CONFERENCE PRE-PRINT

RECONSTRUCTING THE PLASMA BOUNDARY WITH A REDUCED SET OF DIAGNOSTICS

M. STOKOLESOV, M.R. NURGALIEV, I. KHARITONOV, E. ADISHCHEV, D. SOROKIN
Next Step Fusion
Bertrange, Luxembourg
Email: ms@nextfusion.org

Abstract

The applicability of surrogate models that used limited number of diagnostics for plasma boundary reconstruction is presented. Using DIII-D EFIT data, several models were trained. These models differ in their input quantities, ranging from the most informed (incorporating coil currents, plasma current, and all magnetic sensors) to the least informed, which relies solely on coil currents. Intermediate models utilize different combinations of magnetic diagnostics as inputs. More than 7000 shots, comprising over 1.8 million timesteps, were used to train five surrogate models. The most informed model achieves an average accuracy of 1–2 cm. As the amount of diagnostic data decreases, reconstruction quality degrades by up to a factor of two, with the largest errors observed in models that use only coil currents. The maximum absolute errors occur near the X-point region. However, all models perform comparably well at the midplane and are able to capture X-point trends. The most challenging cases exhibit errors of up to ~20 cm, though these are relatively rare. Importantly, even the model trained solely on coil currents achieves satisfactory accuracy, highlighting its potential for future development.

1. INTRODUCTION

Present-day tokamak magnetic control relies on real-time equilibrium reconstruction codes [1–3], which approximate solutions to the Grad-Shafranov equation while fitting magnetic diagnostic measurements. However, the use of magnetic diagnostics in future Fusion Power Plants (FPPs) may be complicated by limited access and increased distance from the plasma due to shielding and blanket structures. Furthermore, the scaling up of FPP size would require an expansion of the computational grid, which could reduce the performance of equilibrium reconstruction codes and make it challenging to satisfy real-time requirements.

Machine learning (ML) techniques can be employed to address these challenges. Significant progress has been made in applying ML approaches to replace EFIT, most notably in the EFIT-AI project [4, 5]. However, the majority of existing studies rely on comprehensive magnetic datasets as input features for reconstructing the plasma boundary or the full magnetic field distribution [6–8].

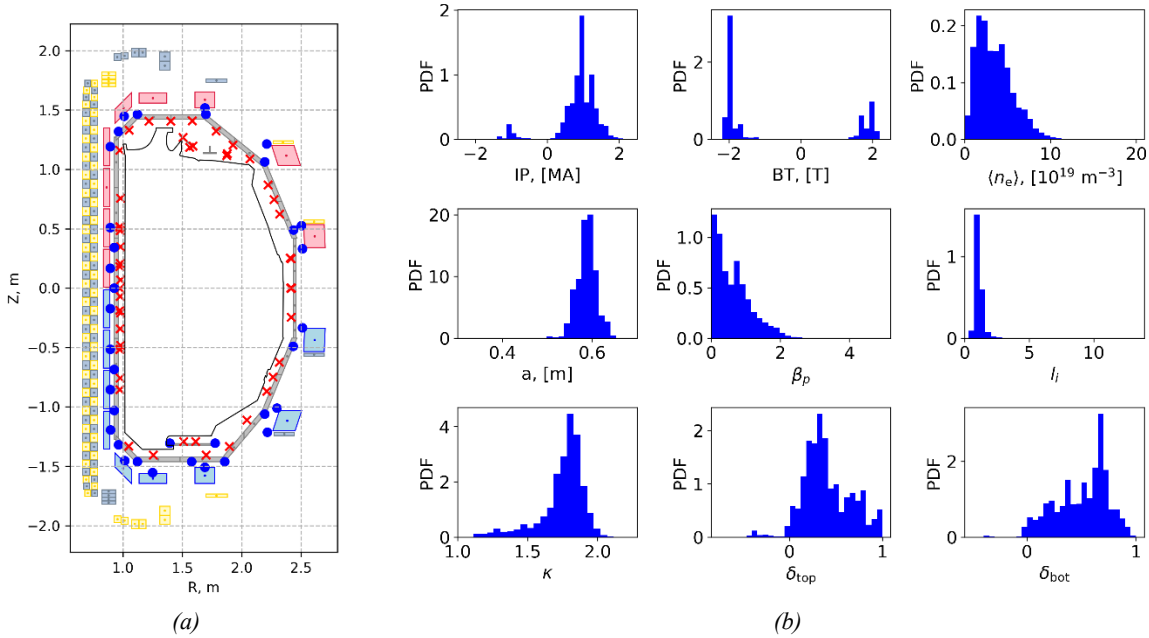
This study focuses exclusively on the reconstruction of the last closed flux surface (LCFS) to support plasma shape and position control. We aim to determine whether data-driven surrogate models, once trained, can provide boundary estimates from a reduced diagnostic set within acceptable error margins and, when designed for sensor robustness, serve as a rapid fallback in the event of diagnostic degradation or temporary unavailability.

2. DATASET

The dataset includes shots starting from 2020 year including shots from the negative triangularity campaign [9]. Shots with pulse duration less than 2 seconds were excluded. The rest of the shots were analyzed for the presence of maximum available trusted magnetic sensors (magnetic probes and flux loops). The RTEFIT algorithm uses 76 magnetic probes and 42 flux loops for equilibrium reconstruction. Some of these sensors can be turned off for a significant amount of time, so the set of used sensors is changing from shot to shot. This information is stored in the database in EFIT-snap files and in the plasma control system configuration. To simplify the training procedure, the set of sensors that fail less often has been chosen. This set of sensors is presented in 70% of the analyzed shots. The final dataset includes >7000 shots and >1.8 million timesteps.

An overview of the final dataset is presented in Figure 1. Figure 1a) shows the position of filtered probes and loops on the poloidal cross-section. Figure 1b) demonstrates histograms of general plasma parameters. The dataset includes shots with normal and inverted plasma current and toroidal field direction. High-density regimes and high βp (ratio of plasma pressure to poloidal magnetic field pressure) are also included. All shot phases (rampup, flattop, and rampdown) are preserved. However, their total duration is shorter than that of flattop phases that can be seen by a lesser presence of low a (minor radius), high li (internal inductance), and low κ (elongation) shots.

The dataset features both positive and negative triangularities. While the strong negative upper triangularity (δ_{top}) is more pronounced, there are also samples with negative lower triangularity (δ_{bot}).



(a) *(b)*

Fig. 1. (a) DIII-D tokamak layout and position of magnetic sensors that are all persist in 70% of shots in 2020-2025 time range; (b) General plasma parameters distribution in selected shots

3. ML MODEL TRAINING

Each model's task is to reconstruct the plasma boundary using the corresponding set of input features at the same timestamp. The models do not rely on information from prior states of the plasma or device and thus perform the same function as the EFIT code.

All the models predict the geometry of the plasma boundary as a vector of size N_c=2×N_p. This vector represents a flattened matrix of shape (N_p, 2), where N_p corresponds to the number of 2D points describing the plasma boundary at a given timestamp, which is set to 90 in our study.

The models are designed to receive the state of a discharge at a specific moment in time and compute the corresponding plasma boundary for that moment. Consequently, the training dataset comprises individual discharge states and can be represented as a matrix of shape (N, D+N_c), where N is the total number of discharge states, D corresponds to the dimensionality of the input feature space, and N_c = 180 represents the dimensionality of the plasma boundary vectors. Before being used for training, both the input features and the target outputs are standardized separately by subtracting their respective means and dividing by their standard deviations.

All the models share the same fully connected neural network (FCNN) architecture, consisting of two hidden layers with 150 and 80 neurons, respectively. To stabilize training and prevent overfitting, we apply batch normalization [10]. ReLU is used as the activation function.

To train the models, we employed the mean squared error (MSE) loss function, which has demonstrated strong performance in predicting plasma parameters [11, 12]:

$$\text{MSELoss} = \frac{1}{N}\sum_{i=1}^{N}\left\|b_i - \hat{b}_i\right\|^2, \quad (1)$$

where b_i is the ground truth boundary vector, $\hat{b}_i$ is the predicted boundary vector, i is the sample index, and N is the size of the training dataset. We used a learning rate of $1\cdot10^{-4}$ and the Adam [13] optimizer for training.

To prevent data leaks, we adopted a shape-based cross-validation approach. Discharge states were first divided into three main groups based on top and bottom triangularity; each group was then further subdivided into four

subgroups according to median values. This process produced 12 distinct sets of plasma shapes. At each cross-validation stage, the training–validation set consisted of 11 of these groups, while the remaining group was used for testing (see Figure 2). To prevent overfitting, we balanced the number of states in each group by subsampling, ensuring that every plasma shape group contained the same number of states. As a result, the training–validation–test ratio was approximately 80–10–10 at each cross-validation stage.

The test subset included discharges conducted in 2025 and in June, July and August of 2024. To prevent overfitting and ensure accurate model evaluation, we balanced the number of states corresponding to different plasma shapes in the training and testing sets using a subsampling approach similar to that used for cross-validation. As a result, the training–validation–test ratio was approximately 70–10–20.

Fig. 2. Distribution of shape parameters in dataset and example of reconstructed shapes on cross-validation subsets

In total 5 surrogate models were trained:

1. [I_{coil}] — least informed model
2. [I_{coil}] + I_p + V_{loop}
3. [I_{coil}] + I_p + [ψ-loops]
4. [I_{coil}] + I_p + [B_p-probes]
5. [I_{coil}] + I_p + [B_p-probes] + [ψ-loops] — baseline model

Here, quantities in square brackets [...] are the arrays, and quantities without them are scalars.

4. MODELS COMPARISION

To evaluate model performance during cross-validation and on the test set, three metrics were used:

- *Maximum Point Displacement (MXD) and Mean Point Displacement (MND)*:

$$MXD = \frac{1}{N}\sum_{i=1}^{N} \max_{j=1,N_p} \|p_{ij} - \hat{p}_{ij}\|, \tag{2}$$

$$MND = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{N_p}\sum_{j=1}^{N_p} \|p_{ij} - \hat{p}_{ij}\|, \tag{3}$$

where p_{ij} and $\hat{p}_{ij}$ are the j-th true and predicted 2D boundary points for the i-th sample, N_p is the number of boundary points, and N is the total number of samples. These two formulas aggregate the sample-wise metric values and are used to compute the values for the histograms shown in Figure 5. To compute the MND metric point-wise (i.e., as a function of poloidal angle), as shown in Figure 3, the following formula is used:

$$MND(\theta_j) = \frac{1}{N}\sum_{i=1}^{N} \|p_{ij} - \hat{p}_{ij}\|, \tag{4}$$

where θ_j is the poloidal angle of the j-th boundary point. Hereafter, all MXD and MND values are reported in meters and calculated in the original scale of the quantities.

- *Coefficient of Determination* R^2:

$$R^2 = \frac{\sum_{i=1}^{N_c}(V_i \cdot R_i^2)}{\sum_{i=1}^{N_c} V_i},\ V_i = \frac{1}{N}\sum_{j=1}^{N}\left(y_{ij} - \overline{y_i}\right)^2,\ R_i^2 = 1 - \frac{\sum_{j=1}^{N}\left(y_{ij} - \widehat{y_{ij}}\right)^2}{\sum_{j=1}^{N}\left(y_{ij} - \overline{y_i}\right)^2}, \quad (5)$$

where R_i^2 is the coefficient of determination for the i-th network output (corresponding to the R or Z coordinate of a boundary point), N_c = 180 (as 90 points are represented by both R and Z coordinates), y_{ij} and $\widehat{y_{ij}}$ are the true and predicted values of the i-th coordinate for the j-th sample, $\overline{y_i}$ is the mean true value of the i-th coordinate, and V_i reflects variance in y_i.

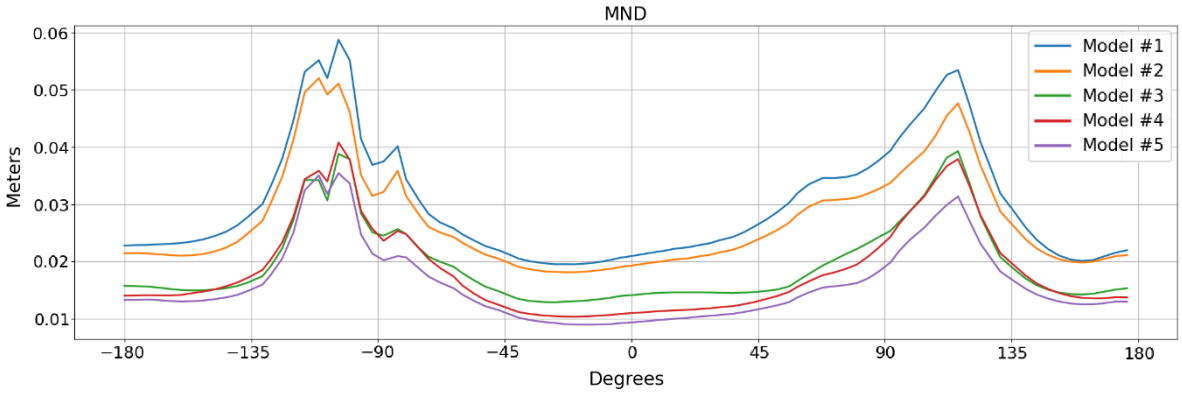


Fig. 3. Poloidal distribution of mean error. Zero angle corresponds to the LFS midplane, the negative value of the poloidal angle corresponds to the lower part of the plasma and the positive value to the upper.

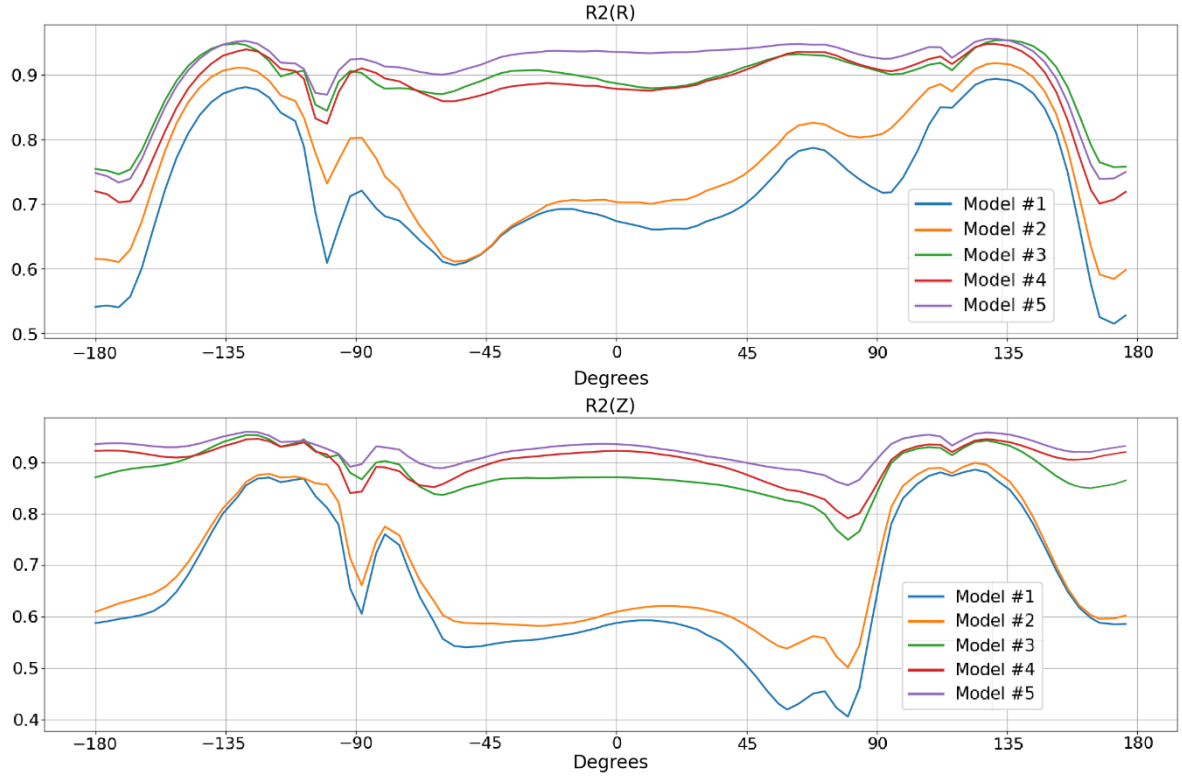


Fig. 4. Poloidal distribution of R^2 metrics for predictions of R and Z coordinates separately. Zero angle corresponds to the LFS midplane, the negative value of the poloidal angle corresponds to the lower part of the plasma and the positive value to the upper.

As expected, the most informed model #5 demonstrates the best performance. As shown in Figure 3, model #5 achieves a mean displacement ranging from 1 to 3.5 cm, with the maximum corresponding to the locations of the lower and upper X-points. Models #3 (flux loops only) and #4 (B_p-probes only) show similar average performance. About twofold deterioration is observed for models #1 and #2.

A more detailed assessment can be made by comparing the models' ability to predict the R and Z coordinates separately. As shown in Figure 4, the coefficient of determination, R^2, remains nearly flat for models #3 – #5 up to $\theta = \pm 135°$. This indicates that these models successfully capture the trends in the LFS and X-point regions. In contrast, the horizontal position of the HFS midplane region (essentially, inner plasma–wall gap) is more difficult to predict, despite exhibiting smaller absolute errors. Additionally, the least informed models #1 – #2 have difficulties with both LFS and HFS midplanes.

The distributions of mean and maximum displacements are shown in Figure 5. The mean error improves by nearly a factor of two when comparing models #1 and #5. In contrast, the maximum displacement – primarily associated with the X-point region – does not exhibit the same trend. The improvement in MXD remains in the range of 20 – 30%, depending on the region of the distribution. This difference originates from MSE loss functions that are used in model training. Due to this, models are optimized more for minimizing average displacement.

A key advantage of these models is their ability to differentiate performance across poloidal regions. Analysis of R^2 indicates that models #3–#5 successfully capture the boundary trends in LFS and X-point regions, whereas the HFS midplane remains more challenging to predict. This spatially dependent performance provides actionable insight for control algorithms: regions of higher uncertainty, such as X-points or the HFS midplane, can be monitored more closely or addressed with adaptive control strategies. In addition, the difference between mean and maximum displacement reveals that training models using MSE loss functions favors overall accuracy while partially neglecting extreme deviations. This suggests that for control applications sensitive to local boundary excursions, tailored loss functions or multi-objective optimization could further enhance ML model performance.

From a plasma control perspective, surrogate ML models offer several practical benefits. First, they provide computationally efficient, real-time compatible plasma boundary estimates, reducing reliance on full equilibrium reconstructions and enabling faster feedback loops. Second, by tracking discrepancies between ML predictions and standard reconstructions, these models can serve as early indicators of off-normal behavior or diagnostic degradation, adding a layer of robustness to control systems. Third, even when operating with a reduced diagnostic set, these models maintain sufficient accuracy in critical regions to support both global shape control and localized interventions, offering flexibility for next-generation tokamak designs where diagnostic access may be limited.

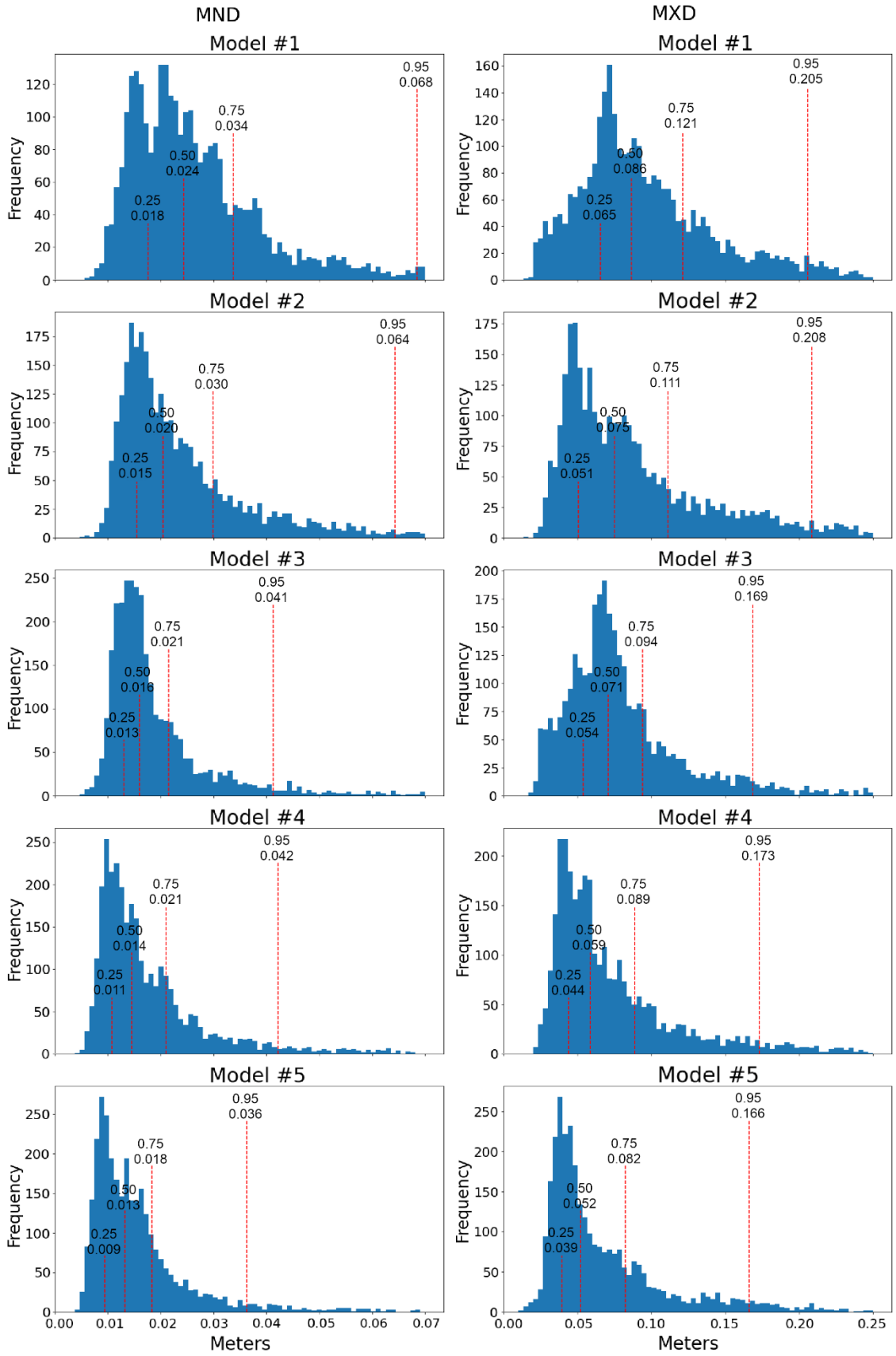


Fig. 5. Poloidal distribution of R^2 metrics for predictions of R and Z coordinates separately. Zero angle corresponds to the LFS midplane, the negative value of the poloidal angle corresponds to the lower part of the plasma and the positive value to the upper.

5. CONCLUSION AND DISCUSSION

Surrogate ML models allow achieving tolerable performance in plasma boundary reconstruction while operating with fewer diagnostic inputs. Beyond serving as lightweight surrogates for equilibrium solvers, these models can also play a supporting role in plasma control by providing early warnings of off-normal behavior. For example, a sudden divergence between ML-based predictions and standard reconstructions can signal diagnostic degradation or plasma state changes, enabling rapid detection and corrective action. In this way, models, that use limited number of sensors, not only maintain boundary reconstruction capability under constrained conditions but also enhance the resilience and safety of real-time control systems.

Overall, the findings suggest that ML models can play a valuable role in real-time plasma control applications. By providing rapid boundary estimates from a limited diagnostic set, these models enable controllers to maintain effective shape and position regulation even when some sensors are unavailable or degraded. The results show that models with more comprehensive diagnostic input (#5) achieve the lowest mean displacement, particularly in the LFS and X-point regions, while models (#3 – #4) maintain comparable performance across much of the plasma, demonstrating that a subset of well-chosen diagnostics can provide sufficient information for practical control tasks. Models with minimal inputs (#1 – #2) exhibit larger errors, particularly in the HFS midplane, highlighting the importance of strategic sensor placement and information content in reduced diagnostic configurations.

ACKNOWLEDGEMENTS

This material is based upon work supported partly by Next Step Fusion S.a.r.l and by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Award(s) DE-FC02–04ER54698.

REFERENCES

[1] J.R Ferron et.al. Real time equilibrium reconstruction for tokamak discharge control. Nuclear Fusion, 38(7):1055–1066, July 1998.

[2] J.-M. Moret et.al. Tokamak equilibrium reconstruction code LIUQE and its real time implementation. Fusion Engineering and Design, 91:1–15, February 2015.

[3] J Blum et.al. Real-time plasma equilibrium reconstruction in a tokamak. Journal of Physics: Conference Series, 135:012019, November 2008.

[4] (L.L. Lao et.al. S 2022 Application of machine learning and artificial intelligence to extend EFIT equilibrium reconstruction. Plasma Physics and Controlled Fusion 64 (7), 2022,074001.

[5] S. Madireddy et al. EFIT-Prime: Probabilistic and physics-constrained reduced-order neural network model for equilibrium reconstruction in DIII-D. Physics of Plasmas 31 (9), 2024,092505

[6] S. Joung et.al. Deep neural network Grad–Shafranov solver constrained with measured magnetic signals. Nuclear Fusion 60 (1), 2019, 016034.

[7] J.T. Wai et al. Neural net modeling of equilibria in NSTX-U. Nuclear Fusion 62 (8), 2022, 086042.

[8] X. Sun et al. Impact of various DIII-D diagnostics on the accuracy of neural network surrogates for kinetic EFIT reconstructions. Nuclear Fusion 64 (8), 2024, 086065.

[9] K E Thome et.al. Overview of results from the 2023 DIII-D negative triangularity campaign. 2024 Plasma Phys. Control. Fusion 66 105018

[10] S. Ioffe and C. Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37, ICML'15 1 p. 448–456, (2015), JMLR.org.

[11] J. Abbate et.al. Data-driven profile prediction for DIII-D. Nuclear Fusion 61 (4), 2021,046027.

[12] C. Wan et.al. Predict the last closed-flux surface evolution without physical simulation. Nuclear Fusion 64 (2), 2024, 026014.

[13] D.P. Kingma and J. Ba. Adam: A method for stochastic optimization. International Conference on Learning Representations abs/1412.6980, 2014.