9th Asia-Pacific Conference on Plasma Physics, 21-26 Sep, 2025 at Fukuoka

Transport model surrogates via Gaussian process regression

M. Honda[1], S. Maeyama[2] and E. Narita[1]
[1] Graduate School of Engineering, Kyoto University
[2]National Institute for Fusion Science
e-mail (speaker): honda.mitsuru.5c@kyoto-u.ac.jp

Neural network (NN) surrogate models are widely used for simulating turbulent transport models like TGLF [1], but they typically require extensive datasets, leading to larger model sizes, longer training times, and a lack of uncertainty quantification. To address these limitations, we explore Gaussian Process Regression (GPR), a nonparametric Bayesian method that provides uncertainty estimates and performs well in small-data regimes. The predictive performance and uncertainty estimates of GPR heavily depend on the kernel choice, which defines input space similarity and acts as a prior for the model function. While data characteristics often guide kernel selection, TGLF surrogate construction involves high-dimensional, complex input-output relationships, making a priori kernel determination challenging.

The connection between deep NNs and GPs, established via the central limit theorem, has enabled closed-form expressions for kernels corresponding to error-function (ERF) and ReLU activations. These form the basis of Neural-Net-induced Gaussian Processes (NNGPs) [2]. NNGPs allow for varying weight variances across input neuron connections and separate tuning of weight/bias variances per layer, parameterizing layer-specific scale factors. The high expressivity of NNGPs, often realized through deep ReLU/ERF kernels, arises from a hierarchical, layer-by-layer iterative construction. Adopting these kernels allows GPR to approximate deep learning behavior while retaining probabilistic interpretability.

Scalability is another challenge for GP-based surrogates. Applying GPR to large datasets is computationally prohibitive due to memory and computational costs. Stochastic Variational Gaussian Process (SVGP) [3,4] offers a scalable alternative by introducing a small number of inducing points through which all predictions are made, significantly reducing these costs.

To develop our GP-based surrogate models, we created “dgpr”, a custom GPR library in Python using JAX for automatic differentiation and GPU acceleration. It supports full GP and SVGP with various kernels, including deep kernels. It also supports the Intrinsic Coregionalization Model for multi-output regression. It shows good agreement in benchmarks against GPy and GPflow. From several numerical tests, it was demonstrated that deep kernels are capable of accurately capturing anisotropic structure without prior consideration tailored to the input data, even when the underlying function exhibits distinct behaviors across input dimensions. This property makes them particularly well-suited for surrogate modeling tasks where the characteristics of the input parameters vary significantly across dimensions.

From 147,529 data points generated by GOTRESS [1] with TGLF simulations, 3,000 points were randomly sampled and then split into training (2,700 points) and test (300 points) sets. Figure 1 shows the temperature profile prediction results of GOTRESS simulations using SVGP surrogates (trained on 2,700 points with deep ReLU, Matérn-3/2, and Matérn-5/2 kernels) compared to an NN surrogate and the original TGLF model. All surrogate models successfully reproduced TGLF's temperature profiles. A simulation was also conducted using a reduced training dataset of 40 points. With only 40 data points, prediction becomes feasible even with a full GP, without the need for SVGP. Remarkably, the full GP and SVGP surrogate models with deep ReLU kernel still accurately reproduced temperature profiles in GOTRESS simulations, whereas the NN surrogate failed due to the shortage of the data points. This demonstrates GPR's effectiveness, particularly with deep kernels, for developing robust surrogates in low-data regimes for complex physical simulations.

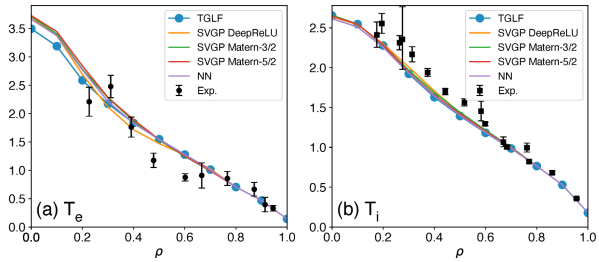

Figure 1. (a) Electron and (b) ion temperature profile predictions using GOTRESS with experimental observations. The results obtained using SVGP surrogate models with three different kernels and the NN surrogate model trained on 2,700 points are compared with those obtained using TGLF. The boundary condition is set at $\rho = 0.8$.

References

[1] M. Honda and E. Narita, Phys. Plasmas 26, 102307 (2019).
[2] G. Pang, L. Yang, and G. E. Karniadakis, J. Comput. Phys. 384, 270–288 (2019).
[3] J. Hensman, N. Fusi, and N. D. Lawrence, “Gaussian processes for big data,” (2013), arXiv:1309.6835 [cs.LG]
[4] J. Hensman, A. Matthews, and Z. Ghahramani, “Scalable variational gaussian process classification,” (2014), arXiv:1411.2005 [stat.ML]