MODEL COMPARISON IN PLASMA ENERGY CONFINEMENT SCALING

R. PREUSS, V. DOSE AND W. VON DER LINDEN
Max-Planck-Institut für Plasmaphysik, EURATOM Association
D-85748 Garching b. München, Germany †

Abstract. Energy confinement data of large fusion experiments have been analyzed in terms of dimensionless form free scaling functions. Several possible physical scenarios lead to different models with a certain number of degrees of freedom. These are used to set up the confinement function as a linear combination of dimensionless power law terms. Then one has to solve the problem, given a particular plasma model which is the optimum number of terms in the expansion, and which model is most likely in the light of the data. Based upon data containing the energy content for a wide variety of variable settings, predictions for single variable scans are made and compared to actual experiments. A very good agreement is obtained.

Key words: Model Comparison, Prior Predictive Value, Invariant Measure

1. Introduction

The plasma behavior in large fusion experiments is now described for over twenty years by energy confinement scaling functions [1]. Since an exact description of energy transport is not possible scaling laws are in common use to give an overview of the plasma machine properties. This allows inter-machine comparisons and the characterization of conditions for enhanced confinement regimes. Even more, for being quite successful in the past in predicting the performance of larger plasma devices from data of small and midsized machines, scaling laws are involved in the design of future fusion reactors.

The functional form of the scaling law is still unknown. Only for reasons of simplicity a power law dependence is assumed and has accumulated considerable credit just by experience. In order to describe the energy content W^{theo} of the confined plasma all the plasma determining quantities in the machine under consideration are fed into the scaling law: For the in this paper examined W7-AS stellarator these are the particle density n, the magnetic field B, the heating power P and the minor radius a. Additional machine parameters like the major radius of torus

†Email: preuss@ipp.mpg.de

R come only into play in comparisons with machines of different size. The scaling law then reads:

$$W_i^{theo} = e^{\alpha_c} n_i^{\alpha_n} B_i^{\alpha_B} P_i^{\alpha_P} a_i^{\alpha_a} + \varepsilon_i \ , \tag{1}$$

where i labels a single measurement with unknown error ε. Also the scaling constant (e^{α_c}) has been expressed in exponential form. Now the above addressed reasons for the power law choice are obvious: Applying the logarithm to both sides of (1) an easy to solve set of linear equations in the sought α's is obtained.

2. Dimensionless Confinement Scaling

A major shortcoming of the above equation is to be solely a data fit and that it does not take care of the physical implications of the input variables. In order to obtain a dimensionally exact scaling function basic plasma models may be used. This was achieved by Connor and Taylor [2] in requiring that the invariances of basic plasma models under similarity transformations should be the same as those of the scaling function belonging to the respective model. Four kinetic models (see table 1) were obtained in considering the Boltzmann equation of motion (describing a collisional plasma) or the Vlasov equation (collisionless plasma) and discriminating further between the electrostatic limit (low-β case) or a self-consistent calculation of the electro-magnetic fields from the Maxwell equations (high-β). By the invariance requirement the plasma variables are combined in three dimensionless factors with new scaling exponents (x_1, x_2, x_3) specified according to table 1.

$$\begin{aligned} W^{theo} &\propto na^4RB^2\left(c_1\frac{P}{na^4RB^3}\right)^{x_1}\left(c_2\frac{a^3B^4}{n}\right)^{x_2}\left(c_3\frac{1}{na^2}\right)^{x_3} && (2)\\ &= c\left(\frac{P}{na^4B^3}\right)^{x_1}\left(\frac{a^3B^4}{n}\right)^{x_2}\left(\frac{1}{na^2}\right)^{x_3} && (3)\\ &= cf\left(n,B,P,a;\boldsymbol{x}\right) \ . && (4)\end{aligned}$$

The constants (c_1, c_2, c_3) in (2) consist of fundamental physical constants and carry physical units. They may be absorbed in an overall constant c in (3). Just the same may be done with the large radius R of the torus which is constant for the here considered dataset of a single machine. Finally, in (4) the scaling terms are comprised by a general function $f(n, B, P, a; \boldsymbol{x})$ where $\boldsymbol{x}$ has components x_1, x_2, x_3.

Already Connor and Taylor proposed to express a general form free energy confinement scaling function as a series of terms of the form (4) with properly chosen $\boldsymbol{x}$ and c. We want to exploit this idea and employ the following ansatz to describe the energy content $\boldsymbol{W}^{theo}$ of N different experiments:

$$\boldsymbol{W}^{theo} = \sum_{k=1}^{E} c_k \boldsymbol{f}\left(\boldsymbol{x}_k\right) = \boldsymbol{F}\boldsymbol{c} \ . \tag{5}$$

$\boldsymbol{F}$ is a $N \times E$ matrix with columns $\boldsymbol{f}(\boldsymbol{x}_k)$. In mathematical terms this is nothing but an expansion in a dimensionally exact basis. In general, N linearly independent

M_j	CT Model	x_1	x_2	x_3	N_{DOF}	$p(M_j\|\boldsymbol{W}^{exp},\boldsymbol{\sigma},I)$
1	collisionless low-β	x	0	0	1	4×10^{-12}%
2	collisional low-β	x	y	0	2	99.7%
3	collisionless high-β	x	0	z	2	0.25%
4	collisional high-β	x	y	z	3	0.025%

TABLE 1. Plasma models according to Connor and Taylor [2].

vectors $\boldsymbol{f}(\boldsymbol{x}_k)$ form a complete basis in the N-dimensional data space and would therefore allow a pointwise reconstruction of the data. This is neither desirable, nor with respect to physics correct, since the corresponding vector of the measured energy content $\boldsymbol{W}^{exp}$ is corrupted by noise. What we really want is an expansion, truncated at some appropriate upper limit E describing the physics, while the residual $N-E$ terms in the expansion (5) would only fit the noise. So we have to look for the most probable expansion order. Furthermore, given the four CT models of table 1 we would like to identify the plasma physics model which describes the data best. Since the number of degrees of freedom N_{DOF} varies between one and three (see table 1) Occam's Razor will have a say in this matter. And finally we would like to estimate the energy content for certain parameter settings. An important topic here are single variable scans (e.g. the variation of the energy content as function of the density alone with all other variables fixed). Such scans are not directly accessible from published databases. On the other hand, single variable scans are experimentally cumbersome and expensive experiments have to be performed for each and every input variable of interest. It is therefore highly desirable to extract single variable scans from existing databases by employing improved data analysis techniques.

3. Bayesian Approach

The likelihood function of our problem reads:

$$p(\boldsymbol{W}^{exp}|\omega,\boldsymbol{c},\boldsymbol{x},E,M_j,\boldsymbol{\sigma},I)=\left(\frac{\omega}{2\pi}\right)^{\frac{N}{2}}\prod_i^N\sigma_i^{-1}$$
$$\cdot\exp\left\{-\omega\sum_{i=1}^N\frac{\left(W_i^{exp}-\sum_k^E c_k f_i(\boldsymbol{x}_k)\right)^2}{2\sigma_i^2}\right\}\ . \tag{6}$$

M_j denotes the plasma kinetic model used to generate the expansion vectors $\boldsymbol{f}(\boldsymbol{x}_k)$. $\boldsymbol{\sigma}$ is the vector of experimental uncertainties associated with the measurement $\boldsymbol{W}^{exp}$. The uncertainty in the energy content $\boldsymbol{W}^{exp}$ contains the direct distributions from the diamagnetic measurement as well as indirect contributions from the finite precision in the input variables (n,B,P,a). Both contributions have been estimated to the best of our knowledge. However, to allow for possible deviations from the true errors we introduce an overall correction factor ω where we assume that the relative error is correctly estimated. The Bayesian analysis yields a posteriori for the most probable model and the optimum expansion order that $\boldsymbol{\sigma}$ was overestimated merely by a factor of 20%.

We are looking for the probability of a model given the data $\boldsymbol{W}^{exp}$. The odds ratio gives

$$\frac{p(M_j|\boldsymbol{W}^{exp},\boldsymbol{\sigma},I)}{p(M_k|\boldsymbol{W}^{exp},\boldsymbol{\sigma},I)}=\frac{p(M_j|\boldsymbol{\sigma},I)}{p(M_k|\boldsymbol{\sigma},I)}\frac{p(\boldsymbol{W}^{exp}|M_j,\boldsymbol{\sigma},I)}{p(\boldsymbol{W}^{exp}|M_k,\boldsymbol{\sigma},I)}\ . \tag{7}$$

The first ratio on the r.h.s., the so-called prior odds, is set to unity since we do not want to favor one model over another. We are left with the determination of the second ratio, the so-called Bayes factor. The global likelihood is given by a discrete sum over all expansion orders

$$p(\boldsymbol{W}^{exp}|M_j,\boldsymbol{\sigma},I)=\sum_E p(E|M_j,\boldsymbol{\sigma},I)p(\boldsymbol{W}^{exp}|E,M_j,\boldsymbol{\sigma},I)\ , \tag{8}$$

where $p(E|M_j,\boldsymbol{\sigma},I)$ is set constant because a priori no expansion order is to be preferred. The remaining marginal likelihood is obtained by

$$\begin{aligned} p(\boldsymbol{W}^{exp}|E,M_j,\boldsymbol{\sigma},I) \quad &= \quad \int p(\boldsymbol{W}^{exp}|\omega,\boldsymbol{c},\boldsymbol{x},E,M_j,\boldsymbol{\sigma},I) \\ &\quad\cdot\quad p(\omega,\boldsymbol{x},\boldsymbol{c}|E,M_j,\boldsymbol{\sigma},I)\ \mu(\omega,\boldsymbol{x},\boldsymbol{c})\ d\omega\ d\boldsymbol{c}\ d\boldsymbol{x}\ , \end{aligned} \tag{9}$$

where we followed the ideas brought up by Rodriguez [3] formulating an invariant measure. $\mu(\omega,\boldsymbol{x},\boldsymbol{c})$ in the marginalization integral. This stems from the requirement that the description of a problem should be the same regardless in which coordinate system it is considered. Then the posterior probability for an infinitesimal change in the parameters is given by

$$p(d\omega,d\boldsymbol{x},d\boldsymbol{c}|E,M_j,\boldsymbol{\sigma},I)=p(\omega,\boldsymbol{x},\boldsymbol{c}|E,M_j,\boldsymbol{\sigma},I)\mu(\omega,\boldsymbol{x},\boldsymbol{c})\ d\omega\ d\boldsymbol{c}\ d\boldsymbol{x}\ , \tag{10}$$

with $\mu(\omega,\boldsymbol{x},\boldsymbol{c})$ describing the Riemannian metricof the problem. The latter is given by

$$\mu(\omega,\boldsymbol{x},\boldsymbol{c})=\sqrt{\det\left[\boldsymbol{g}\right]}\ . \tag{11}$$

$\boldsymbol{g}$ is the Fisher information matrix, which is

$$g_{ij}=-\left\langle\frac{\partial^2\ln p(\boldsymbol{W}^{exp}|\boldsymbol{\theta},E,M_j,\boldsymbol{\sigma},I)}{\partial\theta_i\partial\theta_j}\right\rangle \tag{12}$$

and build up by the second derivative of the log-likelihood function with respect to the parameters $\boldsymbol{\theta}=(\omega,\boldsymbol{x},\boldsymbol{c})$.

The prior function decomposes into

$$p(\omega,\boldsymbol{x},\boldsymbol{c}|E,M_j,\boldsymbol{\sigma},I)=p(\omega,\boldsymbol{x}|E,M_j,I)p(\boldsymbol{c}|\boldsymbol{x},E,M_j,\boldsymbol{\sigma},I)\ . \tag{13}$$

While an uninformative prior is assigned to ω and $\boldsymbol{x}$ we employ what we call the Bessel prior for $p(\boldsymbol{c}|\boldsymbol{x}, E, M_j, \boldsymbol{\sigma}, I)$. This just imposes an upper boundary to the choice of possible coefficients with $\boldsymbol{c}^T \boldsymbol{F}^T \boldsymbol{F} \boldsymbol{c} \leq 2||\boldsymbol{W}^{exp}||^2$. It is equivalent to an entropic prior proposed by Rodriguez, where the hyperparameter α has been marginalized by using the Bretthorst approximation that the prior may be taken out of the integral at the maximum of the likelihood.

While the $\boldsymbol{c}$ and ω integrations may be performed analytically the remaining $\boldsymbol{x}$ integration is carried out numerically by computing the prior predictive value[4] with Markov Chain Monte Carlo(MCMC). This is given by

$$
\begin{aligned}
p(\boldsymbol{W}^{exp}|E, M_j, \boldsymbol{\sigma}, I) &= \int p(\boldsymbol{W}^{exp}|\boldsymbol{x}, E, M_j, \boldsymbol{\sigma}, I) p(\boldsymbol{x}|E, M_j, \boldsymbol{\sigma}, I) \; d\mu(\boldsymbol{x}) && (14) \\
&= \int \Lambda(\boldsymbol{x}) \Pi(\boldsymbol{x}) \; d\boldsymbol{x} \; . && (15)
\end{aligned}
$$

$\mu(\boldsymbol{x})$ is the remaining measure in $\boldsymbol{x}$. In (15) we combined the factors of the integral to new functions Λ and Π, where $\Lambda(\boldsymbol{x})$ contains the pure likelihood and some normalization constants, while $\Pi(\boldsymbol{x})$ comprises the prior and the terms stemming from the Riemannian measure. The latter shall be normalized to $\int \Pi(\boldsymbol{x}) \; d\boldsymbol{x} = 1$. The only way to evaluate (15) is to use $\Pi(\boldsymbol{x})$ as a probability density. Since the likelihood descending term $\Lambda(\boldsymbol{x})$ is generally much more structured than $\Pi(\boldsymbol{x})$ the variance will be large and extremely long Markov chains are needed to obtain the desired accuracy. However, one can solve this problem in defining a new function

$$
Z(\beta) = \int \Lambda^{\beta}(\boldsymbol{x}) \Pi(\boldsymbol{x}) \; d\boldsymbol{x} \tag{16}
$$

with $Z(\beta = 0) = 1$ and $Z(\beta = 1)$ as the sought quantity in (15). The derivative with respect to β gives

$$
\begin{aligned}
\frac{\partial \ln Z(\beta)}{\partial \beta} &= \int \ln \Lambda(\boldsymbol{x}) \rho_{\beta}(\boldsymbol{x}) \; d\boldsymbol{x} \\
&= \langle \ln \Lambda(\boldsymbol{x}) \rangle_{\beta} && (17)
\end{aligned}
$$

with

$$
\rho_{\beta}(\boldsymbol{x}) = \frac{\Lambda^{\beta}(\boldsymbol{x}) \Pi(\boldsymbol{x})}{\int \Lambda^{\beta}(\boldsymbol{x}') \Pi(\boldsymbol{x}') \; d\boldsymbol{x}'} \tag{18}
$$

as the new sampling density. If we integrate both sides of 17,

$$
\begin{aligned}
\int_0^1 \langle \ln \Lambda(\boldsymbol{x}) \rangle_{\beta} \; d\beta &= \int_0^1 \frac{\partial \ln Z(\beta)}{\partial \beta} \; d\beta && (19) \\
&= \ln Z(\beta = 1) - \ln Z(\beta = 0) && (20) \\
&= \ln p(\boldsymbol{W}^{exp}|E, M_j, \boldsymbol{\sigma}, I) \; , && (21)
\end{aligned}
$$

we are back where we started in (14). To obtain the marginal likelihood one therefore has to calculate the integral on the l.h.s. in (19) where the expectation value

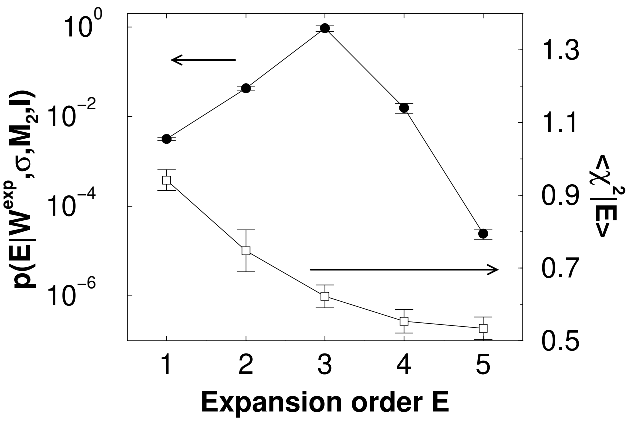

Figure 1. Upper curve: expansion order. Notice the logarithmic scale on the left. Lower curve: misfit of data and model.

$\langle \ln \Lambda(\boldsymbol{x}) \rangle_\beta$ is accessible by Markov Chain Monte Carlotechniques. The proof that this is now feasible may be found in [4].

4. Results

The data which we have used in our calculations are the 153 W7-AS data from the international stellarator energy confinement data base [5]. We have selected only data points with rotational transform of $\iota \approx 1/3$ since the single variable scans which we present further down and which we want to compare with have been performed at this value. Odds ratios obtained from (7) are converted back to probabilities using $\sum_j p(\boldsymbol{W}^{exp}|M_j, \boldsymbol{\sigma}, I) = 1$. The resulting model probabilities are depicted in the last column of table 1. We see that the $\iota \approx 1/3$ W7-AS data are best described by the collisional low-β Connor Taylor model. The high beta models follow in second and third place with much lower probability. The collisionless low-β model is clearly inappropriate to describe the transport physics in W7-AS.

The probability distribution for the expansion order given data and model is now readily obtained:

$$p(E|\boldsymbol{W}^{exp}, M_j, \boldsymbol{\sigma}, I) = \frac{p(E|M_j, \boldsymbol{\sigma}, I)p(\boldsymbol{W}^{exp}|E, M_j, \boldsymbol{\sigma}, I)}{p(\boldsymbol{W}^{exp}|M_j, \boldsymbol{\sigma}, I)} \ . \tag{22}$$

This quantity is displayed for the most probable plasma model, the collisional low beta case, in Fig. 1 as full circles. The error bars indicate Monte Carlo integration uncertainties. The open squares, with associated error bars, depict the misfit (here square of the usual random mean square error) between data and model prediction (5) for each expansion order. We observe quite a typical behavior. Already three terms in the expansion (5) lead to a rapid decrease of the misfit in the order of 35% of the value for $E = 1$. Though it does decrease further with increasing expansion order the probability for a given E decreases rapidly so that contributions of higher expansion orders become very small demonstrating Occam's razor.

Finally we come to determination of observables. Here we want to test the predictive power of the present theory. Such a test is provided by a comparison of measured single variable scans to the predictions from the present theory. The latter is the expectation value of the energy content

$$\langle W \rangle = \frac{\int W \; p(W|\boldsymbol{W}^{exp}, \boldsymbol{v}, M_j, \boldsymbol{\sigma}, I) \; dW}{\int p(W|\boldsymbol{W}^{exp}, \boldsymbol{v}, M_j, \boldsymbol{\sigma}, I) \; dW} \; . \tag{23}$$

The additional condition $\boldsymbol{v}$ specifies the "input data vector" $\boldsymbol{v}^T = (n, B, P, a)$ at which the energy content measurement was performed. $p(W|\boldsymbol{W}^{exp}, \boldsymbol{v}, M_j, \boldsymbol{\sigma}, I)$ is again obtained by marginalization

$$\begin{aligned} p(W|\boldsymbol{W}^{exp}, \boldsymbol{v}, M_j, \boldsymbol{\sigma}, I) = \sum_E \int & p(W|E, \boldsymbol{x}, \boldsymbol{c}, \boldsymbol{v}, M_j, \boldsymbol{\sigma}, I) \\ & \cdot p(E, \boldsymbol{x}, \boldsymbol{c}, \omega|\boldsymbol{W}^{exp}, M_j, \boldsymbol{\sigma}, I) \; d\boldsymbol{x} \; d\boldsymbol{c} \; d\omega \; , \end{aligned} \tag{24}$$

where

$$p(W|E, \boldsymbol{x}, \boldsymbol{c}, \boldsymbol{v}, M_j, \boldsymbol{\sigma}, I) = \delta \left(W - \boldsymbol{g}^T(\boldsymbol{x}, \boldsymbol{v}) \cdot \boldsymbol{c} \right) \; , \tag{25}$$

with $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{v})$ as an E-dimensional vector with elements $f(\hat{n}, \hat{B}, \hat{P}, \hat{a}; \boldsymbol{x})$. The second factor in (24) is just

$$p(E, \boldsymbol{x}, \boldsymbol{c}, \omega|\boldsymbol{W}^{exp}, M_j, \boldsymbol{\sigma}, I) = \frac{p(\boldsymbol{W}^{exp}, E, \boldsymbol{x}, \boldsymbol{c}, \omega|M_j, \boldsymbol{\sigma}, I)}{p(\boldsymbol{W}^{exp}|M_j, \boldsymbol{\sigma}, I)} \tag{26}$$

and we get

$$\langle W \rangle = \frac{\sum_E p(\boldsymbol{W}^{exp}|E, M_j, \boldsymbol{\sigma}, I) \int \boldsymbol{g}^T(\boldsymbol{x}, \boldsymbol{v}) \boldsymbol{c}_{ML} \rho(\boldsymbol{x}) \; d^E \boldsymbol{x}}{\sum_E p(\boldsymbol{W}^{exp}|E, M_j, \boldsymbol{\sigma}, I)} \; . \tag{27}$$

In Fig. 2 we show the result for a density scan obtained on the one hand from the present theory and on the other hand from experiments at W7-AS. Because these data are not included in the W7-AS data base the analysis is based on, this nice agreement shows the predictive power of our approach. The stellarator energy confinement data base is represented by the open circles which are spread all over since they were obtained for various settings of the variables (B, P, a). The full circles in Fig. 2 depict experimental results for a density scan. Here the variables B, P, a are held fixed and only n varies. Representative error bars signify the precision level of these data. But what would be the answer of our present semi-empirical theory based only on the data with open circles? This is shown by the continuous curve along with the confidence range indicated by the gray shaded area. Within the density range of the single variable scan the prediction of the semi-empirical theory runs straight through the data and exhibits clearly a previously supposed density saturation [5,6]. Outside this range the data set $\boldsymbol{W}^{exp}$ is too sparse, which is reflected in the rapidly widening error band and therefore indicates where the extrapolation becomes unreliable. However, it is not a shortcoming of the probabilistic approach but rather an honest outcome that

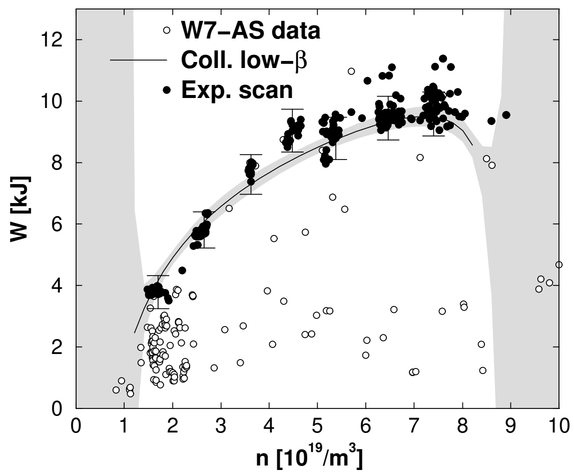

Figure 2. Experimental results of a single variable density scan (full circles) compared to the predictions of the present semi-empirical theory (continuous line, shaded area represents the error) for B = 2.5T, P = 0.45MW, a = 0.176m. The input data (open circles) are shown regardless of additional variations in B, P, a and are therefore spread all over. The histogram accounts for their distribution over the density axis.

an extrapolation beyond the parameter regime supported by the data base is not possible.

5. Conclusion

In summary, form free dimensionally exact energy confinement functions derived from data of the international stellarator data base and treated consistently according to the rules of Bayesian probability theory have identified the collisional low beta Connor Taylor model to be the most probable plasma physics model for W7-AS. Moreover, single variable scans were reproduced in quantitative agreement with experiments. The result of a single variable scan is therefore already hidden in the data obtained for arbitrary variable choices and can be extracted from the latter by a proper data analysis.

References

1. J. Hugill and J. Sheffield *Nucl. Fusion*, **18**, p. 15, 1978.
2. J. W. Connor and J. B. Taylor *Nucl. Fusion*, **17**, p. 1047, 1977.
3. C. Rodriguez, "From euclid to entropy," in *Maximum Entropyand Bayesian Methods*, J. W. T. Grandy, ed., Kluwer Academic, Dordrecht, 1991.
4. W. von der Linden, R. Preuss, and V. Dose, "The prior predictive value." These proceedings.
5. U. Stroth, M. Murakami, R. A. Dory, H. Yamada, S. Okamura, F. Sano, and T. Obiki *Nucl. Fusion*, **36**, p. 1063, 1996.
6. U. Stroth *Plasma Phys. Control. Fusion*, **40**, p. 9, 1998.