The updated ITPA global H-mode confinement database: description and analysis

G Verdoolaege[1], SM Kaye[2], C Angioni[3], OJWF Kardaun[3], M Maslov[4], M Romanelli[4], F Ryter[3], K Thomsen[3], the ASDEX Upgrade Team†, the EUROfusion MST1 Team‡ and JET Contributors§

[1]Department of Applied Physics, Ghent University, 9000 Ghent, Belgium
[2]Princeton Plasma Physics Laboratory, Princeton University, Princeton, NJ 08543, United States of America
[3]Max-Planck-Institut für Plasmaphysik, D-85748 Garching, Germany
[4]United Kingdom Atomic Energy Authority, Culham Centre for Fusion Energy, Culham Science Centre, Abingdon, OX14 3DB, United Kingdom of Great Britain and Northern Ireland

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

CULHAM CENTRE FOR FUSION ENERGY, Abingdon, UK (COMPASS-D, JET, MAST, START)
ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE, Lausanne, Switzerland (TCV)
MULTI-INSTITUTIONAL DIII-D TEAM LED BY 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)

E-mail: `geert.verdoolaege@ugent.be`

January 2021

Abstract. The multi-machine ITPA Global H-mode Confinement Database has been upgraded with new data from JET with the ITER-like wall and ASDEX Upgrade with

†See the author list of 'H. Meyer et al., Nucl. Fusion 59 (2019), 112014'.
‡See the author list of 'B. Labit et al., Nucl. Fusion 59 (2019) 086020'.
§See the author list of 'E. Joffrin et al., Nucl. Fusion 59 (2019), 112021'.

the full tungsten wall. This paper describes the new database and presents results of regression analysis to estimate the global energy confinement scaling in H-mode plasmas using a standard power law. Various subsets of the database are considered, focusing on type of wall and divertor materials, confinement regime (all H-modes, ELMy H or ELM-free) and ITER-like constraints. Apart from ordinary least squares, two other, robust regression techniques are applied, which take into account uncertainty on all variables. Regression on data from individual devices shows that, generally, the confinement dependence on density and the power degradation are weakest in the fully metallic devices. Using the multi-machine scalings, predictions are made of the confinement time in a standard ELMy H-mode scenario in ITER. The uncertainty on the scaling parameters is discussed with a view to practically useful error bars on the parameters and predictions. One of the derived scalings for ELMy H-modes on an ITER-like subset is studied in particular and compared to the IPB98(y, 2) confinement scaling in engineering and dimensionless form. Transformation of this new scaling from engineering variables to dimensionless quantities is shown to result in large error bars on the dimensionless scaling. Regression analysis in the space of dimensionless variables is therefore proposed as an alternative, yielding acceptable estimates for the dimensionless scaling. The new scaling, which is dimensionally correct within the uncertainties, suggests that some dependencies of confinement in the multi-machine database can be reconciled with parameter scans in individual devices. This includes vanishingly small dependence of confinement on line-averaged density and normalized plasma pressure (β), as well as a noticeable, positive dependence on effective atomic mass and plasma triangularity. Extrapolation of this scaling to ITER yields a somewhat lower confinement time compared to the IPB98(y, 2) prediction, possibly related to the considerably weaker dependence on major radius in the new scaling (slightly above linear). Further studies are needed to compare more flexible regression models with the power law used here. In addition, data from more devices concerning possible 'hidden variables' could help to determine their influence on confinement, while adding data in sparsely populated areas of the parameter space may contribute to further disentangling some of the global confinement dependencies in tokamak plasmas.

1. Introduction

The dominant mechanism determining the energy confinement time in tokamaks is heat conduction and convection due to turbulent transport. Reliable predictions of the turbulent thermal transport coefficients in tokamak plasmas from first principles are still not feasible, despite substantial progress in theoretical understanding over the past decades. Consequently, global energy confinement studies based on empirical scaling expressions obtained from multi-machine data sets still represent an important instrument 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-day experiments, they may serve as boundary conditions for modelers and can guide development of theoretical models for heat transport in tokamaks.

Since 1989, the H-Mode Database Working Group (DBWG) 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 [1], a multi-machine H-mode database was established for tokamak confinement scaling, in collaboration with teams from the various data-contributing devices [2]. In 1998, version 2.8 of this database (DB2.8) was used to derive the ITER Physics Basis ELMy H-mode scaling expression IPB98(y, 2), which has been extensively used as a reference for global thermal energy confinement scaling in tokamak ELMy H-mode plasmas [3].

Over the years, the ITPA global H-mode confinement database was 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 [4]. Addition of further data from JET and the low-aspect ratio devices NSTX [5] and MAST [6] led to DB4.3 [7] and DB4.5 [8].

Recently, however, experimental evidence has been collected suggesting that the coverage by the standard subset of the database, on which the IPB98(y, 2) scaling is based, may 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 q_{95} (safety factor at the surface encompassing 95% of the poloidal magnetic flux) and high normalized plasma pressure β. Another important issue is the observed disparity between some of the dependencies found by IPB98(y, 2) and those obtained in corresponding single-machine scans. This is especially the case with the density dependence and the level of power degradation, where single-machine scans sometimes tend to point out weaker dependencies [9, 10]. Discrepancies with scans of individual dimensionless variables have also long been noted [11]. 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 considered to be more reactor-relevant, suggests revisiting the confinement scaling issue.

Within the ITPA framework, an activity has been conducted with the aim to update the confinement database with data closer to ITER baseline and hybrid scenario conditions, to expand the parameter range and include new data from devices with metallic walls, to explore new predictor variables and to employ several advanced regression techniques which aim at increasing the robustness of the scaling expression [12]. The present paper reports on the status of the database and presents a number of results of power law regression analysis for estimation of the global energy confinement scaling in terms of engineering variables and dimensionless quantities. Specific emphasis is laid on the uncertainty on the data and the estimated parameters, as well as on the influence on the scaling expression of data from the machines with fully metallic wall components. We concentrate mainly on confinement scaling in ELMy H-mode plasmas and compare with the IPB98(y, 2) scaling.

The paper is structured as follows. Section 2 presents the database, its additions since 2015 and the variables in the energy confinement scaling. In section 3, the methodology of the regression analysis is introduced in order to derive the global confinement scaling in power law form. The transformation between engineering and

dimensionless variables is discussed in some detail, as well as the three regression techniques that have been used in the paper. The results of the regression analysis, using various subsets of the database, are presented in section 4, with some emphasis on deriving practical uncertainty estimates. In section 5 the main results of the paper are summarized and discussed, while the main conclusions are offered in section 6.

2. Description of the database

2.1. Data size and subsets

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 866 time slices obtained during the stationary phase of H-mode discharges [13, 14], while the new AUG-W data consists of 825 H-mode slices [15]. This newly added data includes hybrid and ITER baseline scenarios, among others. Together with the existing entries from Alcator C-Mod, with its molybdenum first wall components, the new data from JET-ILW and AUG-W comprise 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. Even in the high-Z subset, the wall cannot invariably be characterized as fully metallic, because Alcator C-Mod and ASDEX Upgrade employ boronization for wall conditioning. Along with addition of data from metallic devices, some of the older data from ASDEX Upgrade has been reprocessed, employing enhanced data validation criteria and a more accurate estimate of the fast ion energy content. Starting with these updates, the main version number of the database has been augmented, such that presently it is referred to as 'DB5'. The version of the database used for analysis in this paper is DB5.2.3 and it is publicly accessible online [16].

In total, DB5.2.3 contains 14 153 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' (7 537 points from 18 machines), are similar to those applied to earlier database versions, and can be summarized as follows [2, 4]. The constraints target discharges that are in H-mode plasmas only (with ELMs or ELM-free), without pellet fueling, strong internal transport barriers or excessive MHD activity near the β limit, characterized by a relatively steady energy content ($-0.05 \leq (\mathrm{d}W_{\mathrm{tot}}/\mathrm{d}t)/P_{\mathrm{abs,t}} \leq 0.35$, where W_{tot} is the total plasma stored energy and $P_{\mathrm{abs,t}}$ the total absorbed heating power), with limited radiated power P_{rad} ($P_{\mathrm{rad}}/P_{\mathrm{abs,t}} \leq 0.6$) and limited fast particle energy content W_{f} ($W_{\mathrm{f}}/W_{\mathrm{tot}} \leq 0.40$), and with a machine-dependent minimum safety factor ($q_{95} \gtrsim 2.5$). STD5 consists of the data that satisfied the standard selection criteria in the earlier database version DB3, as well as the majority of the new high-Z discharges (866 from JET-ILW and 767 from AUG-W). A number of complementary low-Z discharges from JET and ASDEX Upgrade with high

gas injection rates have been added as well. Many discharges have contributed multiple time slices. Both large type I ELMs and small ELMs occur in the ELMy H-modes in the database.

Apart from the STD5 subset of DB5, additional selection criteria have led to a restricted subset of STD5, with a view to more ITER-relevant predictions. The ITER-like subset contains 6 202 slices from H-mode plasmas in eight devices (ASDEX Upgrade, Alcator C-Mod, COMPASS-D, DIII-D, JET, JFT-2M, JT-60U and PBX-M) and imposes the following additional constraints: $q_{95} > 2.8$, $1.3 < \kappa < 2.2$, $\epsilon < 0.5$ and $Z_{\rm eff} < 5$. Here, $\kappa = b/a$ is the elongation of the last-closed flux surface (LCFS; with $2a$ (m) the width of the surface in the mid-plane and $2b$ (m) the height between the upper plasma edge and the lower edge (or the X-point)), $\epsilon = a/R_{\rm geo}$ is the inverse aspect ratio (with $R_{\rm geo}$ (m) the radius of the geometric axis) and $Z_{\rm eff}$ is the ion effective charge.

Table 1 lists the number of entries in DB2.8, DB5.2.3-STD5 and various subsets of the latter. A distinction has been made between the data points from ASDEX Upgrade with the full tungsten wall and the other points from this machine (henceforth referred to as points from AUG-W and AUG, respectively). Likewise, for the purposes of this paper, the points from JET with the carbon wall (JET-C) and JET-ILW are distinguished, also because of the differing operational characteristics of the machine during the respective periods.

In comparison with DB2.8, DB5.2.3 contains significantly more points, from various devices, with the majority of new entries originating from ASDEX Upgrade and JET. Some data from COMPASS-D and START have been added as well, which can contribute to further narrowing down the dependence of confinement on machine size. Addition of data from the low-aspect ratio machines NSTX and MAST (large ϵ) can be useful for determining the ϵ scaling [5].

2.2. Database variables

Over 200 variables are defined in the database, including various bookkeeping variables used for selecting data subsets. We here concentrate on the variables that appear in the expression of the global energy confinement scaling (see section 3).

2.2.1. Engineering variables Starting with the 'engineering' form of the confinement scaling using a simple power law, the following variables are of particular interest: the thermal energy confinement time $\tau_{\rm E,th}$ (s), the plasma current $I_{\rm p}$ (MA), the on-axis vacuum toroidal magnetic field $B_{\rm t}$ (T), the central line-averaged electron density $\bar{n}_{\rm e}$ (10^{19} m^{-3}), the thermal power lost due to transport through the LCFS $P_{\rm l,th}$ (MW), the major radius $R_{\rm geo}$ (m), the elongation of the LCFS, defined as $\kappa_{\rm a} = V/(2\pi R_{\rm geo}\pi a^2)$ (with V (m^3) the plasma volume inside the LCFS), the inverse aspect ratio $\epsilon = a/R_{\rm geo}$ and the effective atomic mass $M_{\rm eff}$ of the plasma.

The energy confinement time is given by $\tau_{\rm E,th} = W_{\rm th}/P_{\rm l,th}$, where $W_{\rm th}$ is the thermal stored energy. The latter is derived from the total stored energy $W_{\rm tot}$ (itself

Table 1. Number of data entries in DB2.8 and DB5.2.3 listed per machine, for various data subsets.

	DB2.8	DB5.2.3			
		STD5		STD5 ITER-like	
Device	ELMy H	All H	ELMy H	ELMy H	ELM-free H
ASDEX	431	575	431	0	0
AUG	102	1 385	1 377	1 370	8
AUG-W	0	767	767	767	0
Alcator C-Mod	37	82	45	45	37
COMPASS-D	0	21	16	16	5
DIII-D	270	502	388	383	114
JET-C	246	2 211	1 762	1 606	426
JET-ILW	0	866	866	855	0
JFT-2M	59	348	69	59	197
JT-60U	9	100	100	100	0
MAST	0	43	43	0	0
NSTX	0	230	185	0	0
PBX-M	59	214	59	59	155
PDX	97	119	97	0	0
START	0	8	8	0	0
T-10	0	4	0	0	0
TCV	0	17	11	0	0
TdeV	0	7	7	0	0
TEXTOR	0	0	0	0	0
TFTR	0	2	2	0	0
TUMAN-3M	0	36	0	0	0
Total	1 310	7 537	6 233	5 260	942

determined from diamagnetic loop measurements or from an equilibrium reconstruction) by subtracting the energy content associated to fast particles originating from plasma heating. Alternatively, $W_{\rm th}$ is deduced from kinetic measurements. Likewise, the thermal power loss $P_{\rm l,th}$ is obtained from the total power loss by subtracting the power lost through charge exchange reactions and unconfined ion orbits, as well as the neutral beam shine-through power. It should be noted that the power lost through radiation from inside the LCFS was neglected for calculating $P_{\rm l,th}$ (this cannot be neglected in ITER). Furthermore, as before, a correction factor 'TAUC93' was applied for calculating $\tau_{\rm E,th}$ in the case of ASDEX and PDX, to account for their closed divertor shapes [3]. The definition used for the elongation is similar to the one used in the IPB98(y, 2) expression. It was chosen instead of the standard definition $\kappa = b/a$ to improve predictions for START and TdeV, while maintaining very similar estimates of the impact of ϵ, irrespective of inclusion of data from PBX-M, with its indented plasma shape [3].

In an attempt to clarify the dependence of energy confinement on plasma density and fueling, in DB5 variables have been added corresponding to the electron density near the LCFS (separatrix density $n_{\rm e,sep}$, corresponding variable NESEP) or in the scrape-

Table 2. Main statistics of the engineering variables used for global confinement scaling in ELMy H-mode plasmas in the DB5.2.3-STD5 database. The corresponding statistics for the case of the DB2.8 data set are mentioned between parentheses.

	$\tau_{\mathrm{E,th}}$ (s)	I_{p} (MA)	B_{t} (T)	$\bar{n}_{\mathrm{e}}$ (10^{19} m^{-3})	$P_{\mathrm{l,th}}$ (MW)	R_{geo} (m)	$1+\delta$	κ_{a}	ϵ	M_{eff}
Min.	0.0022	0.16	0.26	1.2	0.15	0.28	1.0	0.92	0.16	1.0
	(0.014)	(0.17)	(0.94)	(1.2)	(0.39)	(0.67)	(1.0)	(0.93)	(0.16)	(1.0)
Max.	1.3	5.1	5.8	43	33	3.4	1.9	2.5	0.78	3.2
	(1.3)	(5.1)	(5.8)	(43)	(21)	(3.4)	(1.9)	(2.5)	(0.41)	(2.0)
Mean	0.18	1.4	2.2	6.0	8.0	2.2	1.3	1.6	0.32	1.9
	(0.15)	(1.1)	(2.2)	(5.5)	(4.5)	(1.8)	(1.2)	(1.4)	(0.29)	(1.8)
Median	0.12	1.0	2.2	5.5	6.6	1.7	1.3	1.6	0.32	2.0
	(0.055)	(0.62)	(2.2)	(4.4)	(2.8)	(1.7)	(1.2)	(1.5)	(0.29)	(2.0)
Std.	0.15	0.81	0.66	3.2	5.3	0.70	0.14	0.22	0.082	0.26
	(0.20)	(1.0)	(0.71)	(4.9)	(3.7)	(0.55)	(0.19)	(0.37)	(0.064)	(0.30)

off layer ($n_{\mathrm{e,sol}}$, corresponding variable NESOL). On the one hand, these parameters can be controlled more directly than the line-averaged density. On the other hand, in AUG-W and JET-ILW relatively strong gas fueling has been necessary to prevent heavy impurity accumulation, in turn affecting plasma confinement [17, 18]. Therefore, including these additional variables may help to improve physics understanding and predictions toward ITER. In the present database they are only provided for some of the data from AUG, AUG-W [15] and JET-ILW [14]. In addition, we investigate the influence on confinement of the average triangularity δ of the LCFS. This shape factor is used here as an alternative to q_{95}/q_{cyl}, where q_{cyl} is the safety factor in the cylindrical approximation [7].

Table 2 provides a summary of the main statistics of the engineering variables used in the confinement scaling, including triangularity, for all ELMy H-mode data across all devices in DB5.2.3-STD5. A comparison with the statistics in DB2.8 is also given. The corresponding histograms for DB5.2.3-STD5 are shown in figure 1. As far as the engineering variables are concerned, the main differences in DB5.2.3-STD5 with respect to DB2.8 are mostly due to addition of data at elevated power, and to some extent also higher current and density. This is related to the larger share of data from ASDEX Upgrade and JET in DB5.2.3. Furthermore, as mentioned above, addition of data from the low-aspect ratio machines MAST and NSTX has substantially increased the maximum value of ϵ in the database. Moreover, apart from deuterium (D) plasmas, several devices have contributed data from plasmas with hydrogen (H), or with a mixture of H and D (ASDEX, AUG, DIII-D, JET, JET-ILW, JFT-2M and PDX). There are also data from tritium (T) plasmas and D-T mixtures in JET. Finally, although the database contains entries obtained in helium plasmas, these have been left out in the present analysis of the database and all confinement scalings.

Dependencies between the predictor variables (also known as independent or

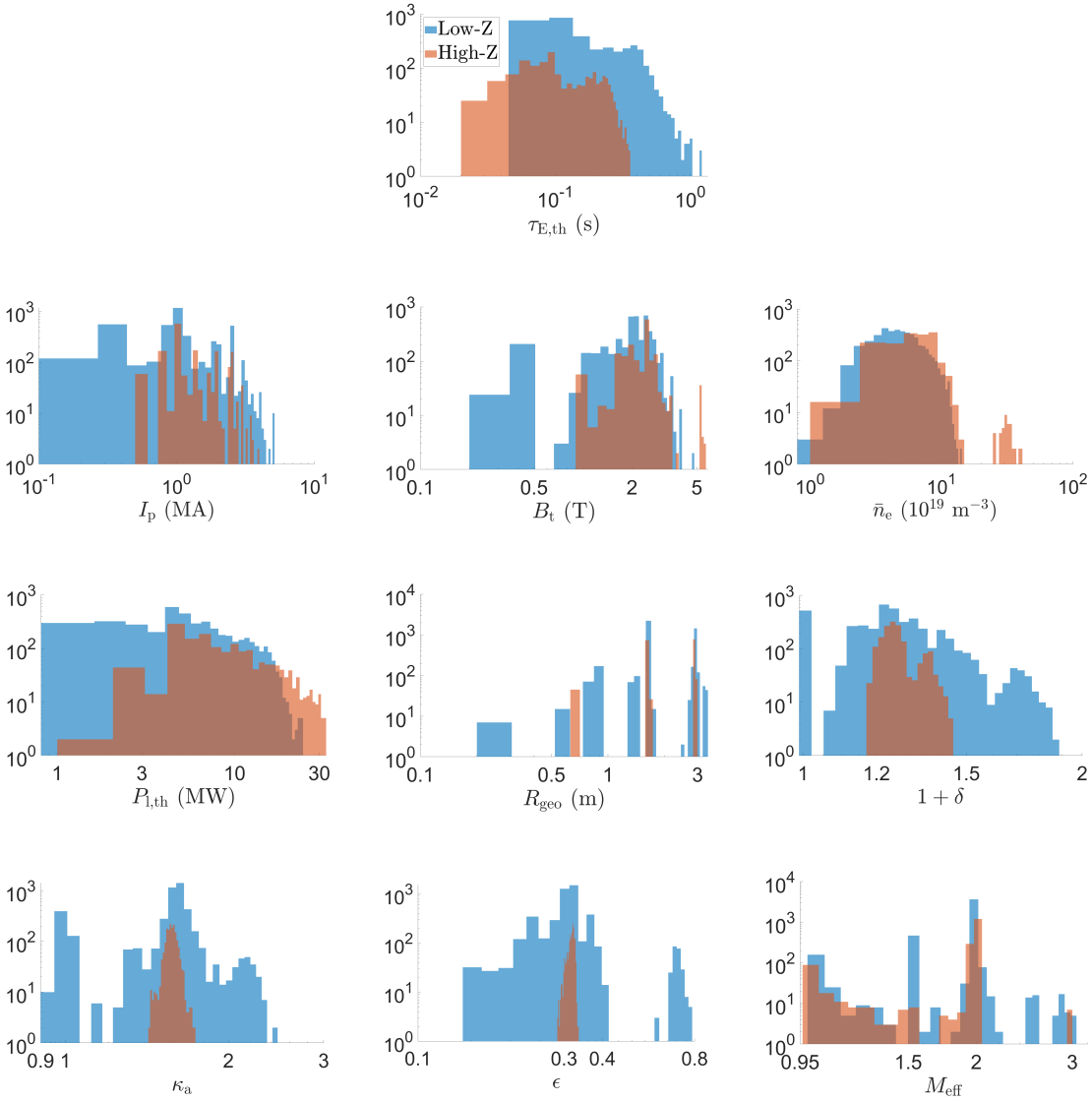

Figure 1. Histograms (logarithmic scale) for the distribution of the engineering variables used for global confinement scaling in ELMy H-mode plasmas in the DB5.2.3-STD5 database. A distinction has been made between the low-Z and high-Z subsets. The vertical axes denote the number of entries in the data set.

explanatory variables, or regressors) for the scaling in DB5.2.3-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, across the database, between the logarithmically transformed variables provide an idea of the strength of such relations (natural logarithms are used throughout this paper). Table 3 shows the correlation matrices for the logarithmic engineering variables in the ELMy H-mode data of DB5.2.3-STD5. If two predictor variables are strongly correlated, then the uncertainty on the estimates of the corresponding individual exponents in the scaling will be high (predictions would not necessarily be affected, however). It

Table 3. Correlation matrix for the logarithmic engineering variables used in global confinement scaling for ELMy H-mode plasmas in the DB5.2.3-STD5 database. Values above 0.6 are highlighted.

	$\ln I_\mathrm{p}$	$\ln B_\mathrm{t}$	$\ln \bar{n}_\mathrm{e}$	$\ln P_\mathrm{l,th}$	$\ln R_\mathrm{geo}$	$\ln(1+\delta)$	$\ln \kappa_\mathrm{a}$	$\ln \epsilon$	$\ln M_\mathrm{eff}$
$\ln I_\mathrm{p}$	**1.0**	0.34	0.083	0.82	0.69	0.34	0.49	0.30	0.33
$\ln B_\mathrm{t}$	0.34	**1.0**	0.27	0.34	0.37	−0.38	−0.26	−0.54	0.20
$\ln \bar{n}_\mathrm{e}$	0.083	0.27	**1.0**	0.20	−0.36	0.21	0.25	0.073	0.35
$\ln P_\mathrm{l,th}$	0.82	0.34	0.20	**1.0**	0.61	0.31	0.44	0.18	0.27
$\ln R_\mathrm{geo}$	0.69	0.37	−0.36	0.61	**1.0**	−0.085	0.020	−0.27	−0.0065
$\ln 1+\delta$	0.34	−0.38	0.21	0.31	−0.085	**1.0**	0.72	0.62	0.21
$\ln \kappa_\mathrm{a}$	0.49	−0.26	0.25	0.44	0.020	0.72	**1.0**	0.54	0.34
$\ln \epsilon$	0.30	−0.54	0.073	0.18	−0.27	0.62	0.54	**1.0**	0.23
$\ln M_\mathrm{eff}$	0.33	0.20	0.35	0.27	−0.0065	0.21	0.34	0.23	**1.0**

can be noted that, in DB5.2.3-STD5 ELMy H-mode plasmas, significant correlation remains between the plasma current and power loss (0.82), current (or power) and major radius (0.69), and also between the shape parameters κ_a, ϵ and $1+\delta$ (always on a logarithmic scale). Although the situation is similar in the version of the database before 2015, there are some improvements with respect to subset DB2.8 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 significantly.

In general, however, bivariate correlations reveal only part of the correlation structure between the total set of variables. Indeed, it may be possible to approximate a specific predictor variable by a combination of two or more other regressors (logarithmic scale). In such a case, the bivariate correlations among this set of variables could be low and still the uncertainty on the estimates of the individual exponents may be high. Therefore, a more detailed analysis of the correlation structure between the predictor variables was carried out using variance decomposition proportions based on singular value decomposition [19, 20]. This revealed evidence for two near-collinearities, involving the intercept, I_p, B_t, R_geo and ϵ, suggesting that, in the database, each of these variables can be expressed moderately well in terms of (a subset of) the remaining predictor variables, i.e. $\bar{n}_\mathrm{e}$, $P_\mathrm{l,th}$, $1+\delta$, κ_a and M_eff. It also indicates the possibility of considerable uncertainty on the parameter estimates corresponding to the variables involved in the near-collinearities.

2.2.2. Dimensionless variables The scaling can be cast in 'dimensionless' form by introducing a number of dimensionless variables. The motivation for this approach lies in dimensionless analysis of the transport of heat and particles, applied in an early stage in the context of magnetically confined plasmas by Kadomtsev [21]; see

also [22].¶ In particular, the energy confinement time is rendered dimensionless through multiplication with the ion cyclotron frequency $\Omega_\mathrm{i} = qB_\mathrm{t}/M_\mathrm{eff}$ (with charge $q = e$ for hydrogenic species). Other dimensionless quantities often used in this context are ρ_* (ion gyroradius ρ_i normalized to the minor radius), β_t (plasma pressure normalized to the magnetic pressure of the toroidal field), collisionality ν_* (ion collision frequency ν_ii normalized to the bounce frequency of trapped particles) and q_cyl (safety factor in the cylindrical approximation). These quantities can be calculated in terms of dimensional variables as follows:‖

$$\begin{aligned}
\rho_* &= \frac{\rho_\mathrm{i}}{a} \simeq \left(\frac{2m_\mathrm{p}}{e}\right)^{1/2} \frac{(M_\mathrm{eff}T)^{1/2}}{B_\mathrm{t}a} \\
&= 1.44 \times 10^{-4} (M_\mathrm{eff}T)^{1/2} (B_\mathrm{t}R_\mathrm{geo}\epsilon)^{-1}, && (1a) \\
\beta_\mathrm{t} &= \frac{p}{B_\mathrm{t}^2/(2\mu_0)} \simeq \frac{4\mu_0 neT}{B_\mathrm{t}^2} = 8.05 \times 10^{-21} nTB_\mathrm{t}^{-2}, && (1b) \\
\nu_* &= \nu_\mathrm{ii} \left(\frac{m_\mathrm{i}}{eT}\right)^{1/2} \left(\frac{R_\mathrm{geo}}{a}\right)^{3/2} q_\mathrm{cyl}R_\mathrm{geo} \simeq \frac{e^2}{2\pi 3^{3/2}\epsilon_0^2} \ln\Lambda \frac{nq_\mathrm{cyl}R_\mathrm{geo}^{5/2}}{T^2 a^{3/2}} \\
&= 5 \times 10^{-11} (\ln\Lambda) nB_\mathrm{t}R_\mathrm{geo}^2 \epsilon^{1/2} \kappa_\mathrm{a} I_\mathrm{p}^{-1} T^{-2}, && (1c) \\
q_\mathrm{cyl} &= \frac{2\pi}{\mu_0} \frac{a^2\kappa_\mathrm{a}}{R_\mathrm{geo}} \frac{B_\mathrm{t}}{I_\mathrm{p}} = 5 \times 10^6 B_\mathrm{t}RI_\mathrm{p}^{-1}\epsilon^2\kappa_\mathrm{a}. && (1d)
\end{aligned}$$

Here, p, n (m^{-3}) and T (eV) stand for the volume-averaged pressure, density and temperature, respectively, assuming equal electron and ion temperature. Furthermore, in these expressions the plasma current I_p is in A, m_i is the ion mass and m_p the proton mass. In the expression for ν_*, $\ln\Lambda$ is the Coulomb logarithm, for which we adopt the expression $\ln\Lambda = 30.9 - \ln(n^{1/2}/T)$. It typically does not vary much over the parameter range of interest. For instance, in the DB5.2.3-STD5 ELMy H-mode ITER-like data set, the average value of $\ln\Lambda$ is 15.5 and the standard deviation is 0.6. For the purpose of transforming the engineering scaling to dimensionless variables (section 3.1), $\ln\Lambda$ can thus be taken as a numerical constant. Moreover, in practice, in this paper β_t has been calculated from the thermal stored energy W_th ($= 3neT$, in J) and the plasma volume in the database. Also, the volume-averaged density has been obtained by multiplying the line-averaged electron density $\bar{n}_\mathrm{e}$ with a factor 0.88, which is an average conversion factor obtained from analysis of the database. In addition, although some of the scaling laws derived in this paper include the triangularity as a predictor variable, we do not consider δ in the definition of the dimensionless variables, in order to facilitate comparison with earlier results. For ease of reference, table 4 summarizes the proportionality relation between the main dimensionless and dimensional variables used in this paper.

Table 5 provides the statistics summary for the dimensionless variables used here, for the ELMy H-mode plasmas in DB5.2.3-STD5, including a comparison with DB2.8.

¶Readers interested in the origin and methods of dimensionless scaling techniques in magnetic confinement fusion may consult the excellent review article by Luce *et al.* [23].

‖The precise definitions of the dimensionless variables vary across the literature. Here, we largely follow [3], with definitions based on the toroidal magnetic field B_t.

Table 4. Relations (approximate) between the main dimensionless variables used in this paper and the dimensional variables.

Dimensionless variable	Proportional to	Description
ρ_*	$(M_{\rm eff}T)^{1/2}(B_{\rm t}R_{\rm geo}\epsilon)^{-1}$	Ion gyroradius normalized to minor radius
$\beta_{\rm t}$	$nTB_{\rm t}^{-2}$	Plasma pressure normalized to magnetic pressure
ν_*	$nB_{\rm t}R_{\rm geo}^2\epsilon^{1/2}\kappa_{\rm a}I_{\rm p}^{-1}T^{-2}$	Collisionality: collision frequency normalized to bounce frequency of trapped particles
$q_{\rm cyl}$	$B_{\rm t}RI_{\rm p}^{-1}\epsilon^2\kappa_{\rm a}$	Safety factor in cylindrical approximation

Table 5. Main statistics of the dimensionless variables used in global confinement scaling for ELMy H-mode plasmas in the DB5.2.3-STD5 database (addition of the radius $R_{\rm geo}$ is explained in section 3.1). The corresponding statistics for the case of the DB2.8 data set are mentioned between parentheses.

	$\Omega_{\rm i}\tau_{\rm E,th}$	ρ_* (10^{-2})	$\beta_{\rm t}$ (%)	ν_*	$q_{\rm cyl}$	$R_{\rm geo}$ (m)	$1+\delta$	$\kappa_{\rm a}$	ϵ	$M_{\rm eff}$
Min.	0.00058	0.23	0.17	0.0012	1.1	0.28	1.0	0.92	0.16	1.0
	(0.019)	(0.31)	(0.23)	(0.0030)	(1.7)	(0.67)	(1.0)	(0.93)	(0.16)	(1.0)
Max.	3.0	4.7	21	2.9	7.9	3.4	1.9	2.5	0.78	3.2
	(3.0)	(2.2)	(3.5)	(2.2)	(5.9)	(3.4)	(1.9)	(2.5)	(0.41)	(2.0)
Mean	0.33	0.64	1.8	0.18	2.8	2.2	1.3	1.6	0.32	1.9
	(0.30)	(0.71)	(1.2)	(0.20)	(2.8)	(1.8)	(1.2)	(1.4)	(0.29)	(1.8)
Median	0.21	0.55	1.3	0.12	2.7	1.7	1.3	1.6	0.32	2.0
	(0.11)	(0.66)	(0.96)	(0.15)	(2.7)	(1.7)	(1.2)	(1.5)	(0.29)	(2.0)
Std.	0.32	0.35	2.2	0.21	0.56	0.70	0.14	0.22	0.082	0.26
	(0.45)	(0.29)	(0.63)	(0.21)	(0.64)	(0.55)	(0.19)	(0.37)	(0.064)	(0.30)

Compared to DB2.8, the updates in version 5.2.3 of the database have the following effect on the range of these variables. Since data have been added from small machines, but still the share of data from the larger machines has increased, ρ_* reaches higher values compared to version 2.8, but the median ρ_* is still lower in DB5.2.3. Likewise, the addition of data from low-aspect ratio devices (MAST and NSTX) has considerably increased the range of $\beta_{\rm t}$, resulting in a higher median over the entire database. Also the range of $q_{\rm cyl}$ has increased (both to lower and higher values), whereas the distribution of ν_* has not changed substantially.

The correlation matrix for the logarithmic dimensionless variables in ELMy H-mode plasmas in DB5.2.3-STD5 is shown in table 6. For ρ_*, in comparison with DB2.8, the correlation with ϵ has come down in DB5.2.3-STD5, but the negative correlation with radius $R_{\rm geo}$ has become stronger, from -0.5 in DB2.8 to -0.76 in DB5.2.3-STD5. Addition of data from low-aspect ratio devices has increased the correlation between $\beta_{\rm t}$ and ϵ. This is confirmed by analysis of the variance decomposition proportions, indicating a single near-dependency involving at least the intercept, $\beta_{\rm t}$ and ϵ. On the other hand, the negative correlation between ν_* and machine size has become weaker.

Table 6. Correlation matrix for the logarithmic dimensionless variables (and radius $R_{\rm geo}$) used in global confinement scaling for ELMy H-mode plasmas in the DB5.2.3-STD5 database. Values above 0.6 are highlighted.

	$\ln\rho_*$	$\ln\beta_{\rm t}$	$\ln\nu_*$	$\ln q_{\rm cyl}$	$\ln R_{\rm geo}$	$\ln(1+\delta)$	$\ln\kappa_{\rm a}$	$\ln\epsilon$	$\ln M_{\rm eff}$
$\ln\rho_*$	**1.0**	0.56	0.00019	−0.26	−0.76	0.18	0.093	0.26	0.032
$\ln\beta_{\rm t}$	0.56	**1.0**	−0.24	−0.59	−0.25	0.64	0.60	0.68	0.17
$\ln\nu_*$	0.00019	−0.24	**1.0**	0.29	−0.36	−0.16	−0.17	−0.28	−0.13
$\ln q_{\rm cyl}$	−0.26	−0.59	0.29	**1.0**	−0.12	−0.20	−0.18	−0.12	0.048
$\ln R_{\rm geo}$	−0.76	−0.25	−0.36	−0.12	**1.0**	−0.085	0.020	−0.27	−0.0036
$\ln 1+\delta$	0.18	0.64	−0.16	−0.20	−0.085	**1.0**	0.72	0.62	0.21
$\ln\kappa_{\rm a}$	0.093	0.60	−0.17	−0.18	0.020	0.72	**1.0**	0.54	0.34
$\ln\epsilon$	0.26	0.68	−0.28	−0.12	−0.27	0.62	0.54	**1.0**	0.23
$\ln M_{\rm eff}$	0.032	0.17	−0.13	0.048	−0.0036	0.21	0.34	0.23	**1.0**

3. Global H-mode confinement scaling: variables and methods

In this section, the methodology is discussed for modeling the dependencies of the global energy confinement time on plasma conditions using a scaling law. In particular, the power law model is introduced to describe the relation between the global confinement and the engineering variables. This relation is then cast in a dimensionless form and compared with the trends estimated directly in dimensionless space. We also touch upon the statistical regression techniques employed to estimate the parameters in these scalings and for making predictions toward other experiments, notably ITER. The discussion is held brief—more details will be published elsewhere.

3.1. Scaling with engineering and dimensionless variables

The IPB98(y, 2) scaling relation was derived by means of regression analysis on the data in DB2.8, using the following power law model [3]:

$$\tau_{\rm E,th} = \alpha_0\, I_{\rm p}^{\alpha_I}\, B_{\rm t}^{\alpha_B}\, \bar{n}_{\rm e}^{\alpha_n}\, P_{\rm l,th}^{\alpha_P}\, R_{\rm geo}^{\alpha_R}\, \kappa_{\rm a}^{\alpha_\kappa}\, \epsilon^{\alpha_\epsilon}\, M_{\rm eff}^{\alpha_M}. \tag{2}$$

In this paper, we will follow the traditional, wide-spread practice of using a simple power-law expression for the confinement scaling, starting with expression (2). We will then compare with a similar scaling including the triangularity as an additional predictor variable. The latter occurs through a factor $(1+\delta)^{\alpha_{1+\delta}}$, in order to enable regression on data sets that include data from circular plasmas ($\delta = 0$).

Employing a power law form for the scaling is convenient from a computational perspective, because the expression is rendered linear on a logarithmic scale, while the distribution around the fitted model becomes more homoscedastic (homogeneous in variance). Another important quality that makes the power law a good model to start the analysis with is related to the robustness of the scaling. Indeed, in addition to goodness-of-fit, model complexity is an important criterion in the regression

methodology. Avoiding an overly complex model can help guarding against overfitting and may ensure robustness of the scaling against (small) perturbations of the data. This is particularly important with a view to extrapolation of the scaling. Nevertheless, it has been pointed out that a power law may not be sufficiently flexible to capture some important dependencies of the confinement time on the plasma parameters over the entire range of plasma conditions to which the scaling is fitted [24, 25]. This may also have implications for extrapolation of the scaling. Several authors have investigated more general, but related functional expressions, such as log-nonlinear scalings (e.g. [7]), or even completely different functional forms [26]. Non-power law scaling will not be considered in this paper, however.

In order to convert the engineering scaling to an expression in dimensionless quantities, it is useful to take the perspective of Connor and Taylor [27] (see also [28]), who investigated the scale-invariance of the equations describing the transport. In particular, the quasi-neutral high-β Fokker-Planck model assumes that plasma quasi-neutrality holds and that the anomalous transport can be described by the Fokker-Planck equation for the electron and ion distribution functions, with two-particle Coulomb collisions governed by the Landau collision operator and with a self-consistent magnetic field satisfying Ampère's equation. If, in addition, it is assumed that the boundary conditions do not appreciably affect the global confinement scaling, then the total conductive heat flux has to be invariant under the set of transformations of the physical variables that leave the equations invariant, and also the confinement time has to transform accordingly. This requirement leads to the following restriction on the exponents in the engineering scaling (2):

$$\alpha_{\mathrm{K}} \equiv 4\alpha_R - 8\alpha_n - \alpha_I - 3\alpha_P - 5\alpha_B - 5 = 0. \tag{3}$$

This is often referred to as the *high-β Kadomtsev constraint* and, if the estimated exponents obey this relation, the scaling is said to be 'dimensionally correct' [23, 29]. In turn, this means that the confinement scaling can be formulated entirely in terms of dimensionless quantities. Although the functional form does not need to be a power law, this does facilitate comparison with the engineering scaling. One common expression for power-law confinement scaling in terms of dimensionless variables is the following:

$$\Omega_{\mathrm{i}}\tau_{\mathrm{E,th}} = \alpha_{0,\mathrm{D}}\, \rho_*^{\alpha_\rho}\, \beta_{\mathrm{t}}^{\alpha_\beta}\, \nu_*^{\alpha_\nu}\, q_{\mathrm{cyl}}^{\alpha_q}\, R_{\mathrm{geo}}^{\alpha_{R,\mathrm{D}}}\, \kappa_{\mathrm{a}}^{\alpha_{\kappa,\mathrm{D}}}\, \epsilon^{\alpha_{\epsilon,\mathrm{D}}}\, M_{\mathrm{eff}}^{\alpha_{M,\mathrm{D}}}. \tag{4}$$

For the exponents of those variables that also occur explicitly in the engineering scaling, a subscript 'D' has been added to avoid confusion. If the dependence on triangularity is also studied, it suffices to add a factor $(1+\delta)^{\alpha_{1+\delta}}$ on the right-hand-side of (4). As the definitions of the dimensionless variables employed in this paper do not include triangularity, the exponent $\alpha_{1+\delta}$ is, at least in principle, identical in the engineering and dimensionless scaling. Furthermore, in (4) we have still allowed for dependence on the physical size R_{geo} of the system, but when the Kadomtsev constraint is fulfilled its exponent $\alpha_{R,\mathrm{D}}$ vanishes. This can be seen by substituting the definitions of the dimensionless variables in terms of the engineering variables into (4). In section 4, regression analysis in dimensionless space will be used to verify to what

extent $\alpha_{R,\mathrm{D}}$ can indeed be neglected in the resulting scalings. In addition, the volume-averaged temperature T can be eliminated from the resulting expression by noting that, under steady-state conditions, $\tau_{\mathrm{E,th}} = W_{\mathrm{th}}/P_{\mathrm{l,th}}$, so that $T \propto \tau_{\mathrm{E,th}} P_{\mathrm{l,th}}/(nV)$, where $V = 2\pi^2 R a^2 \kappa_{\mathrm{a}} \propto R_{\mathrm{geo}}^3 \epsilon^2 \kappa_{\mathrm{a}}$ is the plasma volume inside the LCFS. Equating the resulting exponents to those in (2) allows transforming the exponents in the dimensionless scaling to those in the engineering scaling. This leads to the transformation formulas given in Appendix A.

The scaling (4) is also sometimes written in terms of the Bohm time $\tau_{\mathrm{B}} = a^2/\chi_{\mathrm{B}}$, with $\chi_{\mathrm{B}} \propto T/B_{\mathrm{t}}$ the Bohm diffusivity. Since $\tau_{\mathrm{B}} \propto \Omega_{\mathrm{i}}^{-1} \rho_*^{-2}$, the dimensionless scaling becomes

$$\tau_{\mathrm{E,th}} \propto \tau_{\mathrm{B}} \rho_*^{2+\alpha_\rho}\, \beta_{\mathrm{t}}^{\alpha_\beta}\, \nu_*^{\alpha_\nu}\, q_{\mathrm{cyl}}^{\alpha_q}\, R_{\mathrm{geo}}^{\alpha_{R,\mathrm{D}}}\, \kappa_{\mathrm{a}}^{\alpha_{\kappa,\mathrm{D}}}\, \epsilon^{\alpha_{\epsilon,\mathrm{D}}}\, M_{\mathrm{eff}}^{\alpha_{M,\mathrm{D}}},$$

which is the form used in [3]. The case $\alpha_\rho = -2$ is termed 'Bohm scaling', meaning that the turbulence scale length involves the macroscopic plasma dimensions rather than the gyroradius. When $\alpha_\rho = -3$, the confinement scaling is said to be 'gyro-reduced Bohm', or 'gyroBohm' for short, as the Bohm-normalized diffusivity is proportional to the gyroradius: $\chi_{\mathrm{E,th}}/\chi_{\mathrm{B}} = \tau_{\mathrm{B}}/\tau_{\mathrm{E,th}} \propto \rho_*$. This is the situation that is expected from many turbulence theories. Since in larger-scale future devices ρ_* will be smaller than in present machines ($\rho_* \propto R_{\mathrm{geo}}^{-1} B^{-1}$), gyroBohm scaling leads to more favorable predictions for the confinement time. In fact, in scaling from present to future toroidal devices, the only variable that changes significantly is ρ_*.

3.2. Methodology of the regression analysis

The methodology followed for the regression analysis involves a software workflow that is briefly outlined in Appendix B. The regression model and analysis methods are discussed next.

3.2.1. Regression model The statistical model used for regression analysis comprises, as usual, a deterministic component, i.e. the actual scaling expression, and a stochastic component. The deterministic component has been discussed already above, viz. the power law scaling. Clearly, the deterministic component of any practical regression model can only describe part of the data variability. Any remaining variability is to be described by the stochastic component of the model, often a Gaussian distribution. The present 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. 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.

Another possible reason for concern is the potential occurrence of physical mechanisms that may significantly affect the global confinement, but that are not represented in the scaling model, such as plasma rotation and the electron-to-ion heating ratio. The effect of the associated 'hidden variables' can contribute to additional scatter (uncertainty) in the data. Particularly dangerous is the situation where such a missing variable influences one or more predictor variables, in addition to affecting the response variable $\tau_{\mathrm{E,th}}$. It is then called a 'confounding variable' and it may cause an apparent dependence (a so-called spurious relationship) of confinement on said predictor variables. Similar effects may play in analyzing a pooled data set originating from several groups, e.g. tokamaks in the database, possibly leading to different dependencies in the individual versus the pooled data (Simpson's paradox) [30]. Finally, constraining the data (e.g. the ITER-like data set) or specific data sampling schemes (e.g. depending on operational constraints) may also obfuscate the various confinement dependencies. Ideally, the model would include all predictor variables that have a known causal effect on the confinement through some physical mechanism, and only those, but in practice not all causal relations may be fully clear.

The variability of the data that is not explained by the deterministic regression function can be learned from the data using statistical techniques. On the other hand, we aim to exploit information that is available about the uncertainty on the measured data. Indeed, estimates of the uncertainty on the experimentally measured database variables are provided in the database in terms of percentage errors (i.e. a constant absolute error on a logarithmic scale). This includes statistical, and possibly systematic uncertainty arising from the measurement, but not some of the additional sources of uncertainty just mentioned. For the purposes of the present paper, however, all measurements of database variables are assumed to arise from an underlying normal (Gaussian) distribution, with standard deviation given by the absolute error derived from the corresponding measurement and its percentage error. For the effective mass M_{eff} it was deemed better to assume a fixed absolute error for the data used in the scaling, with a typical value of 0.20.

In estimating the scaling laws, 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 with respect to that on the response variable $\tau_{\mathrm{E,th}}$. This approximation is not valid for the power loss $P_{\mathrm{l,th}}$, for which the percentage error, averaged over the STD5 subset, is 12%, compared to 19% for $\tau_{\mathrm{E,th}}$. Already recognized in [24, 31, 32], 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 involving I_{p}, B_{t}, $P_{\mathrm{l,th}}$ and R_{geo} 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 toward ITER. This may be relevant when exploring new operational scenarios deviating from the correlation structure in the database.

3.2.2. Analysis methods In the past, most results of global confinement scaling, including IPB98(y, 2), were obtained using least squares regression on the log-transformed power law [33]. We will follow the usual practice, where, on a logarithmic scale, the data are assumed to follow an approximately normal (or at least symmetric) distribution around the scaling law. In addition, to avoid devices which contribute large numbers of data points (like JET and ASDEX Upgrade) to dominate the scaling, the measurements can be weighted. This method, called *weighted least squares regression* (WLS), was applied with weights w_{ik} assigned to point i from device k, given by $w_{ik}^{-1} = 2 + \sqrt{n_{\mathrm{obs},k}}/4$, where $n_{\mathrm{obs},k}$ is the number of points contributed by device k (alternative weighting schemes are being evaluated) [34, 32]. In this paper, the results from WLS will be taken as the benchmark, to which results from other techniques will be compared.

While WLS has the advantage of simplicity, it is worth exploring other techniques, as some assumptions underlying WLS might not be fulfilled [35, 36]. In addition, concerns have been raised regarding the common practice of logarithmic transformation of the data prior to regression analysis [31, 37, 36]. While an exact power law becomes linear on a logarithmic scale, in reality there is uncertainty in the data and the model. Therefore, in addition to WLS regression, we have implemented a robustified Bayesian method (referred to here as RBAYES), as well as a recently developed method known as *geodesic least squares regression* (GLS) [36]**; see Appendix C. In contrast to WLS, these methods properly take into account cases where the uncertainty on the predictor variables might not be negligible with respect to that on the response variable. They also exhibit a certain degree of robustness with respect to deviations from the model assumptions, e.g. due to outliers, between-device variability or the logarithmic transformation.††

Apart from parameter estimates, predictions $\hat{\tau}_{\mathrm{E,th,ITER}}$ for the thermal confinement time in ITER were derived from the estimated scaling expressions. It should be noted that, in the Bayesian analysis, predictions were obtained from the posterior predictive distribution (posterior predicted mean), which can be quite different from estimates calculated by direct substitution of the new values of the predictor variables in the estimated scaling expression. For consistency with the predictions by IPB98(y, 2), in this paper the following plasma parameters are used for confinement prediction in the

**Not to be confused with *generalized least squares*, a method that is only distantly related to geodesic least squares [38].

††In this paper, the designation 'robust' is used in a general sense, i.e. weak sensitivity to departures from the model assumptions that are relatively small, but otherwise unspecified.

baseline inductive $Q = 10$ scenario in ITER: $I_\mathrm{p} = 15$ MA, $B_\mathrm{t} = 5.3$ T, $\bar{n}_\mathrm{e} = 10.3 \times 10^{19}$ m^{-3}, $P_\mathrm{l,th} = 87$ MW (based on 40 MW auxiliary heating power and 400 MW fusion power), $R_\mathrm{geo} = 6.2$ m, $\delta = 0.48$, $\kappa_\mathrm{a} = 1.7$, $\epsilon = 0.32$ and $M_\mathrm{eff} = 2.5$ [39, 40]. Using the definitions (1*a*)–(1*d*), with $n = 0.88 \times 10.3 \times 10^{19}$ m^{-3} $= 9.06 \times 10^{19}$ m^{-3} and $T = 8.6$ keV, this yields the following values for the remaining dimensionless quantities in ITER: $\rho_* = 0.0020$, $\beta_\mathrm{t} = 2.24$, $\nu_* = 0.014$ and $q_\mathrm{cyl} = 1.94$. More recent studies relying on 50 MW auxiliary power and 500 MW fusion power suggest a conservative lower limit on $P_\mathrm{l,th}$ of about 100 MW [41, 42]. Predictions based on this higher level of power loss are provided for the main scaling laws derived in this work, as summarized in section 6. Whatever the case, it should be mentioned that extrapolations from global confinement scaling expressions have to be treated with caution due to the various uncertainties involved, including model uncertainties and physics-related uncertainties, e.g. owing to different neutral penetration in ITER, ELM control schemes, etc.

The analysis also intends to report an error bar on every parameter estimate and ITER prediction. Each error bar can be interpreted as (an estimate of) the standard deviation on the parameter or prediction, under the assumption that the underlying distribution is symmetric on the original scale.‡‡ That, for practical purposes, this is not incompatible with the above-mentioned assumption of normality on a logarithmic scale, is verified below. Indeed, only the estimated distributions of the intercepts α_0 and $\alpha_{0,\mathrm{D}}$ will turn out to be considerably skewed (due to the exponential transformation from the logarithmic domain), leading to markedly asymmetric error bars. In this paper, the error bars are consistently reported with two significant digits, and the parameter estimates and predictions with matching precision. Furthermore, in its current form, GLS is an optimization technique returning point estimates for the model parameters. Therefore, confidence intervals for GLS were estimated using a bootstrapping (resampling) technique with 100 samples.

Nevertheless, the value of the error bars based on the classical approaches is relatively limited, because they are derived under the assumption that the simple power law model, with the specific predictor variables used, is exact. In the case of the confinement scaling, it is known that these assumptions are questionable, even at an approximate level [3]. Therefore, alternative definitions of error bars have been proposed, which provide a measure of uncertainty on the estimates that is more useful for various practical purposes [3, 24]. Such alternative definitions typically lead to considerably larger error estimates. In this paper, an approach is followed using the results of the Bayesian method. As shown in more detail in section 4.2.4, this is done by rescaling the posterior covariance.

The quality of a regression fit can be gauged using several means. In classical statistics, the most widespread figures of merit are the root-mean-square error (RMSE)

‡‡Technically, according to the classical, frequentist interpretation of probability, an error bar on a parameter or prediction, as quoted in this paper, is an estimate of the standard deviation of the sampling distribution of the corresponding statistic. According to the Bayesian interpretation of probability, it is an estimate of the standard deviation of the corresponding posterior distribution.

Table 7. Estimates of the parameters (intercept α_0 and exponents) and prediction toward ITER[a] obtained by the IPB98(y, 2) scaling using the DB2.8 database. [3]

α_0	α_I	α_B	α_n	α_P	α_R	α_κ	α_ϵ	α_M	$\hat{\tau}_{98,\mathrm{ITER}}$ (s)
0.0562	0.93	0.15	0.41	−0.69	1.97	0.78	0.58	0.19	3.62
α_0	α_ρ	$\alpha_{\beta_\mathrm{t}}$	α_ν	α_q	$\alpha_{R,\mathrm{D}}$	$\alpha_{\kappa,\mathrm{D}}$	$\alpha_{\epsilon,\mathrm{D}}$	$\alpha_{M,\mathrm{D}}$	
4.24×10^{-7}	−2.69	−0.90	−0.0081	−2.99	−0.0081	3.29	0.71	0.96	

[a]Note that the prediction mentioned in [3] is for an earlier, larger ITER design.

and the coefficient of determination R^2. The RMSE can be interpreted as the standard deviation of the variability in the data that is not explained by the model (the regression function). The RMSE was calculated without weights, on the logarithmic scale, therefore it can directly be compared with the response variable, i.e. $\ln \tau_{\mathrm{E,th}}$ (engineering) or $\ln(\Omega_\mathrm{i}\tau_{\mathrm{E,th}})$ (dimensionless). R^2, also calculated without weights, compares the regression fit to a baseline model that assumes no trend (logarithmic scale). In keeping with the analysis of the confinement scaling throughout much of the literature, we mention both the RMSE and R^2 for all scalings estimated in this work. In addition, the median absolute percentage error (MdAPE) is shown, which is based on the ratio of the residuals over the observations of the response variable (logarithmic scale). The MdAPE and RMSE can be used to gauge the overall dispersion of the data around the fitted hyperplane, either relatively speaking (MdAPE) or in absolute terms (RMSE). The same performance measures are mentioned with the results from the RBAYES and GLS methods, although it should be kept in mind that these methods rely on metrics other than the squared residual for estimating the model parameters.

4. Confinement scaling results

In the remainder of the paper, results are reported of regression analysis for estimating the confinement scaling using the updated confinement database, by means of WLS, robust Bayesian regression (RBAYES) and GLS. For reference, the IPB98(y, 2) parameter estimates and ITER prediction $\hat{\tau}_{98,\mathrm{ITER}}$ are mentioned in table 7 [3]. The high-β Kadomtsev constraint was enforced to obtain IPB98(y, 2), as also the unconstrained scaling satisfied the constraint within the uncertainties [3]. Transforming the scaling to dimensionless form led to the estimates also mentioned in table 7.

Figure 2 shows a plot of the confinement enhancement factor $\mathrm{H}_{98}(\mathrm{y}, 2) = \tau_{\mathrm{E,th}}/\hat{\tau}_{98}$ versus Greenwald fraction $f_{\mathrm{GW}} = \bar{n}_\mathrm{e}/n_{\mathrm{GW}}$ for the DB5.2.3-STD5 ELMy H-mode data set (with $\hat{\tau}_{98}$ the predictions by IPB98(y, 2) and n_{GW} the Greenwald density), differentiating between the various devices. It can be seen that, depending on plasma conditions, IPB98(y, 2) tends to overpredict confinement when approaching the Greenwald limit, particularly in the purely metallic devices [17, 18]. The decreasing trend has been described by a log-quadratic term in [7].

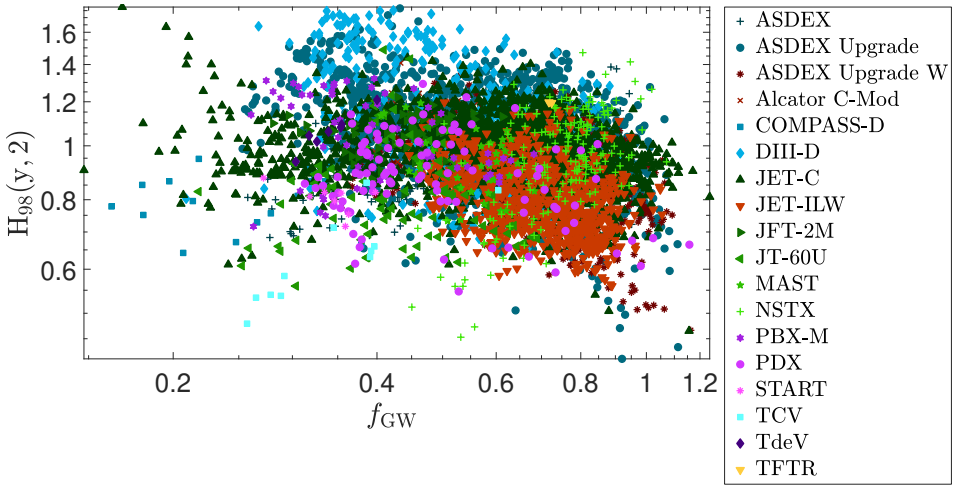

Figure 2. Confinement enhancement factor $H_{98}(y,2)$ versus Greenwald fraction f_{GW} for the DB5.2.3-STD5 ELMy H-mode data.

4.1. Single-device analysis

The analysis 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 data sets, and may lead to the resolution of some of the discrepancies with single-machine scans mentioned above. Table 8 presents estimates and standard deviations for a subset of machines represented in DB5.2.3-STD5 (all H-modes). In the single-device scalings most results of RBAYES and GLS are quite similar to those of OLS, therefore at this point we only present the latter (equivalent to WLS for single-machine regressions). As mentioned before, a distinction is made between the low-Z and high-Z data from ASDEX Upgrade and JET. Each scaling expression only involves those predictor variables that vary sufficiently within the corresponding data set.

The following observations can be made on the basis of these results, with a special focus on the ITER-like devices (Alcator C-Mod, ASDEX Upgrade, 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 and AUG-W ($\alpha_I \sim 1.5$–1.6) [15]. 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 and AUG-W ($\alpha_B \sim -0.3$). The I_p and B_t scaling in ASDEX Upgrade is being investigated [15]. It can be noted though, that the rule-of-thumb $\alpha_I + \alpha_B \gtrsim 1$ is quite well satisfied for all devices.
- 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 [43].
- The power degradation for the carbon-based ITER-like devices is in the range

Table 8. Estimates of the exponents in power law scalings of the global H-mode energy confinement time, expressed in engineering variables, using data from individual devices in the DB5.2.3-STD5 database (all H-modes). The estimates in this table were obtained by means of OLS on the logarithmically transformed data, with error bars corresponding to one standard deviation. $n_{\mathrm{obs},k}$ is the number of observations from device k in the data set.

Device	$n_{\mathrm{obs},k}$	α_I	α_B	α_n	α_P	$\alpha_{1+\delta}$	α_κ	α_M	MdAPE (%)	RMSE	R^2
ASDEX	575	0.692 ±0.057	0.130 ±0.079	0.678 ±0.043	−0.648 ±0.025	—	—	0.781 ±0.061	2.9	0.17	0.71
ASDEX Upgrade	1 385	1.464 ±0.037	−0.263 ±0.038	0.033 ±0.019	−0.658 ±0.013	—	—	—	5.0	0.17	0.73
ASDEX Upgrade W	767	1.558 ±0.036	−0.302 ±0.044	0.055 ±0.029	−0.534 ±0.013	—	—	—	2.8	0.11	0.89
Alcator C-Mod	82	1.149 ±0.080	—	0.101 ±0.091	−0.597 ±0.074	—	—	—	2.7	0.10	0.79
DIII-D	502	1.086 ±0.047	0.085 ±0.058	0.112 ±0.035	−0.677 ±0.018	0.696 ±0.082	—	0.449 ±0.057	5.5	0.18	0.82
JET-C	2 211	1.056 ±0.020	0.126 ±0.023	0.310 ±0.011	−0.743 ±0.010	—	1.112 ±0.089	0.242 ±0.024	10	0.15	0.90
JET-ILW	866	1.166 ±0.048	−0.204 ±0.034	0.084 ±0.025	−0.577 ±0.014	—	—	0.364 ±0.023	4.3	0.11	0.85
JFT-2M	348	0.986 ±0.043	—	0.383 ±0.038	−0.864 ±0.020	—	—	0.114 ±0.032	2.0	0.10	0.93
JT-60U	100	0.78 ±0.15	0.47 ±0.19	−0.177 ±0.085	−0.354 ±0.065	—	—	—	5.4	0.14	0.84
MAST	43	1.20 ±0.33	—	0.14 ±0.16	−0.82 ±0.15	—	—	—	2.5	0.11	0.60
NSTX	230	0.336 ±0.072	1.177 ±0.085	0.557 ±0.073	−0.850 ±0.041	—	0.86 ±0.16	—	2.4	0.13	0.75
PBX-M	214	0.61 ±0.16	—	−0.073 ±0.044	−0.557 ±0.036	—	—	—	2.6	0.12	0.72
PDX	119	0.62 ±0.16	0.63 ±0.16	0.618 ±0.082	−1.144 ±0.078	—	—	—	3.7	0.18	0.68

$\alpha_P \sim -0.74$ (JET-C), to about -0.68 (AUG, DIII-D), but somewhat weaker in the metallic devices: about -0.60 (Alcator C-Mod and JET-ILW) 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 (0.71) between $\ln \bar{n}_\mathrm{e}$ and $\ln P_{\mathrm{l,th}}$.

- The dependence on effective atomic mass varies among the devices, from weak in JET-C and JFT-2M ($\alpha_M \sim 0.1$–0.2), to somewhat stronger in JET-ILW and DIII-D ($\alpha_M \sim 0.4$), to quite strong in ASDEX ($\alpha_M \sim 0.8$). It is also worth noting that this is one dependence where the three regression methods consistently tend to give slightly different answers. RBAYES and GLS usually yield a stronger scaling

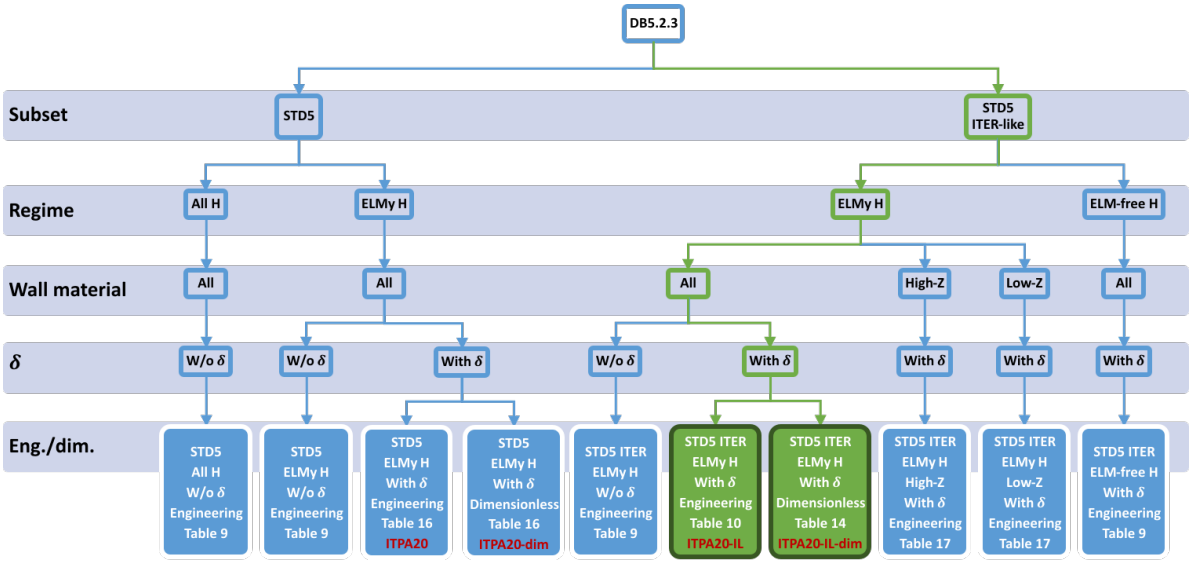

Figure 3. Overview of the multi-machine regressions carried out in this paper. Starting from DB5.2.3, the panels in the background indicate several consecutive choices made in selecting subsets of the data. The actual choices made are shown in the open boxes, eventually leading up to the data subsets used for the scalings, shown in the shaded boxes at the bottom. The tables listing the corresponding regression results are also given in the lower boxes. The green path corresponds to the data selections made to obtain the most ITER-relevant scalings shown in this paper.

with effective mass than WLS, although the confidence intervals are quite large. Nevertheless, similar observations are made for the multi-machine scalings.

4.2. Multi-machine scalings

We next turn to regression analysis on several subsets of the full database, each with data from multiple devices. The various choices made in establishing these subsets are illustrated in figure 3. Starting from the full database DB5.2.3, a first distinction concerns the selection of either the complete standard set STD5 within DB5.2.3, or the ITER-like subset (STD5 ITER-like). In the full STD5 data set, a comparison is made between the scaling with all H-mode plasmas ('All H') and the ELMy H-mode plasmas only ('ELMy H'). Next, we concentrate on the ELMy H-mode points in the STD5 ITER-like set, studying the influence of triangularity δ and wall material (high-Z versus low-Z, both with δ). A final scaling is carried out using only the ELM-free plasmas in the STD5 ITER-like set (with δ). The currently most ITER-relevant scaling is obtained with the STD5 ITER-like ELMy H-mode points, including triangularity. This scaling will be transformed to its dimensionless form and compared with the corresponding scaling carried out directly in dimensionless space.

4.2.1. All H-modes and ELMy H-modes in STD5 The parameter estimates with error bars for the scaling of the global energy confinement in all H-mode plasmas in DB5.2.3-

STD5 are given in table 9. Although the alternative regression techniques (RBAYES and GLS) do show some differences with respect to the parameter estimates obtained with WLS, for clarity and in the interest of comparing with IPB98(y, 2), in this table only the WLS results are shown. A more detailed analysis comprising the results from all methods will be given for one particular scaling in section 4.2.4.

Concentrating on the ELMy H-modes in DB5.2.3-STD5, relatively similar parameter estimates are obtained as in the case of all H-modes, with a slightly reduced power degradation, as shown in table 9. A number of notable differences can be observed in comparison with the estimates by IPB98(y, 2) (table 7).

- The scaling with magnetic field is somewhat stronger in the new scaling, whereas the dependence on plasma current is very similar. However, even though the differences with IPB98(y, 2) lie outside of the error bars on the respective parameter estimates, the significant correlation between several of the predictor variables may easily give rise to the observed changes, as pointed out in section 4.2.4. Hence, these numerical differences should not be interpreted as to indicate significantly different dependencies without additional investigation. In fact, in IPB98(y, 2) the sum of the estimates for α_I and α_B amounts to ~ 1.08. This is about the same as the sum of ~ 1.19 obtained from the present scaling.
- The density dependence is slightly weaker for the ELMy H-modes in DB5.2.3-STD5 and so is the dependence on machine size.
- The dependence on plasma elongation κ_a is somewhat stronger than in IPB98(y, 2), but the ϵ scaling is a bit weaker.

Other dependencies are similar to the those in IPB98(y, 2). As with most fits in this paper, the value of R^2 is high, indicating a clear overall trend. The most noticeable difference with IPB98(y, 2) concerns the predicted global confinement time for the ITER scenario. Indeed, most predictions derived from the DB5.2.3-STD5 data set lie around 3 s, or slightly below. This is a reduction of at least 17% with respect to the predictions derived from DB2.8 (3.62 s).

4.2.2. ELMy H-modes in the STD5 ITER-like subset Restricting the DB5.2.3-STD5 data to the ELMy H-modes based on the ITER-like constraints stated in section 2.1, several dependencies become quite different to the case of ELMy H-modes in the unconstrained STD5 data set.

- The scaling with plasma current has become stronger than in the full STD5 data set, exceeding the estimate by IPB98(y, 2). At the same time, the magnetic field dependence becomes almost vanishingly small or slightly negative. In fact, again it is seen that $\alpha_I + \alpha_B \sim 1.13$, so it should not be concluded that these dependencies are significantly different from before.
- The density dependence has become even weaker compared to the case of the unrestricted STD5 data.

- The degradation of confinement with power loss is weaker than in IPB98(y, 2). This may be related to a similar observation in the metallic devices, as noted in section 4.1. However, also this difference should be interpreted with caution, in view of the significant correlation within the database between power loss, plasma current and major radius.
- The most notable difference with IPB98(y, 2) is the considerably reduced dependence on major radius $R_{\rm geo}$, from almost quadratic to slightly stronger than linear. Due to the correlation with other variables, like $I_{\rm p}$ and $P_{\rm l,th}$, some caution in interpreting this result is again warranted, but nevertheless the weaker size scaling is observed consistently for various subsets of the database.
- The dependence on the shape parameters $\kappa_{\rm a}$ and ϵ is also weaker than in IPB98(y, 2), with inverse aspect ratio playing no role of significance any longer.

At about 2.7 s, the extrapolation toward ITER has become slightly lower than for the previously mentioned scalings in DB5.2.3-STD5. The quality of the fit is similar to that for all H-modes and the unconstrained STD5 ELMy H-modes.

4.2.3. ELM-free H-modes in the STD5 ITER-like subset It is of interest to apply a similar study to the ELM-free H-modes in the database. Here, it is important to stress that these are not steady-state regimes, as they are typically characterized by a strong density increase, often with high radiated power, ending with a very large ELM or a disruption [44]. As such, these plasmas are not particularly relevant for future devices, in possible contrast to more recent approaches to ELM-free operational regimes. The results, given in table 9 for the ITER-like data in DB5.2.3-STD5, are mentioned here just for reference. Nevertheless, the considerably stronger dependence on machine size, compared to the ELMy H scaling in the same ITER-like data set, is remarkable. It is also interesting to note that, whereas the exponents are relatively similar to those obtained for all H-modes in the DB5.2.3-STD5 data set, still the ITER prediction is quite a bit higher, almost at the level of IPB98(y, 2). Indeed, small changes in the exponents of the power law, or in the factor α_0, can result in significantly different extrapolations. This is discussed in some more detail in the next subsection. The results from the RBAYES method (not shown) confirm the WLS results for the ELM-free data.

4.2.4. ELMy H-modes in the STD5 ITER-like subset with δ In clarifying the influence of the plasma shape on confinement, it is interesting to add the triangularity as a predictor variable in the scaling, in the form $1+\delta$. The corresponding regression results are given in table 10. Considering first the results from WLS, the triangularity indeed appears to have some influence on confinement ($\alpha_{1+\delta} \sim 0.56$), comparable in strength to that of the plasma elongation. The exponents for the other predictor variables are similar to those obtained on the same data set, without δ dependence (table 9). Nevertheless, due to the non-negligible influence of δ in the ELMy H-mode ITER-like data set, the scaling with δ is currently proposed as the most ITER-relevant among the

Table 9. Estimates of the parameters (intercept α_0 and exponents) and prediction toward ITER in power law scalings of the global H-mode energy confinement time, expressed in engineering variables, using subsets of the DB5.2.3-STD5 database. All estimates in this table were obtained by means of WLS on the logarithmically transformed data, with error bars corresponding to one standard deviation.

	α_0	α_I	α_B	α_n	α_P	α_R	α_κ	α_ϵ	α_M	$\hat{\tau}_{\mathrm{E,th,ITER}}$ (s)	MdAPE (%)	RMSE	R^2
All H	0.0601 ±0.0011	0.959 ±0.014	0.247 ±0.012	0.2587 ±0.0082	−0.7068 ±0.0054	1.781 ±0.021	0.963 ±0.021	0.453 ±0.027	0.196 ±0.014	2.907 ±0.031	6.4	0.19	0.96
ELMy H	0.0576 ±0.0013	0.984 ±0.015	0.203 ±0.013	0.2530 ±0.0091	−0.6658 ±0.0064	1.698 ±0.023	0.930 ±0.024	0.387 ±0.032	0.175 ±0.016	2.932 ±0.034	6.2	0.18	0.96
ELMy H ITER-like	0.0798 ±0.0024	1.311 ±0.017	−0.178 ±0.017	0.1570 ±0.0090	−0.6339 ±0.0065	1.155 ±0.027	0.670 ±0.041	0.010 ±0.034	0.296 ±0.016	2.675 ±0.036	6.3	0.17	0.95
ELM-free H ITER-like	0.0638 ±0.0037	0.899 ±0.036	0.319 ±0.034	0.282 ±0.024	−0.704 ±0.016	1.910 ±0.061	1.052 ±0.072	0.602 ±0.063	0.164 ±0.037	3.45 ±0.14	6.6	0.16	0.99

Table 10. Estimates of the parameters (intercept α_0 and exponents) and prediction toward ITER in power law scalings of the global H-mode energy confinement time, expressed in engineering variables, using the ELMy H-mode plasmas in the DB5.2.3-STD5 database, restricted to the ITER-like conditions. The estimates were obtained from the logarithmically transformed data using weighted least squares (WLS), a robust Bayesian method (RBAYES) and geodesic least squares (GLS). The error bars correspond to one standard deviation, but for the RBAYES method they are based on a Gaussian approximation with rescaled posterior covariance matrix ($\gamma_{\mathrm{cov}} = 10$; see main text).

	α_0	α_I	α_B	α_n	α_P	α_R	$\alpha_{1+\delta}$	α_κ	α_ϵ	α_M	$\hat{\tau}_{\mathrm{E,th,ITER}}$ (s)	MdAPE (%)	RMSE	R^2
WLS	0.0670 ±0.0021	1.291 ±0.017	−0.134 ±0.017	0.1473 ±0.0088	−0.6442 ±0.0063	1.194 ±0.027	0.560 ±0.032	0.673 ±0.039	−0.000 ±0.033	0.302 ±0.016	2.902 ±0.040	6.0	0.17	0.95
RBAYES	0.071 $^{+0.059}_{-0.032}$	1.40 ±0.16	−0.22 ±0.17	0.067 ±0.097	−0.684 ±0.061	1.00 ±0.27	0.63 ±0.36	0.59 ±0.63	−0.25 ±0.46	0.42 ±0.16	2.58 ±0.46	5.7	0.17	0.95
GLS	0.0231 ±0.0033	1.267 ±0.028	−0.094 ±0.027	0.094 ±0.015	−0.7315 ±0.0067	1.270 ±0.047	0.775 ±0.054	2.03 ±0.16	−0.315 ±0.092	0.633 ±0.051	2.707 ±0.057	6.2	0.20	0.93

various scalings presented in this paper. We therefore investigate this scaling in some more detail.

Table 10 also contains the analysis results by the two more robust methods RBAYES and GLS. The goodness-of-fit in terms of the MdAPE and RMSE, as well as the coefficient of determination R^2, is comparable for the three methods. Figure 4 shows histograms for the individual scaling exponents, obtained by (good) approximation of the joint posterior distribution from the RBAYES method by a multivariate Gaussian distribution (with a rescaled covariance matrix, as described below). The bivariate distributions between each pair of exponents are also depicted.

Table 10 reveals several differences between the results of the various methods, but these should be interpreted in light of the error estimates on the exponents, which we now discuss. First, it is important to note that the correlation between some

Table 11. Pairwise correlations between the exponents in the scaling with engineering variables estimated by WLS using the DB5.2.3-STD5 ELMy H-mode ITER-like data set, including triangularity.

	α_I	α_B	α_n	α_P	α_R	$\alpha_{1+\delta}$	α_κ	α_ϵ	α_M
α_I	**1.0**	−0.76	−0.42	0.14	−0.85	−0.086	−0.45	−0.83	0.10
α_B	−0.76	**1.0**	0.15	−0.33	0.69	0.16	0.56	0.64	−0.30
α_n	−0.42	0.15	**1.0**	−0.42	0.68	−0.054	0.14	0.41	−0.13
α_P	0.14	−0.33	−0.42	**1.0**	−0.53	−0.078	−0.26	−0.31	0.068
α_R	−0.85	0.69	0.68	−0.53	**1.0**	0.099	0.45	0.77	−0.13
$\alpha_{1+\delta}$	−0.086	0.16	−0.054	−0.078	0.099	**1.0**	0.0087	−0.011	0.0073
α_κ	−0.45	0.56	0.14	−0.26	0.45	0.0087	**1.0**	0.50	−0.35
α_ϵ	−0.83	0.64	0.41	−0.31	0.77	−0.011	0.50	**1.0**	−0.18
α_M	0.10	−0.30	−0.13	0.068	−0.13	0.0073	−0.35	−0.18	**1.0**

pairs of exponents is relatively strong. This is related to the dependencies between the predictor variables. Table 11 shows the pairwise correlations between the scaling exponents, obtained by WLS. Most notably, α_R clearly correlates with α_I, α_B, α_n and α_ϵ. Furthermore, it was already noted that α_I correlates with α_B, but it turns out that both also correlate quite strongly with α_ϵ. This is also demonstrated graphically in figure 4 (note that for a bivariate normal distribution the amount of correlation has no effect on the width of the marginal distributions).

In addition, it was already mentioned that an important assumption of the regression analysis, i.e. the power law model, might be questioned. The error estimates on the individual exponents rely on that assumption. A different model might fit the data better in some critical areas of the data space, but the global goodness-of-fit (e.g. measured by the MdAPE) would not necessarily improve significantly, or might even grow worse. Moreover, goodness-of-fit is not the sole criterion to compare two different models; also the model complexity should be considered. Although model selection and model comparison are outside the scope of the present paper, one may get a feeling of the model uncertainty by studying the sensitivity of the goodness-of-fit (MdAPE) to the values of the model parameters. For now, the aim is to arrive at a practically more useful measure of uncertainty on the individual parameter estimates.

To systematically investigate this issue, the posterior covariance matrix obtained by RBAYES was rescaled by a given factor $\gamma_{\text{cov}}^2 > 1$. Next, $N_\text{s} = 4\,000$ samples were drawn from a multivariate normal distribution with mean corresponding to the posterior mean and covariance given by this rescaled covariance matrix. A new MdAPE was then calculated with each of those 4 000 sets of sampled scaling exponents. It was found that the MdAPE is only moderately dependent on the factor γ_{cov}. This is partly caused by the correlations between some exponents in the scaling law. For instance, increasing α_I can be compensated by decreasing α_B without much compromising the

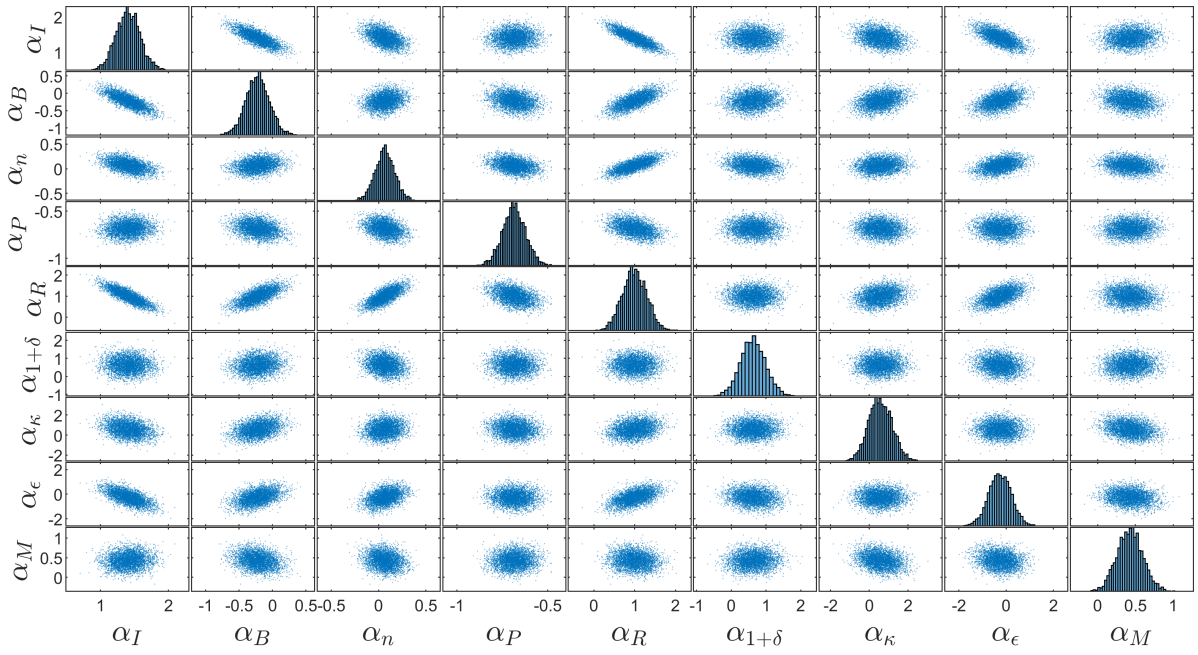

Figure 4. Illustration of the joint posterior distribution for the exponents in the scaling with engineering variables obtained using RBAYES in the DB5.2.3-STD5 ELMy H-mode ITER-like data set with triangularity. The posterior distribution was approximated by a multivariate normal distribution with rescaled covariance matrix ($\gamma_{\rm cov} = 10$; see main text). Panels on the diagonal show histograms depicting the marginal distributions of the individual scaling exponents, while off-diagonal panels illustrate marginal bivariate distributions, for all pairs of scaling exponents.

fit. Indeed, resampling the exponents independently would have a much stronger effect on the MdAPE. As an example, scaling the covariance up by a factor of $\gamma_{\rm cov}^2 = 100$ leads to a median increase of the MdAPE of less than 1 percentage point (from 5.7% to 6.4%) and a maximum increase of less than 3 percentage points in the observed sample. This corresponds to error bars (marginal posterior standard deviations) on the individual parameters that are about 10 times larger than those obtained by WLS and GLS.

Despite the arbitrariness of the precise value of the covariance scale factor $\gamma_{\rm cov}$, the previous study does provide a useful indication of the level of uncertainty on the parameter estimates in the scaling. For illustration, the value $\gamma_{\rm cov} = 10$ is adopted in this paper and the resulting error bars are shown in table 10 for RBAYES. These uncertainties will be used next to interpret the significance of differences in the parameter estimates between the various methods in table 10, or between the results obtained on different data sets in table 9. The following observations can be made (referring to the scaled error bars mentioned with the RBAYES estimates).

- The relative error on the individual exponents is smallest for $P_{\rm l,th}$ (9%), $I_{\rm p}$ (11%) and $R_{\rm geo}$ (27%). The other exponents have a considerably larger relative uncertainty.
- In comparison with WLS and GLS, RBAYES yields a higher exponent for $I_{\rm p}$, but

this is accompanied by the slightly stronger negative dependence on B_t and weaker dependence on R_geo.

- Considering the (absolute) error bar on α_B, the dependence of confinement on magnetic field is vanishingly small.
- Both RBAYES and GLS predict virtually no dependence of confinement on density.
- The power degradation estimated by RBAYES and GLS is stronger compared to WLS. Particularly the difference between GLS and WLS could be significant. In this respect, it is interesting to note that the value of α_P obtained by RBAYES and GLS approaches the one estimated by WLS, if the measurement uncertainty on $P_\mathrm{l,th}$ (on average 10% over the data set) is artificially decreased (see also [31, 32]). Moreover, with ordinary least squares (no weighting) the power degradation is somewhat stronger ($\alpha_P = -0.67$) than with WLS, approaching the level of RBAYES and GLS.
- In view of the error bar on $\alpha_{1+\delta}$, the dependence of confinement on triangularity is rather modest. From the same point of view, there seems to be no clear dependence on elongation and inverse aspect ratio. In fact, with GLS the estimate for the exponent α_κ did not appear reliable, but the dependence was retained to allow a fair comparison with the other methods. However, it should be noted that the lack of clear dependence on the shape parameters $1+\delta$, κ_a and ϵ may be at least partly related to their relatively restricted range in the ITER-like subset.
- Both RBAYES and GLS suggest a non-negligible dependence of confinement on effective mass that is somewhat stronger than that estimated by WLS. The difference appears to be partly related to the weighting scheme in WLS, because unweighted least squares yields a slightly stronger scaling ($\alpha_M = 0.37$).
- The prediction by RBAYES for the confinement time in ITER is about 2.6 s, still somewhat lower than the prediction by WLS, whereas the prediction by GLS is again higher: about 2.7 s. However, these differences are to be seen in light of the considerable extrapolation uncertainty. It is important to note that part of the prediction uncertainty is related to the large uncertainty on the factor α_0.
- The prediction by RBAYES, i.e. the posterior mean of 2.58 s, is about the same as the posterior median. The rescaled posterior standard deviation is 0.46 s. Despite the exponential transformation required to obtain the prediction on the original scale, the posterior predictive distribution is relatively symmetric around the posterior mean (as in all other regressions carried out in this work). Only in the very tails does the distribution, based on the rescaled covariance matrix, become somewhat asymmetric. For instance, the rescaled 95% credible interval with equal tail probabilities is [1.86 s, 3.63 s], or [−28%, +41%], which is quite asymmetric with respect to the posterior mean. However, these bounds have to be interpreted with care, due to the level of arbitrariness of the chosen value of the scale factor γ_cov.

Table 12. Factors γ_k, estimated by RBAYES, and ratios $\sigma_{\mathrm{obs},k}/\sigma_{\mathrm{mod},k}$, estimated by GLS, for each device k in scaling with the engineering variables on the DB5.2.3-STD5 ELMy H-mode ITER-like data set with triangularity. Values above 1.25 have been highlighted.

	AUG	AUG-W	Alcator C-Mod	COMPASS-D	DIII-D	JET-C	JET-ILW	JFT-2M	JT-60U	PBX-M
γ_k	0.96	0.79	0.51	3.8	0.85	1.0	0.89	0.69	1.1	0.85
$\sigma_{\mathrm{obs},k}/\sigma_{\mathrm{mod},k}$	1.2	1.1	1.2	2.0	1.2	1.2	1.2	1.1	1.2	2.0

We still note that the rescaled posterior distribution for the intercept α_0 is considerably skewed, due to the exponential transformation of the estimate on the logarithmic scale. For this case, the role of the standard deviation as an error bar is less suitable. Therefore, in table 10 the error bars on α_0 obtained by RBAYES were chosen to reflect the width of an equal-tailed credible interval around the median, covering about 68% of the total probability (i.e. the probability included within one standard deviation from the mean in a normal distribution).

The estimates of the correction factors γ_k (see Appendix C) for the standard deviation of the likelihood in the Bayesian method are listed in table 12, for each device represented in the scaling. The significance of a particular γ_k being larger (smaller) than 1 is that, on average, the measurements from device k have a wider (narrower) distribution around the fitted hyperplane than expected from the measurement errors. The table also gives, for each device k, the ratio of $\sigma_{\mathrm{obs},k}$ over $\sigma_{\mathrm{mod},k}$ (in fact its median over the data from device k), estimated by GLS (Appendix C). Conceptually, and to a certain extent also quantitatively, these ratios can be compared with the γ_k in the Bayesian method. Although there are some differences between RBAYES and GLS in table 12, both methods agree that the data from COMPASS-D are much more spread around the scaling plane than expected due to measurement error alone. In addition, according to GLS, the points from PBX-M can also deviate considerably from the scaling.

Figure 5 contains visualizations of the fitted scalings. In panel (a), the measured energy confinement time is plotted against the one predicted by the scaling obtained using WLS, RBAYES and GLS. The comparable goodness-of-fit among the three methods is confirmed. Panels (b) to (d) provide residual plots obtained by WLS. To this end, a confinement enhancement factor $\mathrm{H}_{20} \equiv \tau_{\mathrm{E,th}}/\hat{\tau}_{\mathrm{E,th}}$ was calculated using the expression obtained by WLS from table 10. In particular, H_{20} is plotted against Greenwald fraction f_{GW} in figure 5(b), distinguishing the points per device. This can be compared with figure 2 (keeping in mind that this is for a different data set), indicating that the decreasing trend with f_{GW} is alleviated in the new scaling, although not entirely absent. Specifically, the points from the metallic devices AUG-W and JET-ILW still seem to partly contribute to the downward trend, when approaching the Greenwald limit. Figure 5(c) shows a plot of H_{20} against normalized separatrix density $n_{\mathrm{e,sep}}/n_{\mathrm{GW}}$, for those entries in the data set where $n_{\mathrm{e,sep}}$ is available (only for some points from AUG,

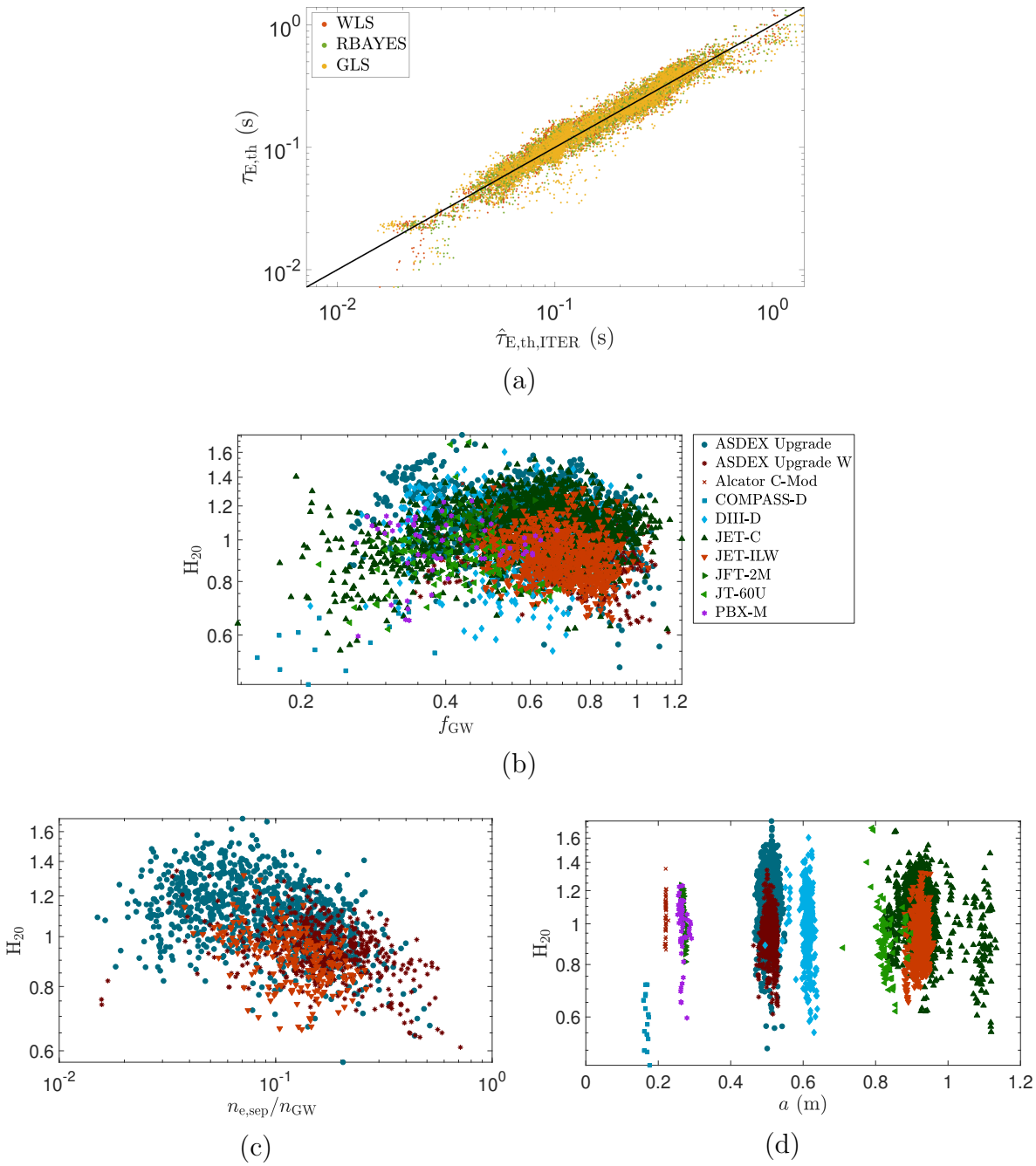

Figure 5. Results of regression analysis for power law scalings of the global H-mode energy confinement time, expressed in engineering variables, using WLS on the DB5.2.3-STD5 ELMy H-mode ITER-like data set with triangularity. (a) Measured versus predicted energy confinement time, obtained using the three methods WLS, RBAYES and GLS. The following three panels contain a plot of the confinement enhancement factor H_{20} obtained using WLS against (b) Greenwald fraction f_{GW}, (c) separatrix density $n_{e,sep}$ normalized by Greenwald density (only for AUG, AUG-W and JET-ILW) and (d) minor radius a.

AUG-W and JET-ILW). There is a clear decreasing trend with increasing separatrix density, indicating a residual role of $n_{\mathrm{e,sep}}$ that is not reflected by the present scaling. As an aside, the influence was investigated of those points from AUG, AUG-W and JET-ILW for which $n_{\mathrm{e,sep}}/n_{\mathrm{GW}} > 0.1$, by performing another regression, leaving out the corresponding database entries at elevated separatrix density from these machines. However, the influence on both the scaling exponents and ITER predictions turned out to be negligible. Finally, figure 5(d) shows the residual distributions for the various machines in the data set. In particular, it is seen that the points from COMPASS-D all have lower confinement than predicted by the scaling, agreeing with the results from RBAYES and GLS in table 12.

It can be instructive to visualize the individual trends of the confinement with the various engineering variables, derived from the fitted scaling law. To that end, figure 6 shows, for each predictor variable, the measured and predicted confinement time (WLS), normalized by the estimated dependence on the other predictor variables. In general, there is a considerable scatter of the data around the fitted trends, although less so for I_{p} and $P_{\mathrm{l,th}}$. In a qualitative way, this agrees with the uncertainties on the parameter estimates and prediction obtained using RBAYES in table 10.

The high-β Kadomtsev constraint is investigated next. Histograms of the quantity α_{K} defined in (3) are shown in figure 7 for the three methods. This yields the following estimates (averages) and standard deviations: for WLS one obtains $\alpha_{\mathrm{K}} = -0.089 \pm 0.066$, for RBAYES $\alpha_{\mathrm{K}} = 0.22 \pm 0.55$ (scaled error bar) and for GLS $\alpha_{\mathrm{K}} = 0.720 \pm 0.084$. WLS appears to well satisfy the constraint, although the scaled error bar suggests that all methods give acceptable results in this respect.

Transforming the engineering form of the scaling, using the parameter estimates from table 10, to dimensionless variables by means of (A.1), one obtains the results in table 13. Note that, in calculating the accompanying error bars, due account has to be taken of the correlation between the parameters. For WLS, these results are only broadly similar to the dimensionless scaling obtained from IPB98(y, 2) (table 7, disregarding the δ scaling), with a considerably weaker ρ_* scaling. Part of the difference is due to our inclusion of a residual dependence on machine size, hence not strictly enforcing the Kadomtsev constraint. Nevertheless, the dependence on R_{geo} is small, so the scaling is almost dimensionally correct. For WLS and GLS, the error bars on the exponents of the dimensionless scaling were obtained from the bootstrap samples. For the RBAYES method, the (scaled) posterior standard deviations are shown. Concentrating again on the latter for a practical error measure, it is seen that the uncertainty on the exponents is quite high. This is in particular due to the factor $(1+\alpha_P)^{-1}$ in (A.1). For instance, the dependence on β_{t} given by WLS, which is even slightly stronger than that in IPB98(y, 2), is only marginally significant, considering the corresponding error bar given by RBAYES. The validity of a transformation (A.1) in practical situations can therefore be questioned, because it is particularly sensitive to the model assumptions.

Finally, regression analysis can be carried out directly in the space of dimensionless variables. In theory, WLS is not suitable for handling this situation, as the uncertainties

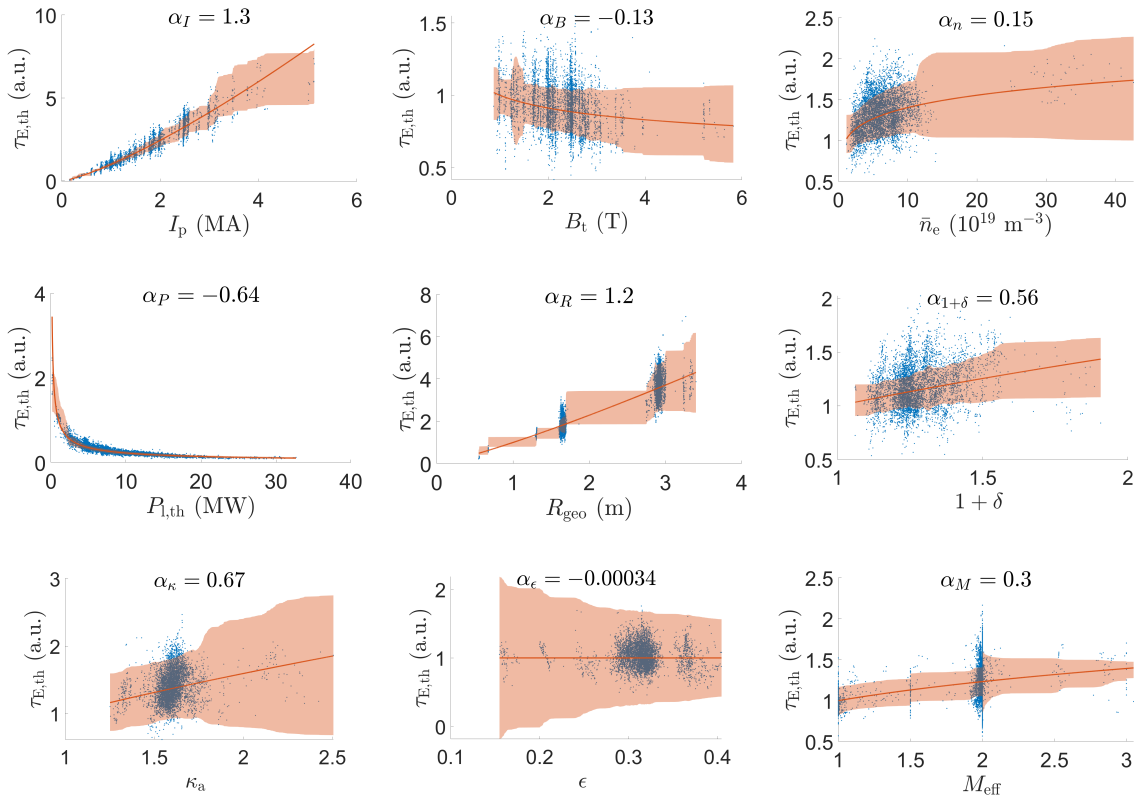

Figure 6. Plots of the normalized measured confinement time (blue points) against the individual engineering variables, within the DB5.2.3-STD5 ELMy H-mode ITER-like data set with triangularity. The predictions obtained with WLS are also shown (red lines), as well as a corresponding confidence band (one standard deviation) provided by the Bayesian analysis (rescaled covariance).

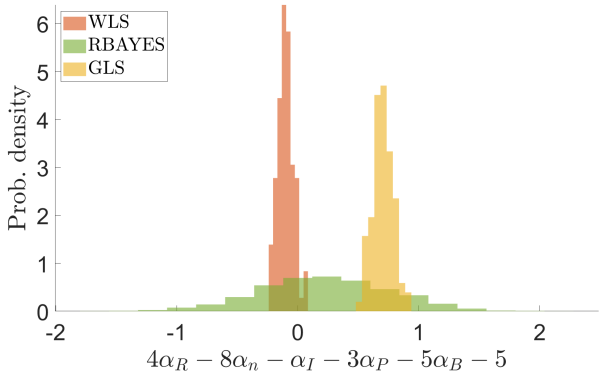

Figure 7. Sampled probability densities of the high-β Kadomtsev constraint α_K for the scaling with engineering variables obtained using WLS, RBAYES and GLS on the DB5.2.3-STD5 ELMy H-mode ITER-like data set with triangularity.

Table 13. Estimates of the exponents in the dimensionless scaling for the global H-mode energy confinement time, obtained by transformation from the exponents in the engineering scaling, using the DB5.2.3-STD5 ELMy H-mode ITER-like data set, including triangularity.

	α_ρ	α_{β_t}	α_ν	α_q	$\alpha_{R,\mathrm{D}}$	$\alpha_{1+\delta}$	$\alpha_{\kappa,\mathrm{D}}$	$\alpha_{\epsilon,\mathrm{D}}$	$\alpha_{M,\mathrm{D}}$
WLS	−1.771	−1.240	−0.157	−3.470	−0.063	1.573	3.71	1.63	0.733
	±0.044	±0.045	±0.011	±0.079	±0.039	±0.097	±0.12	±0.10	±0.041
RBAYES	−0.98	−1.97	−0.08	−4.5	0.21	2.1	4.4	2.7	0.87
	±0.68	±0.74	±0.14	±1.2	±0.48	±1.3	±2.3	±1.7	±0.50
GLS	−0.59	−2.392	0.020	−4.74	0.671	2.88	9.57	2.26	1.65
	±0.10	±0.096	±0.020	±0.16	±0.079	±0.22	±0.63	±0.29	±0.14

on several predictor variables cannot be neglected anymore. Nevertheless, in table 14 the results of WLS are listed, as well as those obtained with RBAYES. The regression with GLS did not converge, so it is not mentioned. These results have to be interpreted in light of the pairwise correlations between the parameter estimates, which are given in table 15 for WLS. In particular, there is considerable correlation between α_ρ and both $\alpha_{R,\mathrm{D}}$ and $\alpha_{\epsilon,\mathrm{D}}$. Specifically, the individual coefficients α_ρ and $\alpha_{R,\mathrm{D}}$ are difficult to disentangle in this scaling (probably this contributes to GLS not converging). Therefore, the regression analysis was also carried out without R_{geo} as a predictor variable. Sizeable correlation also exists between α_β, α_q and $\alpha_{\epsilon,\mathrm{D}}$ (see [5]), as well as between α_ν and $\alpha_{R,\mathrm{D}}$. A number of interesting conclusions can be derived from the scalings with and without R_{geo}. Reference is made to the error bars given by RBAYES, which, for that matter, are generally smaller than those resulting from the direct transformation (table 13).

- The goodness-of-fit of the scaling carried out in dimensionless space is not as large as that of the engineering scaling. Nevertheless, several of the parameter estimates are significantly different from zero (relative to the error bars given by RBAYES) and the value of R^2 is still good.
- Out of the two methods, RBAYES yields the smallest residual dependence on machine size, i.e. it is the most dimensionally correct. However, the uncertainty on the $\alpha_{R,\mathrm{D}}$ exponent is considerable. Leaving R_{geo} out of the model mainly affects the WLS estimates, which approach those obtained by RBAYES. In particular, α_ρ becomes -1.78 with WLS and $\alpha_{\epsilon,\mathrm{D}} = -0.224$.
- While RBAYES is the most dimensionally correct, it also predicts Bohm scaling of the transport ($\alpha_\rho \sim -2.0 \pm 0.3$).
- The dependence on plasma β is negligible (no β degradation), much weaker than in the case of IPB98(y, 2).
- There is a weak inverse scaling with ν_*, whereas no collisionality dependence was seen in IPB98(y, 2).
- There is an inversely linear dependence on safety factor, which is again much weaker than that obtained in IPB98(y, 2).

Table 14. Estimates of the parameters (intercept α_0 and exponents) and prediction toward ITER in power law scalings of the global H-mode energy confinement time, carried out in dimensionless space, using the DB5.2.3-STD5 ELMy H-mode ITER-like data set, including triangularity.

	α_0	α_ρ	α_{β_t}	α_ν	α_q	$\alpha_{R,\mathrm{D}}$	$\alpha_{1+\delta}$	$\alpha_{\kappa,\mathrm{D}}$	$\alpha_{\epsilon,\mathrm{D}}$	$\alpha_{M,\mathrm{D}}$	$\hat{\tau}_{\mathrm{E,th,ITER}}$ (s)	MdAPE (%)	RMSE	R^2
WLS	0.529×10^{-6} $\pm 0.083 \times 10^{-6}$	−2.220 ±0.027	0.157 ±0.019	−0.4595 ±0.0059	−1.161 ±0.037	−0.468 ±0.023	0.382 ±0.058	1.236 ±0.065	−0.828 ±0.054	0.551 ±0.025	2.309 ±0.055	16	0.30	0.88
RBAYES	2.1×10^{-6} $+19 \times 10^{-6}$ -1.9×10^{-6}	−1.97 ±0.33	0.12 ±0.22	−0.426 ±0.063	−1.03 ±0.43	−0.07 ±0.34	0.14 ±0.64	1.89 ±0.85	−0.26 ±0.64	0.62 ±0.37	2.53 ±0.88	16	0.31	0.85

Table 15. Pairwise correlations between the exponents in the scaling with dimensionless variables, estimated with WLS by means of a bootstrapping procedure using the DB5.2.3-STD5 ELMy H-mode ITER-like data set, including triangularity.

	α_ρ	α_{β_t}	α_ν	α_q	$\alpha_{R,\mathrm{D}}$	$\alpha_{1+\delta}$	$\alpha_{\kappa,\mathrm{D}}$	$\alpha_{\epsilon,\mathrm{D}}$	$\alpha_{M,\mathrm{D}}$
α_ρ	**1.0**	−0.56	0.53	−0.19	0.81	0.021	−0.060	0.72	−0.029
α_β	−0.56	**1.0**	−0.30	0.67	−0.37	−0.23	−0.38	−0.64	−0.018
α_ν	0.53	−0.30	**1.0**	−0.24	0.65	0.062	0.0017	0.58	0.084
α_q	−0.19	0.67	−0.24	**1.0**	−0.035	−0.10	−0.25	−0.51	−0.14
α_R	0.81	−0.37	0.65	−0.035	**1.0**	0.072	−0.062	0.56	−0.013
$\alpha_{1+\delta}$	0.021	−0.23	0.062	−0.10	0.072	**1.0**	0.063	−0.017	0.041
α_κ	−0.060	−0.38	0.0017	−0.25	−0.062	0.063	**1.0**	0.19	−0.23
α_ϵ	0.72	−0.64	0.58	−0.51	0.56	−0.017	0.19	**1.0**	−0.092
α_M	−0.029	−0.018	0.084	−0.14	−0.013	0.041	−0.23	−0.092	**1.0**

- The confinement appears to have a residual dependence on elongation κ_a, but no clear scaling with $1+\delta$ and ϵ remains.
- There is a noticeable scaling with effective mass, which appears to be significant.
- In the case of RBAYES, the confinement time prediction in ITER, 2.53 s, is very similar to the prediction obtained from the engineering scaling. The standard deviation is however quite large, about 1 s (95% credible interval [−47%, +88%]).

4.2.5. ELMy H-modes in STD5 with δ A somewhat more general scaling, that takes into account the influence of triangularity, can be given based on all ELMy H-mode plasmas in the STD5 set, not restricting to the ITER-like subset. This scaling is shown in table 16 in engineering form, obtained using WLS, and in dimensionless form, obtained using the RBAYES regression method applied in the space of dimensionless variables.

Concentrating first on the engineering scaling, the main difference with the corresponding scaling on the ITER-like subset (table 10) is the considerably stronger dependence on major radius in the unconstrained data set. This is accompanied by a slightly higher extrapolation of the confinement time to ITER. Apart from the additional

Table 16. Estimates of the parameters (intercept α_0 and exponents) and prediction toward ITER in power law scalings of the global H-mode energy confinement time, using the ELMy H-mode plasmas in the DB5.2.3-STD5 database. The scaling with engineering variables was obtained from the logarithmically transformed data using WLS, while for the scaling with dimensionless variables RBAYES was used.

	α_0	α_I	α_B	α_n	α_P	α_R	$\alpha_{1+\delta}$	α_κ	α_ϵ	α_M	$\hat{\tau}_{\mathrm{E,th,ITER}}$ (s)	MdAPE (%)	RMSE	R^2
WLS	0.0534 ±0.0013	0.976 ±0.015	0.218 ±0.013	0.2442 ±0.0090	−0.6687 ±0.0063	1.710 ±0.023	0.362 ±0.033	0.799 ±0.027	0.354 ±0.031	0.195 ±0.016	3.067 ±0.038	6.1	0.18	0.96
	α_0	α_ρ	α_{β_t}	α_ν	α_q	$\alpha_{R,\mathrm{D}}$	$\alpha_{1+\delta}$	$\alpha_{\kappa,\mathrm{D}}$	$\alpha_{\epsilon,\mathrm{D}}$	$\alpha_{M,\mathrm{D}}$	$\hat{\tau}_{\mathrm{E,th,ITER}}$ (s)	MdAPE (%)	RMSE	R^2
RBAYES	0.087×10^{-6} $^{+0.51 \times 10^{-6}}_{-0.075 \times 10^{-6}}$	−2.24 ±0.34	0.20 ±0.21	−0.469 ±0.061	−0.70 ±0.37	−0.12 ±0.33	0.36 ±0.64	1.24 ±0.47	−1.70 ±0.51	0.53 ±0.30	3.07 ±0.82	16	0.32	0.91

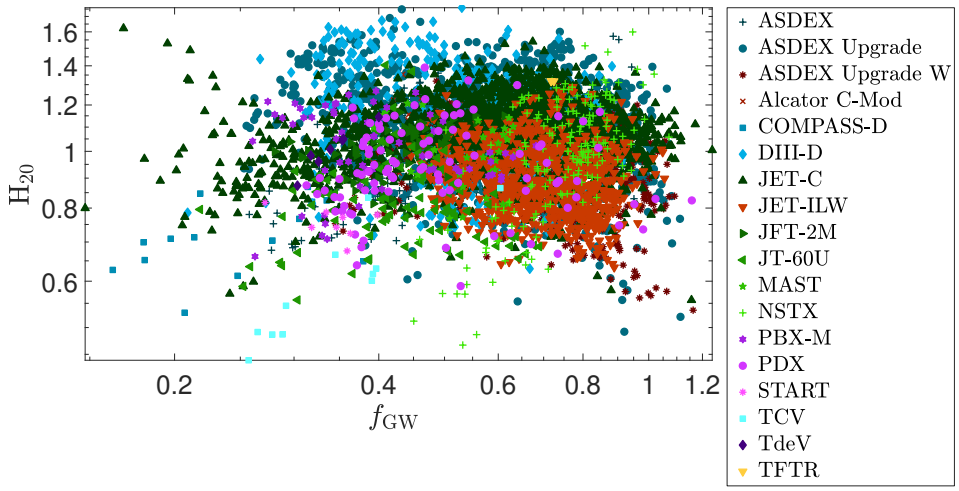

Figure 8. Confinement enhancement factor H_{20} versus Greenwald fraction f_{GW} for the DB5.2.3-STD5 ELMy H-mode data, obtained using WLS.

dependence on δ (although weak), this scaling is very similar to the second scaling provided in table 9, which was performed on the same data set. A plot of the confinement enhancement factor H_{20} against Greenwald fraction is shown in figure 8. This plot can be directly compared to that in figure 2, confirming that, when approaching the Greenwald limit, the new scalings do not produce a downward trend of the confinement enhancement that is as steep as obtained from IPB98(y, 2).

As for the dimensionless scaling, comparing with the RBAYES scaling in table 14, one difference is the somewhat stronger dependence on ρ_* in the scaling on the unconstrained data set, again with an improved ITER prediction. In addition, on the ITER-like subset no clear scaling with κ_{a} and particularly ϵ was observed, considering the approximate error bars given by the Bayesian method, but a modest dependence on δ. In contrast, in the unrestricted data set both κ_{a} and ϵ appear to have a somewhat clearer influence on the confinement, whereas the dependence on δ is much less clear.

Table 17. Estimates of the parameters (intercept α_0 and exponents) and prediction toward ITER in power law scalings of the global H-mode energy confinement time, expressed in engineering variables, distinguishing between type of wall material. The results were obtained by WLS, using the DB5.2.3-STD5 ELMy H-mode ITER-like data set, including triangularity. Estimates obtained using the full data set are shown ('All', cf. table 10), as well as the low-Z and high-Z subsets.

	α_0	α_I	α_B	α_n	α_P	α_R	$\alpha_{1+\delta}$	α_κ	α_ϵ	α_M	$\hat{\tau}_{\mathrm{E,th,ITER}}$ (s)	MdAPE (%)	RMSE	R^2
All	0.0670 ±0.0021	1.291 ±0.017	−0.134 ±0.017	0.1473 ±0.0088	−0.6442 ±0.0063	1.194 ±0.027	0.560 ±0.032	0.673 ±0.039	−0.000 ±0.033	0.302 ±0.016	2.902 ±0.040	6.0	0.17	0.95
Low-Z	0.0667 ±0.0022	1.147 ±0.021	−0.072 ±0.021	0.228 ±0.010	−0.6016 ±0.0080	1.395 ±0.032	0.394 ±0.036	0.661 ±0.045	0.226 ±0.040	0.272 ±0.020	3.208 ±0.053	6.6	0.17	0.96
High-Z	0.189 ±0.012	1.485 ±0.024	−0.356 ±0.025	0.018 ±0.018	−0.6077 ±0.0088	0.671 ±0.046	—	—	—	0.312 ±0.022	1.822 ±0.035	3.6	0.12	0.96

4.2.6. Impact of wall material Clarifying the impact of wall material on confinement has been an important part of the motivation for updating the confinement database. In section 4.1 it was already noted that the full-metal devices AUG-W and JET-ILW may exhibit somewhat weaker power degradation compared to the situation with carbon-based plasma-facing components. In this section, the data from the purely metallic devices ('high-Z', i.e. Alcator C-MOD, AUG-W and JET-ILW) are grouped and the scaling is compared with that using the data from the other devices ('low-Z', i.e. AUG, COMPASS-D, DIII-D, JET-C, JFT-2M, JT-60U and PBX-M). Results of the engineering scaling obtained with WLS on the DB5.2.3-STD5 ELMy H-mode ITER-like data set are shown in table 17. Some predictor variables were left out in the scaling on the high-Z subset, due to an insufficient range. We make the following observations.

- The density dependence, already low in the full data set and the low-Z data, is virtually absent in the high-Z scaling.
- The power degradation is actually similar with the low-Z and high-Z data, possibly due to the very weak dependence in JT-60U and PBX-M (cf. table 8).
- The dependence on plasma size is even weaker than linear in the high-Z data.
- The prediction for ITER is significantly lower for the high-Z subset than for the low-Z data. However, it is important to note that the high-Z subset is made up of data from only three devices, including a large fraction of points from JET-ILW with strong gas fueling. These points have a clear influence of decreasing the ITER prediction in the high-Z subset (this is related to the trend shown in figure 5(c)), without necessarily reflecting the operational conditions aimed for in ITER. The low value of the extrapolation may also be related to the missing predictor variables in the high-Z scaling. As a result, the low-Z and high-Z scalings are not suitable for prediction.

5. Discussion

Analysis of the scaling of global plasma quantities in tokamaks, like the energy confinement time, with other plasma quantities whose settings are, to a certain extent, under the control of the designer or operator of a fusion device, remains a convenient approach to support model building and to make predictions toward new devices. The conceptual simplicity of the method is particularly appealing, but it is important to keep in mind that it is based on various assumptions, while in the implementation a great number of choices have to be made regarding data set, variables, statistical methods, etc. With these caveats in mind, the purpose of the present discussion is to summarize those results of the activity that can be stated with confidence and to highlight the issues for which further investigation may reduce some of the observed uncertainties.

This paper has reported on an update of the ITPA global H-mode confinement database with data from JET with the fully metallic ITER-like wall and ASDEX Upgrade with the full tungsten wall. In the new database, currently referred to as DB5.2.3, we have concentrated on the 'standard set' STD5 with the aim of studying the thermal energy confinement scaling according to a simple power law. Although the immediate motivation for updating the ITPA global H-mode confinement database and re-estimating the confinement scaling has been the availability of new data from fully metallic devices, it has to be stressed that, since the derivation of the IPB98(y, 2) confinement scaling using DB2.8, the database had already grown substantially, leading to DB4 [4]. Hence, it should be clear that, in general, the differences between IPB98(y, 2) and the results presented in this paper are not solely due to the addition of data from AUG-W and JET-ILW since 2015.

With respect to DB2.8, and concentrating on ELMy H-mode plasmas, the database has been extended to higher power, current, density and inverse aspect ratio ϵ. At the same time, some important correlations across the database have been reduced, notably between plasma current (or power loss) on the one hand, and shape parameters and effective mass on the other hand. This is important to help disentangling the respective confinement dependencies. As for the dimensionless variables, the range of ρ_*, β_t and q_cyl has increased, while the correlation between ρ_* and ϵ has decreased, but the negative correlation between ρ_* and R_geo has become stronger, as well as the positive correlation between β_t and ϵ. In general, the dependency structure formed by several of the predictor variables remains a point of attention.

In this paper, on the one hand standard practices have been followed to estimate the confinement scaling, motivated by the aim to compare with IPB98(y, 2). For instance, a simple power law on a logarithmic scale has been used, largely based on the same set of engineering variables calculated according to the standard recipes. Transformation of the scaling to a dimensionless form has also been demonstrated. On the other hand, a number of potential weaknesses have been pointed out of the standard regression methodology based on weighted least squares applied to a log-linear model. Motivated by the aim for robustness of the results with respect to (small) changes in the data and

the power-law model, two alternative regression techniques have been applied in this work. The Bayesian technique (RBAYES) and geodesic least squares (GLS) used here have the additional advantage that they can be applied in situations with considerable uncertainty on the predictor variables in the scaling law. One notable application, demonstrated in this paper, is the estimation of the dimensionless form of the scaling by direct regression analysis in the space of (logarithmic) dimensionless variables. Nevertheless, many regression results are comparable among the three methods used in this paper, therefore most of the reported estimates were obtained with simple WLS.

The uncertainty on the parameter estimates and predictions of the scaling is difficult to quantify precisely. This is partly due to the model uncertainty and the problem of identifiability of the individual parameters due to predictor correlations. Apart from the questions whether all important variables have been included in the scaling, preferably through a causal relationship, and how accurately they have been measured, there is the issue that a power law assumes a single turbulence mechanism characterized by one specific scale length. Strictly speaking, this rules out the possibility of multiple physical mechanisms of energy transport in different regions of the plasma (core versus pedestal) or the operational space (e.g. approaching large β_t). Such considerations warrant more complex models not studied in this paper, e.g. with a term for the core and one for the pedestal (offset linear scaling), or a power law with functional (non-constant) exponents (log-nonlinear scalings). Model uncertainty also provides motivation for the attention paid in this paper to the issue of practically useful uncertainty estimates for the scaling. In [24] (and summarized in [3]), several proposals were made and evaluated to account for violations of the assumptions underlying the classical confidence interval for predictions from the regression model. Not accounting for model uncertainty resulted in a 95% confidence interval of about [−20%,+25%] with respect to the ITER prediction of 3.62 s by IPB98(y, 2) [3]. On the other hand, consideration of the variation of predictions resulting from various log-nonlinear models suggested a confidence interval that is about twice as wide. Admittedly, the latter approach leaves a certain degree of arbitrariness, not unlike the Bayesian approach using a scaled covariance matrix, outlined here in section 4.2.4. In this sense, the width of the prediction intervals provided by the statistical analyses should be regarded as a guideline, rather than a strictly quantitative uncertainty measure.

Analysis of data from individual devices has confirmed considerable differences in the exponents for the energy confinement scaling between machines. Nevertheless, generally the density dependence is weakest in the ITER-like devices (ASDEX Upgrade, Alcator C-Mod, DIII-D and JET) and, among these devices, the level of power degradation is lowest in the metallic machines (Alcator C-Mod, AUG-W and JET-ILW).

Several multi-machine scalings have been carried out using the standard set STD5 in the new database, confirming a number of earlier observations [12, 43, 45]. Based on WLS regression, the ELMy H-mode plasmas in DB5.2.3-STD5 exhibit a slightly weaker scaling with density, power loss and major radius, when comparing with IPB98(y, 2). Restricting to a somewhat more ITER-like data set, with constraints on q_{95}, κ, ϵ and

$Z_{\rm eff}$, an even weaker dependence on density and power loss is seen for ELMy H-modes, possibly owing to the data from the metallic ITER-like devices AUG-W and JET-ILW. In addition, a strikingly reduced dependence on major radius is observed in this data set (slightly above linear) and inverse aspect ratio (no dependence). A weak dependence on the effective isotope mass is noted, although stronger than in IPB98(y, 2). These dependencies result in a prediction toward the baseline ITER $Q = 10$ ELMy H-mode scenario of about 2.7 s—25% lower than the prediction by IPB98(y, 2) under the same operational conditions. On the same data set DB5.2.3-STD5, restricted to ELM-free H-modes with the ITER-like constraints, the density dependence remains weak, but the power degradation is back at the level seen in IPB98(y, 2), while also the scaling with machine size and ϵ is largely restored. This results in a confinement prediction for ITER of about 3.5 s, i.e. similar to what IPB98(y, 2) predicts for ELMy H-modes. However, being very different from more recent approaches to ELM-free operation, it is doubtful whether these historical ELM-free regimes have much relevance for future operational scenarios.

Plasma triangularity appears to have some influence on global confinement in the DB5.2.3-STD5 ITER-like data set with ELMy H-modes, while not appreciably affecting the other exponents. Therefore, we have discussed this scaling in some more detail. Considering the moderate sensitivity of the goodness-of-fit to variations in the scaling exponents that respect their correlation structure, the degree of uncertainty on the parameter estimates and predictions has been illustrated using the results from the robust Bayesian method. It turns out that the dependencies on power loss, plasma current and major radius are relatively well established, while the magnetic field and line-averaged density seem to play a role of little significance. Of the three shape parameters δ, $\kappa_{\rm a}$ and ϵ, it appears that only triangularity δ has some overall influence on the confinement within the data set. The error bar (scaled standard deviation) on the accompanying exponent $\alpha_{1+\delta}$ is more than 50%, however, and, while the exponents of $\kappa_{\rm a}$ and ϵ are significant with respect to their original, unscaled standard deviations, they are well within the scaled error bars. Both RBAYES and GLS yield a sizeable isotope effect ($\alpha_M \sim 0.4$ to 0.6), as does WLS, although it is weaker ($\alpha_M \sim 0.3$). As for the extrapolation to ITER, the Bayesian method predicts a confinement time of about 2.6 ± 0.5 s, which is lower than the prediction by WLS (~ 2.9 s) and GLS (~ 2.7 s), although still within the typical error bar. The somewhat stronger power degradation seen by RBAYES and GLS, compared to WLS, together with a slightly differing size dependence, contributes to this difference.

Still in the DB5.2.3-STD5 ITER-like data set with ELMy H-modes and considering the influence of triangularity, the methods RBAYES and GLS agree that, overall, the experimental points from COMPASS-D are outliers with respect to the main trend estimated by the scaling. According to GLS, also some experimental points from PBX-M are quite remote from the main trend. This is confirmed by calculating the confinement enhancement factor H_{20} with respect to the new scaling, obtained with WLS in this data set. In addition, whereas $H_{98}(y, 2)$ with the new data shows a downward trend

when approaching the Greenwald limit, this is less the case with H_{20}. Nevertheless, considerable residual dependence on the separatrix density is observed in the data from ASDEX Upgrade and JET-ILW, with confinement decreasing toward higher densities.

Considering formulation of the scaling in dimensionless variables, the scalings estimated by the three methods in the DB5.2.3-STD5 ITER-like data set with ELMy H-modes all obey quite well the high-β Kadomtsev constraint. However, direct transformation of the scaling from engineering to dimensionless variables results in large uncertainties on the transformed exponents. Therefore, in this paper regression analysis taking place directly in the space of (logarithmic) dimensionless variables has been explored. Focusing on the results obtained with the Bayesian method, a vanishingly small dependence on machine size has been obtained, suggesting a dimensionally correct scaling. Overall, a Bohm-type scaling of the turbulent transport is observed, which could be a reason for concern, although it has to be kept in mind that the corresponding error bars are relatively large. On the other hand, the strong degradation with plasma β seen in IPB98(y, 2) has disappeared, as also noted in [32]. Other differences with IPB98(y, 2) are the weakly inverse dependence on collisionality and inversely linear dependence on safety factor.

Finally, the impact of wall material has been studied, grouping data from multiple devices in a low-Z and high-Z subset of the DB5.2.3-STD5 ITER-like data with ELMy H-modes. This suggests vanishing dependence of confinement on the line-averaged density in fully metallic devices and a size dependence that is even weaker than in the non-metallic devices. However, the resulting scalings are not intended for prediction.

6. Conclusion

In conclusion, analysis of standard subsets of the updated ITPA global H-mode confinement database has confirmed a number of earlier observations, but has also revealed several differences in the energy confinement time scaling, when comparing to the previous de facto standard IPB98(y, 2). Most notably, the disagreement regarding the strength of some confinement dependencies between the global scaling and single-machine scans has been alleviated to some extent. In engineering form, the presently recommended scaling for ELMy H-modes, based on a power law, is obtained by weighted least squares regression (WLS) on the standard set STD5 of the ITPA DB5.2.3 database, restricted to the ITER-like subset and including triangularity as a predictor variable. The estimates by WLS are selected for the proposed scaling as a compromise between the results from the three regression techniques applied in this work, further motivated by the simplicity of the WLS technique and the relatively minor differences among the estimates from the different methods. The expression, obtained from table 10, is as follows:

$$\tau_{\mathrm{E,th}} = \left(0.067\ {}^{+0.059}_{-0.032}\right) I_{\mathrm{p}}^{1.29\pm0.16}\, B_{\mathrm{t}}^{-0.13\pm0.17}\, \bar{n}_{\mathrm{e}}^{0.147\pm0.097}\, P_{\mathrm{l,th}}^{-0.644\pm0.061}\, R_{\mathrm{geo}}^{1.19\pm0.27} \times (1+\delta)^{0.56\pm0.36}\, \kappa_{\mathrm{a}}^{0.67\pm0.63}\, M_{\mathrm{eff}}^{0.30\pm0.16}. \quad (5)$$

This relation will be referred to as **ITPA20-IL** (ITER-like) and it forms the basis for the definition of the confinement enhancement factor H_{20}, introduced in Section 4.2.4. It can be used for approximate prediction by substituting values for the predictor variables at the desired prediction point. The accompanying uncertainty estimates are derived from the Bayesian analysis and, as such, correspond to a rough estimate of a single standard deviation. However, it is essential to note that, for extrapolation, one must not vary one of the exponents in this expression, even within its error bar, while keeping the others fixed. Doing so increases the risk of an unreliable prediction. This is due to the correlation structure between the exponents: changing one exponent in general requires adapting others as well. For the same reason, for prediction one should generally not ignore dependencies with exponents that are negligible within their error bar. The exception here is the dependence on ϵ, which has been ignored in expression (5), as the exponent is almost zero. Rather, the uncertainty estimates are intended to be considered jointly for estimation of extrapolation uncertainty, for comparison with experimentally observed dependencies and theoretical models, and as an input to modeling activities seeking uncertainty quantification.

The most salient differences between scaling (5) and the IPB98(y, 2) scaling are the nearly vanishing dependence on line-averaged density, the almost linear dependence on major radius, the dependence on triangularity and the lack of dependence on inverse aspect ratio. There is also a slightly positive dependence on isotope mass. In comparison with IPB98(y, 2), the weaker dependence on machine size may be an important cause for the lower confinement time prediction of 2.90 ± 0.46 s for the baseline inductive $Q = 10$ ELMy H-mode scenario in ITER (2.65 ± 0.42 s at $P_{\mathrm{l,th}} = 100$ MW).

Expressed in dimensionless variables, the recommended confinement scaling is obtained by means of the robust Bayesian analysis in dimensionless space, using the same standard set in DB5.2.3, with ITER-like constraints (table 14):

$$\Omega_{\mathrm{i}}\tau_{\mathrm{E,th}} = \left(2.1\ {}^{+19}_{-1.9}\right) \times 10^{-6}\, \rho_*^{-1.97\pm0.33}\, \beta_{\mathrm{t}}^{0.12\pm0.22}\, \nu_*^{-0.426\pm0.063}\, q_{\mathrm{cyl}}^{-1.03\pm0.43} \times (1+\delta)^{0.14\pm0.64}\, \kappa_{\mathrm{a}}^{1.89\pm0.85} \epsilon^{-0.26\pm0.64}\, M_{\mathrm{eff}}^{0.62\pm0.37}. \quad (6)$$

This scaling, not intended for the purpose of prediction, will be denoted **ITPA20-IL-dim**. Notable differences with IPB98(y, 2) are the weaker scaling with normalized gyroradius (Bohm-type scaling), normalized plasma pressure (electrostatic scaling) and cylindrical safety factor, as well as a stronger inverse scaling with collisionality.

Slightly more general scalings are obtained from the entire set of ELMy H-modes in DB5.2.3-STD5, not restricting to the ITER-like subset. Reproduced from table 16 by means of the same respective methodology as ITPA20-IL and ITPA20-IL-dim, these scalings are as follows. The engineering scaling, henceforth referred to as **ITPA20** (error bars from RBAYES), is

$$\tau_{\mathrm{E,th}} = \left(0.053\ {}^{+0.030}_{-0.018}\right) I_{\mathrm{p}}^{0.98\pm0.19}\, B_{\mathrm{t}}^{0.22\pm0.18}\, \bar{n}_{\mathrm{e}}^{0.24\pm0.11}\, P_{\mathrm{l,th}}^{-0.669\pm0.059}\, R_{\mathrm{geo}}^{1.71\pm0.32} \times (1+\delta)^{0.36\pm0.39}\, \kappa_{\mathrm{a}}^{0.80\pm0.38} \epsilon^{0.35\pm0.66}\, M_{\mathrm{eff}}^{0.20\pm0.17}, \quad (7)$$

yielding an ITER confinement prediction of 3.07 ± 0.46 s (2.79 ± 0.42 s at $P_{\mathrm{l,th}} = 100$ MW). Other prediction points can be substituted. Apart from the weaker dependence on density and major radius, and the suggestion of a weak dependence on triangularity, this scaling is relatively similar to IPB98(y, 2). In contrast, the dimensionless scaling that is proposed for this data set, called **ITPA20-dim** (not for prediction), is quite different from the dimensionless form of IPB98(y, 2):

$$\Omega_{\mathrm{i}}\tau_{\mathrm{E,th}} = \left(0.087\ {}^{+0.51}_{-0.075}\right) \times 10^{-6}\, \rho_*^{-2.24\pm0.34}\, \beta_{\mathrm{t}}^{0.20\pm0.21}\, \nu_*^{-0.469\pm0.061}\, q_{\mathrm{cyl}}^{-0.70\pm0.37} \times (1+\delta)^{0.36\pm0.64}\, \kappa_{\mathrm{a}}^{1.24\pm0.47} \epsilon^{-1.70\pm0.51}\, M_{\mathrm{eff}}^{0.53\pm0.30}. \quad (8)$$

It is important to stress again that the parametric dependencies of the confinement suggested by these scaling laws are averages across a multi-machine database and a broad variety of plasma conditions. The dependencies can differ to a varying extent from those obtained in dedicated parameter scans. Indeed, the single-machine scalings obtained from DB5 already point at considerable differences of the exponents between devices.

The analysis of the updated database has also raised a number of issues that call for more detailed analysis. One important point of attention is the regression model used for the global confinement scaling, both concerning the functional form and the choice of predictor variables. Indeed, dependencies between predictor variables in the regression still hamper unambiguous estimation of the individual dependencies. From the statistical point of view, the best remedy is to acquire more data to fill in some sparsely populated regions of the data space. Data from additional devices regarding possible 'hidden variables' affecting the confinement (separatrix density, plasma rotation) would be very useful as well. Collection of additional data should be complemented with variable selection based on physical and statistical arguments. Furthermore, physical constraints may be invoked to further regularize the scaling, e.g. a soft Kadomtsev constraint imposed by a Bayesian prior distribution. Finally, more flexible models based on simple generalizations of the power law (e.g. separating core and pedestal, distinguishing between transport regimes or using a power law with variable exponents) may enable faithful representation of the confinement trends over a wide range of the parameter space, while still allowing confident extrapolation to new devices. Thus, both maintenance and extension of the database are expected to remain important activities.

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 and 2019–2020 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. G.V. would like to thank the Princeton Plasma Physics Laboratory for the hospitality during several research stays, as well as Ondrej Grover for interesting discussion.

Appendix A. Transformation formulas between engineering and dimensionless exponents

The transformation formulas between exponents in the engineering scaling (2) and those of the scaling in dimensionless form (4) are provided in this appendix.§§ From dimensionless to engineering form one finds

$$\boldsymbol{\alpha} = \frac{1}{1-\gamma}\left[\begin{array}{c|c} \boldsymbol{a} & A \end{array}\right]\left[\begin{array}{c} 1 \\ \boldsymbol{\alpha}_{\mathrm{D}} \end{array}\right],$$

where

$$\begin{aligned}
\boldsymbol{\alpha} &= [\alpha_I, \alpha_B, \alpha_n, \alpha_P, \alpha_R, \alpha_{1+\delta}, \alpha_\kappa, \alpha_\epsilon, \alpha_M]^{\mathrm{t}}, \\
\boldsymbol{\alpha}_{\mathrm{D}} &= [\alpha_\rho, \alpha_\beta, \alpha_\nu, \alpha_q, \alpha_{R,\mathrm{D}}, \alpha_{1+\delta}, \alpha_{\kappa,\mathrm{D}}, \alpha_{\epsilon,\mathrm{D}}, \alpha_{M,\mathrm{D}}]^{\mathrm{t}}, \\
\boldsymbol{a} &= [0, -1, 0, 0, 0, 0, 0, 0, 1]^{\mathrm{t}}, \\
\gamma &= \frac{\alpha_\rho}{2} + \alpha_\beta - 2\alpha_\nu,
\end{aligned}$$

and

$$A = \begin{array}{c} \begin{array}{ccccccccc} \rho_* & \beta_{\mathrm{t}} & \nu_* & q_{\mathrm{cyl}} & R_{\mathrm{geo}} & 1{+}\delta & \kappa_{\mathrm{a}} & \epsilon & M_{\mathrm{eff}} \end{array} \\ \left(\begin{array}{ccccccccc} 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1/2 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ -5/2 & -3 & 8 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1/2 & -1 & 3 & 1 & 0 & 0 & 1 & 0 & 0 \\ -2 & -2 & 9/2 & 2 & 0 & 0 & 0 & 1 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array} \begin{array}{l} I_{\mathrm{p}} \\ B_{\mathrm{t}} \\ \bar{n}_{\mathrm{e}} \\ P_{\mathrm{l,th}} \\ R_{\mathrm{geo}} \\ 1{+}\delta \\ \kappa_{\mathrm{a}} \\ \epsilon \\ M_{\mathrm{eff}} \end{array} .$$

This can be inverted for transforming from engineering to dimensionless scalings:

$$\boldsymbol{\alpha}_{\mathrm{D}} = A^{-1}\left(\frac{\boldsymbol{\alpha}}{1+\alpha_P} - \boldsymbol{a}\right), \tag{A.1}$$

§§The transformation for scaling laws without $1+\delta$ is obtained simply by leaving the corresponding rows and columns out of the matrix A and vectors $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}_{\mathrm{D}}$ and $\boldsymbol{a}$.

where

$$A^{-1} = \begin{array}{c} \begin{matrix} I_\mathrm{p} & B_\mathrm{t} & \bar{n}_\mathrm{e} & P_\mathrm{l,th} & R_\mathrm{geo} & 1{+}\delta & \kappa_\mathrm{a} & \epsilon & M_\mathrm{eff} \end{matrix} \\ \begin{pmatrix} -3/2 & -3/2 & -2 & -3 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1 & 3/2 & 0 & 0 & 0 & 0 & 0 \\ -1/4 & -1/4 & 0 & -1/2 & 0 & 0 & 0 & 0 & 0 \\ -3/4 & 1/4 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ -1/4 & -5/4 & -2 & 1/2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1/8 & -15/8 & -2 & -7/4 & 0 & 0 & 0 & 1 & 0 \\ 3/4 & 3/4 & 1 & 3/2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{matrix} \rho_* \\ \beta_\mathrm{t} \\ \nu_* \\ q_\mathrm{cyl} \\ R_\mathrm{geo} \\ 1{+}\delta \\ \kappa_\mathrm{a} \\ \epsilon \\ M_\mathrm{eff} \end{matrix} .$$

In particular, it follows that $4\alpha_{R,\mathrm{D}} = (-\alpha_I - 5\alpha_B - 8\alpha_n + 2\alpha_P + 4\alpha_R)/(1+\alpha_P) - 5$ and forcing this to vanish is indeed equivalent to the Kadomtsev constraint mentioned in the main text.

Appendix B. Regression workflow

All analysis described in this paper was carried out using the MATLAB software [46]. To facilitate the workflow and reproducibility of the results, the information about each specific analysis is kept in a configuration file. These configuration files specify the subset of the data to be loaded from the DB5.2.3 database, as well as details about the regression analysis (model, variables, analysis options, etc.) and the output files. When the analysis is finished, the configuration details, the specific data set used for the analysis and the regression results are stored in a `mat` file. In addition, the analysis results (tables and figures) are automatically written to a LaTeX file for publication purposes.

Two of the regression techniques applied in this work (RBAYES and GLS) require error estimates on the scaling variables. Depending on the device, some of the scaling variables are expressed in terms of other database variables (e.g. the thermal stored energy as a function of the total energy content and the fast particle contribution). Therefore, at the start of each analysis, the data corresponding to the scaling variables is automatically calculated based on their respective expressions in terms of other database variables. The associated error estimates are derived by means of standard Gaussian error propagation (based on assumptions described in the main text). To that end, for each device represented in the database, the expressions for the various scaling variables (engineering and dimensionless) are stored in a device-specific file, together with the information about measurement uncertainty.

Appendix C. Bayesian and geodesic least squares regression methods

Appendix C.1. Regression model

One of the key purposes of the Bayesian (RBAYES) and GLS methods is to take into account the possibility of non-negligible uncertainty on all variables in a general power-law scaling. To do this, one may consider an idealized situation where, on a logarithmic scale, the following linear relation holds between a response variable η and a set of predictor variables $\xi_1, \ldots \xi_{n_{\rm pred}}$:

$$\eta = \ln \alpha_0 + \sum_{j=1}^{n_{\rm pred}} \alpha_j \xi_j. \tag{C.1}$$

In a next step, it is common to assume that the measured data, described by variables y, x_1, …, $x_{n_{\rm pred}}$, are related to those idealized variables by

$$y = \eta + \epsilon_y, \quad x_1 = \xi_1 + \epsilon_{x_1}, \quad \ldots, \quad x_{n_{\rm pred}} = \xi_{n_{\rm pred}} + \epsilon_{x_{n_{\rm pred}}}.$$

Here, the uncertainty with respect to the idealized relation (C.1) is described by $n_{\rm pred}+1$ normally distributed variables with zero mean:

$$\epsilon_y \sim \mathcal{N}\left(0, \sigma_y^2\right), \quad \epsilon_{x_1} \sim \mathcal{N}\left(0, \sigma_{x_1}^2\right), \quad \ldots, \quad \epsilon_{x_{n_{\rm pred}}} \sim \mathcal{N}\left(0, \sigma_{x_{n_{\rm pred}}}^2\right).$$

Appendix C.2. Robust Bayesian regression

Given $n_{{\rm obs},k}$ assumed independent observations from device k, out of a total of $n_{\rm tok}$ tokamaks, the likelihood in a Bayesian approach is obtained by marginalizing (integrating) over the unobserved values $\xi_{j,i_k,k}$ ($j = 1, \ldots, n_{\rm pred}$, $i_k = 1, \ldots, n_{{\rm obs},k}$, $k = 1, \ldots, n_{\rm tok}$) assumed by the variables ξ_j. This results in the following joint distribution of the full set of measured data, conditional on the parameter values:

$$\begin{aligned} &p(\{y_{i_k,k}\}, \{x_{j,i_k,k}\} | \{\alpha_0, \alpha_j\}, \{\gamma_k\}) \\ &\quad = \prod_{k=1}^{n_{\rm tok}} \prod_{i_k=1}^{n_{{\rm obs},k}} \frac{1}{\sqrt{2\pi\gamma_k^2\sigma_{{\rm mod},k}^2}} \exp\left[-\frac{1}{2}\frac{\left(y_{i_k,k} - \eta_{i_k,k}\right)^2}{\gamma_k^2\sigma_{{\rm mod},k}^2}\right]. \end{aligned} \tag{C.2}$$

It can be shown that, for a linear model (C.1), the parameters $\sigma_{{\rm mod},k}$ are given by [47, 48, 49, 36]

$$\sigma_{{\rm mod},k}^2 = \sigma_{y,k}^2 + \sum_{j=1}^{n_{\rm pred}} \alpha_j^2 \sigma_{x_j,k}^2. \tag{C.3}$$

We assume that, in this expression, the standard deviations $\sigma_{y,k}$ and $\sigma_{x_j,k}$ correspond to the measurement uncertainties (percentage errors) provided in the database, which vary among the different machines. The result (C.3) can also be obtained through standard Gaussian error propagation. We call $\sigma_{{\rm mod},k}$ the *modeled standard deviation* for device k, as the expression (C.3) hinges on the model assumptions. At the same time, however, we admit that the actual spread of the data around the fitted hyperplane may differ from that suggested by the known measurement uncertainties. This is modeled

by introducing, for each device k, an additional parameter γ_k that scales the respective modeled standard deviations $\sigma_{\mathrm{mod},k}$. As discussed in the main text, the actual spread can become (significantly) wider ($\gamma_k > 1$) than suggested by the error bars on the measured data, e.g. due to model uncertainty. As such, a certain degree of robustness is achieved by the RBAYES technique, as an alternative to employing for instance a heavy-tailed likelihood distribution [49]. The following maximally uninformative priors were used for the parameters [47]:

$$p(\{\alpha_0, \alpha_j\}) = \left(1 + \sum_{j=1}^{n_{\mathrm{pred}}} \alpha_j^2\right)^{-(n_{\mathrm{pred}}+2)/2}, \qquad p(\{\gamma_k\}) = \prod_{k=1}^{n_{\mathrm{tok}}} \gamma_k^{-1}.$$

The joint posterior distribution of the regression parameters and the factors γ_k was sampled using a Markov chain Monte Carlo (MCMC) technique (Hamiltonian Monte Carlo with $N_{\mathrm{s}} = 4\,000$ samples). In practice, this means that the correction factors γ_k were learned from the data.

Appendix C.3. Geodesic least squares regression

GLS shares with WLS the benefit of simplicity, but, like RBAYES, GLS is relatively robust against deviations from the model assumptions [35, 36]. For instance, the method has been shown to yield consistent results, irrespective of log-transformation of the data. It also takes into account errors in all variables and can accommodate heteroscedastic errors. Somewhat akin to the RBAYES technique, GLS learns for each device the scatter of the data around the fitted hyperplane (or hypersurface in the nonlinear case). The method then minimizes, for each data point, the difference between two conditional probability models for the response variable. On the one hand, a normal probability distribution is considered, centered on the measured data point:

$$p(y|y_{i_k,k}, \sigma_{\mathrm{obs},k}) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{obs},k}^2}} \exp\left[-\frac{1}{2}\frac{(y - y_{i_k,k})^2}{\sigma_{\mathrm{obs},k}^2}\right].$$

The device-dependent standard deviation $\sigma_{\mathrm{obs},k}$, which we call the *observed standard deviation*, models the actual dispersion of the data and is learned from the data. This model is designed with minimal assumptions in mind. On the other hand, the normal distributions used for the likelihood in (C.2) (without the factors γ_k) can be viewed as probability models that would be exact if all the model assumptions were perfectly obeyed. Estimation of the model parameters and the observed standard deviations is achieved by minimizing a sum of squared distances between each pair of distributions, over all data points. The distance measure used is the *Rao geodesic distance* based on the Fisher information metric on the Riemannian manifold of univariate normal distributions [50].

Appendix C.4. Weighting scheme

Both RBAYES and GLS encompass a weighting scheme, although it is different from the weighting implemented in WLS in this paper. Specifically, the likelihood distribution employed by RBAYES and GLS weighs the residuals by the uncertainty on the data, taking into account a device-dependent deviation from the measurement uncertainty (through γ_k or $\sigma_{\mathrm{obs},k}$). Hence, both RBAYES and GLS allow significant heteroscedasticity between devices, on top of that between data points, which is a key ingredient leading to their robustness. No additional weighting based on the number of data points contributed by each device was implemented in the RBAYES and GLS methods. As a result, these methods do not depend on any choice for the weights.

References

[1] Kaye S and Goldston R 1985 *Nucl. Fusion* **25** 65–69
[2] Christiansen J *et al.* 1992 *Nucl. Fusion* **32** 291–338
[3] ITER Physics Basis Expert Groups on Confinement and Transport and Confinement Modelling and Database, ITER Physics Basis Editors 1999 *Nucl. Fusion* **39** 2175–2249
[4] McDonald D *et al.* 2007 *Nucl. Fusion* **47** 147–174
[5] Kaye S *et al.* 2006 *Plasma Phys. Control. Fusion* **48** A429–A438
[6] Valovic M *et al.* 2005 *Nucl. Fusion* **45** 942–949
[7] Kardaun O *et al.* 2006 The tortuous route of confinement prediction near operational boundary. Analysis based on ITERH.DB4/L.DB3 database (application: ITER FEAT) (Proc. 21st Int. Conf. Chengdu, 2006) (Vienna: IAEA) CD-ROM file [IT/P1-10] and http://www-naweb.iaea.org/napc/physics/FEC/FEC2006/html/index.htm
[8] Kardaun O 2017 Reflections on plasma confinement and operation boundaries, ensuing from experimental thermonuclear installation–TOKAMAK-20–preliminary study IPP 2017-05 Max Planck Institute for Plasma Physics Garching, Germany URL `http://hdl.handle.net/11858/00-001M-0000-002E-746D-7`
[9] Challis C *et al.* 2015 *Nucl. Fusion* **55** 053031
[10] Ryter F *et al.* 2001 *Nucl. Fusion* **41** 537–550
[11] Petty C 2008 *Phys. Plasmas* **15** 080501
[12] Verdoolaege G, Kaye S, Angioni C, Kardaun O, Maslov M, Romanelli M, Ryter F, Thomsen K, the ASDEX Upgrade Team, the EUROfusion MST1 Team and JET Contributors 2018 First analysis of the updated ITPA global H-mode confinement database *Preprint: IAEA Fusion Energy Conference, [EX/P7-1]* (Gandhinagar, India)
[13] Maslov M, Romanelli M, Brix M, Flanagan J, Boboc A and JET Contributors 2018 The EUROfusion JET-ILW global confinement database *Proc. 45th EPS Conference on Plasma Physics,* P5.1064 (Prague, Czech Republic)
[14] Maslov M, Boboc A, Brix M, Flanagan J, Peluso E, Price C, Romanelli M and JET Contributors 2020 *Nucl. Fusion* **60** 036007
[15] Ryter F *et al.* 2021 *Nucl. Fusion* **61** 046030
[16] 2021 International Global H-Mode Confinement Database `https://osf.io/drwcq` and `http://arks.princeton.edu/ark:/88435/dsp01m900nx49h`
[17] Nunes I and JET Contributors 2016 *Plasma Phys. Control. Fusion* **58** 014034
[18] Schweinzer J *et al.* 2011 *Nucl. Fusion* **51** 113003
[19] Belsley D, Kuh E and Welsch R 1980 *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity* (New York: John Wiley & Sons)
[20] Belsley D 1991 *Computer Science in Economics and Management* **4** 33–50

[21] Kadomtsev B 1975 *Sov. J. Plasma Phys.* **1** 295–297
[22] Samaras D 1971 *Theory of ion flow dynamics* (New York: Dover Publications)
[23] Luce T, Petty C and Cordey J 2008 *Plasma Phys. Control. Fusion* **50** 043001
[24] Kardaun O 1999 *Plasma Phys. Control. Fusion* **41** 429–469
[25] Garcia J, Cambon D and JET Contributors 2018 *Plasma Phys. Control. Fusion* **60** 025028
[26] Murari A, Peluso E, Gaudio P and Gelfusa M 2017 *Nucl. Fusion* **57** 126017
[27] Connor J and Taylor J 1977 *Nucl. Fusion* **17** 1047–1055
[28] Connor J 1988 *Plasma Phys. Control. Fusion* **30** 619–650
[29] Christiansen J *et al.* 1991 *Nucl. Fusion* **31** 2117–2129
[30] Pearl J 2014 *Am. Stat.* **68** 8–13
[31] McDonald D, Meakins A, Svensson J, Kirk A, Andrew Y, Cordey J and ITPA H-mode Threshold Database WG 2006 *Plasma Phys. Control. Fusion* **48** A439–A447
[32] Cordey J *et al.* 2005 *Nucl. Fusion* **45** 1078–1084
[33] Kardaun O 2005 *Classical Methods of Statistics* (Berlin: Springer)
[34] Kardaun O 2002 Interval estimate of the global energy confinement time during ELMy H-mode in ITER-FEAT, based on the international multi-tokamak ITERH.DB3 dataset IPP-IR-2002/5 1.1 Max Planck Institute for Plasma Physics Garching, Germany URL `http://hdl.handle.net/11858/00-001M-0000-0027-40C4-1`
[35] Verdoolaege G and Noterdaeme J M 2015 *Nucl. Fusion* **55** 113019
[36] Verdoolaege G 2015 *Entropy* **17** 4602–4626
[37] Xiao X *et al.* 2011 *Ecology* **92** 1887–1894
[38] Aitken A 1934 *Proc. Roy. Soc. Edinburgh A* **55** 42–48
[39] Doyle E *et al.* 2007 *Nucl. Fusion* **47** S18–S127
[40] Hawryluk R *et al.* 2009 *Nucl. Fusion* **49** 065012
[41] Kim S, Casper T and Snipes J 2018 *Nucl. Fusion* **58** 056013
[42] Koechl F *et al.* 2020 *Nucl. Fusion* **60** 066015
[43] Kaye S, Verdoolaege G, Angioni C, Kardaun O, Maslov M, Romanelli M, Ryter F and Thomsen K 2018 First analysis of the updated ITPA global H-mode confinement database *60*$^{\text{th}}$ *Annual Meeting of the APS Division of Plasma Physics,* TP11.00104 (Portland, OR, USA)
[44] Ryter F *et al.* 1993 *Nucl. Fusion* **33** 979–990
[45] Sips A *et al.* 2018 *Nucl. Fusion* **58** 126010
[46] MATLAB 2019 *9.7.0.1261785 (R2019b)* (Natick, Massachusetts: The MathWorks Inc.)
[47] Preuss R and Dose V 2005 Errors in all variables *Bayesian Inference and Maximum Entropy Methods in Science and Engineering* (*AIP Conference Proceedings* vol 803) (Melville, NY) pp 448–455
[48] Von Toussaint U 2015 *Entropy* **17** 3898–3912
[49] von der Linden W, Dose V and von Toussaint U 2014 *Bayesian Probability Theory: Applications in the Physical Sciences* (Cambridge: Cambridge University Press)
[50] Amari S and Nagaoka H 2000 *Methods of Information Geometry* (New York: American Mathematical Society)