EFDA–JET–PR(14)07

M. Irishkin, F. Imbeaux, T. Aniel, J.F. Artaud, Tore Supra Team,
FDA Integrated Tokamak Modelling Task Force Contributors
and JET EFDA contributors

Automated Comparison of Experimental Density Profile Reconstruction with Models

Automated Comparison of Experimental Density Profile Reconstruction with Models

M. Irishkin[1], F. Imbeaux[1], T. Aniel[1], J.F. Artaud[1], Tore Supra Team,
FDA Integrated Tokamak Modelling Task Force Contributors
and JET EFDA contributors*

JET-EFDA, Culham Science Centre, OX14 3DB, Abingdon, UK

[1]*CEA, IRFM, F-13108 Saint-Paul-Lez-Durance, France*
** See annex of F. Romanelli et al, "Overview of JET Results",*
(24th IAEA Fusion Energy Conference, San Diego, USA (2012)).

Preprint of Paper to be submitted for publication in
Plasma Physics and Controlled Fusion

ABSTRACT

A systematic and automated comparison between the physical quantities reconstructed from experimental data and the ones resulted from physical models could be a very useful approach for long pulse tokamak experiments producing a large amount of data. A discrepancy in such comparison brings key information for the scientific exploitation of the experiment since it means either that the pulse deviates from usual models i.e. there is interesting physics to analyze in that pulse or that there are issues with the reconstructed data, coming from the measurements or from the way they have been processed, which have to be corrected.

In this paper we propose to develop an expert system that carries out in an integrated way i) a reconstruction of plasma quantities from the measurements, ii) a prediction of the reconstructed quantities, according to validated expectations / models, iii) and an intelligent comparison of the first two steps. The paper shows the first application of this quite general comparison method to the electron density profile reconstruction in the core plasma.

1. INTRODUCTION

Tokamak experiments produce large quantities of data, 50 Gbytes of data per second is expected for an ITER pulse. A systematic analysis of these data thus requires automated processes. The first step of the analysis is usually to compute a number of plasma physical quantities as a post-processing of the various measurements. This is usually carried out by an automated chain of codes, typically known as the Plasma Reconstruction Chain.

What is up to now quite rarely done in a systematic and automated way is to verify whether the reconstructed plasma characteristics are consistent with models that have been verified for many other pulses. A discrepancy in such a comparison brings key information for the scientific exploitation of the experiment since it means either:

a) That the pulse deviates from usual models i.e. there is interesting physics to analyze in that pulse
b) That there are issues with the reconstructed data, coming from the measurements or from the way they have been processed, which have to be corrected

Conversely, agreement between the reconstructed quantities and the models provide an increased confidence in the plasma reconstruction validity and consistency.

To address these useful functionalities, we propose to develop an expert system carrying out in an integrated way:

1) The Plasma Reconstruction from the measurements, using Bayesian methods
2) The prediction of the reconstructed quantities, according to validated expectations/models
3) An intelligent comparison of the first two steps providing an automated analysis and reporting on events of physical interest during the pulse

In such a procedure, it is interesting to use Bayesian methods for the Plasma Reconstruction since

it provides a rigorous framework for quantifying the error bars on the reconstructed quantities and thus for the comparison to a model's prediction.

Model predictions should be in the general case evaluated using Integrated Modelling tools since they provide a self-consistent and close to the measurements approach. Although the Bayesian methods are used quite frequently in fusion physics [6, 12, 7], the novelty of the overall method lies in the development of relevant comparison criteria between predicted and reconstructed quantities, allowing the automation of the physical analysis. Although the method has been primarily designed for experimental data analysis and validation, it can also reciprocally be used for model validation.

We present here the first application of this quite general comparison method to the electron density profile reconstruction in the core plasma. We chose two parameters to compare: a line-average density and a profile peaking factor (ratio of the central to the average density). Using the Bayesian inference, probability histograms were constructed for these quantities. The analysis was performed for about 20 Tore Supra and 14 JET L-mode shots. Two different models for predicting the density profile peaking have been used in the comparison, providing conclusions on their accuracy for this experimental dataset. Comparison criteria have been defined which allow an automated determination of the agreement quality. They also provide a way, when the agreement between the model and the reconstructed quantities is not acceptable, to discriminate between problems in the experimental data and limits of validity of the models.

Section II presents the reconstruction of plasma profiles using Bayesian methods. Section III discusses the models used to predict the electron density profile. Section IV describes the implementation of the whole procedure under the European Task Force on Integrated Tokamak Modelling (ITM-TF) framework. Section V describes how the comparison criteria have been established. Section VI presents the application of the analysis methods to the Tore Supra and JET datasets for model validation. Conclusions and perspectives for future work are presented in section VII.

2. RECONSTRUCTION OF ELECTRON DENSITY PROFILES

2.1. RADIAL PROFILE PARAMETERIZATION

To reconstruct plasma electron density profile we must first parameterize it. We can do it by using simple 3rd order splines on the grid of normalized toroidal flux coordinate (ρ). In total we will have N+2 parameters, where N is the number of grid points for ρ, which will consist of N values of the profile in grid points and first derivatives in the center ($\rho = 0$) and at the edge ($\rho = 1$) of the plasma. Having a toroidal geometry we fix the first derivative in the center to be 0. Thus we have N+1 parameters to describe the radial profile at a given time slice.

2.2. PRINCIPLES OF BAYESIAN ANALYSIS

Bayesian analysis allows obtaining the probability distribution of the parameters of the splines defining the radial profiles, using constraints from the experimental measurements and taking into

account their uncertainties. We remind here the main principles of the Bayesian analysis.

The formula for Bayes' theorem for conditional probabilities can be written as follows:

$$P(A|B) = \frac{P(A,B) \cdot P(A)}{P(B)} \tag{2.1}$$

Where A, B are some events and P(•) is the probability that the event occurs.

We may also notice that marginal probability P(B) is equal to the integral (or a sum in a discrete case) of the joint probability P(A, B) over A, which in turn is equal to the product of P(B|A) and P(A):

$$\begin{aligned} P(B) &= \int P(A, B) \cdot dA \\ &= \int P(B|A) \cdot P(A) \cdot dA \end{aligned} \tag{2.2}$$

Thus we see that the denominator of the formula (2.1) is equal to the integral of its nominator and plays role of normalization constant. We will skip it in our further discussion and will simply write

$$P(A|B) \propto P(B|A) \cdot P(A) \tag{2.3}$$

We may also think of A as of the profiles parameters, whereas B corresponds to the experimental data points (measurements):

$$P(\text{Params}|\text{Measurements}) \propto P(\text{Measurements}|\text{Params}) \cdot P(\text{Params}) \tag{2.4}$$

Suppose that we have several data points which are independent from each other, then

$$\begin{aligned} P(\text{Params}|D_1,D_2,...,D_n) &\propto P(D_1,D_2,...,D_n|\text{Params}) \cdot P(\text{Params}) \\ &\propto \prod_{i=1}^{n} P(D_i|\text{Params}) \cdot P(\text{Params}) \end{aligned} \tag{2.5}$$

We can do the same with the parameters (assuming their internal independence):

$$P(\{\text{Params}_j\}_{j=1}^{k} \mid \{D_i\}_{j=1}^{n}) \propto \prod_{i=1}^{n} P(D_i \mid \{\text{Params}_j\}_{j=1}^{k}) \cdot \prod_{i=1}^{k} P(\text{Params}_j) \tag{2.6}$$

The first term on the right hand side of the equation (2.6) is called likelihood. Generally speaking it is our data acquisition model which expresses a probability of obtaining the measurements we have given a set of parameters. The second term on the right hand side is a prior probability. It incorporates our prior knowledge about the distribution of the parameters (if we do not have it then we can choose so called non-informative prior). The term on the left hand side of the equation (2.6) is called posterior probability and gives a posterior probability of parameters given the measurements.

In fact the posterior probability is a trade-off between the likelihood and the prior: if we have a strong prior belief then we need much experimental evidence to change it and vice versa if we have no prior knowledge then we can rely on whatever the experiment gives us.
Although the equation (2.6) looks simple, its numerical implementation is not so evident. The thing is that all we have is just random number generators which can make a sampling from a given prior distribution. Based on these samples we can calculate likelihood and thus we will have a sample which represents posterior probability distribution. To avoid huge number of iterations we must have an appropriate way for sampling from the posterior distribution. The class of algorithms for doing this is called Markov Chain Monte Carlo methods [1]. These algorithms samples Markov chains which converge to the posterior distributions defined by the equation (2.6).

2.3 DIAGNOSTICS INVOLVED IN THE PROFILE RECONSTRUCTION

Two main diagnostics are used in the Profile Reconstruction, namely interferometry and Thomson Scattering. Since the quality of the measurements is different in Tore Supra and JET, different relative weights between diagnostics are applied in the Profile Reconstruction (related to the likelihood in the section below).

Interferometry [9] provides measurements of electron density integrated along lines of sight, while Thomson Scattering provides local measurements [4].

2.4 ASSUMPTIONS OF THE BAYESIAN ANALYSIS

To perform a Bayesian analysis we must make some assumptions about the prior distributions of parameters and the likelihood. In the chapter IIa we stated that there are N+1 parameters of spline interpolation. In addition to this, we introduced one more parameter for Tore Supra: a recalibration parameter for Thomson scattering measurements as the diagnostic has a large uncertainty in its absolute calibration for density measurement. Thus we have N+2 parameters in total and the goal is to find the posterior distributions of these parameters defined by the equation (2.6).

We assume that the prior probability distributions for the N+1 parameters of spline interpolation are uniform, since we do not have a strong a priori knowledge of the solution. The choice of the boundaries of the uniform prior distribution has been done in two different ways for Tore Supra and JET. For Tore Supra, we use a prior distribution centered on the model prediction because we do not have directly from the measurements a good first guess for the local value of density: interferometry does not give us local measurements and Thomson Scattering data contains a large uncertainty on its calibration coefficient. Further analysis showed the uniform prior distribution with boundaries [0.5*model prediction; 1.5*model prediction] do not restrict significantly the possible parameter space and at the same time gives a reasonable prior. Concerning JET, as we have a number of relatively precise Thomson scattering measurements, we can determine boundaries of the prior distribution based on these measurements.

The prior distribution for the recalibration parameter for Tore Supra was chosen to be uniform

with wide boundaries (from zero to ten) as we have no prior information on it.

The likelihood used in our analysis was normal as we assumed the normality of the errors distribution for all the diagnostics. The error bar for the interferometry data for Tore Supra was assumed to be constant and equal to $5*10^{17}$ $1/m^3$, for JET – 20% of a measurement value. Note that this anomalously high value for the JET interferometer error bar has been chosen to reduce the weight of some lines of sight which apparently have large systematic errors and could not be reconciled otherwise with the measurement of the other ones and the HRTS measurements. The error bar for Thomson scattering was taken to be 20% for Tore Supra and 10%*(central_density/ measurement_density)^2 for JET High Resolution Thomson Scattering. We chose not constant error bar for the HRTS JET data because the examination of the experimental measurements showed that the error bar at plasma edge tends to be higher than the one in the center. Then using Monte Carlo Markov Chain algorithms implemented in a Python module pymc (Pymc) which make sampling from the posterior distribution according to the equation (2.6) we get the samples for all the parameters. Based on the samples obtained, we can calculate statistics on any quantities of our interest (for example, density peaking and line-average density). The statistics also gives us 95% highest probability density intervals which are used in the comparison with the density profile predictions.

3. PREDICTION OF THE ELECTRON DENSITY PROFILE

The novelty of this work consists in the automated and systematic comparison of the Plasma reconstruction results with models. In this application the models consist in a simple scaling expression for one parameter (density profile peaking). Integrated Modelling is a powerful tool for prediction and model validation and has the advantage of enforcing the consistency of the simulated parameters, a common point with Bayesian analysis. Moreover it allows estimating quantities that are difficult to measure directly, while they may be input to some models, e.g. the safety factor. Therefore, even if the first application presented here is relatively simple, we use METIS (Artaud, 2008), a fast integrated modeling transport code, in view of future integrated analysis of many physics quantities simultaneously. The speed and robustness of this simplified integrated modeling tool are key advantages in view of automated analysis of a large amount of data.

In this application, METIS was used in the following conditions: i) current diffusion is predicted ii) electron and ion temperatures are predicted from the heat transport equations, using a simple diffusion coefficient model with fixed radial shape and renormalized to an L-mode scaling iii) the electron density is calculated as follows.

In L mode the density profile $n_e(t,x)$ in METIS is defined with three parameters: the central density ($n_{e,0}$), the edge density ($n_{e,a}$) and the peaking factor ($\nu_n(t)+1$):

$$n_e(t,x) = n_{e,0}(t) - n_{e,a}(t)) \cdot (1-x^2)^{\nu_n(t)} + n_{e,a}(t) \qquad (3.1)$$

where $v_n(t) = \frac{n_{e,0}(t)}{\bar{n}_e(t)} - 1$, t is the time, x is a normalized minor radius. For each magnetic surface, the minor radius is defined as $a = \frac{R_{max} - R_{min}}{2}$, where R_{max} ($resp.$ R_{min}), is the maximum (resp. minimum) major radius of that surface. The minor radius is then normalized to its value at the LCFS a_m, so that $x = \frac{a}{a_m}$,.

The edge density and peaking factors are calculated from scaling expressions. The average density is prescribed from experimental measurements, and the central density is determined so that the calculated profile matches this prescription. Because of this prescription from experiment, the predictive part of METIS electron density profile estimation lies essentially in the scaling expressions for the peaking factor and for the edge density.

The edge density is calculated using one of the following scaling laws for L mode:

- in X point configuration in H mode [10]:

$$n_{e,a} = C_n \cdot n_l^{-2} \text{ with } C_n = 5 \cdot 10^{-21} - 6.7 \cdot 10^{-24} \cdot T_{e,a} \text{ and } C_n = 10^{-21} \tag{3.5}$$

- in X point configuration in L mode [10]:

$$n_{e,a} = 0.00236 \cdot n_l^{-1.08} \cdot K_{ref}^{-1.11} \cdot B_{ref}^{-0.78} \tag{3.6}$$

- with a toroidal limiter (F. Clairet, private communication):

$$n_{e,a} = 1.0 \cdot 10^{-21} \cdot n_l^{-2} \cdot q_{edge} \cdot R_{ref} \tag{3.7}$$

- with a poloidal limiter [15]:

$$n_{e,a} = 5.0 \cdot 10^{-21} \cdot n^{-2} \tag{3.8}$$

For all Tore Supra shots the edge density is calculated using the scaling law (3.7). For JET, the choice of scaling law for JET shots depends on the mode and configuration of the plasma at a particular time.

Two simple models for the peaking factor in L mode have been designed and compared to experiments in this work. The two models are empirical and describe different parametric dependences which have been observed in experiments.

Model 1 is related to the observation that, in the Saturated Ohmic Confinement regime (SOC), the density peaking factor is inversely proportional to the density, as seen in many tokamaks such as ASDEX [3, 13] Tore Supra [8] and more recently C-MOD [11]. The model assumes a dependence on the ratio of the saturation density to the line-integrated density and is a priori valid only in plasmas that have the same type of dominant turbulence as the SOC regime (namely Ion Temperature Gradient dominant modes).

$$v_n(t) = 0.5\,\frac{n_{sat}}{\bar{n}} \tag{3.2}$$

with

$$n_{sat} = 0.06 \cdot 10^{20} \cdot \frac{I_p}{10^6} \cdot R \cdot \frac{\sqrt{A_{main}}}{K \cdot a^{5/2}} \tag{3.3}$$

Model 2 assumes a functional dependence of peaking factor on the plasma internal self-inductance and was derived based on experimental data from JET [14]. We analysed the Figure 7 of [14] where the peaking factor is plotted versus li and we fitted the data to a straight line.

$$v_n(t) = \frac{4}{3} \cdot l_i - \frac{3}{4} \tag{3.4}$$

Note that the internal inductance range of the experimental data reported in the original figure of [14] is relatively small, i.e. 1 to 1.35. The linear fit that we derived from the figure will be applied to a larger range in this work.

METIS solves the current diffusion and uses the resulting li to evaluate the density peaking factor.

4. IMPLEMENTATION UNDER THE EUROPEAN INTEGRATED MODELLING FRAMEWORK

Methods and techniques that are discussed in this article are implemented within the framework of the European Integrated Tokamak Modeling Task Force (ITM-TF) (ITM). The two main motivations for this choice are i) the ITM-TF Data Model is tokamak-generic, thus the method can then be applied to any experiment ii) the link with Integrated Modelling tools, namely equilibrium identification codes and METIS for this particular application. Moreover, the ITM-TF Framework provides also methods for accessing data from various experiments, Tore Supra and JET in this application.
The analysis is done in the following steps:

1) we first run the METIS code to get predictions of electron density profiles.
2) we run an equilibrium identification code (Equinox [5] in this application) to have a description of plasma equilibrium
3) map experimental measurements on the equilibrium
4) run Bayesian analysis
5) Run comparison of predicted (step 1) and reconstructed (step 4) density profiles

5. DEVELOPMENT OF AUTOMATED COMPARISON CRITERIA FOR TORE SUPRA AND JET

A dataset of 20 Tore Supra and 14 JET L-mode shots has been selected to establish automated comparison criteria, covering a wide range of plasma parameters (see tables 1 and 2). We chose one time slice for every pulse to conduct the analysis, in a phase in which the density profile is

stationary (i.e. does not evolve over several characteristic transport times; for the density profile we use the energy confinement time as characteristic transport time which is of the order of 0.03s for Tore Supra and 0.3s for JET), except for the JET shots featuring an H mode transition and for which the time of analysis is just before the L to H mode transition.

To compare the reconstructed and the predicted profiles, we defined three levels of agreement quality:

- "acceptable": the predicted profile lies within or very close to the 95% Highest Probability Density (HPD) interval provided by the Bayesian analysis. An "acceptable" agreement provides a sort of mutual validation of the models used in the simulation and the Bayesian profile reconstruction, since it would be quite fortuitous that both would be "wrong" in the same way;
- "needs investigation": the predicted profile is marginally outside the 95% HPD interval, so that the comparison has to be further checked by human intervention
- "not acceptable": the simulation profile is strongly outside the 95% HPD interval.

The quantities used for the automated comparison are:

- relative squared profile discrepancy: a sum of squares of ratios between predicted profile and the closest boundary of 95% HPD interval (0 if the predicted profile is within the HPD interval boundaries)
- number of lines of sight (LOS) within the 95% HPD range: from MCMC sampling we have statistics for the density integrals along interferometry lines of sight and thus we can compare it to the experimental data
- relative peaking discrepancy : a ratio between minimal distance of predicted peaking factor and one of its 95% HPD boundaries
- relative integral discrepancy: a ratio between minimal distance of METIS line-average integral and one of its 95% HPD boundaries

The exact value of the criteria used to classify each case among the 3 agreement levels have been decided by considering the Tore Supra data set only (see Table 3). Then they have been applied as such to the JET dataset, with only one minor modification related to the second quantity since this one is diagnostic-dependent. They also provided a satisfying classification of the various pulses among the 3 agreement levels for the JET case, which emphasizes the tokamak-generic character of the analysis.

On Figures 1-3 examples of each level of agreement quality are shown. Figure 1 illustrates a case where the predicted density profile perfectly lies within 95% highest probability range and therefore the analysis concludes that the predictions and experimental data are in agreement. On Figure 2 is shown a comparison which raises "need investigation" flag. Figure 3 shows an example of a "not acceptable" agreement.

The automated analysis can go beyond the level of agreement and detect possible problems in the Plasma Reconstruction or question the validity of the models used in the prediction. In case the agreement is not considered "acceptable", two questions arise: whether the reconstruction was correct and whether the profile peaking model used in the simulation was accurate. To answer the former question we should look primarily at two parameters: number of lines of sight within the 95% HPD range (if it is less than 5 for Tore Supra and 3 for JET then likely there is an inconsistency in the experimental data) and convergence of Markov chains (if they have not converged then again likely there is an inconsistency in the experimental data). If we see that the experimental results show no evidence of inconsistency then the discrepancy between experiment and the simulation is attributed to the models used in the simulation (i.e. the problem is considered as a model validation issue). This last step of the automated interpretation of the results allows using the method for model validation, as discussed below.

6. APPLICATION OF THE METHOD FOR MODEL VALIDATION

In this section, we illustrate how the automated comparison between reconstructed profiles and their simulation counterpart can be used for model validation.

Two models have been successively applied in METIS, as described in section III: the former treats the peaking as a function of a ratio of saturation density over line-integrated density and the latter considers the peaking to be a function of internal inductance.

6.1 MODEL VALIDATION FOR TORE SUPRA

Figure 4 shows the results of the automated comparison for Tore Supra, using model 1 (which assumes a dependence between peaking factor and the ration of saturation density over average density). All but one cases of "not acceptable agreement" or "needs investigation" point to an issue in the Plasma Reconstruction, not in the model. The model shows acceptable agreement for 93% of the dataset, which validates the model in this range of parameters. This could be found at first glance surprising, since this model has been applied to a wide range of plasma conditions (mostly non ohmic plasmas) that are a priori very different from its original "design domain"; the model was inspired from the density peaking parametric dependence in ohmic shots above the saturation density. This may suggest that the dominant instability in the analysed dataset is of the same kind as in the SOC regime, i.e. Ion Temperature Gradient modes [14].

Conversely, model 2 shows an acceptable agreement in only one case, i.e. only 7% of the cases for which the Plasma Reconstruction was correct. This shows that model 2 is not adequate to capture the dependences of this dataset.

To complete / confirm these conclusions of the automated analysis, more traditional scatter plots can be made and submitted to human judgement. This is feasible for this application because the comparison focuses essentially on single, scalar parameters (the density profile peaking factor). We may also be interested in plotting peaking factor of METIS versus peaking factor obtained in

the experimental data analysis. Figures 6 and 7 show these plots for two models. We can see that the experimental peaking factor is in good agreement with the METIS one for model 1. Whereas the model 2 tends to overestimate the peaking factor.

We may be also interested in the testing the model 2 assumption of the dependence of peaking factor on internal self-inductance. Figure 8 shows the METIS result in blue points and experimental result in orange points. As we see, there is no evidence of linear dependence between these two quantities.

Based on the figures above we can confirm that for the Tore Supra dataset the model 1 describes the peaking factor better than the model 2 and that there is no dependence of the peaking factor on internal inductance. It does not look surprising as the second model was empirically derived for JET data and a smaller range of li and thus the extrapolation of the model for a wider range of li and another tokamak seems incorrect.

6.2 MODEL VALIDATION FOR JET

To perform the analysis, 14 JET shots were chosen. They were L-mode shots taken during their plateau phase or L-mode parts of hybrid shots analysed just before their L to H mode transition. We applied the criteria discussed in the section V with only one change: we set up the threshold for number of fitted interferometry lines of sight to 3 instead of 5 as the total number of lines of sight taken into account during the analysis for JET is less than the one for Tore Supra (7 for JET and 10 for Tore Supra).

Figures 9 and 10 show the summary of the comparison results for model 1 and model 2.

We may notice that there are only two shots that have data inconsistency issue and the discrepancies in the other cases come from the integrated simulation part.

Figures 11 and 12 show comparison of two peaking factors (METIS vs experiment).

Based on the results of the comparison we conclude that both models show marginal agreement (33% for model 1 and 42% for model 2).

Figure 13 shows the dependence of peaking factor on internal self-inductance (for METIS model 2). Based on our limited dataset, no clear dependence of the density peaking on li is observed.

For the JET case, the conclusions on the validity of the models are less obvious and both models show only marginal agreement. Model 2 shows definitely a better agreement for JET than for Tore Supra, which to some extent was expected since it was empirically determined from a JET dataset (with Carbon wall). Nonetheless the statistics of acceptable agreement for the model remains relatively small (below 50%), showing that the linear fitting of the Figure 7 data from [14] that we did is not reliable for the wider range of internal inductances that we have been studying here. Our JET dataset is mixing discharges with the Carbon wall and the ITER like wall, but the statistics of validity of both models do not show a strong difference between these two subsets. Meanwhile, to derive stronger conclusion further investigation is needed for an extended shot database.

The summary of the analysis for JET and Tore Supra data is presented in the Table 4.

CONCLUSION

The present work was devoted to developing a first application of the automated comparison method discussed in the Introduction. The first simple application of the developed methods was shown for 20 Tore Supra and 14 JET L-mode shots for electron density profile. Such method provides simultaneous validation of experimental data and model validation, with a qualification and quantification of the agreement between a model and the reconstructed profiles from measurements. It provides also statistics of the agreement quality, thus contributing to establish the domain of validity of a given model. The application of the method was carried out for two density peaking factor models in the METIS simulation code. The analysis showed that the first model that assumes dependence of peaking factor on the ration between saturation density and average density works quite well for Tore Supra data (93% of acceptable agreement) while for JET data both models show marginal agreement (below 50% of acceptable agreement). No dependence of peaking factor on internal self-inductance was observed for Tore Supra data and the second model (internal inductance dependence) is clearly not adequate for this dataset. The implementation of the methods is tokamak-generic as was performed using the ITM-TF Framework.

In the future we would like to extend the automated comparison method, which is quite generic, to other applications. The next foreseen application is to reconstruct electron and ion temperatures profiles and perform validation of heat and toroidal momentum transport models. From a technical and operational point of view, a systematic application of the method to the full duration of a pulse will require parallelization of the analysis over many time slices and needs to be implemented in the future.

ACKNOWLEDGEMENTS

The authors would like to thank Rainer Fischer for fruitful discussions on technical implementation details concerning Bayesian analysis.

This work, supported by the European Communities under the contract of Association between EURATOM and CEA, was carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

[1]. Andrieu, C., & al. (2003). An introduction to MCMC for machine learning. Machine Learning, **50,** 5-43.

[2]. Artaud, J. (2008). METIS User's Guide. PHY/NTT-2008.001.

[3]. Becker, G. (1990). Analysis of energy and particle transport and density profile peaking in the improved Ohmic confinement regime. Nuclear Fusion **30**, 11.

[4]. Beurskens, M., & al. (2011). H-mode pedestal scaling in DIII-D, ASDEX Upgrade, and JET. Physics of Plasmas **18**, 056120.

[5]. Blum, J., & al. (2012). Reconstruction of the equilibrium of the plasma in a Tokamak and identification of the current density profile in real time. Journal of Computational Physics **231**, 960.

[6]. Dinklage, A., & al. (2012). Integrated diagnostics design. Fusion Science and Technology **62**, 419-427.

[7]. Fischer, R., & al. (2006). Flexible and reliable profile estimation using exponential splines. Bayesian Inference and Maximum Entropy Methods in Science and Engineering, AIP conference proceedings **872**, 296-303.

[8]. Garbet, X., & al. (1992). Turbulence and energy confinement in Tore Supra Ohmic discharges. Nuclear Fusion **32**, 2147.

[9]. Gil, C., & al. (2009). Diagnostic systems on Tore Supra. Fusion Science and Technology, **56**.

[10]. ITM. (n.d.). Retrieved from http://portal.efda-itm.eu/itm/portal

[11]. Porter, G., & al. (1999). Analysis of separatrix plasma parameters using local and multi-machine databases. Journal of Nuclear Materials 266-269, 917. Pymc. (n.d.). Retrieved from http://pymc-devs.github.io/pymc/index.html

[12]. Rice, J., & al. (2012). Ohmic energy confinement saturation and core toroidal rotation reversal in Alcator C-Mod plasmas. Physics of Plasmas **19**, 056106.

[13]. van Milligen, B., & al. (2011). Integrated data analysis at TJ-II: The density profile . Review of Scientific Instruments **82**, 073503.

[14]. Wagner, F., & al. (1993). Transport in toroidal devices - the experimentalist's view. Plasma Physics and Controlled Fusion **35**, 1321.

[15]. Weisen, H., & al. (2005). Collisionality and shear dependences of density peaking in JET and extrapolation to ITER. Nuclear Fusion **45**, L1-L4.

[16]. Wesson, J. (2004). Tokamaks. In Tokamaks, 3rd edition (p. 737). Clarendon University Press.

Shot	Time, s	Toroidal field, T	Plasma current, MA	Internal self-inductance	Central density, $x10^{19}$ m^{-3}	Saturation density, $x10^{19}$ m^{-3}	NBI power, MW	Ohmic heating power, MW
84541	43.5	1.73	1.57	0.83	2.11	2.77	0.00	0.98
84543	43.5	1.73	1.57	0.84	2.09	2.79	0.00	0.98
84545	43.5	1.73	1.57	0.83	2.10	2.79	0.00	0.93
84792	43.5	1.73	1.58	0.82	2.03	2.97	0.00	0.85
84795	43.5	1.72	1.58	0.82	2.01	2.99	0.00	0.91
84796	43.5	1.72	1.58	0.82	2.02	2.98	0.00	0.91
82541	52.0	2.64	2.46	0.93	5.04	4.48	0.00	2.13
82536	52.9	2.69	2.46	0.97	4.33	4.22	0.00	1.15
82120	47.9	2.20	1.97	0.57	3.63	3.29	0.00	1.07
75225	47.5	2.03	1.69	0.68	3.85	2.92	17.30	0.44
77895	43.0	2.69	1.47	0.84	1.89	2.71	0.00	0.67
77914	45.0	2.32	2.39	0.84	2.02	4.15	0.00	1.32
77922	45.4	2.32	2.25	0.91	2.15	3.90	0.00	0.93
77933	45.8	2.34	2.65	0.85	2.36	4.20	0.00	1.55

Table 1: Summary of the shots characteristics for the Tore Supra dataset: central density is taken from the METIS code, saturation density is calculated by the METIS code as per formula 3.3, internal self-inductance deduced from magnetic measurements using the TPROF code; toroidal field, plasma current, LH, ICRH and Ohmic heating powers are taken from the Tore Supra database.

Shot	Time, s	Toroidal field, T	Plasma current, MA	Internal self-inductance	Central density, $x10^{19}$ m^{-3}	Saturation density, $x10^{19}$ m^{-3}	LH power, MW	ICRH power, MW	Ohmic heating power, MW
47666	11.0	3.78	1.50	1.12	5.23	7.03	4.42	0.00	0.61
47659	5.0	3.78	1.30	1.21	5.14	6.07	2.93	0.00	0.98
47658	4.0	3.78	1.30	1.22	4.90	6.10	2.45	0.00	1.22
47657	10.0	3.78	1.50	1.15	5.49	6.98	0.89	0.00	1.13
47096	4.0	3.77	1.00	1.33	4.19	4.67	0.01	0.00	0.66
47092	10.0	3.77	1.00	1.35	4.51	4.67	0.01	1.03	0.73
47654	12.1	3.78	0.91	1.42	3.71	4.21	4.02	0.00	0.24
46982	11.1	3.84	0.61	1.67	2.58	2.83	1.91	0.00	0.43
47011	10.0	3.84	1.00	1.29	4.66	4.62	0.00	0.99	0.92
47067	10.0	3.75	0.71	1.57	2.91	3.25	3.24	0.00	0.33
45583	9.5	2.30	0.81	1.34	2.30	3.70	1.74	0.00	0.27
45530	22.0	2.78	0.49	1.91	1.85	2.25	2.64	0.00	0.00
45584	9.5	2.30	0.80	1.30	3.89	3.68	1.73	0.00	0.41
45581	9.5	2.30	0.80	1.30	3.63	3.67	0.62	0.00	0.45
45552	9.5	3.85	1.00	1.45	3.53	4.58	0.62	0.00	0.49
45564	9.5	3.84	1.01	1.53	2.58	4.63	1.74	0.00	0.13
45175	9.5	3.85	1.00	1.19	7.17	5.57	1.47	0.00	0.92
48317	10.2	3.76	0.71	1.55	3.66	3.54	3.95	1.99	0.02
45525	22.0	2.78	0.50	1.85	1.86	2.26	2.65	0.00	0.04
47160	3.3	3.84	1.00	1.26	3.77	4.65	0.00	0.00	1.06

Table 2: Summary of the shots characteristics for the JET dataset: toroidal field, plasma current, central density, Ohmic heating power are taken from the METIS code, saturation density is calculated by the METIS code as per formula 3.3, internal self-inductance deduced from magnetic measurements using the EFIT code; NBI heating power is taken from the JET database. The Pulse No's: 75225-77933 are the ones with carbon wall and the Pulse No's: 782120-84796 are the ones with ITER-like wall.

Agreement between Plasma Reconstruction and simulation)		**Quantity**	**Value**
Acceptable		Relative profile error	< 0.05
	AND	Peaking	$< 1.5\%$
	AND	Integral	$< 1.5\%$
Needs investigation		Relative profile error	$[0.05; 0.23]$
	OR	Peaking	$\geq 1.5\%$
	OR	Integral	$\geq 1.5\%$
Not acceptable		Relative profile error	> 0.23
	OR	LOS	< 5 for Tore Supra < 3 for JET

Table 3: Summary of the analysis criteria: the first column indicates the conclusion on the agreement between Plasma Reconstruction and the simulation; columns 2 and 3 indicate the criteria to establish the conclusion of the first column

JET		
Number of shots	14	14
Reconstruction issues	*2*	*2*
Simulation issues	*8*	*7*
Acceptable fit	*4*	*5*
Both models show marginal agreement (33% and 42% success respectively)		

Table 4: Summary of the analysis.

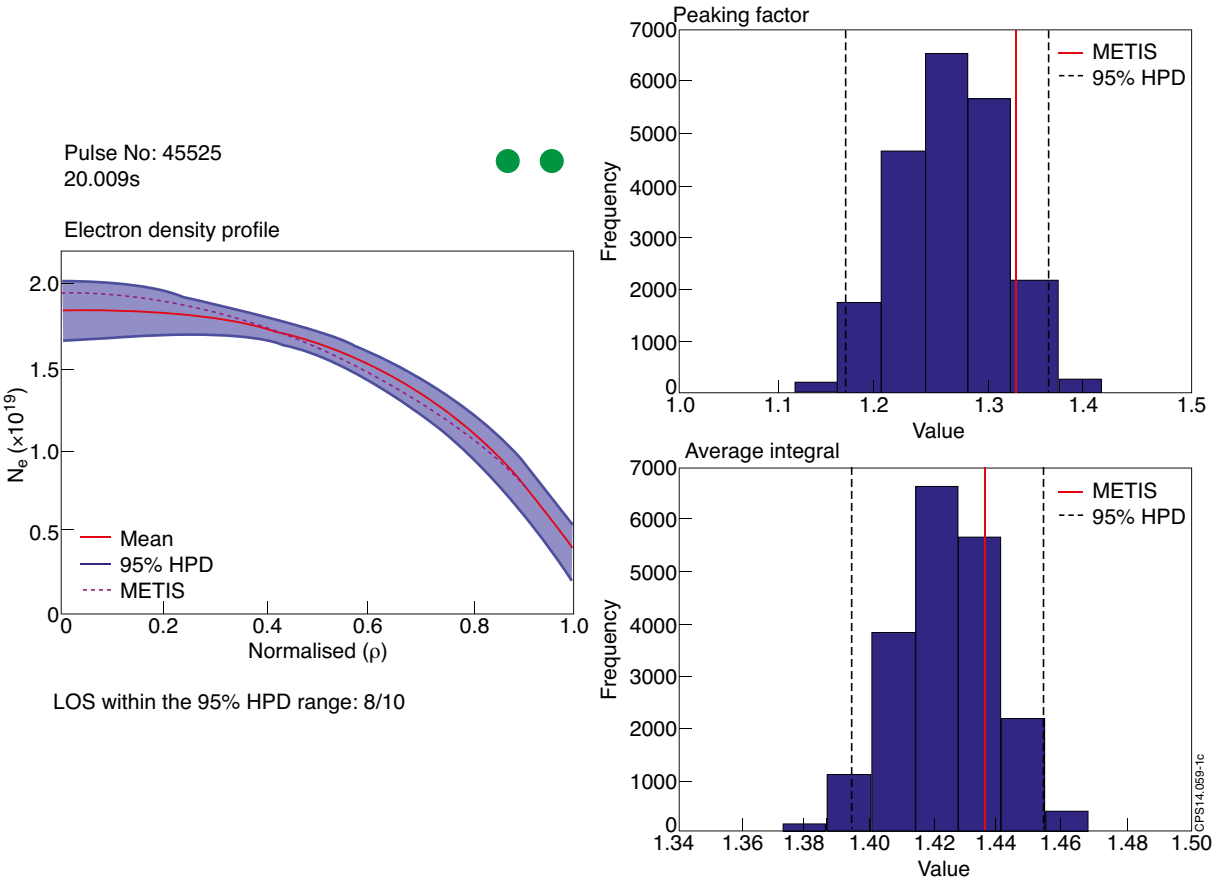

Figure 1: Example of "acceptable" agreement: electron density profile for the Pulse No: 45525. Left plot shows the mean (read) profiles and 95% HPD interval (blue area) obtained by Bayesian analysis and METIS result (dashed magenta line); right plots show distribution for the peaking factor and average integral, their 95% HPD interval (range between dashed lines) and the METIS values (red lines).

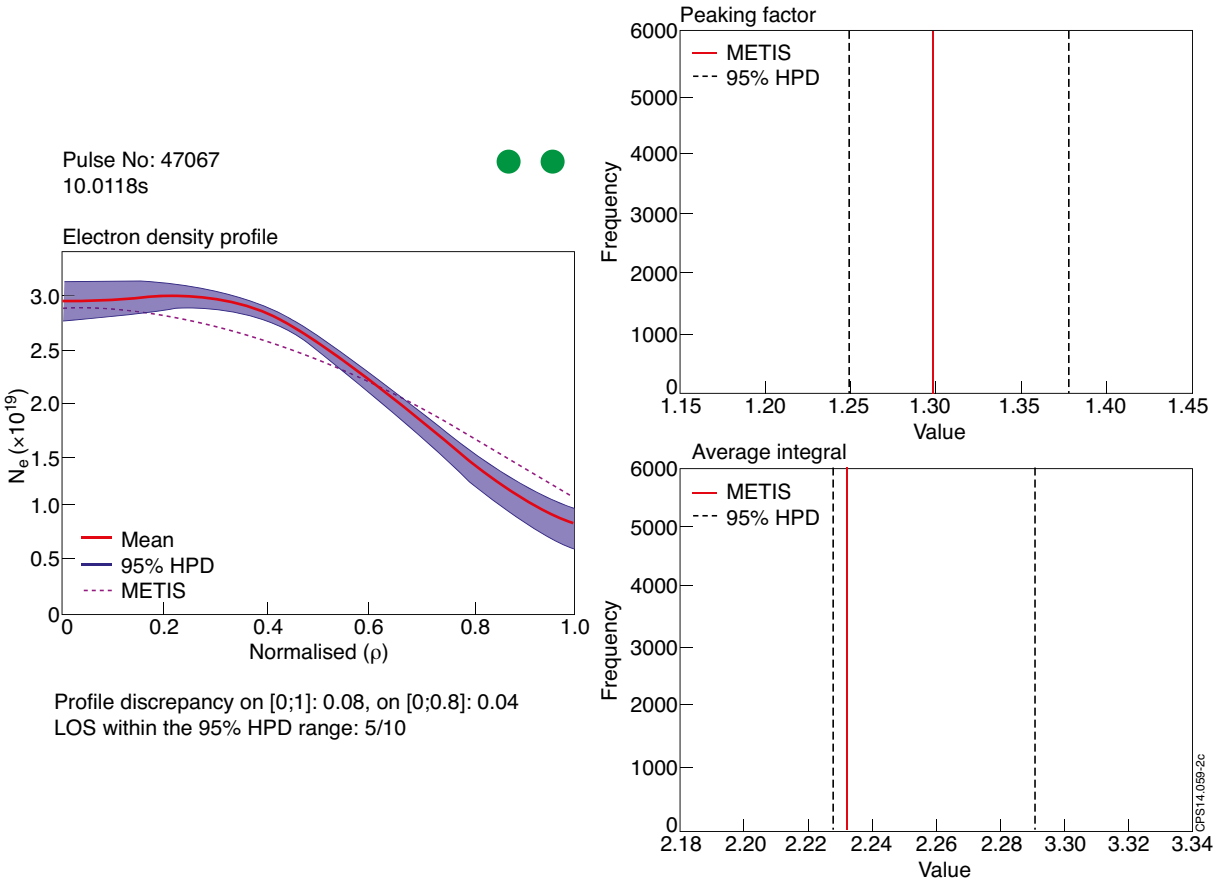

Figure 2: Example of "needs investigation" agreement: electron density profile for the Pulse No: 47067. Left plot shows the mean (read) profiles and 95% HPD interval (blue area) obtained by Bayesian analysis and METIS result (dashed magenta line); right plots show distribution for the peaking factor and average integral, their 95% HPD interval (range between dashed lines) and the METIS values (red lines).

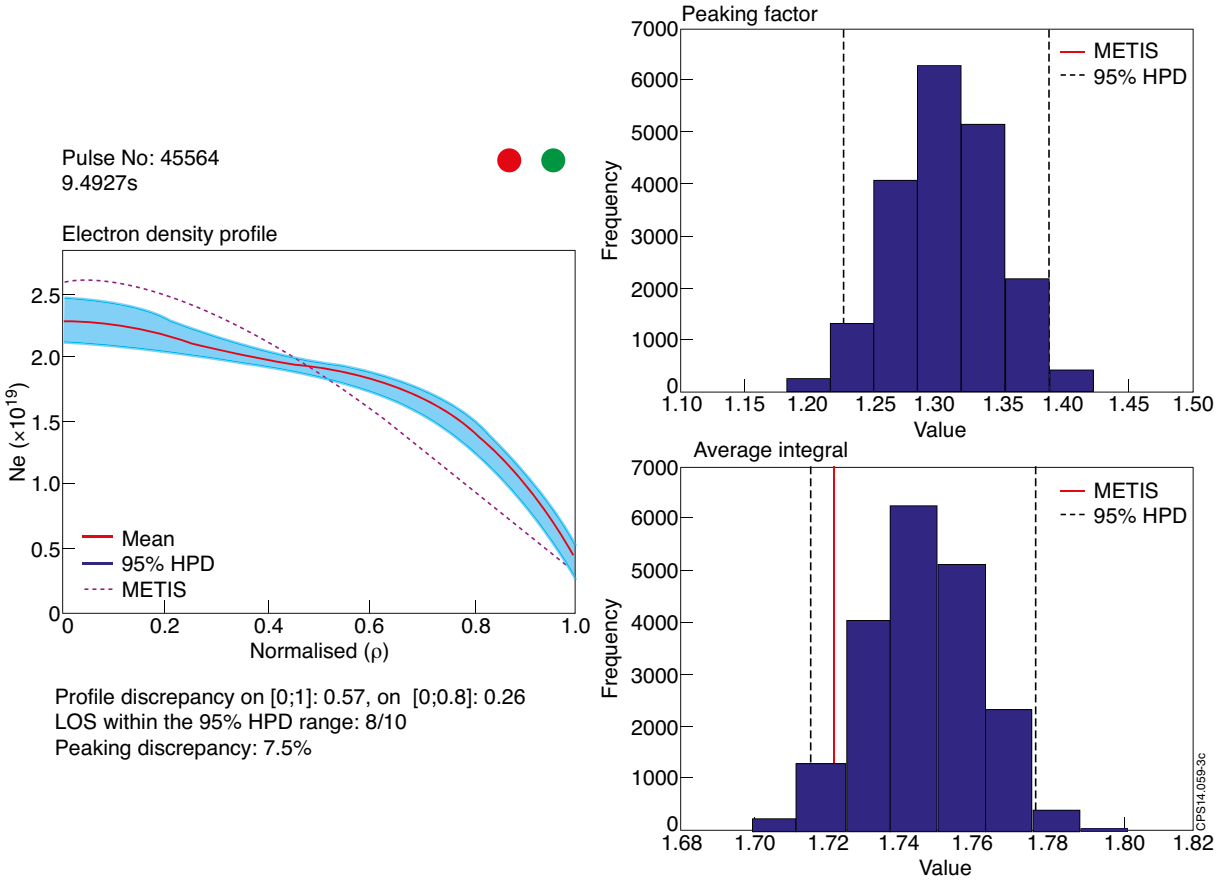

Figure 3: Example of "not acceptable" agreement: electron density profile for the Pulse No: 45564. Left plot shows the mean (read) profiles and 95% HPD interval (blue area) obtained by Bayesian analysis and METIS result (dashed magenta line); right plots show distribution for the peaking factor and average integral, their 95% HPD interval (range between dashed lines) and the METIS values (red lines).

Pulse No:	Conclusion	Issue
47666	OK	
47659	OK	
47658	OK	
47657	OK	
47096	OK	
47092	WARN	Data inconsistency issue
47654	OK	
46982	WARN	Simulation needs investigation
47011	OK	
47067	WARN	Data inconsistency issue
45583	WARN	Data inconsistency issue
45530	OK	
45584	OK	
45581	OK	
45552	OK	
45564	NO	Data inconsistency issue
45175	OK	
48317	WARN	Data inconsistency issue
45525	OK	
47160	OK	

CPS14.059-4c

Figure 4: Results of the automated comparison for Tore Supra, using model 1. OK mean "acceptable agreement", WARN means "needs investigation", NO means "not acceptable agreement".

Pulse No:	Conclusion	Issue
47666	NO	Simulation needs investigation
47659	OK	
47658	NO	Simulation needs investigation
47657	WARN	Simulation needs investigation
47096	NO	Simulation needs investigation
47092	NO	Data inconsistency issue
47654	NO	Simulation needs investigation
46982	NO	Simulation needs investigation
47011	NO	Simulation needs investigation
47067	NO	Data inconsistency issue
45583	NO	Data inconsistency issue
45530	NO	Simulation needs investigation
45584	NO	Simulation needs investigation
45581	NO	Simulation needs investigation
45552	NO	Simulation needs investigation
45564	NO	Data inconsistency issue
45175	NO	Simulation needs investigation
48317	NO	Data inconsistency issue
45525	NO	Simulation needs investigation
47160	NO	Simulation needs investigation

CPS14.059-5c

Figure 5: Results of the automated comparison for Tore Supra, using model 2. OK mean "acceptable agreement", WARN means "needs investigation", NO means "not acceptable agreement".

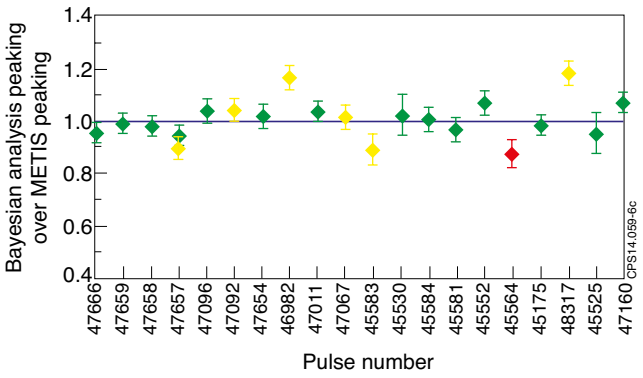

Figure 6: Ratio of peaking factors: experiment over METIS model 1. Green points – "acceptable" agreement; yellow points – "needs investigation" agreement; red points – "not acceptable" agreement.

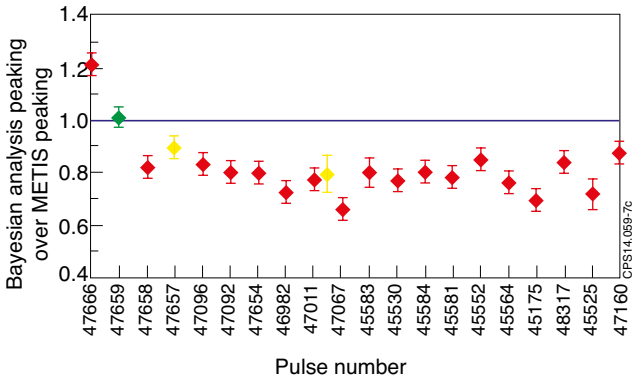

Figure 7: Ratio of peaking factors: experiment over METIS model 2. Green points – "acceptable" agreement; yellow points – "needs investigation" agreement; red points – "not acceptable" agreement.

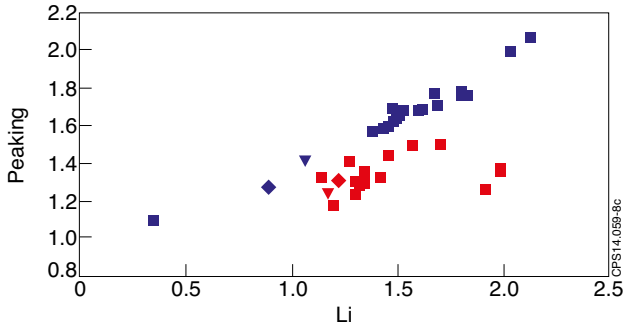

Figure 8: Peaking factor vs internal self-inductance. Blue – METIS result; orange – experimental result. Diamond point means that there is an acceptable fit between the METIS profile and the experimental one, triangles refers to the fit "needs investigation" and squares are used for the profiles with not acceptable fit.

Pulse No:	Conclusion	Issue
84541	NO	Data inconsistency issue
84543	OK	
84545	OK	
84792	OK	
84795	WARN	Simulation needs investigation
84796	WARN	Simulation needs investigation
82541	WARN	Simulation needs investigation
82536	OK	
82120	NO	Data inconsistency issue
75225	WARN	Simulation needs investigation
77895	NO	Simulation needs investigation
77914	NO	Simulation needs investigation
77922	WARN	Simulation needs investigation
77933	WARN	Simulation needs investigation

CPS14.059-9c

Figure 9: Conclusion on every Pulse No: based on application of criteria described in section 5a for METIS model 1 for JET L-mode Pulse No:s.

Pulse No:	Conclusion	Issue
84541	NO	Data inconsistency issue
84543	OK	
84545	OK	
84792	OK	
84795	WARN	Simulation needs investigation
84796	WARN	Simulation needs investigation
82541	NO	Simulation needs investigation
82536	WARN	Simulation needs investigation
82120	NO	Data inconsistency issue
75225	WARN	Simulation needs investigation
77895	WARN	Simulation needs investigation
77914	NO	Simulation needs investigation
77922	OK	
77933	OK	

CPS14.059-10c

Figure 10: Conclusion on every Pulse No: based on application of criteria described in section 5a for METIS model 2 for JET L-mode Pulse No:s.

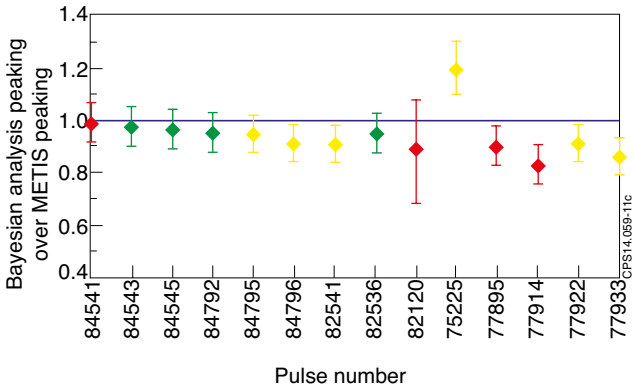

Figure 11: Ratio of peaking factors: experiment over METIS model 1. Green points – "acceptable" agreement; yellow points – "needs investigation" agreement; red points – "not acceptable" agreement.

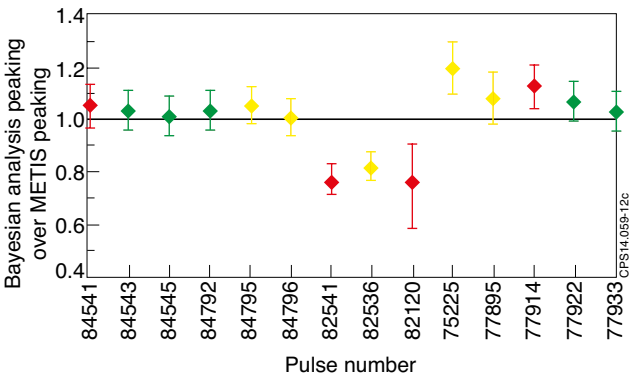

Figure 12: Ratio of peaking factors: experiment over METIS model 2. Green points – "acceptable" agreement; yellow points – "needs investigation" agreement; red points – "not acceptable" agreement.

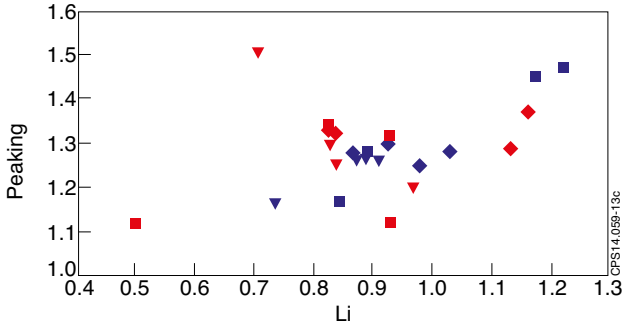

Figure 13: Peaking factor versus internal self-inductance. Blue – METIS result; orange – experimental result. Diamond point means that there is an acceptable fit between the METIS profile and the experimental one, triangles refers to the fit "needs investigation" and squares are used for the profiles with not acceptable fit.