DYNAMIC APERTURE PREDICTION BASED ON MACHINE LEARNING

J. Xu*, J. Xia, University of Science and Technology of China, Hefei, China

Abstract

The dynamic aperture (DA) is one of the most important parameters of nonlinear beam dynamics in storage rings. It describes the transverse phase space region where the motion of a particle can remain stable. In the design and optimization of storage rings, long-term particle tracking is usually required to ensure an sufficient DA. However this process is very time consuming. This study explores the possibility of using machine learning methods for DA prediction. Firstly, several regression models from magnet strengths to resonance driving terms are constructed using different machine learning methods, showing that the use of machine learning can be applied to the nonlinear performance analysis of storage ring lattice. Then predictive regression models from magnet strength to DA are constructed, and the results show that artificial neural network have better prediction accuracy. The method will be further developed for nonlinear analysis and optimization of storage ring.

INTRODUCTION

The Hefei Advanced Light Facility (HALF) is a diffraction-limited storage ring designed to deliver high-brightness synchrotron radiation [1]. Its ultra-low emittance is achieved through a hybrid multi-bend achromat lattice, which inherently introduces strong nonlinearities. The dynamic aperture (DA), a key metric of lattice nonlinear performance, must be sufficiently large to ensure stable beam injection and long beam lifetime. Conventionally, DA evaluation requires extensive particle tracking, which is computationally expensive and limits the scope of lattice optimization.

Machine learning (ML) has emerged as a powerful tool in accelerator physics, with applications in beam diagnostics, control, and lattice design. In this work, we propose an ML-based surrogate model to directly predict resonance driving terms and DA from magnet strengths, bypassing the need for time-consuming analytical calculations or particle tracking. We generate a large dataset of HALF lattice solutions using a localization method combined with multi-objective optimization, then train several regression models to map magnet settings to key nonlinear performance indicators. The results demonstrate that ML can effectively capture the complex relationship between lattice parameters and nonlinear behavior, providing a fast and reliable evaluation tool for lattice design and optimization.

DATASET GENERATION AND FEATURE ANALYSIS

The HALF storage ring lattice consists of 11 variable magnets: 8 quadrupoles and 3 sextupoles. To generate a diverse and representative dataset, we employed a localization method based on NSGA-II, which iteratively optimizes the upper and lower bounds of the variables to cover the feasible solution space. Constraints included:

- Natural emittance $\varepsilon_{\mathrm{nat}}$ <90 pm · rad,
- Tune fractional parts in $[0.1, 0.4]$,
- Dispersion at straight sections $|\eta|$ <0.002 m,
- Corrected chromaticity in $[2, 6]$.

A total of 90,000 feasible lattice solutions were obtained. For each solution, we recorded the 11 magnet strengths (features), and computed resonance driving terms: second- and third-order chromaticities ($\partial \nu_{x,y}/\partial \delta^2$, $\partial \nu_{x,y}/\partial \delta^3$) and amplitude-dependent tune shifts (ADTS), as well as the DA area via particle tracking using. To reduce redundant data for DA tracking, DBSCAN clustering was applied to the feature space, yielding 3,081 representative lattice configurations. Tracking these configurations took approximately one week on a personal computer.

A Spearman correlation analysis reveals several insights into the HALF lattice shown in Fig. 1. Strong negative correlation ($\rho < -0.8$) is observed between Q4 and Q5 in the high-dispersion region, consistent with their opposing roles in dispersion control. Horizontal second-order chromaticity shows strong negative correlation with dispersion at straight sections and with optical functions at sextupole SD2, suggesting that controlling these parameters is beneficial for managing chromatic aberrations. The three ADTS terms exhibit weak linear correlation with individual magnet strengths, reflecting the complexity of their underlying formulas. These correlation patterns provide valuable guidance for linear lattice design and inform the feature selection for ML models.

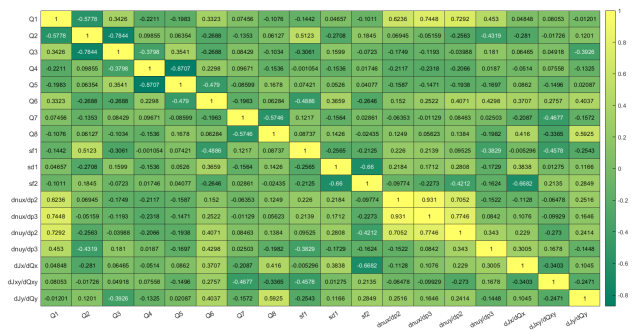

Figure 1: Correlations between single variables.

MACHINE LEARNING METHODS

Problem Formulation and Model Selecting

We formulate two distinct prediction tasks:

Task 1: Prediction of Resonance Driving Terms Predict seven nonlinear parameters (second- and third-order chromaticities, three ADTS terms) from 11 magnet strengths. This represents a multi-output regression problem where the

* xujianhao@ustc.edu.cn

target variables have known analytical expressions but involve complex combinations of lattice functions.

Task 2: Prediction of Dynamic Aperture Predict the DA area from 11 magnet strengths. This is a single-output regression problem with no closed-form expression, requiring surrogate modeling.

All input features (magnet strengths) are standardized using Z-score normalization:

$$x_{\text{new}} = \frac{x - \mu}{\sigma}, \tag{1}$$

where μ is the sample mean and σ is the sample standard deviation. This transformation eliminates scale differences and improves model training stability. Target variables are used without normalization to maintain physical interpretability.

Machine Learning Models

Four regression models were trained to predict each nonlinear parameter from the normalized magnet strengths:

- **Linear Regression**: Baseline model assuming linear relationships.
- **Polynomial Regression**: Degree-2 polynomial to capture quadratic interactions.
- **Random Forest (RF)**: Ensemble of 8 decision trees, max depth 50.
- **Artificial Neural Network (ANN)**: Two hidden layers with 32 neurons each, ReLU activation, trained with batch size 8 for up to 500 epochs.

Models were implemented using Scikit-Learn and Keras. Data were split into 80 % training and 20 % testing, with 5-fold cross-validation. Performance metrics included mean absolute error (MAE), root mean square error (RMSE), and R^2 score.

RESULTS AND DISCUSSION

Prediction of Resonance Driving Terms

Tables 1 through 3 summarize the prediction performance for each resonance driving term. For second-order chromaticities, polynomial regression achieves near-perfect accuracy ($R^2 = 0.99999$, MAE ≈ 1.8 for horizontal and 0.8 for vertical), demonstrating that quadratic interactions sufficiently capture the mapping from magnet strengths to these second-order aberrations. The ANN model also performs excellently ($R^2 = 0.9999$), while linear regression shows slightly larger errors but still high R^2 values above 0.99.

For third-order chromaticities, the ANN model demonstrates superior performance ($R^2 = 0.9999$, MAE = 104.36 for horizontal), significantly outperforming polynomial regression ($R^2 = 0.99$) and linear regression ($R^2 = 0.98$). The random forest model shows the weakest performance ($R^2 \approx 0.84 - 0.91$), suggesting that tree-based methods are less effective for capturing high-order polynomial relationships inherent in these terms.

Table 1: Prediction Performance for Second-Order Horizontal Chromaticity $\partial\nu_x/\partial\delta^2$.

Metric	Linear	Polynomial	RF	ANN
MAE	5.9865	1.8440	26.448	4.6472
RMSE	7.6660	2.4641	33.508	6.1927
R^2	0.9999	0.99999	0.92	0.9999

Table 2: Prediction Performance for Third-Order Horizontal Chromaticity $\partial\nu_x/\partial\delta^3$

Metric	Linear	Polynomial	RF	ANN
MAE	462.37	284.36	973.97	104.36
RMSE	618.69	385.25	1252.4	140.2
R^2	0.98	0.99	0.91	0.9999

The ADTS terms exhibit the highest prediction accuracy overall. For $\partial\nu_x/\partial\mathcal{J}_x$, polynomial regression achieves $R^2 = 0.9999$ with MAE of 1346.8, while ANN achieves comparable accuracy ($R^2 = 0.99$, MAE = 937.24). Similar trends are observed for the coupled ADTS terms and vertical ADTS. These results confirm that ML models, particularly polynomial regression and ANN, can effectively learn the analytical relationships encoded in resonance driving terms, even when intermediate optical functions are not explicitly provided as features.

Table 3: Prediction Performance for ADTS Term $\partial\nu_{x,y}/\partial\mathcal{J}_{y,x}$

Metric	Linear	Polynomial	RF	ANN
MAE	4571.7	1346.8	6449.9	937.24
RMSE	5857.7	1803.2	8573.7	1245.8
R^2	0.97	0.999	0.94	0.9999

Prediction of Dynamic Aperture

DA prediction represents a more challenging surrogate modeling task due to the absence of an analytical expression and the highly nonlinear dependence on lattice parameters. For DA tracking, we first applied DBSCAN clustering to about 90,000 lattice solutions to remove redundant configurations. With epsilon = 0.2 and minPts = 2, we obtained 3,081 representative clusters. Tracking these configurations using ELEGANT required approximately one week of computation time, yielding DA areas for model training.

Table 4 presents the performance of all four models on DA prediction. The ANN model significantly outperforms all other methods, achieving an R^2 of 0.80, MAE of 1.2277 mm^2, and MSE of 3.0178 mm^2 on the test set. Polynomial regression ($R^2 = 0.69$) and random forest ($R^2 = 0.58$) show moderate performance, while linear regression ($R^2 = 0.52$) fails to capture the nonlinear complexity of DA.

Figures 2 and 3 illustrate the ANN model's predictive capability. The mean relative error across test samples is 2.3 %, with approximately 95 % of predictions having relative error below 5 %. The residual distribution shows no systematic

Table 4: Prediction Performance for Dynamic Aperture Area (DA_{area} in mm^2)

Metric	Linear	Polynomial	RF	ANN
MAE	2.0911	1.6336	1.9042	1.2277
MSE	7.233	4.6329	6.5662	3.0178
R^2	0.52	0.69	0.58	0.80

bias, and the correlation coefficient between predicted and true values is 0.92 on the test set, indicating strong linear correlation despite the complex underlying mapping.

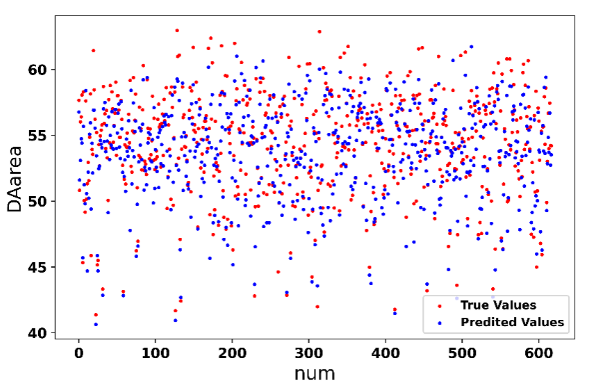

Figure 2: Comparison of the true and predicted values on the test set for ANN model.

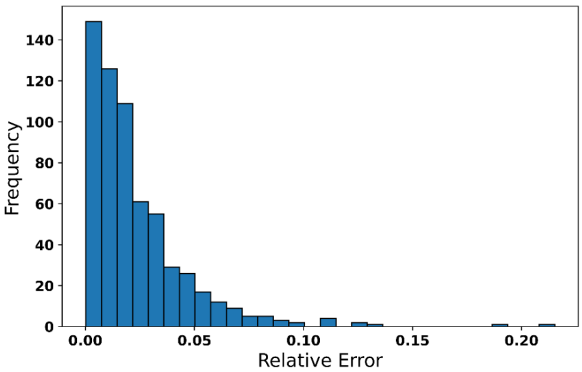

Figure 3: Histogram of relative error between true and predicted values of ANN model.

Figure 4 shows the training and validation MAE evolution during ANN training. The model converges smoothly after approximately 200 epochs, with no significant overfitting observed. The final validation MAE is slightly higher than the training MAE, indicating good generalization.

Comparison with Physics-Based Methods

The trained ANN surrogate model achieves DA prediction in milliseconds per lattice configuration, compared to tens of minutes for conventional particle tracking. This speedup of over 10^4 enables previously infeasible applications. While the prediction accuracy ($R^2 = 0.80$) is lower than for resonance driving terms, this level of accuracy is sufficient for early-stage lattice design and comparative ranking of candidate solutions. For final verification, limited particle tracking can be performed on the top-ranked candidates.

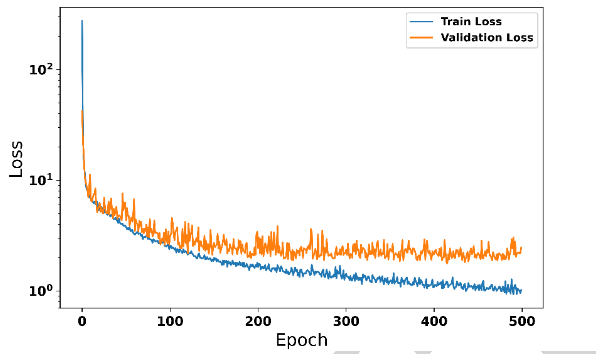

Figure 4: Training and validation MAE during ANN training.

CONCLUSION AND OUTLOOK

This study demonstrates the successful application of machine learning for predicting both resonance driving terms and DA in the HALF storage ring. The key findings are:

1. For resonance driving terms with known analytical expressions, polynomial regression and ANN models achieve near-perfect accuracy ($R^2 > 0.999$), effectively learning the underlying physics.
2. For DA, which lacks a closed-form expression, the ANN surrogate model achieves $R^2 = 0.80$ with a mean relative error of 2.3 %, providing a practical alternative to time-consuming particle tracking.

Future work will focus on three directions: (1) expanding the dataset to include more lattice configurations, particularly those near performance boundaries; (2) exploring more sophisticated deep learning architectures, such as generative models for physics-constrained data augmentation; and (3) integrating the surrogate model directly into multi-objective optimization frameworks for full-scale nonlinear lattice optimization. The methodology developed in this work is general and can be readily applied to other storage ring designs and to additional performance metrics such as momentum aperture and beam lifetime.

REFERENCES

[1] Z. Bai *et al.*, "Progress on the storage ring physics design of Hefei Advanced Light Facility (HALF)", in *Proc. 14th Int. Part. Accel. Conf. (IPAC'23)*, Venice, Italy, pp. 1075–1078, May 2023. doi:10.18429/JACoW-IPAC2023-MOPM038