

Fast Neural-Network Surrogate Model of the Updated Multi-Mode Anomalous Transport Module for NSTX-U

Brian Leard¹, Member, IEEE, Zibo Wang, Member, IEEE, Shira Morosohk², Tariq Rafiq, and Eugenio Schuster³, Member, IEEE

Abstract—The multimode module (MMM) is a theory-based anomalous transport model that is used in integrated codes to predict the electron/ion temperature, electron/impurity density, and toroidal/poloidal rotation profiles of tokamak plasmas. It includes transport driven by a variety of electron and ion scale modes and accounts for the effects of collisions, fast-ion and impurity dilution, noncircular flux surfaces, finite beta, and Shafranov shift. Due to the large number of interconnected physical phenomena captured by MMM, it is a computationally intensive code, unsuitable for control applications that require from fast offline to real time or faster-than-real-time prediction speeds. Therefore, significant effort has been dedicated to the development of neural-network (NN)-based surrogate models for tokamaks like DIII-D and EAST with the goal of reproducing a selected set of MMM's output predictions at a much faster speed. In this work, a NN model for MMM9.1 is trained based on predictions within the National Spherical Torus Experiment Upgrade (NSTX-U) operating regime. This model, referred to as MMMNet, is capable of predicting the ion, electron, and impurity thermal diffusivities, electron particle diffusivity, and toroidal and poloidal momentum diffusivities. The newer MMM9.1 version improves the accuracy, consistency, speed, and physics basis of several components of MMM that are critical for NSTX-U as well as for the spherical-tokamak reactor concept. The MMM training data are based on NSTX experimental shots and adapted to suit the NSTX-U geometry and to cover the entirety of the NSTX-U parameter space. In simulation testing, MMMNet is shown to have a similar accuracy to MMM while running at a fraction of its computation time. This makes MMMNet suitable for real-time control applications.

Index Terms—Control-oriented modeling, data-filtering, machine learning, multimode module (MMM), neural network (NN).

I. INTRODUCTION

FOR nuclear fusion to be considered a viable commercial energy source, it is necessary to reliably achieve advanced tokamak (AT) scenarios. These are characterized by a high fusion gain, magneto-hydrodynamic (MHD) stability,

and improved confinement of the plasma, potentially even allowing steady-state operation [1]. For these AT scenarios to occur, it is necessary to control the spatial distribution of the temperature, momentum, and density for the various plasma species (ions, electrons, and impurities). For successful control of these properties, prediction models are needed that range from fast (for offline control design and scenario optimization), to very fast (for online control, estimation, and optimization). To achieve these required calculation times, control-oriented prediction codes must often use reduced-order or empirical models. The use of these simplified models often sacrifices some prediction accuracy, which can lead to reduced control effectiveness. Data-based models incorporating machine-learning techniques might be able to contribute toward resolving this issue. Machine-learning models can capture complex physical dynamics in the plasma while maintaining fast computational times. Indeed, there have already been numerous examples of these types of models being successfully employed for nuclear-fusion applications. Surrogate models based on neural networks have been trained to reproduce the results of NUBEAM [2], [3], [4], GENRAY [5], EPED [6], and multimode module (MMM) [7]. Additionally, machine learning models have been used for plasma equilibrium reconstruction/computation [8], [9], [10], and disruption prediction [11], [12]. There have even been successful efforts at designing entire control schemes using data-based approaches through reinforcement learning [13], [14]. In this work, a machine learning technique is used to effectively reproduce the results of MMM with a computational time suited for control applications. This work builds on previous efforts [7] by enlarging the set of diffusivity profiles predicted by the surrogate model, using an upgraded version of MMM, simplifying the approach to handling spatial profiles, and gearing the model toward a different tokamak. Additionally, this work fundamentally restructures the data-based model; whereas the previously mentioned work used an average output of multiple NN's to predict all diffusivities simultaneously, this work has trained a different NN to predict each diffusivity separately. Finally, unlike past efforts, the data this model has been trained on have been processed using established data filtering techniques for data-based models, removing nonrealistic data and ensuring that the network only reproduces physical phenomena.

The MMM is a theory-based 1-D turbulent transport model that predicts the thermal, particle, and momentum anomalous

Manuscript received 4 October 2023; revised 2 February 2024; accepted 27 February 2024. Date of publication 2 April 2024; date of current version 9 December 2024. The material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences under Award Number DE-SC-0021385. The review of this article was arranged by Senior Editor R. Chapman. (Corresponding author: Brian Leard.)

The authors are with the Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA 18015 USA (e-mail: brian.leard@lehigh.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TPS.2024.3377611>.

Digital Object Identifier 10.1109/TPS.2024.3377611

0093-3813 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

diffusivities. It contains models for the ion and electron temperature gradient modes, trapped electron modes, peeling, ballooning, microtearing, and high-mode-number MHD modes [15]. MMM has been validated against experimental results on various tokamak geometries [16]. Version 9.1 of MMM is superior to its predecessors due to the inclusion of the microtearing mode [17] and a novel model for electromagnetic electron temperature gradient mode [18]. The incorporation of these models improves MMM's ability to predict electron temperature profiles in conventional and, particularly, low-aspect-ratio high-beta tokamaks such as National Spherical Torus Experiment Upgrade (NSTX-U). Due to the heavily nonlinear nature of the diffusivity evolution, the large scope of the physical phenomena involved, and the numerous methods of turbulent transport, MMM has a significant computational time. This lengthy runtime limits MMM's applicability in plasma control. Instead, control-oriented simulators have often turned to semi-empirical models. For example, models, such as Bohm/Gyro-Bohm, or Coppi-Tang are frequently used to predict anomalous thermal diffusivities. However, the empirical parameters of these models were calibrated and validated on the geometry of conventional tokamaks [19], [20]. A machine-learning model that can recreate the results of MMM for NSTX-U geometry could better capture the physics of spherical tokamaks operating in a high- β regime with a significantly smaller aspect ratio. This approach could offer an alternative path for anomalous diffusivity calculation and potentially improve the control of thermal, particle, and momentum profiles.

This article is organized as follows. Section II discusses the methodology of generating the dataset. Section III overviews the techniques used to filter the dataset for outlier data as well as to smooth output profiles. Section IV describes the neural network architecture. Section V compares the outputs of the neural-network (NN) surrogate model MMMNet to MMM. Section VI draws conclusions and discusses potential future work.

II. CREATION OF DATASET

While it is potentially feasible to train a neural network for MMM that is valid for different tokamak geometries, this would involve creating a dataset that encaptures the entire parameter space of MMM, and would be both time consuming and potentially lead to an overcomplicated NN model. It is more practical to create a dataset geared for a specific machine, a method whose effectiveness has already been partially demonstrated [7]. Since the goal of this model is to recreate MMM results for NSTX-U, the parameter space for the model's dataset was created using experimental data from NSTX-U's 2016 run campaign. To generate the dataset, the simulation code TRANSP [21] was run using the MMM9.1 module to calculate anomalous diffusivities. These TRANSP runs were based on 214 experimental NSTX-U shots. To further enhance the dataset and capture a larger parameters space, multiple TRANSP runs were based on each shot. These runs were diversified by randomly varying certain scalars. The effective charge (Z_{eff}) was assigned a random value between 1 and 4, and the pedestal location was assigned

TABLE I
LIST OF INPUTS AND OUTPUTS TO MMMNet

	Symbol	Description	Units	Grad.
Inputs	Z_{imp}	Average charge of impurities		
	A_{imp}	Average mass of impurities	amu	
	Z_{eff}	mean effective charge		
	B_t	Toroidal magnetic field	T	
	κ	Elongation		
	R	Major radius	m	
	a	Minor radius	m	
	n_e	Electron density	m^{-3}	x
	n_i	Ion density	m^{-3}	x
	n_{imp}	Impurity density	m^{-3}	x
	n_h	Hydrogenic ion density	m^{-3}	
	T_e	Electron temperature	keV	x
	T_i	Ion temperature	keV	x
	q	Safety factor profile		x
	$\omega_{E \times B}$	$E \times B$ shear	rad s^{-1}	
v_ϕ	Toroidal velocity	m s^{-1}	x	
v_θ	Poloidal velocity	m s^{-1}	x	
Outputs	χ_e	Electron thermal diffusivity	$\text{m}^2 \text{s}^{-1}$	
	χ_i	Ion thermal diffusivity	$\text{m}^2 \text{s}^{-1}$	
	χ_z	Impurity thermal diffusivity	$\text{m}^2 \text{s}^{-1}$	
	χ_ϕ	Toroidal momentum diffusivity	$\text{m}^2 \text{s}^{-1}$	
	χ_θ	Poloidal momentum diffusivity	$\text{m}^2 \text{s}^{-1}$	
	χ_d	Electron particle diffusivity	$\text{m}^2 \text{s}^{-1}$	

a value governed by a normal distribution centered at 0.8 with a standard deviation of 0.03. Since data-based models can often be poor at extrapolation, this variation expanded the parameter space in which the model would be viable.

These TRANSP shots used the diffusivities calculated by MMM to predict certain profiles which were chosen from a selection of the electron temperature, the ion temperature, the angular velocity, the electron density, the ion density, and the impurity density. However, predicting all these profiles simultaneously can cause stability issues within the TRANSP simulation. Therefore, only a subset of the aforementioned profiles was predicted during a TRANSP run. The remaining profiles were sourced from the corresponding experimental shot. In other words, while MMM predicted all six diffusivities for each TRANSP shot, only a subset of these diffusivities was used for profile prediction. Varying the combination of profiles predicted using the MMM diffusivities was another method employed to increase the variability of the dataset.

A total of 589 TRANSP runs were conducted, producing over 500 000 time slices that could be used for the training of the surrogate model. The inputs and outputs of MMMNet, shown in Table I, were chosen to be identical to the inputs and outputs of MMM. Since MMM is a gradient-based code, in some cases the spatial gradient of the profile was also included as an input, indicated by the last column of the table. While previous work has often reduced the dimensionality of these profiles using principle component analysis (PCA), this work represents profiles as a discrete series of nodes, which is how profiles are represented in TRANSP.

III. DATA PREPROCESSING

All data-based models are only as accurate as the underlying data that the model has been trained on. Therefore, before the

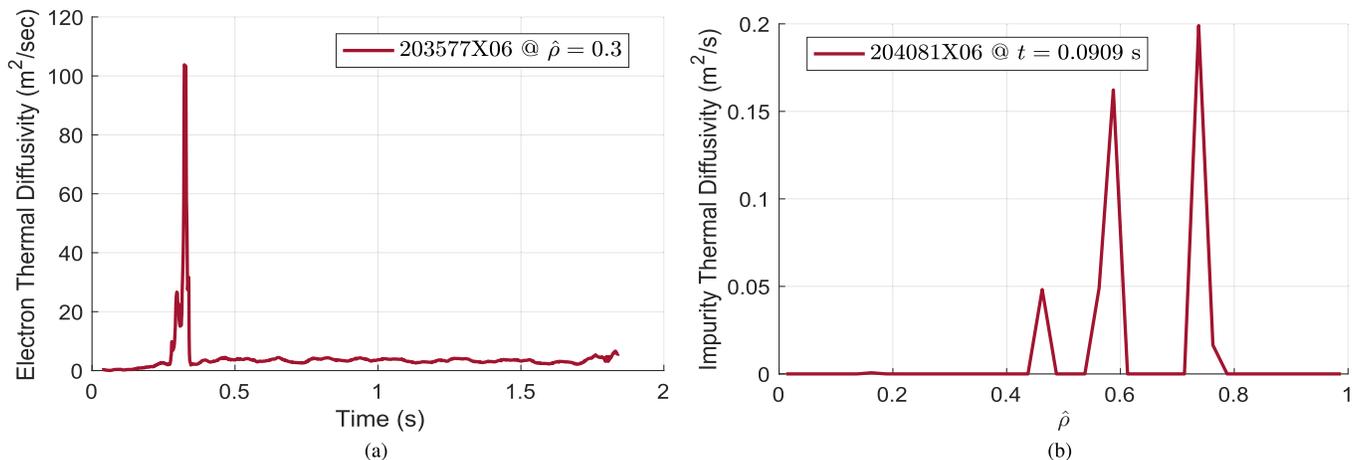


Fig. 1. (a) Temporal spike in the electron thermal diffusivity in TRANSP shot 203577X06. (b) Impurity diffusivity profile in shot 204081X06 composed of spatial spikes, categorized as a “sparse” profile. Note that the variable $\hat{\rho}$ is the normalized mean effective minor radius, which is the spatial coordinate.

NNsurrogate model was trained, the raw data was analyzed to remove any non-realistic or incomplete data. Preprocessing the data requires a methodical approach and some basic knowledge of the underlying system. Without removing poor data, the network could be learning “incorrect” behavior due to the fact that it is trying to recreate unrealistic outputs. However if a subset of “good” or physically viable data is removed, the network’s parameter space becomes limited, and it will not be able to reproduce certain physical phenomena. Therefore, it is pivotal to carefully analyze the data and remove any data that is clearly “non-realistic” without impacting data that is otherwise viable. This was done using a series of filters that removed datasets that exhibited a particular kind of anomalous behavior. Since we are dealing with profiles evolving through time, it is important to examine not just each particular value in the dataset, but also how that value compares to previous and future times in the simulation.

A. Filtering Data for Outliers

There were two types of data anomalies that were considered outliers and removed from the dataset. The first anomaly was the data “spike,” where a certain value or collection of values were exceptionally large relative to their neighboring values. These spikes could occur either spatially or temporally. The second anomaly was when input or output parameters of a time slice were missing values, or had seemingly incomplete profiles.

An example of a temporal spike is shown in Fig. 1(a), where the value of the electron thermal diffusivity is relatively stable, rapidly increases several orders of magnitude, and then quickly decreases back to the initial value. These spikes were identified by taking the temporal gradients of each TRANSP shot. Any time slice where the absolute values of the gradients were higher than a predetermined threshold were removed. As a further precaution, time slices immediately neighboring a temporal spike were also removed.

A slightly different approach was adopted to handle spatial spikes. This is because MMM is a gradient based code, and therefore could produce somewhat “noisy” data with high

localized spatial gradients. Additionally, it was common to see physically realistic profiles that had substantial increases in the diffusivity over a small spatial region. Therefore, large spatial gradients were common, and could not be solely relied upon to filter spikes in the outputs’ profiles. First, a filter was applied to the outputs to detect and eliminate “narrow spikes” where one value was significantly larger than all other values in the profile. Next, the output profiles were smoothed using the procedure discussed in Section III-B. After this smoothing process, a gradient-based filter was applied that removed time-slices higher than a predetermined threshold. However, this threshold was kept purposefully high to prevent disregarding physically viable data.

Finally, time slice that were classified as having missing data were removed. Data was deemed missing when the value of an input or output parameter equaled zero despite this being physically unrealistic for said parameter. Sometimes, the output profiles were not entirely missing, but were rather incomplete or “sparse.” In these cases the profile was predominantly zeros, with only isolated nonzero values, such as the profile shown in Fig. 1(b). However, these profiles were deemed sufficiently unreliable to justify their removal from the dataset.

B. Smoothing Output Profiles

As previously stated, the nature of the MMM code (being gradient based) can lead to noisy data that has sharp peaks or valleys in the output profiles. When these peaks grew large enough to dominate the profile, the time slice was removed. However smaller spikes in the profile were expected and did not necessarily mean the profile was physically unrealistic. Therefore, it was not desirable to remove all data that had spatial spikes. Instead, less significant spikes were mitigated by smoothing out the output profiles using a gaussian weighted moving average filter, overviewed in [22]. This filter has a dual purpose of removing smaller spikes and reducing the noisy nature of the data, capturing only dominant features of the profile. This type of filter uses observations close to the target point x_0 to fit a simple model over the data that is smooth across the spatial domain. This can be represented

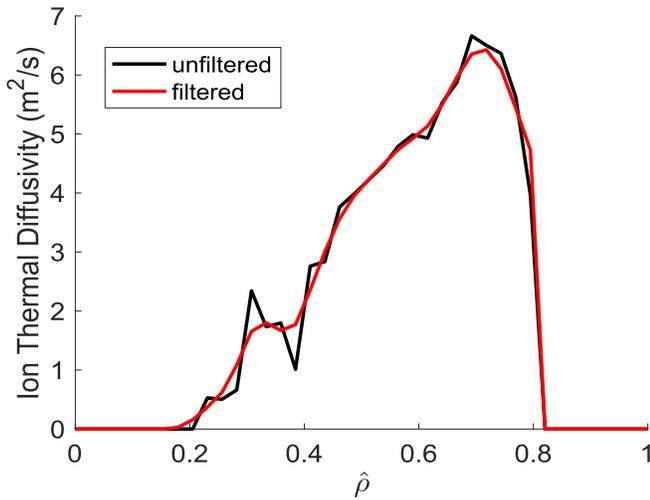


Fig. 2. Unfiltered MMM profile compared to an MMM profile processed through a gaussian moving-average filter.

mathematically as

$$y_0 = \frac{\sum_{i=1}^N K_\lambda(x_0, x_i) \bar{y}_i}{\sum_{i=1}^N K_\lambda(x_0, x_i)} \quad (1)$$

where x_i , $i = 1, \dots, N$, is a particular spatial node, with N being the total number of nodes in the profile. The variable y_0 dictates the profile value at a particular spatial node x_0 after smoothing, and \bar{y}_i is the original profile value at the spatial node i . A kernel function K_λ , which in this case is a Gaussian function, assigns a certain weight to a node x_i based on the distance between the node and the target point x_0

$$K_\lambda(x_0, x_i) = \frac{1}{\lambda} \exp\left[-\frac{\|x_i - x_0\|^2}{2\lambda}\right]. \quad (2)$$

The parameter λ is a constant that determines the width of the region that has significant contribution to the target point. Using this method, the outputs from MMM were smoothed. A comparison between an original MMM output profile and a smoothed profile is shown in Fig. 2.

C. Normalizing the Data

After smoothing the data, the inputs and outputs were normalized to make the mean of each 0 and the standard deviation 1. This was done by applying the transformation

$$\tilde{z} = \frac{z - \mu}{s} \quad (3)$$

where z is the original value, μ is the mean of that value for all time slices in the dataset, s is the standard deviation, and \tilde{z} is the normalized value. This process prevents inputs of large orders of magnitude, such as densities, from having a dominant influence on the cost function and marginalizing other inputs. Normalization guarantees that the relationship between inputs and outputs will be determined based on the physical dependence of one on the other, and not due to their respective scales. It should be noted that this scaling issue can occur within a specific input–output profile as well if there is a large difference in scales within the profile.

TABLE II
NEURAL NETWORK ARCHITECTURE

Final Neural Network Hyperparameters	
Hidden layers	2
Nodes per hidden layer	120
Epochs	50
Batch size	500
Activation func. (hidden layers)	ReLU
Activation func. (output layer)	Linear
Learning Rate	0.001
Solver	Adam
Loss function	Mean squared error

TABLE III
CORRELATION RESULTS BETWEEN MMM AND MMMNet

Output	R^2 : Training	R^2 : Testing	Comp. Time (ms)
χ_i	0.976	0.911	0.287
χ_e	0.987	0.939	0.283
χ_z	0.888	0.802	0.255
χ_ϕ	0.938	0.857	0.246
χ_θ	0.969	0.894	0.253
χ_d	0.966	0.891	0.230

Therefore, the nodes in each profile were individually normalized. Fig. 3 shows the normalized distribution for each input.

Once the data was filtered and normalized, each TRANSP shot in the dataset was classified into one of three groups, training, validation, or testing. The training set (80% of dataset) trains the weights and biases of the neural network, the validation set (10% of dataset) helps determine the hyperparameters of the network, and the testing set (10% of dataset) is used for model evaluation.

IV. NEURAL NETWORK METHODOLOGY

The objective of this neural network-based model was to accurately predict the results of MMM with a sufficiently fast execution time. This model, referred to as MMMNet, is a collection of six individual neural networks that predict each anomalous diffusivity calculated by MMM. These networks were coded in Python using the Tensorflow application programming interface (API) [23]. The multilayer perceptron was chosen as the type of neural network due to its simplistic structure which allows for faster computation times. A grid search weighing both accuracy and computation time was then conducted to determine the optimal neural network hyperparameters. Table II shows the results of this grid search. Definitions of the hyperparameters referenced in this table can be found in [24]. The topology of each neural network was kept the same for all six networks. The identical nature of the networks allows for easier integration into other predictive codes, and tests done to vary hyperparameters of each network individually did not show any significant improvement in the correlation between MMM and MMMNet for the validation data or computational time for the predicted outputs. A schematic of the general workflow for creating the NN-based model is shown in Fig. 4.

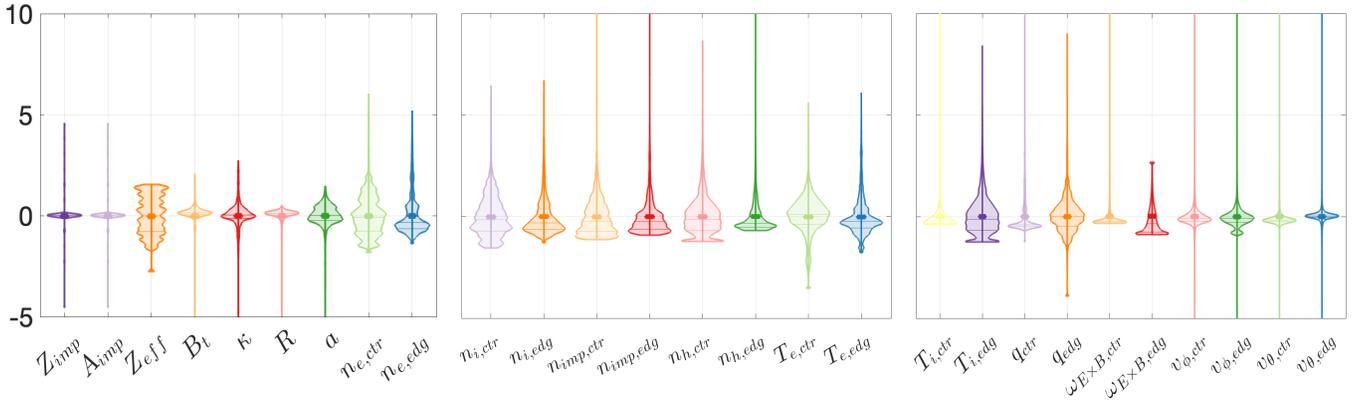


Fig. 3. Butterfly plots showing the distribution of the normalized inputs. The label “ctr” denotes the value of the profile at the center of the plasma and the label “edg” denotes the value of the profile at the plasma edge or boundary.

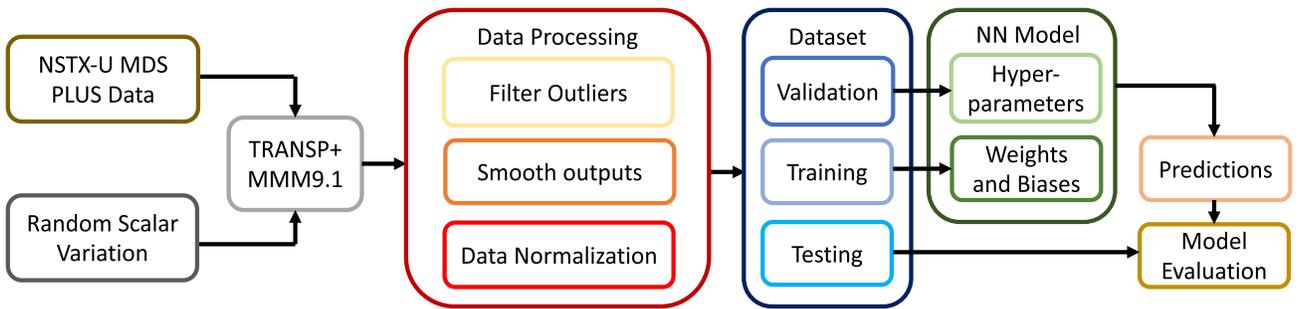


Fig. 4. Workflow diagram for building each neural network model. Experimental data was used to run TRANSP to prepare a database, which was processed and then split into training, validation, and testing groups. The two former groups designed the neural network and the latter tested the NN’s performance.

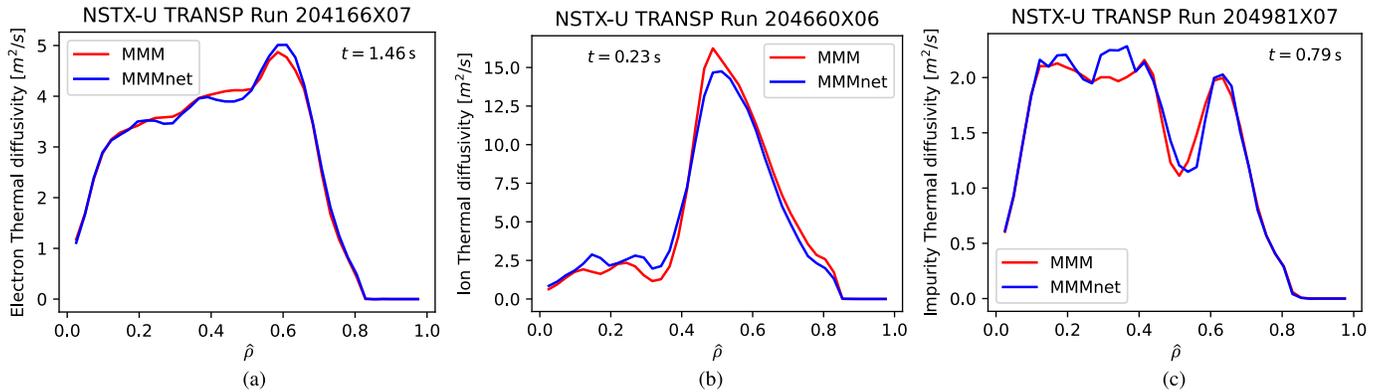


Fig. 5. Comparison of thermal diffusivity profiles calculated by MMM and MMMNet for time slices of shots in the testing dataset. (a) Electron thermal diffusivity profile. (b) Ion thermal diffusivity profile. (c) Impurity thermal diffusivity profile.

V. MODEL EVALUATION

There is a strong correlation between MMM and MMMNet, indicating that the neural network can successfully recreate the results of MMM. This is shown by the high correlation value (R^2) in Table III. Note that there is not a significant difference between the R^2 values of the training and testing data. Since the testing data was not available to the network during training, this suggests that the neural network is learning the underlying behavior of MMM and not simply recreating the training data. Table III also shows that there is a slightly lower validation R^2 score for the impurity diffusivity, χ_z . Tests done to increase this value by varying the hyperparameters did not

have a significant impact. This lower correlation could be due to the underlying model being more difficult to predict and requiring a more complex structure of network to capture the dynamics involved. The last column of Table III shows the average computation times of MMMNet for one prediction. The average computation time per prediction for MMM9.1 is 73 ms, making the total MMMNet computation two orders of magnitude faster than MMM. Fig. 5 shows profile comparisons between MMM and MMMNet for various diffusivities at specific times. Note that these TRANSP shots belonged to the testing dataset and therefore the network did not observe them in training. Fig. 6 shows correlation comparisons for

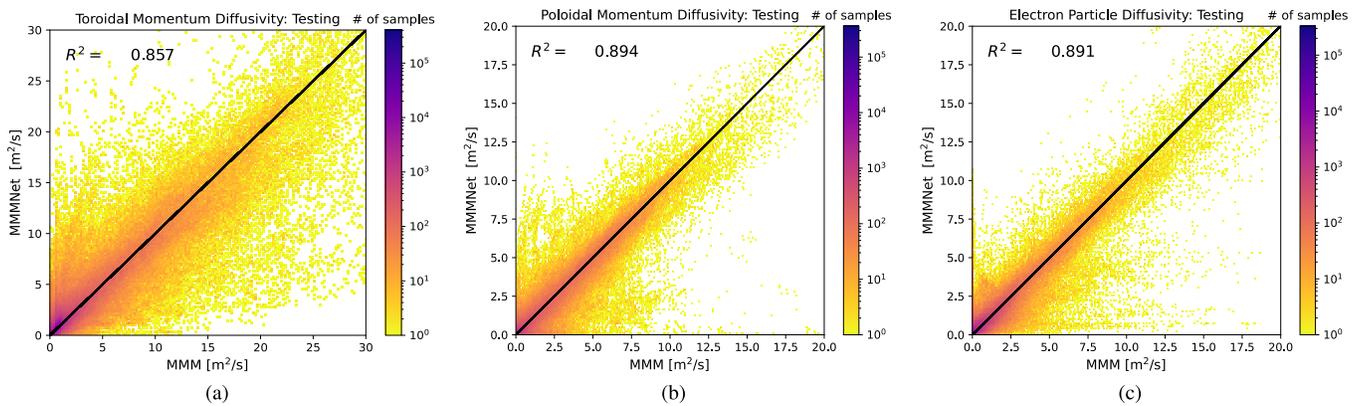


Fig. 6. MMM versus MMMNet correlation for testing data. (a) Toroidal momentum diffusivity. (b) Poloidal momentum diffusivity. (c) Electron particle diffusivity.

the testing data. As expected, diffusivities that have a lower correlation exhibit a larger width on the correlation plot, indicating that there was a greater variance of MMMNet values for a given MMM value.

VI. CONCLUSION AND FUTURE WORK

A neural network surrogate model was trained to recreate the results of MMM with a significantly reduced computation time. Therefore, control-oriented plasma simulation codes can exploit the accuracy of MMM while still running at the speeds required for plasma-control applications. This neural network model, MMMNet, consists of six multilayer perceptrons that predicts one of the anomalous diffusivities output from MMM. Each network had two hidden layers, 120 nodes per layer, and was trained specifically for NSTX-U parameters. The dataset used for the neural network was preprocessed to remove potentially unrealistic results as well as to reduce some of the gradient-sourced noise that can occur in the output profiles. MMMNet produces a strong correlation with MMM while maintaining sufficiently short calculation times. Control-oriented simulation codes, such as the control-oriented transport simulator (COTSIM), could benefit from MMMNet to improve simulation capabilities for density, temperature, and momentum profile evolution without significantly increasing the computation time.

REFERENCES

- [1] T. S. Taylor, "Physics of advanced tokamaks," *Plasma Phys. Controlled Fusion*, vol. 39, no. 12B, pp. B47–B73, Dec. 1997.
- [2] M. D. Boyer, S. Kaye, and K. Erickson, "Real-time capable modeling of neutral beam injection on NSTX-U using neural networks," *Nucl. Fusion*, vol. 59, no. 5, Mar. 2019, Art. no. 056008.
- [3] S. M. Morosohk, M. D. Boyer, and E. Schuster, "Accelerated version of NUBEAM capabilities in DIII-D using neural networks," *Fusion Eng. Des.*, vol. 163, Feb. 2021, Art. no. 112125.
- [4] Z. Wang, S. Morosohk, T. Rafiq, E. Schuster, M. D. Boyer, and W. Choi, "Neural network model of neutral beam injection in the EAST tokamak to enable fast transport simulations," *Fusion Eng. Des.*, vol. 191, Jun. 2023, Art. no. 113514.
- [5] G. M. Wallace et al., "Towards fast and accurate predictions of radio frequency power deposition and current profile via data-driven modelling: Applications to lower hybrid drive," *J. Plasma Phys.*, vol. 88, no. 4, Aug. 2022, Art. no. 895880401.
- [6] O. Meneghini et al., "Self-consistent core-pedestal transport simulations with neural network accelerated models," *Nucl. Fusion*, vol. 57, no. 8, Jul. 2017, Art. no. 086034.
- [7] S. M. Morosohk, A. Pajares, T. Rafiq, and E. Schuster, "Neural network model of the multi-mode anomalous transport module for accelerated transport simulations," *Nucl. Fusion*, vol. 61, no. 10, Oct. 2021, Art. no. 106040.
- [8] L. L. Lao et al., "Application of machine learning and artificial intelligence to extend EFIT equilibrium reconstruction," *Plasma Phys. Controlled Fusion*, vol. 64, no. 7, Jun. 2022, Art. no. 074001.
- [9] S. Joung et al., "Deep neural network Grad-Shafranov solver constrained with measured magnetic signals," *Nucl. Fusion*, vol. 60, no. 1, Dec. 2019, Art. no. 016034.
- [10] Z. Wang, X. Song, E. Schuster, and T. Rafiq, "Neural network-based free boundary equilibrium solver for control-oriented transport simulations," *IEEE Trans. Plasma Sci.*, 2023.
- [11] B. H. Guo et al., "Disruption prediction on EAST tokamak using a deep learning algorithm," *Plasma Phys. Controlled Fusion*, vol. 63, no. 11, Sep. 2021, Art. no. 115007.
- [12] K. J. Montes et al., "Machine learning for disruption warnings on alcator C-mod, DIII-D, and EAST," *Nucl. Fusion*, vol. 59, no. 9, Jul. 2019, Art. no. 096015.
- [13] J. Degraeve et al., "Magnetic control of tokamak plasmas through deep reinforcement learning," *Nature*, vol. 602, no. 7897, pp. 414–419, 2022.
- [14] J. Seo et al., "Development of an operation trajectory design algorithm for control of multiple 0D parameters using deep reinforcement learning in KSTAR," *Nucl. Fusion*, vol. 62, no. 8, Jul. 2022, Art. no. 086049.
- [15] T. Rafiq, A. H. Kritiz, J. Weiland, A. Y. Pankin, and L. Luo, "Physics basis of multi-mode anomalous transport module," *Phys. Plasmas*, vol. 20, no. 3, Mar. 2013, Art. no. 032506.
- [16] T. Rafiq et al., "Validating the multi-mode Model's ability to reproduce diverse tokamak scenarios," *Plasma*, vol. 6, no. 3, pp. 435–458, Jul. 2023.
- [17] T. Rafiq et al., "Microtearing instabilities and electron thermal transport in low and high collisionality NSTX discharges," *Phys. Plasmas*, vol. 28, no. 2, Feb. 2021, Art. no. 022504.
- [18] T. Rafiq et al., "Electron temperature gradient driven transport model for tokamak plasmas," *Phys. Plasmas*, vol. 29, no. 9, Sep. 2022, Art. no. 092503.
- [19] M. Erba, T. Aniel, V. Basiuk, A. Becoulet, and X. Litaudon, "Validation of a new mixed Bohm/gyro-bohm model for electron and ion heat transport against the ITER, tore supra and START database discharges," *Nucl. Fusion*, vol. 38, no. 7, pp. 1013–1028, Jul. 1998.
- [20] G. R. McKee et al., "Non-dimensional scaling of turbulence characteristics and turbulent diffusivity," *Nucl. Fusion*, vol. 41, no. 9, pp. 1235–1242, Sep. 2001.
- [21] J. Breslau et al., "TRANSP. Computer software," USDOE Office Sci. (SC), Fusion Energy Sci. (FES) (SC-24), Jun. 2018, doi: 10.11578/dc.20180627.4.
- [22] T. Hastie, R. Tibshirani, and J. H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, vol. 2. Cham, Switzerland: Springer, 2009.
- [23] B. Pang, E. Nijkamp, and Y. N. Wu, "Deep learning with TensorFlow: A review," *J. Educ. Behav. Statist.*, vol. 45, no. 2, pp. 227–248, 2020.
- [24] J. A. Stegemann and N. R. Buenfeld, "A glossary of basic neural network terminology for regression problems," *Neural Comput. Appl.*, vol. 8, no. 4, pp. 290–296, Dec. 1999.



Brian Leard (Member, IEEE) received the bachelor's degree in mechanical engineering from Lehigh University, Bethlehem, PA, USA, in 2020, where he is currently pursuing the Ph.D. degree with the Department of Mechanical Engineering and Mechanics, focusing on advanced control techniques for tokamak plasmas.

He is currently working with the Princeton Plasma Physics Laboratory (PPPL), Princeton, NJ, USA. His research interests include optimal control, machine learning, scenario optimization, and control-oriented modeling.



Zibo Wang (Member, IEEE) received the B.S. degree in processing equipment and control system from East China University of Science and Technology, Shanghai, China, in 2012, and the M.S. degree in mechanical engineering and mechanics from the University of Pittsburgh, Pittsburgh, PA, USA, in 2018. He is currently pursuing the Ph.D. degree in the plasma control laboratory with the Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA, USA.

His research interests include cover optimal control, machine learning, and scenario planning.



Shira Morosohk was born in Denver, CO, USA, in August, 1996. She received the bachelor's and Ph.D. degrees in mechanical engineering from Lehigh University, Bethlehem, PA, USA, in 2018 and 2023, respectively.

Since 2020, she has been with the DIII-D National Fusion Facility, San Diego, CA, USA, where she is currently a Post-Doctoral Researcher with Oak Ridge Associated Universities, Oak Ridge, TN, USA. Her research focus is in applying machine learning methods to design and implement controllers for tokamak plasmas.



Tariq Rafiq received the M.Phil. degree in plasma physics from Quaid-i-Azam University, Islamabad, Pakistan, in 1998, and the Ph.D. degree in electrical engineering from Chalmers University, Gothenburg, Sweden, in 2004. He completed a Post-Doctoral Fellowship at the University of Wisconsin, Madison, WI, USA.

Currently, he serves as a Research Associate Professor with the Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA, USA. He is a Nuclear Fusion Theorist and a Modeler, specializing in the development of thermal, particle, and momentum transport models. His research interests include validating developed models against tokamak experiments, modeling scenarios, improving tokamak discharge performance, designing new experiments, and extrapolating results to future thermonuclear fusion devices like ITER and the fusion pilot plant.



Eugenio Schuster (Member, IEEE) received the B.Sc. degree in electronic engineering from the University of Buenos Aires, Buenos Aires, Argentina, in 1994, the B.Sc. and M.Sc. degrees in nuclear engineering from the Balseiro Institute, Bariloche, Argentina, in 1998, and the Ph.D. degree in aerospace engineering from the University of California at San Diego, San Diego, CA, USA, in 2004.

He is currently a Professor with the Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA, USA. He is an expert in nuclear-fusion plasma control and leads the Lehigh University Plasma Control Laboratory.

Prof. Schuster was a recipient of the National Science Foundation (NSF) CAREER Award for his work on "Nonlinear Control of Plasmas in Nuclear Fusion." He has been appointed as a Scientist Fellow in Plasma Control by the ITER Organization. Moreover, he has been designated by the U.S. Department of Energy (DOE) as an Expert Member of the Integrated Operation Scenarios (IOS) Topical Group within the International Tokamak Physics Activity (ITPA), which he is currently Chairing. He has served as a Leader of the Operations and Control Topical Group with the U.S. Burning Plasma Organization (BPO) and as the Founder Chair of the Technical Committee on Power Generation with the IEEE Control Systems Society.