2\textsuperscript{nd} IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis

Programme

Book of Abstracts

Boston, USA
30 May – 2 June, 2017
DISCLAIMER:

This is not an official IAEA publication. The views expressed do not necessarily reflect those of the International Atomic Energy Agency or its Member States.

The material has not undergone an official review by the IAEA. This document should not be quoted or listed as a reference.

The use of particular designations of countries or territories does not imply any judgement by the IAEA, as to the legal status of such countries or territories, of their authorities and institutions or of the delimitation of their boundaries.

The mention of names of specific companies or products (whether or not indicated as registered) does not imply any intention to infringe proprietary rights, nor should it be construed as an endorsement or recommendation on the part of the IAEA.
2nd IAEA Technical Meeting on
Fusion Data Processing, Validation and Analysis

30 May – 2 June, 2017
Boston, USA

IAEA Scientific Secretary

International Atomic Energy Agency
Vienna International Centre
Wagramer Straße 5, PO Box 100
A-1400 Vienna, Austria
NAPC Physics Section
Tel: +43-1-2600-21753
Fax: +43-1-26007
E-mail: Physics@iaea.org

Ms. Sehila Maria Gonzalez de Vicente

International Programme Advisory Committee

Chair: N. Howard (USA)
A. Dinklage (Germany), R. Fischer (Germany), S. Kajita (Japan), S. Kaye (USA), D. Mazon (France), D. McDonald (Eurofusion/EU), A. Murari (Eurofusion/EU), J. Vega (Spain), J. M. Noterdaeme (Germany), G. Verdoolaege (Belgium), M. Xu (China)

Local Organisation Committee

Nathan Howard
MIT Plasma Science and Fusion Center
175 Albany St.
NW17-168
Cambridge, MA 02139 USA
E-mail: nthoward@psfc.mit.edu
Tel: +1 617-253-4785

Jessica Coco
MIT Plasma Science and Fusion Center
77 Massachusetts Ave.
NW17-186
Cambridge, MA 02139 USA
E-mail: coco@psfc.mit.edu
Tel: +1 617 253 5400
Fax: +1 617 253 0627

Meeting Website:
https://nucleus.iaea.org/sites/fusionportal/Pages/List%20of-IAEA-Technical-Meeting-on-FDP.aspx
Topics

I. Uncertainty quantification

II. Model selection, validation & verification

III. Probability theory and statistical analysis

IV. Inverse problems & equilibrium reconstruction

V. Integrated data analysis

VI. Real time data analysis

VII. Machine learning

VIII. Signal/image processing & pattern recognition

IX. Experimental design and synthetic diagnostics

X. Data management
## Schedule

### Tuesday, 30 May, 2017

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30-9:00</td>
<td>Registration</td>
</tr>
<tr>
<td>9:00-9:05</td>
<td>Welcome and Opening Address</td>
</tr>
<tr>
<td>9:05-9:35</td>
<td><strong>Session 1: Inverse Problems and Equilibrium Reconstruction</strong></td>
</tr>
</tbody>
</table>
| 9:05-9:35    | **I-1:** Cianciosa M.  
*Uncertainty Analysis in 3D Equilibrium Reconstruction*                      |
| 9:35-10:00   | **O-1:** Xiao C.  
*2D magnetic field diagnosed by Laser-driven Ion-beam Trace Probe*             |
| 10:00-10:25  | **O-2:** Faugeras B.  
*Assimilation of polarimetry Stokes vector measurements in tokamak free-boundary equilibrium reconstruction with application to ITER* |
| 10:25-10:45  | **Coffee Break**                                                                                      |
| 10:45-11:10  | **O-3:** Howell E.C.  
*Development of a Non-Parametric Gaussian Process Model in V3FIT*                |
| 11:10-11:35  | **O-4:** Skvára V.  
*Robust Bayesian linear regression for Tokamak plasma boundary estimation*      |
| 11:35-12:00  | **O-5:** Stagner L.  
*Determining the Population of Individual Fast-ion Orbits using Generalized Diagnostic Weight Functions* |
<p>| 12:00-13:00  | <strong>Lunch Break</strong>                                                                                       |</p>
<table>
<thead>
<tr>
<th>Time</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:00-13:30</td>
<td>I-2: Chilenski M.A.</td>
<td>Bayesian inference of impurity transport coefficient profiles</td>
</tr>
<tr>
<td>14:20-14:45</td>
<td>O-8: Wang T.</td>
<td>Bayesian soft X-ray Tomography on Tore Supra and WEST</td>
</tr>
<tr>
<td>14:45-15:15</td>
<td>Discussion for Inverse Problems &amp; Equilibrium Reconstruction</td>
<td></td>
</tr>
<tr>
<td>15:15-15:35</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>15:35-16:00</td>
<td>O-9: Verdoolaege G.</td>
<td>Benchmarking robust regression techniques for global energy confinement scaling in tokamaks</td>
</tr>
<tr>
<td>16:00-16:25</td>
<td>O-10: Trask E.</td>
<td>Empirical Optimization with the Optometrist Algorithm: Randomization Coupled With Expert Interpretation</td>
</tr>
<tr>
<td>16:50-17:10</td>
<td>Discussion on Probability Theory &amp; UQ</td>
<td></td>
</tr>
</tbody>
</table>

**Chair:** G. Verdoolaege
## Wednesday, 31 May, 2017

### Session 3: Model Selection, Validation and Verification

**Chair: N. Howard**

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
<th>Speaker</th>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00-8:25</td>
<td><strong>O-12:</strong> Ernst D.R. Multichannel Validation of Gyrokinetic Simulations using a Synthetic Diagnostic for Doppler Backscattering based on Full-Wave Simulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8:25-8:50</td>
<td><strong>O-13:</strong> Rodriguez-Fernandez P. Validation of Quasilinear Transport Codes via Machine Learning Strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8:50-9:15</td>
<td><strong>O-14:</strong> Howard N.T. Validating Simulations of Multi-Scale Plasma Turbulence in in ITER-Relevant, Alcator C-Mod Plasmas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:15-9:40</td>
<td><strong>O-15:</strong> Vaezi P. An Improved Approach to Uncertainty Quantification for Plasma Turbulence Validation Studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:40-10:05</td>
<td><strong>O-16:</strong> Michoski C. Global Surrogates for the Upshift of the Critical Threshold in the Gradient for ITG Driven Turbulence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:05-10:25</td>
<td><strong>Coffee Break</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:25-10:50</td>
<td><strong>O-17:</strong> Jacobson C.M. Validation of MHD Models using MST RFP Plasmas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:50-11:15</td>
<td><strong>O-18:</strong> Reusch L.M. Model Validation for Quantitative X-ray Measurements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:15-11:40</td>
<td><strong>O-19:</strong> Mazon D. GEM tomographic measurements for WEST and validation strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:40-12:10</td>
<td><strong>Discussion for Model Selection Validation and Verification</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:10-13:00</td>
<td><strong>Lunch Break</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Speaker(s)</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>13:00-13:30</td>
<td>I-3: Rattá G.A.</td>
<td>AUG-JET cross-tokamak disruption predictor</td>
<td></td>
</tr>
<tr>
<td>13:55-14:20</td>
<td>O-21: Berkery J.</td>
<td>Disruption event characterization and forecasting of global and tearing mode stability for tokamaks</td>
<td></td>
</tr>
<tr>
<td>14:20-14:45</td>
<td>O-22: Rattá G.A. on behalf of Vega J.</td>
<td>Increased warning times in JET APODIS disruption predictor by using confidence qualifiers</td>
<td></td>
</tr>
<tr>
<td>14:45-15:05</td>
<td></td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>15:05-15:30</td>
<td>O-23: Granetz R.S.</td>
<td>Developing Universal Disruption Warning Algorithms Using Large Databases on Alcator C-Mod, EAST, and DIII-D</td>
<td></td>
</tr>
<tr>
<td>15:30-15:55</td>
<td>O-24: Rea C.</td>
<td>Exploratory machine-learning studies for disruption prediction using large databases on DIII-D</td>
<td></td>
</tr>
<tr>
<td>15:55-16:20</td>
<td>O-25: Ho A.</td>
<td>Tokamak profile database construction incorporating Gaussian process regression</td>
<td></td>
</tr>
<tr>
<td>16:45-17:10</td>
<td>O-27: Smith D.R.</td>
<td>Identification of ELM evolution patterns with unsupervised clustering of time-series similarity metrics</td>
<td></td>
</tr>
<tr>
<td>16:45-17:10</td>
<td></td>
<td>Discussion for Disruption Prediction and Pattern Recognition</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Presentation</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>8:00-8:30</td>
<td>I-4: Citrin J.</td>
<td><em>Realtime capable first principle transport modelling for tokamak prediction and control</em></td>
<td></td>
</tr>
<tr>
<td>8:30-8:55</td>
<td>O-28: Meneghini O.</td>
<td><em>Integrated infrastructure for the development of machine-learning models aimed at fusion applications</em></td>
<td></td>
</tr>
<tr>
<td>8:55-9:20</td>
<td>O-29: Grierson B.</td>
<td><em>Interpretive Analysis and Predictive Discharge Modeling with TRANSP</em></td>
<td></td>
</tr>
<tr>
<td>9:20-9:45</td>
<td>O-30: Jakubowski M.W.</td>
<td><em>Thermographic measurements of power loads to plasma facing components at Wendelstein 7-X</em></td>
<td></td>
</tr>
<tr>
<td>9:45-10:10</td>
<td>O-31: Puig Sitjes A.</td>
<td><em>Wendelstein 7-X near real-time image diagnostic system for plasma facing components protection</em></td>
<td></td>
</tr>
<tr>
<td>10:30-10:55</td>
<td>O-32: Logan N.</td>
<td><em>OMFIT Tokamak Profile Data, Fitting and Physics Analysis</em></td>
<td></td>
</tr>
<tr>
<td>10:55-11:20</td>
<td>O-33: Kostuk M.</td>
<td><em>Automatic between-pulse analysis of DIII-D experimental data performed remotely on a supercomputer at Argonne National Laboratory</em></td>
<td></td>
</tr>
<tr>
<td>11:45-12:10</td>
<td>O-35: Xu M.</td>
<td><em>Data processing on application of real-time system and validation of diagnostics in HL-2A</em></td>
<td></td>
</tr>
<tr>
<td>12:10-13:00</td>
<td></td>
<td><strong>Lunch Break</strong></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>13:00-13:30</td>
<td>I-5:</td>
<td>Salewski M. Integrated data analysis of fast-ion measurements by velocity-space tomography</td>
<td></td>
</tr>
<tr>
<td>13:30-13:55</td>
<td>O-36:</td>
<td>Wojenski A. Advanced real-time data quality monitoring concept for GEM detector based SXR plasma diagnostics</td>
<td></td>
</tr>
<tr>
<td>13:55-14:20</td>
<td>O-37:</td>
<td>Nornberg M.D. Incorporating beam attenuation calculations into an Integrated Data Analysis model of plasma impurity content</td>
<td></td>
</tr>
<tr>
<td>14:20-14:50</td>
<td></td>
<td>Discussion for Real Time and Integrated Data Analysis</td>
<td></td>
</tr>
<tr>
<td>14:50-15:10</td>
<td></td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>15:10-15:40</td>
<td>I-6:</td>
<td>Kajita S. Assessment and mitigation of wall light reflection in ITER by ray tracing</td>
<td></td>
</tr>
<tr>
<td>15:40-16:05</td>
<td>O-38:</td>
<td>Yu Y. Data analysis and effect corrections of Phase Contrast Imaging diagnostic on HL-2A tokamak</td>
<td></td>
</tr>
<tr>
<td>16:05-16:30</td>
<td>O-39:</td>
<td>Liu Y. Synthetic diagnostic for interpreting the ECE spectrum in LHW-heated plasmas on EAST</td>
<td></td>
</tr>
<tr>
<td>16:30-16:55</td>
<td>O-40:</td>
<td>Liu C. Explanation of prompt growth of ECE signal in tokamak runaway electron experiments using ECE synthetic diagnostic</td>
<td></td>
</tr>
<tr>
<td>16:55-17:15</td>
<td></td>
<td>Discussion for Experimental Data Analysis and Synthetic Diagnostics</td>
<td></td>
</tr>
<tr>
<td>18:30</td>
<td></td>
<td>Banquet Dinner at Top of the Hub – Boston, MA</td>
<td></td>
</tr>
</tbody>
</table>
# Friday, 2 June, 2017

## Session 7: Data Management/Handling

**Chair:** M. Xu

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
<th>Speaker/Sponsor</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00-8:30</td>
<td><strong>I-7:</strong> Marzouk Y. <em>Computational advances for Bayesian inference and optimal experimental design</em></td>
<td></td>
</tr>
<tr>
<td>8:30-8:55</td>
<td><strong>O-41:</strong> Pinches S. <em>ITER’s Integrated Modelling Infrastructure and Strategy for Data Analysis and Validation</em></td>
<td></td>
</tr>
<tr>
<td>8:55-9:20</td>
<td><strong>O-42:</strong> De Witt S. on behalf of Lupelli I. <em>The SAGE Project: A paradigm shift in the Storage Systems for Data Centric Computing</em></td>
<td></td>
</tr>
<tr>
<td>9:20-9:45</td>
<td><strong>O-43:</strong> Smith S. <em>OMFIT (One Modeling Framework for Integrated Tasks): An Efficient Community Driven Integrated Modeling Framework</em></td>
<td></td>
</tr>
<tr>
<td>9:45-10:10</td>
<td><strong>O-44:</strong> Greenwald M. <em>Navigational Data Management</em></td>
<td></td>
</tr>
<tr>
<td>10:10-10:30</td>
<td><strong>Coffee Break</strong></td>
<td></td>
</tr>
<tr>
<td>10:30-10:55</td>
<td><strong>O-45:</strong> Zabeo L. <em>Diagnostic Data Handling in the PCS</em></td>
<td></td>
</tr>
<tr>
<td>10:55-11:20</td>
<td><strong>O-46:</strong> Emoto M. <em>Improvement of Automatic Physics Data Analysis Environment for the LHD Experiment</em></td>
<td></td>
</tr>
<tr>
<td>11:20-11:45</td>
<td><strong>O-47:</strong> Kim J. <em>Development of unified data analysis in KSTAR</em></td>
<td></td>
</tr>
<tr>
<td>11:45-12:10</td>
<td><strong>O-48:</strong> De Witt S. <em>A Comparison of Data Management Techniques Across Different Science Disciplines</em></td>
<td></td>
</tr>
<tr>
<td>12:10-12:40</td>
<td><strong>Discussion for Data Management &amp; Meeting Closing</strong></td>
<td></td>
</tr>
</tbody>
</table>
Abstracts

List of Invited Orals:

I-1: Cianciosa M., *Uncertainty Analysis in 3D Equilibrium Reconstruction*

I-2: Chilenski M.A., *Bayesian inference of impurity transport coefficient profiles*

I-3: Rattá G.A., *AUG-JET cross-tokamak disruption predictor*

I-4: Citrin J., *Realtime capable first principle transport modelling for tokamak prediction and control*

I-5: Salewski M., *Integrated data analysis of fast-ion measurements by velocity-space tomography*

I-6: Kajita S., *Assessment and mitigation of wall light reflection in ITER by ray tracing*

I-7: Marzouk Y., *Computational advances for Bayesian inference and optimal experimental design*
In the forward equilibrium problem, the (assumed) known properties of an MHD equilibrium model are used to predict the expected experimental observations of the plasma. For example, with a known solution of the Grad-Shafranov equation for an axisymmetric plasma, along with the position of a magnetic diagnostic, a straightforward Biot-Savart integration over the plasma volume predicts the magnetic flux due to the plasma current through the magnetic diagnostic loop. Equilibrium Reconstruction is the inverse problem, where experimental observations are known, and the parameters specifying the MHD equilibrium model are inferred from the observations. A Bayesian analysis, along with convenient assumptions, yields a least squares minimization problem, where the parameter space is searched for those equilibrium model parameters most likely to yield the observed experimental signals.

Errors are unavoidable in the observed signals. They may be due to systematic problems, random fluctuations, or deficiencies in the model. An estimate of the uncertainty in the observed signals is a necessary input to the least-squares formulation, and allows for different weightings of the various observational signals. This input uncertainty in the observed signals, causes an associated uncertainty in the reconstructed results. Tracking this uncertainty from the signal inputs to the reconstructed results is of vital importance in validating a reconstruction. We describe the propagation of this assumed observational uncertainty from the space of observed signals to the parameter space, and from the parameter space to the computation of other desired equilibrium properties.

V3FIT is a reconstruction code built around the VMEC 3D equilibrium code. V3FIT finds the most-likely parameters using a quasi-newton optimization method to search parameter space. In so doing, V3FIT must compute an approximation to the Jacobian - the partial derivatives of the signals with respect to the parameters. Using this Jacobian, uncertainty in the input signals is propagated into parameter space, providing a confidence interval on the optimal parameters. Uncertainty in reconstructed parameters is further propagated into model derived quantities using the equilibrium solution.

To demonstrate the validity of this uncertainty propagation, whole shot equilibrium reconstructions where performed using different models. The random variation (noise) in the reconstructed parameter is compared against the propagated uncertainty. A Bayesian model selection analysis will be used to quantify the better of two different model parameterizations. Examples of this propagated uncertainty including the vital role it played in the confirmation of the first instance of helical core reconstructed in a weakly 3D tokamak, will be presented.

Work supported under U.S. DOE Cooperative Agreement DE-AC05-00OR22725 and U.S. DOE Award number DE-FG02-03ER54692.
Recent attempts to obtain maximum likelihood estimates of impurity transport coefficient profiles \( D_Z, V_Z \) in the Alcator C-Mod tokamak using standard techniques have failed to yield unique solutions. In order to fix these issues, we have recast the problem in the Bayesian framework and applied the MultiNest algorithm [1] to perform both parameter estimation and model selection. For a given level of complexity in the inferred \( D_Z, V_Z \) profiles, MultiNest delivers a set of samples from the posterior distribution for \( D_Z, V_Z \) (used to compute the \( D_Z, V_Z \) profiles and their uncertainties) as well as an estimate of the model evidence (used to select the appropriate level of complexity). Initial work using synthetic data shows that model selection is the critical step which has traditionally been neglected: the model evidence correctly prefers more complicated models which match the true \( D_Z, V_Z \) profiles better, despite the existence of simpler models with comparable goodness-of-fit. Furthermore, use of Gaussian process regression (GPR) to efficiently propagate the uncertainties in the \( n_e \) and \( T_e \) profiles through the analysis reveals that these uncertainties (previously believed to be dominant for Alcator C-Mod) have a negligible impact. Ongoing work to apply these techniques to experimental data will yield an improved understanding of impurity transport and how well gyrokinetic simulations match the impurity channel.

References:
I-3: AUG-JET cross-tokamak disruption predictor

G.A.Rattá¹, J.Vega¹, A. Murari², the EUROfusion MST1 Team* and JET Contributors†

EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK
¹Laboratorio Nacional de Fusión. CIEMAT, Madrid, Spain
²Consorzio RFX (CNR, ENEA, INFN, Università di Padova, Acciaierie Venete Sp A) Corso Stati Uniti 4,35127 Padova. Italy

The application of Machine Learning (ML) techniques to predict disruptions has shown potential to considerably improve the detection rates and the warning times in JET [1] and other tokamaks. However, ML predictors learn from past events, which implicate an already stored database to develop them. Therefore, a significant problem arises at the time of developing ML-based systems for ITER. In this work and to tackle this problem, a Genetic Algorithms-optimized (GAs) predictor based on a previous study [1] was trained using, initially, only AUG data and tested with a wide database of JET. This "smaller to larger" tokamak approach is meant as a test for future extrapolation of this technique to ITER.

The outcomes of the direct application of the cross-predictor derived in a 44.47% of false alarms and more than a 65% of premature alarms, which indicates the need of some information about the target device to achieve reasonable performance.

Then, in a second approach, a new predictor was trained with AUG database plus one disruptive and one non-disruptive pulse of JET. The procedure was repeated twice to ensure the robustness of the results, since the GAs evolves towards different solutions in every run. In both runs the final cross-predictions (over the chronologically first 500 shots after the training) reached an average of a ~94% of total detected disruptions (~90% of them with anticipation times higher to 10 ms). The false alarms in that period were, on average, of 12.95%. An analysis of the ageing effect (assessing the decrease in the model's accuracy with time after the last training) has been performed for a database that contains more than 5000 discharges.

Finally, and to provide a solution to the predictor's ageing, an adaptive approach, based on the system's retraining each time a disruption is missed, has been introduced and applied.

References:

* See the author list of “Overview of progress in European Medium Sized Tokamaks towards an integrated plasma-edge/wall solution” by H. Meyer et al., to be published in Nuclear Fusion Special issue: Overview and Summary Reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17-22 October 2016).
† See the author list of “Overview of the JET results in support to ITER” by X. Litaudon et al. to be published in Nuclear Fusion Special issue: overview and summary reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17-22 October 2016).
I-4: Realtime capable first principle transport modelling for tokamak prediction and control

J. Citrin\textsuperscript{1}, T. Aniel\textsuperscript{2}, C. Bourdelle\textsuperscript{2}, Y. Camenen\textsuperscript{3}, V. Dagnelie\textsuperscript{1}, H. Doerk\textsuperscript{4}, F. Felici\textsuperscript{5}, A. Ho\textsuperscript{1}, D. Hogeweij\textsuperscript{1}, R. Nouailletas\textsuperscript{2}, K. van de Plassche\textsuperscript{4,6}, G. Verdooldaege\textsuperscript{7,8}, D. van Vugt\textsuperscript{6}

\textsuperscript{1}DIFFER - Dutch Institute for Fundamental Energy Research, Eindhoven, the Netherlands
\textsuperscript{2}CEA, IRFM, F-13108 Saint Paul Lez Durance, France
\textsuperscript{3}CNRS, Aix-Marseille Univ., PIIM UMR7345, Marseille, France
\textsuperscript{4}Max Planck Institute for Plasma Physics, Boltzmannstr. 2, Garching, Germany
\textsuperscript{5}Eindhoven University of Technology, Department of Mechanical Engineering, Control Systems Technology Group, PO Box 513, 5600 MB Eindhoven, The Netherlands
\textsuperscript{6}Science and Technology of Nuclear Fusion, Department of Applied Physics, Eindhoven University of Technology, Eindhoven, The Netherlands
\textsuperscript{7}Department of Applied Physics, Ghent University, 9000 Gent, Belgium
\textsuperscript{8}Ecole Royale Militaire / Koninklijke Militaire School (LPP – ERM/KMS), Brussels, Belgium

Present-day first-principle-based tokamak turbulent transport models are characterised by significant computational evaluation times. This reduces viability for wide-scale scenario optimization and realtime applications. Nevertheless, an accurate predictive model for turbulent transport fluxes is vital for interpretation of present-day tokamak experiments, and extrapolation to future machines. We aim to circumvent these conflicting constraints of accuracy and tractability, by applying neural network machine learning methods to emulate existing reduced turbulence models without a loss in model accuracy. The evaluation time of a trained neural network is orders of magnitude faster than the original model. A proof-of-principle of this concept is developed [1], based on a 4-input-dimension (4D) emulation of the QuaLiKiz [2] turbulent transport model, and is coupled to the RAPTOR [3] fast tokamak simulator. Based on this success, the next generation of neural network models is being produced, with a 9D QuaLiKiz database forming a training space with \(\sim 3 \times 10^8\) flux points. Higher input dimensions are vital for capturing extended turbulence regimes and plasma scenarios. For even higher dimensionality (\(\sim 20D\)), corresponding to the full set of local plasma parameters, the training sets must be populated in a subspace capturing the natural correlations observed in experiments. An extensive multi-machine profile database is being constructed for this purpose [4]. Applying the neural network models in fast tokamak simulators such as RAPTOR would open up a plethora of possibilities and innovation in tokamak realtime controller design and validation, scenario preparation, and discharge optimization.

References:
I-5: Integrated data analysis of fast-ion measurements by velocity-space tomography

M. Salewski\textsuperscript{1}, B. Geiger\textsuperscript{2}, M. Nocente\textsuperscript{3,4}, A.S. Jacobsen\textsuperscript{2}, F. Binda\textsuperscript{5}, C. Cazzaniga\textsuperscript{6}, G. Ericsson\textsuperscript{5}, J. Eriksson\textsuperscript{5}, G. Gorini\textsuperscript{3,4}, W.W. Heidbrink\textsuperscript{7}, C. Hellesen\textsuperscript{5}, A. Hjalmarsson\textsuperscript{5}, V.G. Kiptily\textsuperscript{8}, T. Koskela\textsuperscript{9,10}, S.B. Korsholm\textsuperscript{1}, T. Kurki-Suonio\textsuperscript{9}, F. Leipold\textsuperscript{1}, J. Madsen\textsuperscript{1}, D. Moseev\textsuperscript{1}, S.K. Nielsen\textsuperscript{1}, J. Rasmussen\textsuperscript{1}, M. Schneider\textsuperscript{12}, P.A. Schneider\textsuperscript{2}, S.E. Sharapov\textsuperscript{8}, L. Stagner\textsuperscript{7}, M. Stejner\textsuperscript{1}, G. Tardini\textsuperscript{2}, M. Tardocchi\textsuperscript{4}, the ASDEX Upgrade Team, the EUROfusion MST1 Team and JET Contributors

\textsuperscript{1} Technical University of Denmark, Dep. of Physics, Kgs. Lyngby, Denmark
\textsuperscript{2} Max-Planck-Institut für Plasmaphysik, Garching, Germany
\textsuperscript{3} University of Milano Bicocca, Dep. of Physics, Milano, Italy
\textsuperscript{4} Istituto di Fisica del Plasma, Consiglio Nazionale delle Ricerche, Milano, Italy
\textsuperscript{5} Dep. of Physics and Astronomy, Uppsala University, Sweden
\textsuperscript{6} ISIS Facility, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Didcot, UK
\textsuperscript{7} UC Irvine, Dep. of Physics and Astronomy, CA, US
\textsuperscript{8} CCFE, Culham Science Centre, Abingdon, Oxon, UK
\textsuperscript{9} Aalto University, Dep. of Applied Physics, Aalto, Finland
\textsuperscript{10} NERSC, Lawrence Berkeley National Laboratory, CA, USA
\textsuperscript{11} Max-Planck-Institut für Plasmaphysik, Greifswald, Germany
\textsuperscript{12} ITER Organization, St Paul-lez-Durance Cedex, France

Integrated data analysis combines measurements from different diagnostics to jointly measure plasma parameters of interest such as temperatures, densities, and drift velocities. Integrated data analysis of fast-ion measurements has long been hampered by the complexity of fast-ion velocity distribution functions. This has recently been overcome by tomographic inversion techniques. In these methods 2D images of the functions consisting of a few hundreds or thousands of pixels are reconstructed using the available fast-ion measurements. Here we present an overview and current status of this emerging technique. We focus on integrated data analysis of the major fast-ion diagnostics fast-ion D-alpha spectroscopy, collective Thomson scattering, gamma-ray and neutron emission spectrometry and neutral particle analyzers. Attention will be paid to the combination of diagnostics with uncertain calibration. We present examples of measured fast-ion velocity distribution functions in plasmas heated by neutral beam injection and ion cyclotron resonance heating at ASDEX Upgrade and JET. In sawtoothing plasmas no evidence for effects of sawteeth on energetic particles with pitches around zero is found, whereas large fractions of energetic particles with pitches close to one are ejected from the plasma center. The tomographic inversion at JET confirms the theoretical prediction that few particles are accelerated to energies larger than about 2 MeV in a third harmonic ion cyclotron resonance heating scenario.

\textsuperscript{*} See the author list of “Overview of the JET results in support to ITER” by X. Litaudon et al. to be published in Nuclear Fusion Special issue: overview and summary reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17-22 October 2016)
I-6: Assessment and mitigation of wall light reflection in ITER by ray tracing

Shin Kajita¹, Evgeny Veshchev², Maarten De Bock², Robin Barnsley², Manfred von Hellemann²

¹IMaSS, Nagoya University, Nagoya 464-8603, Japan
²ITER Organization, Route de Vinon-sur-Verdon, CS 90 046, 13067 St. Paul Lez Durance Cedex, France

In ITER, reflection of photons on vacuum vessel will make parasitic signals (stray light) for optical diagnostics, and the issue is more serious than present tokamak devices, because the wall is fully metal and the emission profiles will have higher contrast. In ITER H-alpha spectroscopy, brighter divertor emission can form two orders of magnitude greater stray light than the actual signal [1]. In charge exchange recombination spectroscopy (CXRS), in addition to an increase in the photon noise, bright divertor emission may lead to saturation of the detector, and the CXRS signal in the edge region may be considerable superimposed on the weak core signal by reflection [2]. In this study, to estimate and mitigate the effect of the stray light in ITER, ray tracing simulations are performed using a software LightTools.

To reduce the stray light effects, ray transfer matrix [3] were constructed and the emission profiles in the divertor region were reconstructed by considering the reflection on the walls. It was found that better performance was obtained when the sources around the strike points were allocated in detailed manner and the stray light can be reduced significantly for all of Be, He and Ne spectroscopy. Concerning the reflection of the CXRS signal, the diffusive reflection component was found to be important. Reflected low temperature spectrum can be superimposed on the higher temperature core signal by ~10% in maximum. However, since the core signal has large photon noise, it is likely that the reflection can be buried in the photon noise.

“The views and opinions expressed herein do not necessarily reflect those of the ITER Organization.”

References:
I-7: Computational advances for Bayesian inference and optimal experimental design

Youssef Marzouk

Massachusetts Institute of Technology
Department of Aeronautics & Astronautics
Center for Computational Engineering, Statistics & Data Science Center
http://uqgroup.mit.edu

E-mail of Corresponding Author: ymarz@mit.edu

Bayesian statistics provides a coherent framework for integrating experimental data with complex mathematical models, for quantifying uncertainty in parameter estimates and model predictions, and for optimally planning experiments or selecting observations. Yet the computational expense associated with Bayesian approaches can present significant bottlenecks in large-scale systems. This talk will discuss new sampling algorithms and approximation methods designed to make the application of rigorous Bayesian methods more computationally tractable. We will focus on two problems frequently encountered in practice: (1) inference with computationally intensive models, and (2) optimal experimental design.

For the first problem, we introduce a new framework for accelerating posterior sampling, borrowing ideas from deterministic approximation theory and derivative-free optimization. Previous efforts at integrating approximate models into inference typically sacrifice either the sampler’s exactness or efficiency; our approach addresses these limitations by exploiting useful convergence characteristics of local approximations. In particular, we develop an asymptotically exact Markov chain Monte Carlo (MCMC) scheme using local approximations, and describe variations of the algorithm that exploit parallel computation. Results suggest that when the likelihood has some regularity, the number of model evaluations per MCMC step can be greatly reduced without biasing the Monte Carlo average.

For the second problem, we propose an information theoretic framework for “focused” experimental design with nonlinear models, with the goal of maximizing information gain in targeted subsets of model parameters. We also present a new layered multiple importance sampling scheme for estimating expected information gain in this focused setting. This scheme yields significant reductions in estimator bias and variance for a given computational effort, making optimal design more tractable for non-Gaussian and computationally intensive problems.
**List of Regular Orals:**

**O-1:** Xiao C., *2D magnetic field diagnosed by Laser-driven Ion-beam Trace Probe*

**O-2:** Faugeras B., *Assimilation of polarimetry Stokes vector measurements in tokamak free-boundary equilibrium reconstruction with application to ITER*

**O-3:** Howell E.C., *Development of a Non-Parametric Gaussian Process Model in V3FIT*

**O-4:** Skvára V., *Robust Bayesian linear regression for Tokamak plasma boundary estimation*

**O-5:** Stagner L., *Determining the Population of Individual Fast-ion Orbits using Generalized Diagnostic Weight Functions*

**O-6:** Stankunas G., *Accurate Determination of Radiated Power Density Profile Using Bolometer Data for DT Baseline Scenario at JET*

**O-7:** Ferreira D. R., *Full-pulse tomographic reconstruction with deep neural networks*

**O-8:** Wang T., *Bayesian soft X-ray Tomography on Tore Supra and WEST*

**O-9:** Verdoolaege G., *Benchmarking robust regression techniques for global energy confinement scaling in tokamaks*

**O-10:** Trask E., *Empirical Optimization with the Optometrist Algorithm: Randomization Coupled With Expert Interpretation*

**O-11:** Fujii K., *Machine Learning of Noise for LHD Thomson Scattering System*

**O-12:** Ernst D.R., *Multichannel Validation of Gyrokinetic Simulations using a Synthetic Diagnostic for Doppler Backscattering based on Full-Wave Simulations*

**O-13:** Rodriguez-Fernandez P., *Validation of Quasilinear Transport Codes via Machine Learning Strategies*

**O-14:** Howard N.T., *Validating Simulations of Multi-Scale Plasma Turbulence in in ITER-Relevant, Alcator C-Mod Plasmas*

**O-15:** Vaezi P., *An Improved Approach to Uncertainty Quantification for Plasma Turbulence Validation Studies*

**O-16:** Michoski C., *Global Surrogates for the Upshift of the Critical Threshold in the Gradient for ITG Driven Turbulence*

**O-17:** Jacobson C.M., *Validation of MHD Models using MST RFP Plasmas*

**O-18:** Reusch L.M., *Model Validation for Quantitative X-ray Measurements*

**O-19:** Mazon D., *GEM tomographic measurements for WEST and validation strategies*
O-20: Kates-Harbeck J., Disruption Forecasting in Tokamak Fusion Plasmas using Deep Recurrent Neural Networks

O-21: Berkery J., Disruption event characterization and forecasting of global and tearing mode stability for tokamaks

O-22: Vega J., Increased warning times in JET APODIS disruption predictor by using confidence qualifiers

O-23: Granetz R.S., Developing Universal Disruption Warning Algorithms Using Large Databases on Alcator C-Mod, EAST, and DIII-D

O-24: Rea C., Exploratory machine-learning studies for disruption prediction using large databases on DIII-D

O-25: Ho A., Tokamak profile database construction incorporating Gaussian process regression

O-26: Churchill R.M., Finding structure in large datasets of particle distribution functions using unsupervised machine learning

O-27: Smith D.R., Identification of ELM evolution patterns with unsupervised clustering of time-series similarity metrics

O-28: Meneghini O., Integrated infrastructure for the development of machine-learning models aimed at fusion applications

O-29: Grierson B., Interpretive Analysis and Predictive Discharge Modeling with TRANSP

O-30: Jakubowski M.W., Thermographic measurements of power loads to plasma facing components at Wendelstein 7-X

O-31: Puig Sitjes A., Wendelstein 7-X near real-time image diagnostic system for plasma facing components protection

O-32: Logan N., OMFIT Tokamak Profile Data, Fitting and Physics Analysis

O-33: Kostuk M., Automatic between-pulse analysis of DIII-D experimental data performed remotely on a supercomputer at Argonne National Laboratory

O-34: Kocan M., ITER Wide Angle Viewing System: Synthetic Measurements and Challenges of the Real-time Data Processing

O-35: Xu M., Data processing on application of real-time system and validation of diagnostics in HL-2A

O-36: Wojenski A., Advanced real-time data quality monitoring concept for GEM detector based SXR plasma diagnostics

O-37: Nornberg M.D., Incorporating beam attenuation calculations into an Integrated Data Analysis model of plasma impurity content
O-38: Yu Y., *Data analysis and effect corrections of Phase Contrast Imaging diagnostic on HL-2A tokamak*

O-39: Liu Y., *Synthetic diagnostic for interpreting the ECE spectrum in LHW-heated plasmas on EAST*

O-40: Liu C., *Explanation of prompt growth of ECE signal in tokamak runaway electron experiments using ECE synthetic diagnostic*

O-41: Pinches S., *ITER’s Integrated Modelling Infrastructure and Strategy for Data Analysis and Validation*

O-42: Lupelli I., *The SAGE Project: A paradigm shift in the Storage Systems for Data Centric Computing*


O-44: Greenwald M., *Navigational Data Management*

O-45: Zabeo L., *Diagnostic Data Handling in the PCS*

O-46: Emoto M., *Improvement of Automatic Physics Data Analysis Environment for the LHD Experiment*

O-47: Kim J., *Development of unified data analysis in KSTAR*

O-48: Witt S. de, *A Comparison of Data Management Techniques Across Different Science Disciplines*
Laser-driven Ion-beam Trace Probe (LITP) is a new diagnoses method to reconstruction 2D profiles of the poloidal magnetic field (Bp) and radial electric field (Er) in the tokamak devices [1, 2, 3]. The basic idea of LITP is that, an ion beam generated and accelerated by a laser on the chamber wall, injected into the chamber and then flies out in the other side of the wall. The inward and outward ion positions were recorded. The flying traces of ions with different energies in the chamber are used to reconstruct 2D profiles of Bp and Er simultaneously. The preliminary experimental test is also undergoing in the PPT device [3, 4].

To reconstruct 2D profiles of Bp and Er, some new techniques, such as the integration along cures, the critical numerical method to get the coefficient matrixes, etc., are applied in the reconstruction processes. The preliminary numerical tests show that the reconstruction error is less than 15%. In this talk we will present some important steps of data reconstruction, preliminary experimental results. The reconstruction error levels and some application limits of LITP will also be discussed.

References:
O-2: Assimilation of polarimetry Stokes vector measurements in tokamak free-boundary equilibrium reconstruction with application to ITER

Blaise Faugeras

Université Côte d'Azur, CNRS, Inria, LJAD, France

The modelization of polarimetry Faraday rotation measurements commonly used in tokamak plasma equilibrium reconstruction codes is an approximation to the Stokes model. This approximation is not valid for the foreseen ITER scenarios where high current and electron density plasma regimes are expected. In this work a method enabling the consistent resolution of the inverse equilibrium reconstruction problem in the framework of non-linear free-boundary equilibrium coupled to the Stokes model equation for polarimetry is provided. Using optimal control theory we derive the optimality system for this inverse problem. A sequential quadratic programming (SQP) method is proposed for its numerical resolution. Numerical experiments with noisy synthetic measurements in the ITER tokamak configuration for two test cases, the second of which is an H-mode plasma, show that the method is efficient and that the accuracy of the identification of the unknown profile functions is improved compared to the use of classical Faraday measurements.
A non-parametric Gaussian process regression model is developed in the 3-D equilibrium reconstruction code V3FIT. A Gaussian process is a normal distribution of functions that is uniquely defined by specifying a mean function and covariance kernel function. Gaussian process regression assumes that an unknown profile belongs to a particular Gaussian process, and uses Bayesian analysis to select the function that gives the best fit to measured data. The implementation in V3FIT uses a hybrid representation where Gaussian processes are used to infer some of the equilibrium profiles, and standard parametric techniques are used to infer the remaining profiles.

The implementation of the Gaussian process is tested using experimental data from the Compact Toroidal Hybrid experiment (CTH). These reconstructions use Gaussian processes to infer the emissivity profiles for a two-color soft X-ray diagnostic. Standard parametric models are used to represent the density, pressure, and current profiles. A Quasi-Newton iteration is used to simultaneously converge on the optimal set of model parameters and kernel hyper-parameters. The hybrid reconstructions are compared with fully parametric reconstructions.

This work was supported by Auburn University and the U.S. DOE through Award Numbers DE-FG02-03ER54692.
O-4: Robust Bayesian linear regression for Tokamak plasma boundary estimation

V. Škvára¹²⁴, V. Šmíd³, J. Urban⁴

¹ Institute of Information Theory and Automation, Prague, Czech Republic
² Faculty of Nuclear Sciences and Physical Engineering, CTU, Prague, Czech Republic
³ Regional Innovation Center for Electrical Engineering, Pilsen, Czech Republic
⁴ Institute of Plasma Physics, Prague, Czech Republic

The problem of plasma boundary or equilibrium reconstruction in a tokamak is important for control and data analysis. Essentially, this problem involves an inversion of the Grad-Shafranov equation with constraints given by diagnostic observations. This leads generally to a linear regression. However, the number of regressors is potentially infinite and a manual choice of the computed coefficients significantly influences the results. Also, the classical approaches to solving linear regression do not easily account for outliers in data or the fact that the underlying model may not be valid in certain phases of the discharge.

In this paper, we use formal Bayesian estimation technique [1] to handle these issues. We build an alternative to classical least squares solution using a hierarchical model of parameters. This enables us to embed some favorable properties into the final estimate. Firstly, we employ a self-tuning penalization on the norm of each element of the coefficient vector. This leads to an automatic selection of the optimal structure of the model as some elements are ignored. Secondly, we relax the assumption of Gaussian noise and assume that the distribution of the noise is Student’s t. Therefore, the estimate is less sensitive to outliers and can compensate for model inaccuracies or systematic errors.

The resulting algorithm and its performance in tokamak conditions is illustrated on the VacTH code [2] using data from the COMPASS tokamak. This code uses a toroidal harmonics decomposition of the equilibrium poloidal magnetic flux in the vacuum region, which is then used to reconstruct the plasma boundary. We compare different approaches using simulated and real data. Our method demonstrates how the Bayesian approach to linear regression improves the precision of the boundary reconstruction in different phases of a discharge. Also, it enables us to infer further properties of the physical system such as the position of outlier measurements. In addition, the possibility of using our method in the EFIT equilibrium reconstruction code using data from JET is discussed.

References:
Due to the complicated and anisotropic nature of the fast-ion distribution function, diagnostic velocity-space weight functions [1], which indicate the sensitivity of a diagnostic to different fast-ion velocities, are used to facilitate the analysis of experimental data. Additionally, when velocity-space weight functions are discretized, a linear equation relating the fast-ion density and the expected diagnostic signal is formed. In a technique known as Velocity-space Tomography [2], many measurements can be combined to create an ill-conditioned system of linear equations that can be solved using various computational methods [3].

However, when velocity-space weight functions (which by definition ignore spatial dependencies) are used, Velocity-space Tomography is restricted, both by the accuracy of its forward model and also by the availability of spatially overlapping diagnostic measurements. In this work we extend velocity-space weight functions to a full 6D generalized coordinate system and then show how to reduce them to a 3D orbit-space without loss of generality using an action-angle formulation. Furthermore, we show how diagnostic orbit-weight functions can be used to infer the full fast-ion distribution function, i.e. Orbit Tomography. Examples of orbit weights functions for different diagnostics and reconstructions of fast-ion distributions from synthetic data are shown.

References:
O-6: Accurate Determination of Radiated Power Density Profile Using Bolometer Data for DT Baseline Scenario at JET

Gediminas Stankunas\textsuperscript{1}, Fernanda Rimini\textsuperscript{2} and JET Contributors\textsuperscript{*}

\textit{EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK}
\textsuperscript{1}Lithuanian Energy Institute, Laboratory of Nuclear Installation Safety, Breslaujos str. 3, LT-44403 Kaunas, Lithuania
\textsuperscript{2}CCFE, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK

E-mail of Corresponding Author: gediminas.stankunas@lei.lt

The experimental data obtained from the campaign dedicated to Baseline scenario for DT (deuterium-tritium) at Joint European Torus (JET) is being investigated in the frame of EUROfusion’s programme. The development of reliable ~4.0 MA scenario at q95 ~ 2.7-3 compatible with DT operation and pulse length of 5s together with the optimization of the scenario for high performance at 4.0 MA to achieve $P_{\text{fus}} = 15\text{MW}$ in DT are the main experimental goals for the study related to “Baseline scenario for DT” at JET. For this purpose, data analysis of the experiments of the DT campaign needs the bolometric measurements of the energy losses with electromagnetic radiation and neutral particles, which is the essential diagnostic tool for hot plasmas.

At JET, where the plasma has a complex shape, the measurement of the spatial distribution of radiation losses employs several multichannel bolometric arrays with different directions of sight (horizontal and vertical) installed in a poloidal cross-section. Moreover, additional channels are used for obtaining the radiation loss distribution in the region of the divertor for radiative experiments. Tomography is used for reconstruction from the set of line-of-sight integrated measurements of brightness (in $\text{W}\cdot\text{m}^{-2}$) to the local emissivity (in $\text{W}\cdot\text{m}^{-3}$) profiles.

As the result of the data analysis, the radiated power in tokamaks are provided by JET bolometer tomography reconstruction. In addition, energy balance as well as reliance on discharge and the time evolution of radiation loss profiles are delivered for investigation of the plasma and impurity transport in baseline scenario for DT.

\textsuperscript{*} See the author list of “Overview of the JET results in support to ITER” by X. Litaudon et al. to be published in Nuclear Fusion Special issue: overview and summary reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17-22 October 2016)
Plasma tomography consists in reconstructing the 2D plasma profile on a cross-section of the fusion device, based on radiation measurements taken along several lines of sight. In JET, there is a horizontal camera (KB5H) and a vertical camera (KB5V) with 24 bolometers each, and with a sampling rate of 5 kHz. This means that a 30-second pulse could, in principle, yield as much as 150,000 reconstructions. However, the reconstruction process is usually attained through iterative regularization methods, which are computationally intensive. On average, only a few reconstructions are actually computed per pulse.

In previous work [1], we have shown that a deep neural network with several up-convolutional layers (up-sampling + convolution) can approximate the results of tomographic reconstruction with high accuracy. More recently, we improved the design of such network by replacing the up-convolutional layers with deconvolutional layers (i.e. transposed convolutions) in order to obtain the logical inverse of a convolutional neural network (CNN). We have also removed the requirement for any data preprocessing, so the sensor data coming from the bolometers can be fed directly to the network.

The network has been trained on a set of 23,500 sample tomograms collected from all JET campaigns since the installation of the ITER-like wall (ILW) in 2011. We used an adaptive gradient descent algorithm with a small learning rate ($10^{-4}$) and a large batch size (400). After about 1900 epochs (8 hours on an Nvidia Titan X GPU), a minimum loss value was achieved on a holdout validation set comprising 10% of data. With the trained network, the tomographic reconstruction for an entire pulse can be performed in a matter of seconds, producing high-frame-rate videos of plasma heating, disruptions, and other phenomena.

References:

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. IPFN activities received financial support from Fundação para a Ciência e a Tecnologia (FCT) through project UID/FIS/50010/2013. The Titan X GPU used for this research was donated by NVIDIA Corporation.

* See the author list of “Overview of the JET results in support to ITER” by X. Litaudon et al. to be published in Nuclear Fusion Special issue: overview and summary reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17-22 October 2016)
O-8: Bayesian soft X-ray Tomography on Tore Supra and WEST

Tianbo Wang1,2,3, Didier Mazon2, Jakob Svensson4, Dong Li1 and Geert Verdoolaege3,5

1. Southwestern Institute for Physics, CNNC, C-610200 Chengdu, China
2. Institute for Magnetic Fusion Research, CEA, F-13115 Saint-Paul-lez-Durance, France
3. Department of Applied Physics, Ghent University, B-9000 Ghent, Belgium
4. Max-Planck-Institut für Plasmaphysik, D-17491 Greifswald, Germany
5. Laboratory for Plasma Physics, Royal Military Academy (LPP-ERM/KMS), B-1000 Brussels, Belgium

Gaussian Process Tomography (GPT) [1] is a recently developed tomography method based on Bayesian probability theory. This method has been applied successfully for the soft-X-ray (SXR) diagnostic at the W7-X stellarator [2] and for magnetic surface reconstruction at the JET tokamak [1]. By modeling the SXR emissivity field in a poloidal cross-section as a Gaussian process, Bayesian SXR tomography can be carried out in a robust and extremely fast way. Owing to the short execution time of the algorithm, it is an important candidate for providing real-time information on impurity transport and for fast MHD control. In addition, the Bayesian formalism allows quantifying the uncertainty on various profile parameters.

In this paper, Gaussian process tomography is validated using a synthetic data set and results are shown of its application to the reconstruction of SXR emissivity profiles on the Tore Supra and WEST tokamaks. The method is compared with the classical algorithm based on minimization of the Fisher information, in terms of accuracy, robustness and computational load.

References:
O-9: Benchmarking robust regression techniques for global energy confinement scaling in tokamaks

G. Verdoolaege$^{1,2}$

$^1$Department of Applied Physics, Ghent University, B-9000 Ghent, Belgium
$^2$Laboratory for Plasma Physics – Royal Military Academy (LPP – ERM/KMS), B-1000 Brussels, Belgium

Scaling of the global energy confinement in tokamaks has recently gained renewed interest, as several dependencies described by the IPB98 scaling law are not reproduced in experiments. While this may indicate dependence of the confinement on additional parameters not contained in the historical scaling law, the mathematical technique used to estimate the scaling has also been shown to be an important factor [1]. The present contribution concerns work aimed at increasing the confidence in the data-analytical methodology used for deriving the confinement scaling law and regression analysis on fusion data in general. Given the heterogeneous origins of the confinement data and the various uncertainties in both the data and the model, robustness of the regression results is at least as important a criterion as goodness-of-fit. In this contribution, we present results of an extensive benchmarking test carried out using several robust regression techniques on the IPB98 database, using both engineering and dimensionless quantities. It is shown that the highest degree of robustness is obtained with the recent geodesic least squares (GLS) technique, which performs regression on a probabilistic manifold [2]. Visualizations of the regression results are also presented, by means of an intuitive model of the space of univariate normal distributions.

References:
O-10: Empirical Optimization with the Optometrist Algorithm: Randomization Coupled With Expert Interpretation

E. Trask, M. Binderbauer, H. Gota, R. Mendoza

Tri Alpha Energy, Inc., P.O. Box 7010, Rancho Santa Margarita, CA 92688, USA

E-mail of Corresponding Author: axel.winter@iter.org

The configuration space of a plasma physics experiment has countless dimensions. Optimizing performance in such a system is strongly affected by the generally nonlinear dependence of physical variables (magnetic field, plasma density, etc.) on hardware settings, which usually consists of a complex many-to-many map. Development of physical models is an excellent abstraction with which to describe plasma processes, but the complexity necessary to fully describe a whole machine typically precludes a complete and/or analytical formulation. Empirical optimization then becomes the option of choice when fast response time, limited understanding, or complexity precludes construction of complete physical models.

The Optometrist Algorithm was developed to efficiently search through high dimensional configuration spaces while allowing physical model filters to guide experimentation. Markov Chain Monte Carlo (MCMC) randomization of machine subsystem settings is used to generate pairs of experimental settings (shots). Human experts (representing physical model filters) may then choose which outcome produced subjectively better results. After making a choice, the Optometrist Algorithm then proposes new pairs of settings. Use of this innovation led to improved field-reversed configuration (FRC) plasma performance in the C-2U device at Tri Alpha Energy, characterized by temporarily rising plasma temperature and total energy for 3 and 2 milliseconds, respectively.
Many scientific measurement systems exhibit complexly distributed noise caused by non-linear transform of the raw signals as well as the accuracy limitation of the system calibration. Widely adopted method to consider such a complex noise is the uncertainty propagation, where all the noise sources are modelled by hand. However, there are often unknown noise sources and correlations, which make the reliable inference difficult.

We propose more data-driven noise model, where the noise (likelihood) is modelled by a parametric density distribution, the mean and scale of which are further parameterized by neural networks as a function of the latent parameters $f$. The neural network parameter $\theta$ contains information about the noise property specific to the system.

We applied this noise model to the Thomson scattering system in Large Helical Device (LHD). From more than 60,000 sets of data we learned $\theta$ for this system, while the noise origin is left unknown. An example of the data and the posterior distribution with our noise model are shown in Fig.1 (a) and (b), respectively. The uncertainties by manual propagation (errorbars in (a)) are much smaller than the data scatter. On the other hand, our posterior standard deviation is much closer to the mean scatter. Furthermore, the scatter of our posterior mean is greatly suppressed thanks to the nonzero-mean noise model, which corrects the systematic error.

Fig. 1: (a) A typical measurement data produced by LHD Thomson scattering system, electron temperature (green, $T_e$) and density (blue, $N_e$). The estimated uncertainties by propagation methods are shown by errorbars. (b) The variational posteriors with the noise model learned from a lot of data. The posterior mean and standard deviations are shown by markers and their errorbars, respectively.
O-12: Multichannel Validation of Gyrokinetic Simulations using a Synthetic Diagnostic for Doppler Backscattering based on Full-Wave Simulations

D. R. Ernst\textsuperscript{1}, T. L. Rhodes\textsuperscript{2}, W. Guttenfelder\textsuperscript{3}, K. H. Burrell\textsuperscript{4}, R. Bravenec\textsuperscript{5}, A. M. Dimits\textsuperscript{5}, B. A. Grierson\textsuperscript{3}, C. Holland\textsuperscript{2}, C. C. Petty\textsuperscript{4}, L. Zeng\textsuperscript{2} and the DIII-D Team

\textsuperscript{1}MIT Plasma Science and Fusion Center, Cambridge, MA 02139, USA
\textsuperscript{2}University of California Los Angeles, P.O. Box 957099, Los Angeles, CA 90095, USA
\textsuperscript{3}Princeton Plasma Physics Laboratory, P.O. Box 451, Princeton, NJ 08543, USA
\textsuperscript{4}General Atomics, P.O. Box 85608, San Diego, CA 92186, USA
\textsuperscript{5}Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94551, USA
\textsuperscript{6}Fourth State Research, Austin, TX, 503 Lockhart Drive, Austin, TX 78704, USA
\textsuperscript{7}University of California San Diego, 9500 Gilman Dr., La Jolla, California 92093, USA

Using a new synthetic diagnostic based on full-wave simulations, we demonstrate simultaneous agreement between nonlinear gyrokinetic simulations and experiment, using GYRO and GENE, in all transport channels (particles, ion and electron thermal energy, and momentum). Detailed Doppler Backscattering (DBS) density fluctuation spectra in the inner plasma core are simultaneously reproduced. The same close agreement persists with and without strong electron heating, without adjustment, in DIII-D low torque quiescent H-Mode experiments designed to study density gradient driven TEM turbulence \cite{ref1}. We directly observe density gradient driven TEMs for the first time as a band of discrete toroidal mode numbers which intensify during strong electron heating, degrading confinement in the inner core. The experiments were designed using predictive linear gyrokinetic simulations to isolate the density gradient driven TEMs in the inner core. Between-shots linear GYRO simulations were developed and used to guide the experiments, and to determine the desired DBS O-Mode launch angle and cutoff location. For this work we carry out two-dimensional full-wave simulations of DBS using the FDTD2D code \cite{ref2} on NVIDIA Fermi GPUs. A semi-analytic description of the effective spot size is then convolved with nonlinear GYRO simulations in a synthetic diagnostic. Our analytic results reveal new focusing and defocusing effects associated with the departure of the reflecting surface from flux surfaces. We then compare our semi-analytic results with the DBS transfer function obtained from direct full-wave simulations.

References:

\textit{This work was supported in part by the US Department of Energy under DE-FC02-08ER54966\textsuperscript{1}, DE-AC02-05CH11231\textsuperscript{1,3,5,6}, DE-FC02-04ER54698\textsuperscript{4}, DE-AC02-09CH11466\textsuperscript{6}, FG02-08ER54984\textsuperscript{2}, DE-AC52-07NA27344\textsuperscript{3}, DE-FG02-08ER54978\textsuperscript{8}.}
Understanding transport in magnetically confined plasmas is critical for developing predictive models for future devices, such as ITER. Nowadays, special attention is given to understanding perturbative transport and transient phenomena, but it still remains a formidable task. New techniques need to be developed to extract the useful information contained in the temporal behaviour of transport and plasma parameters, which can help the validation process of turbulent transport physics. Thanks to recent progress in simulation and theory, along with enhanced computational power and better diagnostic systems, the direct match of modeling and experimental results is becoming a reality. In this work, a machine learning framework is developed to address the issue of validation consistency of cutting-edge quasilinear transport models, such as TGLF. A surrogate-based strategy using Gaussian processes and sequential parameter updates is used to achieve the combination of plasma input parameters that matches experimental heat fluxes within diagnostic error bars. Several strategies and optimization algorithms are assessed, and preliminary results indicate that these machine learning algorithms are very suitable as a self-consistent and comprehensive validation methodology for plasma transport codes. Future work will focus on employing this procedure with additional constraints, such as incremental diffusivity measurements using perturbative transport techniques, and turbulence levels using dedicated diagnostic systems.

*Work supported by US DOE under grants DE-FC02-99ER54512 (C-Mod) and DE-FC02-04ER54698 (DIII-D) and La Caixa Fellowship.*
O-14: Validating Simulations of Multi-Scale Plasma Turbulence in in ITER- Relevant, Alcator C-Mod Plasmas

N.T. Howard\textsuperscript{1}, C. Holland\textsuperscript{2}, A.E. White\textsuperscript{1}, M. Greenwald\textsuperscript{1}, J. Candy\textsuperscript{3}, and A.J. Creely\textsuperscript{1}

\textsuperscript{1}Massachusetts Institute of Technology, Cambridge, MA, USA
\textsuperscript{2}University of California, San Diego, San Diego, CA, USA
\textsuperscript{3}General Atomics, San Diego, CA, USA

ITER-like conditions obtained from dedicated Alcator C-Mod experiments have been studied using cutting-edge simulations that simultaneously capture both ion and electron-scale turbulence. Operated with ITER-like toroidal field, density, and coupled ions and electrons ($T_i \sim T_e$), these conditions shed light on the possible influence of cross-scale coupling in predictions of ITER and represent one of the only validation exercises ever performed using the multi-scale (coupled electron and ion-scale) gyrokinetic model. Using the GYRO code, these multi-scale simulations utilized all experimental inputs (kinetic profiles and geometry) and realistic electron mass ($m_i/m_e=3600$) to capture the dynamics of electromagnetic ($\delta \phi & \delta A$) turbulence with $k_\theta \rho_s$ up to 42.0 while including e-i & i-i collisions, 3 gyrokinetic species and ExB shear effects. In qualitative agreement with previous investigations in L-mode, this work demonstrates that cross-scale coupling can significantly affect the saturation of the turbulence at ion-scales, resulting in enhanced levels of heat transport. It is demonstrated that experimental ion and electron heat fluxes can be quantitatively reproduced with multi-scale simulation. However, direct comparison of measured incremental diffusivities with simulation suggests that, high-k TEM/ETG turbulence does not appear to play the dominant role in reproducing measured electron incremental diffusivities. Quantitative comparisons of simulation results with experiment, the origin of the measured incremental diffusivities, and the implications of cross-scale coupling for prediction of ITER will be presented.

Supported by US DOE grant DEFC02-99ER54512-CMOD.
O-15: An Improved Approach to Uncertainty Quantification for Plasma Turbulence Validation Studies

P. Vaezi, C. Holland, S. C. Thakur, G. R. Tynan

Center for Energy Research, University of California, San Diego, La Jolla, CA 92093

We present the results of a new validation study of self-regulating drift-wave turbulence/zonal flow dynamics in the Controlled Shear Decorrelation Experiment (CSDX) linear plasma device which contrasts conventional approaches to uncertainty quantification (UQ) within the plasma community to more advanced techniques developed for other fields. Due to the strong nonlinear dependence of plasma turbulence on drive and dissipation mechanisms, uncertainties in experimental inputs can be greatly magnified in simulations of this turbulence. Thus, careful UQ and its inclusion within validation metrics is an integral part of plasma turbulence validation studies. To predict CSDX dynamics, we use a set of nonlocal 3D equations describing which evolve density, vorticity, and electron temperature fluctuations, and include proper sheath boundary conditions. Nonlinear simulations of these equations are carried out using BOUndary Turbulence framework [Dudson, 2009]. To identify the dominant parametric dependencies of the model, an extensive linear growth rate sensitivity analysis is performed using input parameter uncertainties which are taken from the experimental measurements [Thakur, 2013]. For the direct comparison of nonlinear simulation results to experiment, we use synthetic Langmuir probe diagnostics to generate a set of synthetic observables which are in turn used to construct the validation metrics. We observe significant improvement of model-experiment agreement relative to the previous 2D simulations [Holland, 2007]. An essential component of this improved agreement is found to be the strong effect of electron temperature fluctuations on floating potential measurements, which introduces clear amplitude and phase shifts relative to the plasma potential fluctuations. To minimize the number of simulations required for uncertainty quantification, we use the low-discrepancy Hammersley sequence sampling method [Simpson, 2001] for efficient sampling of the input parameter space, and a rapidly converging non-intrusive Probabilistic Collocation method [Webster, 1996] to generate a probabilistic model response. Using this approach is shown to yield significantly better constrained uncertainty estimates than conventional uniform sampling methods for practical numbers of nonlinear simulations. Potential applications to this UQ approach to other plasma simulations will be discussed.

This work is supported by US DoE under DE-FG02-06ER54871.
The suppression of micro-turbulence and ultimately the inhibition of large-scale instabilities observed in tokamak plasmas, e.g. Rogers et al. (2000), is partially characterized by the onset of a global stationary state. This stationary attractor corresponds experimentally to a state of “marginal stability” in the plasma. However, the critical threshold that determines the onset of this state is generally not the same as the critical threshold for linear instabilities. This is partly because the linearly unstable perturbations saturate creating a new stable stationary state in a process referred to as “self-organized criticality”. The critical threshold that characterizes the onset in the nonlinear regime is observed both experimentally and numerically to exhibit an upshift relative to the linear theory. That is, the onset in the stationary state is up-shifted from those predicted by the linear theory as a function of the ion temperature gradient \( R_0/L_T \) as discussed by Dimits et al. (2000); Mikkelsen and Dorland (2008).

Because the transition to this state with enhanced transport and therefore reduced confinement times is inaccessible to the linear theory, strategies for developing nonlinear reduced physics models (i.e. low-fidelity models) to predict the upshift have been ongoing, such as those of Kolesnikov and Krommes (2005). As a complement to these efforts, the principle aim of this work is to establish low-fidelity surrogate models that can be used to predict instability driven loss of confinement using training data from high-fidelity models. For the high-fidelity model, XGC1 Total-f gyrokinetic simulations discussed by Ku et al. (2016) are run in a circular geometry (i.e. low-\( f \)). These simulations are global, full-radius, and electrostatic. By judiciously selecting surrogate training scenarios, we are able to determine low-cost surrogates for the temperature gradient \( R_0/L_T \), as a function of the toroidal magnetic field \( B_0 \) and the major radius \( R_0 \) of the tokamak.

References:


O-17: Validation of MHD Models using MST RFP Plasmas


University of Wisconsin—Madison

Significant effort has been devoted to improvement of computational models used in fusion energy sciences. Rigorous validation of these models is necessary in order to increase confidence in their ability to predict the performance of future devices. MST is a well diagnosed reversed-field pinch (RFP) capable of operation over a wide range of parameters. In particular, the Lundquist number $S$, a key parameter in resistive magnetohydrodynamics (MHD), can be varied over a wide range and provide substantial overlap with MHD RFP simulations. MST RFP plasmas are simulated using both DEBS, a nonlinear single-fluid visco-resistive MHD code, and NIMROD, a nonlinear extended MHD code, with $S$ ranging from $10^4$ to $5 \times 10^6$ for single-fluid runs, with the magnetic Prandtl number $P_m = 1$. Experiments with plasma current $I_P$ ranging from 60 kA to 500 kA result in $S$ from $4 \times 10^4$ to $8 \times 10^6$. Validation metric comparisons are presented, focusing on how magnetic fluctuations $\tilde{b}$ scale with $S$. Single-fluid NIMROD results give $S \sim \tilde{b}^{-0.20}$, and experiments give $S \sim \tilde{b}^{-0.09}$ for the dominant $n=6$ mode. Additionally, simulated mode amplitudes are about half of the experimental measured values. This discrepancy is explored using an external circuit model for the toroidal field and also using two-fluid simulations.
We have begun the validation of a soft x-ray (SXR) continuum model of fusion plasmas. SXR brightness measurements contain a wealth of information on many plasma parameters of interest, however without information from modeling, it is nearly impossible to extract. A quantitatively validated forward model is therefore necessary for the accurate interpretation of SXR measurements, and will be critical in the burning plasma era. Until recently, we have been using a simplified soft x-ray emissivity model and have focused on extracting electron temperature profiles via the two–color soft x-ray technique. This simplified model is well verified and works very well to extract reliable electron temperatures in low–Z machines (RFX-Mod). Furthermore, it has been developed in parallel to the development of associated SXR diagnostics, and can model the absolute brightness from those diagnostics well. However, there are limitations to that agreement. These limitations become particularly apparent when inferring electron temperature in a machine with medium–Z plasma facing components (Madison Symmetric Torus). This simplified model appears to be missing necessary physical effects. The atomic database and analysis structure (ADAS) is a powerful interpretive tool that has been extensively used to model and predict atomic line spectra and level populations based on assumptions that are valid for fusion plasmas. These predictions have been extensively compared against experimental measurements, and are in good agreement. However, continuum radiation in the x-ray range, while also modeled in ADAS, has not been rigorously verified or tested against experimental data. We have therefore begun a systematic comparison of ADAS to the simplified model, adding physics to the simplified model when necessary. Non-constant gaunt factors and dielectronic recombination were systematically added to the simplified model in order to verify that ADAS was using the most up-to-date parameter spaces and calculation methods. Both ADAS and the simplified model were compared to experimental data to assess the importance of various physical effects and to begin the validation of SXR continuum calculations from ADAS.

This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Fusion Energy Sciences program under Award Number DE-SC0015474.
Measuring Soft X-Ray (SXR) radiation (0.1-15 keV) of fusion plasmas is a standard way of accessing valuable information on particle transport, in particular tungsten (W) which could degrade plasma core performances and cause radiative collapses.

The design and implementation of a new SXR diagnostic developed for the WEST project [1], based on a triple Gas Electron Multiplier (GEM) detector [2] is first presented. This detector, tested on ASDEX-Upgrade [3] before the WEST installation and the first experimental campaign in 2017, works in photon counting mode and presents energy discrimination capabilities. The SXR system is composed of two 1D cameras (vertical and horizontal views respectively), located in the same poloidal cross-section to allow for tomographic reconstruction.

This paper is mainly focused on the GEM Soft X-ray measurement validation and on tomography issues, including recent developments made to get more precise information about plasma features from the inverted SXR data like W poloidal distribution. Preliminary simulations of plasma emissivity and W distribution have been performed for WEST using a recently developed synthetic diagnostic [4] coupled to a tomographic algorithm based on the minimum Fisher information (MFI) inversion method [5]. Comparison and validation with first GEM acquisition is performed as well as estimation of transport effect on impurity ionization equilibrium and cooling factor depending on the chosen energy band used for the tomographic inversion.

References:
O-20: Disruption Forecasting in Tokamak Fusion Plasmas using Deep Recurrent Neural Networks

Julian Kates-Harbeck\textsuperscript{1}, Kyle Felker\textsuperscript{2}, Alexey Svyatkovskiy\textsuperscript{2}, Eliot Feibush\textsuperscript{2}, William Tang\textsuperscript{2}

\textsuperscript{1}Harvard University, Cambridge MA 02138
\textsuperscript{2}Princeton University, Princeton NJ 08544

The prediction and avoidance of disruptions in tokamak fusion plasmas represents a key challenge on the way to stable energy production from nuclear fusion. A fusion plasma is a complex dynamical system with some unknown internal state which emits a time series of possibly high dimensional observable data that is captured by sensory diagnostics. Using such diagnostic data from past plasma shots with both disruptive and non-disruptive outcomes, we train a deep recurrent neural network to predict the onset of disruptions in an online setting.

To deal with very large amounts of data and the need for iterative hyperparameter tuning, we also introduce a distributed training algorithm that runs on MPI clusters of GPU nodes and provides strong linear runtime scaling. Our approach demonstrates competitive predictive performance on experimental data from the JET tokamak, and we highlight promising avenues for extending our method to cross-tokamak prediction as well as to high-dimensional diagnostic data such as temperature and density profiles.
The new Disruption Event Characterization and Forecasting (DECAF) code [1] has the goal of characterizing and forecasting events that can lead to disruption of a tokamak plasma. Determining common chains of events that occur can provide insight into how to cue disruption avoidance systems. One example is a resistive wall mode (RWM) detection algorithm which can be either a threshold test on magnetic signals or a more sophisticated kinetic RWM model. In an NSTX dataset with 26 unstable RWMs, the detection of an RWM event by the threshold test within 100ms of the disruption was followed immediately by “vertical stability control” events 15 times. Further, looking at the two-event chains that happened directly after this set of RWMs, we find that even though there are theoretically 56 two-event combinations that could occur from the eight currently tested for, just two two-event chains accounted for 50% of the cases and five accounted for 77%. The importance of kinetic effects, such as resonances between plasma rotation and bounce and precession particle motions, as well as the effects of collisions on RWM stability has recently been shown by analysis with the MISK code [2]. This physics, streamlined for fast computation, has now been incorporated in a “reduced” kinetic model in DECAF [Error! Bookmark not defined.]. The reduced model performed well in its first iteration on NSTX data, finding instability 84% of the time for experimentally unstable cases, and stability in 77% of experimentally stable cases [Error! Bookmark not defined.]. Another key cause of disruptions is the physical event chain that comprises the appearance of rotating magnetohydrodynamic (MHD) modes, their slowing by resonant field drag mechanisms, and their subsequent locking. Therefore a module was written for DECAF which both identifies the existence of rotating MHD modes and tracks certain characteristics that can lead to disruption [3]. This module has been tested with data from both NSTX and its upgrade NSTX-U. Characteristics such as identification of a mode locking time based on a loss of torque balance and bifurcation of the mode rotation frequency are examined to determine the reliability of such events in predicting disruptions.

References:

Supported by U.S. DOE grant DE-FG02-99ER54524
Increased warning times in JET APODIS disruption predictor by using confidence qualifiers

J. Vega\textsuperscript{1}, S. Dormido-Canto\textsuperscript{2}, A. Murari\textsuperscript{3}, G. A. Rattá\textsuperscript{1}, R. Castro\textsuperscript{1} and JET Contributors\textsuperscript{*}

\textit{EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK}

\textsuperscript{1}Laboratorio Nacional de Fusión, CIEMAT. Av. Complutense 40. 28040 Madrid. Spain

\textsuperscript{2}Dpto. de Informática y Automática. Juan del Rosal 16. 28040 Madrid. Spain

\textsuperscript{3}Consorzio RFX. Corso Stati Uniti 4, 35127 Padova, Italy

E-mail of Corresponding Author: jesus.vega@ciemat.es

JET APODIS disruption predictor follows a two-layer architecture of Support Vector Machines (SVM) classifiers. The 1\textsuperscript{st} layer contains three independent SVM classifiers that operate in parallel. As a discharge is in execution, the classifiers use the three most recent time windows of 32 ms to make predictions about the disruptive or non-disruptive plasma behaviour. Of course, the three predictions are not necessarily the same. Therefore, the 2\textsuperscript{nd} layer is used to combine the above three outputs into a single one. The output of the 2\textsuperscript{nd} layer determines whether or not to trigger an alarm. This paper shows a modification of the APODIS 2\textsuperscript{nd} layer classifier to obtain larger warning times, which is essential to put into operation earlier mitigation actions. The modification of the 2\textsuperscript{nd} layer classifier is based on qualifying every 32 ms how good both predictions (disruptive and nondisruptive) are. This qualification is carried out in the framework of conformal predictions. Given a set of examples \{\(z_i = (x_i, y_i), i = 1, ..., n\), where \(x_i \in \mathbb{R}^m\) are feature vectors and \(y_i \in L = \{L_i, ..., L_k\}\) are labels, conformal predictors determine the label of a new feature vector \(x_{n+1}\) by assuming for it all possible labels and predicting the label that makes \(x_{n+1}\) the most conformal to the initial set of \(n\) examples. The conformity measure is determined by computing the p-values \((1/(n + 1) \leq p \text{- value} \leq 1)\) for all possible labels (V. Vovk et al. Proc. 16\textsuperscript{th} Int. Conf. on Machine Learning. (1999). San Francisco, CA), where the minimum value means the least conformal label and the maximum value means the most conformal one. In the case of the APODIS, only two labels are possible \(L = \{\text{disruptive, non-disruptive}\}\). To increase the warning time, an alarm is triggered when simultaneously both assumptions (disruptive and non-disruptive) are qualified with the minimum p-values. In other words, an alarm is triggered when the statistical confidences in both labels are minimum. This has been applied to 789 JET discharges in the range 82429-83793 (81 unintentional disruptions predicted by APODIS and 708 non-disruptive discharges). All disruptions are predicted and 94\% of them have a warning time >10 ms (minimum time in JET to trigger the disruption mitigation valve). The average warning time of the classical APODIS predictor is 428 ms with a standard deviation of 1166 ms. With the new second layer classifier, the average warning time is 606 ms with a standard deviation of 1874 ms. The false alarm rate increases from 1\% to 5\%.

\textsuperscript{*} See the author list of “Overview of the JET results in support to ITER” by X. Litaudon et al. to be published in Nuclear Fusion Special issue: overview and summary reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17-22 October 2016)
Algorithms for providing real-time warning of impending disruptions have been developed for a few tokamaks, but these are usually trained on, and optimised for, specific individual machines, which necessarily requires data from a large number of disruptions on each machine. For ITER, it may not be feasible to accumulate a sufficient training set of unmitigated disruptions, particularly at high performance. Is it possible to develop ‘universal’, i.e. machine-independent, real-time disruption warning algorithms using data from multiple present-day tokamaks, which could then be directly applied to ITER, thereby eliminating the need for accumulating many unmitigated disruptions on ITER? To begin to address this question, we have assembled SQL disruption databases for three different tokamaks, Alcator C-Mod, EAST, and DIII-D, and have used these to examine disruption similarities and differences between the three machines. Each device’s database contains data from all plasmas in their respective 2015 operating campaign, both disruptive and non-disruptive (~2000-3000 discharges). For each shot, about 40 disruption-relevant plasma parameters are sampled at regular intervals during the shot. ‘\(I_p\) error’, ‘Greenwald fraction’, ‘radiated power fraction’, and ‘time until disrupt’ are examples of some of the parameters. Additional sampling is done at a higher rate for a fixed period before each disruption. The ~40 parameters for each time slice for each shot constitute a single record in the SQL database. The databases are accessible using any SQL-enabled software, such as Matlab, Python, or IDL. Examples of statistical analyses of the data in the databases will be shown, focusing on the inherent differences that we have observed between machines, which do not bode well for the goal of developing a universal warning algorithm. These databases are also amenable to the application of so-called ‘machine learning’ techniques, including deep neural networks, support vector machines, and random forests, to the disruption warning problem. These will be described in detail in a follow-on talk.
Using data-driven methodology, we exploit the time series of relevant plasma parameters for a large set of disruptive and non-disruptive discharges from the DIII-D tokamak with the objective of developing a disruption prediction algorithm. Our work is motivated by statistical analysis of disruptions carried out during past years [1,2], and aims at building an algorithm that would guarantee sufficient warning time to trigger disruption mitigation and suppression mechanisms. This will be highly important in the realization of a device like ITER.

We focus on a subset of disruption predictors, mainly non-dimensional and/or machine-independent parameters such as plasma internal inductance and Greenwald density, coming from both plasma diagnostics and equilibrium reconstructions. The utilisation of dimensionless indicators favors a more direct comparison between different tokamak devices. Data preprocessing, and in particular feature selection and feature reduction, is a mandatory step in the machine-learning workflow. We initially discuss the methods adopted in the exploratory data analysis, e.g. correlation analysis on the multivariate dataset during the flattop phase of the discharges. In order to develop a robust disruption warning algorithm, where we can obtain a warning time related to the probability of disruption occurrence, we first discuss the methodology adopted to solve the binary classification problem (disruptive/non-disruptive) on DIII-D data. Then, we present the results coming from a multi-class classification problem, where we include the time dependency through the definition of class labels on the basis of the elapsed time before the disruption (i.e. “far from a disruption”, “within 50 ms of disruption”, etcetera). Disruption prediction can be tackled using a plethora of different machine-learning techniques, each presenting its drawbacks and advantages. Some of these algorithms, such as Artificial Neural Networks, Support Vector Machines and Random Forests, are discussed through their classification/regression application, and their performances are evaluated using accuracy metrics for the purpose of disruption prediction.

References:

This work was supported by the US Dept. of Energy under DE-FC02-04ER54698 and DE-SC0014264.
O-25: Tokamak profile database construction incorporating Gaussian process regression

A. Ho, J. Citrin, C. Bourdelle, Y. Camenen, F. Felici, M. Maslov, K.L. van de Plasche, H. Weisen and JET Contributors*

EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB, UK
1 DIFFER - Dutch Institute for Fundamental Energy Research, Eindhoven, the Netherlands
2 CEA, IRFM, F-13108 Saint Paul Lez Durance, France
3 IIFS/PIIM, CNRS - Université de Provence, Marseille, France
4 Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands
5 EURATOM-CCFE Fusion Association, Culham Science Centre, Abingdon, OX14 3DB, UK
6 Swiss Plasma Centre, EPFL, 1015 Lausanne, Switzerland

Shot preparation and tokamak controller design require accurate, fast, and robust models to predict the plasma state, including the turbulent transport coefficients, and its evolution in time. By emulating quasilinear transport models, such as QuaLiKiz [1], with neural networks [2], the aim is to close the remaining tractability gap for such applications. The neural network model can then be integrated into the fast tokamak simulator, RAPTOR [3], for making predictions of the plasma evolution.

However, generating a training set for large-dimensionality (≈ 20D) neural network models may require a prohibitive number of turbulent transport simulations. By sampling experimental data to derive simulation inputs, this number can be reduced by restricting the parameters to a relevant subspace. The development of a profile database initially focusing on JET-relevant parameter space is currently underway, with ≈ 1000 discharges each divided into ≈ 10 time slices spanning the ramp-up, flat-top, and ramp-down phases. A workflow of kinetic profile extraction has been implemented, including Gaussian process regression fitting [4] of the measured plasma profiles with reported errors. The result is a powerful tool for estimating the profiles, their derivatives, and associated uncertainties with increased statistical rigour. A corollary application of this database is to mark a subset of these discharges, with both high quality kinetic profiles and heat and particle sources, for widespread validation of the neural network transport models through fast integrated modelling frameworks.

References:

* See the author list of Overview of the JET results in support to ITER by X. Litaudon et al. to be published in Nuclear Fusion Special issue: overview and summary reports from the 26th Fusion Energy Conference (Kyoto, Japan, 17-22 October 2016).
Finding structure in large datasets of particle distribution functions using unsupervised machine learning

R.M. Churchill, C.S. Chang, S. Ku

Princeton Plasma Physics Laboratory

Many modern simulation codes that run on high-performance computing resource generate extremely large datasets. A key challenge is how to extract meaningful scientific knowledge from these datasets. Often, the solution is to simply reduce the dataset with pre-existing physics knowledge, e.g. looking only at the moments of particle distribution functions and their fluctuations. Techniques to extract various structures from these datasets are useful to explore physics that may be lost from common reductions.

An example question in tokamak physics that these techniques could help shine light on is determining the generation mechanism of spatial turbulent eddies near the open field-line region (commonly referred to as "blobs"). From a physics perspective, it has been hypothesized [1] that coherent structures in velocity space play a role in generating "blobs". However, other research had suggested that such coherent structures in velocity space should not survive long enough to have an effect [2].

To explore this concept, we apply unsupervised machine learning methods to particle distribution function datasets generated by the extreme-scale gyrokinetic turbulence code XGC1 [3], a multiphysics code focused on simulating the edge of fusion plasmas. K-means clustering is applied to the entire edge region where blobs are generated, and no common velocity-space structure is found across space. However, applying K-means clustering to particle distribution functions from "blob" regions results in ring-like clusters in velocity space at constant speed. Such analysis of the phase space structure can provide insight and next steps to explore in understanding the underlying physics of blob generation.

References:

This work was supported by the U.S. Department of Energy under Contract No. DE-AC02-09CH11466
O-27: Identification of ELM evolution patterns with unsupervised clustering of time-series similarity metrics

David R. Smith\textsuperscript{1*}, R. Fonck\textsuperscript{1}, G. McKee\textsuperscript{1}, A. Diallo\textsuperscript{2}, S. Kaye\textsuperscript{2}, B. LeBlanc\textsuperscript{2}, and S. Sabbagh\textsuperscript{3}

\textsuperscript{1}University of Wisconsin-Madison, Madison, WI, USA
\textsuperscript{2}Princeton Plasma Physics Laboratory, Princeton, NJ, USA
\textsuperscript{3}Columbia University, New York, NY, USA

E-mail of Corresponding Author: drsmith8@wisc.edu

Unmitigated edge localized modes (ELMs) pose a risk for ITER, and predictive capabilities for ELM properties like saturation mechanisms, filament dynamics, and pedestal collapse require validation of nonlinear ELM models. Large data archives and unsupervised machine learning algorithms provide the opportunity for a data-driven approach to ELM evolution dynamics. Here, we identify characteristic ELM evolution patterns with unsupervised cluster analysis of time-series similarity metrics \cite{1}. A database of ELM events was assembled from beam emission spectroscopy (BES) measurements on the National Spherical Torus Experiment (NSTX). The ELM events exhibit pedestal collapse and stored-energy loss, and the ELMs are likely Type I due to screening criteria and NSTX parameter regimes. Similarity metrics like time-lag cross-correlation and dynamic time warping were developed to quantify the similarity of ELM events and time-series data. Next, hierarchical and k-means clustering algorithms were applied to the similarity metrics to identify groups of ELMs with similar evolution patterns. The analysis yielded two or possibly three characteristic evolution patterns within the ELM database. In addition, the identified ELM groups correspond to distinct parameter regimes for several ELM-relevant parameters such as plasma current, elongation, and pedestal height. The identified evolution patterns and distinct parameter regimes suggest genuine variation in the underlying ELM dynamics. To clarify the mechanisms that yield distinct evolution patterns, we are presently pursuing MHD simulations of the identified evolution patterns.

References:

This material is based upon work supported by the US Department of Energy, Office of Science, Office of Fusion Energy Sciences under Award Numbers DE-FG02- 89ER53296, DE-SC0001288, and DE-AC02-09CH11466. This research used resources of the National Spherical Torus Experiment-Upgrade, which is a DOE Office of Science User Facility.
O-28: Integrated infrastructure for the development of machine-learning models aimed at fusion applications

Orso Meneghini, Sterling Smith

*General Atomics*

From a technical standpoint, among the most challenging aspects of developing machine learning models are the data collection, and the data management throughout the stages of post-processing, the deployment of the trained models, and their validation. This paper describes the software and the general computational infrastructure that was developed to support the development of two neural-networks (NN) based models that were used to perform non-linear multivariate regression of theory-based models for the core turbulent transport fluxes, and the pedestal structure. Specifically, these models were able to consistently reproduce the results of the TGLF and EPED1 codes over a broad range of plasma regimes, and with a computational speedup of several orders of magnitudes. Effective data collection was achieved by the development and use of a remote data collection library named HARVEST. The availability of C, FORTRAN, and Python APIs, allows this system to be conveniently embedded within any scientific code, and data to be transferred at run-time to a remote server. Data is transferred asynchronously eliminate the performance impact on time-sensitive numerical codes. On the server side, HARVEST leverages the functionality of no-SQL databases to support rapidly changing schemas of unstructured data with scalable performance. The OMFIT framework provided the infrastructure used to 1) concurrently submit large sets of simulations on high-performance computing clusters to efficiently build the training database 2) orchestrate the entire data pipeline, from accessing the data stored on the remote HARVEST data collection system to the training and benchmarking of the ensemble of NNs; and finally 3) execute the integrated workflows that are needed for the analysis, verification, and validation of the new models. The NNs were implemented based on the open-source FANN library. The techniques described in this paper are generic, and both the HARVEST infrastructure and the data management workflows within the OMFIT framework can be readily adapted to other problems of interest for fusion applications.

*Work supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-SC0012656 (GA AToM SciDAC), DE-FG02-95ER54309 (GA theory), DE-FC02-04ER54698 (DIII-D).*
O-29: Interpretive Analysis and Predictive Discharge Modeling with TRANSP

Brian Grierson\textsuperscript{1}, Nikolas Logan\textsuperscript{1}, Shaun Haskey\textsuperscript{1}, Lang Cui\textsuperscript{1}, Sterling Smith\textsuperscript{2}, Orso Meneghini\textsuperscript{2}, James Buchanan\textsuperscript{3}

\textsuperscript{1}Princeton Plasma Physics Laboratory, Princeton, NJ 08540, USA
\textsuperscript{2}General Atomics, San Diego, CA 92121, USA
\textsuperscript{3}CCFE, Culham Science Centre, Abingdon, OX14 3DB, UK

TRANSP [1] simulations are being used in the OMFIT [2] workflow manager to enable a machine-independent means of experimental analysis, postdictive validation, and predictive time dependent simulations on the DIII-D, NSTX, JET and C-MOD tokamaks. The procedures for preparing input data from plasma profile diagnostics and EFIT equilibrium reconstruction will be presented, as well as processing of the time-dependent heating and current drive sources and assumptions about the neutral recycling. Common settings for TRANSP simulation fidelity will be described contrasting between-shot analysis, power balance, and fast-particle simulations. A series of data consistency metrics are computed such as comparison of experimental vs. calculated neutron rate, equilibrium stored energy vs. total stored energy from profile and fast-ion pressure, and experimental vs. computed surface loop voltage. Discrepancies between data consistency metrics can indicate errors in input quantities such as electron density profile or $Z_{\text{eff}}$, or indicate anomalous fast-particle transport. Measures to assess the sensitivity of the verification metrics to input quantities are discussed, including scans of the input profiles and anomalous fast-ion diffusivity. For predictive calculations, TRANSP uses GLF23 or TGLF to predict core plasma profiles, with user defined boundary conditions in the outer region of the plasma. ITPA validation metrics are provided in post-processing to assess the transport model validity. By using OMFIT to orchestrate the steps for experimental data preparation, selection of operating mode, submission, post-processing and visualization, we have streamlined and standardized the usage of TRANSP.

References:

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, using the DIII-D National Fusion Facility, a DOE Office of Science user facility under award DE-AC02-09CH11466 and DE-FC02-04ER54698.
One of the main aims of Wendelstein 7-X, a newly built stellarator in Greifswald, is the investigation of quasi-steady state operation of magnetic fusion devices, for which power exhaust is a very important issue. The predominant fraction of the energy lost from the confined plasma will be taken by the 10 specially designed discrete island divertor modules, which can sustain up to 10 MW/m². In order to protect the divertor elements from overheating and to monitor power deposition onto the divertor elements 10 state-of-the-art infrared endoscopes will be installed at W7-X and software is under development for real-time analysis.

In the first operational phase, however, most of the graphite armor components and the island divertor were not yet installed. To protect metal parts, in particular divertor frame structure and to guaranty a reasonable performance, a so-called limiter configuration was used. Five graphite limiters, matching the fivefold symmetry of the plasma, were installed in symmetry planes at the inboard side. The magnetic vacuum configuration has been chosen such that it has a smooth scrape-off layer, with no stochastic region and no large magnetic islands, such that the limiters efficiently intercept 99% of the convective plasma heat load in the SOL. Using rather limited set of tools we were able to characterize the heat fluxes on the poloidal limiters in W7-X during the first helium and hydrogen plasmas during OP1.1. A non-uniform distribution of power loads with two heating stripes on each limiter was observered. The measured patterns could be to a large degree explained with modelling. IR thermography and calorimetry using multiple IR camera systems, combined with slow thermocouples to account for toroidal asymmetries, allowed us to estimate that the limiters intercepted up to 60% of the total energy put into the vessel by the ECRH heating system for low power (0.6 MW) long pulse (6 second) shots, and a smaller fraction (~35%) for high power (4MW) short duration (1 second) discharges.
O-31: Wendelstein 7-X near real-time image diagnostic system for plasma facing components protection

A. Puig Sitjes¹, M. Jakubowski¹, A. Ali¹, P. Drewelow¹, F. Pisano², V. Moncada³, T. Ngo⁴, J.M. Travere⁴, G. Kocsis⁵, T. Szepesi⁵ and W7-X Team¹

¹ Max Planck Institute for Plasma Physics, Wendelsteinstrasse 1, 17491 Greifswald, Germany
² University of Cagliari, Piazza d’Armi, 09126 Cagliari, Italy
³ ThemaDIAG, 100 Impasse des Houllières, ZA Le Pontet, F-13590 Meyreuil, France
⁴ CEA, IRFM, F-13108 Saint Paul-lez-Durance, France
⁵ Wigner RCP, RMI, Konkoly Thege 29-33, H-1121 Budapest, Hungary

E-mail of Corresponding Author: aleix.puig.sitjes@ipp.mpg.de

Wendelstein 7-X (W7-X) fusion experiment is aimed at proving that the stellarator concept is suitable for a future fusion reactor. Therefore, it is designed for steady-state plasmas of up to 30 minutes, which means that the thermal control of the Plasma Facing Components (PFC) is of vital importance to prevent damages to the device.

In this paper an overview of the design of the Near Real-time Image Analysis System for PFC protection in W7-X is presented. The goal of the System is to monitor the PFC with high risk of permanent damage due to local overheating during plasma operations and send alarms to the Fast Interlock System. The monitoring of the PFC is based on thermographic and visible video cameras, which are analyzed by means of GPU-based computer vision techniques, to detect the strike-line, hot spots and dangerous thermal events. The video streams and the detected thermal events are displayed online in the control room in the form of a thermal map and permanently stored in the database. In order to determine the emissivity and maximum temperature allowed, a pixel-based correspondence between the image and the observed device part is required. The 3D geometry of W7-X makes the System particularly sensitive to the spatial calibration of the cameras since hot spots can be expected anywhere and a full segmentation of the field of view is necessary, in contrast to other ROI-based systems. A precise registration of the field of view and a correction of the strong lens distortion caused by the wide-angle optical system are then required.

During the next operation phase the un-cooled graphite divertor units will allow to test the System without risk of damaging the divertors in preparation when water-cooled high-heat flux divertors will be used.
One Modeling Framework for Integrated Tasks (OMFIT) [1] has been used to develop a consistent tool for interfacing with, mapping, visualizing, and fitting tokamak profile measurements. The OMFITprofiles tool was motivated by the expanding needs of predictive and whole device modeling efforts, which require trusted time evolving inputs. In addition to the need for many time steps in modeling, the needs of experimental physicists to visualize profile dynamics for decision making during tokamak experiments motivated an efficient and consistent workflow providing immediate physics implications from fundamental measurements. OMFIT is used to integrate the many diverse diagnostics on multiple tokamak devices into a regular data structure, consistently applying spatial and temporal treatments to each channel of data. Tokamak data is fundamentally time dependent and is treated so from the start, with front loaded and logic-based manipulations such as filtering based on the identification of Edge Localized Modes that commonly scatter data. Fitting is general in its approach, and tailorable in its application in order to address physics constraints and handle the multiple spatial and temporal scales involved. Although community-standard 1D fitting is supported, including scale length fitting and fitting polynomial-exponential blends to capture the H-mode pedestal, OMFITprofiles includes 2D fitting using bivariate splines or radial basis functions. These 2D fits produce regular evolutions in time, removing jitter that has historically been smoothed ad-hoc in transport applications. Visualization is built in at each step, so scientists can clearly understand the raw data, its convolution to the time base of interest, mapping to flux coordinates, fits and their physics consequences. Profiles interface directly with a wide variety of models within the OMFIT framework, providing the inputs for TRANSP, kinetic-EFIT 2D equilibrium and GPEC 3D equilibrium calculations. The OMFIT profile tool’s rapid and comprehensive analysis of dynamic plasma profiles thus provides the critical link between raw tokamak data and simulations necessary for physics understanding.

References:

This work is supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, using the DIIID National Fusion Facility, a DOE Office of Science user facility under award DEAC0209CH11466 and DEFC0204ER54698.
O-33: Automatic between-pulse analysis of DIII-D experimental data performed remotely on a supercomputer at Argonne National Laboratory

M. Kostuk\textsuperscript{1}, T. Uram\textsuperscript{2}, T. Evans\textsuperscript{3}, D. M. Orlov\textsuperscript{3}, M. Papka\textsuperscript{2}, D. Schissel\textsuperscript{1}

\textsuperscript{1}General Atomics, 3550 General Atomics Ct, San Diego, CA 92121
\textsuperscript{2}Argonne National Laboratory, 9700 S. Cass Ave, Argonne, IL 60439
\textsuperscript{3}University of California San Diego, La Jolla, CA 92093

For the first time, an automatically triggered, between-pulse fusion science analysis code was run on-demand at a remotely located supercomputer at Argonne National Laboratory (ANL, Chicago, IL) in support of an in-process experiment being performed at DIII-D (San Diego, CA). This represents a new paradigm for combining geographically distant experimental and high performance computing (HPC) facilities to provide enhanced data analysis that is quickly available to researchers. Plasma physics experiments for fusion research occur in a fast-paced cycle; a plasma discharge at DIII-D is typically about 6 seconds, while the total cycle including review of the previous pulse and preparing the next one takes about 10-15 minutes. Enhanced analysis improves the understanding of the current pulse, translating into a more efficient use of experimental resources, and to the quality of the resultant science. The analysis code used here, called SURFMN \cite{1}, calculates the magnetic structure of the plasma using Fourier transform. Improved memory management, file I/O, and parallelization using pure MPI have made between shot analysis possible. Increasing the number of Fourier components provides a more accurate determination of the stochastic boundary layer near the plasma edge by better resolving magnetic islands, but requires 26 minutes to complete using local DIII-D resources, putting it well outside the useful time range for between shot analysis. These islands relate to confinement and edge localized mode (ELM) suppression \cite{2}, and may be controlled by adjusting coil currents for the next shot. Argonne has ensured on-demand execution of SURFMN by providing a reserved queue, a specialized service that launches the code after receiving an automatic trigger, and with network access from the worker nodes for data transfer. Runs are executed on 252 cores of ANL’s Cooley cluster and the data is available locally at DIII-D within three minutes of triggering. The original SURFMN design limits additional improvements with more cores, however our work shows a path forward where codes that benefit from thousands of processors can run between pulses.

References:

This work supported by the U.S. Department of Energy under DE-FG02-95ER54309.
The visible and infrared (VIS/IR) wide angle viewing system (WAVS) in ITER will have important machine protection roles and will contribute to plasma control as well as to physics studies. The interpretation of VIS/IR measurements in the all-metallic environment of ITER will be challenging. For example, the machine protection algorithms can misinterpret the reflections from ITER plasma facing components as a local overheating and trigger a fast termination of plasma. False positives can induce fatigue of the machine and hamper experimental programmes, and are therefore to be kept to a minimum, whilst ensuring that real events are detected. The modelling of synthetic images plays an important role in predicting the performance of the WAVS and for optimizing the diagnostic design. This contribution will summarize new results on the modelling of the WAVS synthetic images, including the impact of the optical blur on the WAVS spatial resolution, the impact of reflections on the apparent surface temperature and the impact of the temperature of the WAVS mirrors on the image quality. The synthetic images are generated by combining 3d temperature fields (produced by the field line tracing combined with the heat transfer simulations) with the raytracing simulations by the SPEOS code. The contribution will also address the challenges of the real-time processing of the WAVS data for the purpose of the plasma control and central interlock. Finally, an overview of the synthetic data production and integrated data analysis for other ITER diagnostics will be given. Disclaimer: The views and opinions expressed herein do not necessarily reflect those of the ITER Organization.
O-35: Data processing on application of real-time systems and validation of diagnostics in HL-2A

M. Xu on behalf of HL-2A team

Recently, real-time control systems for tearing mode control and feedback have been built on the HL-2A tokamak. An FIR polarimeter/interferometer was built to measure the plasma density in real-time and was used for density feedback control. A digital phase comparator technique with real-time dynamic spectrum analysis was implemented. In order to reduce the processing time, all signal processing was carried out in Field Programmable Gate Array (FPGA) chips. The time resolution can be less than 1 μs. ECRH launcher with real-time steering was developed too. A few real-time diagnostics and intelligent controllers allowed to achieve a precise control of the ECRH deposition location. A reliable feedback loop was designed, developed and tested by integration and coordination of several diagnostics with plasma control system (PCS).

The CODAC Framework of HL-2A was established and applied in HL-2A during the experimental campaigns of 2015 and 2016. The framework consisted of four main components: (1) An Simens PLC Ring Network used for slow control interlock and safe interlock; (2) an EPICS network that all of subsystem can get the authorized information of other subsystems using a single protocol; (3) a reflective memory network which acted as the fast control network in ITER CODAC for plasma real time control; (4) and a AVN network for audio and video acquisition and storage.

An integrated analysis of several interdependent diagnostics using Bayesian probability theory was developed and applied in HL-2A, which improved the data analysis. The combination of different probability distribution functions can provide a joint posterior probability over all the physical parameters. In addition, Bayesian graphical models led to a clearly arranged formulation of complex diagnostics models. Details for this method will be introduced.
O-36: Advanced real-time data quality monitoring concept for GEM detector based SXR plasma diagnostics

A. Wojenski\textsuperscript{1}, K. Pozniak\textsuperscript{1}, D. Mazon\textsuperscript{2}, M. Chernyshova\textsuperscript{3}

\textsuperscript{1}Warsaw University of Technology, Institute of Electronic Systems, Nowowiejska 15/19, 00-665 Warsaw, Poland
\textsuperscript{2}CEA, IRFM F-13108 Saint-Paul-lez-Durance, France
\textsuperscript{3}Institute of Plasma Physics and Laser Microfusion, Hery 23, 01-497 Warsaw, Poland

Soft X-ray diagnostics of tokamaks plasma is necessary in order to gather knowledge about physics of the phenomena. One of the most important topics is analysis of plasma impurities. The scientific output can result in new plasma control mechanisms, like real-time magnetic field control based on SXR online analysis feedback. The major difficulty is working with intense plasma radiation. In case of the GEM detector measurements, the analog signals generated by high event rate photos can be malformed. Thus, it is necessary to provide online data quality monitoring of the signals, before performing on measurement data any kind of feedback-based operation. Experiments were performed at CELIA and ASDEX facilities, confirming that very intense radiation will occur in low quality measurement data.

The proposed in the paper unique concept of data quality monitoring will perform several key tasks in term of efficient plasma diagnostics:

- Data filtering – only correct signal data can be further processed by advanced spectra computation algorithms
- Advanced, raw signal GEM detector analysis working in tokamak environment
- Evolution of spectra computation algorithms

The data filtering function is necessary in order to include online SXR analysis to feedback loop of tokamak control systems. However, unique from the SXR spectra computation point of view, function is gathering a large number of raw data events from the GEM detector. The signals can be then analyzed in term of high magnetic field influence, overlapping signals or other phenomena’s. Real data is necessary for development and tests of completely new charge and spectra computation algorithms (also different electronics construction). Important feature is registration of a large number of malformed events in raw data format (analog signals). The valid signals can be streamed directly to data processing blocks. The presented functionality concept is unique for this kind of measurement systems.
We have created a forward model for charge exchange impurity density measurements that incorporates neutral beam attenuation measurements self-consistently. Determining the resistive dissipation of hot plasmas requires knowledge of the impurity content quantified by the effective ionic charge $Z_{\text{eff}}$. Typically $Z_{\text{eff}}$ is determined from visible bremsstrahlung emission, but in limited plasmas with a relatively high edge neutral density, radiation from the neutral particles contributes as much to the visible spectrum as does impurity radiation. Previously, we successfully extracted core impurity density profiles (and thus $Z_{\text{eff}}$) from soft x-ray brightness measurements of bremsstrahlung and impurity recombination emission by using charge exchange impurity density measurements as prior information in an Integrated Data Analysis (IDA). We are now extending the IDA model to include charge exchange impurity density measurements self-consistently. The first step in developing this forward model is to quantify the neutral beam particle density as it traverses the plasma, which itself depends on the total ionic content of the plasma; ion impact and charge-exchange collisions lead to attenuation. We use measurements of the beam emission and particle flux to the opposite wall to constrain the beam attenuation using the model impurity density profiles. While attenuation measurements alone do not provide a unique impurity density measurement in the case of multi-species inhomogeneous plasmas, they do provide significant limits on the possible range of impurity densities. These measurements take advantage of recent detailed calibrations performed during a refurbishing of our 50 kV diagnostic neutral beam to more precisely quantify systematic uncertainties in the impurity density measurements.
O-38: Data analysis and effect corrections of Phase Contrast Imaging diagnostic on HL-2A tokamak

Yi Yu\textsuperscript{1}, Shaobo Gong\textsuperscript{1,2}, Yifan Wu\textsuperscript{1,2}, Boda Yuan\textsuperscript{1,2}, Lin Nie\textsuperscript{2}, Rui Ke\textsuperscript{2}, Ting Long\textsuperscript{2}, Hao Liu\textsuperscript{1,2}, Xuru Duan\textsuperscript{2}, Minyou Ye\textsuperscript{1} and HL-2A team

\textsuperscript{1} School of Nuclear Science and Technology, University of Science and Technology of China, Jinzhai Road 96#, 230026, Hefei, China
\textsuperscript{2} Southwestern Institute of Physics, P. O. Box 432, 610041, Chengdu, China

Phase Contrast Imaging (PCI) diagnostic has recently been developed on HL-2A tokamak to diagnose the fluctuation of plasma density. Laser beam is injected into plasma from bottom port and passes through plasma vertically at a spatial position of 0.6<r/a<0.725. Phase plate with a 1.325 um thick gold coating and with a groove in the middle of it is used to reflect the scattered and unscattered laser beams, respectively. Contrasted signals are detected by 32-channel HgCdTe detectors and data is collected by 5 MHz collectors.

Diffraction effect is carefully discussed in this talk for the port size limitation of PCI diagnostic on HL-2A. The effects of the size of groove on the phase plate, the size of detectors and the phase flicker from thin plasma approximation are calculated and result in a high-wavenumber oscillation and a sudden drop at very small wavenumbers. These will bring great adverse effects on diagnostic data and corresponding data corrections are needed.

To get reliable diagnostic data, absolute calibration of PCI system is carried out by means of using 15 kHz sound waves from a loudspeaker. Data correction factors of phase flicker effect are also measured by frequency scanning. Line integral path effect correction to the phase shift data, which is caused by the spherical wave of the sound waves, is also considered.

Another way to diagnose the plasma density by setting the groove of phase plate outside the unscattered beam center is also presented in this talk, and the PCI system is found to be a spatial band-pass filter.
O-39: Synthetic diagnostic for interpreting the ECE spectrum in LHW-heated plasmas on EAST

Y. Liu\textsuperscript{1}, T. F. Zhou\textsuperscript{1}, A. Ti\textsuperscript{1}, H. L. Zhao\textsuperscript{1} and L. Figini\textsuperscript{2}

\textsuperscript{1}Institute of Plasma Physics, Chinese Academy of Sciences, Hefei 230031, China
\textsuperscript{2}Istituto di Fisica del Plasma, CNR, 20125 Milano, Italy

E-mail of Corresponding Author: liuyong@ipp.ac.cn

Lower hybrid wave (LHW) is the current-drive technique of highest priority on EAST. Over 60s steady-state long-pulse H-mode has been achieved recently under radio-frequency heating on EAST [1]. Characterizing the high-energy electrons driven by the LHW is the foundation of a better understanding of the current-drive physics.

Besides the hard X-ray diagnostic, electron cyclotron emission (ECE) measurement is also a good indicator of the high-energy electrons. ECE diagnostic on EAST has reached a milestone in the past few years [2, 3, 4], and this offers another option to investigate the high-energy electrons. However, it is not straightforward to derive the information of high-energy electrons from the ECE spectrum. In order to obtain the information of the high-energy electrons, SPECE [5] code has been adopted to reproduce the ECE measurements in both ohmic and LHW-heated plasmas on EAST.

References


This work is supported by the National Natural Science Foundation of China (No. 11405211), and the National Magnetic Confinement Fusion Science Program of China under Contract No. 2013GB106002. This work is partly supported by the National Magnetic Confinement Fusion Science Program of China under Contract No. 2015GB101000.
Runaway electrons can be generated in tokamaks in both the flat-top and the disruption phases. The mitigation of runaway electrons is an important topic of tokamak disruption study. The generated runaway electron beam can be highly relativistic and collimated on the parallel direction. To study the behavior of runaway electrons in tokamak, several diagnostic methods have been used in experiments. Recently, it was observed in runaway electron experiments in DIII-D [1] that the Electron Cyclotron Emission (ECE) signal could grow promptly and became much larger than that in normal discharges. The ECE signal was attributed to runaway electron beam rather than thermal electrons, but the mechanism for the strong radiation of runaway electrons in the ECE regime was unclear.

In this work, we present an explanation of this abnormal ECE signal using a newly developed ECE synthetic diagnostic tool for runaway electrons. The tool was developed based on the reciprocal theorem in electrodynamics [2]. Both the emission and the absorption of electron cyclotron waves are calculated using the test particle method. The results of the synthetic diagnostic calculations show that runaway electrons can give significant contributions to the emission of electron cyclotron waves but very little effect to the absorption. Due to the relativistic effect, the radiation of the electron cyclotron wave by runaway electrons is not limited by the resonance layer of thermal electrons. In addition, the excitation of whistler waves, which can scatter the pitch angle of runaway electrons, can strongly increase the radiation power of electron cyclotron wave and lead to the abrupt growth of ECE signal observed in experiments. This work successfully links the abrupt growth of ECE signal with the kinetic instabilities of runaway electron beam, and provides a new approach to diagnose the runaway electron energy and pitch angle distribution. Further work includes applying this tool to other runaway electron experiments and other diagnostic scenarios with more complicated geometries.

References:
O-41: ITER’s Integrated Modelling Infrastructure and Strategy for Data Analysis and Validation


*ITER Organization, Route de Vinon-sur-Verdon, CS 90 046, 13067 St. Paul-lez-Durance Cedex, France*

The ITER Project Requirements (PR) lays down the measurement requirements for ITER operation through a list of plasma parameters to be measured. These measurement requirements are elaborated and delegated to the various diagnostic systems capable of contributing to these measurements through the diagnostics’ System Requirements Documents (SRDs). These SRDs further elucidate the requirements by assigning spatial and temporal resolutions and accuracy to each measurement parameter, which represent the extremes of values for all operational roles (Machine Protection; Basic Machine Control; Advanced Machine Control; Physics) of a particular measurement. The SRDs also assign Diagnostic Contributions to particular diagnostics deemed appropriate for providing relevant measurements toward the particular requirements. Whilst the detailed implementation of the process for combining diagnostic contributions and validating the measurements is yet to begin, the need for a strictly managed, yet flexible, process is recognised, and the infrastructure to support such activities is already maturing.

The ITER Integrated Modelling (IM) Programme will support the ITER Project in the development and execution of the ITER Research Plan (IRP) as well as design basis of the ITER facility during construction, in particular for diagnostics. Strategically, the ITER IM Programme is implemented using expertise and technologies developed within the ITER Members’ research programmes. The Integrated Modelling & Analysis Suite (IMAS) is the software infrastructure that has been developed in response to the needs of the IM Programme and which will support the requirements of both plasma operations and research activities. An agile approach is taken to the development of IMAS and a software management framework consisting of linked issue tracking, source code repositories and a continuous integration server to automatically build and regression test revisions has been established. It is essential that results generated for ITER are reproducible and so software hosting and rigorous version control are prerequisites and already ensured, whilst provenance tracking for handling inputs is still in development.

The unifying element of IMAS is its use of a standardized data model capable of describing both experimental and simulation data. This enables the development of workflows that can flexibly use different software components as well as being independent of the device being modelled. The connection to existing data is facilitated by adopting a plug-in architecture within the IMAS Access Layer. This allows devices to control remote access through existing security arrangements as well manage the mapping of their local data into the standard Data Model and the provision of Machine Description data. Machine Description data is part of the Data Model and is crucial to developing device-independent workflows, especially those involving synthetic diagnostics and elements thereof, which can be validated on existing devices before being used on ITER.
The storage, management and analysis ecosystem of multi-petabyte data sets and multi-terabyte data streams in real time are today’s frontiers of data-intensive applications, led by the proliferation of large scale physics experiments, distributed scientific instruments, sensor networks, web and social media data. Planned large facilities such as the Square Kilometer Array will require the ability to handle and perform analysis on data rates over 1Tb/s. While the data rates for ITER are not anticipated to be at this level, it would be wise to have plans for handling this type of extreme data rate considering the continuous growth in volume, complexity and throughput of fusion relevant diagnostics and sensors. In the near future this scenario will require a paradigm shift in the storage, computing architectures and large scale data analytics to derive timely and innovative insights from fusion data. The SAGE storage system is a novel storage architecture funded by the European Union Horizon 2020 program and developed by Seagate in collaboration with other European partners of the consortium, which includes nuclear fusion facilities and synchrotron sources together with extremely data-centric HPC codes for climate and space weather prediction. The SAGE storage system will be capable of efficiently managing large and complex data sets and data streams with the ability to perform user-defined computations integral to the storage system. The SAGE system is built around the Mero object storage platform and its supporting ecosystem of tools. These all work together to provide the required functionalities and scaling to exascale data analysis pipelines. The SAGE system will seamlessly integrate a new generation of storage device technologies, including non-volatile memories in a tiered hierarchy. In this discussion we will provide an overview of the SAGE architecture, demonstrating the potential strengths of such systems as fusion research becomes more data intensive. We will explain how a co-design approach between industry and researchers has been used to identify requirements across a large cross section of disciplines in order to design a system which meets the needs of a wide range of different users.
OMFIT (One Modeling Framework for Integrated Tasks): An Efficient Community Driven Integrated Modeling Framework

Sterling Smith¹, Orso Meneghini¹, and the OMFIT Team

¹General Atomics, San Diego, CA 92121, USA

In this presentation we provide an overview of the methods and philosophies of OMFIT [1] and show recent uses for fusion analysis and modeling. OMFIT is a Python-based framework for running scientific software on a local or remote workstation and organizing the parsed results in a tree structure GUI (graphical user interface). So far, OMFIT has been employed to run and interpret the results from toroidal fusion plasma related codes and related experimental data, with a large and growing user base. One of OMFIT’s strengths is the ease with which a user can create scripts for carrying out a particular task, and then share that script with other scientists, as attested to by the number of members of the OMFIT team (contributors to the codebase of OMFIT)[2]. Distributed code version control, coordination, and review is managed through GitHub. Stable releases of OMFIT are produced every month or two (the current version is 0.18.4). The collection of scripts used for running and plotting code results are bundled into an OMFIT module. There are over 50 OMFIT modules that interface with many of the sorts of codes that exist for treating the wide range of spatial and temporal scales for toroidal fusion plasmas. OMFIT also interfaces with the major database and file types, including IMAS, allowing ready access to most worldwide device data. In OMFIT, all modeling data and the scripts used to produce the data, as well as all of the settings used, are saved together into a single project, which is useful for being able to reproduce and share results. Among other things, OMFIT has enabled the following: 1) Self consistent pedestal-core calculations; 2) Training of neural networks for core turbulence and edge structure; 3) Discharge design from scratch; 4) Efficient parameter scans for turbulence sensitivity, and power balance and equilibrium uncertainty; 5) Efficient construction of 2D grids for various codes; 6) Edge Deuterium density determination; 7) Consistent treatment of time dependent kinetic profile data fitting; 8) Kinetically constrained equilibrium reconstructions.

References:

Work supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-SC0012656 (GA AToM SciDAC), DE-FG02-95ER54309 (GA theory), DE-FC02-04ER54698 (DIII-D).
Scientific programs are struggling with the size and complexity of their data, in all its forms and formats. One important reason for this is the relative paucity of “navigational” metadata – that is, metadata that explicitly exposes the multitude of relationships between data elements. These sorts of relationships might characterize data source, provenance, the physical properties represented in the data, data versioning, annotation threads, data dictionaries, data catalogs, data shape (which typically determines which applications can consume or display the data) or larger organizational entities like research campaigns, experimental proposals and research products (for example, publications, presentations and public databases). Overall, the current lack of systematic capture and organization of this type of metadata, makes it difficult to navigate complex data structure or to fully utilize or discover available information. It is also clear that data without adequate metadata has limited utility and lifespan, severely limiting the validity and traceability of the scientific insights gained. Filling this gap would have broad applicability in virtually all areas of scientific research and other intellectual pursuits. This project is providing tools that allow data managers to easily develop metadata schemas that represent and expose the multiple and complex relationships that exist between data elements and which are typically not well represented in data systems. Schemas and data are manipulated through a RESTful API (Representational State Transfer - Application Programming Interface). Relationships among the data are represented as mathematical graph structures all built upon a common meta schema. There is an emphasis on recording the full data lifecycle using the provided APIs and granular data object URI schemas, which facilitate instrumenting complex and varied workflows. Upon these same technologies, a modern web based exploration tool is being built. Initial application areas will be developed for plasma physics, ocean monitoring and modeling and Uncertainty Quantification.

Supported by NSF Award 1640829
O-45: Diagnostic Data Handling in the PCS

L. Zabeo, P.C. De Vries, J.A. Snipes, C. Watts, G. Vayakis, A. Winter, M. M. Kocan, M. De Bock

ITER Organization, Route de Vinon-sur-Verdon, CS 90 046, 13067 St. Paul Lez Durance Cedex, France

E-mail of Corresponding Author: luca.zabeo@iter.org

ITER has recently approved the new approach for the path to and through operations, a four-stage approach, which sets out the sequence of ITER installation. The ITER Research Plan that defines the ITER operations is adapted to the machine status, focusing on specific targets compatible with the machine conditions. The first plasma (FP) will be followed by two not-active phases - Pre Fusion Plasma Operation (PFPO-1 and PFPO-2) and the latest Fusion Power Operation (FPO). The development of the Plasma Control System (PCS) will follow the requirements of the four-stage approach. Similar to the PCS, the availability of the various diagnostic systems is spread over the 4 stages, providing the required plasma and machine parameters needed to control the machine, as well as for physics/machine analyses studies.

To ensure adequate control, the necessary measurement and diagnostic requirements should be carefully tracked. This paper focuses on the implementation of the functional interface between diagnostics, measurements and control and event handling functions. The data handling for PCS will be illustrated, through a number of examples and use cases. System interface sheets are used to track information concerning performance for the PCS and the actual diagnostic deliverables over the real-time network. The diagnostic-PCS interface is, however, not one-to-one. The data delivered by the diagnostics may require additional post-processing and validation, before being used as inputs for the PCS control functions. Frequently, data from multiple diagnostics is required to be processed together to provide or improve the actual control input. The diagnostic designer will provide the necessary processing procedure to generate the measurements required by the project, whereas PCS will define some of the supporting functions. CODAC (Control, Data Acquisition and Communication) will guarantee the needed support for the heavy data processing as well as the data distribution among the plant systems. It is important to mention that the PCS not only requires specific plasma measurements but also machine parameters with an appropriate quality and robustness to guarantee plasma control as well as machine protection (as the first line of defence). Examples that will be discussed in here are: magnetic diagnostics for basic control functions like the control of the shape of the plasma requiring the processing of a large number of probes; cameras involved in machine protection functions where complex data interpretation is expected; density for basic control requiring integration of multiple different diagnostics.
Most of all the physics data of the LHD experiment has been managed by the Analysed Data Server system [1], and the total number of the data has increased to more than 7 millions. Some kind of the data are calculated from the raw signal data directly, but others are calculated from other physics data. In order to obtain such a physics data, they must register dependant data in advance. Furthermore, it must be recalculated when the dependant data is updated. In order to launch the calculation program automatically, and to maintain the dependency among the data, the authors have developed Automatic Analysis System (AutoAna) [2]. During the consecutive short pulse experiments, AutoAna provides the latest data soon after the discharge ends by executing the analysis program automatically. It also provides the necessary data to run simulation programs, such as transport analyses. It used to be hard task for the scientists to assemble necessary data to execute simulation program.

In this way, AutoAna has contributed the useful data analysis environments to the scientists. However, being recognized the utility of AutoAna, various requests have aroused. One of them was to run complex analysis programs that took very long time to complete the calculations. Because such programs prevent other existent programs from running, it was impossible to provide the latest summary results during the short pulse experiments while these programs were running. Another request was to provide the user-friendly interface to view the status of the programs, and to run them manually. In this presentation, the recent improvement for overcoming these difficulties and outlook for the future plans of AutoAna will be described in detail.

References:
O-47: Development of unified data analysis in KSTAR

Jayhyun Kim\textsuperscript{1}, Dong Keun Oh\textsuperscript{1}, Insik Choi\textsuperscript{1}, Sang-Hee Hahn\textsuperscript{1}, Sehoon Ko\textsuperscript{1}, Seong-heon Seo\textsuperscript{1}, Jae-Min Kwon\textsuperscript{1}, Byung-ho Park\textsuperscript{1}, and Syun’ichi Shiraiwa\textsuperscript{2}

\textsuperscript{1}National Fusion Research Institute, Daejeon, Republic of Korea
\textsuperscript{2}Massachusetts Institute of Technology, Cambridge, MA, United States

In the experimental magnetic fusion device operated on the shot-base pulses, the experimental data analysis can be largely divided into between-shot analysis for the in-situ adjustment of the experiment and post-session analysis for the intensive computation. Generally, the necessary analysis is added from the data stored in the experimental database (e.g., MDS+ database) to yield valid physical results in each analysis phase. This interpretation requires time-consuming processes such as the extraction of derived data and integrated analysis of the experimental data.

In this work, we established a platform to handle the experimental data of KSTAR device with mainly using πScope [1] which is a python based scientific workbench, in order to perform a series of data interpretation work in a unified workflow [2]. As an application, through the unified analysis system currently being developed, the process from the KSTAR experimental data to the GYRO code for turbulent analysis is established. We will further integrate in-situ analysis and post-processing tool used in KSTAR to consolidate data management from experimental data to final analysis.

References:
Data Management will become increasingly important in the Fusion Community over the next decade as it prepares for ITER. This experiment will generate vast quantities of data - each day generating more data than JET has produced in its entire history. This data rate will bring Fusion into the era of Big Data which will require a paradigm shift in the way data is managed, processed and curated. In common with other large scale experimental facilities ITER will be a truly international experiment which brings additional challenges and opportunities to foster closer collaboration (and competition) between data centres which could host and process fusion data and these opportunities should be grasped.

As an introduction, we discuss why data management is important and what are the key elements to making a successful data management plan. This will bring in the concept of FAIR data (Findable, Accessible, Interoperable and Re-usable) and how we can ensure the data that we produce today will still be a valuable asset for the next generation of scientists.

Experience has shown that there are a number of barriers are common across the different science communities. These barriers are not only technical, but legal and political as well. In almost all geographically distributed projects Federated Identity Management of one of the keys to success, and crosses all three of these barriers. While this is not the only one, we present different solutions to this problem with have been adopted by different experiments and research groups and the key notion of trust in any such system.

Finally, we look at advances in ‘generic’ data management services in different geographic regions (provided by both public and private bodies) and some of the work being done within the Research Data Alliance to try and harmonise and simplify data management in a very standardised way.
Notes
Notes