Confinement scaling with machine size in the updated ITPA global H-mode energy confinement database

J. Hall[1], P. Zhang[1], Y. Zhang[1], C. Angioni[2], G. Verdoolaege[1], JET Contributors* and EUROfusion Tokamak Exploitation Team**

[1]*Department of Applied Physics, Ghent University, 9000 Ghent, Belgium*
[2]*Max-Planck-Institut für Plasmaphysik, Boltzmannstrasse 2, D-85748, Germany*
See the authorlist of C. Maggi et al., Nucl. Fusion **64, 112012 (2024)*
***See the authorlist of E. Joffrin et al., Nuclear Fusion **64**, 112019 (2024)*

Introduction

Energy confinement scaling laws, derived strictly from experimental observations, are important tools for quantifying, comparing and predicting confinement performance in magnetic confinement fusion devices. The International Tokamak Physics Activity (ITPA) global H-mode energy confinement database, which in 1998 yielded the widely adopted IPB98(y,2) scaling [1], has since been substantially expanded. Its most recent version, DB5.2.3, adds data from devices with fully metallic plasma-facing components, specifically JET with the ITER-like wall (JET-ILW) and ASDEX Upgrade with a full tungsten wall (AUG-W). Regression on the "standard set" of this database produced the ITPA20 scaling for ELMy H-mode plasmas [2], which for a representative ITER baseline scenario predicts $\tau_E = 3.07 \pm 0.46$ s, about 15% below the 3.62 s of IPB98(y,2).

Among the most notable distinctions between ITPA20 and IPB98 is a pronounced reduction in the dependence of τ_E on the major radius R_{geo}: the exponent decreases from $\alpha_R = 1.97$ in IPB98(y,2) to 1.71 ± 0.32 in ITPA20, and further to 1.19 ± 0.27 in the ITPA20-IL variant. We show that this reduction forms part of a longer-term trend visible, under certain conditions, well before the latest database iteration. We assess the influence of multicollinearity among predictors on the weakening of α_R and find that, although present, it cannot account for the reduction by itself. Using data-driven optimisation and classification, we then isolate a subset of discharges specifically responsible for weakening the machine-size dependence, show that this subset is best characterised in the dimensionless space of normalised pressure β_N, collisionality ν_*, safety factor q_{95} and gyroradius ρ_*, and provide updated τ_E projections for ITER and SPARC.

Data sets and scaling model

Three versions of the database are compared throughout: DB2.8 [1] (1998), DB4.5 [3] (2004) and DB5.2.3 [2] (2020), restricted to ELMy H-mode plasmas from their standard sets. Spherical tokamaks are excluded owing to their distinct transport physics, at the expense of reducing the

range of inverse aspect ratio ε. Consistent subsets are extracted for direct comparison: $\mathrm{Std2}_{\mathrm{sub}}$ is the intersection of Std2 with both Std4 and Std5 (1244 entries), and $\mathrm{Std4}_{\mathrm{sub}}$ the intersection of Std4 with Std5 (2533 entries); the full Std5 contains 5919 entries across 15 devices. The points newly added at each transition, $\mathrm{Std4}^{\mathrm{new}}$ (1289 entries) and $\mathrm{Std5}^{\mathrm{new}}$ (3386 entries), are the subject of the optimisation described below.

The thermal energy confinement time $\tau_{E,\mathrm{th}}$ is described throughout by the power law

$$\tau_{E,\mathrm{th}}^{\mathrm{fit}} = \alpha_0 \, I_p^{\alpha_I} \, B_t^{\alpha_B} \, \bar{n}_e^{\alpha_n} \, P_{l,\mathrm{th}}^{\alpha_P} \, R_{\mathrm{geo}}^{\alpha_R} \, \kappa^{\alpha_\kappa} \, M_{\mathrm{eff}}^{\alpha_M}, \tag{1}$$

where I_p (MA) is the plasma current, B_t (T) the toroidal field, $\bar{n}_e$ ($10^{19}\,\mathrm{m}^{-3}$) the line-averaged density, $P_{l,\mathrm{th}}$ (MW) the thermal power loss, R_{geo} (m) the major radius, $\kappa = V/(2\pi^2 R_{\mathrm{geo}} a^2)$ the elongation and M_{eff} the effective mass. The inverse aspect ratio ε usually retained in such models is dropped here following the removal of spherical devices. All regressions use ordinary least squares (OLS).

The OLS coefficients shown in Table 1 make the downward trend of α_R clearly visible: it falls from 1.7 in $\mathrm{Std2}_{\mathrm{sub}}$ to 1.4 in $\mathrm{Std4}_{\mathrm{sub}}$ and 1.3 in Std5, while model quality stays high ($R^2 \geq 0.96$).

Multicollinearity analysis

Multicollinearity, defined as a strong linear dependence among predictor variables, can inflate the variances of estimated coefficients and thereby destabilise the regression estimates. We quantified its severity via eigenvalue decomposition of the Gram matrix $X^{\mathsf{T}}X$ for each dataset, computing condition indices and variance decomposition proportions [4]. In all three model variants, two principal directions exhibit substantial multicollinearity (CI ≥ 30), involving predominantly the intercept, I_p, $\bar{n}_e$, and R_{geo}; the corresponding variance inflation factors for I_p and R_{geo} are consistently greater than 5. Despite these elevated VIFs, the standard errors of the coefficients remain small, attributable to the large sample sizes and low residual variances. Multicollinearity therefore contributes to parameter instability but cannot, by itself, account for the observed progressive decrease in α_R. This interpretation is corroborated by the similar multicollinearity structure found in the influential and non-influential data subsets identified below.

Identification of influential subsets

Recognizing that multicollinearity is not the sole determinant of the observed effect, our data-driven procedure isolates the subset of observations that most strongly contributes to the reduction in α_R. Specifically, we partition $\mathrm{Std4}^{\mathrm{new}}$ into "influential" (I) and "non-influential" (NI) components such that, when combined with $\mathrm{Std2}_{\mathrm{sub}}$, the NI subset preserves a size dependence consistent with the original scaling law, whereas the I subset accounts for the observed deviation. An analogous procedure is applied to $\mathrm{Std5}^{\mathrm{new}}$ in conjunction with $\mathrm{Std4}_{\mathrm{sub}}$.

Each data point is assigned a weight that quantifies its probability of belonging to the influential subset; further methodological details are provided in [5]. Candidate influential subsets

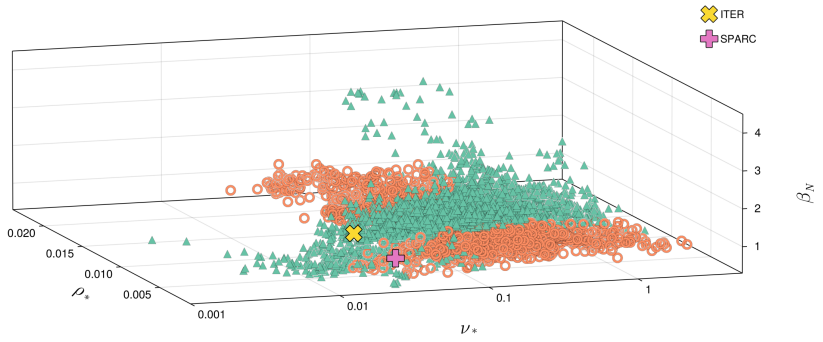

Figure 1: Std5 in the dimensionless space of ρ_*, ν_* and β_N, colour-coded by classifier prediction: non-influential points as green triangles, influential points as red open circles. The operational points of ITER (gold $\times$) and SPARC (magenta $+$) are included.

are subsequently generated via weighted sampling, and a multi-objective optimization framework is employed to identify solutions that concurrently minimize both the influential scaling exponent α_R^{I} and the cardinality of the influential subset [5].

Table 1: OLS regression coefficients for the base datasets $\mathrm{Std2_{sub}}$, $\mathrm{Std4_{sub}}$ and Std5 under the log-linear model (1). The major-radius exponent α_R (bold) decreases progressively from 1.7 to 1.3 while model quality stays high.

Dataset	α_0	α_I	α_B	α_n	α_P	α_R	α_κ	α_M	R^2
$\mathrm{Std2_{sub}}$	0.040	1.1	0.085	0.32	-0.62	**1.7**	0.69	0.11	0.98
$\mathrm{Std4_{sub}}$	0.055	1.2	-0.11	0.27	-0.57	**1.4**	0.57	0.21	0.97
Std5	0.071	1.2	-0.096	0.17	-0.65	**1.3**	0.67	0.30	0.96

The non-influential sets preserve a size dependence close to the base dataset, while the influential points drive α_R down to ≈ 1.2 in both transitions.

Classification in dimensionless space

To characterise the influential points, a random forest classifier was trained to distinguish between the I and NI subsets. Prior application of the Synthetic Minority Oversampling Technique (SMOTE) [6] was employed to mitigate the pronounced class imbalance. Feature selection was conducted independently for each transition. Remarkably, both transitions, $\mathrm{Std2_{sub}} \to \mathrm{Std4_{sub}}$ and $\mathrm{Std4_{sub}} \to$ Std5, yielded the same four optimal predictors (β_N, ν_*, q_{95}, ρ_*, in decreasing order of importance), despite being derived from non-overlapping datasets. The classifier achieved balanced accuracies of 77% and 74%, respectively.

Figure 1 shows the Std5 classification on the (ρ_*, ν_*, β_N) axes, chosen to emphasise spatial clustering. Two well-localised regions of influential points emerge: one at low β_N with intermediate-to-high ν_*, and another at elevated β_N and low collisionality. Non-influential points

lie predominantly between these clusters, with both ITER and SPARC closest to the non-influential regime [5].

Scaling laws and predictions

Regressing the two classes of the full Std5 separately yields distinct relations. The non-influential set gives

$$\tau_{E,\mathrm{th}}^{\mathrm{NI}} = 0.04\, I_p^{1.1} B_t^{0.065} \bar{n}_e^{0.35} P_{l,\mathrm{th}}^{-0.71} R_{\mathrm{geo}}^{1.7} \kappa^{0.77} M_{\mathrm{eff}}^{0.14}, \quad \mathrm{RMSE} = 0.14, \tag{2}$$

preserving the near-quadratic size dependence of IPB98(y,2), while the influential set gives

$$\tau_{E,\mathrm{th}}^{\mathrm{I}} = 0.177\, I_p^{1.5} B_t^{-0.25} \bar{n}_e^{-0.19} P_{l,\mathrm{th}}^{-0.62} R_{\mathrm{geo}}^{0.56} \kappa^{0.73} M_{\mathrm{eff}}^{0.54}, \quad \mathrm{RMSE} = 0.15, \tag{3}$$

with a substantially weaker dependence ($\alpha_R = 0.56$) and notable changes in several other exponents consistent with findings in other published works [7]. The influential scaling should be regarded as descriptive rather than predictive.

The classifier assigns both ITER and SPARC an 84% probability of belonging to the non-influential class. The non-influential scaling predicts $\tau_{E,\mathrm{th}}^{\mathrm{NI}} = 3.8$ s for ITER and 0.88 s for SPARC, comparable to or slightly above the corresponding IPB98(y,2) predictions of 3.6 s and 0.76 s and more optimistic than ITPA20, whereas the influential scaling reduces both by roughly a factor of two (1.9 s and 0.46 s).

Conclusion

Across three database versions α_R falls monotonically from 1.7 to 1.3, a trend that multi-collinearity cannot explain alone. A multi-objective optimisation isolates a compact influential subset driving α_R to ≈ 1.2. A random forest classifier then demonstrates the influential subsets are localised in a region in $(\rho_*, \nu_*, q_{95}, \beta_N)$ space. ITER and SPARC are classified non-influential (with 84%) probability, giving predictions of 3.8 s and 0.88 s for standard operational points. These results delineate a restricted domain of applicability for the IPB98(y, 2) sclanning and offer an explanation for the reduced size dependence of ITPA20-IL.

Acknowledgements

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 10105220 – EUROfusion). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission.

References

[1] ITER Physics Basis Editors, Nucl. Fusion **39**, 2175 (1999)

[2] G. Verdoolaege et al., Nuclear Fusion **61**, 076006 (2021)

[3] D.C McDonald et al., Nuclear Fusion, **47** 147, (2007)

[4] D.A. Belsley, *Computer Science in Economics and Management* **4**, 33 (1991)

[5] J. Hall et al., Nuclear Fusion (2026) (under review)

[6] N.V. Chawla et al., J. Artif. Intell. Res. **16**, 321 (2002)

[7] C. Angioni et al., Nuclear Fusion **66**, 026006 (2026)