Scalable simulation-based inference framework for large-scale validation in fusion

Citation for the original published paper (version of record):

Järvinen, A., Kit, A., Bruncrona, A. et al (2025). Scalable simulation-based inference framework for large-scale validation in fusion. 51st Eps Conference on Plasma Physics Eps 2025: 73-76

(article starts on next page)

Scalable simulation-based inference framework for large-scale validation in fusion

A.E. Järvinen[1], A. Kit[1], A.M. Bruncrona[1], D. Jordan[1], A. Niemelä[1], L. Acerbi[2], A. Bharti[3], T. Fülöp[4], M. Hoppe[5], E. Nardon[6], S. Silburn[7], L. Zanisi[7], and JET Contributors*

[1] *VTT Technical Research Centre of Finland, Espoo, Finland*

[2] *University of Helsinki, Helsinki, Finland*

[3] *Aalto University, Espoo, Finland*

[4] *Chalmers University of Technology, Gothenburg, Sweden*

[5] *KTH Royal Institute of Technology, Stockholm, Sweden*

[6] *CEA, Institute for Magnetic Fusion Research, Saint-Paul-lez-Durance, France*

[7] *UK Atomic Energy Authority, Abingdon, UK*

Simulation-based inference (SBI) is a rapidly developing field aiming to address the inference challenges that are typically related to complex, high-fidelity simulators in science [1]. Computationally costly models, with a selection of uncertain input parameters, are ubiquitous in science, including magnetic confinement fusion research (MCF) [2 – 5]. Due to the input uncertainties, multiple forward passes are typically required in model validation to quantify the input parameter distributions that best reproduce the experimental observations [6]. Solving this inverse problem is one of the key challenges in model validation and must be done algorithmically, such as using SBI, in large-scale validation workflows.

Data-efficient, inverse inference workflows have been demonstrated in model validation applications within MCF [6 – 9]. However, a large-scale adoption of these approaches in model validation activities in MCF has not emerged yet. One of the key entry barriers is the software infrastructure that is needed for large-scale applications. In addition to the SBI algorithms, a large-scale validation workflow on high-performance computing (HPC) platforms requires solutions for overall distributed task orchestration, management of the simulation database, and for processing failed simulations without human-in-the-loop. These requirements align with those needed for data generation for general machine learning surrogate modelling of computationally expensive models [10 – 13]. Therefore, in this work, steps are taken to build a scalable SBI framework on top of the *Enchanted-surrogates*, originally designed for simulation data generation for surrogate models [13].

Enchanted-surrogates software package was initiated in late 2024 to streamline development of surrogate models for computationally demanding physics simulators with HPC platforms (Fig 1). The first physics models that were targeted by this package were the

---

* *See the author list of C.F. Maggi et al. 2024 Nucl. Fusion 64 112012*

core transport model TGLF and linear, ideal pedestal MHD with HELENA and MISHKA [15 – 17]. The present list of models that the development team is actively working on includes also GENE, DREAM, and SOFT [2, 5, 18]. The software package consists of four primary classes: (1) Executors, (2) Samplers, (3) Parsers, and (4) Runners.

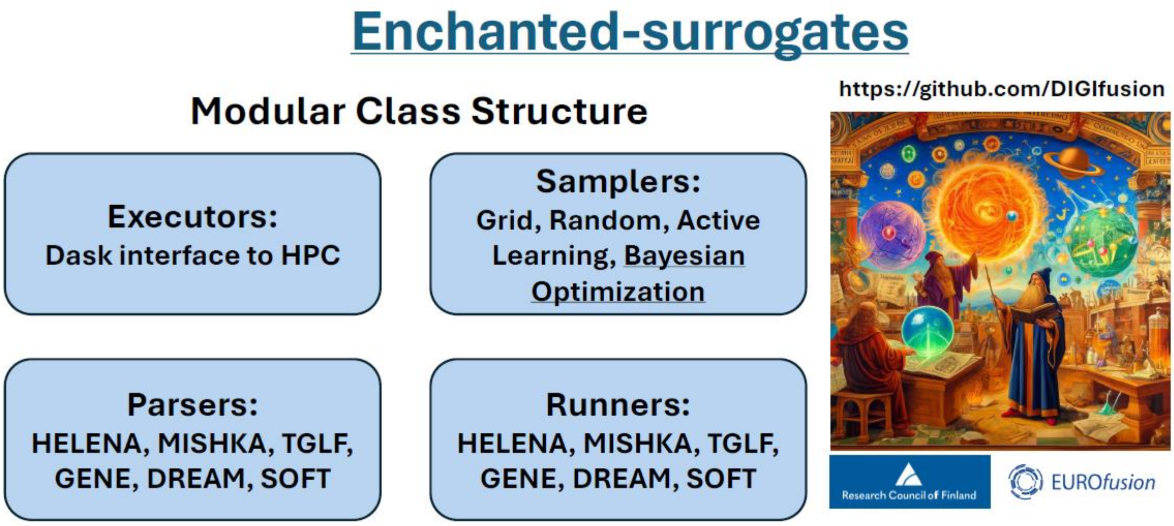

Figure 1. Overview of the Enchanted-surrogate software package [14].

Executors provide the primary interface with the underlying HPC platform and hardware. The present implementation is relying on Dask [19]. The task scheduling systems on HPC platforms are typically not intended for workflows requesting several hundred or more individual tasks in parallel. To solve this challenge, the basic functionality in Enchanted-surrogates is such that Dask reserves the necessary hardware and establishes a virtual cluster for the targeted large-scale sampling procedure. As a result, the HPC platform will not see each sample as a standalone task, but instead the sampling procedure is conducted as a significantly smaller number of larger tasks. Due to the modular structure of Enchanted-surrogates, the present Dask implementation can be supplemented by any other approach that provides the necessary functionality, such as SLURM arrays.

Samplers contain sampling procedures, such as standard grid and random sampling. The present active development is focused on active learning, sparse grids, and Bayesian optimization (BO) samplers [14, 21]. The BO sampler was implemented, using the BoTorch software package, offering a good selection of readily implemented options for Gaussian process regression (GPR) models and acquisition strategies [22]. The procedure to extract the posterior distribution, such as in the BOLFI framework of the ELFI software package [23], is on the list to be implemented soon. Simulation failure can be problematic for any active sampling strategy, such as BO, as the workflow is relying on results from previously launched simulations to guide the acquisition. A simplistic strategy would be to replace the failed sample with a placeholder value far away from the optimum. However, this procedure would severely hamper the GPR fitting procedure by introducing a large, local discontinuity. Instead, the strategy in Enchanted-surrogates is to maintain separate result-dictionaries for successful and failed simulations. Then the primary GPR can be trained only on successful simulations, while the dictionary of failed simulations can be used to gradually reduce the sampling probability in those parts of the operational space that yield a high probability of simulation failure, as was discussed by Chakrabarty et al. in [24]. The latter feature is on the list to be implemented soon.

Parsers and Runners contain everything that is needed to interface Executors and Samplers with the actual physics simulator. This is where custom solutions are needed for each new physics simulator. Parses contains all the necessary input- and output-parsing routines to generate the input files and to post-process the output. In the case of BO sampler, Parsers collect also the simulator output and compare those to the user-provided optimization metric, such as distance to experimental observations. Runners contain the necessary functions to run the simulator executable on the virtual cluster provided by Executors. The entire workflow is controlled by a user-given configuration file.

The developed framework is applied for parameter inference of runaway electron (RE) transport simulations of JET Pulse Number (JPN) 95135 with argon induced disruption and RE beam (Fig. 2) [6, 25]. The study is initialized by simulating the current quench and early plateau with DREAM assuming no radial RE transport with background electron temperature, argon assimilation fraction, and RE seed magnitude optimized with BO to match the evolution of total plasma current, similar to [6]. RE seed profile is assumed to have the same shape as the plasma current prior to the onset of disruption. At the end of this initialization (Current quench) simulations, the synthetic synchrotron image, obtained with SOFT, is qualitatively in agreement with the experimental observation (Fig. 2a).

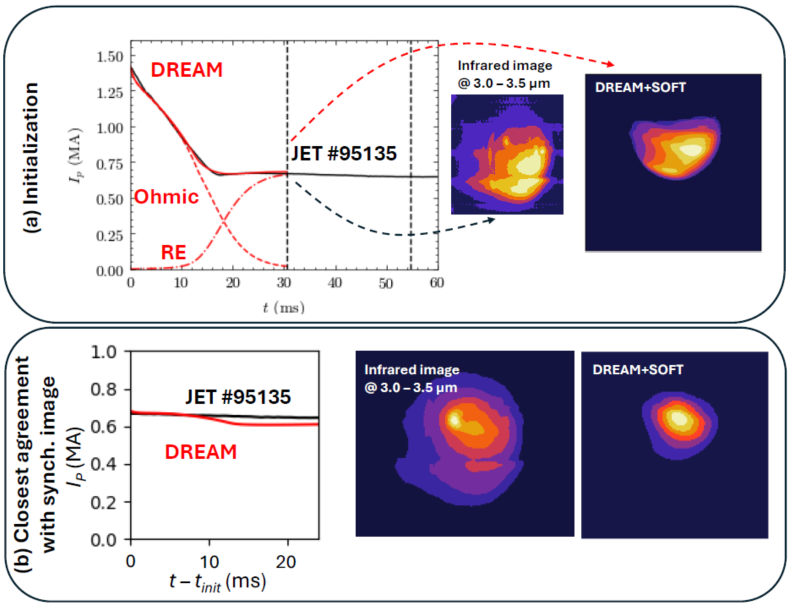

Figure 2. (a) Comparison of the observed and simulated plasma current, I_p, as well as the observed and predicted synchrotron image at the end of the initialization (Current quench) simulations. (b) Comparison of the observed and predicted I_p and synchrotron image at the end of the early RE plateau simulation for the case with the closest agreement with the synchrotron emission pattern.

Starting from this initialized simulation, the early RE plateau is simulated allowing radial diffusion of the RE population via the Rechester-Rosenbluth model [26]. The magnetic field fluctuations are parameterized as magnitude (δB/B) and radial profile proportional to $r^{\alpha-1}e^{-\beta r}$, where r stands for the radial coordinate and α and β are free parameters. In addition, the toroidal loop voltage at the wall is considered a free parameter within the inference workflow. With this setup, the algorithm finds a solution with relatively centrally peaked RE density profile through RE loss at the edge of the plasma, resulting in qualitative shift of the synchrotron pattern as centrally peaked rather than peaking at the high-field side as in the initialization simulation. This shows promising capability of the established workflow in conducting hundreds of kinetic simulations on HPC platform within an inference workflow aiming to match an experimental observation.

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. The work of Aaro Järvinen, Adam Kit, and Amanda Bruncrona has been partially supported by the Research Council of Finland grant no. 355460, and the work of Daniel Jordan and Anna Niemelä by the Research Council of Finland grant no. 358941. The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources. The relevant CSC project numbers are 2004389, 2009007, and 462000451.

References

[1] K. Cranmer, J. Brehmer, and G. Louppe, Proc. Natl. Acad. Sci. **117** (2020) 30055.
[2] F. Jenko, et al. Phys. Plasmas **7** (2000) 1904.
[3] S. Wiesen, et al. J. Nucl. Mat. **463** (2015) 480.
[4] M. Hoelzl, et al. Nucl. Fusion **61** (2021) 065001.
[5] M. Hoppe, et al. Comp. Phys. Commun. **268** (2021) 108098.
[6] A.E. Järvinen, et al. J. Plasma Phys. **88** (2022) 905880612.
[7] E. Crovini, et al. Phys. Plasmas **31** (2024) 063901.
[8] I. Ekmark, et al. J. Plasma Phys. **90** (2024) 905900306.
[9] I. Pusztai, et al. J. Plasma Phys. **89** (2023) 905890204.
[10] A. Bruncrona, et al. submitted to Phys. Plasmas (2025).
[11] A. Kit, et al. This conference.
[12] A. Bruncrona, et al. This conference.
[13] A. Niemelä, et al. This conference
[14] D. Jordan, et al. This conference.
[15] https://github.com/DIGIfusion/enchanted-surrogates/
[16] G.M. Staebler, J.E. Kinsey, and R.E. Waltz, Phys. Plasmas **14** (2007) 055909.
[17] G.T. Huysmans, J. Goedbloed, W. Kerner, Proc. CP90 Conf. on Comp. Phys. Proc. (Amsterdam) (1991).
[18] A.B. Mikhailovskii, G.T.A. Huysmans, Plasma Physics Reports **23** (1997) 844.
[19] M. Hoppe, et al. Nucl. Fusion **58** (2018) 026032.
[20] Dask Development Team (2016). Dask: Library for dynamic task scheduling. http://dask.pydata.org
[21] L. Zanisi, et al. Proceedings of the 30th IAEA-FEC, Chengdu, China, Oct 13 – 18, 2025.
[22] M. Balandat, et al. NeurIPS 2020, Vancouver, Canada.
[23] J. Lintusaari, et al. JMRL **19** (2018) 1 – 7.
[24] A. Chakrabarty, et al. IEEE **53** (2023) 2629.
[25] C. Reux, et al. Phys. Rev. Lett. **126** (2021) 175001.
[26] A.B. Rechester and M.N. Rosenbluth, Phys. Rev. Lett. **40** (1978) 38.