EFDA–JET–PR(11)33

S. González, J. Vega, A. Murari, A. Pereira, S. Dormido-Canto,
J. M. Ramírez and JET EFDA contributors

Automatic Location of L/H Transitions: Time Determination, Uncertainty Interval and Validation

Automatic Location of L/H Transitions: Time Determination, Uncertainty Interval and Validation

S. González[1], J. Vega[1], A. Murari[2], A. Pereira[1], S. Dormido-Canto[3], J. M. Ramírez[3] and JET EFDA contributors*

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

[1]Asociación EURATOM/CIEMAT para Fusión, Madrid 28040, Spains
[2]Consorzio RFX, Associazione EURATOM/ENEA per la Fusione, Padova 4-25127, Italy
[3]Departamento de Informática y Automática, UNED, Madrid 28040, Spain,
** See annex of F. Romanelli et al, "Overview of JET Results",*
(23rd IAEA Fusion Energy Conference, Daejon, Republic of Korea (2010)).

Preprint of Paper to be submitted for publication in
Fusion Engineering and Design

ABSTRACT

Completely automatic techniques to estimate and validate L/H transition times can be essential in L/H transition analyses. The generation of databases with hundreds of transition times and without human intervention is an important step to accomplish (a) L/H transition physics analysis, (b) validation of L/H theoretical models and (c) creation of L/H scaling laws. An entirely unattended methodology is presented in this article to build large databases of transition times in JET using time series. The proposed technique has been applied to a dataset of 551 JET discharges between campaigns C21 and C26. The prediction with discharges that show a clear signature in time series is carried out through the locating properties of the wavelet transform. It is an accurate prediction and the uncertainty interval is ±3.2ms. The discharges with a non clear pattern in the time series use an L/H mode classifier based on discharges with clear signature. In this case, the estimation error shows a distribution with mean and standard deviation 27.9ms and 37.62ms respectively. As an additional original development, the level of significance of the predictions can be used to determine a variable time segment around the transition, in which there is a guaranteed probability to find the time of the L to H transition. Its temporal length correlates with the input power: increasing the input power, the plasma moves to a clearer H-mode regime, reducing the length of this guaranteed probability confidence interval.

1. INTRODUCTION

Plasma resistivity behaves as $\rho \sim T_e^{-3/2}$, where T_e is the electron temperature. As a consequence, the plasma resistivity saturates with the temperature and, in Tokamaks, although the plasma current is increased, the plasma heating due to the Joule effect tends to level off. As a result, the simple ohmic power is not enough to achieve thermonuclear conditions and auxiliary methods (for instance neutral beam injectors, electron cyclotron resonant heating or ion cyclotron resonant heating) are required. It is known that after going beyond a certain threshold of this additional injected power, the plasma shows high confinement properties (H-mode) characterised by the presence of an external transport barrier. The name of H-mode distinguishes this confinement regime from a worse one that takes place with auxiliary heating but before crossing such a threshold: the low confinement mode (L-mode).

Nowadays, there is not a single theory to describe the transition from the L to the H confinement regime. There are many competing theoretical models (a summary can be found in [1]) but no one has proven to completely explain the transition. A typical approach to acquire knowledge about the physics involved in the L/H transition has been to determine the time instants when it occurs and to generate scaling laws with several physical quantities [2]. Unfortunately, the transition time estimation can be a very tedious and time consuming process, because it requires human intervention to perform Visual Data Analysis (VDA).

To alleviate these problems, machine learning methods (in particular classification systems) have been developed in previous works with several purposes, for instance: regime identification in JET

[3], accurate determination of transition times (both L/H and H/L) in JET [4], dependence of the JET L/H and H/L boundary on local edge plasma parameters [5] and estimation of the H-mode threshold from a multimachine L/H database (http://efdasql.ipp.mpg.de/HmodePublic/) for a selected plasma quantity [6].

In the general formulation of a classification problem, an input sample $x = (x_1, x_2, ..., x_d)$ needs to be classified to one (and only one) of M groups (or classes) $C_1, C_2, \ldots, C_M$. To this end, a decision rule is used [7]. The values $x_j, j = 1,..., d$ are features of distinctive nature, which allows distinguishing the class that an input sample belongs to. Fig. 1 shows the decision function for a binary case (M = 2). To generate the frontier between the classes in a supervised manner, a dataset of input samples with a wellknown class for each sample is needed. This dataset is called training dataset. There are several methodologies to get the decision rule (also called model) from the training dataset and some of them have been used in the previous references: Support Vector Machines (SVM) [3, 4, 6], neural networks [5], fuzzy logic [3] and Bayesian statistics [4]. Typically, a test dataset (with known classes for its samples) is used to verify the success rate of the classifier. This is accomplished by classifying the test samples with the model obtained in the training process. Then, a comparison between the predicted values and the real ones is carried out.

In all the above references, the automatic estimation of the confinement mode has required human intervention to create the training and test datasets (identification of L and H samples and determination of transition times). As mentioned previously, this is an annoying task and for this reason, the datasets generated by a reduced number of people include only dozens of discharges ('discharge' or 'shot' will be used with the same meaning in this article): for example, 23 discharges were used in [3], 42 discharges for L/H transitions and 38 for H/L transitions in [4] and, finally, 67 shots in [5]. Also, it should be noted that the manual determination of supervised datasets are subject to human errors.

This article describes a novel technique for the determination of L/H transition times in a completely automatic way without human intervention. It has been applied to the JET database but the method is quite general and could be applied to other fusion devices. As it is explained later, the methodology uses very recent developments in machine learning for massive data analysis.

The use of a completely automatic technique makes possible the generation of datasets with hundreds or thousands of discharges and hundreds or thousands of samples per shot with an arbitrary number of features. When a reduced number of discharges and feature vectors are used to create classifiers, computations are tractable in nowadays personal computers. However, to solve the problem with hundreds of thousands of input samples, high performance computing is needed. Therefore, to avoid computational limitations in the methodology described in this article, recent parallel versions [8] of machine learning software have been used.

Completely automatic techniques to estimate and validate transition times can be essential in L/H transition analyses. Five reasons can be emphasised:

- No manpower is necessary.

- All discharges are processed exactly in the same way by the corresponding computer codes, thereby ensuring reproducibility and avoiding human errors.
- All transition times are estimated, even the ones of discharges impossible to distinguish by visual data analysis of time series.
- Large databases of transition times can be generated to tackle analysis with large statistical significance.
- The database of transition times can be updated after each new discharge in an unattended manner.

After summarising the objective of this paper in the introduction, section 2 provides an overview of the methodology to implement the automatic system. Section 3 explains the automatic way of recognizing discharges with H-mode. Section 4 sums up the software processes to distinguish between transitions with clear and non-clear signatures. In the first case, the transition instants can directly be determined. Machine learning methods are developed in section 5 to create a classifier to discern the confinement regime of arbitrary samples. The procedures for the automatic determination of transition times for discharges with non-clear signatures are described in section 6. Section 7 explains the validation of the transition times with clear patterns carried out using the models created in section 5. The results obtained in a database of 551 JET discharges are shown in Section 8. Finally, section 9 includes a discussion about the methods and results introduced in this paper. An appendix with the JET signals required by the classifier system has been included at the end of the article.

2. THE AUTOMATIC TRANSITION DETECTOR (ATD): OVERVIEW

In plasma physics, diagnostics produce similar signals for reproducible behaviours. This means that diagnostics translate physical properties into patterns with a direct correspondence between the physical state of the plasma and the structural shape that is generated in the signals. According to this, experts locate the L/H transition by recognizing a morphological pattern in time series data. This is accomplished through visual data analysis. Sometimes, the pattern appears very clearly and completely isolated from other plasma signatures. However, very often, the structural form to search for is hidden among other patterns and, in the worst case, it is impossible to locate it by simple VDA of time series. Of course profiles of relevant plasma parameters, such as the electron temperature, could be investigated but this would present several difficulties. First of all, the profiles are not always available with the appropriate time resolution. Moreover the interpretation of profile shapes, given the typical accuracy of the measurements, can be of prohibitively difficulty.

This work, therefore, translates into a computational environment the whole set of manual actions that performs a human analyst to locate L/H transitions using time series signals. The direct effect is the immediate location of transitions without the need of manpower. On the other hand, the methodology described in this paper can also determine transition times even in the situations where the human expert is not able to locate the transition by means of VDA of time series.

Figure 2 summarises the ATD methodology. It consists of 4 software modules ('Readout loop', 'Discrimination loop', 'Model creation' and 'Prediction module') that generate the database of transition times together with the uncertainty intervals in the estimations.

For an unattended determination of transition times in past discharges, the first step is to know whether the plasma achieves the H-mode. The H-mode is recognised by the presence of Edge Localised Modes (ELMs). Therefore, the module 'Readout loop' in fig.2 reads the database of discharges (in an arbitrary selected range) and processes the data for the automatic recognition of ELMs. If ELMs are found, an intermediate database is created with the discharge list of H-mode plasmas. Specific details are given in section 3.

To achieve the H-mode confinement regime, high auxiliary input power is necessary. However, the threshold that ensures the change of confinement is known within a wide margin. In addition, it is known that it also depends on the ion mass number [9]. Moreover the phenomenon of the transition can be different from discharge to discharge. Sometimes, the L/H transition takes place in such a way that the plasma enters directly into a Type I ELM H-mode (fig.3a). In this case, the transition time is typically identified by a fast drop in the D_α signal which becomes clearly visible within the L/H Transition Interval (LHTI). The LHTI is defined in this article as the temporal segment between the auxiliary power injection time instant and the first Type I ELM (fig.3b). This kind of transition shows a clear signature (or pattern) that is easily identified by means of visual data analysis. However, there is at least another class of transitions. The plasma attains a Type III ELM regime that finally, when the power is further increased, reaches the Type I ELM regime. When this type of transition happens, the drop in the D_α signal is not clear. It may appear mixed with oscillations in the D_α waveform or even completely hidden inside the oscillations (fig.4).

The second ATD module ('Discrimination loop') is explained in depth in section 4. This module performs a loop to read a range of H-mode discharges from the database generated with the 'Readout loop' module. It classifies the transitions into two groups according to their signature in the D_α waveform. In case of a clear signature, the time instant of the D_α drop is determined (and also an uncertainty interval) in an automatic way. A database with the transition times and uncertainty intervals is generated. On the other hand, the discharges showing a non-clear signature need a different data processing. Therefore, an auxiliary database with the corresponding shot list is created. A number of 551 JET discharges have been analysed in this work and 74.05% of them show a non-clear pattern. Therefore, only a reduced set of transitions can be located with the 'Discrimination loop' module. To include shots with non-clear signatures automatically, machine learning methods are used. A supervised classifier system (SVM based) is generated using the discharges showing a clear pattern in the transition. It is a binary classifier (samples are in either L-mode or H-mode) that determines a decision rule that allows the classification of other samples. This is accomplished with the third module of fig.2: 'Model creation'. The model described by the decision rule is stored for later use ('Data driven model' box of fig.2). Full details are given in section 5.

The last module of the automatic detector is the 'Prediction module'. This module reads data

from the shots with non-clear signatures. A subset of samples around the L/H transition is chosen per discharge. These samples correspond to the LHTI segment that delimits the transition (fig.4). The samples are classified (either L-regime or H-regime) according to the decision rule generated in the third module. By following the temporal evolution of the sample classification, the transition time instant is determined. A complete explanation is provided in section 6.

3. READOUT LOOP

The first step of the ATD methodology is the 'Readout loop'. Its aim is the automatic generation of a database of discharges that show H-mode transitions. This module replaces the visual data analysis processes that a human carries out to identify intervals in which the plasma is in the H mode of confinement. The manual identification is based on the recognition of ELMs in a discharge. Since ELMs only take place in H-mode plasmas, ELM detection means that the plasma has achieved the Hmode, and therefore, the plasma contains an L/H transition. ELMs are recognized by humans through the observation of typical peaks in the D_α waveform, which appear synchronously with drops in the diamagnetic energy [10].

The 'Readout loop' implements an unattended H-mode recognition system. The focus is exclusively centred on the D_α signal. H-mode discharges are identified through the pattern shown in fig.5 that corresponds to the one used in visual data analysis. But, how to recognize such a pattern in an automatic way using computers? To this end, a Universal Multi-Event Locator (UMEL) [11] has been used. UMEL is based on Support Vector Regression (SVR) [12] and it was previously applied to the location of ELMs in JET [10]. In [10], UMEL allows the recognition of discharges with L/H transitions and the determination of the time instant in which every single ELM takes place. Exactly the same methodology has been implemented in this article to identify H-mode discharges and to locate the time of the first ELM. Therefore UMEL allows identifying the interval LHTI.

4. DISCRIMINATION LOOP

The objective of the 'Discrimination loop' module is to distinguish the L/H transitions with clear signatures from the ones with non clear signatures. Necessarily, the transition instant takes place within the LHTI. Therefore, the 'Discrimination loop' module analyzes the structural shape of the D_α inside the LHTI, looking for a clear drop in the signal.

The wavelet transform [13] has been applied to the D_α samples in the LHTI to locate the transition pattern. On the one hand, the approximation coefficients of the wavelet provide the high-scale, low-frequency components of the waveform. On the other hand, the detail coefficients of the wavelet represent the low-scale, high-frequency components of the signal. Clear signatures are characterised by an unambiguous drop in the D_α emission (fig. 6a). In contrast, non clear signatures do not show such an obvious drop (fig.7a).

Due to the location properties of the Wavelet Transform (WT), a Haar WT has been chosen to locate the drop in the D_α and to decide whether the L/H signature is clear. However, it should be

noted that the decision about the type of signature has to be determined in an automated way without human intervention. Again, UMEL is used for this purpose. A SVR fit of the D_α Haar transform in the LHTI is computed. If the detail coefficients show a well localized peak (fig.6b), the transition has a clear pattern. In this case, the L/H transition instant is determined as the time corresponding to the maximum detail coefficient. On the contrary, if no well localized peak appears (fig.7b), the transition has a non clear signature and, therefore, the transition time instant cannot be estimated with the 'Discrimination loop' module.

It should be mentioned that different decomposition levels have been tested for the Haar transform. Taking into account that the drops in the D_α emission can be significantly different from each other, the best results have been achieved with a decomposition level of 5.

The transitions with a clear signature are included in a database of L/H transitions times and uncertainty intervals (fig.2). The uncertainty interval (error bar) corresponding to the transition time estimation is related not only to the sampling period (TS) of the D_α signal but also to the number of points that are reduced in the wavelet decomposition. Due to the fact that each decomposition level of the Haar WT reduces by a factor 2 the number of samples, the uncertainty interval is $\pm 2^L TS$ around the prediction (or equivalently an error bar of $2^{L+1}TS$) where L is the decomposition level of the Haar WT. As $T_S = 0.1ms$, the uncertainty interval for all determinations with a clear pattern is $\pm 3.2ms$.

5. MODEL CREATION

The L/H transition times cannot be located using the wavelet transform in shots with a non clear drop in the D_α signal inside the LHTI. Therefore, an alternative method has been used. It consists of developing a supervised classifier with the discharges identified in the 'Discrimination loop' module that show a clear signature in the D_α waveform. For these discharges, outside the uncertainty interval, the time instants are known to be in either the L-mode or H-mode and, therefore, the classifier can be built with any number of training samples in an automatic way. The goal of the 'Model creation' module of fig.2 is the automatic generation of a classification system to distinguish L and H confinement modes. The resulting decision function will allow the classification of new samples. Given the LHTI samples of a discharge, it is possible to present them in ascending temporal order to the classifier and to detect the change of confinement from L to H [4].

The first step to build the classifier is to define what features (signals) best represent the type of confinement (L or H) at any time instant. To this end, the signals have been chosen in an automatic way by using the results of the feature extractor described in [14]. It starts with a large dataset of signals and discards one by one the less relevant features up to the identification of an appropriate set of waveforms to properly distinguish the L and H confinement modes.

Features relevant to distinguish the L and H confinement modes are different in temporal segments close to the transition (let's say within ±100ms around the transition) and far from it (for example ±500ms about the transition point). In the final system, to classify any time instant of a discharge

as L or H, the temporal distance to the transition is unknown. Therefore, it is not possible to discern the best type of features (either close or far to the transition) to be used in the classification process. However, a unique classifier must be created.

In similar cases, a combination of classifiers in a multi-layer model has achieved good results. For example, a multi-layer classifier has allowed the achievement of the highest success rates ever in the prediction of JET disruptions between campaigns C1 and C19 [15].

For the L/H transition in JET, a two-layer model (SVM based) has been developed (fig.8). The first layer consists of a classifier (fig.8, upper left box) trained in a symmetric interval of 100ms around the transition with a resolution of 1000 samples/s (i.e. 201 samples per pulse with a sampling period of 1ms). On the other hand, a second classifier in the first layer appears (fig.8, upper right box). It is trained with different features than the previous one and a symmetric interval of 500ms around the transition time. The resolution is again 1000 samples/s which implies the use of 1001 samples per discharge. The respective objective of the first and second classifiers is to emphasize the automatic learning near to and far from the transition.

The signals used in each classifier are described in Appendix A and also in fig.8. As it was previously pointed out, both sets have been selected by an automatic feature extractor [14]. It is important to note that the signals determined as relevant for the ±500ms classifier have also been identified as relevant in the ±100ms classifier. It means that these signals are suited enough for the phenomenological description of the decision function. However, the classifier closer to the transition needs more information to obtain a high success rate because the prediction just around the transition is more complex.

SVM has been chosen to implement the classifiers. SVM is a binary classification system that determines the decision function to separate both classes. SVM uses a set of training samples (x_i, y_i), i = 1, ..., N, where x_i are the feature vectors and $y_i \in \{-1, +1\}$ (in this article the L-mode and H-mode will be identified by class $\{-1\}$ and class $\{+1\}$ respectively). In general, a classification problem is not linearly separable in the input space. SVM maps the input space into a feature space (typically of much higher dimensionality) by means of a non-linear function called kernel. This mapping can always separate data from two categories by means of a hyper-plane (fig.9). SVM allows estimating the following decision function:

$$D(x) = \sum_{i=1}^{N} \alpha_i y_i H(x_i, x)$$

$H(x_i, x)$ is the kernel function and the parameters α_i, a = 1,..., N are the solutions of the optimization problem solved by SVM [7]. The data points $\mathbf{x}_i$ associated with the nonzero α_i are called support vectors. Once the support vectors have been determined, the SVM decision function has the form

$$D(x) = \sum_{\text{support vectors}} \alpha_i y_i H(x_i, x) \qquad (1)$$

$D(\mathbf{x})$ is the distance (with sign) from $\mathbf{x}$ to the hyper-plane that separates the two classes and, hence,

the hyper-plane points satisfy D(x) = 0 . It should be noted that D(**x**) actually is only proportional to the real distance. However this is not a problem, because all data points are rescaled by the same normalization factor.

The rule to classify a feature vector **u** as L mode (class {−1}) or H mode (class {+1}) is given by:

$$\begin{array}{c} \textit{if sgn } (D(u)) \geq 0 \\ \mathrm{u} \in \{+1\} \\ \textit{othewise} \\ \mathrm{u} \in \{-1\} \end{array} \tag{2}$$

where sgn(t) is the sign function, i.e.

$$sgn(t) = \begin{cases} 1 \textit{ if } t \geq 0 \\ -1 \textit{ if } t < 0 \end{cases}$$

Both first layer classifiers use a linear kernel ($H(\mathbf{x}, \mathbf{x}') = \mathbf{x} \cdot \mathbf{x}'$) to separate the two confinement modes. Therefore, the decision function is linear in its components and satisfies

$$D_1(\boldsymbol{u}) = C_0 + \sum_{i=1}^{11} C_i \cdot u_i \tag{3}$$

for the ±100 ms classifier. The constants C_j, j = 0, 1, ..., 11 come from (1) and u_i are the different plasma quantities: Bt, LI, ..., LAD4 (see fig.8). For the ±500ms classifier, the hyper-plane becomes

$$D_2(\boldsymbol{v}) = B_0 + \sum_{i=1}^{5} B_i \cdot v_i \tag{4}$$

where, again, the constants Bi are determined from equation (1) and the physical quantities correspond to Bt, AD36, ..., LAD4 (see fig.8).

The training of the second layer classifier (it has been also based on SVM with a linear kernel) is carried out with the outputs of the first layer classifiers (D1(u) and D2(v)). Its decision function has the form

$$D(\mathbf{u}, \mathbf{v}) = A_0 + A_i \cdot D_1(\mathbf{u}) + A_2 \cdot D_2(\mathbf{v}) \tag{5}$$

It is important to mention that equation 5 defines the model to predict the confinement mode of a discharge at any time:

The plasma is in L mode ⇔ D(u,v) < 0.Otherwise, the plasma is in H mode

6. PREDICTION MODULE

The last step of the methodology is the prediction of L/H transition times in discharges with non clear transition patterns. To classify a time instant (t_0) of a discharge as either L mode or H mode, two samples at this time instant are necessary. The samples are represented by their respective feature vectors $\mathbf{u}(t_0) \in R^{11}$ and $\mathbf{v}(t_0) \in R^5$ to be classified respectively by the ±100ms and ±500ms

classifiers. To this end, $D_1(\mathbf{u}(t_0))$ (equation (3)) and $D_2(\mathbf{v}(t_0))$ (equation (4)) have to be computed. These distances (with sign) are introduced in equation (5) to obtain:

$$D(t_0) = A_0 + A_i \cdot D_1(\mathbf{u}(t_0)) + A_2 \cdot D_2(\mathbf{v}(t_0)) \tag{6}$$

This expression is exactly equation (5) but it has been emphasized that the prediction corresponds to time t_0 and is carried out according to the following rule:

if sgn $(D(t_0)) < 0$
    the discharge is in L mode at t_0
otherwise
    the discharge is in H mode at t_0

Given a discharge, the samples within the LHTI are classified one by one in ascending temporal order with the above procedure. Fig.10(a) shows the D_α signal around the transition of JET Pulse No: 74636. The typical drop is not observed and, therefore, the transition is determined with the 'Prediction module'. Figure 10(b) represents the distances to the separating hyper-plane (6) of the samples. The bold red vertical line determines the transition time. It corresponds to the time instant t_{TRANS} in which $D(t_{TRANS}) = 0$.

At this point, it is important to note that the transition time has to include an uncertainty interval. This uncertainty interval can be directly connected with the sampling period of the feature vectors, like in [4]. However, this approach could be rather optimistic and a more conservative option should be identified.

A possible way to determine the uncertainty interval could be closely related to the transition time estimation method. In the present case, machine learning techniques (in particular, classifiers) have been used. By following the distance evolution of equation (6) around the transition, (actually $\mathrm{sgn}(D(t))$), the transition is located. But the question is: are comparable the levels of significance of the classifications near and far from the transition? In other words, how certain are the classifications close and faraway?

To address this point it has to be taken into account that a binary classifier is being used to follow the temporal evolution of the confinement mode. The predictions take place just in the boundary of both classes and it is clear that predictions in the frontier between classes have less reliability. The classifier predicts a label (either L or H) but it is unknown how good such prediction is. Therefore, for a proper classification, the inclusion of a level of significance for each prediction is required. Intuitively, thinking of compact clusters, one expects that predictions for samples near to the cluster centres have a high level of significance. However, this level of significance should decrease for samples approaching the border between the classes. Intuition tells that the level of significance is minimum just at the border. Taking into account that SVM has been the base classifier in this article,

the above interpretation means that the classification of samples near the separating hyper-plane shows a lower reliability than samples far from the hyper-plane. Figure 11 gives a visual image of this reasoning.

Coming back to the L/H problem, by following the temporal evolution of the reliability as samples are classified as L or H in the LHTI, a threshold in this quantity can be defined. This threshold defines a level of reliability below which the prediction is assumed to be not very trustworthy. The time instants when the reliability crosses the threshold could be considered a Probability Confidence Interval (PCI) of the prediction (fig.11).

To provide the prediction with a level of reliability, conformal predictors have been used [16]. Conformal predictors are a type of predictive models that hedge the output labels with a couple of values, confidence and credibility, for each classification. Both parameters vary between 0 and 1. Confidence indicates how relevant the classification given by the predictor is (0 means not relevant at all and 1 means completely relevant). On the other hand, credibility represents the quality of the data on which to base the decision [16]. In other words, it reflects how representative the training data is to perform the classification of a new sample. A low credibility means the sample to classify is not conformal to the training dataset, whereas a high value implies the opposite.

Conformal predictors can be developed in a straightforward manner from any method of point prediction, including for example SVM and neural networks. The SVM combiner of fig. 8 has been adapted as a conformal predictor to provide an evaluation of its own accuracy and reliability.

For each sample to classify, the values of confidence and credibility are computed using a measure of non-conformity [16]. A non-conformity measure is a measure of the supportiveness of this sample. A high value of the non-conformity measure indicates that this sample is strange and unlikely to occur. However, nonconformity measures are not absolute. The numerical value for sample i, let's say α_i, does not, by itself, tell how unusual the sample is. For that, it is necessary a comparison of α_i to the other α_j. A convenient way of making this comparison is to compute the fraction

$$\frac{\#\{\, j = 1,\ldots,n : \alpha_j \geq \alpha_i \,\}}{n}$$

where n is the number of samples. This fraction, which lies between $1/n$ and 1, is called the *p-value* of sample i. It represents the fraction of samples as non-conforming as sample i. If the p-value is small (close to $1/n$) then the sample i is very non-conforming, ie it is an outlier. If it is large (close to its upper bound 1), then the sample i is very conforming.

The conformal prediction in a binary classifier implies to compute the *p-values* of the sample to classify for each possible class. The label of the classification will be the class with the largest *p-value*. The credibility is given by the largest *p-value* and the confidence is 1 – the second p-value. A very detailed description and an example of conformal predictors in multi-class problems can be found in [17].

There are many non-conformity measures for conformal predictors. In this paper, SVM is used as

base classifier. But, as it was mentioned before, the nearer the sample to the separating hyper-plane, the lower its significance level. Therefore, to take into account this fact in the calculation of the *p-values*, the non-conformity measure has been chosen proportional to the absolute value of the distance between the sample and the separating hyper-plane.

Among the several variants of conformal predictors, an off-line version of Inductive Conformal Predictors (ICPs) [18] has been selected in this article. The main advantage of the ICPs versus the transductive conformal predictors is its computational efficiency: the bulk of the computation is performed only once no matter the number of samples to classify [18]. Typically, in the off-line version of the ICPs, the training dataset is split into two different sets: a proper training set and a calibration set. In the present case, the former is used to compute the separating hyper-plane and the latter is utilised to compute the values of confidence and credibility through the non-conformity scores.

Figure 12 shows again the example of fig.10. The level of reliability in the prediction can be seen in fig.12c. In this article, instead of using two parameters (confidence and credibility) to qualify the prediction, the product of both quantities has been chosen. The idea of selecting the product of confidence by credibility has already been used in [19]. In the case of the L/H transitions, a reliability threshold of 0.3 has been chosen after an appropriate series of tests. This means that, on average, the values of confidence and credibility are greater than 0.5 (to be more exact, if the value of both parameters is around 0.55, the product is 0.3). The horizontal solid line in fig.12c gives such reliability threshold and the vertical dashed lines delimit the PCI. It is important to note that, firstly, the PCI might not be a symmetric time interval around the L/H transition and, secondly, its estimation criterion completely relies on the classification method.

7. VALIDATION OF THE L/H TRANSITION TIMES OF DISCHARGES WITHCLEAR PATTERNS

This section describes the validation of the transition times found in the 'Discrimination loop' module for transitions with a clear signature in the D_α. With the present technique, this validation is extremely important because the lack of human intervention makes the detection of errors impossible. If the structural shape of the D_α is such that an abrupt change of amplitude is erroneously taken as a transition pattern, then some samples will be misclassified and, therefore, the training dataset for the twolayer model will use false information. It is clear that a software process has to be in charge of the validation to avoid wrong identifications.

The two-layer model explained in the previous sections has been trained with shots that show a clear pattern in the D_α signal. Therefore, if any of these discharges is used as input to the two-layer classifier, its resulting PCI should include the transition time window of 3.2ms around the identified transition obtained with the 'Discrimination loop' module (it should be remembered that ±3.2ms is the uncertainty interval determined by the wavelet location method). The inclusion of the shorter interval, obtained with the wavelet locator, inside the larger interval, achieved with the conformal prediction of the two-layer model, is a consequence of the conservative character of the

second method. Therefore, if the shorter interval is not included in the larger interval, it is possible to conclude that the transition pattern was not properly located and, therefore, this shot must not be included to create the two-layer model.

This validation method is appropriate for two main reasons. The first one is the low misclassification rate of the 'Discrimination loop' module. This means that the statistical weight of misclassified discharges is very low. Therefore, in generating the two-layer model, it is possible to have a limited number of samples with a wrong classification of their confinement mode (L or H). But being a reduced number of samples, their impact is expected to be negligible in the model determination. On the other hand, these samples should be detected as misclassified by the model.

The second reason is the use of linear kernels to build the SVM multi-layer model. Linear kernels have very high generalization ability and their sensitivity to errors is lower than other kernels. In contrast, they usually have lower success rates (this is not an issue in the present case, as it is shown in the next section). So, again, a reduced number of misclassified samples should have a low influence in the training process.

Therefore, if the validation process detects misclassifications, the corresponding discharges are excluded from the database of discharges with a clear signature and they are included in the database of a non-clear pattern. Then, the two-layer model is retrained using the new database of discharges that show a clear pattern.

8. RESULTS FOR JET

A database of 551 discharges from JET campaigns C21 to C26 has been used to test the methodology introduced in this article. The L/H transition times have been determined by experts and, therefore, the error of the technique can be estimated.

The 'Discrimination loop' has located a total of 143 transitions (25.95%) with a clear L/H transition pattern. With these discharges, the two-layer model has been built. The first layer models of fig.8 have been determined with 28743 and 143143 feature vectors respectively (with respective dimensions of 11 and 5). To reduce the computation time, the SVM parallel version described in [8] has been used. To implement the conformal predictor in the second layer, the set of 143 discharges has been split into a proper training set (containing a random selection of 75% of the 143 shots) and a calibration set (with the remaining 25%).

The success rate of the SVM classifier with a linear kernel in the interval ±100ms has been 96.65%. The corresponding one for the SVM classifier with training samples ±500ms around the transition has been 97.72%. This seems to indicate that the latter classifier is better. However, this is not necessarily true in the case of the L/H transition. Points far from the transition are easier to classify than points close to it and, therefore, a better success rate does not mean a better classifier.

As it has been mentioned previously, the goal of the combiner of fig.8 is to provide a combination of the distances of the first layer SVM classifiers to reduce the classification errors. Figure 13 shows a data processing example in JET Pulse No: 74570. In this case, the readout loop has enclosed the

transition in the time interval [3.606, 4.219]. The first layer classifiers have been applied to this LHTI obtaining the output distances that can be seen in fig. 13. The ±100ms classifier misclassifies the points in the interval [3.656, 3.740]. These misclassified points are far from the real L/H transition (3.977s) and thus, the classifier does not work properly (it should be noted that the classifier has been trained in an interval of ±0.1s around the transitions). In contrast, it predicts other L/H transition at time 3.999 seconds. This time is closer to the real transition time than the time predicted by the ±500ms classifier (4.012s). The output of the combiner model is a combination of both distances that avoids the misclassification of the points in the interval [3.656, 3.740] and gives a transition time of 4.010s. The error of this time compared with the real L/H time is worse than the one obtained by the ±100ms classifier, but it is better than the one given by the ±500ms classifier. Therefore, the combiner uses the information from both classifiers to minimize the error of the predicted L/H transition and to avoid the misclassification of points far from it.

After generating the two-layer model, the first operation has been the validation of the training set of 143 discharges to make sure that the automatic recognition process of L/H transitions worked properly. In two discharges it has been found that the transition time predicted by the wavelet method is outside the PCI estimated by the twolayer classifier. As explained in the previous section, the interpretation is that the structural pattern of the transition is not clearly isolated in their Da signals and, therefore, these shots cannot be used to train the two-layer model. So, the discharges have been deleted from the database of clear signatures and they have been included in the database of non-clear patterns.

With a total number of 141 (25.59%) discharges with clear signature, the twolayer model is built again. A 75% of the discharges (106) is randomly chosen to be used as training dataset in the first layer models and as proper training set in the second layer model. The rest of discharges (35) are used as test datasets for the ±100 and ±500ms classifiers as well as calibration set for the combiner classifier to compute confidence and credibility.

Table 1 summarizes the hyper-plane coefficients corresponding to equations (3) and (4) respectively.

The hyper-plane of equation (6) corresponding to the SVM combiner is

$$D(t_0) = -14.41 + 18.81 \cdot D_1(\mathbf{u}(t_0)) + 19.61 \cdot D_2(\mathbf{v}(t_0))$$

Table 2 summarizes the success rates achieved with these datasets in the confinement mode prediction with the different classifiers.

In table 2, it is important to note that the success rates of the training sets are a bit higher than the ones of the test/calibration set. This is normal taking into account that the success rates achieved with the training data are overestimated. In addition, it should be emphasised that the classification rates obtained by the combiner classifier are higher than the other classifiers. This confirms the fact that a combiner classifier uses the information provided by the others to increase the system performance.

The 141 discharges are classified with the two-layer classifier and, in this set of discharges, all

transition times determined with the wavelet method are within the PCI that has been estimated with the conformal predictor. After this validation, the two-layer model is used to predict the transition times of the 410 remaining discharges that do not show a clear transition pattern.

The conformal predictor has been applied to estimate the PCI in the prediction of the L/H transition times. As previously established as a reasonable choice, the reliability threshold has been set to 0.3. Table 3 summarizes the results with the 410 discharges of non clear transitions.

Two different groups of discharges have been identified. The first group corresponds to discharges whose real transition time is within the estimated PCI (fig.14). The success rate is 96.83%. Fig. 15a shows the histogram of the PCIs corresponding to the 397 shots (the bin size is 20ms). The mean PCI is 225.2ms. It is important to mention that there are 5 discharges (Pulse No's: 74759, 76519, 76521, 76688 and 77634) whose uncertainty intervals vary between 710ms and 3.91s. These special cases are discharges with moderate input power (in all cases less than 10MW) and/or discharges that achieve their maximum value in a non-abrupt way, with a long period of low power injection. Figure 16 shows one of these examples. The plots from top to bottom are the input power, the D_α emission, the output of the classifier and the level of reliability for discharge 76521. These shots with a long period of low injected power are not JET standard discharges and greatly increase the average PCI in table 3. If they are not taken into account, the mean PCI becomes 206ms.

Figure 15b is also related to the first group of discharges of table 3. It represents the distribution of the differences between the real and the estimated transition time of the 410 discharges. The absolute mean value is 27.9ms and the standard deviation is 37.6ms. These statistical parameters mean that the prediction is quite accurate. The uncertainty interval for non clear transitions is also 27.9ms.

The second group of discharges in table 3 is made up of shots (Pulse No's: 73559, 73576, 74611, 74622, 75066, 75222, 75740-75743, 77009, 77054, 77179) whose real transition time is not inside the estimated PCI (fig.17). Only 13 discharges (3.17%) belong to this group. The average(|real L/H – estimated L/H|) is 93.8ms and the mean PCI is 453.5ms. However, it is important to emphasize that in all these cases, the real transition time is very close to the extremes of the PCI. Figure 18 shows the distribution of times between the real transition and the closest interval extreme. The average value is 5.96ms. Figure 17 represents the Pulse No: 77009, in which a moderate input power is maintained for a quite long period after the transition. As a consequence, the PCI is 4.172s and, therefore, it highly contributes to have a very large value of the PCI average (453.5ms) as it is shown in table 3. Again, if this shot is ignored due to its non standard character in JET, the mean PCI is 164.3ms that is comparable to the statistical values of the first group of discharges.

9. DISCUSSION

ATD is a methodology to determine L/H transition times in a completely automatic way. ATD can be used to create large databases of L/H transition times (the larger database the more reliable results) that can be used for several purposes, for example, L/H transition physics analysis, validation of

L/H theoretical models or creation of L/H scaling laws. ATD can be executed in an unattended way after every discharge to automatically include the new transition time in the database.

At least two different transitions can be found in JET according to their signature in temporal evolution signals. The first one shows the clear transition pattern that generally is used in visual data analysis to recognise when the transition takes place ie a drop in the D_α signal. Approximately, only 25% of JET discharges show this pattern. The rest of discharges can require very time consuming VDA and also the transition recognition can be impossible even for an expert human eye. Due to the fact that the L/H transition time of 75% of JET shots can be of difficult determination, machine learning methods have to be used. The difference of ATD with previous methods is the complete absence of human intervention in the process, not only for selecting a training dataset but also for prediction, uncertainty interval estimations and validation.

Usually in machine learning, tasks that can be manually carried out by a human without much effort (for instance some event identification or pattern recognition), can also be automated with a high success rate. Therefore, the question was: can the discharges that show a clear transition signature be used to predict the transition time of the remaining 75% discharges? Even more, can the results be validated in an unattended manner?

ATD has shown that the answer to both questions is 'yes'. Before focusing the attention on validation aspects, it is important to emphasize how the prediction of new discharges is carried out. The previous step consists of discriminating the clear/non clear character of the new discharge. If it is considered to have a non clear signature, the transition instant, $t_{NONCLEAR}$, has to be determined by the conformal classifier of section 6. On the contrary, if a clear drop in the Da is detected, the transition time t_{CLEAR} is determined with the Haar transform and the UMEL technique (see section 4). It should be noted that the uncertainty interval here is 6.4 ms which is smaller than the average of the first group of discharges in table 3 (13 ms). The validation of this prediction is performed with the existing two-layer model presented in section 5. If the PCI determined with the conformal predictor encloses the temporal segment of 6.4ms resulting from the Haar transform and UMEL, then $t_{CLEAR} \pm 3.2$ms is considered the transition time. (It should be remembered that the training dataset of shots with a clear pattern has been determined in section 7 to have an uncertainty interval of 6.4ms and, therefore, the procedure is justified). Otherwise, if the PCI determined with the conformal predictor does not enclose the temporal segment of 6.4ms, the transition time is the one predicted by the conformal classifier.

The unattended prediction of transition times and their validation are original developments presented in this article. Another new tool is the PCI obtained with the conformal predictors. By properly choosing the threshold, it is possible to determine a time interval PCI within which the presence of the transition is guaranteed with a certain probability (determined by the choice of the threshold). Conformal predictors have been used exactly because they have the capability to guarantee these hedged predictions. Moreover, among other advantages, they do not need to assume any type of data distribution (as it happens for example with Bayesian estimators in which a prior

information is required). In addition, conformal predictors allow the determination of a PCI on a discharge by discharge basis and, therefore, each shot has its own PCI independently of other discharges.

The length of the PCI is linked to the operational management of the discharges. As it was previously explained, as time increases in the LHTI, the confinement mode prediction evolves from a high reliable L-mode towards a low confidence L-mode. The low reliability can be maintained during more (figs.16 and 19) or less (fig.12) time. At a certain time instant, the classifier starts to predict H-mode with increasing reliability as time passes. From a mathematical point of view, the interpretation of this behaviour has been suggested previously in fig.11: the feature vectors go through the frontier between both classes. However, it can be also described in terms of the physics how the plasma begins to leave the L confinement mode to achieve the H-mode. The transition process takes certain time, in which the plasma moves from the low confinement mode to the high mode, crossing an intermediate state during a variable period of time. At each time instant, the intermediate plasma state is characterized as L- or H-mode but the prediction credibility is low. The length of this transition period is not arbitrary and it is related to the manner in which the input power is injected into the plasma.

Figures 16 and 19 show discharges in which the input power is injected in steps and the plasma takes its time to achieve the H-mode. Fig. 20 plots the power injection in dPulse No: 74636. In this case, the injection is not carried out in steps and the uncertainty interval is 104ms (fig.12). Other discharge numbers in which this happens are Pulse No's: 74438, 76635, 76914, 77050, 77781, ...

This interpretation of the PCI as an interval in which the transition takes place with a certain probability and its link with the use of the additional heating power is confirmed by the following observations: 1) The level of reliability of the predictions starts to decrease in a monotonic way. 2) A minimum level is maintained during certain time. 3) The level of reliability begins to increase also in a monotonic way up to high values.

Future work should be devoted to the analysis of the opposite H to L transition. Moreover, the ATD opens the possibility to perform studies using databases of thousands of discharges. This will permit to reformulate scaling laws with a much higher statistical confidence in the results. It is also worth pointing out that the propped method is based on absolutely general tools, such as UMEL, SVM and conformal predictors. The approach can therefore be applied to any form of abrupt phase transition, whose signature can be founding time series signals.

ACKNOWLEDGEMENT

This work was partially funded by the Spanish Ministry of Science and Innovation under the Project No. ENE2008-02894/FTN. This work, supported by the European Communities under the contract of Association between EURATOM/CIEMAT, was carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

[1]. J.W. Connor, H.R. Wilson. "A review of theories of the L-H transition". Plasma Physics and Controlled Fusion **42** (2000) R1-R74.

[2]. Y.R. Martin, T. Takizuka and ITPA CDBM H-mode Threshold Database Working Group. Journal of Physics : Conference Series 123 (2008) 012033 (12 pp).

[3]. A. Murari, G. Vagliasindi, M. K. Zedda, R. Felton, C. Sammon, L. Fortuna, P. Arena. IEEE Transactions on Plasma Science. **34**, 3 (2006) 1013-1020.

[4]. J. Vega, A. Murari, G. Vagliasindi, G. A. Rattá and JET-EFDA Contributors. Nuclear Fusion. **49** (2009) 085023 (11pp).

[5]. A. J. Meakins, D. C. McDonald and EFDA-JET contributors. Plasma Physics and Controlled Fusion**52** (2010) 075005 (24 pp).

[6]. A.A. Lukianitsa, F. M. Zhdanov, F. S. Zaitsev. Plasma Physics and Controlled Fusion **50** (2008) 065013 (14 pp).

[7]. V. Cherkassky, F. Mulier. "Learning from data". John Wiley & Sons, Inc. (1998).

[8]. J. Ramírez, S. Dormido-Canto, J. Vega and JET-EFDA Contributors. Parallelization of automatic classification systems based on support vector machines: Comparison and application to JET database. Fusion Engineering and Design. **85** (2010) 425-427.

[9]. E. Righi, D.V. Bartlett, J.P. Christiansen, G.D. Conway, J.G. Cordey, L.-G. Eriksson, H.P.L. De Esch, G.M. Fishpool, C.W. Gowers, J.C.M. de Haas, P.J. Harbour, N.C. Hawkes, J. Jacquinot, T.T.C. Jones, W. Kerner, Q.A. King, C.G. Lowry, R.D. Monk, P. Nielsen, F.G. Rimini, G. Saibene, R. Sartori, B. Schunke, A.C.C. Sips, R.J. Smith, M.F. Stamp, D.F.H. Start, K. Thomsen, B.J.D. Tubbing, N. Zornig. "Isotope scaling of the H mode power threshold on JET". Nuclear Fusion **39**, 3 (1999) 309-319.

[10]. S. González, J. Vega, A. Murari, A. Pereira, M. Beurskens, and JET-EFDA contributors, Automatic ELM location in JET using a Universal Multi-Event Locator, Fusion Science and Technology

[11]. J.Vega, A. Murari, S. González and JET-EFDA contributors, A universal support vector machines based method for automatic event location in waveforms and video movies: applications to massive nuclear fusion databases, Review of Scientific Instruments Vol. **81** (2010) 023505 (11 pp) doi: 10.1063/1.3302629

[12]. V. N. Vapnik, Statistical Learning Theory, John Wiley & Sons (1998)

[13]. S. Mallat, A wavelet Tour of Signal Processing, 2nd ed. Academic Press, New York (2001)

[14]. S. González, J. Vega, A. Murari, A. Pereira, J. M. Ramírez, S. Dormido-Canto and JET-EFDA contributors, SVM-based feature extractor for L/H transitions in JET, Review of Scientific Instruments **81**, 10E123 (2010); doi:10.1063/1.3502327

[15]. G.A. Rattá, J. Vega, A. Murari, G. Vagliasindi, M.F. Jonhson, P.C. de Vries and JET-EFDA contributors, An advanced disruption predictor for JET tested in a simulated real time environment, Nuclear Fusion Vol. **50** (2010) 025005 (10 pp) doi: 10.1088/0029-5515/50/2/025005

[16]. C. Saunders, A. Gammerman, and V. Vovk. Transduction with confidence and credibility. In Proceedings of IJCAI'99, volume **2**, pages 722–726, 1999.

[17]. J. Vega, A. Murari, A. Pereira, S. González and I. Pastor. "Accurate and reliable image classification by using conformal predictors in the TJ-II Thomson Scattering". Review of Scientific Instruments **81**, 10E118 (2010) 4pp.
[18]: Alexander Gammerman, Vladimir Vovk, Hedging Predicitions in Machine Learning, The Computer Journal, V.**50** n.2, p 151-163, March 2007
[19]. L. Makili, J. Vega, S. Dormido-Canto, I, Pastor, A. Murari. Computationally efficient SVM multi-class image recognition with confidence measures. Fusion Engineering and Design (in press).

APPENDIX A. LIST OF SIGNALS REQUIRED BY THE SYSTEM

0.1 seconds model: Classification success rate: 96.65%

1. Toroidal Magnetic Field (BT).
2. $D\alpha$ inner view (AD36).
3. Core electron density line average (LAD3).
4. Edge electron density line average (LAD4).
5. Electron temperature at PSI = 0.8, normalised radius (TE08). Calculated from ECE measurements.
6. Plasma inductance (LI).
7. Safety factor at PSI = 0.95, normalised radius (Q95).
8. Radial inner gap (RIG).
9. Top outer gap (TOG).
10. Radiated power (RAD).
11. Electron temperature at PSI = 0.2, normalised radius (TE02). Calculated from ECE measurements.

0.5 seconds model: Classification success rate: 97.72%

1. Toroidal Magnetic Field (BT).
2. $D\alpha$ inner view (AD36).
3. Core electron density line average (LAD3).
4. Edge electron density line average (LAD4).
5. Electron temperature at PSI = 0.8, normalised radius (TE08). Calculated from ECE measurements.

C_j/ B_j	(l/H-0.1,L/H+0.1)	(l/H-0.5,L/H+0.5)
BT (T)	4.13	3.35
LI	4.82	–
Q95	2.07	–
RIG (m)	3.40	–
AD36 (s^{-1} cm^{-2} sr^{-1})	−2.67e-16	−1.16e-16
TOG (m)	−16.62	–
RAD (W)	−6.28e-7	–
TE02 (eV)	−2.44e-3	–
LAD3 (m^{-3})	−1.30e-19	−3.48e-19
LAD4 (m^{-3})	5.07e-19	5.56e-19
TE08 (eV)	1.66e-2	9.84e-3
bias	−11.45	−7.53e-2

Table 1: Coefficients of the first layer hyper-planes.

Model	Training/Proper training set	Test/calibration set
±100ms model	98.62%	96.65%
±500ms model	98.80%	97.72%
Combiner	99.17%	98.73%

Table 2: SVM multi-layer model classification results.

Number of non clear signature transitions	410 shots
Number of properly located transitions (real transition time inside the PCI)	397 shots (96.83 %)
Average(|real L/H – estimated L/H|)	27.9 ms
std(|real L/H – estimated L/H|)	37.6 ms
Average(real L/H – estimated L/H)	12.5 ms
std(real L/H – estimated L/H)	45.1 ms
Average PCI	225.2 ms
Number of not properly located transitions (real transition time outside the PCI)	13 shots (3.17 %)
Average(|real L/H – estimated L/H|)	93.8 ms
std(|real L/H – estimated L/H|)	55.4 ms
Average(real L/H – estimated L/H)	45.7 ms
std(real L/H – estimated L/H)	101.7 ms
Average PCI	453.5 ms

Table 3: Results with discharges that show a non clear transition pattern. It should be noted that the results of the second group are consistent with the average temporal length of the LHTI: 533ms

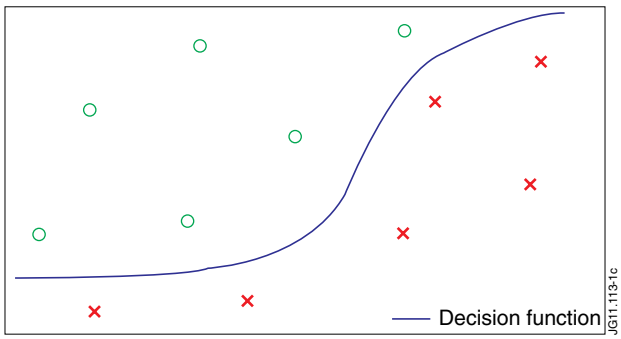

Figure 1: Example of a binary classification problem.

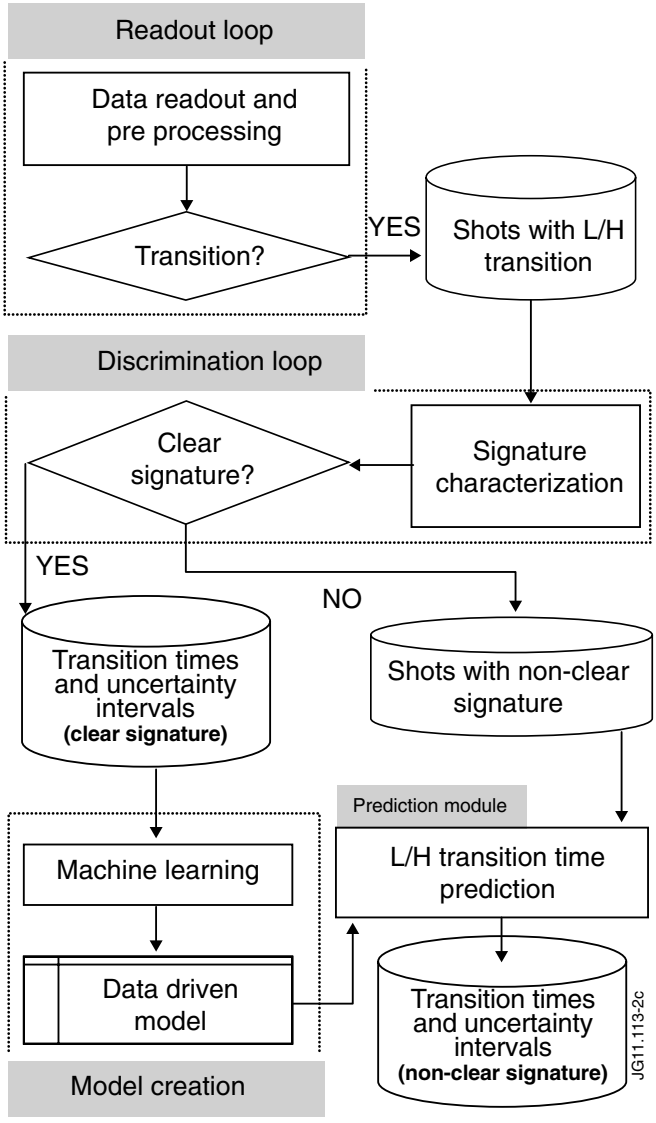

Figure 2: Software modules for the automatic estimation of transition times with their corresponding uncertainty intervals.

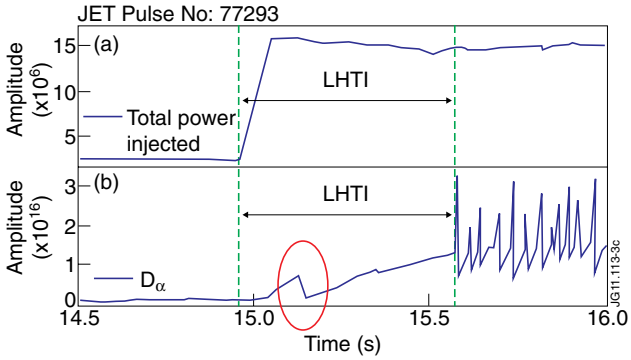

Figure 3: Total injected power (a) and D_α (b) in a L/H transition with a clear signature (drop in the D_α signal).

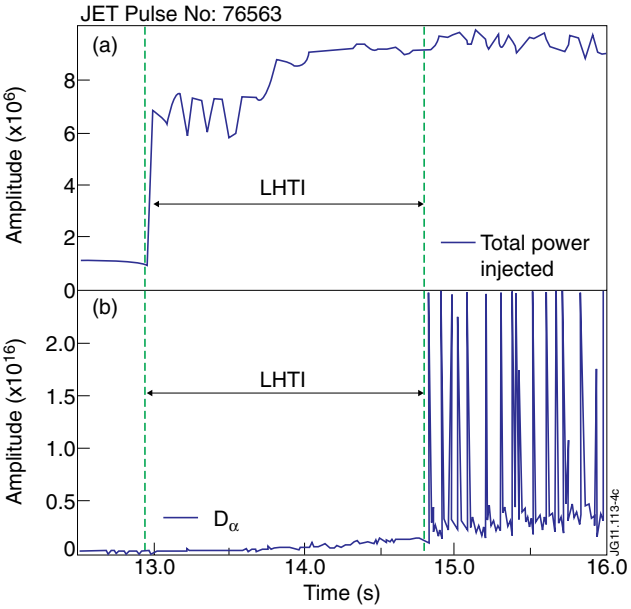

Figure 4: Total injected power (a) and D_α (b) in a L/H transition with a non-clear signature.

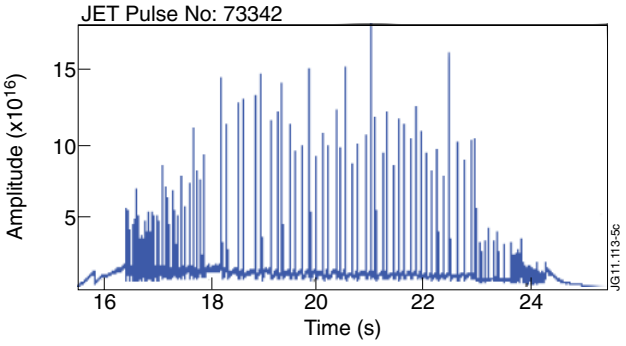

Figure 5: D_α signal showing the typical pattern of ELMs.

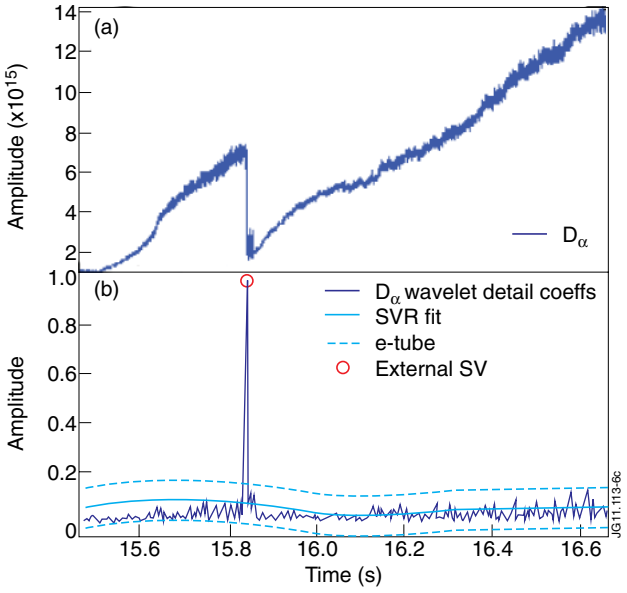

Figure 6: (a) L/H transition with a clean signature in the LHTI (JET Pulse No: 73337). (b) UMEL locates the peak by means of an external support vector [11].

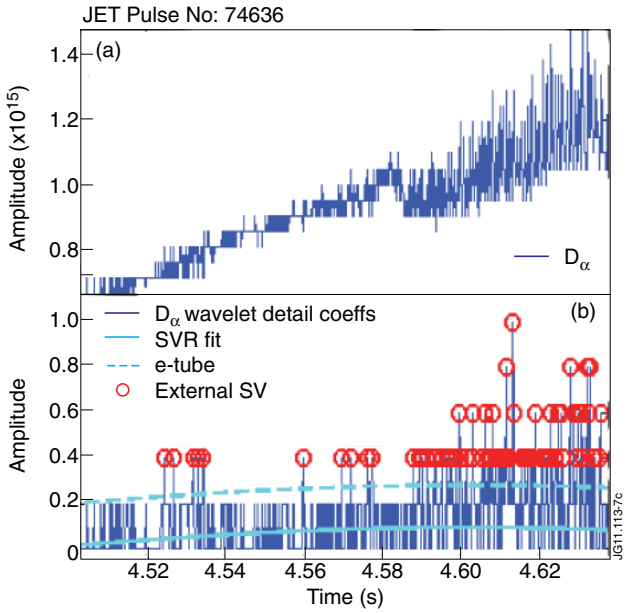

Figure 7: (a) L/H transition with a non clean signature (JET Pulse No: 74636). (b) UMEL does not locate a single peak.

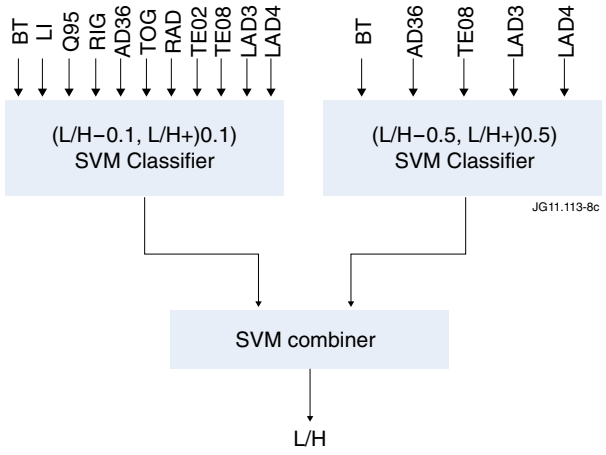

Figure 8: SVM two-layer classifier.

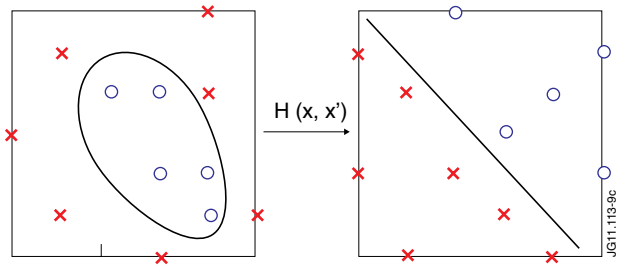

Figure 9: SVM technique. H(x,x') is the kernel function.

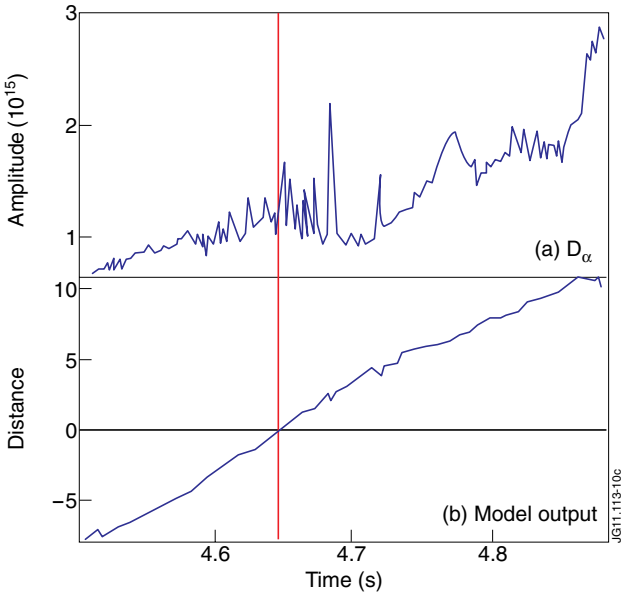

Figure 10: Discharge with a non-clear signature. The time instant when the distance is 0 is the transition time.

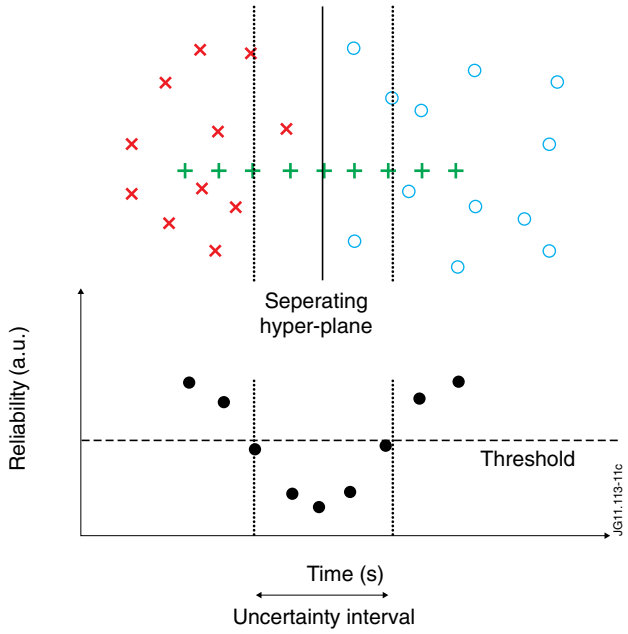

Figure 11: Training samples are 'x' and 'o'. Samples to classify are '+'. The bottom plot represents the reliability of the samples to classify.

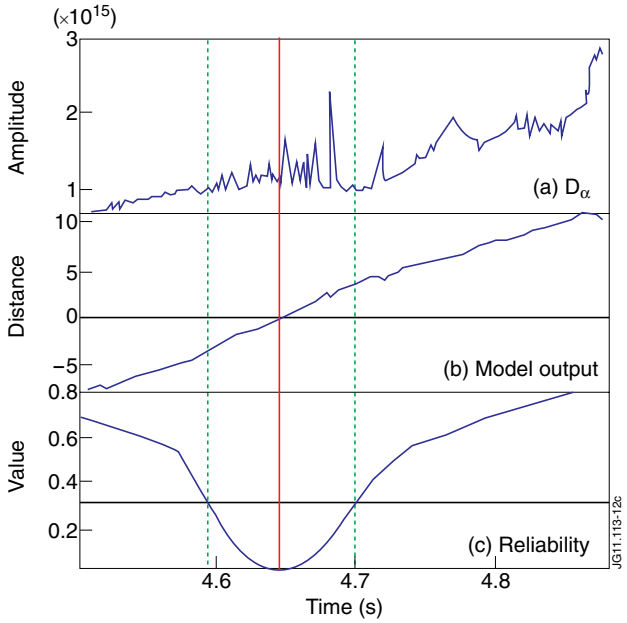

Figure 12: By establishing a reliability threshold of 0.3, the uncertainty interval is 104ms.

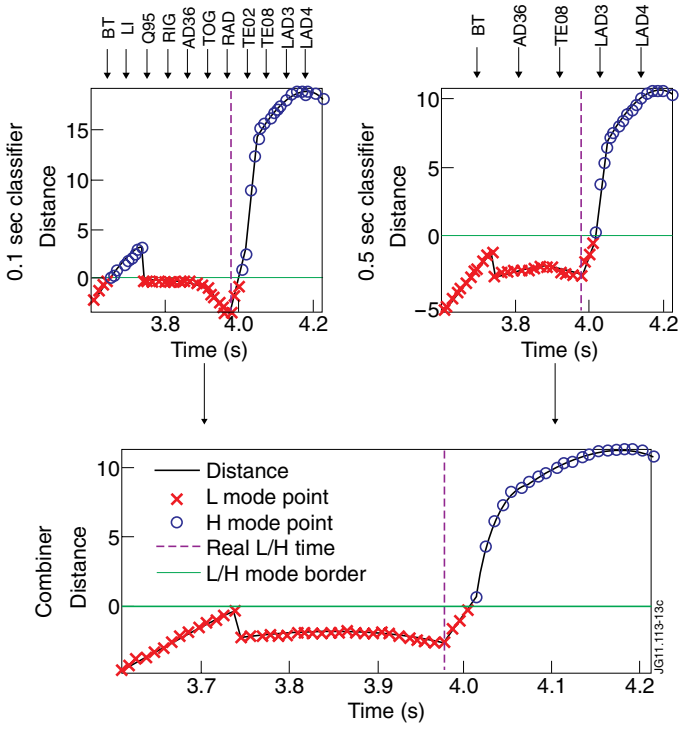

Figure 13: Example of the two-layer classifier in JET Pulse No: 74570. The real transition, t = 3.997s ,is represented by the dashed red vertical line. All times in s.

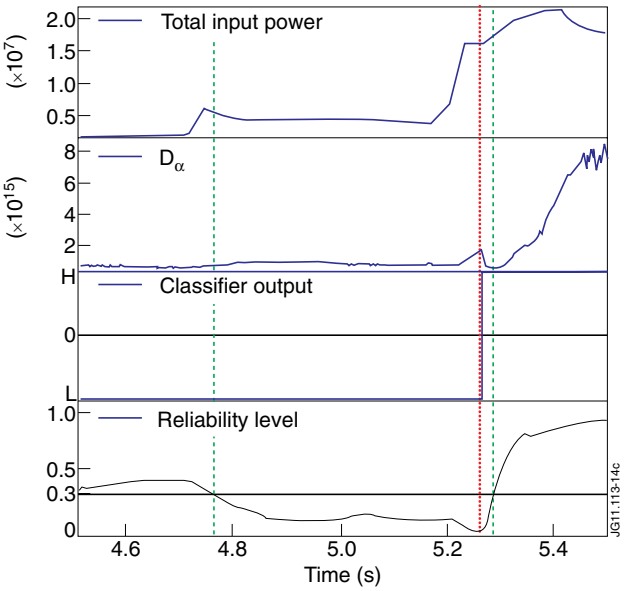

Figure 14: Accurate L/H transition time determination of JET Pulse No: 76493. The classifier output is correct and the vertical dashed lines show the uncertainty interval.

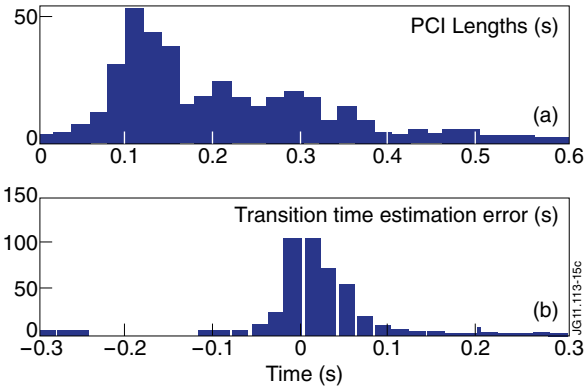

Figure 15: Distribution of the PCI (top) and the average(real L/H – estimated L/H) (bottom).

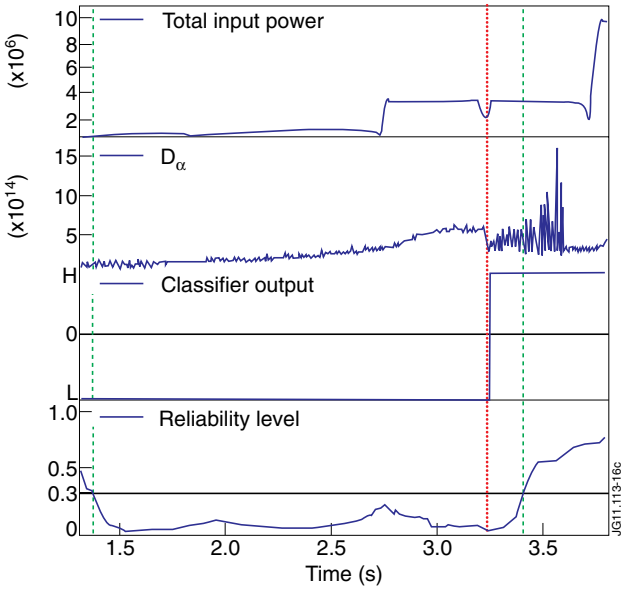

Figure 16: Example of discharge with moderate input power and a long step before injecting the total power.

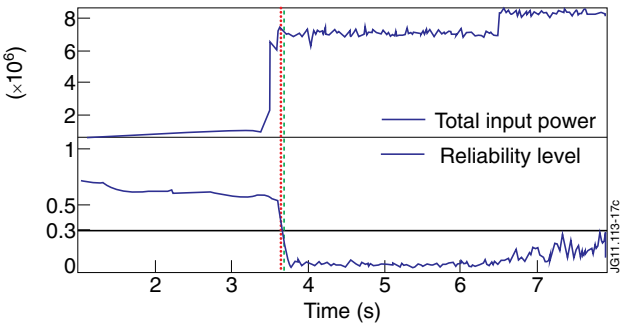

Figure 17: Example of the second group of discharges. The real transition takes place 12ms before the left extreme of the uncertainty interval.

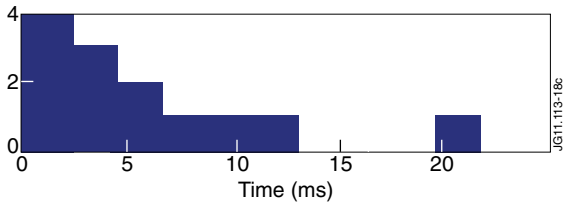

Figure 18: Distribution of times between the real transition and the closest extreme of the uncertainty interval for the second group of discharges in table 3.

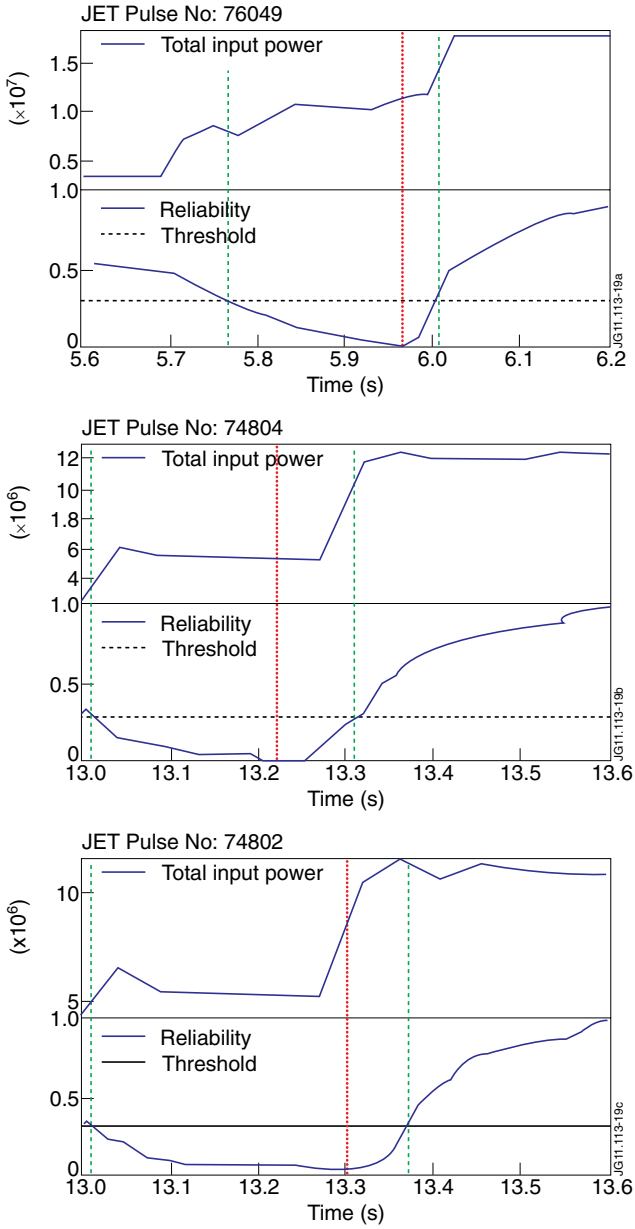

Figure 19: Different uncertainty interval lengths and its relation to the input power.

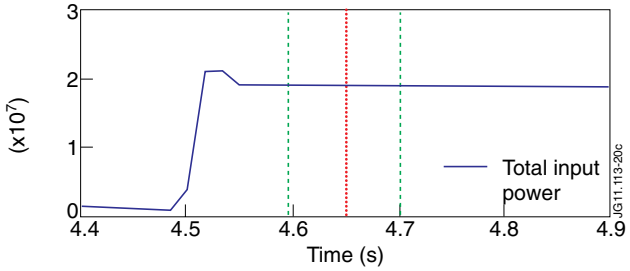

Figure 20:Abrupt power injection seems to generate narrower uncertainty intervals.