A statistical methodology to derive the scaling law for the H-mode power threshold using a

large multi-machine database

A statistical methodology to derive the scaling law for the H-mode power threshold using a large multi-machine database

A. Murari[1]**, I. Lupelli**[2]**, P. Gaudio**[2]**, M. Gelfusa**[2] **and J. Vega**[3]

[1] Associazione EURATOM-ENEA per la Fusione, Consorzio RFX, 4-35127 Padova, Italy
[2] Associazione EURATOM-ENEA - University of Rome 'Tor Vergata', Roma, Italy
[3] Asociación EURATOM-CIEMAT para Fusión, CIEMAT, Madrid, Spain

Abstract

In this paper, a refined set of statistical techniques is developed and then applied to the problem of deriving the scaling law for the threshold power to access the H-mode of confinement in tokamaks. This statistical methodology is applied to the 2010 version of the ITPA International Global Threshold Data Base v6b(IGDBTHv6b). To increase the engineering and operative relevance of the results, only macroscopic physical quantities, measured in the vast majority of experiments, have been considered as candidate variables in the models. Different principled methods, such as agglomerative hierarchical variables clustering, without assumption about the functional form of the scaling, and nonlinear regression, are implemented to select the best subset of candidate independent variables and to improve the regression model accuracy. Two independent model selection criteria, based on the classical (Akaike information criterion) and Bayesian formalism (Bayesian information criterion), are then used to identify the most efficient scaling law from candidate models. The results derived from the full multi-machine database confirm the results of previous analysis but emphasize the importance of shaping quantities, elongation and triangularity. On the other hand, the scaling laws for the different machines and at different currents are different from each other at the level of confidence well above 95%, suggesting caution in the use of the global scaling laws for both interpretation and extrapolation purposes.

(Some figures may appear in colour only in the online journal)

1. Introduction: scaling laws and regime transitions in tokamaks

Scaling laws are a central tool in physics, engineering, biology and finance and have been extensively used since the beginning of modern science in association with multivariate regression techniques, mainly for interpretation and extrapolation purposes of complex problems starting from experimental evidence. In recent times, this approach has been used to understand phase transitions, to interpret the correlation functions near critical points and the resulting relations among the exponents that occur in those functions [1, 2]. Regime transitions, sudden changes from one confinement type to another, are very common even in thermonuclear plasmas. In the ASDEX machine it was discovered in 1982 that, when the input power was increased sufficiently, the plasmas tended to transit spontaneously to an enhanced confinement mode called the high confinement or H-mode [3]. L–H transitions are experimentally observed to occur when the heating power applied to the plasma exceeds a critical value, which is now known as the threshold power for H-mode access (P_{Thresh}) [4]. The H-mode is typically reached transiting first through a low-confinement regime the so-called L-mode. The transition to the H-mode can be considered a sort of phase transition and is routinely achieved [4] in the vast majority of tokamaks. The main physical difference of the H-mode confinement regime, with respect to the L-mode, is a steep transport barrier at the plasma edge, known as an edge transport barrier (ETB). Once the correct L-mode conditions are met, the transition from L- to H-mode occurs with the spontaneous formation of an ETB. When the ETB is starting to develop, the confinement at the edge of the plasma improves, which consequentially sustains the further growth of the H-mode pedestal. With regard to the time scales, the formation of an ETB occurs on a typical time scale of 1 ms, as can be seen by its effects on particle transport. Since the ETB increases significantly the overall confinement properties of the plasma, a lot of

experimental efforts have been devoted to identifying the scaling law expressions which govern the access to the H-mode. A series of theoretical models exist, which try to interpret the onset of an ETB as the interplay between plasma instabilities and various stabilizing factors. These models consider a wide range of different physical phenomena and provide some testable criteria for the onset of the H-mode in terms of measurable plasma parameters [5]. These criteria are expressed in terms of a critical electron temperature T_{ec}, which is considered to be the relevant order parameter, to use a language typical of phase transitions and auto-organization studies. On the other hand, even the best theoretical models have not been able to properly predict the access to the H-mode. Moreover, from an engineering and an operational point of view, the P_{Thresh} represents a more useful parameter than the electron temperature in the design of new machines. Therefore various empirical models are available to determine the power threshold between the L- and H-mode of confinement [6]. Various studies have focused on collecting and assembling measurements of P_{Thresh} from multiple machines into a single database and extracting an empirical scaling law. These scaling laws are expressed in the form of power-law monomials [6]:

$$Y = f(X_n, \beta_p) = \beta_0 X_1^{\beta_1} X_2^{\beta_2} X_3^{\beta_3} \ldots X_n^{\beta_p}, \quad (1)$$

where Y represent the dependent variable (P_{Thresh}), X_n the set of n candidate physical quantities and β_p the set of $P + 1$ unknown coefficients. This is the functional form of the scaling laws which will be considered in this paper. Power-law relations are attractive in many fields of science for theoretical reasons, but mainly because they are widely used as a rough assessment of complex multivariate problems. Although various physical quantities are believed to have a potential influence on the L to H transition, it is also widely accepted, for practical purposes, that the P_{Thresh} mainly depends on the plasma line integrated density (n_{e20}), the strength of the torodial field (B_t) and the plasma surface area (S) (all quantities are generally known with a precision of between $\pm1\%$ and $\pm10\%$). On the other hand, the plasma shape, described by the two main macroscopic quantities elongation (k) and triangularity (δ), has also a quite significant impact on confinement and its influence has therefore also to be considered. In general, therefore the scaling laws considered in this paper are of the form

$$P_{Thresh}(MW) = \beta_0 a(m)^{\beta_1} R(m)^{\beta_2} S(m)^{\beta_3} \delta^{\beta_4} k^{\beta_5} B_t(T)^{\beta_6} n_{e20}$$
$$\times (10^{20} \, m^{-3})^{\beta_7} q_{95}^{\beta_8} I_p(MA)^{\beta_9}, \quad (2)$$

where I_p is the plasma current, a the minor radius, R the major radius and q_{95} the safety factor at 95% of the plasma radius. For continuity with previous studies, the same physical quantities and notation as in [6] are adopted. Over the last years, significant efforts have been devoted in the fusion community to the collection of a multi-machine database for the systematic study of the power threshold to achieve the H-mode. The results reported in this paper have been obtained using the 2010 version of the ITPA database (IGDBTHv6b), whose variables and main characteristics are described in the official website[4] (see also section 2). The extraction of a scaling law from such a database is quite a challenging statistical
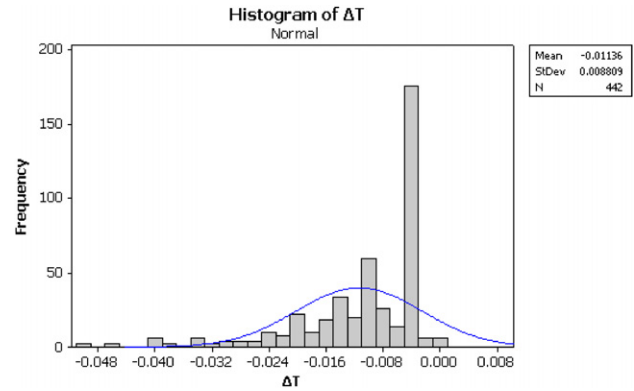

Figure 1. Distribution of the difference (ΔT in seconds) between the L–H transition time and the time-slice considered in the statistical methodology.

activity which implies first to identify the most representative variables, second to perform adequate forms of regressions and finally to select the best regression formula from the various reasonable candidates. The main statistical tools, which have been adapted and used to analyse the IGDBTHv6b database, are described in detail in section 3. The results of the statistical analysis are reported in section 4 for the entire database, providing results in agreement with previous investigations but showing the importance of shaping parameters, elongation and triangularity. In section 5 the analysis has been particularized for individual machines, whose data are best interpreted by scaling laws which are statistically different from each other. Different current regimes ($I_p < 2.5$ MA for low current and $I_p > 2.5$ for high current regimes) would also indicate different optimal scaling laws, as reported in section 6. Conclusions and suggestions of further investigations are the subject of section 7 of the paper.

2. Overview of the database

The statistical analysis has been carried out using the 2010 version of the ITPA database (IGDBTHv6b) [6]. As in [6], the SELEC2007 criteria has also been applied. Furthermore, it has been decided to filter the database and keep only the measurements taken very close to the time of the L–H transition. Therefore only measurements in the interval of 50 ms before the L–H transition have been retained in order to reduce the possible occurrence of undesired spurious events. The distribution of the difference (ΔT in seconds), between the L–H transition time and the time-slice (ts) when the data are sampled, is shown in figure 1.

The times considered in this paper are mainly distributed around 11 ± 9 ms before the transition. The database is also cleaned from duplicate data and randomized. A graphic overview of the operational space and considered tokamaks, for the geometrical (a, R, S), shape (k, δ) and operative ($B_T, n_{e20}, I_p, q_{95}$) quantities versus P_{Thresh} with the same dimensions of equation (2) is provided in figures 2, 3 and 4.

Only the P_{Thresh} dependence on geometrical quantities is quite linear whereas n_{e20}, B_t and shape quantities show a strong nonlinear behaviour as do the operative quantities. The final dataset includes 442 time slices from the main tokamaks. The operational range covered by this set of discharges is
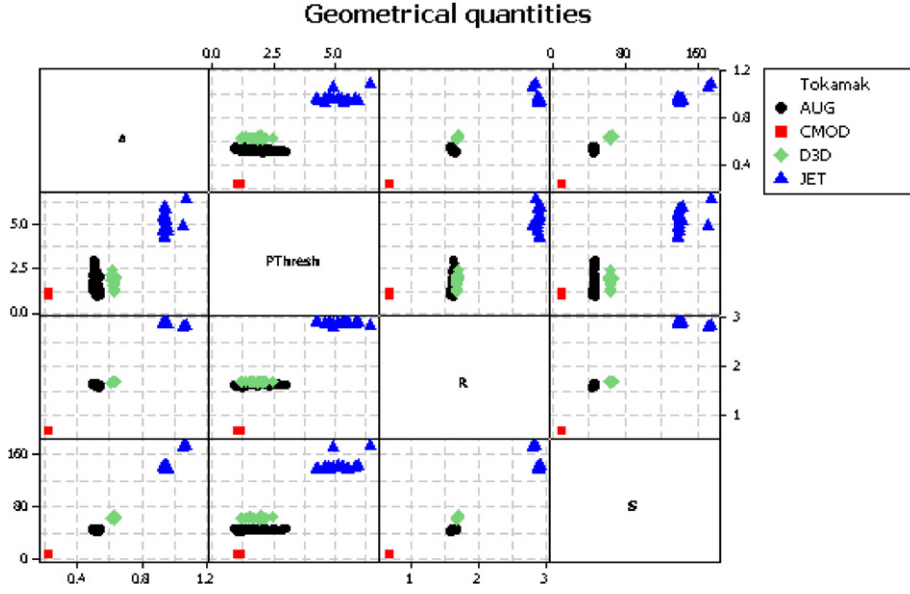
Figure 2. Geometrical quantities versus experimental threshold power for the time slices considered in the statistical analysis.
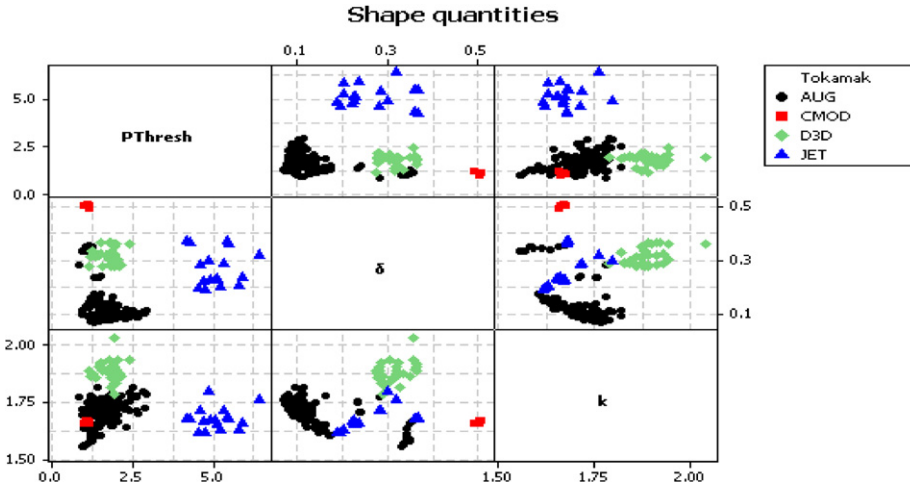

Figure 3. Shape quantities versus experimental threshold power for the time slices considered in the statistical analysis.

$1.29 < B_t[T] < 5.37$, $0.20 < n_{e20}[10^{20}\,\mathrm{m}^{-3}] < 1.19$, $0.67 < R[m] < 2.92$, $0.216 < a[m] < 1.03$, $7.32 < S[m^2] < 174.10$, $0.0696 < \delta < 0.508$, $1.56 < k < 2.04$ $2.51 < q_{95} < 6.78$, $0.54 < I_p[MA] < 3.22$ and $0.831 < P_{\mathrm{Thresh}}[MW] < 6.466$.

3. The statistical methodology

The identification of the best empirical scaling law from a large multi-machine database can be conceptually divided into three phases: (1) feature extraction: the selection of the most appropriate macroscopic quantities available in the database to describe the phenomenon without assumption about the functional form of the scaling law, (2) regression method: application of the regression analysis in order to provide the best fit of the available data, (3) statistical model selection: identification of the most performing model among the candidate ones. Therefore, first an appropriate set of criteria for choosing between the contending subsets of independent variables is to be devised. Then the most appropriate regression techniques have to be deployed. Finally, statistically sound indicators have to be defined to determine the model, which best interpret the experimental evidence. With regard to the first aspect, the approach adopted is the one of agglomerative hierarchical clustering and correlation analysis (see section 3.1). For the studied database, the most appropriate regression method proves to be nonlinear regression (see section 3.2). The identification of the best model will be attacked with the use of two widely accepted model selection criteria: the Akaike information criterion (AIC) [7] and the Bayesian information criterion (BIC) [8] (see section 3.3).

3.1. Feature extraction

The choice of predictive variables to include in the candidate models is carried out by an agglomerative hierarchical clustering [9], which identifies groups of variables that behave
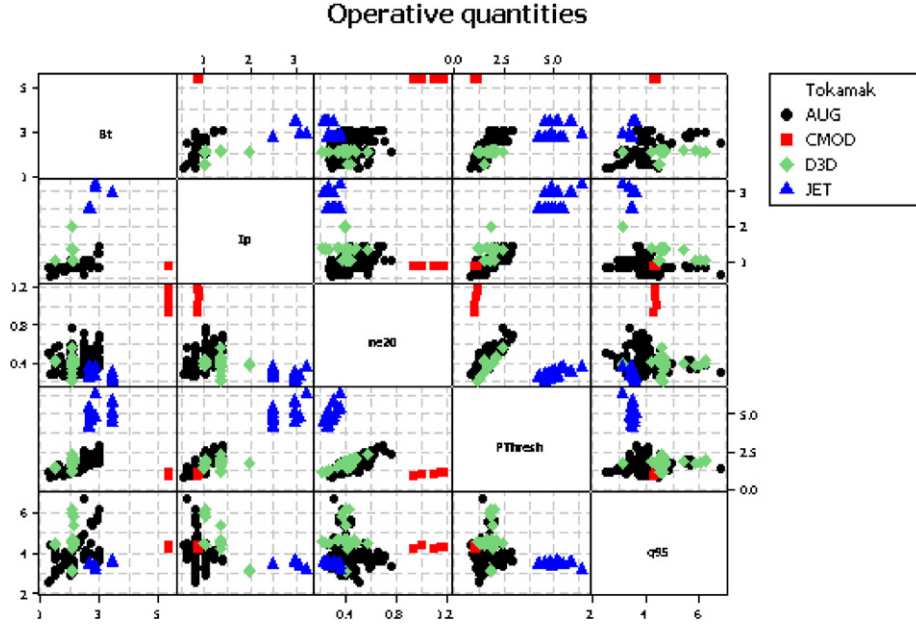
Operative quantities

Figure 4. Operative quantities versus experimental threshold power for the time slices considered in the statistical analysis.

similarly or show similar characteristics in the dataset without assumption about the functional form of the scaling law. The objective is to reduce the number of variables in the model, using only the most significant variable for each cluster. Therefore the clustering is applied to the variables and not to the observation vectors. The adopted procedure begins from the evaluation of the original $n \times n$ distance matrix, $\boldsymbol{D}$, (where n is the number of variables) with all variables separate, each forming its own cluster. The correlations method for distance measures has been used. With the correlation method, the i, j entry of the distance matrix is $d_{i,j} = 1 - \rho_{ij}$ where ρ_{ij} is the (Pearson product moment) correlation between variables i and j. Thus, the correlation method will give distances between 0 and 1 for positive correlations, and between 1 and 2 for negative correlations. At the beginning, when each variable constitutes a cluster, the distance between clusters is simply the inter-variables distance. In the first step, the two variables closest together are joined if the similarity threshold level is reached. The similarity, s_{ij}, between two clusters i and j, is given by $s_{ij} = 100(1 - d_{ij}/d_{\max})$, where $d_{\max}$ is the maximum value in the original distance matrix $\boldsymbol{D}$. The specific threshold level of similarity of 95% has been used in this paper. This value of the threshold is the one typically used in the literature, in analogy to the 95% confidence usually adopted in statistical treatments. Such a level of similarity seems appropriate also in this application, since it has been checked that small variations, of a few per cent, around this threshold do not change the results appreciably.

A linkage rule is necessary for calculating inter-cluster distances when there are multiple variables in a cluster. In the next step, either a third variable joins the first two, or two other variables join together into a different cluster. This process will continue until all clusters are joined into one.

The linkage rule determines how the distance between two clusters is defined. In this paper, Ward's linkage has been used. The objective of Ward's linkage is to minimize the within-cluster sum of squares. With this linkage method, the distance between two clusters is the sum of squared deviations from points to centroids. In detail, the formula for the distance matrix, implementing Ward's linkage, can be written as

$$d_{mj} = \frac{(N_j + N_k)d_{kj} + (N_j + N_l)d_{lj} - N_j d_{kl}}{N_j + N_m}, \qquad (3)$$

where N_j, N_k, N_l and N_m are the number of variables in clusters j, k, l and m. Relation (3) refers to the case of cluster k and l being merged into cluster m. In Ward's linkage, it is possible for the distance between two clusters to be larger than $d_{\max}$. If this happens, the similarity is negative. The decision about the final grouping is shown in a graphical way called a dendrogram (see section 4). The dendrogram shows the manner in which the clusters are formed and suggests which variables are to be included in the candidate models. The candidate models are formed considering only one variable in each cluster except for the n_{e20} and B_t that are always retained in the candidate models.

3.2. Regression techniques

The predicted H-mode P_{Thresh} is generally estimated by ordinary least-squares (OLS) fitting of the power law via its corresponding log-expression. The problem with this method is that the transformation mainly 'distorts' the experimental errors and this violates one of the basic assumptions of the OLS technique [7]. Linearization procedures should be avoided unless the errors are multiplicative log normally distributed because they can otherwise give misleading results. This comes from the fact that, whatever the experimental errors on a dependent variable might be, the errors on the logarithm of the dependent variable are different. Therefore, when the transformed sum of squares is minimized, different results will be obtained for both the parameter values and their calculated standard deviations. Only with multiplicative errors that are log normally distributed, this procedure gives unbiased and

consistent parameter estimates. The validity of the statistical model assumptions for the IGDBTH has been studied in detail [10, 11], indicating that the assumptions of OLS are weakly satisfied [11] and suggesting that an alternative fitting technique would be more appropriate. As a consequence, given the candidate laws from the feature extraction process, a direct nonlinear regression technique has been carried out because it is more suitable for extrapolation purposes, and the multicollinearity between variables is mainly due to the lack of data. Multicollinearity is present if two or more independent variables in a model are highly linearly related. Multicollinearity in linear regression causes problems in using regression models to draw conclusions about the relationships between independent and dependent variables. More details and information about the type of nonlinear regression, used in this paper, including many data examples, are given by Bates and Watts [12]. Both linear and nonlinear regressions minimize the sum of the squared residuals (SSE) to estimate the parameters. However, they use very different approaches. For linear regression, the numerical algorithm can mathematically derive the minimum SSE by solving equations. However, for nonlinear regression, there is no direct solution for minimizing the SSE. Thus, an iterative algorithm estimates the parameters by systematically adjusting their estimates to reduce the SSE. The algorithm uses a starting value for each parameter to calculate the initial SSE. In this paper, the starting values are free to move from an interval centred in -100 to 100. At each iteration, the algorithm adjusts the parameter estimates in a manner that reduces the SSE compared with the previous iteration. The iterations continue until the algorithm converges on the minimum SSE. The Gauss–Newton method has been used in order to determine the best parameter estimation. The code achieves the convergence criteria in few iterations (10 iterations for the most complex case), using the Gauss–Newton algorithm with the parameter range specified above. However, fast and easy convergence, by itself, does not guarantee that a model is the best among the candidate ones. Therefore, it is crucial to perform a model selection process and examine the plots of the residuals to ensure that the model adequately fits the data.

3.3. Model selection

Two well established and independent model selection criteria, based on the concept of information entropy (Akaike information criterion, AIC) and Bayesian formalism (Bayesian information criterion, BIC), are then used for model selection in order to identify the most efficient scaling law from candidate models. The AIC developed by Hirotsugu Akaike under the name of 'an information criterion' in 1971, is a measure of the goodness of fit of an estimated statistical model. The AIC is an operational way of trading off the complexity of an estimated model against how well the model fits the data and can be written as

$$\text{AIC} = 2P + N \times \ln\left(\frac{\text{RSS}}{N}\right), \tag{4}$$

where RSS is the residual sum of squares of the model, N the number of observations, or equivalently, the sample size and P the number of free parameters to be estimated. Given

any number of candidate models, the model with the lower value of AIC is the one to be chosen. Increasing the number of free parameters to be estimated improves the goodness of fit, regardless of the number of free parameters in the data generating process. Hence AIC not only rewards goodness of fit but also includes a penalty that is an increasing function of the number of estimated parameters. The AIC methodology attempts to find the model that best explains the data with a minimum of free parameters.

The BIC criterion can be formulated as (under the assumption that the model errors or disturbances are normally distributed)

$$\text{BIC} = N \cdot \ln\left(\frac{\text{RSS}}{N}\right) + P \times \ln(N). \tag{5}$$

Given any two estimated models, the model with the lower value of BIC is the one to be preferred. The BIC is an increasing function of RSS and an increasing function of P. That is, unexplained variation in the dependent variable and the number of explanatory variables increases the value of BIC. The BIC penalizes free parameters more strongly than the AIC does. These two indicators implement a form of 'Occam Razor' by penalizing models with a higher number of parameters. It is important to keep in mind that the BIC and AIC can be used to compare estimated models only when the dataset is identical for all estimates being compared.

4. Results for the full multi-machines database

Starting with the process of variable selection, the dendrogram, corresponding to the whole database used in this paper, is shown in figure 5.

It is worth noticing that the dendrogram groups the variables in a way, which is very coherent with the expectation from the design of the machines and the everyday operation of the experiments. The geometrical quantities (a, R and S) are grouped together (blue coloured cluster), reflecting the fact that they are strongly correlated (at the level of the design of the machines). Given the very strong relation between these quantities, only one of them is to be used in the scaling laws. The clustering also shows that both shaping parameters (k, δ) are grouped together, revealing the practice of trying to increase the plasma confinement by raising both these shaping quantities. Also the safety factor q_{95}, which is operatively normally linked to the plasma shape, is indeed shown to be strongly correlated with k and δ. B_t and I_p are in contrast quite independent variables, reflecting the relative freedom of the operator to determine their values in the experiments. On the other hand, n_{e20} seems to be highly correlated with the magnetic field. This is spurious correlation due to the way many machines are operated and it is clear that the best regressions are obtained if the density is retained in the scaling laws. The analysis of the dendrogram summarizes visually the similarity between the variables and provides clear indications about the good candidate scaling laws to be considered in the regression analysis. Choosing only one of the variables in each cluster, the remaining candidate scaling laws are summarized in table 1. In this table, the most recent scaling laws available in the literature (Scal2008) and their extension, considering
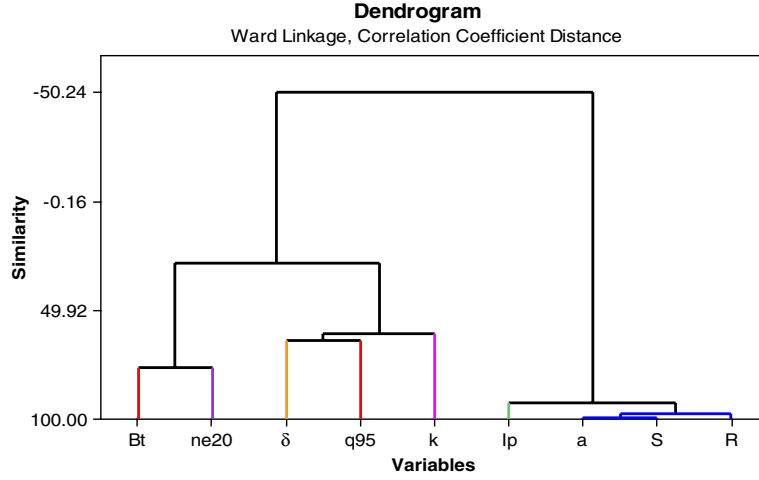
Dendrogram
Ward Linkage, Correlation Coefficient Distance

Figure 5. Dendrogram used for variable selection in the full multi-machine database.

Table 1. Candidate scaling laws provided by dendrogram analysis for the full multi-machine database.

Model	Equation
NLM-1	$P_{Th} = \beta_0 a^{\beta_1} \delta^{\beta_4} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7} q_{95}^{\beta_8} I_p^{\beta_9}$
NLM-2	$P_{Th} = \beta_0 S^{\beta_3} \delta^{\beta_4} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7} q_{95}^{\beta_8} I_p^{\beta_9}$
NLM-3	$P_{Th} = \beta_0 R^{\beta_2} \delta^{\beta_4} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7} q_{95}^{\beta_8} I_p^{\beta_9}$
NLM-4	$P_{Th} = \beta_0 \delta^{\beta_4} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7} q_{95}^{\beta_8} I_p^{\beta_9}$
Scal2008	$P_{Th} = \beta_0 S^{\beta_3} B_t^{\beta_6} n_{e20}^{\beta_7}$
ICFDT2011	$P_{Th} = \beta_0 S^{\beta_3} \delta^{\beta_4} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7}$
NLM-5	$P_{Th} = \beta_0 S^{\beta_3} \delta^{\beta_4} B_t^{\beta_6} n_{e20}^{\beta_7}$
NLM-7	$P_{Th} = \beta_0 S^{\beta_3} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7}$
NLM-9	$P_{Th} = \beta_0 S^{\beta_3} B_t^{\beta_6} n_{e20}^{\beta_7} q_{95}^{\beta_8}$
NLM-11	$P_{Th} = \beta_0 S^{\beta_3} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7} q_{95}^{\beta_8} I_p^{\beta_9}$
NLM-13	$P_{Th} = \beta_0 S^{\beta_3} k^{\beta_5} B_t^{\beta_6} n_{e20}^{\beta_7} q_{95}^{\beta_8}$

also the shape quantities [13] proposed by the authors of this paper, are included.

The obtained coefficients (the multiplying constant and the various exponents for the independent variables) of the scaling laws are shown in table 2.

Also the equivalent linear models are reported in table 3. In this way, it can be verified that the AIC and BIC criteria manage to properly identify the nonlinear models as the most appropriate ones (confirming the consideration of section 3.2). The score assigned to the various scaling laws by these model selection criteria are reported in the previous table (the red colour has been used to indicate the best models). A strict evaluation on the basis of AIC and BIC parameters would indicate quite clearly that the best scaling law is the nonlinear one number 11 (NLM-11). This law reads (with 95% confidence interval for the parameters):

$$P_{\text{Thresh}} = 0.068_{0.063}^{0.074} S^{1.0.35_{1.020}^{1.050}} K^{-1.152_{-1.220}^{-1.084}} B_t^{0.764_{0.741}^{0.787}} n_{e20}^{0.765_{0.754}^{0.775}}$$
$$\times q_{95}^{-0.051_{-0.072}^{-0.031}} I_p^{-0.034_{-0.056}^{-0.012}}. \tag{6}$$

A simpler scaling law, involving seven instead of six measurements, has also quite good performance in terms both

of BIC and AIC scores. For the reader's convenience this alternative is reported in the following:

$$P_{\text{Thresh}} = 0.067_{0.061}^{0.073} S^{1.041_{1.024}^{1.058}} \delta^{0.004_{-0.002}^{0.011}} k^{-1.135_{-1.207}^{-1.063}} B_t^{0.778_{0.748}^{0.808}}$$
$$\times n_{e20}^{0.764_{0.754}^{0.774}} q_{95}^{-0.064_{-0.092}^{-0.037}} I_p^{-0.049_{-0.080}^{-0.018}}. \tag{7}$$

The superior performance of the scaling law NLM-11, identified by the statistical indicators BIC, is due to a lower number of free unknown parameters included in the model. A simple inspection of the residuals indeed shows that equation NLM-11 provides a very similar pattern of residuals, i.e. provides the same information, compared with NLM-2, as can be seen in figure 6.

A comparison between the selected model, NLM-11, and the most used version of the scaling law, Scal2008, is shown in figure 7.

Inspection of figure 7 reveals immediately that the distribution of the residuals is narrower, more symmetric and better catered with respect to zero for the scaling law NML-11. This equation therefore interprets significantly better than the main alternatives the available data. A comparison between the selected model, NLM-11, and the most used version of the scaling law [6]

$$P_{th} = 0.049 S^{0.717 \pm 0.035} B_t^{0.803 \pm 0.032} n_{e20}^{0.941 \pm 0.019} \tag{8}$$

is shown in figure 8. The P_{Thresh} predicted by this NLM-11 model for ITER ($B_t = 5.37$(T) $n_{e20} = 1$ (10^{20} m^{-3}) $S = 678$ (m^2) $k = 1, 86$ $q_{95} = 3$ $I_p = 15$ (MA)) is 88 MW.

Coming to a more physical interpretation of the derived equations, the NLM-11 scaling law shows a very strong dependence of the power threshold on the shape of the plasma. Indeed, the effect of elongation k has one of the highest exponents. This is in agreement with MHD theory, since the increased shaping of the plasma is indeed expected to improve stability and confinement. Another interesting remark relates to the effects of the plasma current and q_{95}. Both these quantities are retained in the scaling law NLM-11 with a very low exponent. On the other hand, their effect is quite important. The distribution of the residuals is significantly affected, if these two quantities are not included in the regression, as can

Table 2. BIC and AIC values and coefficient for nonlinear regression models. The last row gives the estimated threshold power for a device of the ITER class: the parameters used to evaluate this power are: $I_p = 15$ MA, $B_t = 5$ T, $n_{e0} = 10^{20}$, $a = 2$ m, $S = 678$, $R = 6.2$ m, $\delta = 0.5$, $k = 1.86$, $q_{95} = 3.5$.

Mod. sel.criterion	Candidate models (nonlinear regression models)										
	NLM-1	NLM-2	NLM-3	NLM-4	Scal2008	ICFDT2011	NLM-5	NLM-7	NLM-9	NLM-11	NLM-13
BIC	−2106	−2793	−2457	−1241	−1888	−2783	−1971	−2784	−1993	−2797	−2794
AIC	−2138	−2826	−2489	−1270	−1905	−2807	−1991	−2805	−2014	−2826	−2819
Parameter						Coefficients					
Constant	9.349	0.067	0.809	7.863	0.034	0.072	0.027	0.074	0.050	0.068	0.077
a	2.01	—	—	—	—	—	—	—	—	—	—
R	—	—	1.66	—	—	—	—	—	—	—	—
S	—	1.041	—	—	0.987	1.023	1.023	1.02	0.953	1.035	1.014
δ	−0.06	0.004	0.002	−0.221	—	−0.004	−0.057	—	—	—	—
k	−0.62	−1.135	−0.278	−3.89	—	−1.247	—	−1.266	—	−1.152	−1.223
B_t	0.797	0.778	0.406	−0.734	0.868	0.728	0.838	0.728	0.861	0.764	0.733
n_{e20}	0.775	0.764	0.726	0.753	0.678	0.774	0.674	0.776	0.652	0.765	0.768
q_{95}	−0.116	−0.064	0.291	1.047	—	—	—	—	−0.191	−0.051	−0.03
I_p	−0.034	−0.049	0.346	1.692	—	—	—	—	—	−0.034	—
P_{Thresh} (MW)	81.89	88.32	97.71	73.69	91.10	89.17	90.49	88.57	86.00	88.25	88.82

Table 3. AIC an BIC value and coefficient for the equivalent linear regression models. The last row gives the estimated threshold power for a device of the ITER class: the parameters used to evaluate this power are: $I_p = 15$ MA, $B_t = 5$ T, $n_{e0} = 10^{20}$, $a = 2$ m, $S = 678$, $R = 6.2$ m, $\delta = 0.5$, $k = 1.86$, $q_{95} = 3.5$.

Mod. sel.criterion	Candidate models (linear regression models)										
	LM-1	LM-2	LM-3	LM-4	SCAL 2008 (linear)	ICFDT2011 (linear)	LM-5	LM-7	LM-9	LM-11	LM-13
BIC	−1872	−2700	−2396	−1159	−1772	−2763	−1810	−2755	−1954	−2716	−2759
AIC	−1904	−2732	−2429	−1187	−1788	−2788	−1830	−2776	−1975	−2744	−2784
Parameter						Coefficients					
Constant	7.983	0.050	0.903	5.073	0.040	0.065	0.035	0.070	0.052	0.053	0.073
a	2.510	—	—	—	—	—	—	—	—	—	—
R	—	—	1.666	—	—	—	—	—	—	—	—
S	—	1.078	—	—	0.957	1.017	0.961	1.019	0.942	1.063	1.011
δ	−0.020	0.006	0.0005	−0.287	—	−0.016	−0.045	—	—	—	—
k	0.695	−0.797	−0.575	−3.916	—	−1.09	—	−1.154	—	−0.845	−1.078
B_t	1.239	0.871	0.369	−0.739	0.806	0.750	0.831	0.737	0.829	0.847	0.748
n_{e20}	0.752	0.748	0.744	0.677	0.667	0.762	0.646	0.775	0.64	0.748	0.76
q_{95}	−0.518	0.141	0.36	1.239	—	—	—	—	0.171	−0.118	−0.051
I_p	−0.486	0.123	0.358	1.666						−0.096	
P_{Thresh} (MW)	86.46	241.08	96.13	55.89	79.42	89.29	76.74	90.59	117.42	90.20	90.54

be seen in figure 9 (scaling laws NLM-7 and NLM-13). Indeed, in the table inserted in the figure, it can be seen how, even if the standard deviations of the three scaling laws are not terribly different, the mean of the residual distribution is much closer to zero for the scaling law NLM-11.

To summarize, the method of selecting the more significant variables based on dendrograms provides quite good results which are systematically confirmed by analysis of the residuals. The statistical analysis also indicates consistently that nonlinear regression should be used, since the results obtained with the traditional log scaling are not statistically significant. In terms of interpretation of the selected scaling law, the results for the entire database indicate that the dependence from the main quantities B_t, n_{e20} and S confirms the results of previous analyses [6]. The exponents agree within the error bars with the most established version of the scaling law [6]. On the other hand, the present analysis reveals a very strong dependence of the results on the plasma shaping and in particular on the plasma elongation. Current and q_{95} have a much weaker but nonetheless non-negligible effect of the quality of the fit.

5. Results for individual machines

One can first of all try to verify the quality of the previously derived scaling law NLM-11 by analysing the residual distribution of each machine, for which enough data were available to provide a sufficient statistical basis of the conclusions. These machines are JET, AUG and DIII-D. Since each individual machine typically operates at the same minor and major radii, the dependence from the plasma geometry has been eliminated in the regressions discussed in the following. This is of course a perfectly legitimate step since, if the general scaling law already derived was perfect, the contribution of the individual machines should be recovered by fixing their dimensional parameters. The statistical analysis of the database, in any case, confirms that these two quantities do not change over a sufficient interval to provide good regressors. To shed light on the behaviour of the individual machines, the same approach used to analyse the whole database has also been followed to investigate the scaling laws for the various individual machines. Therefore the analysis of the dendrograms, the nonlinear regression and
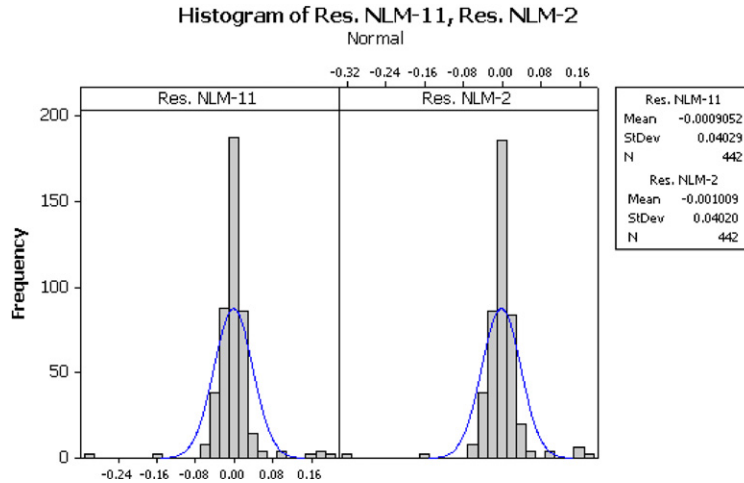
Figure 6. Histogram of the residuals for the two competing relations NLM-11 and NLM2.
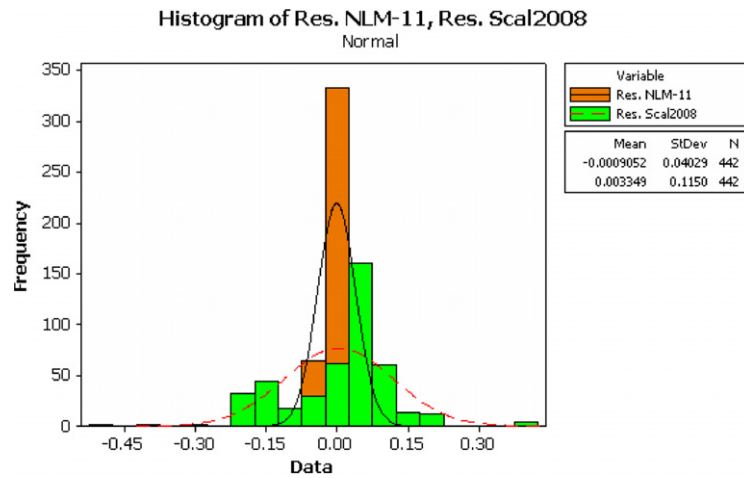

Figure 7. Comparison between the proposed scaling law and the most used version of the scaling law available.
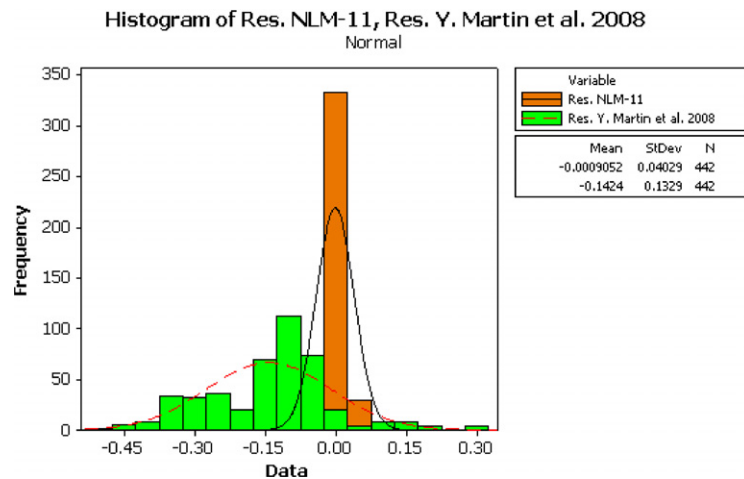

Figure 8. Comparison between the proposed scaling law and thescaling from equation (8).

the evaluation of the residual distributions have been repeated for the individual machines. The dendrograms for the three machines investigated are reported in figure 10.

The dendrograms are very similar between JET and DIII-D, apart from some not very relevant minor variations.

They basically confirm the coherence of the database of these two machines and the agreement of the analysis results with the expectations based on the design and operation of the machines. It is indeed interesting to note that, once the dependence on the plasmas surface is eliminated, the
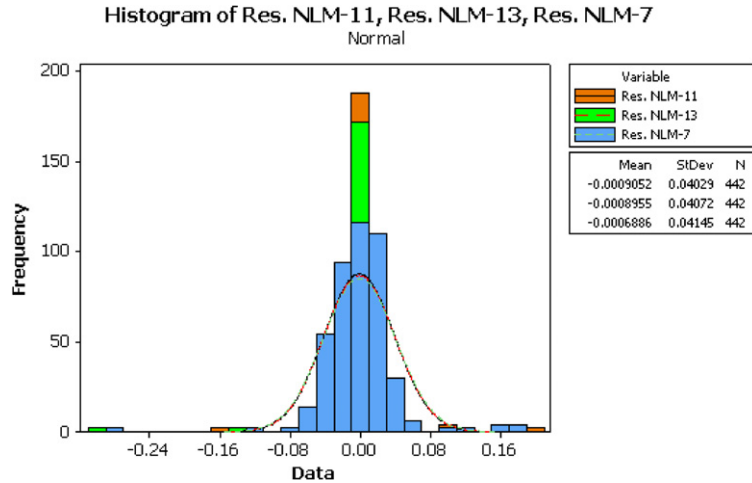
Figure 9. Histogram of the residuals for scaling law NLM-11 without I_p and q_{95}.
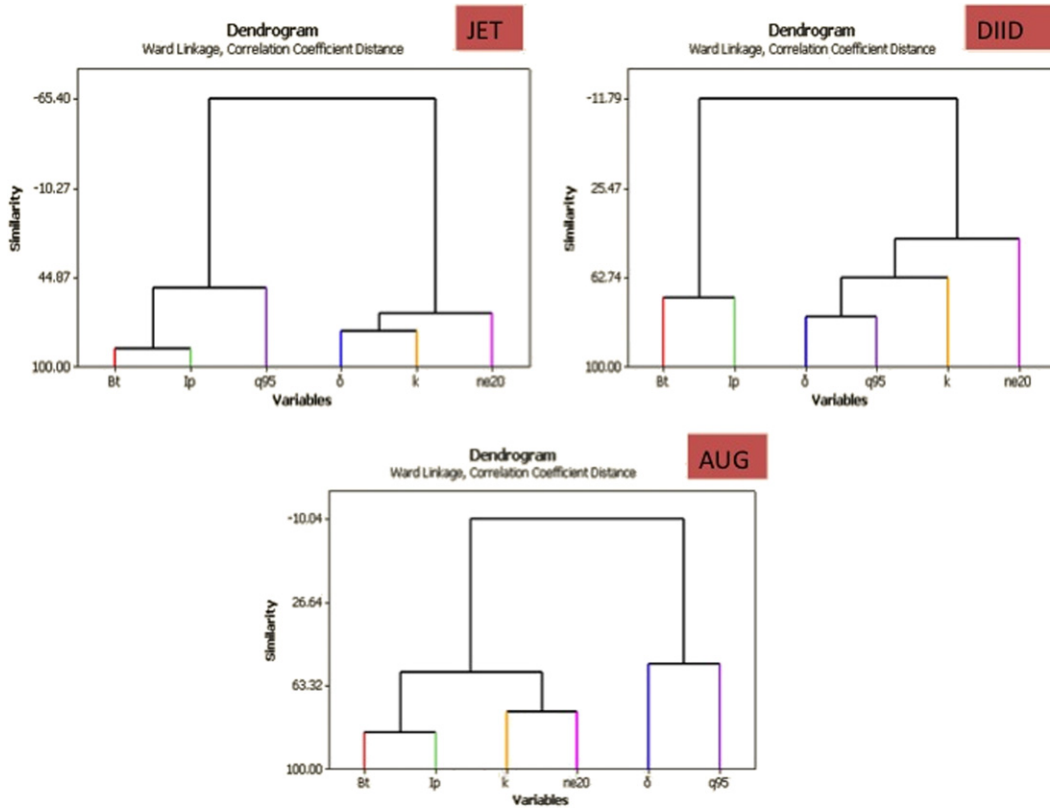

Figure 10. Dendrograms for the machines investigated: JET, DIID and AUG.

correlation between the toroidal field and the plasma current becomes even more significant than for the entire database. AUG, on the other hand, seems to have significantly different characteristics already at the level of the database. In any case, the final selection of the combination of variables, to apply regression analysis to, follows the same method used for the entire database. Again the analysis of the dendrograms, the nonlinear regression and the evaluation of the residual distributions have been repeated for the individual machines. Only one of the strongly correlated variables is selected for each individual regression. The main set of scaling laws tested is reported in table 4.

The scores of the AIC and BIC criteria for the various candidate models are reported in table 5.

Candidate scaling laws for the individual machines. For the reader's convenience, the best scaling laws for each machine, as selected by the AIC and BIC criteria, are reported in the following together with the 95% confidence intervals for the parameters. It is worth mentioning that the classification of the various regression formulae, as provided by the AIC and BIC criteria, is confirmed by a detailed analysis of the residuals. The residuals of the three scaling laws JET-4, AUG-1 and DIII-D-4 are clearly better centred around zero and narrower

Table 4. Scaling laws considered for the various machines on the basis of the dendrogram analysis.

Model	Equation
JET-1	$P_{\mathrm{Th}} = \beta_0 \delta^{\beta_4} k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8} I_{\mathrm{p}}^{\beta_9}$
JET-2	$P_{\mathrm{Th}} = \beta_0 \delta^{\beta_4} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8}$
JET-3	$P_{\mathrm{Th}} = \beta_0 k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8}$
JET-4	$P_{\mathrm{Th}} = \beta_0 k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8} I_{\mathrm{p}}^{\beta_9}$
DIII-D-1	$P_{\mathrm{Th}} = \beta_0 \delta^{\beta_4} k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8} I_{\mathrm{p}}^{\beta_9}$
DIII-D-2	$P_{\mathrm{Th}} = \beta_0 \delta^{\beta_4} k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7}$
DIII-D-3	$P_{\mathrm{Th}} = \beta_0 k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8}$
DIII-D-4	$P_{\mathrm{Th}} = \beta_0 k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8} I_{\mathrm{p}}^{\beta_9}$
AUG-1	$P_{\mathrm{Th}} = \beta_0 \delta^{\beta_4} k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8} I_{\mathrm{p}}^{\beta_9}$
AUG-2	$P_{\mathrm{Th}} = \beta_0 \delta^{\beta_4} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7}$
AUG-3	$P_{\mathrm{Th}} = \beta_0 B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8}$
AUG-4	$P_{\mathrm{Th}} = \beta_0 k^{\beta_5} B_{\mathrm{t}}^{\beta_6} n_{\mathrm{e20}}^{\beta_7} q_{95}^{\beta_8} I_{\mathrm{p}}^{\beta_9}$

than those of the other scaling laws. Moreover, in the case of JET and DIII-D, the scaling laws selected are the most parsimonious because they use two variables less than the next ones in the classification obtained with the AIC and BIC criteria:

$$P_{\mathrm{Thresh}}(\mathrm{JET}-4) = 6.705^{8.419}_{5.342} k^{-0.215^{-0.021}_{-0.408}} B_{\mathrm{t}}^{0.376^{0.501}_{0.250}} n_{\mathrm{e20}}^{0.744^{0.764}_{0.723}}$$
$$\times q_{95}^{-0.124^{0.095}_{-0.344}} I_{\mathrm{p}}^{0.524^{0.655}_{0.394}}, \tag{9}$$

$$P_{\mathrm{Thresh}}(\mathrm{DIIID}-4) = 2.259^{2.436}_{2.096} k^{-0.766^{-0.620}_{-0.913}} B_{\mathrm{t}}^{0.341^{0.440}_{0.243}} n_{\mathrm{e20}}^{0.757^{0.771}_{0.743}}$$
$$\times q_{95}^{0.419^{0.521}_{0.317}} I_{\mathrm{p}}^{0.391^{0.488}_{0.294}}, \tag{10}$$

$$P_{\mathrm{Thresh}}(\mathrm{AUG}-1) = 1.876^{1.936}_{1.818} \delta^{-0.011^{-0.006}_{-0.016}} k^{-0.313^{-0.255}_{-0.371}}$$
$$\times B_{\mathrm{t}}^{0.586^{0.615}_{0.557}} n_{\mathrm{e20}}^{0.746^{0.752}_{0.741}} q_{95}^{0.153^{0.181}_{0.124}} I_{\mathrm{p}}^{0.140^{0.169}_{0.110}}. \tag{11}$$

The obtained results are particularly interesting and deserve some comments. The analysis of the dendrograms indicates that almost the same variables are the most relevant for all the three machines (except triangularity in the case of DIII-D as discussed later). On the other hand, the scaling laws for the various machines are indeed different from each other with a probability typically well in excess of 95%. Basically the only exponent, which is not statistically different for the various machines, is the one of the density, whose dependence is well established around 0.75. Even the dependence on the magnetic field is statistically different at least for AUG. All the other exponents differ outside the 95% confidence intervals and the one of q_{95} has even a different sign for JET-4 with respect to the other machines. Moreover, the best scaling law for AUG presents even an additional variable, triangularity, with respect to the other machines. Its exponent is low but significant (on the basis of the residual analysis and the model selection criteria) and consistently negative. It is worth mentioning that the same differences between machines remain even if a different set of regressors is arbitrarily chosen. So the converge on statistically different scaling laws is not so much an effect of the application of the AIC and BIC criteria but more an output of the nonlinear regression analysis (which is an absolutely well established technique).

6. Analysis at different plasma currents

The significant differences found between the scaling laws obtained for the individual machines are statistically significant, as discussed in detail in the last section. On the other hand, the application of regression to derive scaling laws for the power threshold to access the H-mode has a main practical aspect, to help planning of experiments and design machines. Therefore it is appropriate to investigate how extrapolations based on the derived scaling laws perform. In particular, given the fact that the plasma current is one of the main design parameters, of both experiments and machines, it has been decided to investigate how a general scaling law obtained at low plasma currents, using the contribution of all the machines, is effective in describing the behaviour at higher currents. To this end, the analysis performed for the entire database, and described in sections 3 and 4, has been repeated adding the condition for the plasma current to be below 2.5 MA. The dendrogram of the variables is identical to the one shown in figure 5. The best scaling law, according again to the score of the AIC and BIC criteria is

$$P_{\mathrm{Thresh}}(I_{\mathrm{p}} < 25\,\mathrm{MA}) = 0.042^{0.044}_{0.040} S^{1.127^{1.136}_{1.118}} \delta^{0.023^{0.026}_{0.020}} k^{-0.728^{-0.691}_{-0.766}}$$
$$\times B_{\mathrm{t}}^{0.937^{0.952}_{0.923}} n_{\mathrm{e20}}^{0.750^{0.754}_{0.745}} q_{95}^{-0.207^{-0.193}_{-0.221}} I_{\mathrm{p}}^{-0.189^{-0.173}_{-0.205}}. \tag{12}$$

The distribution of the residuals is of quite good quality as can be seen in the left graph of figure 11.

The same scaling law has been applied to the discharges with plasma current higher than 2.5 MA. The fitting properties of this equation unfortunately extrapolate quite poorly to higher currents. This can be seen in the right graph of figure 11. The distribution of the residuals is clearly of inferior quality. To confirm the discrepancy of the results obtained at high and low currents, the same thorough analysis of the database, as described in section 3, has been performed also for the discharges at current higher than 2.5 MA. Again the analysis of the dendrograms, nonlinear regression, the scoring with the AIC and BIC criteria and the evaluation of the residual distributions have been repeated. The scaling law best fitting the experimental data is now found to be

$$P_{\mathrm{Thresh}}(I_{\mathrm{p}} > 25\,\mathrm{MA}) = 0.024^{0.105}_{0.005} S^{1.287^{1.626}_{0.949}} k^{-1.608^{-1.163}_{-2.054}}$$
$$\times B_{\mathrm{t}}^{0.744^{0.844}_{0.645}} n_{\mathrm{e20}}^{0.775^{0.790}_{0.759}} q_{95}^{0.293^{0.459}_{0.128}} I_{\mathrm{p}}^{-0.400^{-0.174}_{-0.627}}. \tag{13}$$

This scaling law is clearly different from the general one, except for the established dependence on the density. Particularly different are the exponents of the shaping parameters and q_{95}. But there is also no statistical agreement even on the toroidal field and plasma current dependences. The extrapolation capabilities of the derived scaling law can therefore be considered quite limited. It is worth mentioning that scaling laws obtained with linear regression extrapolate even worse than the ones obtained with nonlinear regression. This is expected from statistical considerations but has been clearly confirmed in practice in our application.

7. Conclusions

A new approach to identify the scaling law for the power threshold to access the H-mode has been described in detail. The method is based on the combination of various established

Table 5. Scaling laws and AIC and BIC scores for the individual machines (in different colours the best candidate models for each machine).

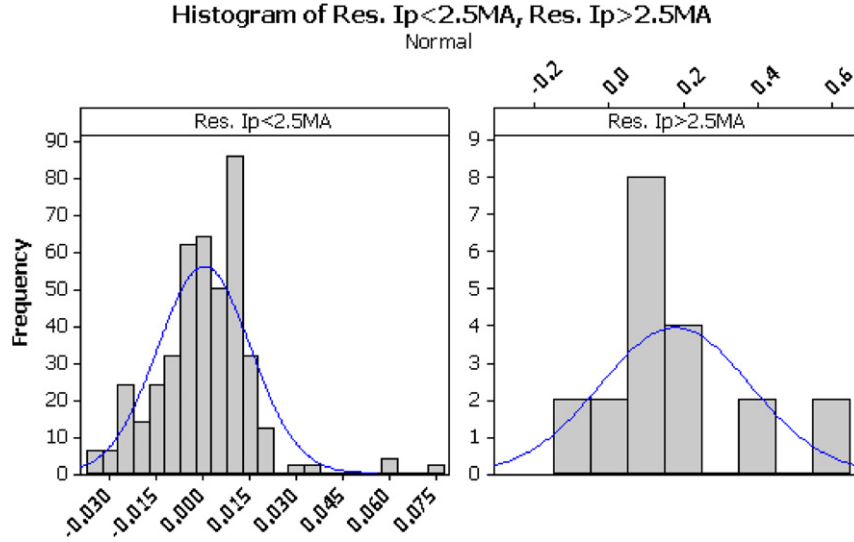|	JET-1	JET-2	JET-3	JET-4	AUG-1	AUG-2	AUG-3	AUG-4	DIII-D-1	DIII-D-2	DIII-D-3	DIII-D-4
BIC	−239	−191	−201	−240	−2537	−2374	−2418	−2518	−681	−639	−638	−685
AIC	−250	−199	−209	−250	−2563	−2388	−2433	−2544	−698	−651	−651	−699

Figure 11. Histogram of the residuals for the best scaling law obtained at low current regimes ($I_p < 2.5$ MA) applied to discharges at higher current regimes.

statistical tools, so the conclusions are to be considered quite robust in statistical terms. The inspection of the dendrograms, based on Ward's linkage, has proved to provide quite reliable results. Indeed, the most relevant variables are consistently identified also for individual machines and at different current levels. A statistical analysis indicates quite clearly that nonlinear regression is the only appropriate tool to be applied to the database and this is confirmed by the inspection of the residual distribution functions. The selection of the various candidate models has been based on the AIC and BIC criteria and again confirmed by the analysis of the residuals. The main results of previous studies have been reproduced. The main novelty of this part of the investigation is the relevance of the shape quantities, elongation and triangularity, and of q_{95}, which emerges quite clearly. Including these quantities has a quite relevant effect on the obtained scaling laws and the quality of the fit. The same approach applied to the main individual machines (JET, DIII-D and AUG) reveals some unexpected results. Apart from the exponent of the density, which remains quite stable around 0.75 (figure 12), practically all the scaling laws for the single machines have different dependences on the regressor with a statistical confidence higher than 95%, as can be appreciated visually from figure 13.

This suggests treating with great caution the overall scaling laws derived from the whole database. Since the individual experiments seem to obey different laws, it cannot be excluded that the physics of the transition to the H-mode is different in different machines. In that case, of course, the entire approach of multi-machine scaling should be revisited. Another possible alternative is of course the fact the set of macroscopic variables chosen could not be optimal and that other, mesoscopic or microscopic, quantities could play a relevant role in the transition. The impact of missing variables
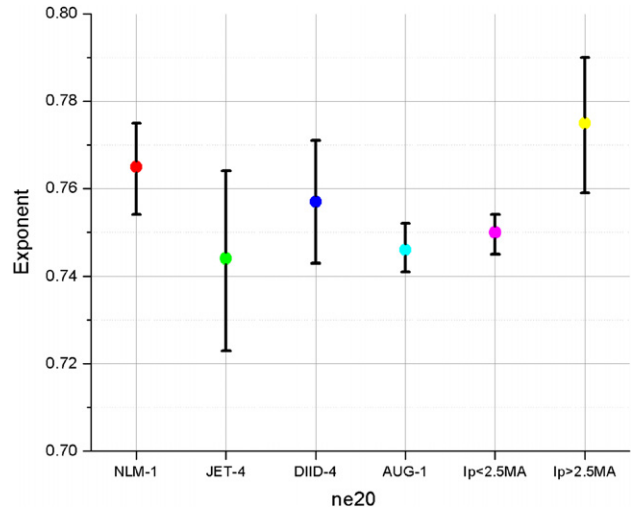

Figure 12. Exponents of the density for the proposed scaling law, with the 95% confidence interval.

could be different in different devices, resulting in different scaling laws if they are ignored.

In any case, the results reported in this paper are more aimed at describing a methodology than at providing a definitive answer to this physics question. In order to clarify this point and determine to what extent the physics of the L to H transition is the same on different machines, a more thorough investigation is required and in particular a specific database should be built for this purpose. The delicate interpretation of the obtained scaling laws is stressed by the analysis at different currents. The scaling law obtained from the entire database at plasma currents below 2.5 MA does not extrapolate very well to higher currents. The number of high current
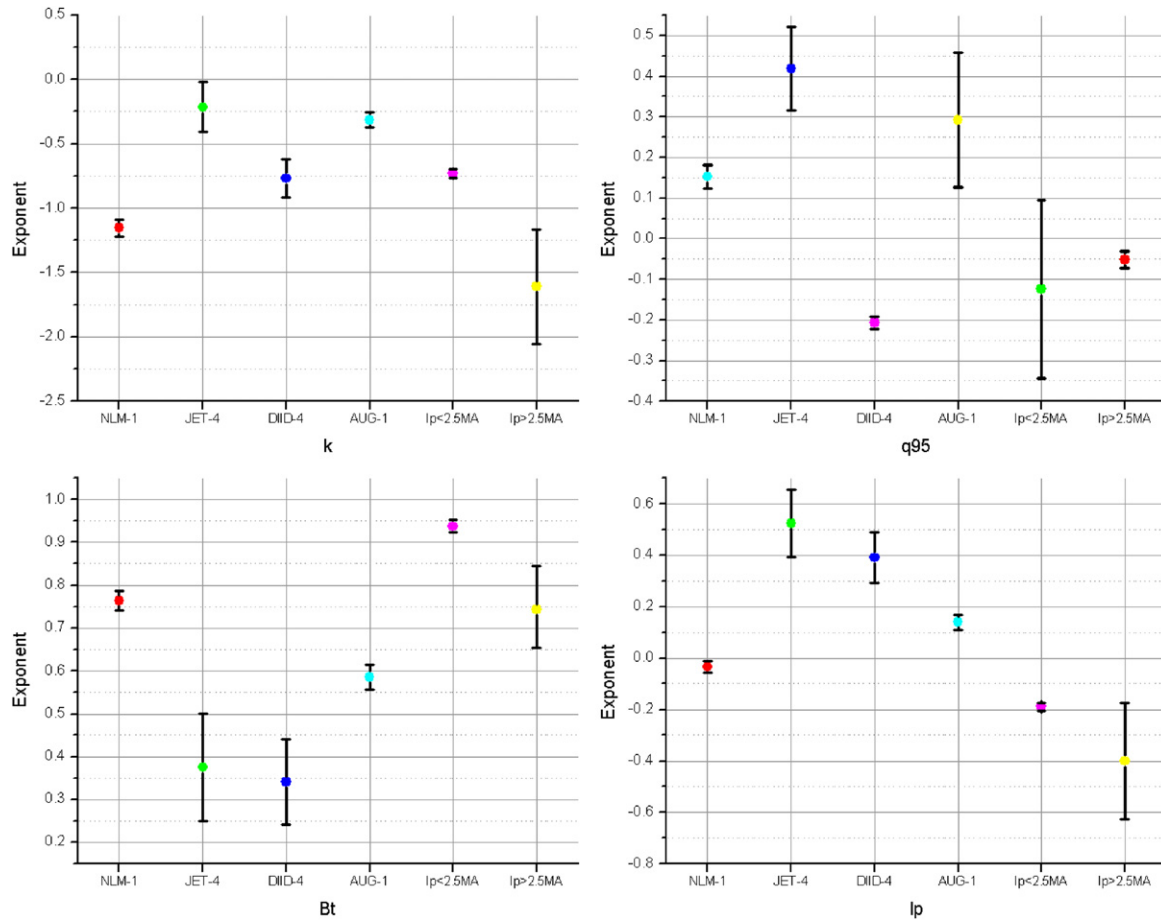
Figure 13. Exponents of k, q_{95}, B_t and I_p for the proposed scaling laws, with the 95% confidence interval.

discharges in the database is not huge but on the other hand the extrapolation is over a very small range. Therefore the extrapolation capability of scaling laws obtained from multi-machine databases remains a point to be further assessed.

References

[1] Sornette D. 2004 *Critical Phenomena in Natural Sciences* (*Springer Series in Synergetics*) (Zuerich: Springer) p 206
[2] Barenblatt G.I. 2003 *Scaling* (Edinburgh: Cambridge University Press)
[3] Wagner F. *et al* 1982 *Phys. Rev. Lett.* **49** 1408
[4] Wesson J. 2004 *Tokmaks* 3rd edn (Oxford: Oxford University Press)
[5] Connor J.W. and Wilson H.R. 2000 *Plasma Phys. Control. Fusion* **42** R1
[6] Martin Y.R. *et al* 2008 *J. Phys.: Conf. Ser.* **123** 012033
[7] McQuarrie A.D.R. and Tsai C.-L. 1998 *Regression and Time Series Model Selection* (London: World Scientific)
[8] Dillon W. and Goldstein M. 1984 *Multivariate Analysis: Methods and Applications* (New York: Wiley)
[9] Rencher A.C. 1995 *Methods of Multivariate Analysis* (New York: Wiley)
[10] McDonald D.C. *et al* 2006 *Plasma Phys. Control. Fusion* **48** A439
[11] Meakins A.J. *et al* 2005 *32nd EPS Conf. on Plasma Physics* vol 29C (ECA) p P-2.012 http://epsppd.epfl.ch/Tarragona/pdf/P2_012.pdf
[12] Bates D.M. and Watts D.G. 1988 *Nonlinear Regression Analysis and its Applications* (New York: Wiley)
[13] Lupelli I., Murari A., Gaudio P. and Gelfusa M. 2012 Statistical analysis of plasma shape influence on the power threshold to access the H-mode *Nucl. Instrum. Methods* A at press