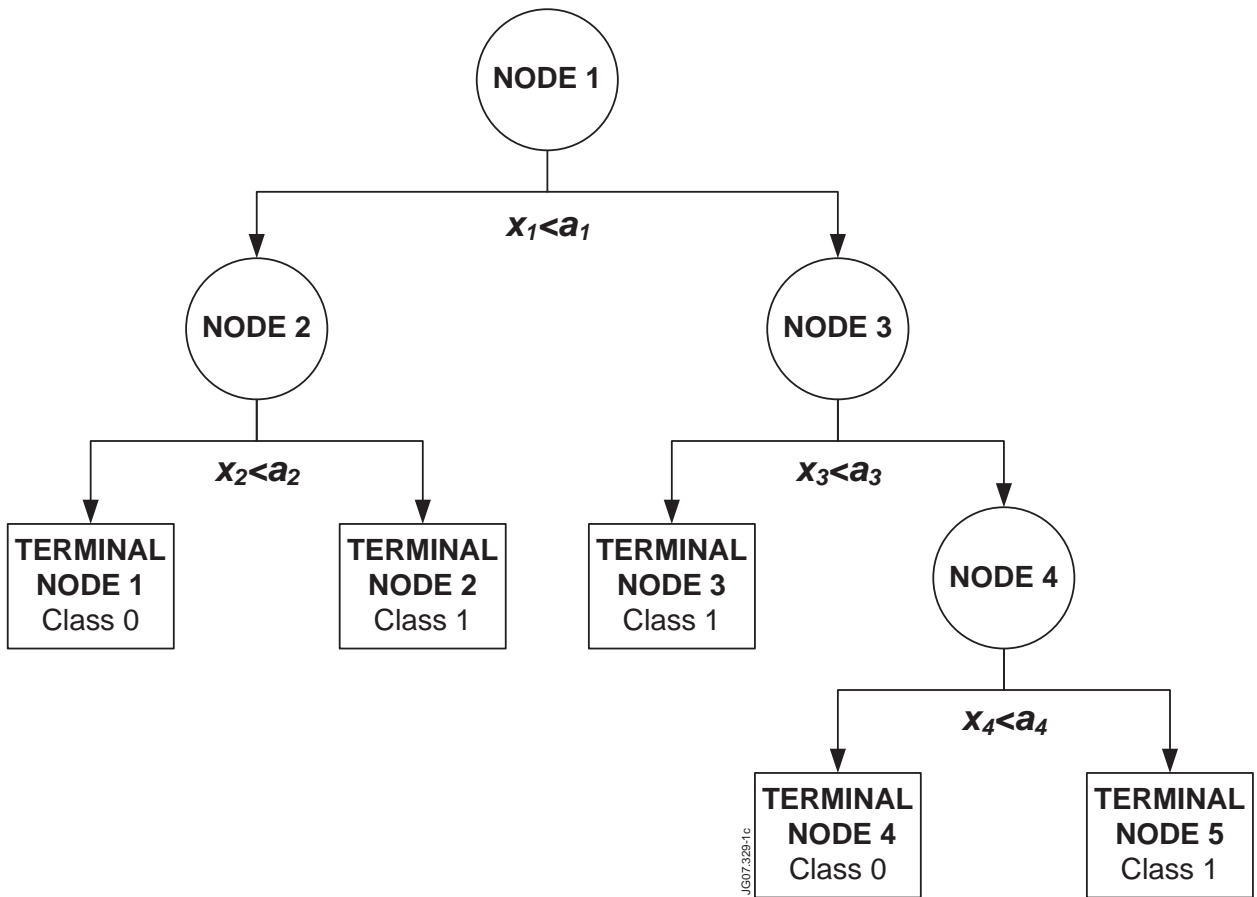


Figure 1: General structure of a classification tree for a binary target variable: 0 and 1. x_i is the selected split variable (splitter) and a_i is the selected split value.