

PAPER • OPEN ACCESS

Surrogate model of transport quantities of gyrokinetic simulations on HL-2A

To cite this article: H.J. Zhao *et al* 2026 *Nucl. Fusion* **66** 106034

View the [article online](#) for updates and enhancements.

You may also like

- [Neural networks for turbulent transport prediction in a simplified model of tokamak plasmas](#)
L M Pomârjanschi
- [Quasilinear turbulent particle and heat transport modelling with a neural-network-based approach founded on gyrokinetic calculations and experimental data](#)
E. Narita, M. Honda, M. Nakata et al.
- [Neural operator surrogate models of plasma edge simulations: feasibility and data efficiency](#)
N. Carey, L. Zanisi, S. Pamela et al.



HIDEN
ANALYTICAL

Trusted in Research
for over 40 years





Scan to read the
Physics World Feature

Mass Spectrometry Delivers Fusion Insights

Hidden Analytical featured in **physicsworld**

Discover how Hiden Analytical's specialist mass spectrometry technology, developed through collaboration with Chris Marcus at Oak Ridge National Laboratory, is supporting nuclear fusion research.

H/He Isotope Analysis • Light Gas Separation • Residual Gas Analysis • Fusion Diagnostics

[Click to read the Physics World Feature](#)

www.HidenAnalytical.com

Surrogate model of transport quantities of gyrokinetic simulations on HL-2A

H.J. Zhao¹ , S.M. Li¹ , J.Q. Li², Y.X. Wei³ and Y. Xiao^{1,*} 

¹ Department of Physics, Zhejiang University, Hangzhou 310027, China

² Southwestern Institute of Physics, PO Box 432, Chengdu 610041, China

³ Zhejiang Lab, Hangzhou 311121, China

E-mail: yxiao@zju.edu.cn

Received 25 February 2026, revised 23 August 2026

Accepted for publication 25 August 2026

Published 8 September 2026



Abstract

Artificial intelligence techniques, particularly machine learning, are increasingly being employed to reduce the computational cost of high-fidelity fusion simulations. In this work, nonlinear gyrokinetic simulations are performed to systematically scan the plasma gradient parameter space relevant to HL-2A discharge #27055 plasmas, generating a database of electrostatic drift-wave turbulence for the development of transport surrogate models. A local surrogate model for turbulent transport is first constructed using a feedforward neural network (FNN) trained on 4267 nonlinear local simulation data points, achieving coefficients of determination of up to $R^2 \simeq 0.9$ for key transport coefficients. Building upon this, a second-stage global surrogate model based on support vector regression (SVR) is developed to map local transport predictions to global turbulent transport, trained on 40 global gyrokinetic simulation cases. The resulting global surrogate model is subsequently coupled to an reduced transport solver, which successfully reproduces the steady-state plasma profiles in good agreement with experimental observations. This study provides a practical workflow for predicting global turbulent transport using local surrogate models and machine learning techniques. The proposed approach offers an accurate and computationally efficient strategy for integrated transport simulations, with potential applications in scenario development and control-oriented real-time simulations for present and future fusion devices.

Keywords: turbulent transport, surrogate model, machine learning, integrated modeling

(Some figures may appear in colour only in the online journal)

1. Introduction

Efficient, accurate and reliable integrated modeling [1] is essential for both the design and the operational control of magnetic confinement fusion devices. Among the various physical processes involved, turbulent transport [2] remains

one of the most challenging components to predict, owing to its intrinsically nonlinear, multiscale and kinetic nature [3–10]. In present integrated modeling frameworks, turbulent transport is typically evaluated using theoretically derived quasilinear models [11, 12], empirical or semi-empirical formulae [13], or numerical simulations using reduced models [14]. While these approaches enable practical applications, they generally rely on quasilinear theory or gyrofluid model [9, 15–17], which may not adequately capture key kinetic effects in strongly driven or multiscale turbulence regimes and may also neglect important global physics effects.

Nonlinear gyrokinetic simulations are widely regarded as the most physically self-consistent and reliable framework for modeling turbulent transport in fusion plasmas. However,

* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

their extremely high computational cost severely restricts their direct use in integrated simulations of plasma profile evolution, particularly in scenarios relevant to real-time plasma control. This limitation motivates the development of reduced-order or surrogate models that retain the essential physics of gyrokinetic turbulence while achieving orders-of-magnitude reductions in computational expense.

In recent years, artificial intelligence (AI) and machine learning techniques have been increasingly explored in fusion research, including applications to disruption prediction [18], magnetohydrodynamic instabilities [19, 20], microinstabilities [21, 22], and turbulent transport [2, 14]. Most existing transport surrogate models, however, are trained on data generated from experimental data [23], quasilinear transport model [24, 25] or reduced fluid model [26]. Surrogate models directly based on nonlinear gyrokinetic turbulence simulations remain scarce, primarily due to the prohibitive computational cost of high-fidelity global simulations and the effectively infinite degrees of freedom associated with plasma profile variations, which render brute-force parameter scans impractical.

Local gyrokinetic simulations offer a partial remedy by reducing dimensionality and enabling systematic parameter scans. Nevertheless, a substantial gap remains between local and global gyrokinetic simulations, and the direct application of local-transport-based surrogate models often leads to significant discrepancies when global profile evolution is considered. Bridging this local–global disparity therefore represents a key challenge in developing reliable gyrokinetic-data-driven transport models.

In this work, we propose a methodology that establishes a quantitative mapping between turbulent transport obtained from local profile-setting simulations and that from global profile-setting simulations. This approach enables the construction of a computationally efficient surrogate model that preserves the essential global gyrokinetic physics while reducing the effective computational cost of turbulent transport prediction from millions of CPU-hours to the millisecond level. The methodology provides a practical pathway for incorporating high-fidelity gyrokinetic information into integrated transport modeling and may also offer guidance for the systematic use of reduced-model or local-simulation data in global applications.

High-fidelity turbulence data are generated using the global gyrokinetic code GTC (Gyrokinetic Toroidal Code) [27], based on experimental configurations of the HL-2A tokamak [28]. As a massively parallel particle-in-cell code, GTC has demonstrated excellent scalability and predictive capability in simulations of a wide range of micro- and macro-instabilities in toroidal plasmas, including ion temperature gradient (ITG) mode [29], trapped electron mode (TEM) [30, 31], kinetic ballooning mode (KBM) [32], toroidal Alfvén eigenmode (TAEs) [33], and internal kink mode [34]. The code consistently captures essential nonlinear physics such as turbulent fluctuation characteristics [35], zonal-flow-turbulence interaction [27, 29] and neoclassical effects [36], and has shown good agreement with experimental observations across multiple tokamak regimes [37, 38].

To construct the surrogate models, two widely used machine learning approaches are employed: neural networks (NNs) and support vector regression (SVR). Neural networks provide flexible nonlinear function approximation through multi-layer architectures optimized via backpropagation [39], while SVR, grounded in statistical learning theory, offers strong generalization capability in high-dimensional and limited-sample regimes through kernel-based mappings and structural risk minimization [40].

Based on HL-2A plasma parameters [28], a dataset comprising 762 nonlinear simulation cases is generated using local nonlinear gyrokinetic simulations with systematic plasma gradient scans. In this work, within the parameter space of HL-2A discharge #27055, we construct and validate a surrogate model for turbulent transport using simulation data generated by the first-principles global gyrokinetic code GTC. To the best of our knowledge, this is the first surrogate model trained on first-principles global gyrokinetic simulation data. To address the discrepancy between local and global simulations, an SVR-based local-to-global mapping is constructed, enabling reliable prediction of global turbulent transport from local inputs. The resulting transport surrogate model is further coupled to transport equations to form an reduced transport code. Using particle and heat source profiles inferred from experimental measurements as prescribed inputs, the reduced transport code evolves the density and temperature profiles to a converged steady state that qualitatively agrees with the experimental observations.

The remainder of this paper is organized as follows. Section 2 describes the HL-2A simulation setup and the construction of the local and global gyrokinetic databases. Section 3 presents the neural-network-based surrogate model for local turbulent transport and its validation. Section 4 introduces the SVR-based local-to-global mapping and evaluates its performance. Section 5 applies the surrogate transport model within an reduced transport solver to predict plasma profile evolution. Finally, section 6 summarizes the main results and discusses implications for future integrated modeling efforts.

2. Simulation parameters

We perform GTC gyrokinetic turbulence simulations for a variety of plasma profiles based on the plasma parameters of Shot 27055 from the HL-2A tokamak [28]: on-axis density $n_0 = 3.0 \times 10^{19}/\text{m}^3$, on-axis temperature $T_e = T_i = 1.5 \text{ keV}$, on-axis toroidal magnetic field $B_t = 1.7 \text{ T}$, major radius $R = 1.65 \text{ m}$, minor radius $a = 0.40 \text{ m}$. The safety factor profile, electron density, and ion and electron temperature profiles within the core region at 675 ms in this experiment are shown in figure 1. In this work, the GTC simulations employ a circular cross-section tokamak model, which is sufficient to handle the geometry in the core plasma of HL-2A [41, 42]. We perform gyrokinetic GTC simulations that include kinetic electrons to study ITG and TEM turbulence. Collisions, electromagnetic effects, equilibrium flow and the influence of fast ions are also not considered in this work, as we aim to study

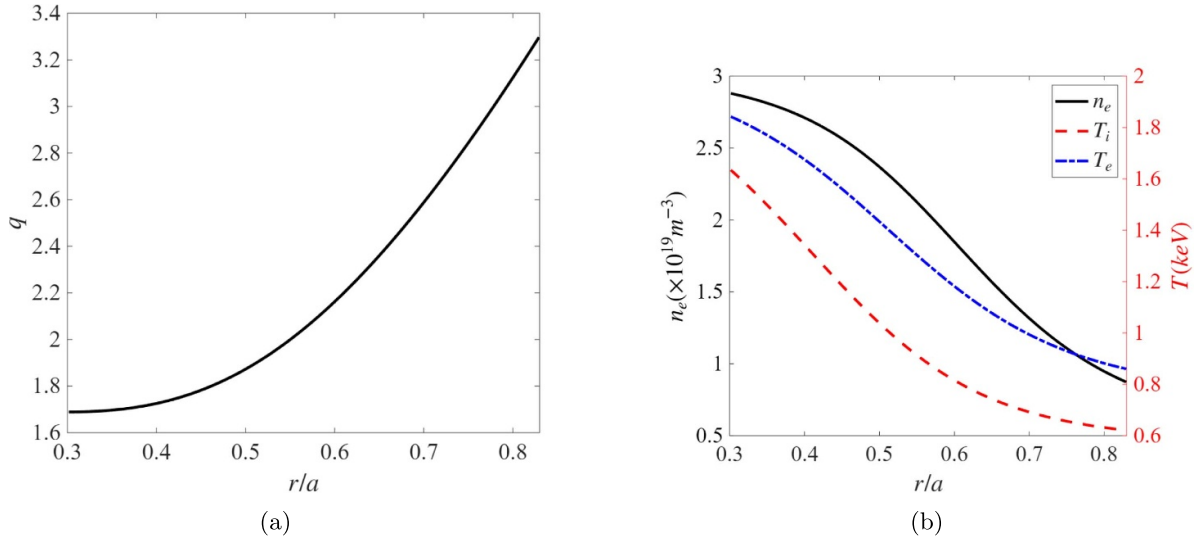


Figure 1. The fitted safety factor q profile (a) and the profiles of electron density n_e , ion temperature T_i , and electron temperature T_e (b) at 675 ms in shot 27 055 on the HL-2A tokamak.

electrostatic drift-wave turbulence in the low beta plasmas of HL-2A. We establish a three-dimensional plasma gradient space by varying the length scale for the three plasma profiles that crucially determine the drift wave instability, i.e. normalized density gradient $R/L_n = -R d \ln(n)/dr$ and normalized temperature gradients $R/L_{T_i} = -R d \ln(T_i)/dr$ and $R/L_{T_e} = -R d \ln(T_e)/dr$, while keeping other plasma parameters unchanged. For each simulation, we calculate the flux-surface average of turbulent transport coefficients at eight equidistant radial points $r_1, r_2, \dots, r_8$, which represent the transport fluxes at these radial positions. To reduce statistical fluctuations of turbulent transport, a temporal average is applied during the saturated phase of nonlinear simulations.

Sensitivity analysis is a crucial step in the construction of surrogate models [18, 19, 43]. Owing to limited computational resources, the present sensitivity analysis is performed in the vicinity of the reference operating point. Specifically, we vary the density gradient scale length R/L_{ne} , electron temperature gradient scale length R/L_{Te} , ITG scale length R/L_{Ti} , temperature ratio τ , safety factor q , device size R_0 , and magnetic shear $\hat{s}$ (reference shear versus reversed shear). For all parameters except $\hat{s}$, five values are considered: $0.5\times$, $0.75\times$, $1.0\times$, $1.25\times$, and $1.5\times$ the reference value. As shown in the figure 2, the three gradient scale lengths exert the strongest influence on the predicted global transport coefficients near the reference operating point, whereas the effects of the other parameters are comparatively weak. This trend is physically reasonable, since the gradient scale lengths directly provide the free energy that drives microinstabilities and turbulent transport, making them the most important features for accurate global transport predictions. Although the present study does not provide an exhaustive exploration of the full parameter space, this local sensitivity analysis nevertheless offers useful insight into the relative importance of different input features around the reference operating point. A more systematic parameter scan and expansion of the training database will be pursued in future work.

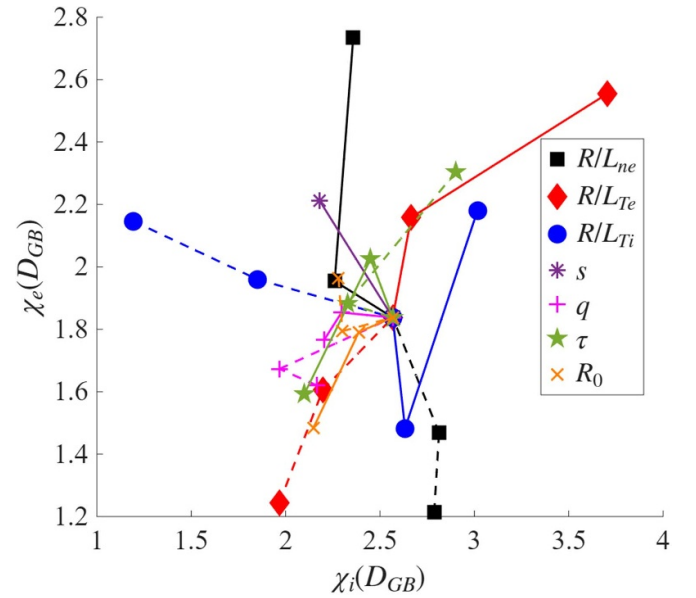


Figure 2. Sensitivity test performed by varying parameters near the experimental reference point. The GTC nonlinear simulation results are shown for the χ_i and χ_e . The scanned parameters include R/L_{ne} , R/L_{Te} , R/L_{Ti} , $\hat{s}$, q , τ , and R_0 . Arrows in the figure indicate the direction of increasing parameter values.

Two methods are used for setting up the plasma profiles in the GTC gyrokinetic simulation, namely the local simulation setting and the global simulation setting. For this local simulation setting, a super-exponential profile [44] $R/L_s = A \exp[-((r-r_0)/w)^6]$ is adopted for the scale lengths of plasma density and temperature profiles, as shown in figure 3, to eliminate boundary effects in the simulation. With this local setting, the sole degree of freedom for the profile variation is the height of the flat top of the plasma gradients. We select 8 values uniformly from 2.2 to 9.9 in the directions of the density gradient and the electron temperature gradient, and select

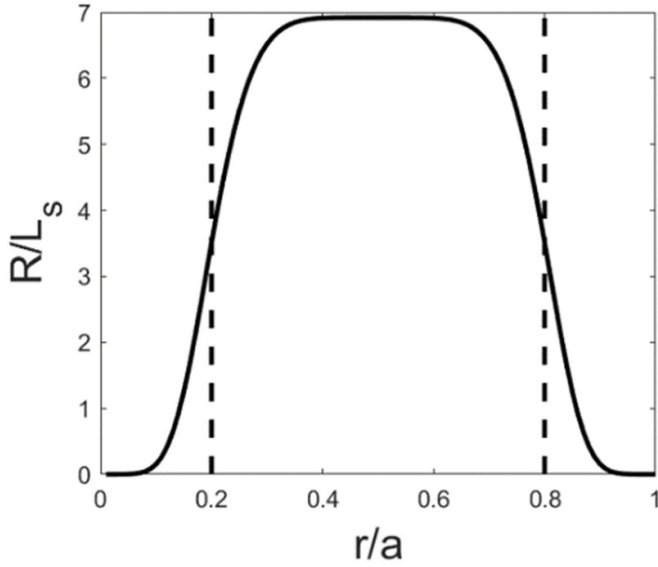


Figure 3. Radial profiles of plasma temperature/density scale length employed in the local gyrokinetic simulation. The vertical dotted line marked the simulation domain.

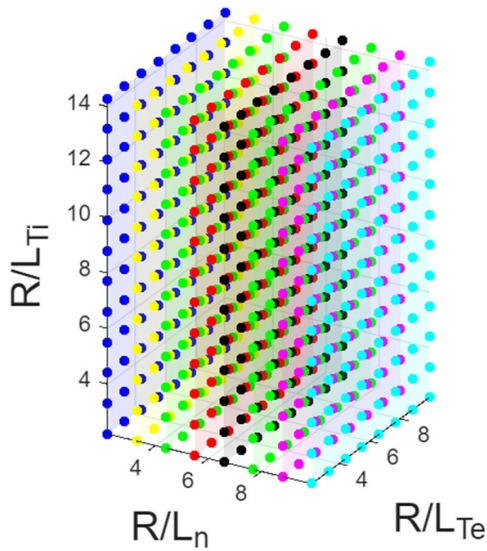


Figure 4. The plasma gradient space used for local simulation scan, with each point represents one particular nonlinear gyrokinetic simulation.

12 values uniformly from 2.2 to 14.3 in the direction of the ITG, forming a 3D plasma gradient space comprising 768 sample points ($8 \times 8 \times 12$), as shown in figure 4. For the selected sample points, some simulation cases turn out to be stable and thereby are excluded from the simulation dataset. The final number of the effective simulation cases is then found to be 762.

For the global simulation setting, the HL-2A experimental plasma profiles are approximated by the function $A(\hat{\psi}) = 1 + c_1(\tanh((c_2 - \hat{\psi})/c_3) - 1)$ distributions, where $A = n_e, T_i, T_e$, and $\hat{\psi}$ is the normalized poloidal magnetic flux, c_1, c_2 , and c_3 are the fitted values to be determined. The corresponding plasma gradient profiles are shown in figure 5. The fitted

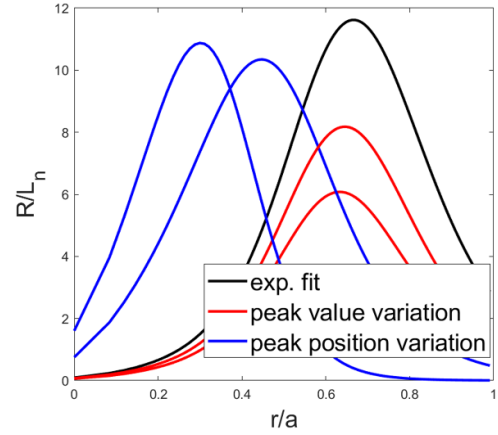


Figure 5. Representative radial profiles for the plasma temperature and density gradients employed in the global GTC gyrokinetic simulations.

analytical gradient profiles exhibit at least 2 degrees of freedom, e.g. peak amplitude and peak position. We note that the variation in the width of the plasma gradient profile is ignored for simplicity. A more extensive exploration of global profiles is left for future work. Consequently, we generated a dataset of global gradient profiles comprising 50 nonlinear simulation cases by randomly varying the peak amplitudes and positions. The real experimental profiles may be more complex than these analytical global profiles. Therefore, it is useful to build a mapping between the transport from the local simulation and that from the global simulation. In this way, we may use the transport from the local simulation to predict turbulent transport in the global simulation. We note that the local-profile and global-profile simulations considered in this work have comparable computational costs, since both involve nonlinear GTC simulations. The purpose of introducing the two-stage surrogate is not to reduce the cost of individual simulations, but to decompose a high-dimensional global learning problem into two lower-dimensional learning tasks. This hierarchical strategy provides a practical route for constructing surrogate models that incorporate global physics while mitigating the curse of dimensionality associated with arbitrary global profile variations. It is conceptually related to the pretraining framework proposed in [45], in which a simpler model is used to facilitate learning of a more complex one.

3. Surrogate model for transport in local simulation

In the beginning, eight radial positions are sampled in each local simulation, with dimensionless parameters $(R/L_n, R/L_{Ti}, R/L_{Te}, r/a, q)$ as input variables and transport coefficients χ_i, χ_e, D_i as target variables. This strategy generates a dataset with a total of $762 \times 8 = 6096$ data points, where all transport coefficients are normalized by the GyroBohm unit $D_{GB} = (\rho_i/R_0)(T_e/eB)$. Following common practice in machine learning [46], the dataset is randomly split into training, validation, and test sets in a 7:1.5:1.5 ratio, corresponding

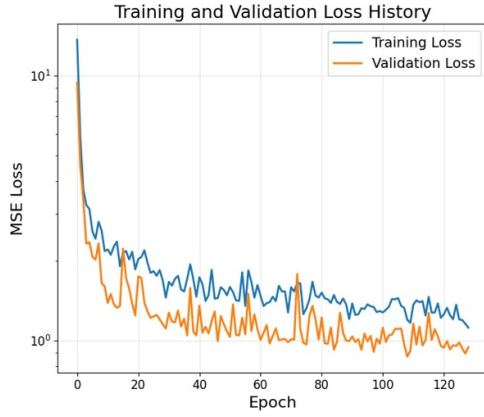


Figure 6. Training history of the loss function on the training and validation datasets for the local profile transport simulation surrogate model.

to 4267, 914, and 915 data points, respectively. The validation set is employed to iteratively evaluate model performance, mitigate overfitting, optimize hyperparameters, and identify the optimal convergence state. This plays a critical role in ensuring the generalization performance of the surrogate model. A local transport surrogate model is constructed and optimized using a feedforward neural network (FNN) for the 1D transport coefficients χ_i , χ_e , and D_i . The FNN has an input dimension of 5 and an output dimension of 3, with four hidden layers, each containing 535 neurons. A ReLU activation function is employed, together with a dropout rate of 0.033 15 for regularization. To ensure generalization and guard against overfitting, training and validation losses were monitored throughout the process, and their evolution is presented in figure 6. An early-stopping strategy was applied to retain the checkpoint with the lowest validation loss. Hyperparameters, including architectural choices and optimization settings, were tuned using a Tree-structured Parzen Estimator within the Optuna framework, with the validation loss as the objective. The consistent convergence of the two loss curves, together with the strong performance on an independent test set, supports the reliability of the trained model. As shown in figure 7, the FNN model achieves a regression coefficient $R^2 > 0.9$ for all target transport coefficients in the test set, demonstrating that the model can accurately predict turbulent transport from local GTC gyrokinetic simulations.

In particular, figure 8 compares the transport levels predicted by the local surrogate model, indicated by red triangles, with those obtained from local gyrokinetic simulations using the GTC code, shown by blue solid curve. The predicted transport exhibits excellent agreement with the simulated results, as shown in figure 8, thereby further validating the reliability of the local surrogate model. The mean absolute percentage error (MAPE) quantifies the prediction error for a given case or sample point, which is defined as

$$\text{MAPE}(y, \tilde{y}) = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{|y_i - \tilde{y}_i|}{|y_i|} \quad (1)$$

where N_s is the number of radial points, y_i denotes the simulated value at the i th radial point, while $\tilde{y}_i$ represents the

corresponding predicted value. The MAPEs in figure 8 are found to be consistently below 8% for all three transport coefficients, demonstrating the high accuracy of the local surrogate model.

4. Local–global mapping

Next, a machine-learning-based mapping scheme is proposed to extend the local surrogate model to global transport prediction for arbitrary plasma profiles. The workflow, illustrated in figure 9, consists of the following five steps: (1) perform a sufficient number of local gyrokinetic simulations to generate the training dataset for the FNN-based local surrogate model, as described in Sections 2–3; (2) carry out global gyrokinetic simulations to produce a smaller dataset covering a range of global plasma profiles; (3) identify the essential input features; (4) establish the local-to-global mapping using SVR, with global GTC transport coefficients as the target outputs; (5) verify the accuracy of the local-to-global mapping using an independent test dataset.

As elaborated in the aforementioned workflow, the SVR approach is adopted to establish a Local–Global mapping that bridges the local simulation outputs and global simulation results. For the SVR Local–Global mapping, we choose the input features as radial coordinates, plasma gradients, and transport quantities predicted by the local surrogate at the eight fixed radial sampling positions, i.e. $R/L_n(r_j)$, $R/L_T(r_j)$, $R/L_{Te}(r_j)$, r_j/a , $q(r_j)$, $\chi_{i,\text{local}}(r_j)$, $\chi_{e,\text{local}}(r_j)$, $D_{i,\text{local}}(r_j)$ for $j = 1, 2, \dots, 8$. Correspondingly, the output targets are defined as the radial transport coefficients obtained from global GTC gyrokinetic simulations, i.e. $\chi_{i,\text{global}}(r_j)$, $\chi_{e,\text{global}}(r_j)$, and $D_{i,\text{global}}(r_j)$. Different from the training strategy employed in the local surrogate model, the present method incorporates all eight radial sampling points from each global profile case as a single integrated training sample, which enables the model to better characterize the physical correlations across radial locations. In total, 50 global profile simulation cases are utilized to build the SVR Local–Global mapping, among which 40 cases are allocated to the training set and the remaining 10 cases form the independent test set. The SVR model employs a radial basis function (RBF) kernel with hyperparameters $C = 6$, $\epsilon = 0.05$, and γ set to the ‘scale’ option implemented in the scikit-learn library. As illustrated in figure 10, the constructed global surrogate model using SVR mapping achieves satisfactory predictive accuracy for global transport coefficients over the test set with $R^2 > 80\%$.

Although the regression performance on the full gyrokinetic dataset is not particularly high compared with datasets derived from fluid simulations or quasilinear theory [2, 24], the model nevertheless provides relatively accurate predictions of the transport coefficients, both qualitatively and quantitatively. Figure 11 presents a representative comparison in which the transport coefficients predicted by the global surrogate model at eight preselected radial locations are evaluated against results from global GTC simulations. The global surrogate model predictions (red dots) show excellent agreement with

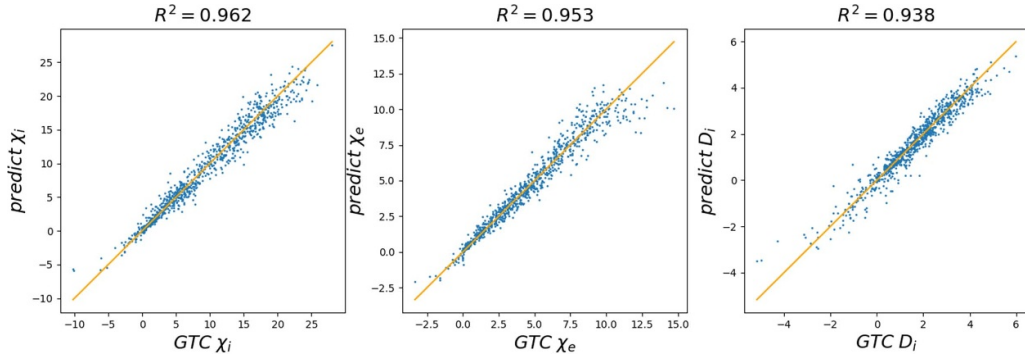


Figure 7. Regression surrogate model for transport coefficients χ_i, χ_e , and D_i , of local simulation with $R^2 > 0.9$ for all the quantities.

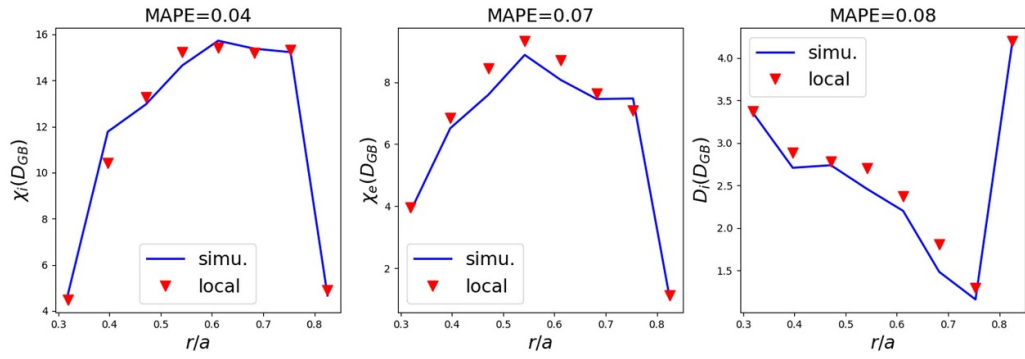


Figure 8. Transport predicted by the local surrogate model (red triangles) is compared to that given by the GTC local simulation in the test dataset (blue solid curve) for transport quantities (a) χ_i , (b) χ_e , (c) D_i .

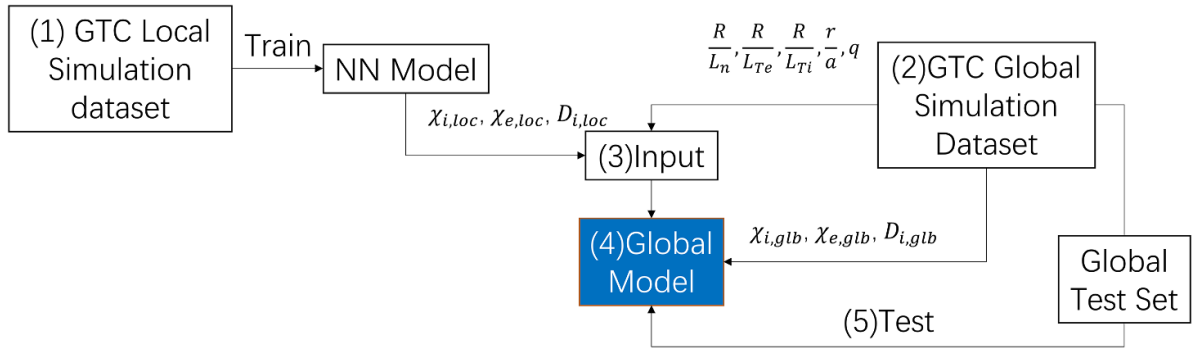


Figure 9. Workflow for constructing the global model.

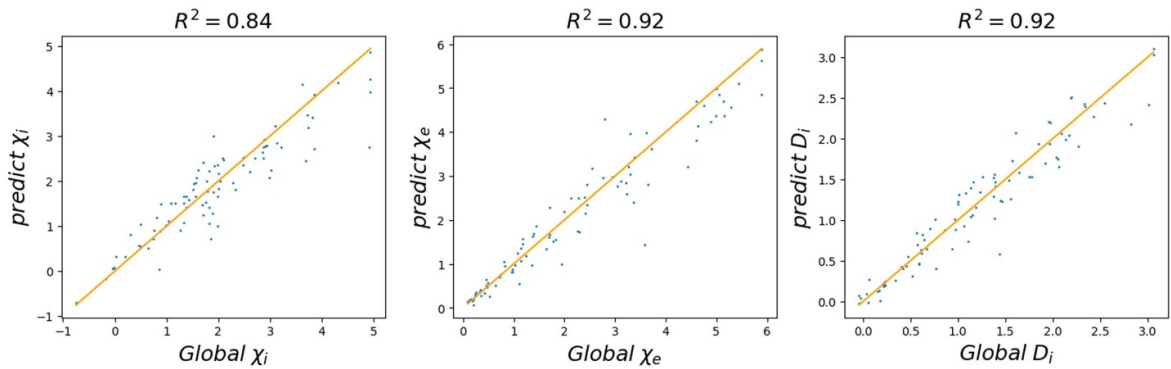


Figure 10. Regression for transport coefficients χ_i, χ_e, D_i from global surrogate model using the global surrogate model.

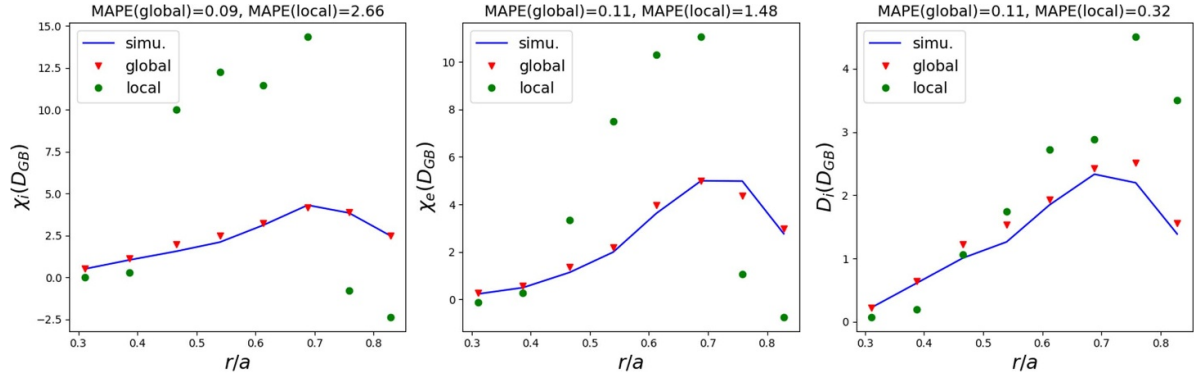


Figure 11. Comparison of transport coefficient predictions from surrogate models with results from global gyrokinetic simulations for a representative case.

the global simulation results (blue curves). For all three transport coefficients, the MAPEs are around 10%, indicating the high predictive accuracy of the global surrogate model.

By contrast, predictions obtained using the local surrogate model alone (green dots) exhibit substantial deviations from the simulation results, with MAPEs exceeding 100%, as shown in figure 11. These results not only validate the effectiveness of the proposed mapping model but also, more importantly, reveal a pronounced discrepancy in turbulent transport characteristics between local and global simulations, even when evaluated at identical radial positions with identical gradient scale lengths. This discrepancy is physically reasonable, as global transport inherently involves nonlocal effects and complex global physics, such as turbulent spreading [29]. The observed differences therefore clearly demonstrate both the necessity and the efficacy of developing the proposed mapping workflow.

We progressively increased the number of samples in the training set and evaluated the corresponding regression performance of the transport coefficients on the test set. As shown in figure 12, the coefficient of determination (R^2) converges when the training set size reaches approximately 35, indicating that a large number of global gyrokinetic simulation samples is not required to effectively reproduce the essential global transport characteristics. This result strongly supports the feasibility of constructing a global surrogate model using only a small number of global gyrokinetic simulations in combination with the local surrogate model.

5. Prediction of plasma profile evolution

Building upon the established global surrogate model, this study develops an reduced transport solver to predict plasma profile evolution. The solver employs coupled transport equations in cylindrical coordinates for the density $N(r)$, electron temperature $T_e(r)$, and ion temperature $T_i(r)$. Turbulent transport coefficients derived self-consistently from the global surrogate model are incorporated into the transport equations, as shown in equations (1)–(3). The formulation is similar to that adopted in the neoclassical transport theory of [47]. The particle source $S_N(r)$, electron and ion heat source

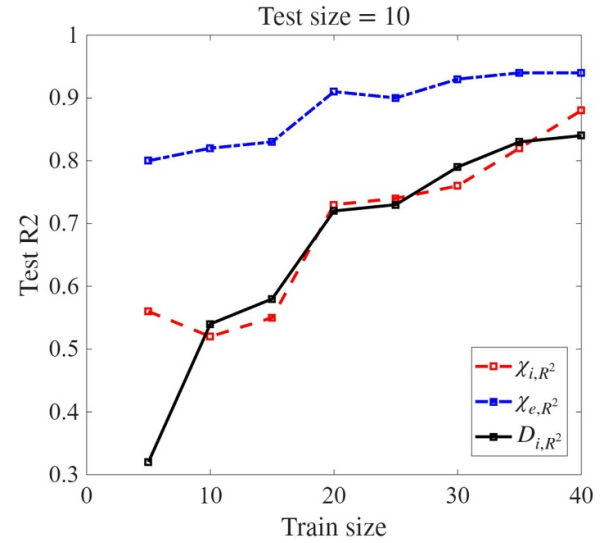


Figure 12. Coefficient of determination (R^2) of global transport coefficients with the size of the training dataset, shows some convergence when training set size reaches 35.

$S_{pe}(r)$, $S_{pi}(r)$ are implemented to sustain the prescribed plasma pressure profile. Order-of-magnitude estimates indicate that collisional time-scale processes—such as Ohmic heating, neo-classical transport, collisional heat exchange, and viscosity—are approximately one order of magnitude smaller than turbulent transport in the core region ($r/a < 0.8$). Therefore, the present solver retains only turbulent transport effects. A circular cross-section magnetic configuration is assumed, which justifies the use of approximate cylindrical coordinates. In addition, magnetic flux surfaces are taken to remain stationary over the time scale of interest. Consequently, this work focuses primarily on the influence of turbulent transport and external heat sources on plasma profile evolution.

The transport solver adopts the energy flux formulation used in the GTC code,

$$Q_E = -N\chi\partial_r T + 3/2\Gamma T,$$

where Q_E is the radial energy flux, and the particle flux is expressed as $\Gamma = -D\partial_r N$.

Based on the preceding physical considerations and discussion, the transport equations can be written as follows:

$$\frac{\partial N}{\partial t} - \frac{1}{r} \frac{\partial}{\partial r} \left(rD(r) \frac{\partial N}{\partial r} \right) = S_N(r) \quad (2a)$$

$$\frac{3}{2} N \frac{\partial T_i}{\partial t} + \frac{3}{2} T_i \frac{\partial N}{\partial t} - \frac{1}{r} \frac{\partial}{\partial r} \left(rN\chi_i(r) \frac{\partial T_i}{\partial r} + r \frac{3}{2} T_i D(r) \frac{\partial N}{\partial r} \right) = S_{pi} \quad (2b)$$

$$\frac{3}{2} N \frac{\partial T_e}{\partial t} + \frac{3}{2} T_e \frac{\partial N}{\partial t} - \frac{1}{r} \frac{\partial}{\partial r} \left(rN\chi_e(r) \frac{\partial T_e}{\partial r} + r \frac{3}{2} T_e D(r) \frac{\partial N}{\partial r} \right) = S_{pe}. \quad (2c)$$

In the preceding equations, the heating source profiles are computed with the NUBEAM [48] Monte Carlo code using HL-2A experimental inputs such as neutral beam injection (NBI) parameters and plasma profiles. The particle source, which is difficult to determine directly from diagnostics or simulations, is instead derived inversely: given the transport coefficients from GTC simulations, the particle source is adjusted such that the density profile reaches a steady state. The source profiles are shown in figure 13. Owing to the complexity of edge physics, which lies beyond the scope of the present study, the experimental measurements are imposed as fixed boundary conditions. Spatial discretization is performed using a second-order central difference scheme, while temporal advancement is treated implicitly. The resulting finite-difference form of the transport equation is given as follows,

$$\frac{N_j^{n+1} - N_j^n}{\Delta t} - \frac{1}{r_j \Delta r^2} \left[r_{j+1/2} D_{j+1/2}^n (N_{j+1}^{n+1} - N_j^{n+1}) - r_{j-1/2} D_{j-1/2}^n (N_j^{n+1} - N_{j-1}^{n+1}) \right] = S_{N,j} \quad (3a)$$

$$\begin{aligned} & \frac{3}{2} N_j^{n+1} \frac{T_{i,j}^{n+1} - T_{i,j}^n}{\Delta t} + \frac{3}{2} T_{i,j}^{n+1} \frac{N_j^{n+1} - N_j^n}{\Delta t} \\ & - \frac{1}{r_j \Delta r^2} \left[r_{j+1/2} N_{j+1/2}^{n+1} \chi_{i,j+1/2}^n (T_{i,j+1}^{n+1} - T_{i,j}^{n+1}) - r_{j-1/2} N_{j-1/2}^{n+1} \chi_{i,j-1/2}^n (T_{i,j}^{n+1} - T_{i,j-1}^{n+1}) \right] \\ & - \frac{1}{r_j \Delta r^2} \left[r_{j+1/2} T_{i,j+1/2}^{n+1} D_{j+1/2}^n (N_{j+1}^{n+1} - N_j^{n+1}) - r_{j-1/2} T_{i,j-1/2}^{n+1} D_{j-1/2}^n (N_j^{n+1} - N_{j-1}^{n+1}) \right] = \frac{3}{2} S_{pi} \end{aligned} \quad (3b)$$

$$\begin{aligned} & \frac{3}{2} N_j^{n+1} \frac{T_{e,j}^{n+1} - T_{e,j}^n}{\Delta t} + \frac{3}{2} T_{e,j}^{n+1} \frac{N_j^{n+1} - N_j^n}{\Delta t} \\ & - \frac{1}{r_j \Delta r^2} \left[r_{j+1/2} N_{j+1/2}^{n+1} \chi_{e,j+1/2}^n (T_{e,j+1}^{n+1} - T_{e,j}^{n+1}) - r_{j-1/2} N_{j-1/2}^{n+1} \chi_{e,j-1/2}^n (T_{e,j}^{n+1} - T_{e,j-1}^{n+1}) \right] \\ & - \frac{1}{r_j \Delta r^2} \left[r_{j+1/2} T_{e,j+1/2}^{n+1} D_{j+1/2}^n (N_{j+1}^{n+1} - N_j^{n+1}) - r_{j-1/2} T_{e,j-1/2}^{n+1} D_{j-1/2}^n (N_j^{n+1} - N_{j-1}^{n+1}) \right] = \frac{3}{2} S_{pe}, \end{aligned} \quad (3c)$$

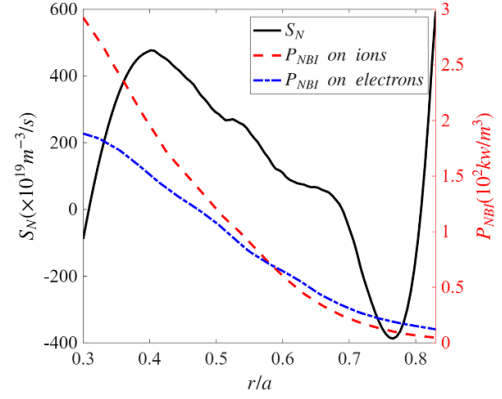


Figure 13. The particle source profiles $S_N(r)$ inferred from both experimental and simulation data, along with the NUBEAM-simulated NBI power deposition profiles $P_{NBI}(r)$ on ions and electrons based on the experimental profiles, were provided by the experimental group.

where n denotes the time-step index and j the radial grid index. Here, Δt and Δr represent the temporal and spatial step sizes, respectively. The notation $f_{j\pm 1/2}$ is defined as $(f_{j+1} + f_j)/2$, where f denotes an arbitrary physical quantity.

As shown in figure 14, the simulated steady-state density, electron temperature and ion temperature profiles obtained using the global surrogate model exhibits good qualitative agreement with the experimental measurements. Quantitatively, the deviation between the predicted and measured profiles remains within an acceptable range $\text{MAPE} \approx 10\%$, thereby demonstrating the predictive capability of the reduced transport solver. In contrast, the local model employed in figure 15 produces pronounced deviations from the experimental data in the evolution of the ion and electron temperature profiles, while the density profile shows slightly better agreement. Overall, the local model exhibits poor numerical convergence. Notably, even without incorporating complex edge physics, the global surrogate model successfully reproduces the essential transport characteristics in the plasma core region, highlighting both its physical reliability and the numerical robustness of the transport scheme.

Nevertheless, discrepancies persist between the simulated profile evolution and the experimental observations. These differences are primarily attributed to incomplete experimental input. In the present study, only the heat deposition profile associated with NBI was available, whereas particle fueling was not included.

Despite these limitations, the proposed global transport model demonstrates considerable promise in capturing macroscopic transport behavior. Future work will aim to develop a unified and fully integrated simulation framework that couples the present transport model with additional physics modules calibrated against a broader set of experimental measurements. In particular, the enhanced framework will incorporate multiple heat and particle sources consistent with experimental conditions, together with neoclassical and radiative transport effects, and will be systematically benchmarked against other integrated simulation codes. These developments will further improve the model's physical fidelity and strengthen

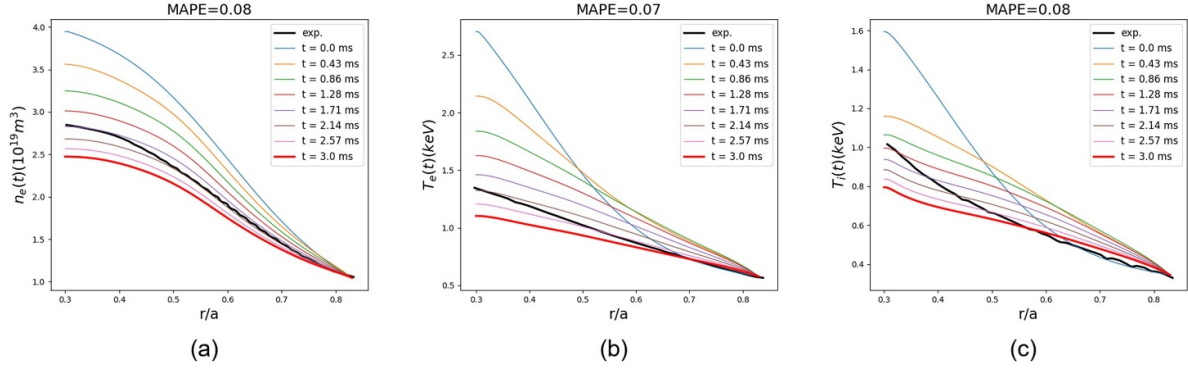


Figure 14. Evolution of (a) the electron density n_e , (b) the electron temperature T_e , and (c) the ion temperature T_i relative to their respective experimental profiles, obtained from the global surrogate model.

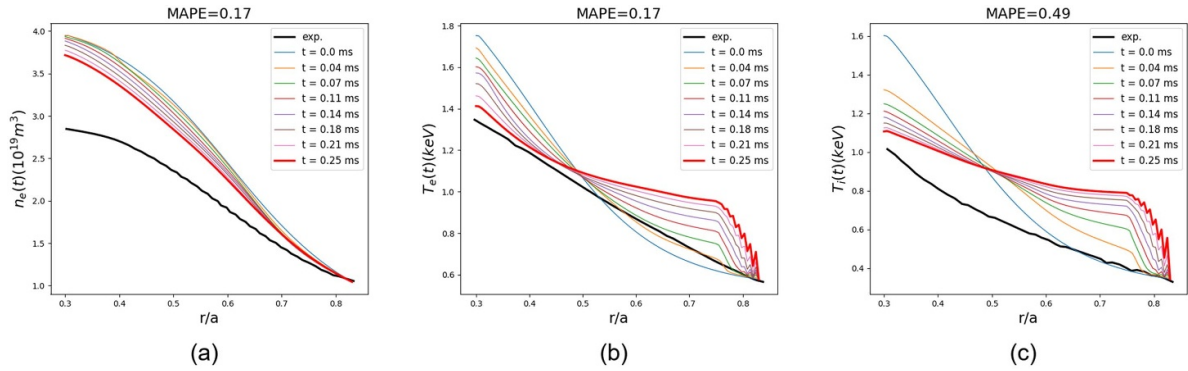


Figure 15. Evolution of (a) n_e , (b) T_e , and (c) T_i relative to the experimental profiles, evolved with the local surrogate model in the reduced transport code.

its quantitative predictive capability across an expanded range of operating regimes.

Our surrogate transport model can be integrated as a drop in replacement for empirical models in a real time framework like RAPTOR [49]. At each time step, it evaluates local profiles T_e, T_i, n_e and returns χ_i, χ_e, D_i in microseconds, advancing the transport PDEs. The solver runs inside a state estimator such as an extended Kalman filter [50], fusing predictions with diagnostics; the estimated state then drives model based controllers. This preserves the existing loop while improving the physical fidelity of the transport model.

6. Conclusion and discussion

This study presents a practical surrogate modeling approach for predicting turbulent transport in HL-2A tokamak plasmas by integrating high-fidelity gyrokinetic simulations with machine learning techniques. A two-stage modeling strategy is proposed to bridge the gap between the large freedom in plasma profiles and accurate global transport predictions. Specifically, a neural-network surrogate is developed for local simulations, achieving $R^2 > 0.9$. An SVR-based surrogate model is then constructed to map local transport to

global transport. This two-stage modeling strategy effectively overcomes the limitations inherent in purely local and purely global surrogate models.

The proposed methodology offers substantial computational advantages by significantly reducing reliance on computationally expensive full-fidelity gyrokinetic simulations while preserving essential physical accuracy. The local surrogate model is trained, validated and tested on 762 nonlinear simulation cases, while the training and testing of the global surrogate model requires only 50 nonlinear simulation cases, highlighting the efficiency of this methodology and its potential for accelerating integrated simulation workflows. The significant differences between the turbulent transport results from local simulations and those from global simulations demonstrate the necessity of developing global surrogate model, which can incorporate global turbulent effects such as turbulence spreading. The reduced transport code based on the global surrogate model exhibit good agreement with experimental measurements on the steady-state plasma profiles, particularly in the core region where turbulent transport dominates, as shown in figure 14. This agreement supports the physical relevance of the global surrogate model, despite simplifications such as the neglect of neoclassical transport, the latter being justified by its contribution being at least one order of magnitude smaller in the core plasmas.

Although the machine learning models demonstrate good predictive performance, further developments could enhance their physical interpretability. Future work may incorporate physics-guided constraints or symbolic regression to extract underlying transport scalings and physical relationships. While the current models perform reliably within the explored parameter space, extrapolation to significantly different operational regimes may present grand challenges. Active learning strategies could be employed to improve model robustness and generalization. The proposed surrogate modeling method also shows promises for real-time applications, although additional validation over broader operational regimes is warranted. Extensions of the parameter space, including variations in the safety factor profile $q(r)$ and magnetic shear s , as well as the inclusion of electromagnetic effects to address high- β scenarios, represent important directions for further research. The present surrogate model is limited to HL-2A discharges resembling shot #27055. Given the narrow operating range of HL-2A (concentric core flux surfaces, limited variations in magnetic field, density, and pressure), we expect that the model would perform similarly for other typical HL-2A discharges or other devices, with possible differences easily mapped using the proposed Local–Global approach. Extension to other devices and scenarios is left for future work. It should be noted that our reduced transport solver is developed in cylindrical geometry with specified source profiles, and does not include self-consistent equilibrium, neoclassical effects, detailed heating, or edge physics. Furthermore, due to the lack of time-resolved experimental data, we are limited to examining profile evolution toward steady state and cannot obtain time-dependent results for direct experimental comparison. Moreover, the fixed boundary condition approximation adopted here could also be refined in future studies. Hence the present model serves as a reduced framework rather than a fully integrated transport solver.

Overall, this work establishes a valuable approach for developing turbulent transport surrogate models by combining first-principles simulations with machine learning in fusion research. The proposed methodology addresses the immediate need for efficient and accurate transport modeling. The demonstrated accuracy and efficiency suggest strong potential for applications in scenario development, real-time control, and uncertainty quantification in future fusion devices.

Acknowledgments

We thank Xinliang Xu and Zhengji Li from SWIP for providing us with data and reference information on the HL-2A experiments, as well as for their valuable discussions. This work is partially supported by the National Magnetic Confinement Fusion Energy R & D Program of China under Grant No. 2019YFE03060000, as well as by the National Natural Science Foundation of China (NSFC) under Grant No. 12375225. This work is also Supported by the National Magnetic Confinement Fusion Energy R & D Program of China with Grant No. 2024YFE03020001, the

National Natural Science Foundation of China with Grant No. 12275071.

ORCID iDs

H.J. Zhao  0000-0002-9826-4754

S.M. Li  0009-0009-6834-1083

Y. Xiao  0000-0001-5333-2867

References

- [1] Qilong R. *et al* 2009 Momentum Transport in DIII-D Discharges with and Without Magnetohydrodynamics (MHD) Activity *Plasma Sci. Technol.* **11** 127–30
- [2] Li H., Wang L., Fu Y., Wang Z., Wang T. and Li J. 2025 Surrogate model of turbulent transport in fusion plasmas using machine learning *Nucl. Fusion* **65** 016015
- [3] Qi L., Kwon J.-M., Hahm T., Yi S. and Choi M. 2019 Characteristics of trapped electron transport, zonal flow staircase, turbulence fluctuation spectra in elongated tokamak plasmas *Nucl. Fusion* **59** 026013
- [4] Hahm T.S., Beer M.A., Lin Z., Hammett G.W., Lee W.W. and Tang W.M. 1999 Shearing rate of time-dependent $E \times B$ flow *Phys. Plasmas* **6** 922–6
- [5] Garbet X., Idomura Y., Villard L. and Watanabe T. 2010 Gyrokinetic simulations of turbulent transport *Nucl. Fusion* **50** 043002
- [6] Kishimoto Y., Tajima T., Horton W., LeBrun M.J. and Kim J.Y. 1996 Theory of self-organized critical transport in tokamak plasmas *Phys. Plasmas* **3** 1289–307
- [7] Choi G., Diamond P. and Hahm T. 2024 On how fast ions enhance the regulation of drift wave turbulence by zonal flows *Nucl. Fusion* **64** 016029
- [8] Lu Z.X., Wang W.X., Diamond P.H., Tynan G., Ethier S., Gao C. and Rice J. 2015 Intrinsic torque reversals induced by magnetic shear effects on the turbulence spectrum in tokamak plasma *Phys. Plasmas* **22** 055705
- [9] Li J. and Kishimoto Y. 2002 Zonal flows in gyrofluid simulations of slab electron temperature gradient turbulence *Phys. Plasmas* **9** 1241–54
- [10] Jenko F., Dorland W., Kotschenreuther M. and Rogers B.N. 2000 Electron temperature gradient driven turbulence *Phys. Plasmas* **7** 1904–10
- [11] Bourdelle C., Garbet X., Imbeaux F., Casati A., Dubuit N., Guirlet R. and Parisot T. 2007 A new gyrokinetic quasilinear transport model applied to particle transport in tokamak plasmas *Phys. Plasmas* **14** 112501
- [12] Citrin J. *et al* JET Contributors 2017 Tractable flux-driven temperature, density and rotation profile evolution with the quasilinear gyrokinetic transport model QuaLiKiz *Plasma Phys. Control. Fusion* **59** 124005
- [13] Rafiq T., Kritz A.H., Weiland J., Pankin A.Y. and Luo L. 2013 Physics basis of multi-mode anomalous transport module *Phys. Plasmas* **20** 032506
- [14] Li H., Fu Y., Li J. and Wang Z. 2021 Machine learning of turbulent transport in fusion plasmas with neural network *Plasma Sci. Technol.* **23** 115102
- [15] Staebler G.M., Kinsey J.E. and Waltz R.E. 2007 A theory-based transport model with comprehensive physics *Phys. Plasmas* **14** 055909
- [16] Beer M.A. and Hammett G.W. 1996 Toroidal gyrofluid equations for simulations of tokamak turbulence *Phys. Plasmas* **3** 4046–64
- [17] Dorland W.D. 1993 Gyrofluid models of plasma turbulence *PhD Dissertation* (Princeton University)

- [18] Kates-Harbeck J., Svyatkovskiy A. and Tang W. 2019 Predicting disruptive instabilities in controlled fusion plasmas through deep learning *Nature* **568** 526–31
- [19] Dong G., Wei X., Bao J., Brochard G., Lin Z. and Tang W. 2021 Deep learning based surrogate models for first-principles global simulations of fusion plasmas *Nucl. Fusion* **61** 126061
- [20] Liu Z., Qiu H., Fu G., Xiao Y., Chen Y., Wang Z. and Wei Y. 2025 Prediction of fishbone linear instability in tokamaks with machine learning methods *Nucl. Fusion* **65** 016007
- [21] Meneghini O. et al 2017 Self-consistent core-pedestal transport simulations with neural network accelerated models *Nucl. Fusion* **57** 086034
- [22] Zhao H., Guo Z., Wu X. and Xiao Y. 2025 Machine learning for electrostatic plasma turbulence classification in tokamaks *Plasma Sci. Technol.* **27** 075105
- [23] Chung H., Lee C., Choi G., Kaye S., LeBlanc B., Berkery J. and Na Y.-S. 2025 Development of a data-driven neural network model for electron thermal transport in NSTX *Nucl. Fusion* **65** 086028
- [24] Citrin J., Breton S., Felici F., Imbeaux F., Aniel T., Artaud J., Baiocchi B., Bourdelle C., Camenen Y. and Garcia J. 2015 Real-time capable first principle based modelling of tokamak turbulent transport *Nucl. Fusion* **55** 092001
- [25] van de Plassche K.L., Citrin J., Bourdelle C., Camenen Y., Casson F.J., Dagnelie V.I., Felici F., Ho A. and Van Mulders S. (JET Contributors) 2020 Fast modeling of turbulent transport in fusion plasmas using neural networks *Phys. Plasmas* **27** 022310
- [26] Ma C., Zhu B., Xu X.-Q. and Wang W. 2020 Machine learning surrogate models for landau fluid closure *Phys. Plasmas* **27** 042502
- [27] Lin Z., Hahn T.S., Lee W.W., Tang W.M. and White R.B. 1998 Turbulent transport reduction by zonal flows: Massively parallel simulations *Science* **281** 1835–7
- [28] Duan X. et al HL-2A team 2017 Overview of recent HL-2A experiments *Nucl. Fusion* **57** 102013
- [29] Lin Z. and Hahn T.S. 2004 Turbulence spreading and transport scaling in global gyrokinetic particle simulations *Phys. Plasmas* **11** 1099–108
- [30] Xiao Y. and Lin Z. 2009 Turbulent transport of trapped-electron modes in collisionless plasmas *Phys. Rev. Lett.* **103** 085004
- [31] Zhao C., Zhang T. and Xiao Y. 2017 Gyrokinetic simulation of dissipative trapped electron mode in tokamak edge *Phys. Plasmas* **24** 052509
- [32] Holod I. and Lin Z. 2013 Verification of electromagnetic fluid-kinetic hybrid electron model in global gyrokinetic particle simulation *Phys. Plasmas* **20** 032309
- [33] Zhang W., Holod I., Lin Z. and Xiao Y. 2012 Global gyrokinetic particle simulation of toroidal Alfvén eigenmodes excited by antenna and fast ions *Phys. Plasmas* **19** 022507
- [34] McClenaghan J., Lin Z., Holod I., Deng W. and Wang Z. 2014 Verification of gyrokinetic particle simulation of current-driven instability in fusion plasmas. I. Internal kink mode *Phys. Plasmas* **21** 122519
- [35] Xiao Y., Holod I., Zhang W., Klasky S. and Lin Z. 2010 Fluctuation characteristics and transport properties of collisionless trapped electron mode turbulence *Phys. Plasmas* **17** 022302
- [36] Li S., Yang D. and Xiao Y. 2025 Gyrokinetic simulation of isotope effect on zonal flow response in tokamaks *Phys. Plasmas* **32** 023902
- [37] Wei X. et al 2025 Gyrokinetic simulations of the effects of magnetic islands on microturbulence in KSTAR *Nucl. Fusion* **65** 026026
- [38] Yang H., Zhou T. and Xiao Y. 2021 Gyrokinetic simulation of turbulent transport for I-mode edge plasmas *Nucl. Fusion* **61** 056006
- [39] LeCun Y., Bengio Y. and Hinton G. 2015 Deep learning *Nature* **521** 436–44
- [40] Vapnik V.N. 1995 *The Nature of Statistical Learning Theory* (Springer)
- [41] Yudong P., Enyao W. and Yi L. 2003 HL-2A tokamak edge modeling with B2 *Plasma Sci. Technol.* **5** 2023–6
- [42] Shen Y., Dong J., Shi Z., He H., Zhao K., Peng X., Qu H., Li J. and Sun A. 2025 Study of circular cross-section plasmas in HL-2A tokamak: MHD equilibrium, stability and operational β limit *Acta Phys. Sin.* **74** 135204
- [43] Landreman M., Choi J.Y., Alves C., Balaprakash P., Churchill M., Conlin R. and Roberg-Clark G. 2025 How does ion temperature gradient turbulence depend on magnetic geometry? Insights from data and machine learning *J. Plasma Phys.* **91** E120
- [44] Rewoldt G., Lin Z. and Idomura Y. 2007 Linear comparison of gyrokinetic codes with trapped electrons *Comput. Phys. Commun.* **177** 775–80
- [45] Wan C., Cho Y., Qu Z., Camenen Y., Varennes R., Lim K., Li K., Li J., Li Y. and Garbet X. 2025 A high-fidelity surrogate model for the ion temperature gradient (ITG) instability using a small expensive simulation dataset *Nucl. Fusion* **65** 054001
- [46] Gholamy A., Kreinovich V. and Kosheleva O. 2018 Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation *Computer Science* p 7
- [47] Hinton F.L. and Hazeltine R.D. 1976 Theory of plasma transport in toroidal confinement systems *Rev. Mod. Phys.* **48** 239–308
- [48] Pankin A., McCune D., Andre R., Bateman G. and Kritiz A. 2004 The tokamak Monte Carlo fast ion module NUBEAM in the national transport code collaboration library *Comput. Phys. Commun.* **159** 157–84
- [49] Felici F., Citrin J., Teplukhina A., Redondo J., Bourdelle C., Imbeaux F. and Sauter O. 2018 (JET Contributors and the EUROfusion MST1 Team) Real-time-capable prediction of temperature and density profiles in a tokamak using RAPTOR and a first-principle-based transport model *Nucl. Fusion* **58** 096006
- [50] Felici F., De Baar M. and Steinbuch M. 2014 A dynamic state observer for real-time reconstruction of the tokamak plasma profile state and disturbances 2014 *American Control Conf.* (IEEE) pp 4816–23