# Applying Advanced Statistical Techniques to Tokamak L-H Threshold Data

A.J.Meakins [1] , D.C.McDonald [2] , Y.Andrew [2], J.G.Cordey [2], M.Coppins [1],
ITPA H-mode Threshold Database WG [3] and contributors to EFDA JET workprogramme [*]

*Work partly funded by Euratom and the UK EPSRC*

1. *Imperial College, Prince Consort Road, London, SW7 2BZ, UK*
2. *EURATOM-UKAEA Fusion Association, Culham Science Centre, Abingdon, UK*
3. *Alcator C-Mod: J.A.Snipes, M.Greenwald; ASDEX/ASDEX Upgrade: F.Ryter, O.J.W.F.Kardaun, J.Stober; COMPASS-D/MAST: M.Valovic; DIII-D: J.C.DeBoo; JET/EFDA: Y.Andrew, J.G.Cordey, R.Satori, K.Thomsen; JFT-2M/JT-60U: T.Takizuka, Y.Miura, T.Fukuda, Y.Kamada, K.Shinohara, K.Tsuzuki; PBX-M/NSTX: S.M.Kaye, C.Bush, R.Maingi; TCV: Y.R.Martin; TUMAN-3M: S.Lebedev.*

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

When planning experiments on a given tokamak empirical log-linear power law fits are often used to predict parameters by interpolating data from past experiments. Predicting the threshold heating power required to achieve a transition from the L-mode to the H-mode of operation is one such area where these scalings are used frequently. The JET L-H threshold data, however, is known to contain regions of operating space (e.g. at low density) where power law fits are unsuitable. In this paper we present a Neural Network (NN) approach to the analysis of JET L-H data. NNs have a more flexible functional form allowing them to describe more complex trends in the data. Fitting the NNs is performed using a novel code, DyNE, which explicitly prevents overfitting of the data. Results show an improvement in predictive ability compared to power laws fit to the same data (13% improvement in Root Mean Squared Error) and features such as the low density behaviour of the L-H threshold are correctly described. The application of this method to next step devices, such as ITER, would provide a useful tool that should allow for significant savings in experimental time.

As it is not possible to extrapolate from NNs, extrapolation of L-H threshold data to next step devices power law scalings remain the more appropriate tool. The majority of power law scalings used in multi-machine comparisons are fit using Ordinary Least Squares (OLS). A key assumption in the derivation and application of OLS is that any random scatter in the data must be present only on the dependent variable otherwise the fits become biased. If errors on the data do not agree with this assumption then more sophisticated fitting techniques must be applied. We present an analysis of the errors in the International Global H-mode Threshold Database (IGDBTH) which shows that the errors on the dependent and independent variables are comparable. As such we apply Errors-in-Variables Orthogonal Regression (EVOR) which treats the errors in the dependent and independent variables in an equal manner. For the IGDBTH the errors in one of the independent variables was found to be of similar magnitude to the dependent variable suggesting a skewed OLS fit. Refitting using EVOR produced a modified scaling. This scaling predicts the threshold power for ITER to be 37-38MW, a small increase of 3-4MW over previous estimates, however it still falls within the ITER design parameters.