

FIRST ANALYSIS OF THE UPDATED ITPA GLOBAL H-MODE CONFINEMENT DATABASE

G. VERDOOLAEGE
Ghent University and LPP – ERM/KMS
Ghent, Belgium
Email: geert.verdoolaege@ugent.be

S.M. KAYE
Princeton Plasma Physics Laboratory
Princeton, NJ, USA

C. ANGIONI, O.J.W.F. KARDAUN, F. RYTER, K. THOMSEN,
THE ASDEX UPGRADE TEAM* and THE EUROFUSION MST1 TEAM†
Max Planck Institute for Plasma Physics
Garching, Germany

M. MASLOV, M. ROMANELLI and JET CONTRIBUTORS‡
UKAEA, Culham Science Centre
Abingdon, UK

with acknowledgement to the contributors to the International H-Mode Confinement Public Domain Database:

CULHAM CENTER FOR FUSION ENERGY, Abingdon, UK (COMPASS, JET, MAST, START)
ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE, Lausanne, Switzerland (TCV)
GENERAL ATOMICS, San Diego, CA, USA (DIII-D)
HYDRO-QUEBEC – CENTRE CANADIEN DE FUSION MAGNETIQUE, Varennes, Canada (TdeV)
IOFFE INSTITUTE, St. Petersburg, Russia (TUMAN-3M)
JAPAN ATOMIC ENERGY AGENCY, Naka, Japan (JT-60U, JFT-2M)
KURCHATOV INSTITUTE, Moscow, Russia (T-10)
MAX PLANCK INSTITUTE FOR PLASMA PHYSICS, Garching, Germany (ASDEX, ASDEX Upgrade)
PLASMA SCIENCE AND FUSION CENTER, MIT, Cambridge, MA, USA (Alcator C-Mod)
PRINCETON PLASMA PHYSICS LABORATORY, Princeton, NJ, USA (NSTX, PDX, PBX-M, TFTR)

Abstract

First results are presented of the ongoing analysis of the latest version of the ITPA global H-mode confinement database, following addition in 2017 of new data from JET with the ITER-like wall and ASDEX Upgrade with the full tungsten wall. The main characteristics of the data are presented and the regression methods are discussed that are used for estimation of the global energy confinement scaling in terms of engineering parameters. Employing a simple power law model for the scaling and several different methods, regression analysis has been performed of the latest version of the database and various constrained subsets. Initial results of the scaling analysis are presented and compared with the IPB98(y,2) scaling expression.

1. INTRODUCTION

Global energy confinement studies based on empirical scaling expressions obtained from multi-machine datasets are one of the main approaches for extrapolating plasma performance to new machines such as ITER. In addition, these scaling expressions provide a reference for assessing the quality of global confinement in present experiments, they may serve as boundary conditions for modelers and can guide development of theoretical models. Since 1989, the H-Mode Database Working Group has developed and maintained the Global H-Mode Confinement Database, moving into the framework of the International Tokamak Physics Activity (ITPA) in 2001. Following the example of the L-mode database, a multimachine H-mode database was established for tokamak confinement scaling, in collaboration with teams from the various data-contributing devices. In 1998, version 3 of this database (DB3) was used to derive the ITER Physics Basis ELM_y H-mode scaling expression IPB98(y,2), which has been extensively used as a reference for global thermal energy confinement scaling in

* See the author list of “A. Kallenbach *et al.* Nuclear Fusion **57** (2017) 102015”.

† See the author list of “H. Meyer *et al.*, Nucl. Fusion **57** (2017), 102014”.

‡ See the author list of “X. Litaudon *et al.*, Nucl. Fusion **57** (2017), 102001”.

tokamak ELMy H-mode plasmas [1]. The IPB98(y,2) scaling relation was derived by means of regression analysis on one of the standard subsets, called DB2v8, of the full database DB3, using the following power law model:

$$\tau_{E,\text{th}} = \alpha_0 I_p^{\alpha_I} B_t^{\alpha_B} \bar{n}_e^{\alpha_n} P_{1,\text{th}}^{\alpha_P} R^{\alpha_R} \kappa^{\alpha_\kappa} \epsilon^{\alpha_\epsilon} M_{\text{eff}}^{\alpha_M}. \quad (1)$$

Here, $\tau_{E,\text{th}}$ (s) is the thermal energy confinement time, I_p the plasma current (MA), B_t the on-axis vacuum toroidal magnetic field (T), \bar{n}_e the central line-averaged electron density (10^{19} m^{-3}), $P_{1,\text{th}}$ the thermal part of the power lost through the last-closed flux surface (MW), R the major radius of the geometric axis (m), $\kappa = b/a$ the plasma elongation (where $2a$ (m) is the width of the plasma cross-section in the mid-plane and $2b$ (m) its height between the upper plasma edge and the lower edge (or the X-point)), ϵ the inverse aspect ratio a/R , and M_{eff} the effective atomic mass of the plasma. In the IPB98(y,2) expression, the plasma elongation was defined as $S/\pi a^2$ (with S the area of the plasma cross-section) to improve predictions for START and TdeV, while maintaining a consistent estimate of α_ϵ , irrespective of inclusion or exclusion of PBX-M data. However, in this paper we use the standard definition – a difference to be kept in mind when comparing with IPB98(y,2). Depending on the goals of a particular study, the global confinement scaling can either be expressed in terms of global engineering parameters, as in (1), or using dimensionless plasma parameters, by transformation of the exponents $\alpha_0, \alpha_I, \dots, \alpha_M$ of the engineering scaling relation.

Over the years, the ITPA global H-mode confinement database has been extended and, since about a decade, it contains data from 19 devices of different sizes and shapes. The latest fully public version of the database, DB3v13F, was described in [2] and can be accessed online [3]. Addition of further data from JET and the low aspect ratio devices NSTX [4] and MAST [5] led to DB4v3 [6] and DB4v5 [7]. Recently, however, experimental evidence has been collected that the coverage by the standard subset of the database, on which the IPB98(y,2) scaling is based, could be improved in certain regions of the parameter domain expected to be relevant for operation of future fusion reactors. This particularly concerns regimes with high density, low safety factor q_{95} and high normalized pressure β . Another important issue is the observed discrepancy between some of the dependencies found by IPB98(y,2) and corresponding single-machine scans. This is especially the case with the density dependence and the level of power degradation, where single-machine scans tend to point out weaker dependencies [8, 9]. Furthermore, as the majority of the data in the existing database was obtained in carbon-based machines, recent availability of data from devices with fully metallic walls, which are more reactor-relevant, suggests revisiting the confinement scaling issue.

Within the ITPA framework, an activity is ongoing with the aim to update the confinement database with data closer to ITER baseline and hybrid conditions, to expand the parameter range and include new data from devices with metallic walls, to explore predictor variables and possibly decouple core and pedestal scaling, and to employ more advanced regression techniques with an emphasis on increasing the robustness of the scaling expression. The present paper reports on the status of the database and presents a first set of results of power law regression analysis for estimation of the global energy confinement scaling in terms of engineering parameters. The analysis is following a stepwise approach, starting from simple models. Here, focus lies on the regression methodology and on the influence on the scaling expression of data from the machines with full metallic wall components.

2. DESCRIPTION OF THE DATABASE

Since the start of the initiative to update the H-mode confinement database in 2015, data has been added from JET with the ITER-like wall (JET-ILW) and from ASDEX Upgrade (AUG) with the full tungsten wall (AUG-W). The new JET data comprises 627 time slices obtained during the stationary phase of H-mode discharges [10, 11], while the new AUG data consists of 825 H-mode slices [12]. This newly added data also includes hybrid and ITER baseline scenarios. Together with the existing entries from Alcator C-Mod, with its molybdenum first wall components, the new data from JET and AUG contain the sole entries in the database from metal-only devices (i.e. metallic components for the wall, limiters and divertor). For simplicity, this will be referred to here as the ‘high- Z ’ subset, with the rest of the data making up the ‘low- Z ’ set.¹ In addition, some of the older data has been reprocessed, employing enhanced data validation criteria and more accurate estimates of fast-particle losses. Starting with these updates, the main version number of the database has been augmented, and it is presently being referred to as DB5v7.

¹ This is not a strict distinction (and not directly related to Z_{eff}), as some of the ‘low- Z ’ data originates from devices with *partly* metallic components, while C-Mod and (mostly) AUG operate with boronized wall conditioning.

In total, DB5v7 contains 13913 records (data points) with data from 19 tokamaks, but only dedicated subsets are considered for scaling analysis. The selection criteria for the ‘standard’ DB5 subset, referred to here as ‘STD5’ (7294 points from 18 machines), are similar to those applied to earlier database versions, and can be summarized as follows [2]. The data pertains to H-mode plasmas only (with ELMs or ELM-free), without pellet fueling and without strong internal transport barriers, characterized by relatively steady energy content, with limited radiated power, fast particle energy content and internal inductance, with electron and ion temperatures not differing more than by a factor of 2.5, and with a machine-dependent minimum safety factor. STD5 consists of the low-Z data that satisfied the standard selection criteria in the earlier database version, as well as the new high-Z discharges and a number of complementary low-Z discharges from JET and AUG with high gas injection rates. Many discharges have contributed multiple time slices. Apart from the STD5 subset of DB5, additional selection criteria have led to two more, restricted subsets of STD5, with a view to more ITER-relevant predictions. Subset STD5-SEL1 contains 5956 points from eight devices (AUG, C-Mod, COMPASS, DIII-D, JET, JFT-2M, JT60-U and PBX-M) and imposes the following additional constraints: $q_{95} > 2.8$, $1.3 < \kappa < 2.2$, $\epsilon < 0.5$ and $Z_{\text{eff}} < 5$, for the ion effective charge Z_{eff} . Subset STD5-SEL2 further concentrates on the current ITER standard reference parameters by imposing $2.8 < q_{95} < 3.5$, $1.6 < \kappa < 2.0$, $0.28 < \epsilon < 0.385$ and $Z_{\text{eff}} < 3$, resulting in 1674 data points from five machines (AUG, C-Mod, COMPASS, DIII-D and JET).

Over 200 variables are defined in the database, including various bookkeeping variables used for selecting data subsets. In an attempt to clarify the dependence of energy confinement on plasma density and fueling, in DB5 a number of variables have been added pertaining to the neutral gas pressure and electron density near the last-closed flux surface or in the scrape-off layer. On the one hand, these parameters can be controlled more directly than the line-averaged density. On the other hand, in JET with the ILW and in AUG-W, relatively strong gas fueling has been necessary to prevent heavy impurity accumulation, in turn affecting plasma confinement [13, 14]. Therefore, including these additional variables may help to improve physics understanding and predictions toward ITER. In the present database they are only represented in the newly contributed data from JET and AUG. Additional variables that are being considered for confinement scaling pertain to plasma rotation and torques applied externally by the neutral beams, while in this paper we investigate the influence of the average triangularity δ , as an alternative to q_{95}/q_{cyl} [6].

Dependencies between the scaling predictor variables in STD5 can guide data selection and provide some insight into the results of regression analysis. Motivated by the linearity of a power law on a logarithmic scale, linear correlations between the logarithmically transformed variables provide an idea of the strength of such relations. In STD5, significant correlation remains between plasma current and loss power (0.83), current (or power) and major radius (0.69), and also between the shape parameters κ and δ (0.73). Although the situation is similar in the version of the database before 2017, there are some improvements w.r.t. subset DB2v8 used for the IPB98 scaling expression. Specifically, the substantial correlation between plasma current (or power) on the one hand and shape parameters and effective mass on the other hand, has decreased. Another reduction concerns the correlation between current and inverse aspect ratio ϵ , which has come down from 0.79 in DB2v8 to 0.49 in STD5.

3. GLOBAL CONFINEMENT SCALING

The statistical model used for regression analysis comprises, as usual, a deterministic component, i.e. the actual scaling expression, and a stochastic component, both of which we now discuss.

3.1. Deterministic model component

In pursuing the analysis of the updated database, the classic power-law dependence of the confinement time has been assumed – more complex models will be considered later. One of the key criteria used is robustness of the scaling relation, in addition to goodness-of-fit, especially with a view to extrapolation. As such, a relatively simple model like the power law, being linear on the logarithmic scale, is a good starting point for capturing the data pattern. At this stage of the analysis, the usual set of predictor variables of Eq. (1) has been selected, in addition to the average triangularity δ . The latter occurs through a factor $(1 + \delta)^{\alpha\delta}$, in order to enable regression including data from circular plasmas. Scaling with dimensionless quantities will be considered at a later stage.

3.2. Stochastic model component

The stochastic model component serves to model the data variability not explained by the scaling expression. The analysis is complicated by the heterogeneity of the data, involving measurements from multiple devices. In

particular, it is important to note that there is variation in the data at various levels: within individual discharges, within the data from individual machines, and between devices. Clearly, the deterministic component (i.e. the scaling expression) of any practical regression model can only describe part of this variability. Any remaining variability is to be described by the stochastic component of the model, often a Gaussian distribution. An important aspect of the between-device variability – also noted in the analysis results below – is that the dependence of confinement on plasma parameters, as modeled by the scaling expression, is only broadly similar across devices. In addition to uncertainty owing to the simplified description by the scaling expression of the physics behind global energy confinement scaling in tokamak plasmas, this contributes to regression model uncertainty. Estimates of the uncertainty on the measured database variables are provided in the database. This includes statistical, and possibly systematic uncertainty arising from the measurement, but not some of the additional sources of uncertainty just mentioned. Errors on derived variables have been calculated by Gaussian error propagation.

Additional complications arise from the fact that ordinary least squares (OLS) regression assumes that there is no uncertainty on the predictor variables, or, in practice, that their uncertainty is negligible w.r.t. that on the response variable $\tau_{E,th}$. This approximation is not valid for the loss power $P_{l,th}$, for which the percentage error, averaged over the STD5 subset, is 12%, compared to 19% for $\tau_{E,th}$. Already recognized in [15] and [16], this observation motivates application of regression techniques that can account for uncertainty in all variables.

Correlations between the predictor variables and, in some cases, limited ranges of predictor variables, may further complicate the scaling analysis. In the case of a power law, this adds uncertainty to the estimates of the corresponding exponents. In particular, correlations between I_p , B_t , $P_{l,th}$ and R often result in noticeable constraints on (linear) combinations of their exponents (e.g. $\alpha_I + \alpha_B \sim 1$). Fortunately, it follows from regression theory that this does not need to have a large effect on predictions from the scaling expression – a fact confirmed by the majority of observations in this work. This does not prevent that even small changes in some of the scaling exponents, while keeping the others fixed, can yield significantly different predictions towards ITER. This may be relevant when exploring new operational scenarios deviating from the correlation structure in the database.

3.3. Analysis methods

In the past, most results of global confinement scaling, including IPB98(y,2), have been obtained using OLS regression on the log-transformed power law. In addition, since the number of data points contributed by each machine varies greatly, concerns have been raised that devices which contribute many similar discharges may unduly influence the fit. For this reason, apart from regular OLS, weighted least squares (WLS) has been used as well, with weights for point i from device j given by $w_{ij} = 2 + \sqrt{n_j}/4$, where n_j is the number of points contributed by device j [1]. Alternative weighting schemes are being evaluated.

While OLS has the advantage of simplicity, it is worth exploring other techniques, as some assumptions underlying OLS might not be fulfilled [17, 18]. In addition, concerns have been raised regarding the common practice of logarithmic transformation of the data prior to regression analysis, as this skews the likelihood distribution [15, 18]. We have implemented a robust Bayesian method (its results will be reported elsewhere), as well as geodesic least squares (GLS) with a Gaussian likelihood. GLS is a recently developed technique, sharing with OLS the benefit of simplicity, but being robust against deviations from the model assumptions and taking into account errors in all variables [18]. The method also accommodates heterogeneous errors and has been shown to yield consistent results, irrespective of log-transformation of the data. GLS learns for each device the scatter of the data around the hypersurface corresponding to the scaling expression, and compares this with the error bar given in the database, considered as a single standard deviation of a Gaussian distribution. As such, GLS contains an intrinsic weighting scheme, self-consistently accommodating deviation of the measurements from the pattern suggested by the data as a whole. Most error bars in the database are given as a percentage error, but for GLS to converge the absolute error bars on κ and M_{eff} had to be fixed at a typical value (0.05 for κ , 0.20 for M_{eff}).

Apart from parameter estimates, predictions $\hat{\tau}_{10}$ for the thermal confinement time in ITER (the subscript referring to the density value) have been derived by means of the estimated scaling expressions. The plasma parameters for the standard inductive $Q = 10$ scenario in ITER are $I_p = 15$ MA, $B_t = 5.3$ T, $\bar{n}_e = 10.3 \times 10^{19} \text{ m}^{-3}$, $P_{l,th} = 87$ MW, $R = 6.2$ m, $\delta_{95} = 0.33$, $\kappa_{95} = 1.7$, $\epsilon = 0.32$ and $M_{eff} = 2.5$ [19]. It is stressed that these predictions should be regarded in the context of the ongoing analysis activities. In any case, extrapolations from global confinement scaling expressions have to be treated with caution due to the various uncertainties involved, including physics-related uncertainties owing to different neutral penetration in ITER, ELM control schemes, etc.

The analysis also intends to report error bars on the parameter estimates and ITER predictions, in the form of 95% confidence intervals. It should be mentioned that more realistic confidence intervals are probably larger and remain to be studied in more detail. For this reason, in this paper the parameter estimates and predictions are consistently reported with two significant digits, and the confidence intervals with matching precision, or at least one significant digit. Furthermore, in its current form, GLS is an optimization technique returning point estimates for the model parameters. Therefore, confidence intervals for GLS estimates have been estimated using a bootstrapping (resampling) technique with 100 samples.

3.4. Confinement scaling analysis and results

In the remainder of the paper, first results are reported of regression analysis for estimating the confinement scaling using the updated confinement database, by means of unconstrained OLS, WLS and GLS. The analysis has followed the usual practice of log-linear scaling, considering that nonlinear GLS regression on the original scale yields very similar results. The IPB98(y,2) parameter estimates and ITER prediction $\hat{\tau}_{10}$ are mentioned in Table 1 as a reference [1]. Fig. 1 shows a plot of the confinement enhancement factor $H98(y, 2) = \tau_{E,th}/\hat{\tau}_{10,98}$ vs. Greenwald fraction \bar{n}_e/n_{GW} for the STD5-SEL1 data set (with $\hat{\tau}_{10,98}$ the predictions by IPB98(y,2)), highlighting the purely metallic devices. It can be seen that, depending on plasma conditions, IPB98 tends to overpredict confinement when approaching the Greenwald limit, for both low-Z and purely metallic devices [12, 13]. The decreasing trend has been described by a log-quadratic term in [6].

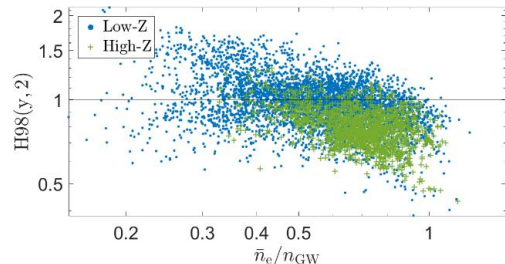


FIG. 1. *H-factor w.r.t. the IPB98(y,2) scaling vs. Greenwald fraction in STD5-SEL1.*

TABLE 1. THE IPB98(y,2) PARAMETER ESTIMATES AND ITER PREDICTION

	α_0	α_I	α_B	α_n	α_P	α_R	α_κ	α_ϵ	α_M	$\hat{\tau}_{10}$ (s)
IPB98(y,2)	0.0562	0.93	0.15	0.41	-0.69	1.97	0.78	0.58	0.19	3.62

3.4.1 Single-device analysis

The analysis has started with examination of individual device scalings, both for low-Z and high-Z walls. This can give insight into similarities and differences between these individual datasets, and may lead to the resolution of some of the discrepancies with single-machine scans mentioned above.

Table 2 presents estimates and 95% confidence intervals for a subset of machines represented in STD5. The number of observations in STD5 from each device is also mentioned, together with two (frequentist) quantities that can be used to compare the quality of different fits. As a measure of goodness-of-fit, the root-mean-square error from the OLS analysis is reported. In addition, the coefficient of determination R^2 is provided as a measure of how much of the data variability is explained by the scaling model, compared to a ‘null’ model containing no dependence on the predictor variables. Similar quantities can be calculated for the other methods. In the single-device scalings most results of GLS are quite similar to those of OLS, therefore at this point we only present the latter. Also, a distinction is made between the low-Z and high-Z data from AUG and JET. Each scaling expression only involves those predictor variables that vary sufficiently within the corresponding dataset. The following observations can be made on the basis of these results, with a special focus on the ITER-like devices (AUG, C-Mod, DIII-D and JET).

- The scaling with plasma current is similar for the ITER-like devices ($\alpha_I \sim 1.1$), except for a somewhat stronger scaling in AUG ($\alpha \sim 1.5$). The scaling with current is generally weaker for the other machines.
- The B_t -dependence is weak in the ITER-like devices, with the strongest (negative) dependence for AUG ($\alpha_B \sim -0.3$). The I_p and B_t scaling in AUG are being investigated [12].
- The density dependence is weak in the ITER-like devices, although slightly positive in JET-C ($\alpha_n \sim 0.3$). In the smaller or more circular devices (ASDEX, JFT-2M, PDX), the dependence is stronger ($\alpha_n \sim 0.38 - 0.68$), which may have influenced the non-negligible density dependence in the IPB98(y,2) scaling [20].
- The power degradation for the carbon-based ITER-like devices is in the range $\alpha_P \sim -0.76$ (JET-C), to -0.66 (AUG-C, DIII-D), but somewhat weaker in the metallic devices: -0.60 (C-Mod) to -0.53 (AUG-

W). The strong dependence in JET-C may be linked to its strong density dependence, as for this device there is a sizeable correlation (0.65) between $\ln \bar{n}_e$ and $\ln P_{l,th}$.

One difference between OLS and GLS that is worth mentioning is the dependence on isotope mass, which GLS often estimates to be stronger than the dependence given by OLS, although the confidence intervals are quite large. Nevertheless, similar observations can be made in case of the multi-machine datasets.

TABLE 2. REGRESSION RESULTS FOR INDIVIDUAL DEVICES IN STD5

Device	N_{obs}	α_I	α_B	α_n	α_P	α_δ	α_κ	α_M	RMSE	R^2
ASDEX	575	0.69 ± 0.11	0.13 ± 0.16	0.68 ± 0.08	-0.65 ± 0.05	-	-	0.78 ± 0.12	0.17	0.71
AUG-C	1384	1.5 ± 0.07	-0.26 ± 0.08	0.033 ± 0.037	-0.66 ± 0.03	-	-	-	0.17	0.73
AUG-W	767	1.6 ± 0.07	-0.30 ± 0.09	0.055 ± 0.057	-0.53 ± 0.03	-	-	-	0.11	0.89
Alcator C-Mod	82	1.1 ± 0.2	-	0.10 ± 0.18	-0.60 ± 0.15	-	-	-	0.10	0.79
DIII-D	502	1.1 ± 0.09	0.080 ± 0.11	0.10 ± 0.06	-0.67 ± 0.04	0.70 ± 0.16	-	0.45 ± 0.11	0.20	0.80
JET-C	2235	1.1 ± 0.04	0.16 ± 0.04	0.31 ± 0.02	-0.76 ± 0.02	-	1.2 ± 0.2	0.053 ± 0.044	0.16	0.89
JET-ILW	600	1.1 ± 0.1	-0.16 ± 0.08	0.072 ± 0.057	-0.57 ± 0.03	-	-	0.40 ± 0.05	0.11	0.86
JFT-2M	348	0.99 ± 0.08	-	0.38 ± 0.08	-0.86 ± 0.04	-	-	0.11 ± 0.06	0.10	0.93
JT60-U	100	0.78 ± 0.29	0.47 ± 0.38	-0.18 ± 0.17	-0.35 ± 0.13	-	-	-	0.14	0.83
MAST	43	1.1 ± 0.9	-	0.17 ± 0.30	-0.86 ± 0.31	-	-	-	0.12	0.60
NSTX	230	0.29 ± 0.14	1.2 ± 0.2	0.58 ± 0.15	-0.84 ± 0.08	-	0.81 ± 0.35	-	0.14	0.74
PBX-M	214	0.61 ± 0.31	-	-0.073 ± 0.086	-0.56 ± 0.07	-	-	-	0.12	0.72
PDX	119	0.62 ± 0.32	0.63 ± 0.32	0.62 ± 0.16	-1.1 ± 0.15	-	-	-	0.18	0.68

3.4.2 Multi-machine scalings

Next, results are presented based on several multi-machine datasets, starting with the full STD5 set in Table 3. With reference to the IPB98(y,2) scaling expression, the following observations can be made (keeping in mind the distinct elongation definitions and the fact that IPB98(y,2) pertains to ELM y H-modes only).

- The scaling with plasma current is somewhat stronger in STD5 compared to IPB98, accompanied by an almost vanishing dependence on toroidal field.
- The density dependence is somewhat weaker in STD5 than IPB98, with a similar level of power degradation (OLS and WLS).
- Compared to IPB98, there is a clearly weaker dependence on R and ϵ in STD5, with α_ϵ close to zero (this is likely to be at least partly related to the different choice of elongation definition [1]).
- Within STD5, WLS gives a slightly weaker dependence on I_p and $P_{l,th}$, but a stronger dependence on \bar{n}_e , in comparison with the other two methods. In addition, GLS gives the strongest power degradation and isotope dependence. The predictions by the three methods are similar, but up to 25% lower than that by IPB98(y,2).

TABLE 3. REGRESSION RESULTS FOR STD5 (RMSE = 0.19, $R^2 = 0.96$)

Method	α_0	α_I	α_B	α_n	α_P	α_R	α_κ	α_ϵ	α_M	$\hat{\tau}_{10}$ (s)
OLS	0.049	1.1	0.085	0.19	-0.71	1.5	0.80	-0.043	0.25	2.7
	± 0.002	± 0.02	± 0.020	± 0.02	± 0.01	± 0.04	± 0.04	± 0.046	± 0.03	± 0.1
WLS	0.040	0.99	0.11	0.29	-0.64	1.7	0.79	0.093	0.25	2.9
	± 0.002	± 0.03	± 0.02	± 0.02	± 0.01	± 0.04	± 0.04	± 0.046	± 0.03	± 0.1
GLS	0.042	1.2	0.068	0.21	-0.78	1.6	0.88	-0.052	0.47	2.7
	± 0.003	± 0.02	± 0.016	± 0.01	± 0.01	± 0.03	± 0.06	± 0.027	± 0.07	± 0.03

In the following step, the analysis focuses on the ITER-like subset STD5-SEL1, yielding the numerical results listed in Table 4. Fig. 2 shows a plot of the experimental confinement times in STD5-SEL1 vs. those predicted by OLS, confirming the overall good fit. The main differences with the total STD5 set are the even stronger dependence on current, even weaker scaling with density and radius, but stronger dependence on the shape parameters κ (note the considerable error bar, however) and ϵ – particularly according to WLS and GLS – but now with negative β_ϵ . Quite remarkable is also the dependence on isotope mass, which has become weaker in the results from WLS, but stronger according to GLS. The ITER predictions remain more-or-less the same. However, as the subset STD5-SEL1 does not include spherical or circular machines, these scaling expressions cannot be used for confinement prediction in devices with values of κ or ϵ outside the specified range (which encompasses the ITER shape parameters).

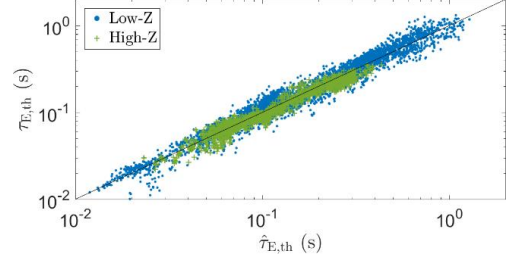

 FIG. 2. Experimental $\tau_{E,th}$ vs. predictions $\hat{\tau}_{E,th}$ using OLS regression on STD5-SEL1.

 TABLE 4. REGRESSION RESULTS FOR STD5-SEL1 (RMSE = 0.18, $R^2 = 0.96$)

Method	α_0	α_I	α_B	α_n	α_P	α_R	α_κ	α_ϵ	α_M	$\hat{\tau}_{10}$ (s)
OLS	0.045	1.3	-0.10	0.13	-0.71	1.2	1.1	-0.32	0.24	2.6
	± 0.005	± 0.03	± 0.04	± 0.02	± 0.01	± 0.06	± 0.1	± 0.05	± 0.04	± 0.1
WLS	0.030	1.3	-0.069	0.19	-0.64	1.3	1.3	-0.46	0.094	3.0
	± 0.005	± 0.04	± 0.056	± 0.05	± 0.03	± 0.1	± 0.2	± 0.08	± 0.055	± 0.2
GLS	0.023	1.3	-0.018	0.17	-0.79	1.5	1.9	-0.38	0.33	2.5
	± 0.007	± 0.04	± 0.067	± 0.03	± 0.02	± 0.1	± 0.4	± 0.08	± 0.13	± 0.1

Bringing in the average triangularity as a predictor variable (in practice $1 + \delta$) slightly increases R^2 from 0.96 to 0.97. Estimates of the α_δ exponent vary between 0.6 and 0.8 (± 0.06), suggesting a general beneficial influence of increased triangularity on the energy confinement. The other exponents and ITER predictions retain a similar value within the STD5-SEL1 set.

Regressions have also been performed on the low-Z and high-Z subsets of STD5-SEL1. The correlations in the high-Z subset are somewhat too strong to allow solid conclusions, but comparison of the regression results between the full STD5-SEL1 set and its low-Z subset (not shown here) appears possible. The main differences are a slightly weaker dependence on I_p , δ and ϵ in the low-Z subset, and a somewhat stronger dependence on density (although unchanged for WLS).

When considering the ITER-like STD5-SEL2 set, regressions (without κ and ϵ dependence due to insufficient variation) point out a slightly stronger dependence on density and R compared to STD5-SEL1, roughly at the level of the full STD5 set. There is also a stronger power degradation ($\alpha_P \sim -0.7$ to -0.8), but a weaker dependence on δ ($\alpha_\delta \sim 0.2$ to 0.4). The power degradation estimated by WLS is the most favorable of the three methods, largely explaining an ITER confinement prediction of 3.5 s, which is roughly at the same level of IPB98(y,2). Similar results can be obtained with GLS, using a separate offset for the JET-ILW data.

4. CONCLUSIONS

First results have been presented of an ongoing activity involving an update of the ITPA global H-mode confinement database and analysis of global confinement scaling in tokamaks. Data has been added from recent operation of JET with the fully metallic ITER-like wall and ASDEX Upgrade with the full tungsten wall. Various issues have been reviewed that may complicate fitting of global thermal energy confinement scalings, potentially adding to the uncertainty of parameter estimates and predictions towards ITER. The present work has focused on data weighting and robust estimation techniques to increase confidence in the scaling results, as well as on the influence of the new data on confinement scaling expressions. The analysis reveals considerable differences in the scaling exponents between individual machines, although generally the density dependence is weakest in the ITER-like devices and the level of power degradation lowest in the metallic machines. Multi-machine scaling expressions on the constrained STD5-SEL1 subset confirm several of the confinement dependencies given by the IPB98(y,2) scaling expression, while others call for more detailed analysis. In addition to the plasma shape parameters already used as predictor variables in IPB98(y,2), the plasma triangularity appears to have a non-negligible influence on confinement. On the other hand, the dependence on density and magnetic field turns out to be weak in the results of multi-machine power law scalings reported here. However, the B_t dependence is compensated by a generally stronger dependence on plasma current. Additional analysis is ongoing, including studies of the dependence of the regression results and ITER predictions on selection criteria, regression model and regression variables. Further extension of the database by inclusion of recent data from other tokamaks would also be most valuable.

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 or the ITER Organization. This work was conducted under the auspices of the ITPA Topical Group on Transport and Confinement. The research was also partly supported by U.S. DOE Contract DE-AC02-09CH11466.

REFERENCES

- [1] ITER Physics Basis: Chapter 2, Nucl. Fusion **39** 12 (1999) 2175–2249.
- [2] MCDONALD, D.C. *et al.*, Nucl. Fusion **47** 3 (2007), 147–174.
- [3] International Global H-Mode Confinement Database (2009), <http://efdasql.ipp.mpg.de/HmodePublic>.
- [4] KAYE, S.M. *et al.*, Plasma Phys. Control. Fusion **48** 5A (2006), A429–A438.
- [5] VALOVIC, M. *et al.*, Nucl. Fusion **45** 8 (2005), 942–949.
- [6] KARDAUN, O.J.W.F. *et al.*, Proc. 21st IAEA Fusion Energy Conference (Chengdu, PRC), IT/P1-10 (2006). <http://www.ipp.mpg.de/~Otto.Kardaun/netreports/hmws2.pdf>.
- [7] KARDAUN, O.J.W.F., IPP-Report 2017-05 (2017).
- [8] CHALLIS, C.D. *et al.*, Nucl. Fusion **55** 5 (2015), 053031 (18 pp.)
- [9] RYTER, F. *et al.*, Nucl. Fusion **41** 5 (2001), 537–550.
- [10] MASLOV, M. *et al.*, Proc. 45th EPS Conf. on Plasma Physics (Prague, Czech Republic), ECA, Vol. 42A, P5.1064 (2018).
- [11] MASLOV, M. *et al.*, Nucl. Fusion, in preparation.
- [12] RYTER, F. *et al.*, Nucl. Fusion, in preparation.
- [13] NUNES, I. *et al.*, Plasma Phys. Control. Fusion **58** 1 (2016), 014034 (10 pp.).
- [14] SCHWEINZER, J., Nucl. Fusion **51** 11 (2011), 113003 (7 pp.).
- [15] MCDONALD, D.C. *et al.*, Plasma Phys. Control. Fusion **48** 5A (2006), A439–A447.
- [16] CORDEY, J.G. *et al.*, Nucl. Fusion **45** 9 (2005), 1078–1084.
- [17] VERDOOLAEGE, G. *et al.*, Nucl. Fusion **55** 11 (2015), 113019 (19 pp.).
- [18] VERDOOLAEGE, G., Entropy **17** 7 (2015), 4602–4626.
- [19] HAWRYLUK, R.J. *et al.*, Nucl. Fusion **49** 6 (2009), 065012 (15 pp.).
- [20] KAYE, S.M. *et al.*, 60th Annual Meeting of the APS Division of Plasma Physics (Portland, OR, USA, 2018).