A linear equation based on signal increments to predict disruptive behaviours and the time to disruption on JET

J. Vega[1], A. Murari[2], S. Dormido-Canto[3], F. Hernández[3], T. Cruz[3],

D. Gadariya[1], G. A. Rattá[1] and JET Contributors*

EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK

[1]Laboratorio Nacional de Fusión, CIEMAT, Madrid, Spain

[2]Consorzio RFX (CNR, ENEA, INFN, Universitá di Padova, Acciaierie Venete SpA), Corso Stati Uniti 4, 35127 Padova, Italy

[3]Dpto. Informática y Automática - UNED, Madrid, Spain

* See the author list of E. Joffrin et al. accepted for publication in Nuclear Fusion Special issue 2019, https://doi.org/10.1088/1741-4326/ab2276

Abstract

This article describes the development of a generic disruption predictor that is also used as basic system to provide an estimation of the time to disruption at the alarm times. The mode lock signal normalised to the plasma current is used as input feature. The recognition of disruptive/non-disruptive behaviours is not based on a simple threshold of this quantity but on the evolution of the amplitudes between consecutive samples taken periodically. The separation frontier between plasma behaviours (disruptive/non-disruptive) is linear in such parameter space. The percentages of recognised and false alarms are respectively 98% and 4%. The recognised alarms can be split into valid alarms (90%) and late detections (8%). The experimental distribution of warning times follows an exponential model with average warning time of 443 ms. On the other hand, the prediction of the time to the disruption has been fitted to a Weibull model that relates this predicted time to the distance of the points to the diagonal in the parameter space of consecutive samples. The model shows a very good agreement between predicted times and warning times in narrow time intervals (between 0.01 s and 0.06 s) before the disruption.

1. Introduction

Nowadays, important efforts are exerted to develop disruption avoidance methods [1-3]. In general, avoidance techniques should detect that the plasma is close to

a stability limit and, therefore, that the probability of disruption is not negligible. Hence, actions from the tokamak control system are required to steer the plasma to safe conditions. Generally speaking and thinking of ITER operation, '*to steer the plasma to safe conditions*' means either to continue the discharge with different plasma parameters (for example, density, heating power, plasma current or magnetic field) or to terminate the discharge in a controlled way.

However, in the presence of an impending disruption, avoidance actions cannot be applied and, therefore, mitigation methods are mandatory. These mitigation methods can be aimed at reducing forces, at alleviating heat loads during the thermal quench and at avoiding runaway electrons. Examples of amelioration measures can be the injection of a significant amount of gases through fast valves [4-5], killer pellets [6-8] or Electron Cyclotron Resonance Heating injection [9-10]. It is important to emphasise that disruption mitigation techniques crucially depend on reliable disruption predictors, whose alarms have to be triggered with enough anticipation time (or warning time) to the disruptive event.

The objective of this article is twofold. A first goal is to show a simple method to develop a generic disruption predictor without the complexity of traditional machine learning models (for example, Support Vector Machines (SVM), artificial neural networks, regression trees or random forests) and whose outputs have a direct interpretation. The predictor is not based on amplitude thresholds but on modelling the signal increments between consecutive samples of relevant signals. Modelling the variations of amplitudes between successive samples is a novel approach to the disruption prediction problem. This article shows that a simple linear equation in two variables is powerful enough to predict disruptions in JET.

A second goal is to provide an estimation of the time to the disruption (TTD) at alarm times, by means of the above predictor and without the need of black box models deduced from machine learning methods.

Section 2 summarises general concepts of disruption prediction. Sections 3 and 4 review disruption predictors and predictors of the time to disruption respectively. Section 5 shows the basis of the present predictor to recognise disruptive behaviours and section 6 describes the particular parameter space that has been used in this article. Section 7 summarizes the training process and section 8 presents and discusses prediction rates in JET. Section 9 implements the predictions of the time to the disruption and analyses its results. Finally, section 10 contains a discussion about the methods shown in the paper.

2. Disruption prediction concepts and definitions

So far, neither theory deduced from first principles nor empirical physics models are able to completely explain the causes and variety of disruptive events. However, at present, big efforts are being devoted to developing physics models of disruptions. This

is a difficult task and only partial models (typically based on strong assumptions and unphysical boundary conditions [11-12]) are available.

An alternative to physics models for disruption prediction is the use of machine learning methods. Machine learning methods allow the creation of binary classifiers in multi-dimensional feature spaces to determine the plasma status (disruptive/non-disruptive) at any time. A training dataset, with examples (feature vectors) of both disruptive and non-disruptive behaviours, allows determining the separation frontier (i.e. an equation) between disruptive and non-disruptive states (fig. 1). To this end, an underlying machine learning method (for example, SVM, neural networks, fuzzy logic, Bayesian classifiers or deep learning) is required. The objective of these equations is to derive data-driven models to explain and predict disruptive events directly from the available empirical evidence. However, so far, equations with direct physics interpretation (*i.e.* simple form equations) are not reliable enough (*i.e.* success and false alarm rates are not close to 100% and 0% respectively) to cope with the description and prediction of disruptions. In present applications, the equations derived from general machine learning methods provide non-trivial relationships (complex equations) between several signals to describe the separation frontier. Although these complex equations are not suitable to interpret the physics causes of the onset of disruptive events, however, they are very useful to forecast imminent disruptions. In this way, during the discharges, feature vectors are used on periodic basis as inputs to the classifiers to determine in which part of the separation frontier the plasma is (fig. 1).

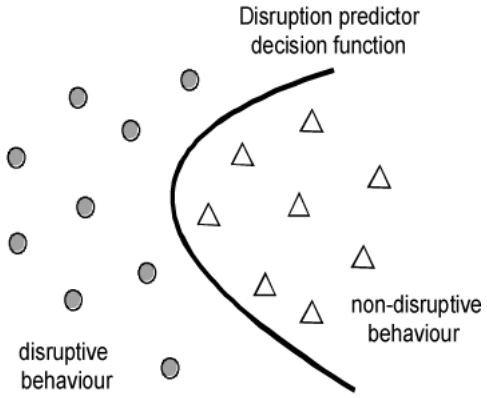

Fig. 1: The separation frontier (or decision function) splits the operational space into two zones: disruptive and non-disruptive. The location of feature vectors at each time determines the plasma state.

The assessment of predictors performance is carried out through three indicators that are obtained from a dataset of discharges different from the one used for the training process. The first output is the success rate (its complementary output is the missed alarm rate). The success rate is the fraction of disruptive discharges successfully recognised by the predictor over the total number of disruptive discharges in the test dataset. The second output is the false alarm rate. It is computed as the number of non-disruptive shots that are recognised as disruptive ones divided by the number of non-disruptive discharges in the test set. The third output is the distribution of warning times obtained from the test dataset. For test purposes, the warning time of a discharge is

defined as the difference between the disruption time and the alarm time of the predictor. With this definition, positive warning times correspond to disruptive behaviours that are recognised with anticipation. However, a negative warning time means that the predictor identifies a disruptive condition after the disruption (tardy detection).

In principle, disruption predictors could be trained to satisfy either avoidance or mitigation requirements. Typically, avoidance actions require longer anticipation times than mitigation methods. Therefore, in general, predictors for avoidance would have to provide warning time distributions that are peaked far from the disruption times. Fig. 2 shows the conceptual difference between predictors for avoidance (fig. 2a) and mitigation (fig. 2b) in terms of the warning time distributions. The black line at warning time 0 is the disruption time reference of the discharges, *i.e.* an alarm with warning time equal 0 means that the alarm has been triggered at the disruption time.

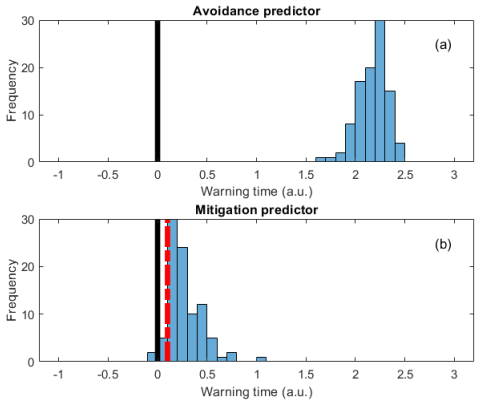

Fig. 2: Conceptual view of warning time distributions for avoidance (a) and mitigation predictors (b). The red dashed line represents the minimum warning time to apply a specific mitigation technique.

Focusing the attention on fig. 2b, another important time is shown (red dashed line to the right of warning time equal 0). For real-time applications, it should be emphasised that mitigation alarms are only useful if the anticipation time is greater than the characteristic reaction time of a particular technique. Due to this fact, the success rate can be a biased indicator to qualify a predictor. The term ‘success rate’ includes three elements: disruptions recognised with negative warning times (that have been called tardy detections), disruptions predicted with a positive warning time but shorter than the reaction time required for a successful mitigation action and, finally, disruptions identified with a larger warning time than the reaction time needed for a successful mitigation action. The fraction of disruptions recognised with warning times (either positive or negative ones) less than the reaction time will be called ‘late detections’. The fraction of disruptions whose warning time is greater than the reaction time will be called ‘valid alarms’.

For example, the time in JET to trigger the disruption mitigation valve (DMV) is 10 ms and this is the reaction time that has been considered in this article. However, the cumulative fraction of detected disruptions with positive warning times (CUMUL0) is

also important. This qualifier gives an idea about the potential increase of valid alarms with the availability of faster mitigation systems, i.e. mitigation systems with characteristic times less than 10 ms in the present case.

It is important to note that the performance of avoidance and mitigation predictors strongly depends on the patterns used as features to train the classifiers. Avoidance predictors are much more difficult to develop because early precursors of potential disruptions also can appear in non-disruptive behaviours and they can generate high rates of false alarms.

3. Review of disruption predictors

In view of the difficulties to devise physics models for disruption prediction, a practical approach to protect tokamaks from disruptions has been the use of the mode lock (*ML*) signal as disruption predictor. During operation, magneto-hydrodynamic (MHD) instabilities can be locked to the wall and, therefore, the amplitude of the *ML* signal can increase. This raise is produced by two facts: the deceleration of the rotating mode and/or the growth of the MHD mode. In a running experiment, usually, if the *ML* amplitude crosses a certain threshold, (which is set-up prior to each discharge), an alarm is triggered to warn about an incoming disruption. In general, this threshold is selected in a manual way (within certain limits determined by experience) and it is chosen depending on the characteristics of the experimental programme: the *ML* threshold is set lower or higher depending on the potential danger of the possible disruptions.

Disruption predictors based on the *ML* signal have been universally accepted in the nuclear fusion community due to its direct relation to MHD activity [13-14]. In JET, the *ML* signal is calculated as the n = 1 combination of saddle loop signals. More specifically, JET uses the *ML* signal normalised to plasma current (*ML*/*Ip*) as predictor. However, an optimal threshold (either *ML* or *ML*/*Ip* based) to trigger alarms cannot be determined. The manual selection of multiple thresholds in the execution of discharges has triggered alarms that do not correspond to real disruptive events (actually, they are false alarms). On the other hand, there are discharges in which the threshold was set-up excessively high and, therefore, disruptions occurred (missed alarms). However, these missed alarms cannot be used to determine a good threshold because there are non-disruptive discharges whose maximum amplitudes are higher than amplitudes in disruptive discharges. This is a clear symptom that thresholds in the *ML* or *ML*/*Ip* signals are not optimal disruption predictors.

As a consequence, historically, the simple *ML* (or *ML*/*Ip*) threshold has not been enough to achieve success rates close to 100%. The reason for this, apart from the problems associated to the lack of an optimal threshold, resides in the fact that not all disruptions show a non-rotating mode close to the disruptive event (for example, disruptions linked to vertical displacement events). To improve disruption prediction, general machine learning methods have been used with, typically, several of the following signals: *ML*, plasma current, poloidal beta, poloidal beta time derivative, safety factor, safety factor time derivative, total input power, plasma internal

inductance, plasma internal inductance time derivative, plasma vertical centroid position, plasma density, stored diamagnetic energy time derivative and net power [15]. These quantities have been combined in multiple ways (to form various multi-dimensional spaces) with different underlying machine learning methods: SVM [16], fuzzy logic [17] or Artificial Neural Networks [18-19].

The use of multi-dimensional spaces with amplitudes exclusively in the time domain is outperformed by the use of features either exclusively in the frequency domain [15, 20-21] or with combinations of both domains [22-24]. Again, in these cases, the resulting predictors are very useful to recognise incoming disruptions with enough anticipation time (on average, hundreds of ms in JET). However, the physics interpretation of the models implemented by these predictors are not clear at all.

Last but not least, it is worth to mention a recent disruption predictor that uses deep learning techniques [25]. Deep learning uses neural network architectures and it is a technology that is being developing very fast during the last years. The predictor in [25] requires a very powerful computational environment for training purposes.

4. Review of methods to predict the time to the disruption

The TTD prediction is a difficult problem with no practical solution so far. There have been few attempts to deal with this but all of them use, first, a complex structure of artificial neural networks for the prediction, second, a limited number of discharges and, third, particular types of disruptions. The first point prevents the possibility of identifying physics reasons and the other two imply that these predictors cannot be generalised to a large number of disruption types. A brief description of three TTD predictors follows.

A first work [26] related to TTD prediction used a fuzzy framework only to achieve a suitable clustering of the input space. The proposed predictor defined a complex structure of neural networks (NN). A first processing layer based on the radial basis function NN scheme was basically used to decompose the original database into four subsets. The layer outputs were used to activate four multilayer perceptrons, trained exclusively on a subset of the original database. The output of the system was a single node linear layer that provided the estimated time to disruption as well as an alarm when appropriated. The predictor generated a temporal evolution signal (time period of 2.5 ms) and the alarm was triggered when the time to disruption was below a certain threshold. An on-line test of the predictor was carried out in open loop, but the total number of discharges is not specified in [26]. In these tests, an alarm was triggered when the TTD was maintained in the range 250 ms – 350 ms for ten consecutive predictions (to avoid false alarms).

A second work is a proposal from Pautasso et al. [27] that is based on a two-layer artificial neural network. The first layer contains 20 neurons and the network output (second layer) provides the time interval to the disruption. Therefore, as in [26], the predictor is a temporal evolution signal whose value at each time instant is the time

to disruption. To filter false positives, an alarm is triggered when the predictor output is less than or equal to 50 ms during three consecutive samples (sampling period is 2.5 ms). The predictor was tested on-line in open loop with 128 discharges (28 of them were disruptions). The success and the false alarm rates with reduced types of disruptions were 79% (22/28) and 7% (7/100) respectively.

A third work about predicting the time to disruption [28] was applied to the JET database and was also based on artificial neural networks. The best network configuration was composed of nine inputs, two hidden layers with six and five hidden neurons respectively, and one output. Only signals available in real-time were taken into account (this means that data relying on off-line equilibrium reconstruction or off-line processed data were not used). The neural network output was a real number between 0 and 1 representing the risk of disruption, which is obtained every 20 ms. An alarm was triggered when the neural network output is above a certain threshold that was chosen minimizing a detection error function. The predictor was tested with 62 and 132 disruptive and non-disruptive shots respectively. The neural network output was analysed in a time interval between 440 ms and 100 ms before the disruption. A success rate of 83.9% was reported.

From these brief descriptions, the complexity of the potential solutions and also the lack of physics basis in the developed TTD predictors are more than evident.

5. Rationale of a generic disruption predictor based on centroids

In view of the next DT campaign in JET and of the future operation of ITER and DEMO, disruption predictors need to be implemented as simple equations with direct physics meaning. This section describes a simple and general method to develop predictors based on a very general principle.

Let's consider a multi-dimensional parameter space $\mathbb{S} \subset \mathbb{R}^m$ where each dimension is a physics quantity. The general method proposed here to develop generic predictors can be summarised in two simple rules:

1. Both the *disruptive* and *non-disruptive* physics knowledge in $\mathbb{S}$ can be compressed into two single points C_D and C_N respectively (fig. 3).
2. At any time t in a running discharge, the plasma state is represented by a point $P(x_1, x_2, ..., x_m) \in \mathbb{S}$ (fig. 3). If the Euclidean distance from P to C_D $\left(d_{P,C_D}\right)$ is shorter than the Euclidean distance from P to C_N $\left(d_{P,C_N}\right)$, the plasma is recognised as being in a '*disruptive* state'. Otherwise, the plasma is in a '*non-disruptive* state'. It should be noted that this reasoning is consistent from a physics point of view: the plasma state is associated to the closest behaviour summarised by the two points C_D and C_N.

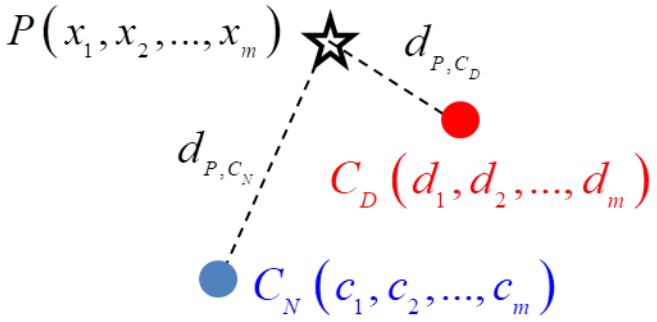

Fig. 3: A general predictor can be based on the well-known nearest centroid technique. From *m* physics quantities, C_D and C_N are centroids, which condense the relevant information about disruptive and non-disruptive behaviours respectively into two points in $\mathbb{S}$.

The points $C_D(d_1, d_2, ..., d_m)$ and $C_N(c_1, c_2, ..., c_m)$ are the centroids of a dataset of disruptive and non-disruptive examples in the parameter space $\mathbb{S}$ respectively. In other words, given a training dataset made up of both disruptive examples $\Delta_i(\delta_{1,i}, \delta_{2,i}, ..., \delta_{m,i}) \in \mathbb{S}, i = 1, ..., n_D$ and non-disruptive examples $\mathrm{X}_i(\chi_{1,i}, \chi_{2,i}, ..., \chi_{m,i}) \in \mathbb{S}, i = 1, ...n_N$, the respective coordinates of C_D and C_N are

$$(d_1, d_2, ..., d_m) = mean(\delta_{1,i}, \delta_{2,i}, ..., \delta_{m,i}), i = 1, ..., n_D$$

and

$$(c_1, c_2, ..., c_m) = mean(\chi_{1,i}, \chi_{2,i}, ..., \chi_{m,i}), i = 1, ..., n_N$$

where $mean(\circ)$ returns a vector containing the average value of each column $\delta_{j,i}$ or $\chi_{j,i}, j = 1, ..., m$.

This general predictor assigns a disruptive behaviour to point $P(x_1, x_2, ..., x_m)$ when

$$d_{P,C_D} < d_{P,C_N}$$

that can be written explicitly as

$$\sqrt{\sum_{i=1}^{m}(x_i - d_i)^2} < \sqrt{\sum_{i=1}^{m}(x_i - c_i)^2}$$

By simple algebraic manipulations, the condition of disruptive behaviour can be expressed

$$2\sum_{i=1}^{m}(d_i - c_i) \cdot x_i > \sum_{i=1}^{m}(d_i^2 - c_i^2) \qquad (1)$$

Taking into account that the centroid coordinates are fixed for each given training set, equation (1) is linear in the physics quantities x_i:

$$\sum_{i=1}^{m} A_i \cdot x_i > K \tag{2}$$

where $A_i, i = 1,...,m$ and K are constants that are determined in the computation process of the centroids. Equation (2) summarises the first result of this article: the formulation of a very general disruption predictor that can be implemented in an easy way.

To compare the predictor represented by equation (2) with other machine learning predictors, five points have to be emphasised:

a. Equation (2) has been deduced from a very simple criterion (the shortest distance to a centroid) without any other kind of hypothesis. No assumptions about layers, neurons, kernels, tree branches etc. have been required. In addition, no time consuming optimization algorithms are necessary for training. Moreover, new re-trainings (centroid computations) can be carried out in a fast way whenever new training examples have to be added.
b. Equation (2) is linear in the physics quantities, which facilitates the physics interpretation.
c. Equation (2) allows a very simple recognition of disruptive behaviours given the coordinates of point $P(x_1, x_2, ..., x_m)$ at any time instant.
d. Due to the general validity of the equation, it could be used for both avoidance and mitigation. The distinction between both applications resides in the precursor capability of the quantities used in the development of the predictor. Of course, multiple predictors (either for avoidance or mitigation) can be obtained by simply changing the physics quantities to form the parameter space $\mathbb{S}$.
e. The prediction is accomplished through a linear inequality, which is very efficient from a computational point of view, even in the case of large dimensions (large *m*). Therefore, the implementation under real-time requirements is not a problem (for example, JET real-time network has a characteristic time of 2 ms and this time is sufficiently long to make predictions with equation (2)).

6. Application to JET of the centroid method for mitigation purposes

This section shows a specific implementation of a generic disruption predictor for JET based on equation (2). To accomplish this implementation, several considerations have been taken into account:

- Only quantities in the time domain have to be used in order to make easier the physics interpretation of the predictions. This prevents one from using features in the frequency domain which, in addition to having a more difficult meaning, also impose the need of data processing in time windows with a minimum number of samples.
- So far, linear predictors based on general purpose machine learning methods have not achieved high success rates (> 95% is required for ITER) and low false alarm rates (<5 % for ITER) in a simultaneous way. However, the challenge is to find the simplest linear predictor based on the centroid method that is able to approach ITER requirements.

- The use of equation (2) in one-dimensional spaces would be connected to the existence of simple thresholds in individual signals to identify disruptive conditions. However, experience tells that this is not optimal as mentioned in section 3. Therefore, the simplest predictor not linked to a threshold that is based on equation (2) defines a separating frontier between disruptive/non-disruptive behaviours of the form: $A_1 \cdot x_1 + A_2 \cdot x_2 = K$. This means that the corresponding parameter space $\mathbb{S}$ has dimension 2.
- The developed predictor has been optimised to recognise disruptive behaviours close to the disruptions (mitigation purposes).
- The ML signal is one of the most relevant quantities in JET to create predictors for mitigation. On the one hand, it is known that most of the JET disruptions generate a ML previous to the disruption [1]. On the other hand, the lack of the *ML* signal produces very bad predictions in JET [29]. Therefore, the *ML* signal will be present in the application of the present predictor.

The above considerations motivate the development of a predictor in a two-dimensional space and the use of the *ML* signal. So, next step would be the search for a second signal to be used together with the *ML*. However, if possible, the fewer signals the simpler the predictor. Therefore, the idea is to use only the *ML* signal but after mapping it into a two-dimensional feature space. This mapping is an evolution of a recent approach focused on anomaly detection to recognise disruptive events [30-32].

In the case of the present predictor (that is based on centroid computations), and assuming that the *ML* signal has a sampling period τ, the two-dimensional parameter space is defined by the amplitudes of consecutive samples of the *ML* signal. Fig. 4 is an example of how the amplitudes of consecutive samples define points in this feature space.

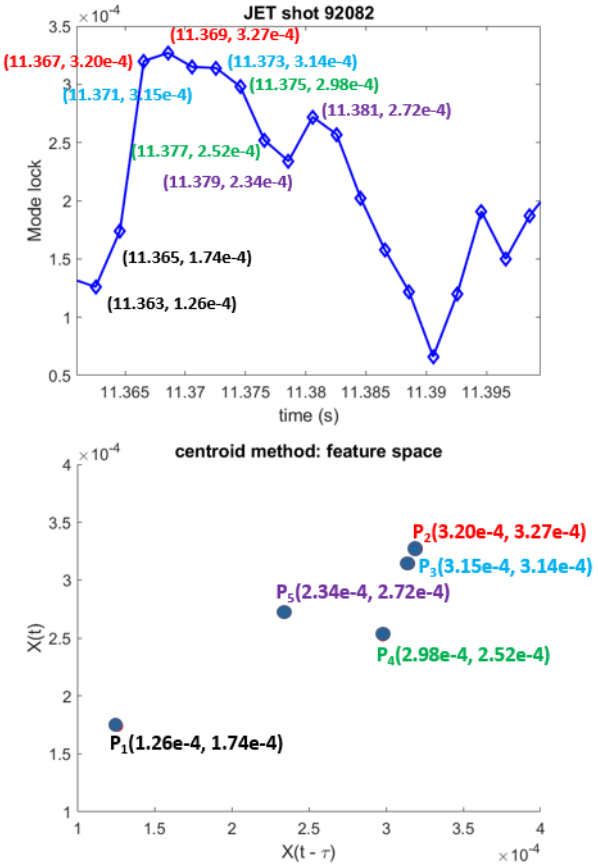

Fig. 4: *ML* amplitudes in black, red, cyan, green and purple diamonds are respectively points P1, P2, P3, P4 and P5 in the feature space. X(t) are the *ML* amplitudes (in T).

Once the two-dimensional feature space is defined, it is necessary to apply the procedure described in section 3. Given a training dataset of n_N non-disruptive discharges with respective non-disruptive examples $\mathrm{X}_i\left(\chi_{1,i},\chi_{2,i}\right)\in\mathbb{S}, i=1,...n_N$ and n_D disruptive shots with corresponding disruptive examples $\Delta_i\left(\delta_{1,i},\delta_{2,i}\right)\in\mathbb{S}, i=1,...,n_D$, the coordinates of the non-disruptive centroid $C_N\left(c_1,c_2\right)$ are

$$\left(c_1,c_2\right)=mean\left(\chi_{1,i},\chi_{2,i}\right), i=1,...,n_N$$

and, in a similar way, the coordinates of the disruptive centroid $C_D\left(d_1,d_2\right)$ are

$$\left(d_1,d_2\right)=mean\left(\delta_{1,i},\delta_{2,i}\right), i=1,...,n_D$$

where $mean(\circ)$ returns a vector containing the average value of each column.

At this point, it is necessary to specify how disruptive/non-disruptive examples of individual discharges are obtained.

In the case of a non-disruptive shot, the time interval considered for the training starts when the plasma current is above a certain threshold, let's say $\mathrm{I_{T1}}$, until the plasma

current decreases below a possible different threshold I_{T2}. In this article, I_{T1}=0.9 MA and I_{T2}=0.7 MA.

Examples of disruptive discharges in $\mathbb{S}$ are determined in a different way. In general, the selection of a number of disruptive feature vectors to train predictors is not unique because disruption precursors do not appear always with the same anticipation. However, for mitigation purposes, it is reasonable to assume that the closer the feature vector to the disruption the more reliable the selection is. In the present case, only one example per disruptive discharge is considered in $\mathbb{S}$ and it has to be as close as possible to the disruption.

7. Training process description: computation of the centroids

The datasets of discharges used in this work correspond to JET operations with the ILW, from the beginning in 2011 until November 2016 (that is the last campaign previous to the preparation of the DT campaigns). The specific information appears in table 1. Disruption times have been defined as the time when the current quench starts. Approximately one third of the available discharges have been used for training (i.e. centroid computations) and the rest for test. It is important to note that no bias exist in the selection of discharges. Practically all disruptive and non-disruptive discharges have been chosen with the only exception of removing intentional and mitigated disruptions. Therefore, the disruptive discharges from table 1 are unintentional and unmitigated ones.

In this article, the predictor based on centroids has been tested with the mode lock normalised to the plasma current. In this respect, it is important to note that the reasoning carried out in the previous section about the *ML* signal (in particular, fig. 4) is valid for *ML/Ip*. As mentioned previously, the *ML/Ip* signal is typically used in the JET control system to trigger disruption alarms when its amplitude is above a selected threshold. Typical threshold in JET is when *ML*/*Ip* is above a value usually between 0.400 – 0.520 mT/MA [33].

Table 1: Datasets of disruptive discharges (D) and non-disruptive discharges (ND) to test a predictor based on centroids in JET. ‘Training’ means computation of centroids. The unbalance between disruptive and non-disruptive discharges reflects the fact that the number of disruptions in JET is about 8% of the discharges.

Type/use	*Number of shots*	*Range*
D/training	113	80181-82504 SEP 2011-MAR 2012
ND/training	1397	80176-82550 SEP 2011-MAR 2012
D/test	277	82569-92410 MAR 2012-NOV 2016
ND/test	3027	82552-92504 MAR 2012-NOV 2016

Fig. 5 shows the disruptive and non-disruptive centroids obtained with the training discharges of table 1:

$$\text{disruptive centroid: } \left(d_1 \pm stdx_D, d_2 \pm stdy_D\right)$$

$$\text{non-disruptive centroid: } \left(c_1 \pm stdx_N, c_2 \pm stdy_N\right)$$

where

$$\left(stdx_D, stdy_D\right) = std\left(\delta_{1,i}, \delta_{2,i}\right), i = 1,\ldots,n_D$$

$$\left(stdx_N, stdy_N\right) = std\left(\chi_{1,i}, \chi_{2,i}\right), i = 1,\ldots,n_N$$

and $std(\circ)$ returns a vector containing the standard deviations.

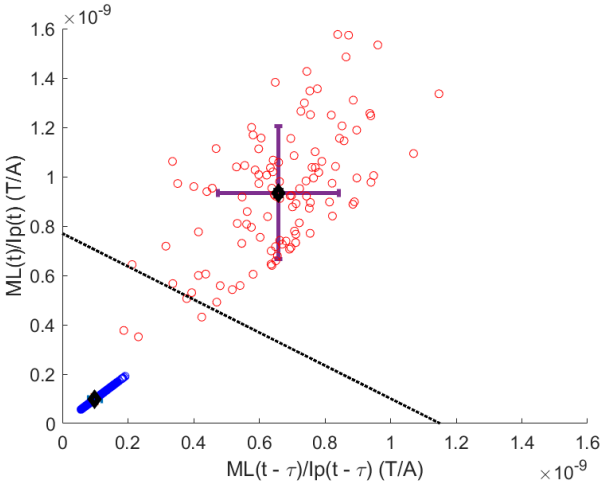

Fig. 5: Black points with error bars are the centroids. Blue and red points are non-disruptive and disruptive examples of individual discharges respectively. The plot shows the centroids for the mode lock normalised to plasma current. The sampling period of the *ML*/*Ip* signal has been τ = 0.002 s. The dashed straight line is the separation line between disruptive and non-disruptive behaviours.

The dashed line in fig. 5 is the linear separation frontier between disruptive and non-disruptive behaviours in the parameter space of the centroids. Fig. 5 also shows that non-disruptive examples form very compact clusters but the disruptive examples are scattered. Also, it should be emphasised that the error bars of the disruptive centroids do not overlap with the cluster of non-disruptive examples.

The error bars in the case of the non-disruptive centroid do not have any influence in the frontier estimation due to the high compact cluster structure. This is a consequence of the fact that $stdx_N \approx stdy_N \approx 0$. However, the displacement of the disruptive centroid within the square box defined by the error bars (fig. 6) changes both the slope and interception of the straight line. Therefore, to take into account the potential impact of the error bars, these ones have been divided in quarters. In this way, 9 points are considered in each direction and, therefore, 81 different disruptive centroids

have been considered (fig. 6). Consequently, 81 different models have to be tested and a criterion to choose the best one has to be established. In the present case, receiver operating characteristics (ROC) analysis has been implemented. Given different models for the same training and test datasets, the ROC curve [34] plots the success rate of each model versus their corresponding false alarm rates. The best model is defined by the point (*false alarm, success rate*) that is closest to the point (0, 100).

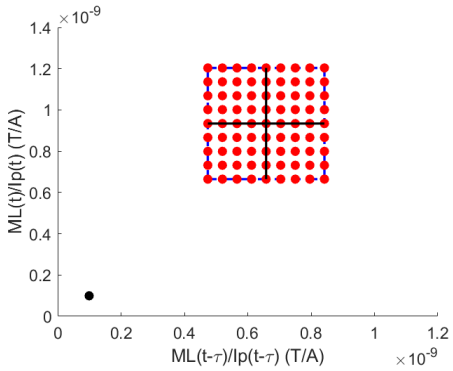

Fig. 6: In addition to the initial disruptive centroid $C_D\left(d_1, d_2\right)$, 80 new disruptive centroids are obtained with positive/negative increments of its coordinates in relation to the error bars in each dimension. The 81 centroids are the red points.

8. Results in JET with ILW discharges

Each one of the 81 models obtained by considering the displacement of the disruptive centroid has been qualified in terms of the valid alarm and false alarm rates together with the distribution of warning times according to the definitions provided in section 2. Fig. 7 shows the ROC curve of the 81 different models, where the valid alarm rates are plotted versus the false alarm rates. The closest predictor to (0, 100) is the optimal predictor and corresponds to the disruptive centroid located in the point $\left(d_1 - 0.75 \cdot stdx_D, d_2 - stdy_D\right)$. The separation frontier between disruptive and non-disruptive behaviours is

$$X\left(t\right) = -0.7441 \cdot X\left(t - 0.002\right) + 6.1243e - 10$$

where $X\left(t\right) = ML\left(t\right) / Ip\left(t\right)$, 0.002 s is the sampling period, *ML* is in Tesla and *Ip* in A. Therefore, a disruptive behaviour is recognised with this model and with this sampling period when

$$X\left(t\right) > -0.7441 \cdot X\left(t - 0.002\right) + 6.1243e - 10. \qquad (3)$$

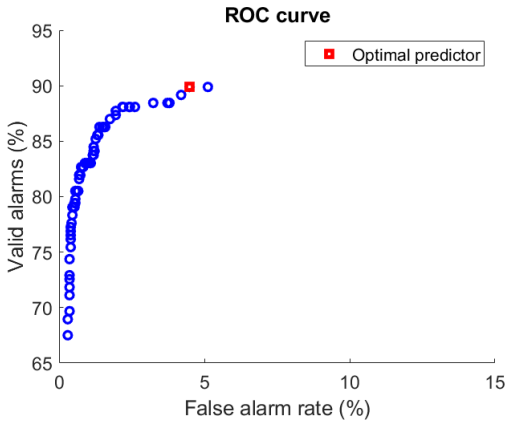

Fig. 7: ROC curve for *ML/Ip*. The red square is the closest points to (0, 100) and determines the best predictor.

By applying the model of equation (3) to the test dataset of table 1, the results are shown in table 2. Fig. 8 shows the distribution of warning times. This distribution can be fitted to an exponential model of the form

$$f(w) = f_0 \cdot \exp\left(-w/T\right)$$

where $f(w)$ is the fraction of detected disruptions with a warning time w, f_0 is the fraction of detected disruptions with positive warning times and T is the average warning time of the predictions. The resulting model parameters are $f_0 = 0.8963 \pm 0.0055$ and $T = 0.4428 \pm 0.0039$ s, where the estimations have been performed with 95% of confidence bounds and the R-square factor of the fit is 0.9778.

Table 2: Results of the predictor selected by the ROC curve. The average warning time (AVEWT) is determined by the fit of the warning times distribution to an exponential model.

CUMUL0 (%)	98
Valid alarms (%)	90
False alarms (%)	4
AVEWT (ms)	443

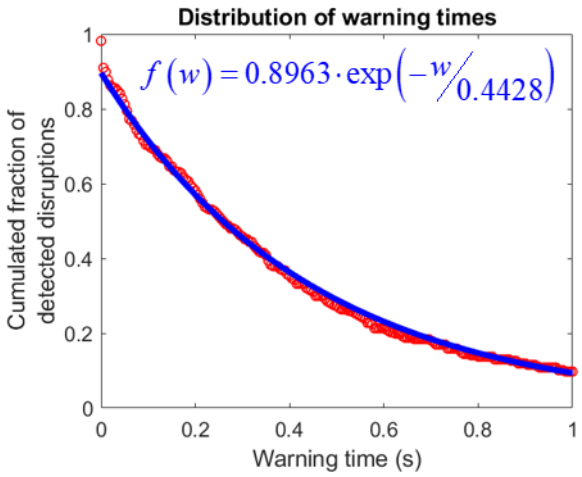

Fig. 8: The distribution of warning times (w) follows an exponential model with an average warning time of 443 ms.

The summary of the best predictor obtained by the ROC analysis is shown in table 2. Obviously, the rest of predictors in the ROC analysis are implemented in a very similar way to equation (3). All of them follow the expression

$$X(t) > A \cdot X(t - 0.002) + B \tag{4}$$

where A and B are constants that are different for each predictor and depend on the coordinates of the respective centroids.

It should be emphasised that a simple linear inequality in two variables can predict with high reliability (high valid alarm and low false alarm rates) disruptive behaviours in JET. The inequality is the outcome of condensing the disruptive/non-disruptive character of JET ILW shots into two centroids. During a running discharge, points in the feature space are obtained on a periodic basis and its nearest centroid determines the plasma state. The signal used in the predictions (ML/Ip) is a common quantity used in JET to recognise disruptive behaviours but, in this article, instead of basing the prediction on a threshold, a particular data processing has been performed.

Fig. 9 shows scatterplots of a non-disruptive discharge (fig. 9a) and two successful disruption predictions (fig. 9b and 9c). Green crosses represent non-disruptive behaviours and are located below the separation frontier (red line). Red circles are disruptive behaviours and appear above the red line. It is important to emphasise that non-disruptive points are quite concentrated around the diagonal of the parameter space. However, disruptive points are very spread from the diagonal.

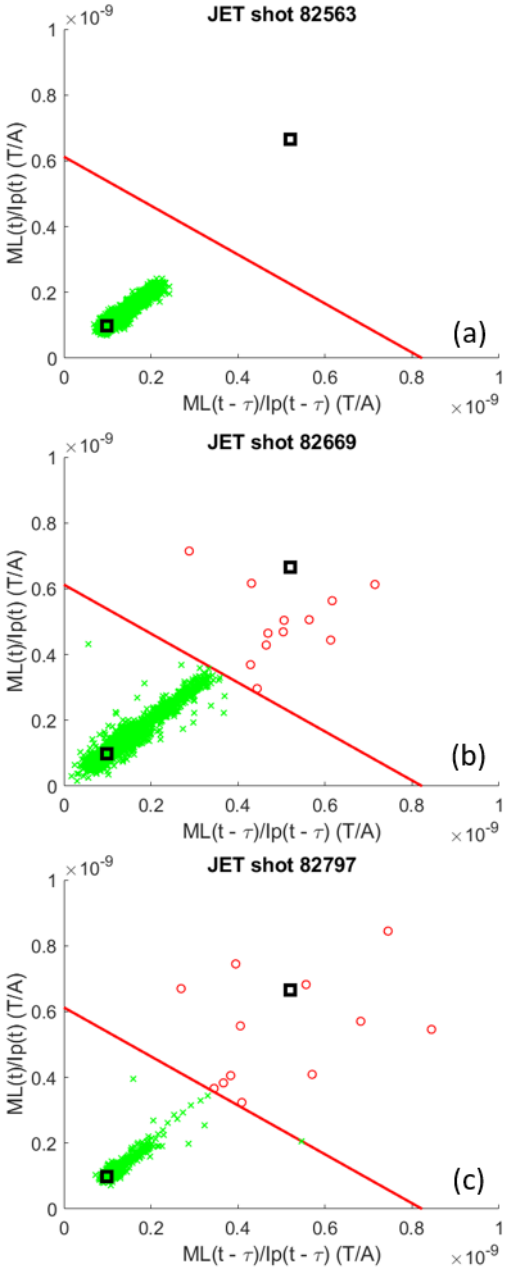

Fig. 9: Scatterplots in the parameter space of consecutive samples for one non-disruptive discharge (a) and two disruptive discharges (b, c). Black squares are the centroids. It should be noted that $\tau = 0.002$ s.

9. Distance to the diagonal in the parameter space: physics interpretation

An off-line analysis of the disruptive ILW discharges in JET and the corresponding alarms triggered by eq. (3) allows establishing that disruptive points

close to the diagonal show larger warning times than the ones far from the diagonal. It is important to note that the term 'warning time' is the actual time interval between the alarm time and the disruption time. The empirical fact that the closer the points to the diagonal the larger the warning times suggests the hypothesis that when an alarm is triggered, the time to the disruption TTD is related to the distance d_D to the diagonal, *i.e.* $TTD = f\left(d_D\right)$.

To model the form of the distribution, a first set of data is required to perform a fit. Once the fit has been carried out, a second set of data (different from the first one) is necessary to test the hypothesis.

To generate the first set, the 113 disruptive discharges of table 1, which have been used to determine equation (3), are considered. The predictor represented by equation (3) is applied to each one of these 113 discharges to obtain all pairs $\left(\textit{distance to the diagonal}, \textit{warning time}\right)$ whenever a disruptive behaviour is detected. This means that every discharge can possibly contribute with several pairs. It is important to note that the use of equation (3) for mitigation purposes only needs the first recognition of a disruptive behaviour to trigger an alarm. However, the situation to model the *warning times* is different as all possible pairs can contribute to fit the data to a specific distribution model.

With regard to the test procedure, a second set of pairs $\left(\textit{distance to the diagonal}, \textit{warning time}\right)$ is generated. This set is made up of all pairs corresponding to disruptive behaviours that are obtained with equation (3) and the 277 test disruptive discharges of table 1.

Among the various alternatives investigated, the best pdf to relate distances to TTDs is a Weibull distribution model of the form

$$TTD\left(d_D\right) = \alpha \cdot \beta \cdot {d_D}^{\beta-1} \cdot \exp\left(-\alpha \cdot {d_D}^{\beta}\right), d_D > 0 \qquad (5)$$

where $TTD(d_D)$ is an estimation of the time to the disruption, d_D is the distance to the diagonal and, finally, α and β are the model coefficients.

However, at this point, some considerations have to be taken into account. The first one is related to the selection of pairs from the training set to determine the model parameters (α and β). Due to the fact that 10 ms is the minimum time in JET to mitigate disruptions with the DMV, a disruption time predictor for JET should be trained to make estimations of TTDs greater than 10 ms. Therefore, only training pairs whose warning time is greater than or equal to 10 ms are used.

The second consideration is related to the goal of finding the largest warning time for which the model of equation (5) provides good results. To this end, 14 intervals of warning times have been defined to create 14 different models whose training warning times are between the following time limits

$$\left[0.01, 0.02 + \left(k-1\right)\cdot 0.01\right], k = 1,...,14 \quad (6)$$

where the times are in s. It should be noted that the temporal resolution between intervals is 0.01 s.

With this definition of intervals, the largest interval with the best results obtained with the test set will define the best model fit. It is important to note that the minimum width is 0.01 s and the maximum one is 0.14 s.

The last point refers to the specifications for building the test set. The test pairs *(d_D, warning time)* are generated with equation (3) for all the test discharges. From this set of pairs, 14 subsets are chosen in such a way that the corresponding test pairs are grouped according to their warning times. In this way, each subset includes the pairs *(d_D, warning time)* within the intervals established by equation (6).

Once the test subsets are defined, the d_D coordinates of the test points that recognise a disruptive behaviour are used as inputs to equation (5) to get the TTD estimations. To assess the model fit, a comparison between the warning times from equation (3) and the TTDs from equation (5) is carried out. The TTDs are classified into one of three classes. The first class contains the fraction of test discharges whose TTDs are below the limit of 10 ms. This corresponds to discharges in which the model fit produces smaller TTDs than the real warning times. The second class includes the fraction of discharges whose TTDs are within the warning time interval used for train/test. This is the ideal case. The third class comprises the fraction of discharges whose TTDs exceed the right limit of the warning time interval. This situation takes place when the model fit provides greater TTDs than the real warning times. Therefore, according to this classification criterion, a successful model would produce very low rates in classes 1 and 3 and a high rate in class 2.

Figure 10 shows the three test rates corresponding to classes 1, 2 and 3 respectively of the 14 models. It can be seen that models from k = 1 to k = 6 (equation (6)) provide the best possible results, i.e. class 1: 0%, class 2: 100%, class 3: 0%. This means that any one of these 6 models determine good predictors of the time to the disruption because the test warning times and the test TTDs show the same values in narrow time intervals.

In particular and according to the second consideration above concerning the largest warning time that produces the best results, the best choice corresponds to the model with k = 6. In this case, the model parameters α and β of equation (5) with 95% confidence bounds are respectively $\alpha = 0.0145 \pm 0.0041$ and $\beta = 0.7038 \pm 0.1060$. Appendix I shows an example of Weibull fit together with the α and β parameters for k = 1, …, 6.

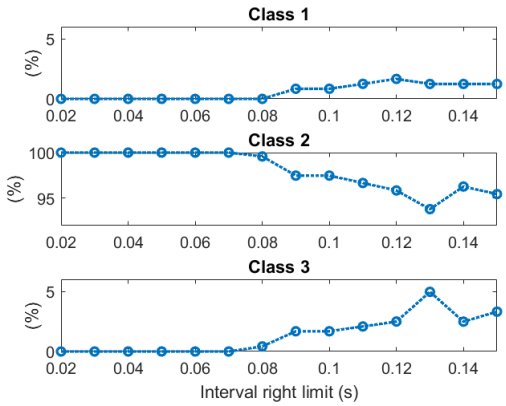

Fig. 10: Test rates in the respective classes for each one of the 14 models. The x axis shows the right limit of the respective time intervals.

10. Discussion

The centroid method is a reliable disruption predictor for JET that has been developed by means of a single signal: mode lock normalised to the plasma current. The method compresses into two single points the disruptive and non-disruptive information of ILW discharges and the prediction is based on the nearest centroid approach. The disruptive/non-disruptive plasma state is determined by a simple linear inequality in two variables, where the variables are the amplitude of consecutive samples that are acquired every 2 ms on JET. It should be mentioned that a preliminary version of the centroid method was tested with a small database of the JT-60U tokamak [35].

The linear dependency of equation (3) means that a simple threshold is not optimal to recognise disruptive behaviours. An additional condition has to be fulfilled. For example, the typical *ML*/*Ip* thresholds in JET between 0.400 mT/MA and 0.520 mT/MA [33], used to recognise a disruptive state at time t, are actually false alarms if, according to equation (3), the amplitudes 2 ms earlier are below 0.285 mT/MA and 0.124 mT/MA respectively. In the same way, small amplitudes (let's say 0.300 mT/MA) do not ensure a non-disruptive behaviour. In the case $X(t) = 0.300\, mT/MA$, a disruptive behaviour is present if $X(t-\tau) > 0.420 mT/MA$. Recently, following a different line of reasoning, the use of simple thresholds in the *ML* signal has been proven to be unreliable [36].

However, taking into account that the *ML*/*Ip* signal is always a positive quantity, the intercept of equation (3) defines a critical value above which the plasma is always in a disruptive state. In other words, if $X(t) > 6.1243e-10$ T/A, equation (3) is always true regardless of the amplitude $X(t-0.002)$.

Moreover, it is important to note that the centroid method is flexible enough to operate at different risk levels of disruptions without manual selection of thresholds. The centroid method allows a more flexible operation by selecting a specific predictor on the ROC curves (fig. 7). A ROC curve permits the selection of a particular predictor

to balance the requirement of a high success rate at the price of increasing the false alarm rate or to admit a smaller success rate but ensuring very low rate of false alarms. According to this, instead of choosing a signal threshold, a different predictor can be selected for each discharge because all of them implement equation (4), with A and B known in each case.

With regard to the distribution of points in the space of consecutive samples, the common property of the points close to the diagonal is that the difference of amplitudes between consecutive samples is really small. But it should be noted that the amplitude of the *ML/Ip* signal is related to the rotation braking of an MHD mode or to the increasing amplitude of the mode. Therefore, a small difference between consecutive samples of the *ML/Ip* signal can be interpreted like small variations either in the rotation or in the amplitudes of MHD modes and then, when these small variations take place they cannot be associated to disruptive behaviours. However, disruptive behaviours can be identified when the points in the parameter space are '*far enough*' from the diagonal, which means that there is '*enough difference*' between the amplitudes of consecutive samples. So, following the previous reasoning, it is possible to conclude that abrupt changes in the amplitude of consecutive samples means abrupt changes either in the rotation or in the amplitudes of MHD modes and, therefore, the fast variations are related to disruptive behaviours.

In addition to this, taking into account the bi-dimensional space of consecutive samples of the *ML*/*Ip* signal, the distance to the diagonal of a point in this space has been related to the time to the disruption by means of a Weibull model.

The estimation of the time to the disruption is a very relevant achievement that has been tested with good results. The most important aspect to emphasise is the simple empirical reasoning to determine the TTD instead of using very complex expressions deduced from general machine learning methods like in [26 - 28].

By considering three classes to compare the TTD predictions with the real warning times at the alarm times produced by the centroid method, the best possible results (class 1: 0%, class 2: 100%, class 3: 0%) are obtained at time intervals close to the disruptions (the largest interval is $[0.01, 0.07]$ s). However, larger time intervals (up to 0.15 s) also provide good results (fig. 10). It should be noted that the fraction of points whose predicted TTD and warning time belong to the same interval (class 2) is above 95% (except in one case). In the same way, the fraction of pairs whose predicted TTD is less than the real warning time (class 1) is below 2% in all cases. Finally, the fraction of pairs in which the predicted TTD is greater than the corresponding warning time (class 3) is always less than 5%.

With regard to methodological developments, it is a subject for future investigations to address the potential of the described techniques (centroid method and TTD estimation) to provide also classification of the disruption types. This is indeed a very important aspect for the optimisation of avoidance and mitigation strategies and

can require also more sophisticated metrics, such as the Geodesic Distance on Gaussian Manifolds already pioneered in [37].

Appendix I

Fig 11 shows, as an example, all the pairs with warning times less than 0.4 s from the training set of 113 disruptive discharges and the fit to a Weibull model.

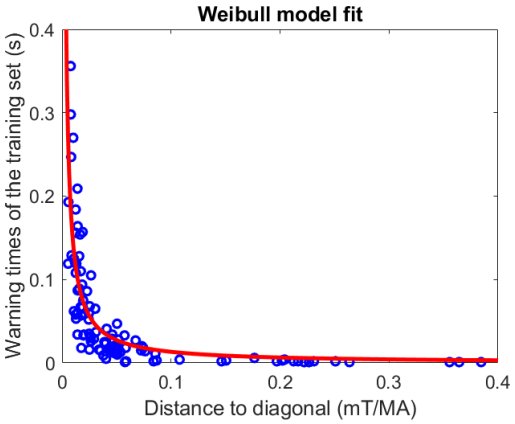

Fig. 11: Circles are the pairs *(distance to diagonal, warning time)* that are obtained by recognizing disruptive behaviours with equation (3) in the dataset of 113 disruptive discharges shown in table 1. The continuous red line is the fit to a Weibull model.

Fig. 12 shows the model parameters α and β of equation (5) with 95% confidence bounds for k = 1, …, 6. It is important to note that the first parameter remains around 0.015 but the second one decreases from 1 to 0.7. Equation (5) with β = 1 reduces to an exponential model (that is able to explain only the case for k = 1). However, the Weibull models explain the 6 cases in which the test warning times and the test TTDs have the same values inside narrow time intervals with the best possible prediction rates.

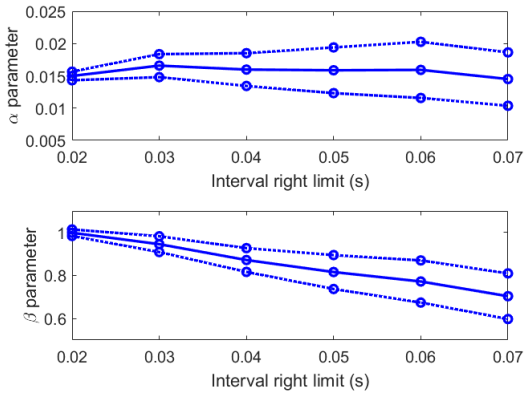

Fig. 12: Variation of the α and β parameters of the Weibull fit with 95% confidence bounds for the cases k = 1, …, 6 of the time intervals (6). The horizontal axis shows the right limit of the intervals.

Acknowledgments

This work was partially funded by the Spanish Ministry of Economy and Competitiveness under the Projects No. ENE2015-64914-C3-1-R and ENE2015-64914-C3-2-R.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] P.C. de Vries, G. Pautasso, E. Nardon, P. Cahyna, S. Gerasimov, J. Havlicek, T.C. Hender, G.T.A. Huijsmans, M. Lehnen, M. Maraschek, T. Markovič, J.A. Snipes and the COMPASS Team, the ASDEX Upgrade Team and JET Contributors. "Scaling of the MHD perturbation amplitude required to trigger a disruption and predictions for ITER". Nuclear Fusion 56 (2016) 026007 (10pp).

[2] M. Maraschek et al. "Path-oriented early reaction to approaching disruptions in ASDEX Upgrade and TCV in view of the future needs for ITER and DEMO". Plasma Phys. Control. Fusion 60 (2018) 014047 (11 pp).

[3] N.W. Eidietis et al. "Implementing a finite-state off-normal and fault response system for disruption avoidance in tokamaks". Nuclear Fusion 58 (2018) 056023 (9 pp).

[4] M. Lehnen, A. Alonso, G. Arnoux, N. Baumgarten, S.A. Bozhenkov, S. Brezinsek, M. Brix, T. Eich, S.N. Gerasimov, A. Huber, S. Jachmich, U. Kruezi, P.D. Morgan, V.V. Plyusnin, C. Reux, V. Riccardo, G. Sergienko, M.F. Stamp and JET EFDA contributors. "Disruption mitigation by massive gas injection in JET". Nuclear Fusion 51 (2011) 123010 (12pp).

[5] M. Bakhtiari, G. Olynyk, R. Granetz, D.G. Whyte, M.L. Reinke, K. Zhurovich and V. Izzo. "Using mixed gases for massive gas injection disruption mitigation on Alcator C-Mod". Nuclear Fusion 51 (2011) 063007 (9pp).

[6] G. Pautasso, K. Buchl, J. C. Fuchs, O. Gruber, A. Herrmann, K. Lackner, P. T. Lang, K. F. Mast, M. Ulrich, H. Zohm and ASDEX Upgrade Team. "Use of impurity pellets to control energy dissipation during disruption". Nuclear Fusion, 36, 10 (1996) 1291-1297.

[7] N. Commaux, L. R. Baylor, S. K. Combs, N. W. Eidietis, T. E. Evans, C. R. Foust, E. M. Hollmann, D. A. Humphreys, V. A. Izzo, A. N. James, T. C. Jernigan, S. J. Meitner, P. B. Parks, J. C.Wesley and J. H. Yu. "Novel rapid shutdown strategies for runaway electron suppression in DIII-D". Nuclear Fusion 51 (2011) 103001 (9pp).

[8] L. R. Baylor, C. C. Barbier, J. R. Carmichael, S. K. Combs, M. N. Ericson, N. D. Bull Ezell, P. W. Fisher, M. S. Lyttle, S. J. Meitner, D. A. Rasmussen, S. F. Smith, J. B.

Wilgen, S. Maruyama, G. Kiss. “Disruption Mitigation System Developments and Design for ITER”. Fusion Science and Technology. 68, 2 (2015) 211-215.

[9] B. Esposito, G. Granucci, P. Smeulders, S. Nowak, J. R. Martín-Solís, L. Gabellieri, FTU and ECRH teams. “Disruption Avoidance in the Frascati Tokamak Upgrade by Means of Magnetohydrodynamic Mode Stabilization Using Electron-Cyclotron-Resonance Heating”. Phys. Rev. Lett. 100, 045006 (2008).

[10] B. Esposito, G. Granucci, M. Maraschek, S. Nowak, A. Gude, V. Igochine, E. Lazzaro, R. McDermott, E. Poli, J. Stober, W. Suttrop, W. Treutterer, H. Zohm, D. Brunetti and ASDEX Upgrade Team. “Avoidance of disruptions at high β_N in ASDEX Upgrade with off-axis ECRH”. Nuclear Fusion 51 (2011) 083051 (9pp).

[11] A. H. Boozer. “Theory of tokamak disruptions”. Physics of Plasmas. 19, 058101 (2012).

[12] A. H. Boozer. “Runaway electrons and ITER”. Nuclear Fusion 57 (2017) 056018 (14pp).

[13] J.A. Snipes et al. “Large amplitude quasi-stationary MHD modes in JET”. Nuclear Fusion 28, 6 (1988) (1085-1097).

[14] M.F.F. Nave, J.A. Wesson. “Mode locking in tokamaks”. Nuclear Fusion 30, 12 (1990) (2575-2583).

[15] G. A. Rattá, J. Vega, A. Murari, G. Vagliasindi, M. F. Johnson, P. C. de Vries and JET-EFDA Contributors. “An Advanced Disruption Predictor for JET tested in a simulated Real Time Environment” Nuclear Fusion. 50 (2010) 025005 (10pp).

[16] B. Cannas, R.S. Delogu, A. Fanni, P. Sonato, M. K. Zedda and JET-EFDA contributors. “Support vector machines for disruption prediction and novelty detection at JET”. Fusion Engineering and Design 82 (2007) 1124–1130.

[17] 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”. Nuclear Fusion 48 (2008) 035010 (10pp).

[18] G. Pautasso, C. Tichmann, S. Egorov, T. Zehetbauer, O. Gruber, M. Maraschek, K.-F. Mast, V. Mertens, I. Perchermeier, G. Raupp,W. Treutterer, C.G.Windsor, ASDEX Upgrade Team. “On-line prediction and mitigation of disruptions in ASDEX Upgrade”. Nuclear Fusion 42 (2002) 100-108.

[19] B. Cannas, A. Fanni, G. Pautasso, G. Sias and P. Sonato. “An adaptive real-time disruption predictor for ASDEX Upgrade”. Nuclear Fusion 50 (2010) 075004 (12pp).

[20] G. A. Rattá, J. Vega, A. Murari, M. Johnson and JET-EFDA Contributors. “Feature extraction for improved disruption prediction analysis at JET”. Review of Scientific Instruments. 79, 10F328 (2008).

[21] G.Sias, R. Aledda, B. Cannas, R. S. Delogu, A. Fanni, A. Murari, A. Pau. “A multivariate analysis of disruption precursors on JET and AUG”. 1st IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis, Nice 1st-3rd June 2015.

[22] J. Vega, S. Dormido-Canto, J. M. López, A. Murari, J. M. Ramírez, R. Moreno, M. Ruiz, D. Alves, R. Felton and JET-EFDA Contributors. “Results of the JET real-time disruption predictor in the ITER-like wall campaigns”. Fusion Engineering and Design 88 (2013) 1228-1231.

[23] S. Dormido-Canto, J. Vega, J. M. Ramírez, A. Murari, R. Moreno, J. M. López, A. Pereira and JET-EFDA Contributors. “Development of an efficient real-time disruption predictor from scratch on JET and implications for ITER”. Nuclear Fusion. 53 (2013) 113001 (8pp).

[24] J. Vega, A. Murari, S. Dormido-Canto, R. Moreno, A. Pereira, A. Acero and JET-EFDA Contributors. “Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks”. Nuclear Fusion. 54 (2014) 123001 (17pp).

[25] J. Kates-Harbeck, A. Svyatkovskiy, W. Tang. “Predicting disruptive instabilities in controlled fusion plasmas through deep learning”. Nature 568 (2019) 526-531.

[26] F. C. Morabito, M. Versaci, G. Pautasso, C. Tichmann, ASDEX Upgrade Team. “Fuzzy-neural approaches to the prediction of disruptions in ASDEX Upgrade”. Nuclear Fusion 41, 11 (2001) 1715-1723.

[27] G. Pautasso, C. Tichmann, S. Egorov, T. Zehetbauer, O. Gruber, M. Maraschek, K.-F. Mast, V. Mertens, I. Perchermeier, G. Raupp,W. Treutterer, C.G.Windsor, ASDEX Upgrade Team. “On-line prediction and mitigation of disruptions in ASDEX Upgrade”. Nuclear Fusion 42 (2002) 100-108.

[28] B. Cannas, A. Fanni, E. Marongiu, P. Sonato. “Disruption forecasting at JET using neural networks”. Nuclear Fusion 44 (2004) 68-76.

[29] G.A. Rattá, J. Vega, A. Murari and JET EFDA Contributors. “Simulation and real-time replacement of missing plasma signals for disruption prediction: an implementation with APODIS”. Plasma Physics and Controlled Fusion. 56 (2014) 114004 (11 pp).

[30] J. Vega, R. Moreno, A. Pereira, S. Dormido-Canto, A. Murari and JET Contributors. “Advanced disruption predictor based on the locked mode signal: application to JET”. Proceedings of Science. ECPD 2015, 028.

[31] J. Vega, A. Murari, S. Dormido-Canto, R. Moreno, A. Pereira, G. A. Rattá and JET Contributors. “Disruption Precursor Detection: Combining the Time and Frequency Domains”. Proc. of the 26th Symposium on Fusion Engineering (SOFE 2015). May 31st-June 4th, 2015. Austin (TX), USA.

[32] S. Esquembri, J. Vega, A. Murari, M. Ruiz, E. Barrera, S. Dormido-Canto, R. Felton, M. Tsalas, D. Valcarcel and JET Contributors. “Real-Time Implementation in JET of the SPAD Disruption Predictor Using MARTe”. IEEE Transactions on Nuclear Science. 65, 2, (2018) 836-842.

[33] C. Reux, M. Lehnen, U. Kruezi, S. Jachmich, P. Card, K. Heinola, E. Joffrin, P. J. Lomas, S. Marsen, G. Matthews, V. Riccardo, F. Rimini, P. de Vries, JET-EFDA Contributors. “Use of the disruption mitigation valve in closed loop for routine protection at JET”. Fusion Engineering and Design 88 (2013) 1101 – 1104.

[34] K. J. Cios, W. Pedrycz, R. W. Swiniarski, L. A. Kurgan. 2007. “Data Mining. A Knowledge Discovery Approach”. Springer.

[35] J. Vega, F. Hernández, S. Dormido-Canto, A. Isayama, E. Joffrin, G. Matsunaga, T. Suzuki. “Assessment of linear disruption predictors using JT-60U data”. Fusion Engineering and Design. 146 (2019) 1291-1294 (https://doi.org/10.1016/j.fusengdes.2019.02.061).

[36] A. Murari, M. Lungaroni, E. Peluso, P. Gaudio, J. Vega, S. Dormido-Canto, M. Baruzzo, M. Gelfusa and JET Contributors. “Adaptive predictors based on probabilistic SVM for real time disruption mitigation on JET”. Nuclear Fusion. 58 (2018) 056002 (16 pp).

[37] A. Murari, P. Boutot, J. Vega, M. Gelfusa, R. Moreno, G. Verdoolaege, P.C. de Vries “Clustering based on the geodesic distance on Gaussian manifolds for the automatic classification of disruptions” Nuclear Fusion 53 (2013) 033006 (9 pp). https://doi.org/10.1088/0029-5515/53/3/033006.