EUROFUSION WPJET1-CP(16) 15330

W Tang et al.

Big Data Machine Learning for Disruption Predictions

Preprint of Paper to be submitted for publication in
Proceedings of 26th IAEA Fusion Energy Conference

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Big Data Machine Learning for Disruption Predictions

William Tang[1], Matthew Parsons[1], Eliot Feibush[1], Jong Choi[2] , Tahsin Kurc[3] (USA)

(1) Princeton Plasma Physics Laboratory; (2) Oak Ridge National Laboratory; (3) Stony Brook University

Andrea Murari[4], Jesus Vega[4], Augusto Pereira[4], Isabel Nunes (IPFN, Portugal), and JET contributors[4]

[4](EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK)

Email of Main Author: tang@pppl.gov

Abstract. Building the scientific foundations needed to develop fusion power in a timely way can be facilitated not only by familiar "hypothesis-driven"/ first principles approaches but also by engaging modern big-data-driven statistical methods featuring machine learning (ML) -- an exciting R&D approach that is increasingly deployed in many scientific and industrial domains. An especially time-urgent and very challenging problem facing the development of a fusion energy reactor today is the need to deal reliably with large-scale major disruptions in magnetically-confined tokamak systems such as the Joint European Torus (JET) today and the burning plasma ITER device in the near future. Significantly improved methods of prediction with better than 95% predictive capability are required to provide sufficient advanced warning for disruption avoidance or mitigation strategies to be effectively applied before critical damage is done to the machine. The supervised machine learning classification technique featured in Support Vector Machines (SVM's) has been further advanced to this end and will be presented in this paper together with early results from ML studies utilizing multi-dimensional signal data to initiate the development of cross-machine-portable predictors. Working on it's repository of the most important and largest (nearly a half petabyte and growing) data base of fusion-grade plasmas, JET's statistical scientists have successfully deployed ML software interfaced with the large JET data base over the course of the past 7 years. This has produced encouraging results involving primarily the application of the SVM approach. The goals for the present investigations are to: (i) achieve greater predictive reliability by improving the physics fidelity of the classifiers within the "supervised" ML workflow; and (ii) establishing the cross-machine portability of the associated software beyond JET to other current tokamak systems and to ITER in the future. In order to so, it will be necessary to address the more realistic multi-dimensional, time-dependent, and much larger complex data instead of the simpler zero-dimensional, temporal data considered at present in all of the JET ML studies.

1. Introduction

The deployment of modern big-data-driven statistical methods featuring machine learning (ML) provides a powerful and exciting complement to traditional "hypothesis-driven"/ first principles approaches to delivering the scientific foundations needed to develop fusion power in a timely way. An especially time-urgent and very challenging problem facing the development of a fusion energy reactor today is the need to reliably mitigate and avoid large-scale major disruptions [1-3] in magnetically-confined tokamak systems including the Joint European Torus (JET), which achieved near-breakeven (0.8) conditions and the burning plasma ITER device currently under construction with the goal of exceeding breakeven by a factor of 10 or more. Disruptions are major macroscopic events generating massive thermal and electromagnetic loads as the plasma's thermal energy and current dissipate in a time on the order of tens of milliseconds (ms). The potential damage from severe impulsive heat loads to the surfaces of the machine can cost the ITER project hundreds of millions of dollars to remediate. Avoiding or at least mitigating them is critical because this device can sustain at most a very small number of full current disruptions. The international fusion mission must accordingly accelerate progress toward achieving the capability to reliably avoid such events with better than 95% predictive capability [2].

Promising tools from advanced ML methods for successfully preventing damage from disruptions come from the two associated categories: (i) *avoidance* as the plasma evolves toward an unstable state where a disruption would occur; and (ii) *mitigation* to minimize damage to the machine as the plasma heads toward a disruption that cannot be controlled to entirely avoid it. For example, mitigation techniques include ejecting large quantities of impurities into the plasma to radiate away some of the energy before it is expelled from the core onto the walls [4]. Of course, avoidance and mitigation strategies are only effective if an oncoming disruption is predicted enough in advance for the appropriate action to be taken.

While a large quantity of experimental disruption data exists, no "hypothesis-based"/first-principles models can currently capture the complex evolution of a disruption with sufficient accuracy and early enough warning (at least 30 ms before the event). In addressing this formidable challenge, increasing attention is being directed toward modern methods of statistical prediction via Machine Learning (ML) which have become increasingly prominent in many domains of science and industry. The ML Support Vector Machine (SVM) approach in particular has exhibited significant promise with analysis of zero-D time traces of the very large disruption-relevant JET data base [5,6]. In the current paper, we initiate for the first time consideration of multi-dimensional profile information as a predictor input. Initial results utilizing the 1D electron temperature profiles from electron-cyclotron emission (ECE) measurements from JET are provided. In addition, progress is reported from the testing of dimensionless parameters to assess the potential for developing a cross-machine-portable predictor.

More generally, this paper will present results from new developments in the testing of ML-based-methodologies – an exciting R&D approach that is increasingly deployed in many scientific and industrial domains -- to help provide much-needed guidance for disruption avoidance with focus on JET. Working on this repository of the most important and largest (nearly a half petabyte and growing) data base of fusion-grade plasmas, JET statistical scientists have successfully deployed ML software interfaced with the large JET data base over the course of the past several years [5,6]. This has produced encouraging results involving primarily the application of the support vector machine (SVM) approach [7,8]. The goals for the present investigations are to: (i) achieve greater predictive reliability by improving the physics fidelity of the classifiers within the "supervised" ML workflow; and (ii) establishing the portability of the associated software beyond JET to other current tokamak systems and to ITER in the future. In order to so, it will be necessary to address the more realistic multi-dimensional, time-dependent, and much larger complex data instead of the simpler zero-dimensional, temporal data considered at present in all of the JET ML studies. Associated challenges in delivering higher physics-fidelity classifiers needed to enable establishing portability of the predictive software when applied to MFE systems different than JET will be presented. In addition, it is expected that deployment of such improved ML software will need to be accordingly upgraded from current modern clusters to much more powerful leadership class supercomputers. This paper will include: (i) a description of a new workflow developed for our "supervised" ML SVM approach; (ii) results from associated applications – including profile information -- to the increasingly larger JET disruption-relevant data base; and (iii) highlights of current progress as well as key obstacles for this major big-data machine-learning problem. Scientific/technical progress achieved that will be discussed include:

• Systematic exploration (via MDS+ tree) of the JET disruption data base of associated signal and video data – enabled by formal approval of the EUROfusion JET leadership;

• Rewrite of SVM cross-validation routines now self-contained within Matlab, eliminating excessive file I/O and improving performance time by 100x;

• Description of PPPL's SVM software "Disruption Predictor Feature Developer (DPFD)" that is interfaced with the JET data base, including positive results from benchmarking vs. results obtained using JET's "Advanced Predictor of Disruptions (APODIS)" for zero-D time traces of both Carbon-wall and ITER-like-wall (ILW) cases.

• Description of collaborative studies with the CS/Applied Math experts at ORNL and Stony Brook University in exploring alternative methods for improving the selection of classifiers that have indicated: (i) clustering based on deterministic annealing is a promising approach in that it does not require a pre-determined number of clusters; and (ii) systematic examination of multi-dimensional image data (e.g., from ECEI measurements and also fast cameras) can provide additional information that improves prediction results since it contains spatial information that should be exploited.

• Discussion of new results from initial ML studies that include electron temperature profile information indicating that since latent patterns emerge when additional signals (different from the set used to date), significant improvements to the algorithms can be achieved.

2.0 SVM Model

The classification methodologies of "supervised" (as opposed to "unsupervised") machine learning provide an attractive interdisciplinary connection between advanced statistical analysis and an important domain physics application such as the problem of disruption prediction. A supervised ML approach is a natural choice because our plasma physics knowledge base provides us with insights into the development of appropriate classifiers for the ML workflow. For supervised classification we start with a set of data that we know corresponds to a disruptive plasma and another set of data that we know corresponds to a non-disruptive plasma. When combined, the data forms our training set. This training set in turn is used to generate a model that can classify new data as being either disruptive or non-disruptive. In the context of our disruption problem, the disruptive or non-disruptive state of a plasma can be described by a combination of signals known as a "feature vector." The objective of classification is to determine a decision function that will tell us whether a particular feature vector represents a disruptive or non-disruptive plasma. After using the training set to generate the decision function, a testing set is then used to assess the function's ability to classify new data. The SVM approach provides a very suitable methodology to tackle such a problem [7,8]. In a simple case the decision function is a hyper-plane which separates the disruptive and non-disruptive points in the feature space. However, this would only be appropriate if the data were linearly separable. Since this is clearly not the real-world situation, a mapping function is used to transform the data to a higher dimensional space where it theoretically can be linearly separated. The goal here is to find the hyper-plane that evenly divides the widest margin of separation between the disruptive and non-disruptive points. To do this, the SVM optimization process identifies a set of support vectors, which bound the maximized margin between the two states. Since it is still unlikely, even in this higher dimension, that there exists a hyper-plane capable of perfectly separating the data, a penalty is placed on points near the boundary. Decreasing the penalty allows a larger number of points to exist inside the margin, but is potentially more robust in making predictions with larger amounts of new data. Multiple values of this so-called "box constraint" parameter need to be tested to find the most accurate model.

3.0 Previous Work

The use of SVM's for tokamak disruption prediction by Cannas *et al*, [9] was motivated by the aim of identifying a statistical method more robust than neural networks at that time [10]. This involved 9 diagnostic signals from JET that included: plasma current, locked mode amplitude, radiated power, plasma density, input power, internal inductance, safety factor, poloidal beta, and plasma centroid vertical position. As a standard, each signal is normalized individually over the entire data set to the interval [0,1] so that no particular signal plays a more significant role in the analysis to be completed. The SVM training was carried out only with signals from disruptive shots in these investigations – working under the premise that a non-disruptive shot would look the same as the stable phase of a disruptive shot. These initial investigations of the application of SVM methodology were subsequently improved in the research carried out at CIEMAT for predicting disruptions in JET by Rattá *et al*. [11] who increased the set of diagnostic signals from the original nine to include time derivatives of the poloidal beta, the plasma internal inductance, and the stored diamagnetic energy. Additionally, instead of including the radiated power, the net power (total - radiated) was used. They examined 30 ms time windows of signal data with a 1 ms (interpolated) sample rate "as a compromise between time resolution and capability to show plasma tendencies," -- thereby forming feature vectors with 390 attributes each. The data set used included an equal number (220) of disruptive and non-disruptive shots, which were randomly sampled for the training and the remainder used for the testing. It is also significant to note that these studies introduced the idea of considering the time evolution of the feature vectors in the prediction; i.e., separate SVM models were trained to identify consecutive time intervals preceding a disruption. For example, a set of three models would be trained to identify disruptions [-120,-90] ms, [-90,-60] ms, and [-60,-30] ms before the disruption. After sequential data analysis by this first tier of SVM, the output of each was fed to a second tier SVM model. This second tier then gives a final decision on whether the plasma is disruptive or not. While the first-tier models are trained using the radial basis function (RBF) kernel, the second-tier model is trained with a linear kernel. After testing in sequences of up to 8 first-tier models, it was decided that a sequence of 3 predictors is optimal. This new predictor proved to be a significant advance since it was shown to outperform the mode-lock sensor on JET and can be considered the foundational step for the current Advanced Predictor of Disruptions (APODIS) software [11,12]. Genetic algorithms were employed as a method of feature extraction, and it was determined that in addition to the FFT representations used in earlier work, the mean of the signal samples on each 30 ms interval was also useful.

At present, APODIS [12,13] feature vectors now contain 14 elements with the signal set reduced to seven signals including plasma current, mode lock amplitude, plasma internal inductance, plasma density, time derivative of stored diamagnetic energy, radiated power, and total input power – with each signal having two representations [i.e., the mean value and the standard deviation of positive FFT components (excluding the first)]. The time interval of the feature vectors was also changed to 32 ms (still with 1 ms sampling). APODIS has reported that ~90% of disruptions are correctly identified and 5% of non-disruptive shots giving false alarms – all at 30 ms before an observed disruptive event.

4.0 Disruption Predictor Feature Developer (DPFD): A New Development Tool

In order to help efficiently assess the ability of diagnostic signals to enhance the SVM prediction of disruptions, a new code has recently been developed at PPPL. This Disruption Predictor Feature Developer (DPFD) is a set of Matlab scripts which amount to about 2000 lines of code, and therefore is very accessible for new developers and users and has recently been made available to those interested. These scripts can be used for everything from producing feature vectors from raw signal data to actually training and testing a two-tiered SVM disruption predictor. Since DPFD can be used to quickly evaluate whether or not the use of new features improves the performance of a predictor, there are two simplifications made to the APODIS approach. DPFD offers a reduced training time compared to APODIS by training with a much smaller data set. By selecting 3 disruptive and 3 non-disruptive training samples from each of several hundred disruptive shots, DPFD uses approximately 1000 samples for disruptive and non-disruptive classes each. APODIS uses about 1000 disruptive samples and 2 million non-disruptive samples from non-disruptive shots [12]. The SVM training process is very compute-intensive, and training with such a large data set requires a cluster to work on. The use of a small training set allows DPFD to be run on a personal computer in a few minutes. Positive results from benchmarking DPFD vs. results from APODIS are illustrated on Fig. 1.

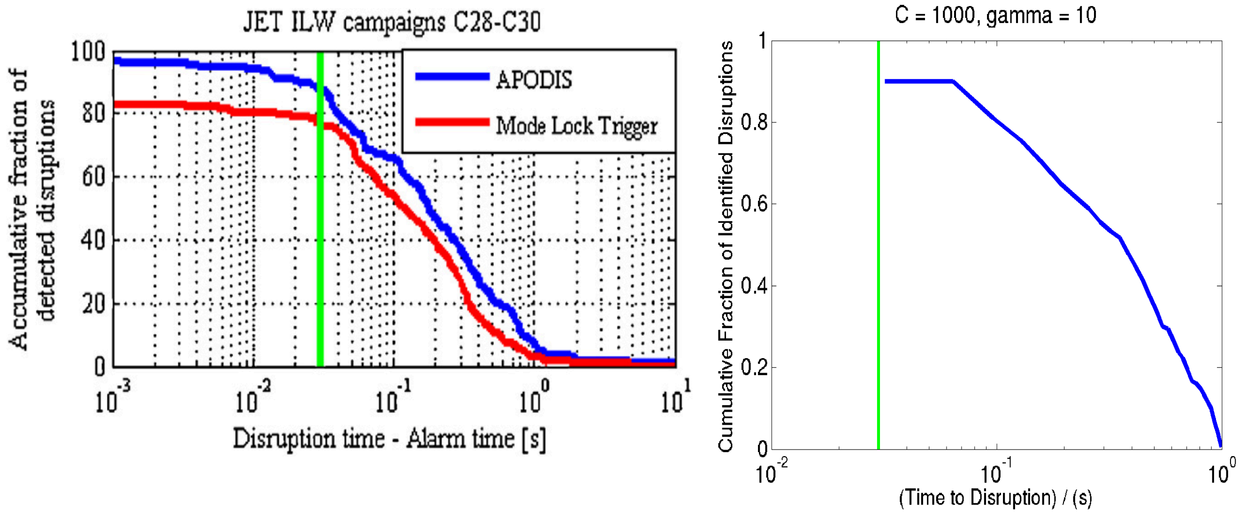

FIG. 1: New DPFD ML Software trained with ~2,000 data samples from JET with similar prediction performance to JET's APODIS code trained with ~2,000,000 data samples. Increased "false" alarms allow associated chains to be analyzed for useful additional information on disruption precursors.

5. Analysis of Dimensionless Parameters

To develop a machine-portable predictor, it is important to work with dimensionless quantities. Toward this objective, prediction performance comparisons are were initiated between the 5 normalized (N5) signals listed in Table 1 vs. the original 7 (O7) signals used by APODIS. We began with the O7 signals and tested 189 combinations of the two SVM free parameters. Selecting a combination with reasonable performance, we then used these parameters with the N5 signal set. In all cases, we trained with data from the Carbon-Fiber Composite (CFC) campaigns and tested with data from the ITER-like wall (ILW) campaigns.

Overall, it was observed that there was a degradation in the disruption prediction rate by about 20% in going from the original (O7) to the normalized N5 results -- and also a small increase in the false alarm rate. As a first test, this could be expected, since in moving from 7 signals to 5 signals, some of the important disruption features were likely to have been lost. To assess this more carefully, we began with the O7 signals and replaced them individually by their normalized counterparts. In doing so we were able to acquire a number of useful "lessons learned" conclusions about the normalization choices made in developing the N5 signals.

1. Mode Lock Amplitude Fraction (f_{ML})	MLA/B_T
2. Normalized Plasma Current (I_N)	$I_p/(aB_T)$
3. Plasma Internal Inductance (l_i)	l_i
4. Greenwald Density Fraction (f_{gw})	$n/(I_p/\pi a^2)$
5. Radiated Power Fraction (f_{rad})	$P_{rad}/(P_{in} - \dot{W}_{dia})$

TABLE 1: Normalized parameters used as inputs to DPFD

First, the mode lock normalization proved a poor choice that significantly degraded the predictor's ability to identify disruptions. In follow-on work, a more appropriate normalization based on the poloidal field will be investigated. Regarding the choice of normalization for the plasma current, neither significant improvement nor degradation of predictive performance was observed -- thereby suggesting that this may be an adequate/benign choice. Using the Greenwald density fraction as a substitute for both the density and plasma current proved to be a very favorable filter for disruptions. While the overall disruption prediction rate decreased by 25% (due to some of the other normalization choices just noted), the false alarms were strongly suppressed. Finally, results of these tests indicated that the radiated power fraction was not a good substitute for the three power signals used previously. This is not particularly surprising since the information from the previously used three signals is now contained in one. However, when we substituted the radiated power fraction for each of the three power signals individually, it was still found that important features of those signals are not captured by the fraction. This suggests that the radiated power fraction is not a viable substitution for the radiated power and that other dimensionless normalization choices need to be considered for selection.

6. 1D ECE Temperature Profile Studies

Since previous work in disruption prediction had only utilized 0D “time trace” signals as inputs, one of the key objectives in our work is to explore enhancing the physics content of the supervised ML classifiers via possible incorporation of higher dimensional profile information into a predictor. We began by examining 1D radial profiles of electron temperature obtained form the JET data base. In order to systematically develop features from this profile information for every shot examined, we used the temperature projected onto the major radius ranging from just inside the magnetic axis to about half the plasma radius. This radial domain was then split into 13 individual temperature measurements from which to produce features. In the same way as the previous signals were examined, we took the mean of the signal and also the standard deviation of the positive FFT components. This was done both for the temperature as well as the temperature gradient computed at 11 points in this domain.

Using this temperature information, we performed two types of tests. For the first type, we took the O7 signals and added a temperature and a temperature gradient measurement for all 13 temperatures and all 11 gradients. The second test was to take the O7 signals and add all 13 temperature and/or all 11 gradients at the same time. In the first attempt with the temperature measurements, it was found that none of the individual temperatures had any significant affect on the prediction performance. However, when all 13 measurements were added at once the false alarm rate spiked considerably. This is evidence of the so-called “curse of dimensionality,” where the addition of many new dimensions to a problem can decrease the statistical significance of any given data point and make it extremely challenging for the predictor to create a meaningful model. After this first attempt, a stricter box constraint parameter was applied and the test redone. With these new parameters it was found that in certain specific radial locations, the temperature and/or the temperature gradient *improved* the prediction performance as evidenced by increased suppression of alarms. Near future investigations will target more careful examination of the significance of these radial locations in terms of disruption characteristics.

7.0 Alternative Approaches

In addition to improving the SVM-centric R&D advances discussed in this paper, our overall work-scope for predictive disruption investigations, will increasingly engage very promising alternative approaches including the following: (i) Deep Learning [14]: Promising new results that use the recurrent neural net (RNN) version of deep learning have now been achieved with comparable success level to SVM -- with the associated progress benefiting from cross-benchmarking results against those from extensive SVM-based investigations based on the same set of scalar signals as used in our on-going SVM DPFD studies; (ii) Deterministic

Annealing Generative Topographic Mapping [15]: As a complementary approach to SVM and convolutional neural networks, DA-GTM has been developed as a method to *reduce* dimensionality and enable discovery of the abstract or hidden structures of high-D observational data. Progress made in applying DA-GTM to fusion disruption data have already yielded improved understanding of clustering features and accordingly provide insights into building improved classifiers dealing with higher-D signals. This provides an alternative approach whereby the outputs from DA-GTM can be used as an input for SVM; and (iii) Large-Scale Image Analysis [16]: Fast-image recognition statistical methods have been applied to fast camera JET data that have produced some promising early results that were enabled by leveraging successful R&D carried out by our collaborators on challenging biomedical imaging tasks. A more detailed discussion of these results from the three stimulating alternative approaches described here will be included in planned future reports.

8. Concluding Comments

Exciting opportunities involving the application of advanced statistical techniques such as machine learning (ML) hold great promise for accelerating progress in dealing with some of the biggest challenges facing fusion energy development – highlighted at present by disruptions. The combination of the increasing variety of ML methodologies enabled by modern technology together with the huge amount of experimental data (~ petabyte and beyond globally) clearly suggest that proper utilization of current and future supercomputing facilities is an exciting resource to leverage. Within the context of progress highlighted in this paper, there exist very promising physics-based classification opportunities in identifying higher dimensionality features to improve predictive capability as well as portability of the associated software. While previous work based on a combination of intuition and basic statistics have fuelled the existing development of features used for predicting disruptions via, for example, SVM methodology, there is significant head-room for improvement in delivering physics-based disruption identification methods to help provide a much-needed highly reliable toolset for ITER -- which can only tolerate a small number of disruptions during high-performance DT operation [17]. Looking into the near future, we believe that advances in disruption prediction should include the R&D goal/vision of leading to development of new experimental control paradigms that will be needed to monitor and maintain large-scale thermonuclear plasma stability in fusion reactors.

References

[1] FC Schuller. Disruptions in Tokamaks. Plasma Physics and Controlled Fusion, 37(11A):A135, 1995.

[2] PC DeVries, Survey of Disruption Causes at JET, Nuclear Fusion 51: 053018, 2011.

[3] TC Hender, JC Wesley, J Bialek, A Bondeson, AH Boozer, RJ Buttery, A Garofalo, TP Goodman, RS Granetz, Y Gribov, et al. MHD stability, operational limits and disruptions. Nuclear Fusion, 47(6):S128, 2007.

[4] EM Hollmann, PB Aleynikov, T Fülöp, DA Humphreys, VA Izzo, M Lehnen, VE Lukash, G Papp, G Pautasso, F Saint-Laurent, et al. Status of research toward the ITER disruption mitigation system. Physics of Plasmas (1994- present), 22(2):021802, 2015.

[5] GA Rattá, J Vega, A Murari, G Vagliasindi, MF Johnson, PC De Vries, and JET EFDA Contributors. An advanced disruption predictor for JET tested in a simulated real-time environment. Nuclear Fusion, 50(2):025005, 2010.

[6] S Dormido-Canto, J Vega, JM Ramírez, A Murari, R Moreno, JM López, A Pereira, JET-EFDA Contributors, et al. Development of an efficient real-time disruption predictor from scratch on JET and implications for ITER. Nuclear Fusion, 53(11):113001, 2013.

[7] William H Press. Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press, 2007.

[8] Chih-Chung Chang and Chih-J Lin. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1– 27:27, 2011.

[9] B Cannas, RS Delogu, A Fanni, P Sonato, MK Zedda, JET-EFDA contributors, et al. Support vector machines for disruption prediction and novelty detection at JET. Fusion Engineering and Design, 82(5):1124–1130, 2007.

[10] C.G. Windsor, G. Pautasso, C. Tichmann, R.J. Buttery, T.C. Hender, JET EFDA Contributors and the ASDEX Upgrade Team. A cross-tokamak neural network disruption predictor for the JET and ASDEX Upgrade tokamaks. Nuclear Fusion 45(5): 2005.

[11] GA Rattá, J Vega, A Murari, JET-EFDA Contributors, et al. Improved feature selection based on genetic algorithms for real time disruption prediction on JET. Fusion Engineering and Design, 87(9):1670–1678, 2012.

[12] J Vega, S Dormido-Canto, JM López, A Murari, JM Ramírez, R Moreno, M Ruiz, D Alves, R Felton, JET-EFDA Contributors, et al. Results of the JET real-time disruption predictor in the ITER-like wall campaigns. Fusion Engineering and Design, 88(6):1228–1231, 2013.

[13] J Vega, A Murari, S Dormido-Canto, R Moreno, A Pereira, A Acero, JET-EFDA Contributors, et al. Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks. Nuclear Fusion, 54(12):123001, 2014.

[14] Y LeCun, Y Bengio, and G Hinton. Deep Learning. Nature 521.7553, 436-444, 2015.

[15] JYChoi, J. Qiu, M. Pierce, and G. Fox. Generative topographic mapping by deterministic annealing. Procedia Computer Science, 1(1):47--56, 2010.

[16] G. Teodoro, T. Kurc, T. Pan, L. Cooper, J. Kong, P. Widener, et al., "Accelerating Large Scale Image Analyses on Parallel, CPU-GPU Equipped Systems," in 26th IEEE International Parallel and Distributed Processing Symposium (IPDPS), 2012, pp. 1093-1104. 2012.

[17] PC De Vries, G Pautasso, D Humphreys, M Lehnen, S Maruyama, JA Snipes, A Vergara, and L Zabeo. Requirements for triggering the ITER disruption mitigation system. Fusion Science and Technology, 69(2):471–484, 2016

Acknowledgements: The authors express their gratitude for the support provided for our research efforts at PPPL by the DOE-SC contract DE-AC02-09CH11466. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.