7th Asia-Pacific Conference on Plasma Physics, 12-17 Nov, 2023 at Port Messe Nagoya

# Convolutional neural network models for forecasting heat fluxes calculated by nonlinear gyrokinetic simulations

E. Narita[1], M. Honda[1], S. Maeyama[2] and T.-H. Watanabe[2]
[1] Graduate School of Engineering, Kyoto University,
[2] Graduate School of Science, Nagoya University
e-mail (speaker): narita.emi.5s@kyoto-u.ac.jp

Nonlinear gyrokinetic simulations are performed to calculate heat and particle fluxes caused by turbulent fluctuations in fusion plasmas. In such simulations, the time evolution of the perturbed distribution function $\tilde{f}$ is solved in the five-dimensional phase space. The fluxes required for transport study are estimated in the saturation phase that follows the growing phase. A calculation including the growing and saturation phases takes a few days or more on supercomputers, and the growing phase can consume about one-third of the computational resources. We have developed machine learning based models that investigate processes leading to turbulent saturation using data provided in the growing phase, aiming to infer the fluxes in the saturation phase with a low computational cost.

We first visualized $\tilde{f}$ calculated by the flux-tube gyrokinetic code GKV [1] in the wavenumber-space $(k_x, k_y)$ as shown in the image placed on the left side of figure 1. Such images are generated at every calculation time step. Next, we constructed a convolutional neural network (CNN) model that reads the image and predicts the simulation time at which the image was processed [2]. The CNN model was built by employing transfer learning and fine-tuning techniques based on the EfficientNet-B4 [3], which is a state-of-the-art CNN model trained on a huge number of real-world images. The simulation time predicted by the CNN model can be converted into the time when the saturation phase commences. The saturation time changes due to the initial condition and an early saturation time is desirable for an effective use of the computational resources. We can utilize the CNN model to optimize the initial condition by forecasting the saturation time.

Our first CNN model was unable to predict fluxes since the images fed into the model show $|\tilde{f}|^2$ normalized with the maximum value at each time and do not provide information on the fluctuation amplitude to the model. To extend the capability of the model, we have developed a multimodal model, adding the magnitude of the electrostatic potential $|\tilde{\phi}|^2$ as an input parameter as shown in figure 1. The image and the value of $|\tilde{\phi}|^2$ are handled by the CNN and the multilayer perceptron, respectively, and the generated feature vectors are concatenated to forecast the simulation time and the electron and ion heat fluxes. The multimodal CNN model was trained on data produced by the simulation for the Cyclone base case (CBC), which is a de facto standard DIII-D parameter set for the gyrokinetic simulation benchmarking test. When the CBC-based predictor is applied to the simulation for a JT-60U plasma parameter set, it successfully forecasts the heat fluxes as well as the simulation time as shown in figure 2 [4]. Such high predictability can be attributed to the fact that turbulence of both the CBC and the JT-60U case is dominantly driven by the ion temperature gradient mode/trapped electron mode. If a test case has other dominant instability, a predictor trained on the equivalent case should be applied.

To forecast fluxes ahead of the input image and value, an extended model including a recurrent NN is under development. The preliminary model shows the ability to infer the saturated heat flux from data given in the growing phase, and it could cut the computational cost by about half.

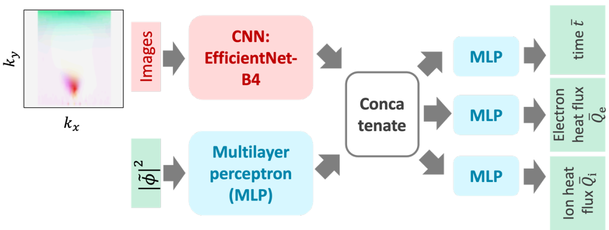


Figure 1 Architecture of the multimodal CNN model.

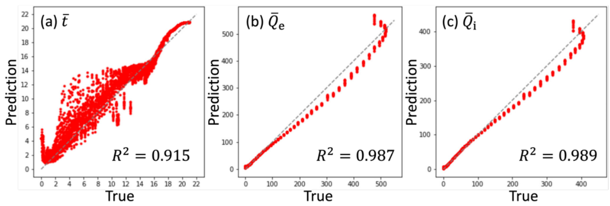


Figure 2 Regression plots of the predicted (a) simulation time and (b) electron and (c) ion heat fluxes against the true values for test data. The determination coefficients $R^2$ are embedded in the plots.

References

[1] T.-H. Watanabe and H. Sugama, Nucl. Fusion, **46** 24 (2006).
[2] E. Narita, M. Honda, S. Maeyama and T.-H. Watanabe, Nucl. Fusion, **62** 086037 (2022).
[3] M. Tan and Q. V. Le, arXiv: 1905.11946 (2019).
[4] M. Honda, E. Narita, S. Maeyama and T.-H. Watanabe, Contrib. Plasma Phys., e202200137 (2023).