Fusion Engineering and Design 123 (2017) 717–721

Contents lists available at ScienceDirect

Fusion Engineering and Design

journal homepage: www.elsevier.com/locate/fusengdes

Classification of ELM types in Joint European Torus based on global plasma parameters using discriminant analysis

Aqsa Shabbir[a], Gregoire Hornung[b], Jean-Marie Noterdaeme[b,c], Geert Verdoolaege[b,d,*], and JET contributors[1]

[a] Department of Electrical Engineering, Lahore College for Women University, Lahore, Pakistan
[b] Department of Applied Physics, Ghent University, B-9000 Ghent, Belgium
[c] Max Planck Institute for Plasma Physics, D-85748 Garching, Germany
[d] Laboratory for Plasma Physics, Royal Military Academy, B-1000 Brussels, Belgium

ARTICLE INFO

Article history:
Received 3 October 2016
Received in revised form 19 May 2017
Accepted 22 May 2017
Available online 20 June 2017

Keywords:
Edge-localized modes
Discriminant analysis
Classification
JET

ABSTRACT

In this work, discriminant analysis is used as the main approach for building a physics based automated classifier for the discrimination of the edge-localized mode (ELM) plasma instability. The classifier is then applied for distinguishing type I and type III ELMs from a set of carbon-wall plasmas at JET. This provides a fast, standardized classification of ELM types which is expected to significantly reduce the effort of ELM experts in identifying ELM types. Further, the classifier yields a separation hyperplane in terms of global plasma parameters, which provides an insight into the range of conditions under which specific ELM behaviors occur.

1. Introduction

Edge-localized modes (ELMs) are magnetohydrodynamic instabilities occurring in the edge region of high confinement fusion plasmas. ELMs lead to the ejection of energy and particles from the plasma core and onto the plasma-facing components (PFCs). While they are beneficial for impurity regulation, in future devices, such as ITER, large unmitigated ELMs will lead to intolerable heat loads on the PFCs.

A first characterization of ELMs is the identification of their type. In this work, a statistical method, discriminant analysis (DA) is employed for developing a simple predictive algorithm for distinguishing ELM types. As an application of our analysis, we discriminate between type I and type III ELMs in a set of carbon-wall (CW) plasmas from the Joint European Torus (JET) tokamak. Previously, several efforts have been made to statistically characterize [1,2] and provide an automated classification scheme for ELMs [3–5]. Herein, the advantage with respect to earlier ELM classification works is twofold. First, we rely on routinely measured global plasma parameters. Secondly, compared to other classifiers that may be more accurate, DA yields a separation hyperplane between ELM types. Thus we obtain an analytical expression enabling verification and prediction of the ELM regime in terms of these parameters, as well as quantification of the significance of each parameter. Our approach is intended as a tool to support planning and analysis of experiments.

2. Discriminant analysis

In this section, background on the application of DA [6] to the discrimination of ELM types is given. DA enables prediction of the class membership (ELM type) based on a linear or a quadratic combination of plasma parameters. DA is a parametric method and assumes that the distribution of the plasma parameters within each class is multivariate normal.

In this work, two classes ($k = 1, 2$) of plasmas with type I and type III ELMs are considered. Each class k with n_k plasmas is denoted by a $n_k \times p$ data matrix, where p is the number of plasma parameters and plasma $\mathbf{x}$ is a $(1 \times p)$ vector from this matrix. The class-specific probability density of $\mathbf{x}$ belonging to class $k = r$ is denoted as $f_r(\mathbf{x})$. Further, π_r denotes the prior probability of plasma $\mathbf{x}$ belonging to

* Corresponding author at: Department of Applied Physics, Ghent University, B-9000 Ghent, Belgium.
E-mail address: geert.verdoolaege@ugent.be (G. Verdoolaege).

[1] See the author "list of Overview of the JET results in support to ITER" by X. Litaudon et al. to be published in Nuclear Fusion Special issue: overview and summary reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17–22 October 2016).

http://dx.doi.org/10.1016/j.fusengdes.2017.05.101
0920-3796/

class r, with $\Sigma_{r=1}^{k}\pi_r = 1$. The posterior probability of a plasma $\mathbf{x}$ belonging to class $k = r$ is obtained by applying Bayes' theorem:

$$P(r|\mathbf{x}) = \frac{f_r(\mathbf{x})\pi_r}{\Sigma_{s=1}^{k} f_s(\mathbf{x})\pi_s}. \quad (1)$$

The denominator is consistent across all classes; hence it suffices to estimate class-specific densities $f_r(\mathbf{x})$ for each of the classes. It follows that we classify $\mathbf{x}$ in class r if $f_r(\mathbf{x})\pi_r$ is maximal. Each of the class densities is modeled as a multivariate normal density:

$$f_r(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}\sqrt{|\Sigma_r|}} \exp\left(-\frac{1}{2}d_r^2(\mathbf{x})\right). \quad (2)$$

The *Mahalanobis* distance of a plasma $\mathbf{x}$ to class r is given as

$$d_r(\mathbf{x}) = \sqrt{(\mathbf{x}-\mu_r)^t(\Sigma_r)^{-1}(\mathbf{x}-\mu_r)}. \quad (3)$$

2.1. Linear discriminant analysis (LDA)

All classes are considered to be sharing a common covariance matrix. Hence, $\Sigma_r = \Sigma$ for all classes r. Taking the logarithm of $f_r(\mathbf{x})\pi_r$ and after simplifying we obtain for each class, the class scores $l_r(\mathbf{x})$ are given by

$$l_r(\mathbf{x}) = \mathbf{x}^t\Sigma^{-1}\mu_r - \frac{1}{2}(\mu_r)^t\Sigma^{-1}\mu_r + \log(\pi_r). \quad (4)$$

The score $l_r(\mathbf{x})$ is a linear function of $\mathbf{x}$ and the decision boundary between two classes is the collection of points $\mathbf{x}$ for which $l_r(\mathbf{x}) = l_s(\mathbf{x})$. In p dimensions the boundary between two classes is thus a hyperplane. The class centers and the common covariance matrix for the classes are estimated from the plasmas for which the ELM type is known (training data). The standard estimates are:

$$\mu_r = \frac{1}{n_r}\sum_{k_i=r}\mathbf{x}_i, \quad (5)$$

$$\Sigma = \frac{1}{n-k}\sum_{r=1}^{k}\sum_{k_i=r}(\mathbf{x}_i-\mu_r)(\mathbf{x}_i-\mu_r)^t. \quad (6)$$

Prior class probabilities are also estimated from the data and are given as:

$$\pi_r = \frac{n_r}{n}, \quad (7)$$

where n_r is the number of plasmas belonging to class r and n is the total number of plasmas in the data.

2.2. Quadratic discriminant analysis (QDA)

Classes do not share a common covariance matrix and the class score $q_r(\mathbf{x})$ is a quadratic function of $\mathbf{x}$. Hence, the decision boundary between any two classes $q_r(\mathbf{x}) = q_s(\mathbf{x})$ is also quadratic. Again, the covariance matrix for each class is estimated by the sample covariance matrix of the training samples in that class.

3. Classification of ELM types

3.1. Dataset

A dataset comprising 74 type I and 26 type III ELMy plasmas spanning over the shot range [50564–76483] was assembled from the JET CW experiments. This is an extension of the data set used earlier by Webster et al. [1] and is the same dataset that was used earlier for the visualization of the tokamak operational space in [7]. The analysis, in this work, has been restricted to time intervals in which the plasma conditions are quasistationary with approximately constant heating, gas fueling and central density. Further, all experiments dealing with ELM control and mitigation techniques have been excluded. The global plasma parameters considered herein are: vacuum toroidal field at $R = 2.96$ (B_t, T), plasma current (I_p, MA), line integrated edge density (n_e, 10^{19} m^{-2}), gas fueling (Γ_{D_2}, 10^{22}s^{-1}), input power (P_{input}, MW) and average triangularity (δ_{avg}). Histograms of plasma parameters for each class are presented in Fig. 1. From a visual inspection of Fig. 1 it can be observed that while B_t, I_p and n_e appear to have symmetric distributions, Γ_{D_2}, in particular, deviates from the normality assumption. However, the normality assumption is only a criterion for optimality and LDA works reasonably well even if this assumption is violated [8,9]. Fig. 1 also indicates a non-trivial classification problem as considerable overlap amongst the two classes can be readily observed.

Table 1
Predictive capability of individual plasma parameters using linear discriminant analysis.

Plasma parameters	Leave-one-out CV success rate (%)	DV
B_t (T)	73.0	2.35
I_p (MA)	74.0	2.34
n_e (10^{19} m^{-2})	74.0	6.45
P_{input} (MW)	82.0	13.7
Γ_{D_2} (10^{22} s^{-1})	81.0	2.99
δ_{avg}	74.0	0.384

3.2. Performance assessment

Leave-one-out cross validation is used for assessing generalization capability of the classifier. For a dataset with N plasmas, N iterations are performed where $N-1$ plasmas are used for training and the remaining sample is used for testing. The leave-one-out cross validated (CV) success rate, which here is quoted as the percentage of ELMy plasmas correctly classified, is in effect, an estimation of the expected performance of the classifier on an unknown independent dataset.

3.3. Predictive capability of each plasma parameter

LDA is performed on the plasmas represented by each global plasma parameter individually. In this case, the estimated covariance matrices coincide with the variances of the two classes and the discriminant function is reduced to a discriminating value (DV). This DV, derived under the assumption of equal variances, is then applied for classification. The DV is given as:

$$DV = \frac{1}{2}(\mu_{class\ 1} + \mu_{class\ 2}). \quad (8)$$

The leave-one-out CV success rates (%) and DVs are presented in Table 1. Under the assumption of unequal variances (QDA), the analysis produces success rates similar to those presented in Table 1 (differences are $\sim 1\%$). Further, Table 1 reveals that the parameters P_{input} and Γ_{D_2} yield the highest success rates and hence may play the main role in the classification between the two classes. This corresponds with established knowledge regarding ELM occurrence.

3.4. Combinations of plasma parameters

DA is then performed on the linear and quadratic combinations of the plasma parameters, in order to further improve the success rate. The average and class-wise leave-one-out CV success rates are given in Table 2. It can be noted that a linear combination of P_{input} and Γ_{D_2} improves the average leave-one-out CV success rate to 91.0% from (81.0–82.0)% yielded by each of them individually. On the other hand, a quadratic combination of P_{input} and Γ_{D_2} increases the average success rate to 89.0%. This is further illustrated in Fig. 2. It can be readily observed that the vertical and horizontal dashed

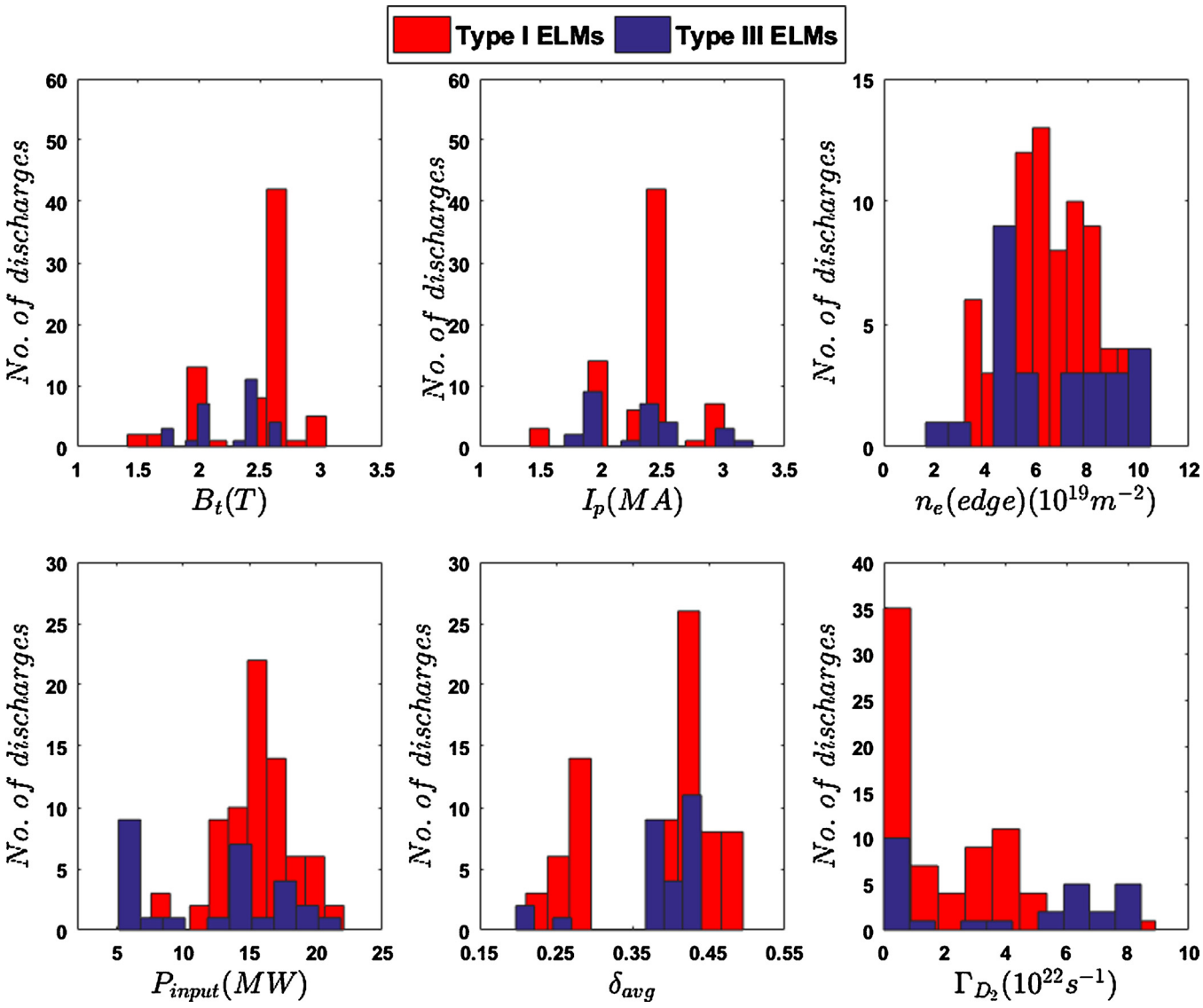

Fig. 1. Histograms of plasma parameters in the analyzed dataset.

lines discriminate the two classes poorly, whereas the solid lines, which are a function of P_{input} and Γ_{D_2}, better separate the two classes. Further, it can be seen that for 10.4 MW $\leq P_{input} \leq$ 16.5 MW the difference between the quadratic and linear boundary is small ($\Delta\Gamma_{D_2} \leq 1.0\,s^{-1}$). However, for $P_{input} >$ 16.5 MW, this difference is substantial.

Fig. 3, presents the decrease in error rate (%) with the addition of other plasma parameters. An addition of the remaining 4 plasma parameters, B_t, I_p, n_e and δ_{avg} to P_{input} and Γ_{D_2} reduces the average error rate to 8% (alternatively, average success rate improves to 92%) for the linear combination of parameters and to 6% for the quadratic case. While the addition of Γ_{D_2} to P_{input} had reduced the error rate by a factor of ~2, the addition of the remaining 4 parameters only lowers it further by 1% for LDA and 5% for QDA. It is noteworthy that the error rate for type III ELMs reduces by ~4% for both LDA and QDA whereas the error rate for type I ELMs remains unchanged for LDA and lowers by ~ 5% for QDA. However, this reduction in error rates comes at the expense of an increased model complexity brought about by an increase in the number of parameters in the discriminant function [10].

From the various models analyzed, the linear combination of P_{input}, Γ_{D_2}, B_t, I_p along with either n_e or δ_{avg}, can be considered best models as they yield high average and class-wise success rates with the least number of parameters. The quadratic combination of P_{input}, Γ_{D_2}, B_t, I_p and n_e gives a slightly higher success rate amongst all analyzed models. However, the quadratic model is significantly more complex, less intuitive and less tractable, than the linear counterpart. However, if the primary goal is correct classification of a new plasma, then this quadratic model can be slightly advantageous compared to the linear ones.

3.5. Separation hyperplane for type I and type III ELMs

The mathematical form for the linear discriminant functions derived for the classification of type I and III ELMs is presented in Table 3. The classification success rates for these linear separating hyperplanes (boundary) are provided in Table 2. For each of the discriminant functions, given in Table 3, type III ELMs are expected if the left-hand side of the expression is less than the constant on the right-hand side. Otherwise, type I ELMs are expected. Further, the goodness-of-fit of each discriminant function is indicated by the Wilks' Λ test statistic [11]. Wilks' Λ is the ratio of within class variability to the total variability in the discriminator variables. A small value (closer to 0) indicates that almost all of the variability

Table 2
Average and class-wise (type I and type III ELMy plasmas) leave-one-out CV success (%) for a linear and quadratic combination of plasma parameters obtained by LDA and QDA, respectively.

Plasma parameters		Leave-one-out CV success (%)		
		I	III	Avg
P_{input}, Γ_{D_2}	LDA	94.6	80.8	91.0
	QDA	90.5	84.6	89.0
P_{input}, Γ_{D_2}, I_p	LDA	94.6	76.9	90.0
	QDA	91.9	80.8	89.0
P_{input}, Γ_{D_2}, δ_{avg}	LDA	94.6	80.8	91.0
	QDA	91.9	73.1	87.0
P_{input}, Γ_{D_2}, n_e	LDA	93.2	76.9	89.0
	QDA	90.5	80.8	88.0
P_{input}, Γ_{D_2}, B_t	LDA	90.5	80.8	88.0
	QDA	93.2	84.6	91.0
P_{input}, Γ_{D_2}, δ_{avg}, B_t, I_p	LDA	94.6	84.6	92.0
	QDA	94.6	88.5	93.0
P_{input}, Γ_{D_2}, n_e, B_t, I_p	LDA	94.6	84.6	92.0
	QDA	94.6	92.3	94.0
P_{input}, Γ_{D_2}, δ_{avg}, B_t, I_p, n_e	LDA	94.6	84.6	92.0
	QDA	96.0	88.5	94.0

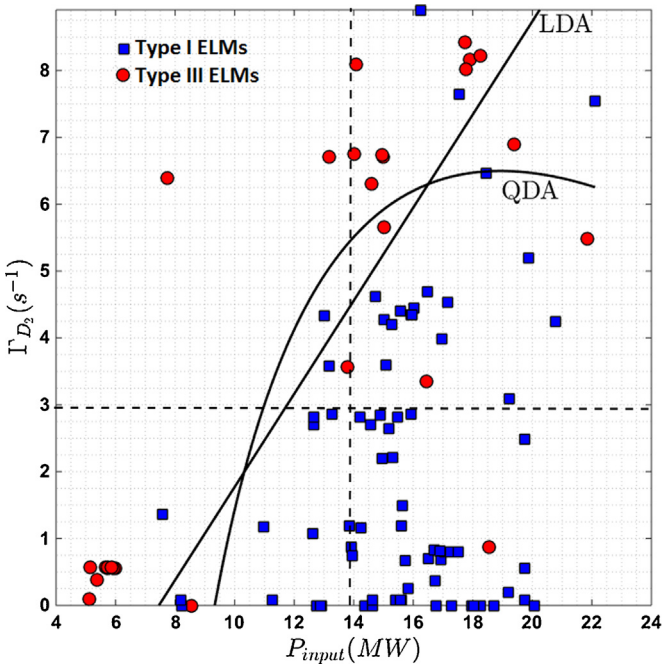

Fig. 2. The solid line and curve indicate the linear (LDA) and the quadratic (QDA) discriminant functions for type I and type III ELMs from the analysed plasmas. Vertical and horizontal dashed lines mark the discriminating values for P_{input} and Γ_{D_2}, respectively.

Table 3
Linear separation hyperplanes (boundary) for type I/III ELMs, in terms of global plasma parameters. Type III ELMs are expected if the left-hand side of the expression is less than the constant on the right-hand side. The corresponding classification success rates (%) are provided in Table 2.

	Linear discriminant functions	Wilks' Λ
L1	$P_{input} - 1.41\Gamma_{D_2} = 7.47$	0.60
L2	$P_{input} - 1.25\Gamma_{D_2} + 7.06B_t - 8.81I_p + 0.70n_e = 8.75$	0.53
L3	$P_{input} - 0.765\Gamma_{D_2} + 12.4B_t - 10.7I_p - 26.1\delta_{avg} = 3.96$	0.47

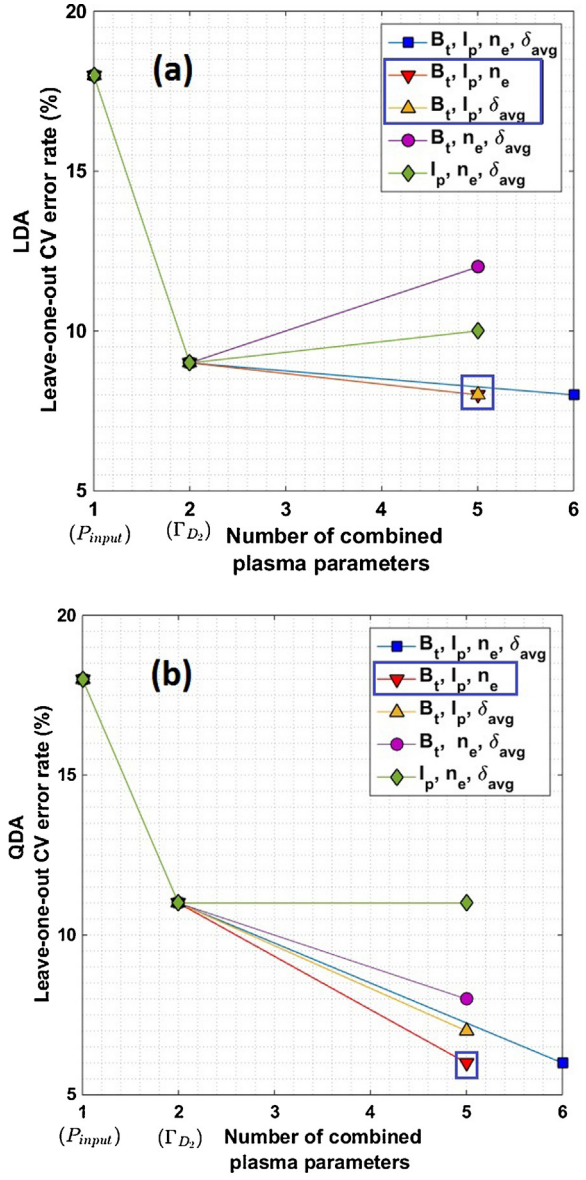

Fig. 3. Leave-one-out CV error rate versus the number of combined plasma parameters using (a) LDA, (b) QDA.

captured by the discriminant function is due to the class differences. It can be observed from Table 3 that while each discriminant function appears to provide a reasonably good fit, discriminant function L3 provides the best separation of the ELMy plasmas.

4. Conclusions

In this work, a simple, high-accuracy, standardized automated classifier has been presented which can considerably reduce the effort of ELM experts in identifying ELM types. Further, the classifier provides a separation hyperplane in terms of plasma parameters which reflects underlying physics and can also aid in determining the operational boundaries for ELMy regimes during experimental planning.

The future work will involve an expansion of the dataset as well as the use of normalized global plasma parameters for rendering a machine independent classifier of ELM types.

Acknowledgements

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 under grant agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] A.J. Webster, R.O. Dendy, J.-E. contributors, Statistical characterisation and classification of edge localised plasma instabilities, Phys. Rev. Lett. 110 (15) (2013) 155004.
[2] A. Murari, F. Pisano, J. Vega, B. Cannas, A. Fanni, S. Gonzalez, M. Gelfusa, M. Grosso, Extensive statistical analysis of ELMs on JET with a carbon wall, Plasma Phys. Controll. Fus. 56 (11) (2014) 114007, http://dx.doi.org/10.1088/0741-3335/56/11/114007.
[3] N. Duro, R. Dormido, J. Vega, S. Dormido-Canto, G. Farias, J. Sánchez, H. Vargas, A. Murari, Automated recognition system for ELM classification in JET, Fus. Eng. Des. 84 (2–6) (2009) 712–715, http://dx.doi.org/10.1016/j.fusengdes.2008.12.003.
[4] A. Shabbir, G. Verdoolaege, J. Vega, A. Murari, ELM regime classification by conformal prediction on an information manifold, IEEE Trans. Plasma Sci. 43 (12) (2015) 4190–4199, http://dx.doi.org/10.1109/TPS.2015.2489689.
[5] A. Shabbir, G. Hornung, G. Verdoolaege, A classification scheme for edge-localized modes based on their probability distributions, Rev. Sci. Instrum. 87 (11) (2016) 11D404, http://dx.doi.org/10.1063/1.4955479.
[6] O.J.W.F. Kardaun, Classical Methods of Statistics, Springer-Verlag, Berlin/Heidelberg, 2005, http://dx.doi.org/10.1007/3-540-29288-8.
[7] A. Shabbir, G. Verdoolaege, O.J.W.F. Kardaun, J.M. Noterdaeme, Visualization of the operational space of edge-localized modes through low-dimensional embedding of probability distributions, Rev. Sci. Instrum. 85 (11) (2014) 1–5, http://dx.doi.org/10.1063/1.4892866.
[8] T. Li, S. Zhu, M. Ogihara, Using discriminant analysis for multi-class classification: an experimental investigation, Knowl. Inf. Syst. 10 (4) (2006) 453–472, http://dx.doi.org/10.1007/s10115-006-0013-y.
[9] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 1998, http://dx.doi.org/10.1007/BF01237942.
[10] K. Burnham, D. Anderson, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, vol. 172, 2nd ed., 2002, http://dx.doi.org/10.1016/j.ecolmodel.2003.11.004.
[11] K. Mardia, J. Kent, J. Bibby, Multivariate Analysis, adsabs.harvard.edu.