EFDA–JET–CP(11)03/04

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

H/L Transition Time Estimation in JET Using Conformal Predictors

H/L Transition Time Estimation in JET Using Conformal Predictors

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, Spain
[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 Proceedings of the
8th IAEA Technical Meeting on Control, Data Acquisition and Remote Participation for Fusion Research,
San Francisco, USA
(20th June 2011 - 24th June 2011)

ABSTRACT
Recent advances in data mining allow the automatic recognition of physical phenomena in the databases of fusion devices without human intervention. This is important to create large databases of physical events (thereby increasing the statistical relevance) in an unattended manner. Important examples are the L/H and H/L transitions. In this contribution, a novel technique is introduced to automatically locate H/L transitions in JET by using conformal predictors. The focus is on H/L transitions because typically there is not a clear signature in the time series of the most widely available signals to recognize the change of confinement. Conformal predictors hedge their prediction by means of two parameters: confidence and credibility. The technique has been based on binary supervised classifiers to separate the samples of the respective confinement modes. Results with several underlying classifiers are presented.

1. INTRODUCTION
The H-mode (high-energy mode) is one of the main confinement regimes in present and future tokamaks and stellerators. The sudden variation of the plasma parameters from the L-mode (low-energy mode) to the H-mode is known as L/H transition. The L/H transition is characterized by the creation of an Edge Transport Barrier (ETB). When the ETB is lost, the plasma returns to L-mode. This transition is known as the H/L transition.

The H mode was firstly detected in the ASDEX tokamak [1]. The transition to access the H mode can be recognized using the Dα emission signal. A fast drop of this signal is observed between the two confinement modes. Previous effort has been done to develop automatic methods to recognize the L/H transition [2,3] using exiting databases of L/H transition times.

The case of the H/L transition is more complex. Typically, there is not a clear signature in the time series of the most widely available signals to recognize the change of confinement. Previous work in the automatic recognition of the H/L transition [2,3] uses data around an H/L transition (previously located by experts) to train a system capable of recognizing H/L transitions in new pulses. The problem of this methodology is the lack of large databases of H/L transition times and the problematic of the manual location of H/L transitions in new pulses (intensive expert knowledge is required).

Two different approaches to the automatic location of H/L transitions using machine learning methods (MLMs) have been developed: on the one hand, a model is trained using a database of L/H transition times. On the other hand, a model is trained using an H/L transition times database. Support Vector Machines (SVM) has been used to build these models.

SVM is a MLM to build predictive models in classification problems [4]. It is particularly advantageous in the case of high-dimensional data. SVM maximizes the distances between the samples of different classes.

Most MLMs do not provide measures of their reliability. In contrast, conformal predictors [5] give measures of confidence and credibility for each prediction. Confidence means how reliable the classification of the predictor is and credibility shows how representative is the training set for the new sample to classify. In this paper, a conformal version of SVM has been implemented.

2. PREDICTION OF H/L TRANSITION TIMES USING A L/H TRANSITION TIME DATABASE

A set of 355 JET pulses has been manually analyzed by experts to locate the L/H and H/L transition times. It has been divided into a training set (100 pulses -80% proper training set, 20% calibration set-) and a test set (255 pulses).

An off-line version of Inductive Conformal Predictors (ICPs) [4] has been selected as the most suitable to predict H/L transitions. ICPs use the proper training set to build a SVM model so the bulk of the computations is performed only once. Conformal measures are computed using the samples in the calibration set. In the on-line version, new samples classified by the model are included into the calibration set. In contrast, in the off-line version new samples are not included in the calibration set and therefore the conformal measures of a sample are the same despite the ones that have been classified before.

The models have been trained using the 11 most relevant signals for the L/H transition. These signals have been selected from a set of 27 signals. The list of signals and the description of the method is reported in [6].

Two different types of kernels have been used to build models using SVM: linear and Radial-Basis Function (RBF) kernels. 5 different values of the C parameter have been tested for the linear kernel. For the RBF kernel, all the combinations of 21 values of the C parameter and 21 values of the σ parameter have been tested (441 combinations).

The SVM models have been trained using an interval of ±0.5 seconds around the L/H transitions. They have been tested using the complete temporal evolution of the pulses in the test set. The best results using a RBF kernel have been obtained using the parameters $C = 5500$, $\sigma = 3$ (test success rate -TSR-= 94.82 %). The linear kernel obtains the best results using the parameter $C = 100$ (TSR: 89.24 %).

The SVM models have been applied to the prediction of L/H and H/L transitions of the test set. The absolute mean errors (*mean(|treal-tpredicted|)*) obtained in the prediction of the L/H transition are: 22ms (linear kernel) and 15ms (RBF kernel).

These models have been also used to predict H/L transitions. The absolute mean errors obtained are: 236ms (linear kernel) and 203ms (RBF kernel). The distribution of the mean errors using a model trained with L/H data can be seen in Fig.1a).

3. PREDICTION OF L/H TRANSITION TIMES USING A H/L TRANSITION TIME DATABASE

The same process described above has been repeated training the models with a H/L transitions time database.

Using a RBF kernel with the parameters $C = 100$, $\sigma = 7.5$ (TSR = 97.60%), the mean absolute error obtained in the prediction of the H/L transition is 10ms. The error in the prediction of the L/H transition is 33ms. Fig.1b) shows the distribution of these errors. In the case of a linear kernel (C = 100, TSR = 97.70%) the error obtained in the prediction of the H/L transition is 102 ms and the one obtained in the prediction of the L/H transition is 28 ms.

4. CONFORMAL MEASURES

Conformal measures, i.e. confidence and credibility, give key information about the accuracy and reliability of the predictions made by the different models. The mean values of the confidence and credibility of the test samples are shown in Table 1.

It can be seen that the credibility values of the models trained using the H/L data are significantly larger that the ones of the models trained with L/H data. It means that the training data used in these cases is more suitable to make predictions. As a consequence, the errors in the predictions of the H/L increase by a factor of 2 comparing to the errors of the H/L trained models (Table 1).

CONCLUSIONS

The lack of reliability to predict one type of transition using data of the other is a clear statistical conclusion of the present analysis. This result has been indeed observed in two directions: the models trained with L/H (H/L) data obtain a mean error predicting H/L (L/H) transitions that is the double of the one obtained using systems trained with H/L (L/H) data (Table 1). One possible explanation of the observed behaviour could be the hysteresis of the L/H – H/L transition [3] but this point must be investigated with more adapted statistical tools. In any case, conformal measures have proven their usefulness to validate data-driven models.

ACKNOWLEDGMENTS

This work was partially funded by the Spanish Ministry of Science and Innovation under the Project No. ENE2008-02894/FTN. This work was supported by EURATOM and 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] F. Wagner, G. Becker, K. Behringer, D. Campbell, A. Eberhagen, W. Engelhardt et al. Phys. Rev. Lett. 49 (1982) 1408-1412.

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

[3] A. J. Meakins, D. C. McDonald and JET-EFDA contributors et al. Plasma Phys. Control. Fusion 52 (2010) 075005 24 pp

[4] V. Vapnik, Statistical Learning Theory, John Wiley & Sons, INC. 1998

[5] A. Gammerman and V. Vovk, The Computer Journal 50 (2) (2007) 151-163 13 pp

[6] S. González, J. Vega, A. Murari, A. Pereira, J. M. Ramírez, S. Dormido-Canto and JET-EFDA contributors, Review of Scientific Instruments 81 (2010) 10E123 3 pp

		Linear	RBF
L/H training	Confidence	99.86 %	99.84 %
	Credibility	64.10 %	66.85 %
	L/H abs. mean error	22 ms	15 ms
	H/L abs. mean error	236 ms	203 ms
H/L training	Confidence	97.02 %	97.53 %
	Credibility	73.25 %	73.77 %
	L/H abs. mean error	28 ms	33 ms
	H/L abs. mean error	102 ms	106 ms

Table 1: Experiment's results summary

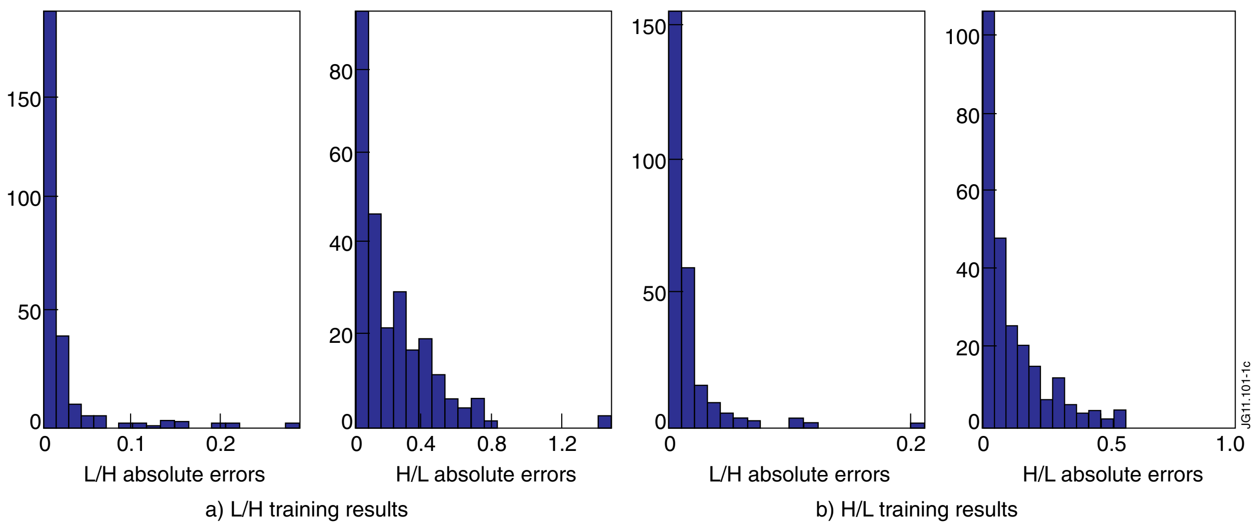

Figure 1. Distributions of absolute mean errors