Scaling of density peaking in H–mode plasmas based on a combined database of AUG and JET observations

C. Angioni[1], H. Weisen[2], O.J.W.F. Kardaun[1], M. Maslov[2], A. Zabolotsky[2], C. Fuchs[1], L. Garzotti[3], C. Giroud[3], B. Kurzan[1], P. Mantica[4], A.G. Peeters[1], J. Stober[1], ASDEX Upgrade Team and contributors to the EFDA-JET workprogramme*

[1] *Max-Planck Institut für Plasmaphysik, IPP–Euratom Association, D-85748 Garching bei München, Germany*

[2] *Centre de Recherches en Physique des Plasmas, Association Euratom–Confédération Suisse, EPFL, 1015 Lausanne, Switzerland*

[3] *UKAEA-Fusion, United Kingdom Atomic Energy Authority, Abingdon, UK*

[4] *Istituto di Fisica del Plasma, Associazione Euratom–ENEA–CNR, Milano, Italy*

* *See Appendix of J.Pamela et al, Fusion Energy 2004 (Proc. 20th Int. Conf. Vilamoura, 2004) IAEA, Vienna (2004).*

Abstract. For the first time, scalings for density peaking in tokamaks are obtained from a database consisting of observations from two devices, ASDEX Upgrade and JET. The investigation relies on an inversion method for the interferometer signals which grants consistent reconstructions despite differences in interferometer geometries. By combining observations from these devices, correlations between physics parameters investigated for their role in determining density peaking is reduced. Multiple regression analyses show that in the combined database collisionality is the most relevant parameter. The particle source provided by neutral beam injection provides a contribution to the peaking, which, although not negligible, is not large enough to explain the whole observed variation of density peaking. The device size, introduced as an alias for possible systematic differences between the devices not captured by the regression parameters, is found to only play a small role in regressions which include collisionality. Device size becomes relevant in scalings which exclude collisionality and include the ratio of the density to the Greenwald density limit. This indicates that density peaking is more likely to be a function of collisionality rather than of the fraction of the density limit. All the scalings which include collisionality in the regression variables predict a peaked density profile for the ITER standard scenario.

1. Introduction and motivations

The ability to extrapolate from present plasma scenarios to ITER partly depends on whether the same shape of the density profile will be realized also in burning plasmas. The shape of the density profile has important consequences on both the plasma confinement and the plasma stability. In a burning plasma, with the same temperature profiles and the same volume averaged density, a peaked density profile produces a larger amount of fusion power and bootstrap current with respect to a flat profile. On the other hand, too peaked a density profile may have negative consequences on both the MHD stability and central accumulation of heavy impurities. Recent experimental results in ASDEX Upgrade (AUG) and JET H–mode plasmas indicate that the density peaking is correlated with the plasma collisionality [1, 2, 3]. This observation may lead to the prediction that density profiles in the ITER standard scenario will not be flat, as usually assumed [4], but peaked, since ITER collisionality is expected to be as low as the lowest collisionalities achieved in present devices. However, as long as results from a single device are considered, collisionality is correlated with other plasma parameters, in particular the Greenwald fraction, the normalized ion Larmor radius ρ_* and the fuelling provided by the beams. Here, we extend the approach which has been undertaken in [3] with a database consisting exclusively of JET observations and, for the first time, we present empirical scalings for the density peaking using a database of observations from two devices of different size. By adding dimensional size to an otherwise non-dimensional set of regression parameters, a valuable test for the consistency and completeness of the set is obtained. In particular, this method allows us to compare the relevances of collisionality and Grennwald fraction. Multiple regression analyses confirm that in the combined database of AUG and JET observations, collisionality is the most relevant parameter.

The database is described in Section 2, while the regression variables are defined in Section 3. We show that by combining observations from different devices, while some correlations are indeed reduced, some additional uncertainties are introduced. The way we have adopted to overcome the limitations encountered is discussed in Section 4. Bivariate correlations are shown in Section 5, while a multiple regression analysis is presented in Section 6. Section 7 proposes scalings for density peaking and discusses the projections for ITER. Finally, in Section 8 the main conclusions of this work are summarized.

2. The combined database of AUG and JET observations

The combined database is composed of 277 JET observations and 343 AUG observations of ELMy H–mode plasmas, of which 99 JET plasmas and 312 AUG plasmas are auxiliarly heated by neutral beam injection (NBI) only, while 33 JET plasmas and 9 AUG plasmas are heated by ion cyclotron resonance heating (ICRH) only. All the other plasmas are heated by a combination of the two heating systems. The ranges

of engineering parameters covered by the two devices in the combined database are presented in Table 1. Shot numbers in AUG are between 8000 and 17000, whereas in JET are between 42000 and 64000.

3. Definition of the regression variables

Our purpose is to express the density peaking in the form of a multivariable regression in terms of dimensionless plasma parameters. The physics plasma parameters ρ_*, ν and β, are considered with the following definitions,

$$\rho_* = 4.37\ 10^{-3}\ (m_{\rm eff}\ \langle T_e \rangle)^{0.5}\ /B_T/a,$$

$$\nu_{\rm eff} = 0.2\ \langle n_e \rangle\ R_{geo}/\langle T_e \rangle^2,$$

$$\beta = 4.02\ 10^{-3}\ \langle p \rangle / B_T^2.$$

In this formulae, densities are in 10^{19} m^{-3}, temperatures in keV, magnetic fields in Tesla, the total plasma pressure p in keV$\times 10^{19}$ m^{-3}, $m_{\rm eff}$ in A.M.U. and the symbol $\langle\ \rangle$ denotes a volume average. Geometrical plasma parameters like q_{95}, the edge triangularity δ are also considered. Given the small variation of aspect ratio and elongation in AUG and JET, these two parameters are not included. Note that in AUG and JET these parameters are very close to those of ITER.

Moreover, the plasma size (the major radius R_{geo}), despite being dimensional, is also included in part of the analysis as a device label, in order to check its significance and relevance in the regressions. Since a dimensionless parameter like density peaking should not be expected to depend on a dimensional parameter like R_{geo}, the parameter R_{geo} is introduced on purpose in the analysis as an intruder, in order to test whether the other set of variables provide a consistent and sufficiently complete regression model. If, as a consequence of the addition of R_{geo} to a set of regression variables, a strong dependence on R_{geo} is found in the regression, with a related significant reduction of the root mean square error (RMSE), this provides the indication that the original set of variables does not provide a consistent and complete set of scaling parameters for the density peaking.

The analysis also considers the Greenwald fraction $F_{GR} = \overline{n}_{e\ lin\ 20}\ \pi a^2/I_p$, where $\overline{n}_{e\ lin\ 20}$ is the line average density in 10^{20} m^{-3} and I_p is the plasma current in MA. We remind that the Greenwald fraction is not a dimensionless parameter of a fully ionised plasma, and therefore could be considered out of place in the present analysis. Nevertheless, since collisionality and Greenwald fraction extrapolate in opposite directions for ITER, we believe that an important goal remains the experimental assessment of whether density peaking is a function of collisionality or of the fraction of the density limit.

Finally dimensionless variables to describe the particle source are considered. The particle source provided by wall neutrals is neglected in the present analysis, in agreement with the result that its contribution can be ignored in the particle balance

equation in the confinement region [5]. Instead we have included parameters to describe the neutral beam fuelling. The neutral beam heating and particle source profiles are computed for all the observations in the database by the steady–state Fokker–Planck PENCIL code [6, 7, 8] for JET data and the Monte Carlo FAFNER code [9] for AUG data. Both these codes take into account the beam injection geometry and the beam energies, as well as the specific plasma equilibrium and plasma profiles. Two different parameters are considered to describe the beam particle source. The first is directly the peaking of the profile of the electron source caused by the neutral beams. The second provides more precisely a quantification of the contribution to the density peaking provided by the beam particle source. Namely, by recasting the general steady state diffusive law for the particle flux in the form

$$-\frac{1}{n}\frac{dn}{dr} = \frac{\Gamma}{nD} - \frac{V}{D}, \tag{1}$$

the local slope of the density profile at the left hand side is expressed as the sum of the particle source contribution and the particle pinch contribution. Since the diffusion coefficient D is difficult to measure and unavailable in the dataset, whereas the heat conductivity χ is more routinely available from the power balance analysis, we parametrize the source contribution due to the beams in the form,

$$\Gamma^*_{NBI} \doteq \frac{R\Gamma_{NBI}}{nD}\frac{D}{\chi} = 2\,T\,\frac{\Gamma_{NBI}}{Q_{NBI}}\,\frac{Q_{NBI}}{Q_{TOT}}\,\left|\frac{R}{T}\frac{dT}{dr}\right| \tag{2}$$

where in the right hand side the average heat conductivity $\chi = (\chi_i + \chi_e)/2$ is evaluated by $\chi = Q_{TOT}/(2\,n\,dT/dr)$, assuming no large difference between electron and ion temperatures. In Eq. (2) Γ_{NBI} is the particle flux produced by the beams, Q_{NBI} is the heat flux produced by the beams and Q_{TOT} is the total heat flux. All the terms on the right hand side of Eq. (2) can be evaluated using the parameters available in the database, like beam deposition profiles (or beam energy), total and beam heating powers, and the temperature profile peaking. If the parameter χ/D is a weak function of the plasma parameters and can be considered costant over the full dataset (this is the strongest assumption in this procedure), then Γ^*_{NBI} is an appropriate scaling parameter describing the beam source contribution to the density peaking. We note that the value of χ/D does not need to be assumed a priori in our procedure. As it will be explained later, the average value of χ/D consistent with the full dataset is actually determined by the regression procedure itself. In this work, all particle and heating fluxes have been computed at $r/a = 0.5$, assuming that all coupled RF power is fully absorbed inside that radius, consistently with the central position of the RF resonance for the plasmas selected in the database. Here, as it is motivated in the next section, the peaking factor $n_e(\rho_{pol} = 0.2)\,/\,\langle n_e\rangle_{Vol}$ is used as response variable. For consistency, the normalized logarithmic temperature gradient $R/L_T = -R/T\,(dT/dr)$ in Eq. (2) is replaced by the temperature peaking factor $T_e(\rho_{pol} = 0.2)\,/\,\langle T_e\rangle_{Vol}$. A linear regression over a subset of well diagnosed 150 AUG and 200 JET temperature profiles reveals that the normalized logarithmic temperature gradient at mid–radius

can be expressed by $R/L_T = 3.23\,(T_2/\langle T\rangle - 0.37)$ with RMSE normalized to the mean value of 9.72% ($T_2 = T(\rho_{pol} = 0.2)$). Therefore, for consistency with the definition adopted for the density peaking, in the statistical analysis over the full AUG and JET database, we have replaced the logarithmic temperature gradient in Eq. (2) with the quantity $(T_2/\langle T\rangle - 0.37)$. We note that, by using Eq. (2) as definition of the beam source parameter in a linear regression of the logarithmic density gradient R/L_n, the regression coefficient can be actually regarded as an empirical estimate of the ratio χ/D [3]. In our approach, in which the regressed variable is the density peaking $n_{e2}/\langle n_e\rangle$, the regression coefficient of Γ^*_{NBI} can still be interpreted as the average value of χ/D provided that it is renormalized to the appropriate factors relating $T_{e2}/\langle T_e\rangle$ and $n_{e2}/\langle n_e\rangle$ to the corresponding logarithmic gradients R/L_{Te} and R/L_n at mid–radius ($n_{e2} = n_e(\rho_{pol} = 0.2)$).

Table 2 shows the mean values, the standard deviations and the full ranges of variation, namely minimum and maximum values, of all the plasma parameters considered in the multivariable regression analyses. Values of the combined database, as well as values of the subsets of AUG and JET data separately, are quoted.

4. Definition of the response variable

The main challenge encountered in combining the observations from AUG and JET is to obtain a consistent definition and measurement of the response variable, namely the density peaking (as well as of the regression variables). Different diagnostics of the density profiles may have systematic errors which do not involve large uncertainties in the ITER prediction when one device is considered alone, but which may cause extremely large uncertainties in the ITER predictions when combined with diagnostics from another device having systematic errors in different directions. Such systematic errors may introduce spurious parametric dependences, in particular in the ρ_* dependence. As an example to this point, let us assume that systematically JET density profiles are measured slightly more peaked than they actually are and AUG density profiles slightly less peaked than they actually are. Of course as long as observations of a single device are considered these small systematic errors are reflected in a small overestimate or underestimate of the ITER peaking. If the measurements from the two devices were considered together in this form, they would artificially increase the ρ_* dependence of the peaking, leading to projections for ITER which would be much more peaked than they should actually be.

To overcome this problem, we have adopted a procedure to obtain values of density peaking from both AUG and JET derived with exactly the same method. Such a procedure starts with the observation that density profile measurements in JET show a better agreement between the Thomson scattering diagnostics and the interferometer line integrals than in AUG. On this basis, we have assumed that JET profiles obtained by the singular value decomposition inversion (SVD–I) method [2, 10], which uses basis functions extracted from the LIDAR Thomson scattering profiles, were more

reliable than AUG measurements based on simple inversion of the interferometer. The consistency between the use of SVD–I profiles or directly the LIDAR Thomson scattering profiles on JET has been verified in Ref. [3].

The steps followed in the adopted procedure can be described as follows.

- For a fixed AUG equilibrium, we have computed the line integrals along the 5 lines of sight of the AUG interferometer of all the JET SVD–I profiles of the database. Fig. 1 shows the chosen equilibrium (AUG shot ♯20661 at 6.0 s) and the geometry of the AUG DCN interferometer. The mapping of each JET density profile onto the AUG equilibrium has been performed by keeping the same relationship between the density and the normalized poloidal flux.
- Still considering the same AUG equilibrium, we have reinverted the computed AUG interferometer line integrals of the JET SVD–I profiles. This inversion was obtained by expressing the JET profiles as a linear combination of a fixed set of 5 basis functions describing the profile shape (as many as the number of lines of sight of the AUG interferometer). Although there is a degree of arbitrariness in the choice of the set of 5 basis functions the results are not sensitive to the choice (within reason), as we have ascertained using different sets of basis functions. Among the different sets of basis functions, we have chosen the one which at the same time accurately describes the original JET profiles and provides a set of sufficiently regular monotonic density profiles in the inversion of the measured AUG line integrals. We also note that the basis functions do not need to be orthogonal. Fig. 2(a) shows the set of 5 basis functions adopted. By the same inversion method, all the AUG profiles are reconstructed from the measurements of line integrals of the AUG interferometer.
- We also considered different definitions of density peaking and chose the definition of density peaking which was most strongly constrained by the type of measurements we had available. We found that, independently of the choice of basis functions, the ratio of the central value to the volume average is more strongly constrained by the condition of matching the line integrals of the interferometer than, e.g., the ratio between two local values. For this reason, we have adopted the definition of density peaking $\mathrm{pk}_{ne} = n_e(\rho_{pol} = 0.2) \,/\, \langle n_e \rangle_{Vol}$ throughout this work. We underline, that the aim of this procedure is not to obtain a precise reconstruction of the exact shape of the density profiles (which would be impossible with only 5 line integrals), but to extract a single parameter, the density peaking factor, with the best possible accuracy. This aim was achieved as demonstrated in Fig. 2(b), where we have plotted the values of density peaking obtained from the reinverted JET profiles as a function of the values of density peaking computed directly on the original SVD-I JET profiles. We find that the RMSE between the density peaking of the original JET profiles and the density peaking of the profiles obtained by our inversion procedure, based on the remapping on the AUG interferometer geometry, is as small as 0.018. The mean value and standard deviation of the density peaking

of the original profiles of the JET subset are 1.454 and 0.193, very close to those obtained at the end of the re–inversion procedure (see 1st column of Table 2). A stronger difference is found instead for the density peaking values of AUG, as compared with those obtained from the AUG density profiles reconstructed with a different inversion method, namely that routinely applied on AUG, based on the Abel inversion method taking into account the edge profile measurements of a lithium beam diagnostics. AUG density profiles are on average significantly more peaked when they are reconstructed with our procedure (mean value 1.37, standard deviation 0.201) as compared to the case when they are reconstructed with the standard AUG inversion method (mean value 1.31, standard deviation 0.175). and therefore with a different inversion method with respect to that used for the JET subset. The adoption for the AUG subset of the density peaking values obtained from the standard AUG density profile reconstruction procedure would have led in all regressions to a much stronger dependence of density peaking on plasma size than that found adopting a set of density peaking values for the two devices reconstructed consistently with the same inversion method as done in the present work.

In this way a set of values of density peaking is obtained for the full set of profiles of AUG and JET we have considered. These values of density peaking have been reconstructed with exactly the same inversion method, starting from the values of the line integrals of the AUG interferometer, directly measured in the case of the AUG densities, or computed by the described remapping in the case of the JET densities. As already mentioned, such a procedure has been applied in order to reduce the effects of possible different systematic errors in the measurements of density peaking in the two devices.

5. Bivariate correlations

Fig. 3 shows a selection of scatter plots among plasma parameters which turn out to have the largest correlations. The corresponding correlation coefficients are quoted in the figure (color online), in black (first value) for the combined database, in red (second value) for only AUG data, in blue (third value) for only JET data (values in smaller fonts indicate the correlation coefficients for the subset in which $P_{NBI} \,/\, P_{TOT} > 0.7$, namely when NBI heating is dominant). Both the peaking of the beam particle source, namely $S_{NBI}(\rho_{pol} = 0.2) \,/\, \langle S_{NBI} \rangle_{Vol}$, where S_{NBI} is the source of electrons due to the neutral beams by ionization and charge exchange per unit volume and time, and the beam source parameter Γ^*_{NBI} have been considered. We observe that while correlations with ρ_* are strongly reduced by combining observations from the two devices, the correlation between $\nu_{\rm eff}$ and the Greenwald fraction remains rather large. Collisionality turns out to be the parameter which has the largest bivariate correlation with density peaking in the combined dataset. However, both the Greenwald fraction and the beam particle source parameter Γ^*_{NBI} show very large correlations with density peaking. A very

strong correlation coefficient (-0.91) between collisionality and the beam particle source parameter in AUG plasmas heated with NBI only is found. This correlation is reduced by considering plasmas from the two devices. At the same value of collisionality, JET plasmas have a particle source parameter Γ^*_{NBI} which is on average smaller than AUG plasmas. Finally, we mention that similarly to the JET subset of observations [3], also in the combined database a very low correlation is found between density peaking and parameters describing the peaking of the current density profile. In particular, correlation coefficients as low as -0.24 and 0.12 are found in the combined database between the density peaking and l_i and q_{95} respectively.

6. Multivariable statistical analysis

Let us consider the vector of observations of the regressed variable Y and N vectors of regression variables X_j. A linear or logarithmic multivariable regression expresses Y in the forms

$$Y = c + \sum_j \hat{a}_j X_j \qquad \text{or} \qquad Y = C\,\Pi_j X_j^{\hat{a}_j},$$

where $\hat{a}_j$ are the estimated regression coefficients. According to [11], we define the following parameter to describe the statistical relevance StR_j of the parameter X_j in the linear regression for Y, $StR_j = \hat{a}_j \times \mathrm{STD}(X_j)$, where with STD we denote the usual standard deviation. Analogously, for a logarithmic regression, $StR_j = \hat{a}_j \times \mathrm{STD}(log(X_j))$. In this way StR_j estimates the variation of the (logarithm of the) regressed variable for one standard deviation variation of the (logarithm of the) regression variable X_j, keeping fixed all the other regression variables. The larger StR_j is, the higher is the relevance of the variable X_j in the regression for Y. Besides this parameter, we have also considered an estimate of the statistical significance of each regression variable, $StS_j = \hat{a}_j/\mathrm{STD}(a_j)$, where $\mathrm{STD}(a_j)$ is one standard deviation, namely 66.67% confidence interval, of the estimated regression coefficient $\hat{a}_j$. In the present work, all the regressions are performed with a robust fit algorithm, which uses iteratively reweighted least squares with the bisquare weighting function.

We note that in the study of density peaking, linear regressions are considered more appropriate since they conserve the physical form of Eq. (1). By including Γ^*_{NBI} in the regression model, the regression coefficient of Γ^*_{NBI} is directly connected with the value of χ/D, as discussed before, while the other terms in the scaling describe the main dependences of the pinch term. Our statistical analysis has been performed applying both linear and logarithmic regression forms. The same results have been obtained in the two cases, leading to the same conclusions. Since linear regressions have a more direct physical interpretation, in the present paper we report only the results obtained with linear regressions.

In Tables 3 and 4 the statistical significance and the statistical relevance obtained in each linear regression for a set of plasma parameters are shown. Different regression

models are considered. Regressions which include the collisionality and exclude the Greenwald fraction F_{GR}, and which include the Greenwald fraction and exclude the collisionality, as well as regressions which include both these plasma parameters, are considered. Moreover, for comparison, models which, besides the dimensionless variables, include as well a device label (namely the geometrical major radius) are analysed. As discussed above, the exercise of including or excluding the major radius allows one to quantify its influence on the statistical significance and relevance of the other variables.

A set of considerations and conclusions can be drawn.

- In all the regression models which include collisionality, collisionality is found to be the parameter with the largest statistical relevance. Furthermore, it is always found to be highly significant.
- In nested models which include or exclude the major radius, it is found that the inclusion of the major radius provides a larger reduction of the RMSE in regressions using F_{GR} rather than in regressions using collisionality. The statistical significance and relevance of R_{geo} turns out to be the largest when ν_{eff} is excluded, whereas is the smallest when F_{GR} is excluded.
- In regression models which include collisionality and exclude the major radius, ρ_* is found to have negligible statistical significance and negligible statistical relevance. On the other hand, in regression models which include the Greenwald fraction and exclude collisionality, the device size is found to play a more important role, through a larger statistical relevance of ρ_* and/or the major radius. The signs of the regression coefficients indicate that at the same Greenwald fraction, the density peaking is larger in JET than in AUG, namely at fixed Greenwald fraction, the density peaking increases with increasing size of the device. On the other hand, in regressions which include collisionality, the device size plays a negligible role. We note that in regressions which include ν_{eff}, ρ_* and R_{geo}, the signs and magnitude of the regression coefficients in front of ρ_* and R_{geo} are such that the effects of these two parameters balance each other, as indicated by the very small residual statistical relevance of ρ_* which is found when R_{geo} is excluded. These results provide the important indication that the density peaking is more likely to be a function of collisionality rather than of the Greenwald fraction. Finally, in regression models which include both collisionality and the Greenwald fraction, density peaking is found to increase with increasing Greenwald fraction at fixed collisionality. We note however that the statistical significance of the Greenwald fraction in this case is small.
- In regression models which exclude collisionality, the beam particle source parameter is found to have a larger statistical relevance. The contribution of the beam particle source can be quantified to not exceed 30% in regressions which include collisionality, which is in agreement with the estimate of 20% presented in [5] for JET alone. Instead, this estimate appears to be smaller than that of a

recent work [12] based on a set of transport simulations of JET H–mode plasmas. However it should be added that the results presented in the latter reference are determined by the specific assumptions on the ratio χ/D made in a transport model. This is generally the case for any analysis based on transport simulations of stationary regimes. On the other hand, as already mentioned, in regressions using the Greenwald fraction, the contribution of the beam particle source is found to be larger, around 40%. We underline that in no regression the contribution of the beam particle source is found to be able to describe alone the observed variation of the density peaking. Otherwise a much stronger statistical relevance (and significance) of the beam particle source would have been found in the regressions. From the physics standpoint, this indicates that it is not possible that the observed variation of density peaking in the database is caused exclusively by effects of the beam fuelling. As mentioned in Sec. 3, the regression coefficient in front of the beam source parameter Γ^*_{NBI} can be used to evaluate the average value of the ratio χ/D over the full set of data. In regressions including collisionality this is found to be close to 1.5, in agreement with previous estimates obtained on the set of JET data alone [3], whereas it is found to be larger (around 2.5) in regressions which include the Greenwald fraction and exclude collisionality, and which is closer to the values assumed by the transport model adopted in [12].

- In all regression models q_{95} is found to have small significance and relevance. A similar small role in the regressions is found for the parameter l_i, both in case l_i is added or in case l_i is used at the place of q_{95}.

Similar conclusions are drawn in the case that logarithmic rather than linear regressions are made, and in the case the beam source parameter Γ^*_{NBI} is replaced by the peaking of the beam particle source, namely by $S_{NBI}(\rho_{pol} = 0.2) \,/\, \langle S_{NBI} \rangle_{Vol}$. By this replacement, it is found that the statistical significance (as well as the statistical relevance) of the peaking of the beam particle source is smaller than that of the beam source parameter Γ^*_{NBI}. For this reason, as well as for the more direct physical interpretation, in the next section, scalings for density peaking are proposed with the inclusion of the beam source parameter Γ^*_{NBI} in the regression variables. We mention however that analogous scalings and very close projections for ITER are obtained in case Γ^*_{NBI} is replaced by the peaking of the beam particle source.

Moreover, analogous results are found on the subset of data with dominant NBI heating. Finally, if the weight of ICRH points is increased in the regression, the RMSE of regressions which exclude collisionality increase more than those of regressions which include collisionality. For instance, in case points with ICRH only are given the same weight as the total subset of the points with beam heating, it is found that the RMSE is 0.174 for a regression which includes collisionality and excludes the Greenwald fraction, while it is 0.201 for a regression which exclude collisionality and include both the Greenwald fraction and the major radius. From the last column of Table 3, we observe that increasing the weight of the ICRH points implies a smaller increase of the RMSE,

from 0.114 to 0.174, in the regression which includes $\nu_{\rm eff}$, with respect to the increase of the RMSE, from 0.122 to 0.201, in the case of the regression which includes F_{GR} and R_{geo}. This shows that points with ICRH only are better described in regressions which include collisionality rather than in regressions which include the Greenwald fraction. This is confirmed by the comparison of the RMSE of the subset of points with ICRH only in regressions obtained over the full set of data. In regressions which use the collisionality, the RMSE of the points with ICRH only is regularly below 0.08, whereas it is around 0.09 or larger in scalings which use the Greenwald fraction. These considerations support the indication mentioned above that density peaking is statistically more likely a function of collisionality rather than a function of the Greenwald fraction.

6.1. Comparison between the statistical significances of $\nu_{\rm eff}$ and F_{GR}

Table 3 shows that, by the replacement of $\nu_{\rm eff}$ with F_{GR}, while keeping the same remaining regression variables in the regression model, the statistical significance of $\nu_{\rm eff}$ is larger than that of F_{GR} by a factor 1.75 in models which include R_{geo} and only by a factor 1.2 in models which do not include R_{geo}. The Greenwald fraction can be interpreted as a dimensionless parameter in the framework of atomic physics [13, 14]. However is not a dimensionless parameter of a fully ionized plasma, namely it cannot be expressed in terms of ρ_*, ν_* and β. Therefore, it can be argued that a more appropriate comparison between the statistical significances of collisionality and the Greenwald fraction is obtained in case the statistical significance of $\nu_{\rm eff}$ is compared with the statistical significance of the combination of F_{GR} with another dimensional parameter, and in particular of the pair $[F_{GR}\ ,\ R_{geo}]$ [11]. Such a comparison is presented in this subsection.

The test statistic $T_{(N,R)}$ for this pair of variables is defined by

$$T_{(N,R)} = (\hat{a}_N\ ,\ \hat{a}_R)^t\ \ \Sigma^{-1}\ \ (\hat{a}_N\ ,\ \hat{a}_R),$$

where $(\hat{a}_N\ ,\ \hat{a}_R)$ is the column vector containing the estimated regression coefficients $\hat{a}_N$ of F_{GR} and $\hat{a}_R$ of R_{geo}. The 2×2 matrix Σ is the covariance matrix of the regression coefficients $(a_N\ ,\ a_R)$. We remind the reader that in case of the null hypothesis, the test statistic $T_{(N,R)}$ would have a χ^2 probability distribution with two degrees of freedom, while the square of the regression coefficient a_ν of $\nu_{\rm eff}$ would have a χ^2 distribution with one degree of freedom. We find that the test statistic $T_{(N,R)} = 92.2$, which corresponds to 46.1 standard deviations of the two degrees of freedom χ^2 distribution, and to 15.4 times the critical value of the χ^2 distribution corresponding to a probability of 95%. For comparison, $(\hat{a}_\nu\ /\ STD(a_\nu))^2 = 103.6$, which is 73.3 standard deviations of the one degree of freedom χ^2 distribution and 27.0 times the critical value of the χ^2 distribution corresponding to a probability of 95% (namely 1.8 times larger than the corresponding number for the pair $[F_{GR}\ ,\ R_{geo}]$). In conclusion, the statistical significance of collisionality considered alone is larger than that of F_{GR} considered alone, as well as of that of the pair of regression variables $(F_{GR}\ ,\ R_{geo})$, but in all cases by less than a factor of 2.

An analogous comparison is performed with the pair of regression variables [Γ^*_{NBI} , F_{GR}], for which the test statistics $T_{\Gamma,N} = 147.4$, namely 24.6 times the critical value of the two degrees of freedom χ^2 distribution. This number is comparable with the corresponding value 27.0, obtained for the regression variable $\nu_{\rm eff}$ considered alone. Therefore it can be stated that the statistical significance of collisionality considered alone is as large as the statistical significance of the pair [Γ^*_{NBI} , F_{GR}]. Of course, the pair of regression variables [Γ^*_{NBI} , F_{GR}] is found to have a test stastistic which is smaller than the pair (Γ^*_{NBI} , $\nu_{\rm eff}$) ($T_{\Gamma,\nu} = 173.6$).

In conclusion, collisionality is found to have a statistical significance, which is larger than that of F_{GR} considered alone, as well as of F_{GR} considered in combination with other parameters, like R_{geo} or the beam fuelling parameter Γ^*_{NBI}, but in any case by less than a factor of 2. Although all these results support the indication that collisionality rather than the Greenwald fraction is the appropriate scaling parameter for the density peaking, they are such that, from the statistical standpoint, on the basis of the present dataset, the Greenwald fraction cannot be eliminated among the possible proper scaling parameters, although it requires to be used in combination with a parameter describing the plasma size. Moreover, the Greenwald fraction remains a highly significant parameter in regression models in which collisionality is excluded. For this reason, in the next section, we deem more appropriate to not exclude the possibility of a dependence on F_{GR}, rather than on collisinality, and to propose separate scalings in terms of $\nu_{\rm eff}$ and F_{GR}.

7. Proposed scalings and ITER predictions

In this section we propose three linear regressions, one which includes collisionality and excludes the Greenwald fraction, and two which exclude the collisionality and include the Greenwald fraction, in one case in combination with dimensionless parameters, and in a second case in combination with the major radius R_{geo}. Of course, in these proposed scalings, only the statistically significant regression variables are used. These has been selected by an iterative procedure [11]. Starting from the regression models including all variables, as those presented in Tables (3) and (4), the least significant variable is eliminated, and the regression remade with the remaining variables. The procedure is iterated up to the moment that neglecting the next variable provides a significant increase of the RMSE. The variation $\Delta RMSE$ of the RMSE due to the exclusion of the least significant variable becomes important when it is large as compared with the ratio $RMSE/N$, where N is the number of observations.

The regression without using the Greenwald fraction reads

$$pk_{\rm scl\,\nu} = 1.347 \pm 0.014 - (0.117 \pm 0.005)\, log(\nu_{\rm eff}) + \\ + (1.331 \pm 0.117)\,\Gamma^*_{NBI} - (4.030 \pm 0.810)\,\beta, \qquad (3)$$

with RMSE = 0.115 (66.7% confidence intervals for the regression coefficients, corresponding to one standard deviation, are quoted). The regression without using

the collisionality reads

$$pk_{\rm scl\,FGR} = 1.849 \pm 0.044 - (0.636 \pm 0.035)\, F_{GR} + \\ + (1.911 \pm 0.151)\, \Gamma^*_{NBI} - (22.54 \pm 3.73)\, \rho_* + \\ - (0.083 \pm 0.015)\, T_{e2}/\langle T_e\rangle_{Vol} + (0.292 \pm 0.069)\, \delta, \qquad (4)$$

with RMSE = 0.127. Here T_{e2} stands for $T_e(\rho_{pol} = 0.2)$.

Besides these two scalings, a very simple engeneering oriented scaling which includes both the Greenwald fraction and the major radius, but does not include the collisionality, can also be given,

$$pk_{\rm scl\,FGR\&R} = 1.253 \pm 0.037 - (0.499 \pm 0.030)\, F_{GR} + \\ + (2.094 \pm 0.137)\, \Gamma^*_{NBI} + (0.117 \pm 0.009)\, R_{geo}, \qquad (5)$$

with RMSE = 0.123.

Density peaking as a function of the three proposed scalings is plotted in Fig. 4. These regressions, as well as analogous regressions in the logarithmic form, are applied for ITER predictions. We consider the ITER standard scenario, with the plasma parameters described in [4], and in particular $\langle T_e\rangle_{Vol} = 8$ keV and $\langle n_e\rangle_{Vol} = 10^{20}$ m^{-3}, namely $\log(\nu_{\rm eff}) = -1.64$ and $\rho_* = 1.84\ 10^{-3}$, Greenwald fraction 0.85, and taking the beam particle source equal to zero (log is the natural logarithm).

The scaling $pk_{{\rm scl}\,\nu}$ in Eq. (3) predicts the peaking $n_{e2}/\langle n_e\rangle_{Vol} = 1.46 \pm 0.04$. More in general, *all* linear or logarithmic regressions which include collisionality in the regression variables predict a peaked density profile for ITER, more precisely values of the peaking above 1.35. We remind that the corresponding scaling based on the database of JET only observations predicts a density peaking for ITER of 1.6 [3].

The scaling $pk_{\rm scl\,FGR}$ in Eq (4) predicts the ITER peaking $n_{e2}/\langle n_e\rangle_{Vol} = 1.20 \pm 0.14$, namely a rather flat profile. More in general, scalings which exclude collisionality from the regression variables, predict flat density profiles for ITER, namely values of peaking around 1.2 or below. In these scalings, the main reason for which the ITER density peaking is not predicted to be exactly equal to one, is the negative regression coefficient in front of ρ_*, which reflects the fact that, at the same Greenwald density, JET profiles are slightly more peaked than AUG profiles. A stronger effect of this kind is obtained in scalings which include both the Greenwald fraction and the major radius. In the extreme case of the scaling $pk_{\rm scl\,FGR\&R}$ in Eq. (5), the ITER density profile is predicted to be peaked, $n_{e2}/\langle n_e\rangle_{Vol} = 1.54 \pm 0.12$, namely even above the prediction given by the scaling with collisionality.

8. Conclusions

A statistical analysis has been performed on a combined dataset of AUG and JET H–mode observations in order to identify the main dependences of density peaking and propose scalings to be applied for ITER predictions. The present work has reached the following conclusions.

- Collisionality is the statistically most relevant variable in regressions for density peaking in our combined AUG and JET dataset.
- All scalings of the density peaking which include collisionality in the regression variables, predict a peaked density profile for ITER, more precisely $n_{e2}/\langle n_e \rangle_{Vol} \simeq 1.5$.
- Collisionality has a statistical significance which is higher than that of the Greenwald fraction, and higher, although comparable, than that of the pair of variables Greenwald fraction and major radius. Although the Greenwald fraction is a highly significant and relevant regression parameter when collisionality is excluded, it is found that it requires to be used in combination with a parameter like R_{geo}, or ρ_*, describing the plasma size in order to obtain a scaling of the density peaking which has a RMSE which is comparable to that obtained using collisionality alone at their place. Moreover, density peaking is found to increase with plasma size at fixed Greenwald fraction.
- The beam fuelling, although non–negligible, cannot explain alone the full observed variation of density peaking.

These conclusions, obtained exclusively on the basis of presently available experimental observations of AUG and JET, look for confirmation and extensions by means of more complete empirical investigations on density peaking, involving possibly an increasing number of devices and new dedicated experiments testing specific dependences. More specifically, a device with a smaller major radius and operating at high field and high current could provide the observations needed in order to decorrelate more effectively collisionality and Greenwald fraction, providing the required test to the scalings proposed in this work. Very recent experimental results obtained in Alcator C-Mod [15] are found in agreement with the present conclusions, and provide a conclusive confirmation of the indication presented here that density peaking is statistically more likely a function of collisionality than of the Greenwald fraction [16]

Comparisons of the present empirical results with recent theoretical studies on particle transport [17, 18, 19, 20, 21, 22] are discussed in Ref. [23].

In particular, there remains an open issue with regard to the dependence on the magnetic shear or more generally, on the peaking of the current density profile, observed in L–mode plasmas [24, 25, 26, 27, 28] and predicted by several theoretical models [24, 25, 29]. Such a dependence is not found in the present study dedicated to H–mode plasmas, nor was it found in the similar analysis specific to JET only [3]. We mention that in these works the current density profiles are usually more peaked in L–mode plasmas than in H–mode plasmas, namely the intervals of l_i covered by the plasmas in the two confinement modes do not overlap (e.g. see Fig. 2 and Fig. 7 in Ref. [2]). To this purpose, it is certainly of interest to investigate experimentally whether a dependence of density peaking on the peaking of the current density profile occurs in H–mode plasmas when the current density profile peaking is varied over the same window as that of L–mode plasmas.

A parameter, which has not been considered in the present analysis, since not reliably available over a rather large subset of observations, is the electron to ion temperature ratio. This was found to play some role in a subset of JET data in Ref. [3], and with a dependence which is in qualitative agreement with the theoretically predicted effect of thermodiffusion [30, 31, 29, 32]. Its inclusion in a multi–machine database is planned to be attempted in the near future. This may slightly modify the role of some parameters. In Ref.[3], it was found to reduce the importance of the beam source term.

Finally, a quantitative agreement between non–linear gyrokinetic simulations and the empirically identified collisionality dependence of density peaking is still missing. Such an agreement would provide the required theoretical support to the prediction of a peaked density profile for ITER, as obtained in scalings which include collisionality in the present empirical work.

References

[1] Angioni C. *et al* 2003 *Phys. Rev. Lett.* **90** 205003.
[2] Weisen H. *et al* 2005 *Nucl. Fusion* **45** L1.
[3] Weisen H. *et al* 2006 *Plasma Phys. Control. Fusion* **48** A457.
[4] Mukhovatov V. *et al* 2003 *Nucl. Fusion* **43** 942.
[5] Zabolotsky A. *et al* 2006 *Nucl. Fusion* **46** 594.
[6] Bailey D. 1998 Report JET-R (1998) 04.
[7] Stubberfield P.M. and Watkins M.L. 1987 Report JET-DPA(06)/87.
[8] Cox M. 1984 Culham Report for JET KR5-33-04.
[9] Lister G. G. 1985 *"FAFNER - A Fully 3-D Neutral Beam Injection Code Using Monte Carlo Methods"* IPP-Report 4–222, Garching, Germany.
[10] Furno I. *et al* 2005 *Plasma Phys. Control. Fusion* **47** 49.
[11] Kardaun O.J.W.F., *Classical Methods of Statistics*, Springer–Verlag Berlin Heidelberg 2005.
[12] Garzotti L. *et al* 2006 *Nucl. Fusion* **46** 994.
[13] Lackner K. 1990 *Comments Plasma Phys. Control. Fusion* **13** 163.
[14] Lackner K. 1994 *Comments Plasma Phys. Control. Fusion* **15** 359.
[15] M. Greenwald, D. Ernst, A. Hubbard, J.W. Hughes, Y. Lin, J. Terry, S. Wukitch, K. Zhurovich, Bull. Am. Phys. Soc. Vol. **51**, no 7, 139 (2006), 48th A.P.S. – D.P.P. Meeting 2006, JO1 3.
[16] Greenwald M. *et al* submitted to *Nucl. Fusion* (2007).
[17] Angioni C. *et al* 2003 *Phys. Plasmas* **10** 3225.
[18] Angioni C. *et al* 2005 *Phys. Plasmas* **12** 112310.
[19] Pereverzev G.V. *et al* 2005 *Nucl. Fusion* **45** 221.
[20] Estrada–Mila C. *et al* 2005 *Phys. Plasmas* **12** 022305.
[21] Jenko F. *et al* 2005 *Plasma Phys. Control. Fusion* **47** B195.
[22] Estrada–Mila C. *et al* 2006 *Phys. Plasmas* **13** 074505.
[23] Angioni C., Carraro L., Dannert T. *et al* 2007 *Phys. Plasmas* **14** 055905.
[24] Isichenko M.B., Gruzinov A.V. and Diamond P.H. 1996 *Phys. Rev. Lett.* **74** 4436.
[25] Baker D.R. and Rosenbluth M.N. *et al* 1998 *Phys. Plasmas* **5** 2936.
[26] Baker D.R. *et al* 2000 *Nucl. Fusion* **40** 1003.
[27] Weisen H. *et al* 2002 *Nucl. Fusion* **42** 136.
[28] Weisen H. *et al* 2004 *Plasma Phys. Control. Fusion* **46** 751.
[29] Garbet X. *et al* 2003 *Phys. Rev. Lett.* **91** 035001.
[30] Coppi B. and Spight C. 1978 *Phys. Rev. Lett.* **41** 551.
[31] Weiland J., Jarmén A.B. and Nordman H. 1989 *Nucl. Fusion* **29** 1810.

[32] Angioni C. *et al* 2004 *Nucl. Fusion* **44** 827.

	$\overline{n}_{e\,lin}$ [10^{19} m^{-3}]	I_p [MA]	B_T [T]	P_{tot} [MW]
AUG	3.80 - 12.74	0.60 - 1.21	1.5 - 3.10	2.40 - 16.0
JET	1.78 - 10.14	0.88 - 3.67	0.86 - 3.72	5.5 - 21.4

Table 1.

Table 1. *Minimum and maximum values of the ranges of engineering parameters covered by the combined database of AUG and JET observations.*

	pk_{ne}	Γ^*_{NBI}	$\log(\nu_{\mathrm{eff}})$	F_{GR}	$\rho_*\times10^3$	$\beta\times10^2$	q_{95}	δ	k	$T_{e2}\,/\,\langle T_e\rangle$
mean	1.409	0.076	-0.149	0.546	5.51	1.47	3.66	0.214	1.75	2.24
mean AUG	1.369	0.094	0.291	0.573	6.30	1.57	3.75	0.168	1.79	2.43
mean JET	1.457	0.054	-0.693	0.513	4.53	1.35	3.56	0.271	1.70	2.01
STD	0.200	0.044	1.099	0.181	1.64	0.603	0.81	0.088	0.06	0.39
STD AUG	0.201	0.044	1.064	0.188	1.35	0.452	0.56	0.058	0.03	0.32
STD JET	0.188	0.031	0.875	0.167	1.42	0.730	1.04	0.084	0.03	0.36
min AUG	0.997	0.010	-2.021	0.280	2.68	0.526	2.83	0.101	1.69	1.82
min JET	1.094	0.003	-2.732	0.215	2.42	0.443	2.25	0.181	1.61	1.18
max AUG	2.058	0.230	2.957	1.087	9.73	3.32	6.56	0.399	1.86	4.51
max JET	2.010	0.133	1.494	0.966	8.82	3.86	6.35	0.505	1.78	3.03

Table 2.

Table 2. *Mean values, standard deviations as well as min and maximum values of the parameters used in the multivariable regression analysis. Values of the full AUG and JET combined database, as well as, for comparison, of the subsets of AUG and JET observations are provided (log is the natural logarithm).*

	Γ^*_{NBI}	$\log(\nu_{\mathrm{eff}})$	F_{GR}	ρ_*	β	q_{95}	δ	$T_{e2}\,/\,\langle T_e\rangle$	R_{geo}	RMSE
no ν_{eff}	7.22		-2.96	0.98	-1.53	0.24	-1.14	-0.94	4.10	0.122
no F_{GR}	5.09	-5.23		1.03	-2.46	-1.01	-1.22	0.09	1.99	0.113
All variables	5.07	-4.35	0.97	1.42	-2.60	-1.18	-1.54	0.15	2.19	0.113
no ν_{eff} & R_{geo}	6.05		-8.43	-2.44	0.86	-0.02	1.81	-2.37		0.127
no F_{GR} & R_{geo}	4.63	-10.13		-0.16	-1.77	-1.47	-0.26	-0.45		0.114
no R_{geo}	4.55	-5.37	0.43	0.04	-1.71	-1.52	-0.45	-0.46		0.114

Table 3.

Table 3. *Values of the statistical significance StS for various plasma parameters used as regression variables for the density peaking in different regressions and corresponding value of the RMSE.*

	Γ^*_{NBI}	$\log(\nu_{\rm eff})$	F_{GR}	ρ_*	β	q_{95}	δ	$T_{e2}\,/\,\langle T_e\rangle$	R_{geo}
no $\nu_{\rm eff}$	0.096		-0.058	0.028	-0.035	0.003	-0.017	-0.012	0.094
no F_{GR}	0.077	-0.098		0.027	-0.048	-0.013	-0.016	0.001	0.049
All variables	0.077	-0.115	0.028	0.042	-0.063	-0.015	-0.025	0.002	0.057
no $\nu_{\rm eff}$ & R_{geo}	0.083		-0.120	-0.053	0.017	-0.000	0.022	-0.030	
no F_{GR} & R_{geo}	0.067	-0.127		-0.003	-0.032	-0.018	-0.003	-0.006	
no R_{geo}	0.066	-0.136	0.012	0.001	-0.037	-0.019	-0.006	-0.006	

Table 4.

Tablee 4. *Values of the statistical relevance StR for various plasma parameters used as regression variables for the density peaking.*

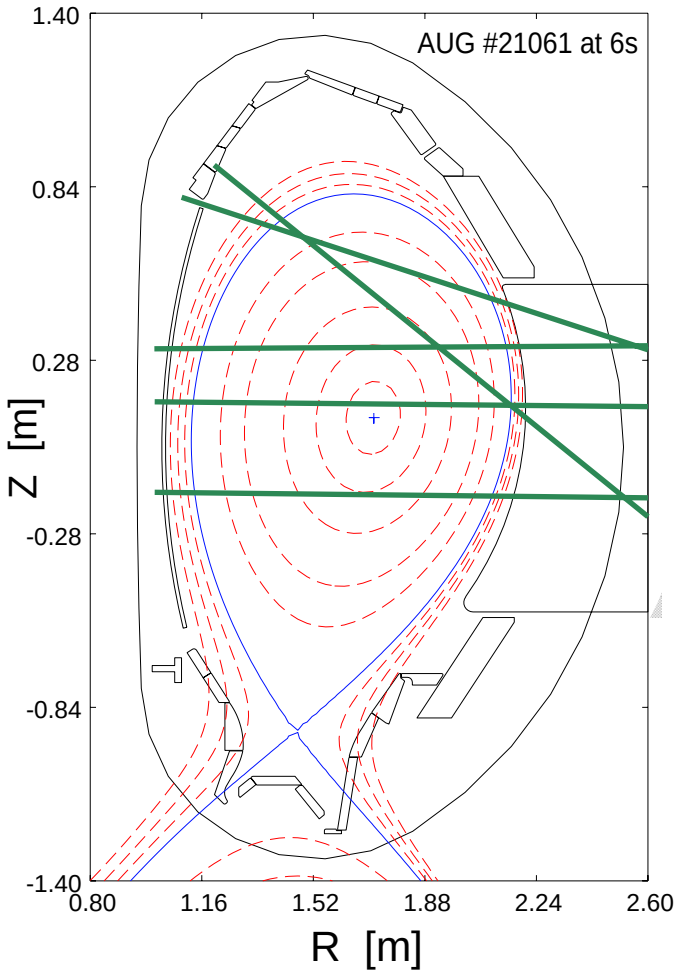

Fig. 1. *(color online) Geometry of the lines of sight ot the AUG DCN interferometer, and the AUG equilibrium used to re–map the JET SVD–I density profiles.*

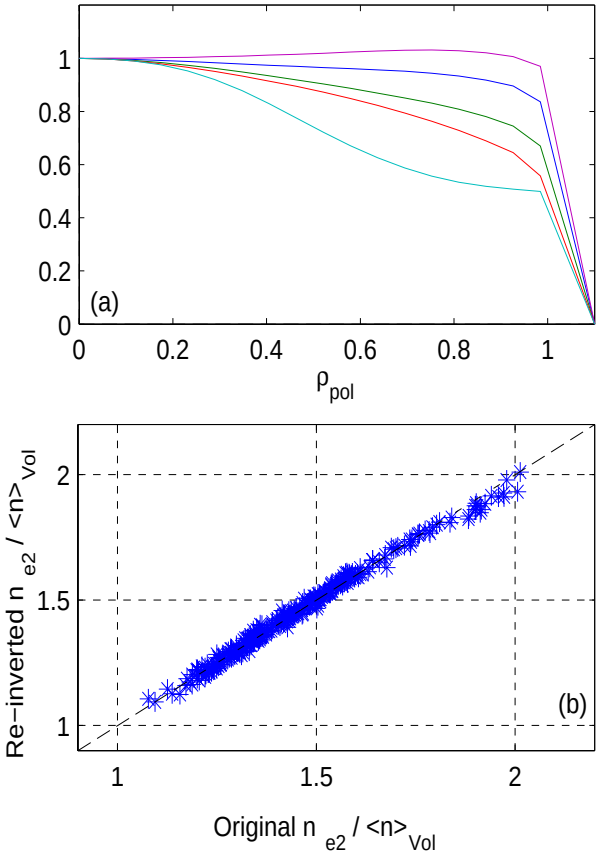

Fig 2. *(color online) (a) The 5 basis functions chosen to describe the profile shape in the inversion of the measured AUG and the computed JET line integrals of the AUG interferometer, and (b) density peaking values obtained from the re–inverted JET profiles against the values of density peaking computed directly on the original JET SVD–I profiles.*

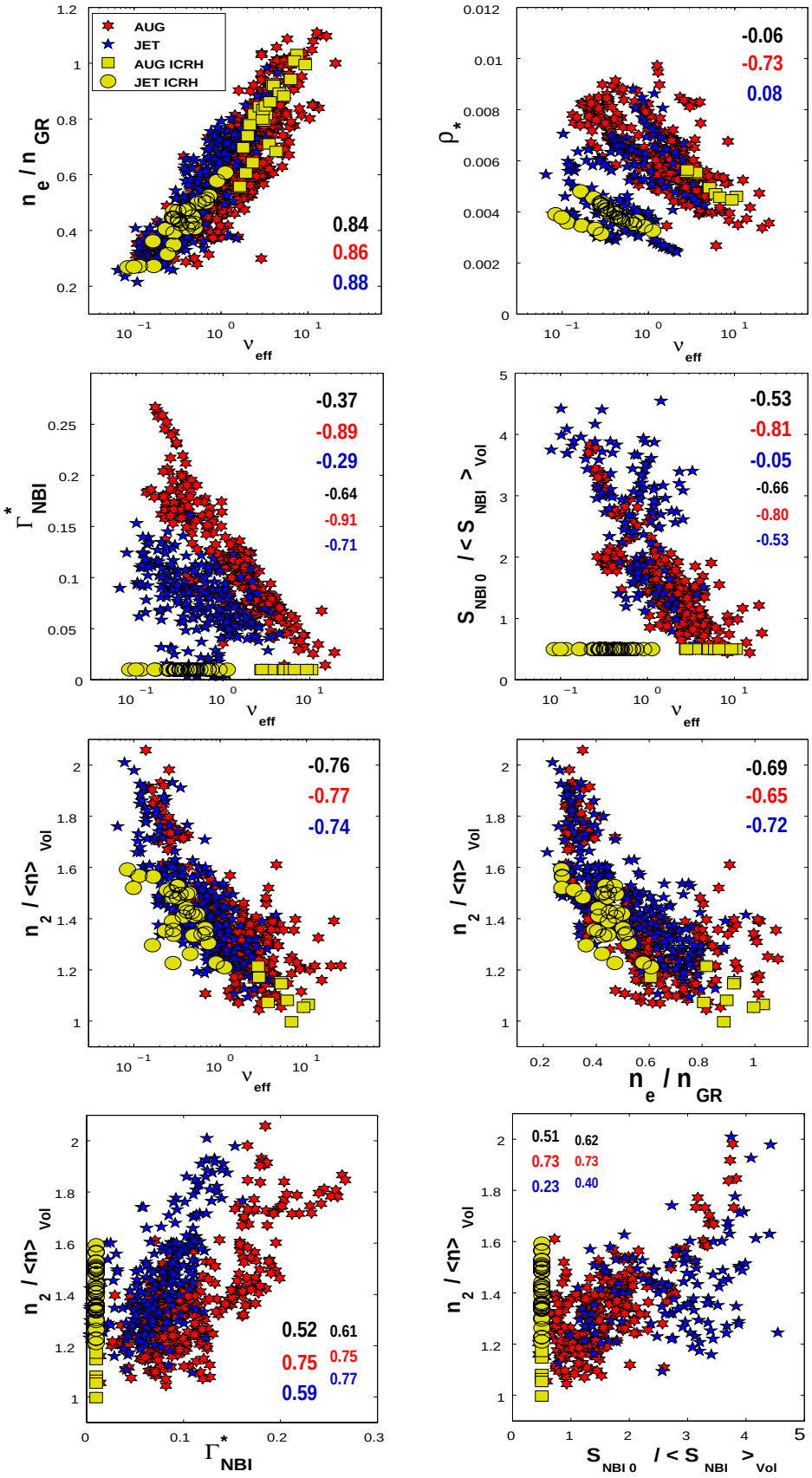

Fig. 3. *(color online) Univariable scatter plots among various plasma parameters. The numbers in the plots provide the related correlation coefficients, in black (top value) for the*

combined dataset, in red (central value) for the AUG subset, in blue (bottom value) for the JET subset. Smaller fonts used in plots with the beam source parameters indicate the correlation coefficients over the subset of observations with $P_{NBI} / P_{TOT} > 0.7$.

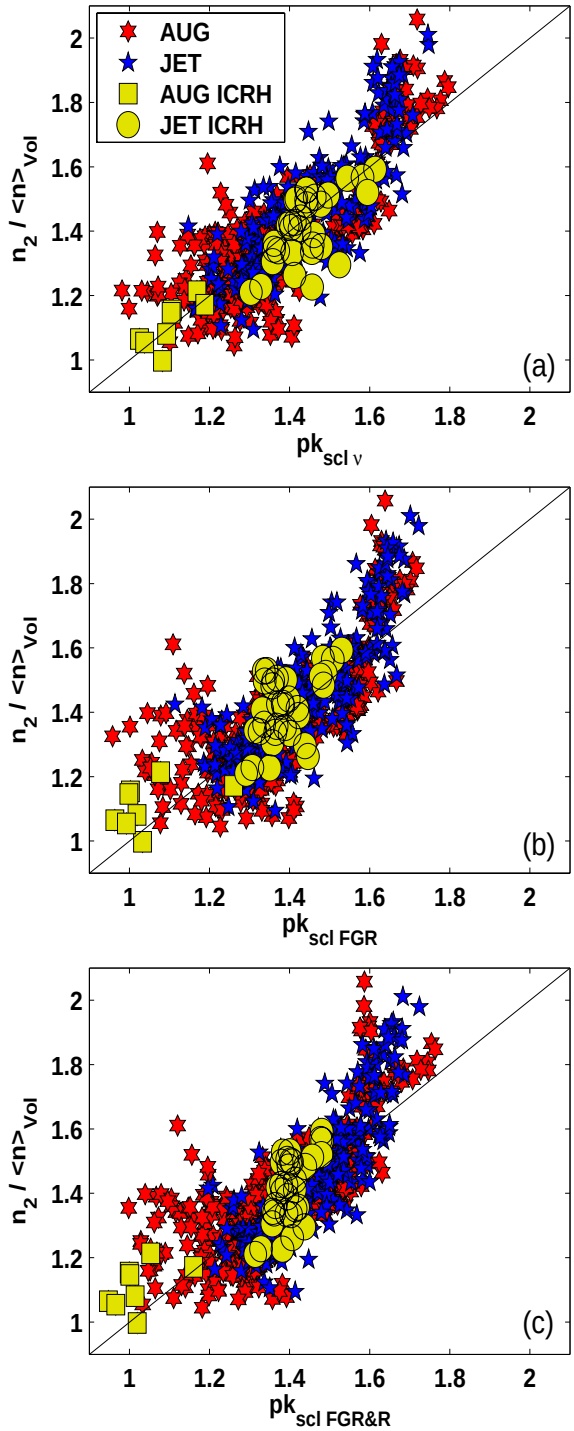

Fig 4. *(color online) Density peaking as a function of the three proposed scalings, (a)* $pk_{\mathrm{scl}\,\nu}$ *in Eq. (3), (b)* $pk_{\mathrm{scl\,FGR}}$ *in Eq. (4), and (c)* $pk_{\mathrm{scl\,FGR\&R}}$ *in Eq. (5).*