WHITE PAPER:
Machine Learning Application to Fusion Energy Predictive Disruption Studies

William Tang (wtang@princeton.edu)

Princeton University
Dept. of Astrophysical Sciences, Plasma Physics Section,
Princeton Plasma Physics Laboratory, and
Princeton Institute for Computational Science & Engineering

January, 2015

INTRODUCTION

As the global energy economy makes the transition from fossil fuels toward cleaner alternatives, fusion becomes an attractive potential solution for satisfying the growing needs. Fusion energy, which is the power source for the sun, can be generated on earth in magnetically-confined laboratory plasma experiments (called "tokamaks") when the isotopes of hydrogen (e.g., deuterium and tritium) combine to produce an energetic helium "alpha" particle and a fast neutron – with an overall energy multiplication factor of 450:1. Building the scientific foundations needed to develop fusion power in a timely way requires properly responding to extreme scale computing and big data challenges that will enable effective new predictive capabilities addressing the complex dynamics governing MFE (Magnetic Fusion Energy) systems -- including ITER, a multi-billion dollar international burning plasma experiment supported by 7 governments representing over half of the world's population. This will involve deployment of both familiar "hypothesis-driven"/first principles approaches as well as new big-data-driven statistical approaches featuring machine learning.

An especially time-urgent and challenging problem facing the development of a fusion energy reactor today is the need to reliably avoid or mitigate the onset of disruptions -- large-scale macroscopic events in tokamak plasmas that lead to a rapid termination of the discharges with accompanying major damage to the confining vessel. During disruptions, the machine is subjected to massive thermal and electromagnetic loads as the plasma's thermal energy and current dissipate in a time on the order of a millisecond. The associated damage from a small number of such events renders a sustainable tokamak fusion reactor impossible to achieve -- especially for ITER-scale burning plasma devices. The temporal sequence for these highly deleterious phenomena can be characterized as a "precursor" phase during which the plasma pressure and current build up to conditions that approach a threshold which triggers large-scale magneto-hydrodynamic (MHD) instabilities. Once exceeded, the "thermal quench" phase begins -- accompanied by a huge loss of thermal energy from the plasma to the first wall. A "current quench" then follows during which the plasma current rapidly falls to zero – with this rapid change in the electric current inducing massive magnetic loads on the machine. Mitigation techniques – such as introducing large amount of impurities into the plasma to radiate away some of the energy before it is expelled from the core region onto the walls – are only effective if deployed sufficiently in advance of the actual onset of the disruptions. It would be especially valuable if machine-learning-based methods could be developed and tested to help provide timely guidance for disruption avoidance in the Joint European Torus (JET) – located in the UK and the repository of the most important data-base of fusion-grade plasmas. JET had previously achieved the world-record "near breakeven" delivery of 10 MW of fusion power and is embarking on a new 5-year mission to revisit the high-powered Deuterium-Trititum (DT) campaign for the first time since 1997. Hypothesis-driven projections for disruption mitigation and avoidance have mostly been based on familiar computational Magnetohydrodynamic (MHD) models of limited physics fidelity that have not proven to be adequate for JET now or likely to be for ITER in the near future.

The benefit to Fusion Energy Science (FES) of machine-learning (ML) approaches to disruption prediction is the attractive potential for delivering the capability to significantly shorten the time to establish reliable operation in ITER at high performance. If ITER has to determine the operating boundary by trial and error, this would prolong the non-nuclear phase of this very expensive program and introduce additional schedule risk into the nuclear phase. It would be especially valuable if advanced statistically-based, large-data-driven methodologies could be developed and tested in a timely way to help provide in-transit guidance for disruption avoidance in JET and in currently operating long pulse international experiments (KSTAR in Korea, EAST in China, …) of relevance to ITER. As just noted, hypothesis-driven projections for disruption mitigation and avoidance have mostly been based on familiar computational MHD models of limited physics fidelity. For example, the influx or accumulation of impurities that can be disruption-relevant are not included in such models. Large-data-driven statistical methods can be viewed as complementary but a fundamentally different approach from hypothesis-driven methods. In view of the observed potential to date demonstrated in many other application domains (e.g., Google, bio-informatics, etc.) it can be concluded that progress toward reliable operation and efficient experimentation in fusion experiments would benefit significantly from more accurate data-driven statistical modeling of the integrated physics relevant to the onset of disruptions first in JET and then in ITER. It is also relevant to note at this point that the work on JET and NSTX – the "National Spherical Torus Experiment" located at the Princeton Plasma Physics Laboratory -- has focused on predicting the occurrence of the aforementioned current quench disruptions. On ITER, it is also important to predict the occurrence of disruptive events that trigger a thermal quench -- which in present-day machines may not always lead to a current quench. In other words, an important predictive tool development challenge would be to be to deliver reliable predictive software capable of providing adequate warning for a thermal quench on ITER.

In the past few years, there has been some promising machine-learning-based predictive software developed using statistical data mining approaches at JET on their disruption-relevant data-base. After first investigating classification & regression tree approaches [1], they moved on to support vector machine (SVM) methodologies [2]. Subsequent efforts have examined the selection of features that are used in the classifiers, using "genetic" algorithms to help select the sub-set of classifiers that have the strongest influence on the actual predictions. Associated publication [3, 4] have looked at predicting the type of disruptions -- an approach with promise of being more useful in comparing with conventional first-principles-based simulation results. It is also relevant to comment at this point that it might well be of interest to examine general unsupervised "deep learning" types of algorithms to select classifiers/features different from the aforementioned "supervised" genetic algorithms that have been deployed to date.

It is also valuable to view the prediction of disruption-relevant events from an experimentalist's perspective – for example, by Peter DeVries on JET [5] and by Stefan Gerhardt on NSTX [6]. In particular, DeVries' paper provides a comprehensive survey of possible causes of disruptions on JET, and Gerhardt's paper describes the prediction of disruptions based on diagnostic data from the high-beta (ratio of plasma to magnetic pressure) spherical torus experiment (NSTX). The disruptive threshold values on many signals were examined where: (i) raw diagnostic data were used as a signal for disruption prediction in some cases, while in others (ii) the deviations of the plasma data from simple models provided the information used to determine the proximity to disruption. Not surprisingly, there was no single signal or calculation and associated threshold value which could be found in these studies to form the basis for disruption prediction in NSTX. A novel means of combining multiple threshold tests was introduced in an algorithm that was applied to a database of ~2000 disruptions during the current flat-top phase of the discharges collected from three NSTX run campaigns. After proper tuning, this algorithm produced a false-positive rate of 2.8%, with a late plus missed warning rate of 3.7%, and thus a total failure rate of 6.5%. Many of these false positives were triggered by near-disruptive MHD events that could possibly have been disruptive in larger plasmas (such as JET) with more stored energy. However, the algorithm is less efficient at detecting the MHD events that actually trigger the disruption process of interest.

While the JET statistical team has achieved progress in predicting disruptions using machine-learning algorithms, there is significant room for improvement with respect both to the software as well as the more powerful hardware on which the possible new tools might need to be deployed. JET is accordingly very interested in collaboratively exploring advances/new machine-learning methodologies for improved disruption predictions – including the exciting new capabilities possibly enabled by access to more powerful computational hardware resources in the US. This will facilitate dealing with more realistic but more complex multidimensional data, instead of the much simpler zero-dimensional data considered at present in all of their studies. Targeted new capabilities will aim to: (i) improve the ability to handle much larger observational data sets, including the large image datasets from fast cameras; and (ii) building new predictors capable of incorporating the multidimensional features of the data together with possible associated access to powerful HPC hardware at, for example, the US DOE Leadership Computing Facilities (LCF's). This will require better understanding of the challenges of multidimensional analysis of the huge time-dependent disruption-relevant data base – first at JET and then including other tokamak facilities worldwide. As a rough estimate, the magnitude of the current (still growing !) multi-dimensional time-dependent signals in the JET data base can clearly exceed a petabyte. It will be a quite formidable but very interesting challenge to develop an optimal strategy/roadmap for how best to identify what information/features can be readily extracted and how to do so, including: (i) investigation of more advanced SVM methods – as well as the exploration of alternative approaches such as using Deterministic Annealing (DA) techniques [7, 8]; and (ii) improved feature extraction deployment of current Genetic Topographical Mapping (GTM) methods with consideration of signal representations different from the mean (averages over a chosen time segment at a given sampling rate) and standard deviation (std) approaches using Fast Fourier Transforms (FFT's).

A prioritized collaborative plan for this research activity is currently under development together with the JET project -- which *has formally agreed to provide the Princeton-based U.S. team with access to their huge disruption-relevant data base*. This will require further work on the following R& D topics:

(1) It is of interest to explore new ensemble and consensus methods that can combine a number of machine learning methods to address more multi-dimensional properties of the data than currently focused upon in the JET studies [1-4]. An associated issue is the cost of computation since multiple methods might be required, and the combination of the results might demand expensive computational operations. However, powerful open-science computational resource centers in the U.S. – such as the Oak Ridge National Laboratory's Leadership Class Facility (OLCF) -- can be a major dedicated asset in helping us further examine ensemble and consensus methods. As highlighted earlier, we will also explore alternative methods [e.g., Ref. 7] based on the deterministic annealing approach.

(2) Regarding signals used in predictors in the previous JET studies [1-4], about 13 signals were selected. There are of course possibilities to explore for improvement – especially since latent patterns may emerge when other signals (including multi-dimensional features) are used. Indeed, from an experimental applications perspective, other measurements will be needed to address, for example, the thermal quench physics noted earlier in this White Paper. In addition, as already discussed in previous studies [1-4], the algorithms that work well for a set of signals may not perform so well when other signals and other instruments are used. For example, future JET diagnostics will have capabilities to image the main chamber in the fusion instrument – deploying, for example, advanced diagnostics capabilities such as ECEI (electron cyclotron emission imaging). It would indeed be very interesting to systematically examine how multi-dimensional image data – with the deployment of modern visualization capabilities to help interpret/understand the more complex data -- could provide additional information that result in improved prediction results. In this regard, we focus on the fact that image data – for example from fast cameras – can potentially provide spatial information that will be very interesting to exploit. Since the current set of diagnostic signals studied to date [1-4], are mainly time series measurements of the various zero-D features (i.e., without spatial information) there is clearly significant headroom for improvement.

(3) Feature extraction and selection methods are also improving at a significant pace in the Computer Science/Applied Math community. As noted earlier in this White Paper, the previous JET studies (1-4) used two variables (mean and variance) calculated from the FFT's of the signals. We intend to examine how the additional features (for example, emerging from the multi-dimensional spatial characteristics of the signals) can be extracted with the goal of improving prediction accuracy. Also, use of methods such as linear regression for selection of signals and features could prove to be useful. In general, our supervised ML approach will draw on the expertise and experience of our fusion domain scientists in producing and analyzing plasma discharges to add a greater level of engagement of "human guidance" to the classification software.

In summary, a significant amount of research by JET scientists together with collaborating European institutions such as CIEMAT in Spain have successfully deployed machine learning software interfaced with the large JET data base over the course of the past 6 years at the JET facility. This has produced encouraging results involving primarily the application of the relatively straightforward support vector machine (SVM) approach. Our perspective is that this challenging problem is an attractive target for exploring whether more large-data-dependent, supervised machine learning methodologies -- either beyond the straightforward SVM approach or upgraded to deal with much larger, more complex multi-dimensional data -- can have a demonstrably positive impact in accelerating progress on this very important demonstration problem in the Fusion Energy Science application domain.

REFERENCES:

(1) “Unbiased and Non-Supervised Learning Methods for Disruption Prediction at JET,” A. Murari, et al., Nuclear Fusion, Vol. 49, 055028 (2009).
(2) “An advanced Disruption Predictor for JET Tested in a Simulated Real-time Environment,” G. A. Rattá, et al., Nuclear Fusion 50.2: 025005 (2010).
(3) “Clustering Based on the Geodesic Distance on Gaussian Manifolds for the Automatic Classification of Disruptions,” A. Murari, et al., Nuclear Fusion 53: 033006 (March, 2013).
(4) “Development of an Efficient Real-time Disruption Predictor on JET & Implications for ITER,” S. Dormido-Canto, J. Vega, et al., Nuclear Fusion 53: 113001 (March, 2013).
(5) “Survey of Disruption Causes at JET,” P. C. DeVries, Nuclear Fusion 51: 053018 (April, 2011).
(6) “Detection of Disruptions in the High-β Spherical Torus NSTX,” S.P. Gerhardt, et al., Nuclear Fusion 53.5: 063021 (May, 2013).
(7) “Generative Topographic Mapping by Deterministic Annealing,” J. Y. Choi, J. Qiu, M. Pierce, and G. Fox, Science Direct, Proc. Computer Science 00, 1-10 (2010).
(8) “Detailed Results of Evaluation of DAVS © Deterministic Annealing Clustering and its Application to LC-MS Data Analysis,”Geoffrey Fox, et al., Parallel Processing Letters, May 17, (2013).