DATA-EFFICIENT DIGITAL TWINNING STRATEGIES AND SURROGATE MODELS OF QUASILINEAR TURBULENCE IN JET AND STEP

[1]L. ZANISI, [1]J. BUCHANAN, [1,2]T. BROWN, [3]A. BRUNCRONA, [1]F. J. CASSON, [4]J. CITRIN, [5]S. CORRENTI, [1]H. DUDDING, [1]F. ERIKSSON, [6]A. HO, [3]D. JORDAN, [3]A. KIT, [3]A. JÄRVINEN, [5]V. LOMONACO, [7]O. MENEGHINI, [7]T. NEISER, [1]S. PAMELA, [1]B. S. PATEL, [5]L. QUARANTIELLO, [1]C.M ROACH, [1]C. SIDDLE, [1]A. SNODIN, [8]G. SNOEP, A. FERNANDO[4] , T. NORMAN[4], P. HAMEL[4] AND [9]JET CONTRIBUTORS

[1]UK Atomic Energy Authority, Culham Campus, Abingdon, Oxfordshire OX14 3DB, UK; [2]University College London, Gower Street, London, WC1E 6BT, UK; [3]VTT, Kivimiehentie 3, Espoo, Finland; [4]DeepMind, 5 New Street Square, London EC4A 3TW
[5]Universita' degli studi di Pisa, Lungarno Pacinotti 43, 56126 Pisa, IT 56126 Pisa; [6]MIT-PSFC, 77 Massachussets Ave, Cambridge, MA, 02139; [7]GA, 3550 General Atomics Court, San Diego, CA; [8]DIFFER, Eindhoven, NL; [9]See the author list of C. Maggi et al., Nuclear Fusion 64 112012 (2024)

Email: lorenzo.zanisi@ukaea.uk

This work summarises completed and ongoing activities to produce fast neural network (NN) surrogate models of quasilinear models of core turbulent transport in tokamaks. In the parameter space of JET, we complement existing surrogates for the QuaLiKiz model with a NN surrogate for the TGLF model, which captures poloidal shaping and electromagnetic effects. NN surrogates for TGLF have also been generated for the equilibrium space that will be spanned in the I_p ramp up of the STEP power plant. Previous work on Active Learning [8] for data-efficient surrogate construction is extended by means of a new bespoke Active Learning library suitable to a streaming setting which will be directly applicable and relevant to digital twinning of a fusion power plant. Strategies to alleviate the data storage requirements in such a streaming scenario are also explored.

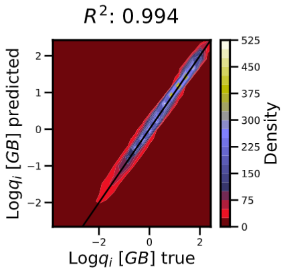

Figure 1. Test performance of TGLF-NN in the JET operating

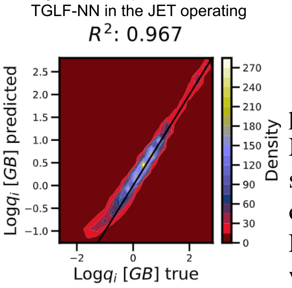

Figure 2. Test performance of TGLF-NN in the STEP ramp up space for the ion heat flux.

Model-based plasma scenario development often excludes core turbulence models, deeming them too slow for iterative applications. Such limitations hinder the optimization of plasma in ITER and DEMO-like devices. For instance, a single transport simulation for the STEP ramp-up (>2000s) with the TGLF quasilinear model can take several weeks. Neural network (NN) surrogate models of quasilinear core turbulence have been devised in the past years. In particular, QLKNN [1], a surrogate model of QuaLiKiz [2] based on the JET operating space, has shown a factor 10^4 prediction speedup. However, the QuaLiKiz model assumes a simple s-alpha geometry and is electrostatic, which is a poor approximation for strongly shaped plasmas. TGLF [3] does not make these assumptions and is therefore, in principle, a more general tool. While TGLF has been validated on some JET discharges, a large-scale validation study on JET will be greatly facilitated by the availability of a suitable surrogate model. Moreover, a large database validation performance comparison in flux-driven integrated modelling of QLKNN and a TGLF NN would help define the respective domains of validity of the models. Finally, a surrogate of TGLF for the STEP ramp up would enable higher fidelity trajectory optimisation compared to current capabilities [4].
Figure 1 shows validation plots for a TGLF NN trained in the JET operating space, with production-ready performance. The JET data from [1] were mapped via IMAS to TGLF inputs, and triangularity, elongation and Shafranov shift were estimated from equilibria using MegPy [5]. The saturation rule SAT1 **[**6**]** was adopted, but extensions to other saturation rules are planned. This NN is directly comparable to the work of [1] for QuaLiKiz but includes three additional shaping dimensions, enabling large scale validation of TGLF for JET and comparisons to QuaLiKiz in flux-driven integrated simulations. Demonstration integrated runs will be presented. Figure 2 shows validation plots for TGLF-NN trained in the STEP I_p ramp up space, where a TGLF dataset was constructed from a database of JETTO simulations with Bohm-gyro-Bohm transport [7]. Figure 2 shows a production-ready surrogate performance. Initial results show that flux-driven integrated runs using the surrogate complete in under 4 hours (down from several weeks when using TGLF) thus opening up higher-fidelity ramp up optimisation studies for STEP.

The generation of large simulation databases is unfeasible for computationally expensive models. Previous work by the authors [8] adopted Active Learning, a strategy that queries the physical model only in regions where additional data would improve the NN performance, and demonstrated a factor of 10 in data efficiency gains compared to random sampling. However, that study was limited to sampling according to the NN uncertainty, which suffers from lack of diversity in each acquisition batch. A more flexible Active Learning library, *Enchanted Surrogates,* based on the recent AI literature [9], has been devised, where simulation runs, surrogate training and Active Learning are integrated in a single workflow. An application is discussed below.

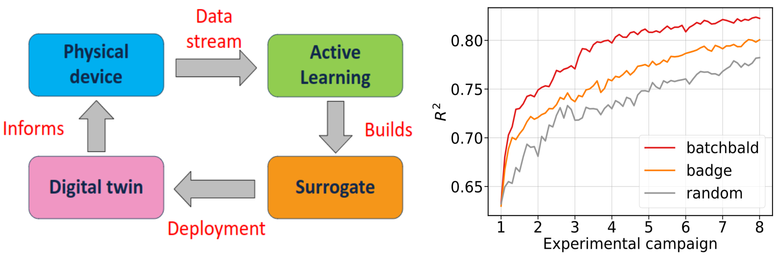

Fig. 3. **Left**: Streaming Active Learning for data-efficient digital twins of tokamak power plants. A time-dependent data stream from a physical device can be used as the input for simulators to support data-efficient surrogate model development via streaming active learning. The surrogates are used in a digital twin to inform the next campaign in, e.g., optimisation tasks. **Right**: An application of Enchanted Surrogates to the dataset by [1] on experiments with progressively higher confinement time and a time-dependent data distribution. Different acquisition functions are shown [9], which achieve higher accuracy than randomly sampling the space and with less data. Results shown for the ion heat flux in QuaLiKiz.

Although constrained by empirical and theory-based extrapolations, the range of plasma states achieved by a machine is in general not fully known at the outset of operations. On the route towards maturation of operations in fusion power plants and reactor-relevant experiments such as ITER, a data stream is produced, which can be exploited in a digital twin infrastructure (Figure 3, left). In this framework, surrogate models based on previous experiments in the stream can be used to inform future campaigns for, e.g., optimisation tasks. The subsequent experimental achievement of more complex and performant plasma scenarios may result in both an increasing volume in parameter space and changing data distributions, thus requiring the acquisition of new simulation training data to update the surrogate. The Active Learning capability of *Enchanted Surrogates* has been adapted to this setting. As a demonstration of a reactor relevant scenario, a mockup sequence of experimental campaigns was devised by sampling historical JET data [4] at increasing values of W_{MHD}, and where QuaLiKiz is the base model. Preliminary results are shown in Figure 3 (right), where the benefit of Active Learning strategies over random sampling in terms of data efficiency is demonstrated.

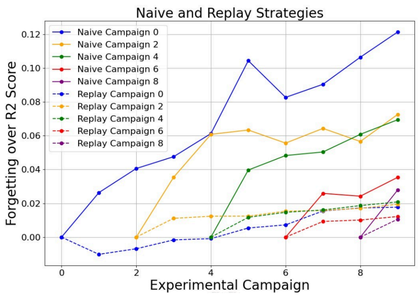

Fig. 5. A Replay strategy with a minimal memory buffer (10000 simulations) achieves lower forgetting compared to a Naive strategy where data from previous campaigns is destroyed. Forgetting is shown only every two experimental campaigns for clarity, and colour indicates the forgetting of a model which is trained starting at that campaign and being evaluated at later times.

Finally, simulations based on experimental campaigns are also expensive to retain in memory. Data storage requirements may imply that some simulations based on earlier experiments need to be destroyed, which would make them unavailable as training data, causing neural surrogate models to be inapplicable to plasma configurations spanned in earlier campaigns (“catastrophic forgetting” [10]). The resulting forgetting, defined as $R^2[data_i, NN_i] - R^2[data_i, NN_j] > 0 \; for \; j > i$, is undesirable as simulation-based discharge planning may need to query low-performance states in the exploration phase of the optimisation. Repeatedly retraining with all previous data can also be costly. This study uses QuaLiKiz, a low memory-footprint model, to benchmark current Continual Learning strategies [11], which identify which simulations should be retained on disk to minimise accuracy loss over the data stream at previous times, as shown in Figure 5 for the total ion heat flux. This study paves the way to applications of Continual Learning to more storage-intensive simulations.

Acknowledgements

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

[1] A. Ho, et al, Physics of Plasmas, 28, 032305,(2021) ; [2] C. Bourdelle et al.,, Plasma Physics and Controlled Fusion, 58,1 (2015); [3] G. M. Staebler et al. Physics of Plasmas, 14, 055909 (2007); [4] J. Mitchell et al., Fusion Engineering and Design, Volume 192 (2023); [5] G. Snoep et al., Phys. Plasmas 30, 063906 (2023); [6] G.M. Staebler et al., Phys. Plasmas, 23:062518, 2016. [7] F. Eriksson et al., 48th European Conference for Plasma Physics (2024),[8] L. Zanisi et al., Nuclear Fusion 64 036022 (2024); [9] D. Holzmueller et al., arXiv:2203:09410 [10] R. M. French and N. Chater, Trends in cognitive sciences, 3 (4):128–135, 1999 [11] L. Wang et al., arXiv:2302.00487 (2023)