

Real-time capable first principle based modelling of tokamak turbulent transport

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 Nucl. Fusion 55 092001

(<http://iopscience.iop.org/0029-5515/55/9/092001>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 69.62.157.224

This content was downloaded on 12/09/2015 at 14:36

Please note that [terms and conditions apply](#).

Letter

Real-time capable first principle based modelling of tokamak turbulent transport

J. Citrin^{1,2}, S. Breton², F. Felici³, F. Imbeaux², T. Aniel², J.F. Artaud²,
B. Baiocchi⁴, C. Bourdelle², Y. Camenen⁵ and J. Garcia²

¹ FOM Institute DIFFER: Dutch Institute for Fundamental Energy Research, PO Box 6336, 5600 HH Eindhoven, The Netherlands

² CEA, IRFM, F-13108 Saint Paul Lez Durance, France

³ Eindhoven University of Technology, Department of Mechanical Engineering, Control Systems Technology Group, PO Box 513, 5600 MB Eindhoven, The Netherlands

⁴ Istituto di Fisica del Plasma 'P. Caldirola', Associazione Euratom-ENEA-CNR, 53 20125 Milano, Italy

⁵ Aix-Marseille Université, CNRS, PIIM UMR 7345, 13397 Marseille, France

E-mail: J.Citrin@diffier.nl

Received 25 February 2015, revised 22 May 2015

Accepted for publication 14 July 2015

Published 30 July 2015



Abstract

A real-time capable core turbulence tokamak transport model is developed. This model is constructed from the regularized nonlinear regression of quasilinear gyrokinetic transport code output. The regression is performed with a multilayer perceptron neural network. The transport code input for the neural network training set consists of five dimensions, and is limited to adiabatic electrons. The neural network model successfully reproduces transport fluxes predicted by the original quasilinear model, while gaining five orders of magnitude in computation time. The model is implemented in a real-time capable tokamak simulator, and simulates a 300 s ITER discharge in 10 s. This proof-of-principle for regression based transport models anticipates a significant widening of input space dimensionality and physics realism for future training sets. This aims to provide unprecedented computational speed coupled with first-principle based physics for real-time control and integrated modelling applications.

Keywords: turbulence, tokamak transport, control, integrated modelling

(Some figures may appear in colour only in the online journal)

1. Introduction

Particle, momentum, and heat transport in the tokamak core is dominated by turbulence driven by plasma microinstabilities [1, 2]. An accurate predictive model for turbulent transport fluxes is thus vital for the interpretation and optimization of present-day experiments, and extrapolation to and control of future machines.

Direct numerical simulation with massively parallel nonlinear gyrokinetic codes has provided tremendous insight to the underlying transport physics and success in reproducing

experimental fluxes in many regimes. However, the computational cost—typically 10^5 CPU hours for a local flux calculation at a single radial point—precludes the routine use of such codes for integrated tokamak transport simulations which demand $\sim 10^3$ flux computations per 1 s of plasma evolution on JET scale devices.

Reduced turbulent transport models have been constructed to increase tractability. They are based on the quasilinear approximation, which is proven to be largely valid in the core of tokamak plasmas [3–5]. These rely on nonlinear simulations for validating their ansatzes and normalizing factors.

These models have proven successful in reproducing experimental profiles in many cases. Examples are TGLF [6] and QuaLiKiz [7]. A ~ 6 orders of magnitude speedup is gained in quasilinear calculations compared to nonlinear simulations. However, while extremely useful, the tractability of such models is still marginal for convenient large-scale scenario development over discharge timescales. For example, 1 s of JET plasma evolution can take up to 24 h with 20 processors, depending on the integrated modelling platform used. This speed is also insufficient for applications such as trajectory optimization, and simulations for developing real-time controllers. Furthermore, any increase in physics fidelity in the models often results in a trade off with further decrease in tractability.

This Letter illustrates an approach to overcome these challenges. The central point is to relegate the expensive flux calculations to a stage precedent to its use in a transport simulation. Instead, analytical formulae are to be used in the simulation, based on a neural network (NN) nonlinear regression of quasilinear fluxes previously compiled in a database. The advantage is twofold: (1) the numerical resolution of analytical formulae is orders of magnitude faster than original flux calculation; (2) the computation time required for compiling the database is independent from the computation time spent during the tokamak simulation itself, hence the training set for NN regression can include results from more complete codes than used in contemporary integrated transport modelling.

Neural networks have found multiple applications in tokamak research, including: nonlinear regression for energy confinement scaling [8]; neoclassical transport [9]; rapid determination of equilibria [10], electron temperature profiles [11], and charge exchange spectra [12]; classification of disruption [13–15] and L-H transition onsets [16]. Most related to this work is a regression of DIII-D heat fluxes from experimental power balance databases [17].

2. Quasilinear transport model and training set

The QuaLiKiz quasilinear gyrokinetic transport model [5, 7, 18–20] was employed in this work. QuaLiKiz solves a linear gyrokinetic dispersion relation for calculating wavenumber spectra of instability growth rates and frequencies. Then, integrating over the spectra, the transport fluxes are calculated via quasilinear flux integrals and nonlinear saturation rules. The bulk of the computational time is spent in the first stage, the dispersion relation solver. QuaLiKiz has been coupled to the CRONOS [21] integrated modelling suite, and has successfully reproduced temperature and density profiles of JET and Tore-Supra discharges [22, 23]. Following recent upgrades [24], the computational time for the QuaLiKiz eigenvalue solver at a single wavenumber in QuaLiKiz is on the order of ~ 1 s.

A database of QuaLiKiz solutions was constructed, in the ion temperature gradient (ITG) instability regime. This instability is often the primary driver of tokamak microturbulence. The code was run with adiabatic electrons for simplicity,

Table 1. Summary of input parameters for the QuaLiKiz adiabatic electron ITG database employed in this work.

Parameter	Min value	Max value	No. of points
R/L_{T_i}	2	12	30
T_i/T_e	0.3	3	20
q	1	5	20
\hat{s}	0.1	3	20
$k_{\theta}\rho_s$	0.05	0.8	16
Total no. of points			3 840 000

which also decreases the computational time to ~ 300 ms. The database covers four input parameters known to have significant impact on ITG transport fluxes in this regime: the driving normalized logarithmic ion temperature gradient R/L_{T_i} , the ion to electron temperature ratio T_i/T_e , the safety-factor q , and the magnetic shear $\hat{s} \equiv \frac{r}{q} \frac{dq}{dr}$. In addition, the input normalized wavenumber $k_{\theta}\rho_s$ was scanned, constricted to above ion-Larmor-radius scales, where $\rho_s \equiv \sqrt{T_e m_i / (Z_i q_e B)}$. The following parameters were maintained fixed: the normalized logarithmic density gradient $R/L_n = 3$, normalized radial location $r/a = 0.5$. No Shafranov shift was assumed in the geometry. The database consists of a dense grid of points summarized in table 1, from which the training sets for the neural network were sifted. The QuaLiKiz outputs we investigate are: growth rates and frequencies, which correspond to 5D input space; ion heat flux, which corresponds to 4D input space due to integration over wavenumbers. The database includes cases corresponding to unstable modes, and cases where no instabilities were found by QuaLiKiz, and the outputs are set to 0.

A regression of the ion heat flux has immediate application for transport modelling. However, a regression of the more primitive linear output has its own specific applications. For example, since the dispersion relation solver is the slowest part of the code, a fast reproduction of growth rates, frequencies and eigenfunctions would allow rapid tests of various saturation rule formulations throughout parameter space. These saturation rules typically evolve following continuous comparisons with nonlinear simulations and experiments. In this sense, a database consisting of the complete outputs of linear codes does not become obsolete, while a quasilinear flux database can.

3. Neural networks

The goal is to find analytical formulae which robustly reproduce the various QuaLiKiz outputs. To this end, a multilayer perceptron neural network is used, which is a nonlinear function with tunable variables (weights and biases), with the property of universal approximation [25, 26]. For an overview, with an emphasis on applications for fusion, see [27]. Linear combinations of the inputs and biases are propagated through a series of nonlinear transfer function vectors (named ‘hidden layers’), until eventually linearly combined to an output layer. With two hidden layers and a single output value (as used in this work), this is represented as:

$$y = b_3 + \sum_i^N w_i^2 g \left(b_1^2 + \sum_j^M w_{ij}^1 g \left(b_j^1 + \sum_k^I w_{jk}^{\text{in}} x_k \right) \right) \quad (1)$$

where y is the output ‘neuron’ containing the output value (i.e. growth rate, frequency, or ion heat flux), x_k the vector of input values, b^n the bias vectors, w^{in} the $M \times I$ weight matrix connecting the input vector to the 1st hidden layer, w^1 the $N \times M$ weight matrix connecting the two hidden layers, and w^2 the weight vector connecting the 2nd hidden layer to the output neuron. g is the nonlinear transfer function, defined as a sigmoid in this work:

$$g(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (2)$$

following a series of optimization tests, two hidden layers, as shown in equation (1), were employed here. The hidden layer sizes M and N were set to 40. The input layer size, I , is 4 for ion heat fluxes, and 5 for growth rates and frequencies.

The key stage is the determination of the optimized values of the weights and biases. This is done by minimizing a cost function consisting of the average squared error between the network output and known target output. This set of target output, known as the ‘training set’, is a subset of the QuaLiKiz output values from the database. The BFGS algorithm [28], implementing a quasi-Newton method, was used for the weight and bias optimization. All NN training in this work was carried out with the MATLAB neural network toolbox [29]. Following training, the network output then emulates the original model within the database input parameter envelope. This is validated by comparison to validation sets sifted from the database, which are different from the training set.

To avoid overfitting the data, regularization techniques were used in the regression. This corresponds to adding a penalty term in the cost function related to the sum of squares of the network weights and biases, leading to smoother output. The use of regularization ensures that the NN response is smooth (e.g. without strong oscillations) in sparse regions of training set parameter space or when extrapolating beyond the training set envelope.

The analytic form of the nonlinear regression function allows for the calculation of analytical gradients of the outputs with respect to the inputs. This is vital for the efficient solution of fast implicit schemes in real-time capable core transport simulators such as RAPTOR [30]. The regularization also ensures smooth gradients throughout parameter space, important for the stability of such implicit schemes.

4. Regression results

We focus on the ion heat flux NN regression, due to its direct relevance for transport modelling applications. Successful regressions of the growth rates and frequencies were also obtained but for brevity not discussed here.

To capture the instability thresholds with high fidelity, the regression was only carried out for a training set corresponding to unstable modes. The NN output for the stable regions in the validation set was then negative, since the regularized

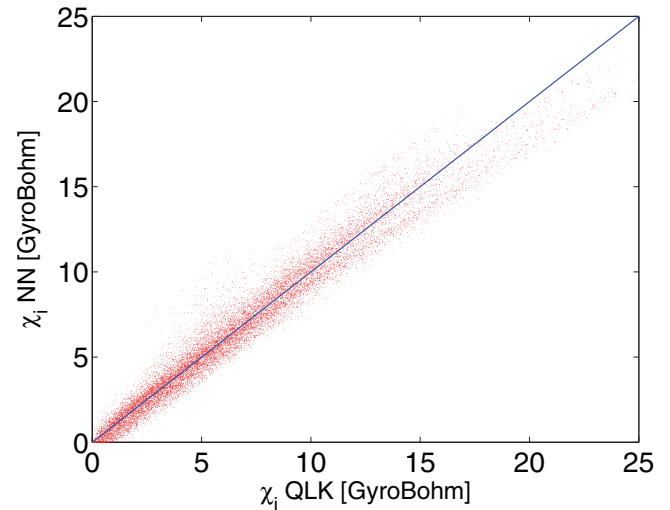


Figure 1. Comparison between normalized ion heat fluxes obtained directly from QuaLiKiz (x -axis) and those from its NN regression (y -axis).

network tends to smoothly extrapolate the trends observed towards the training set envelope. For the final heat flux output, these negative values were then set to zero to represent stability. This scheme avoids having the regularized regression network attempt to directly fit the discontinuous gradients at the instability thresholds, which would be performed poorly due to the regularization constraint. This is an important point since tokamak transport often tends to be maintained near the critical temperature gradient thresholds, especially in high temperature regimes.

The network was trained with a training set of 35 000 points chosen randomly from the set of unstable modes in the database. A comparison between the regression NN and QuaLiKiz outputs for a validation set of 10 000 unstable cases (different from the training set) is shown in figure 1. The regression network has an RMS error of 0.77 in gyroBohm units ($\chi_{i\text{GB}} = \frac{T_i^{3/2} m_i^{1/2}}{(Z_i q_i B)^2 r}$) when compared to the validation set. This RMS error is similar for the training set itself and is primarily due to the regularization constraint. The impact of this error on the simulated profiles is minor. This is due to stiffness, defined here as the local gradient of χ_i with respect to the driving R/L_{T_i} . To quantify this, a comparison was made between the R/L_{T_i} values predicted by the NN and QuaLiKiz to balance a representative $\chi_{i\text{GB}} = 1$. This was done for all values of q , \hat{s} , and T_i/T_e in the database. The RMS error was $\Delta(R/L_{T_i}) = 0.29$, which corresponded to an average relative error of only 4.2%.

The typical quality of the fits can be seen in figure 2, displaying scans of the 4 separate input parameters while the others remained fixed. Negative outputs of the NN network are set to zero. Note the resulting excellent fit of the instability thresholds. In addition, extrapolating the NN scans beyond the range of the training set maintains the trend observed in the data, due to the regularization. This is very encouraging with regard to extension of this approach to more sparse datasets in higher dimensions. However, we do not intend to routinely use NN models in poorly represented regions of parameter space, as the quality of extrapolation cannot be determined a priori.

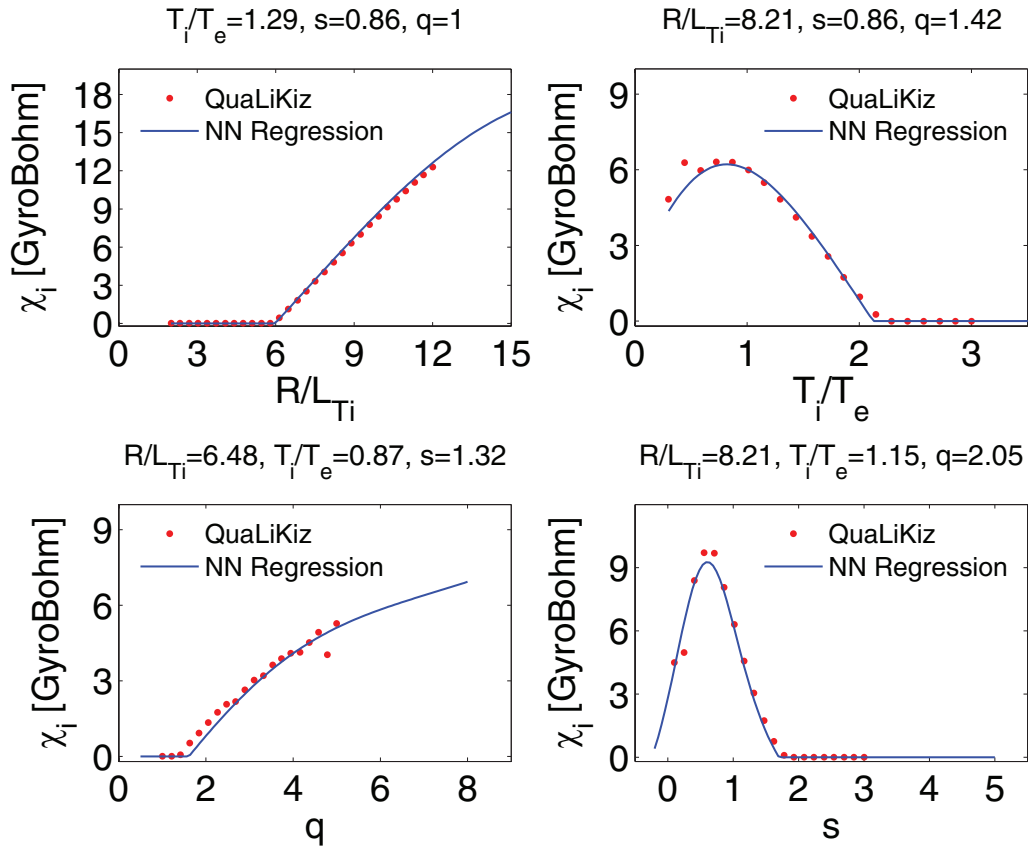


Figure 2. Comparison of NN parameter scans (blue solid lines) versus the original QuaLiKiz ion heat flux calculations (red dots). The scans are in R/L_{Ti} (top left panel), T_i/T_e (top right panel), q (bottom left panel) and s (bottom right panel).

Rather, the training sets should be continuously expanded to cover such encountered sparse or empty regions, and the NN then periodically retrained. Nevertheless, the smoothness of the regularized NN response when extrapolating ensures its robustness and stability during practical use as a transport model, including during phases when such sparse regions are encountered.

Each NN output is calculated on a sub 10 μ s timescale in MATLAB on a Intel(R) Xeon(R) E5450 CPU @ 3.00 GHz. This is a 5 order of magnitude speedup in comparison to the original QuaLiKiz calculations.

5. Application in transport codes

A transport model based on the trained neural network was constructed, and implemented both in the CRONOS [21] and RAPTOR [30] integrated modelling codes.

In CRONOS, the validity of the NN transport model was assured by a successful comparison with a JET baseline H-mode shot 73342 with ion and electron heat transport previously simulated [23] with the full QuaLiKiz model. For brevity we will not focus further on this case. Rather, we focus on the real-time simulation capabilities offered by coupling the NN model to RAPTOR.

Presently, RAPTOR only models electron heat transport. The NN model output was thus modified to roughly approximate ITG regime electron heat transport with kinetic

electrons. This was done by assuming that heat fluxes in ITG kinetic electron cases are higher by factor 3 compared with adiabatic electron cases, and furthermore assuming an ion to electron heat flux ratio of $q_i/q_e = 3$. These approximations are based on typical nonlinear and quasilinear observations in the ITG regime [5, 31].

In figure 3, we compare a RAPTOR simulation of an ITER hybrid scenario, using the QuaLiKiz NN model for electron heat transport, with a simulation of the same case originally carried out [32] using CRONOS and the GLF23 [33] transport model. Using GLF23 allows to compare over ITER-scale discharge times of >100 s, which is less tractable using the original QuaLiKiz model. For heat transport in a pure ITG regime, GLF23 and QuaLiKiz predictions are expected to be similar, as illustrated in specific single-time-slice comparisons [23].

The RAPTOR simulation uses all the same actuator (source) inputs and density evolution as the CRONOS simulation. Ion temperatures were held fixed at $T_i/T_e \sim 0.8$ in L-mode and $T_i/T_e \sim 0.9$ in H-mode. The NN model was operational within a normalized toroidal flux coordinate (ρ) range of 0.25–0.95. For $\rho > 0.95$, χ_e was feedback controlled to maintain a prescribed edge pedestal temperature of 4 keV. For $\rho < 0.25$, a constant χ_e was assumed to maintain a reasonable level of transport, since GLF23 and QuaLiKiz both predicted stability within that region. A RAPTOR simulation of an entire 300 s ITER discharge took 10 s on a single CPU, corresponding to 30× faster than real-time. This combination of simulation

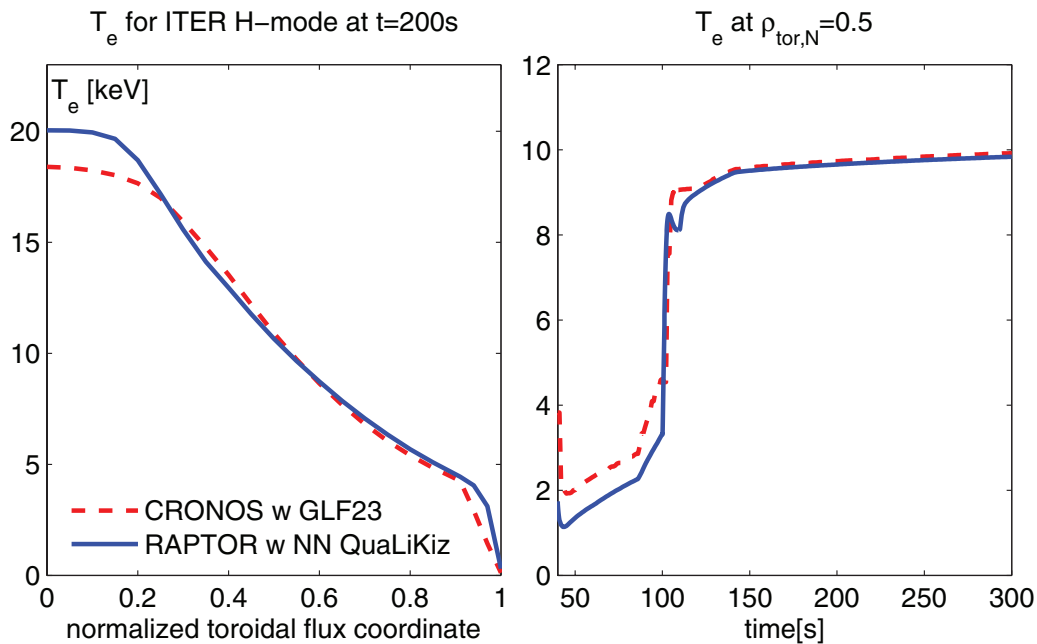


Figure 3. Comparison between the T_e predictions for an ITER hybrid discharge carried out with CRONOS/GLF23 [32] (red curve) and a RAPTOR simulation using the QuaLiKiz NN transport model (blue curve). A typical H-mode profile (left panel) and time dependence at mid-radius (right panel) are shown. The LH transition was set at 100 s.

speed and first-principle modelling is unprecedented. With CRONOS/GLF23, the same simulation took 24h.

6. Conclusions and outlook

A neural network fit to a restricted subspace of quasilinear gyrokinetic transport model calculations, relevant in the ITG regime, was carried out and applied as a transport model for integrated modelling. While the quasilinear model, QuaLiKiz, is 6 orders of magnitude faster than nonlinear simulations, the NN regression leads to a further 5 order of magnitude speedup. This model is thus real-time capable while still being based on first-principles, which is unprecedented. This model has been coupled to the CRONOS integrating modelling suite, and validated against a full QuaLiKiz simulation in the ITG regime. The model is also coupled to the real-time capable RAPTOR tokamak simulation code, and can model a 300 s ITER discharge within 10 s, with good agreement with previous modelling using CRONOS and the GLF23 transport model.

This opens up many new possibilities for real-time controller design and validation, scenario preparation and optimization, and real-time discharge supervision. Such models can be used to design controllers for the plasma profiles using model-based controller design methods (e.g. [34] or [35]). The transport model can be used in closed-loop simulations to validate the designed controllers. Recent work on plasma ramp-up trajectory optimization [36] was carried out with an *ad-hoc* transport model, and can now be improved using this first-principle-based transport model. Also, this transport model can be used in real-time simulations to verify the measured plasma evolution and warn a supervisory control system of any unexpected deviations during the discharge [37]. Specifically for ITER, the faster-than-real-time opens up the

possibility of (on-line) real-time optimization of the discharge evolution in response to such unexpected events.

While applications in the ITG transport regime are already feasible with this model, there remains much scope for expanding the number of input dimensions in the databases used for the fits, as well as employing slower yet more complete linear gyrokinetic codes for populating the database. Neural network topology complexity favourably scales linearly with input dimensionality. However, we estimate that uniform density population of the input dimensions, as carried out in this work, is feasible up to $N \sim 10$. This is due to constraints on the NN training and quasilinear database calculation times. For higher dimensionalities, a training set which captures the natural correlations of the input parameters is then vital. This can be done by basing the training sets on experimental parameters and reasonable extrapolations thereof. This is a feasible goal, as also evidenced in [17], and this work is ongoing.

Acknowledgments

This work is part of the research programme ‘Fellowships for Young Energy Scientists’ (YES!) of the Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organisation for Scientific Research (NWO). The authors greatly thank O. Meneghini, S. Smith, U. von Toussaint and P. Xanthopoulos for inspiring discussions.

References

- [1] Horton W. 1999 *Rev. Mod. Phys.* **71** 735
- [2] Doyle E.J. *et al* 2007 Progress in the ITER physics basis chapter 2: plasma confinement and transport *Nucl. Fusion* **47** S18

- [3] Lin Z. et al 2007 *Phys. Rev. Lett.* **99** 265003
- [4] Dannert T. and Jenko F. 2005 *Phys. Plasmas* **12** 072309
- [5] Casati A. et al 2009 *Nucl. Fusion* **49** 085012
- [6] Staebler G.M., Kinsey J.E. and Waltz R.E. 2007 *Phys. Plasmas* **14** 055909
- [7] Bourdelle C. et al 2007 *Phys. Plasmas* **14** 112501
- [8] Allen L. and Bishop C.M. 1992 *Plasma Phys. Control. Fusion* **34** 1291
- [9] Wakasa A. et al 2007 *Japan J. Appl. Phys.* **46** 1157
- [10] Lister J.B. and Schnurrenberger H. 1991 *Nucl. Fusion* **31** 1291
- [11] Clayton D.J. 2013 *Plasma Phys. Control. Fusion* **55** 095015
- [12] Svensson J., von Hellermann M. and König R.W.T. 1999 *Plasma Phys. Control. Fusion* **41** 315
- [13] Wroblewski D., Jahns G.L. and Leuer J.A. 1997 *Nucl. Fusion* **37** 725
- [14] Cannas B., Fanni A., Marongiu E. and Sonato P. 2004 *Nucl. Fusion* **44** 68
- [15] Vega J. et al 2014 *Nucl. Fusion* **54** 123001
- [16] Gaudio P., Murari A., Gelfusa M., Lupelli I. and Vega J. 2014 *Plasma Phys. Control. Fusion* **56** 114002
- [17] Meneghini O., Luna C.J., Smith S.P. and Lao L.L. 2014 *Phys. Plasmas* **21** 060702
- [18] Citrin J. et al 2012 *Phys. Plasmas* **19** 062305
- [19] Cottier P. et al 2014 *Plasma Phys. Control. Fusion* **56** 015011
- [20] Bourdelle C. 2015 Turbulent transport in tokamak plasmas: bridging theory and experiment CCSD <http://tel.archives-ouvertes.fr/tel-01113299>
- [21] Artaud J.F. 2010 *Nucl. Fusion* **50** 043001
- [22] Casati A. 2009 A quasi-linear gyrokinetic transport model for tokamak plasmas *PhD Dissertation* Universite de Provence France
- [23] Baiocchi B. et al 2015 *Plasma Phys. Control. Fusion* **57** 035003
- [24] Citrin J. et al 2014 *ITPA Transport and Confinement Topical Group (France, 20–22 October 2014)* www.iter.org/org/team/fst/itpa/tc
- [25] Bishop C.M. et al 1995 *Neural Networks for Pattern Recognition* (Oxford: Clarendon)
- [26] Haykin S. 1998 *Neural Networks: a Comprehensive Foundation* (Englewood Cliffs, NJ: Prentice Hall)
- [27] Bishop C. 1994 *Rev. Sci. Instrum.* **65** 1803
- [28] Dennis J.E. Jr and Schnabel R.B. 1996 *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* vol 16 (Philadelphia: SIAM)
- [29] MATLAB and Neural Network Toolbox Release 2012b The MathWorks, Inc., Natick, Massachusetts, USA
- [30] Felici F. and Sauter O. 2012 *Plasma Phys. Control. Fusion* **54** 025002
- [31] Kinsey J.E., Waltz R.E. and Candy J. 2005 *Phys. Plasmas* **12** 062302
- [32] Citrin J. et al 2010 *Nucl. Fusion* **50** 115007
- [33] Waltz R.E., Staebler G.M., Dorland W. and Hammett G.W. 1997 *Phys. Plasmas* **4** 2482
- [34] Maljaars E. et al 2015 *Nucl. Fusion* **55** 023001
- [35] Moreau D. et al 2013 *Nucl. Fusion* **53** 063020
- [36] van Dongen J., Felici F., Hogeweij G.M.D., Geelen P. and Maljaars E. 2014 *Plasma Phys. Control. Fusion* **56** 125008
- [37] Felici F. et al 2012 *Proc. 24th Int. Conf. on Fusion Energy (San Diego, CA, USA, 8–13 October 2012)* EX/P3-12 www-web.iaea.org/napc/physics/FEC/FEC2012/papers/169_EXP312.pdf