ISBN: 978-3-95450-262-2 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2025-MOPCO02

PREDICTION OF FEL PERFORMANCE USING BPM MEASUREMENTS AND MACHINE LEARNING

C. Liu, X. Yang, Y. M. Deng, Y. B. Leng†
NSRL, University of Science and Technology of China, Hefei, P. R. China

Abstract

This study developed and validated a machine learning approach to analyze the correlation between beam position monitor (BPM) measurement data and output laser power in the Hefei Infrared Free-Electron Laser (FEL) facility. Using transverse position, charge, and longitudinal phase information from 280 individual bunches collected by BPM probes upstream of the undulator, we successfully constructed a high-precision predictive model, demonstrating that BPM measurements can effectively predict the output laser power of the infrared FEL.

Based on the trained predictive model, we further deconstructed the neural network architecture to accurately identify key bunches and sensitive parameters that most significantly influence laser power output. This provides a clear and targeted optimization basis for subsequent beam tuning experiments. The data-driven strategy employed in this method significantly reduces the workload associated with traditional experience-based tuning, offering an effective technical means to enhance accelerator operational stability.

INTRODUCTION

Free-electron laser (FEL) facilities, as advanced fourth-generation coherent light sources, play an indispensable role in cutting-edge scientific fields. FELiChEM, a major research infrastructure in China, relies critically on the quality and stability of its radiation for successful user experiments[1, 2]. However, direct diagnostics of FEL radiation often involve destructive methods or lack real-time capability, introducing interference or delays during operation.

The core challenge lies in electron beam quality, which is highly sensitive to accelerator conditions yet poorly quantitatively understood. Traditional tuning approaches depend heavily on operator experience and trial-and-error, resulting in low efficiency and poor reproducibility.

Recent advances in beam position monitor (BPM) systems enable non-invasive, simultaneous measurement of multiple bunch parameters (such as transverse position, longitudinal phase, bunch length, etc.), providing rich diagnostic data [3]. However, the vast data volume exceeds the capabilities of manual analysis.

This study develops a data-driven machine learning approach to address these limitations. Using multi-parameter bunch data from upstream BPMs, we establish a virtual diagnostic model for predicting radiation quality in real-time, without disruption to user experiments. Beyond online monitoring, the trained neural network is interpreted to identify the most influential bunches and physical parameters. This provides explicit optimization guidance—such as which BPM parameters to adjust or which bunch regions to stabilize—enabling precise beam tuning via correctors or RF parameters. Ultimately, this work facilitates a shift from experience-dependent operation to a data-driven intelligent paradigm for accelerator control.

METHODOLOGY

This chapter details the data-driven modeling approach used to establish the relationship between FEL laser power and BPM measurements. The overall research framework consists of three core components: data preprocessing and feature engineering; construction and training of a neural network prediction model; and model interpretation and feature importance analysis.

Data Preprocessing and Feature Construction

The data used in this study, benefiting from groundbreaking advances in bunch-by-bunch diagnostic techniques, were obtained from the Hefei Infrared Free-Electron Laser facility[4,5]. Specifically, they were collected from six upstream beam position monitor (BPM) probes (numbered 1, 2, 3, 7, 8, 9) and one downstream BPM probe (numbered 10) of the undulator. Each BPM probe measured parameters for 280 individual bunches within a macropulse, including transverse position (X, Y), Sum signal (Sum), and longitudinal phase (Phase)[6,7]. These parameters were calculated by the HOTCAP program[8-10], with one turn of data shown in Fig. 1.

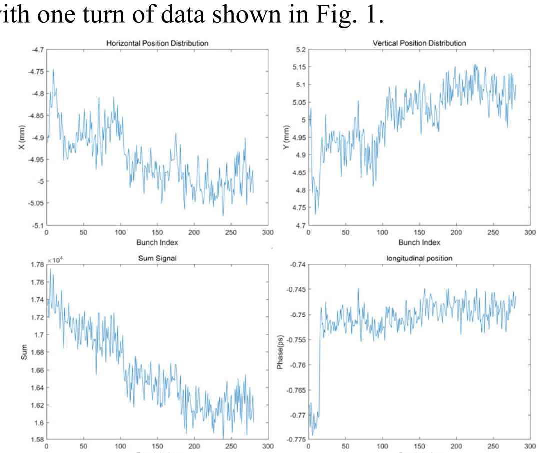

Figure 1: BPM measurement parameters display.

† e-mail address: lengyb@ustc.edu.cn

ISBN: 978-3-95450-262-2 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2025-MOPCO02

A schematic diagram of the typical installation positions of the BPMs along the undulator beamline is shown in Fig. 2. highlighting their coverage of key regions from the linear accelerator exit to the undulator entrance. The raw data were stored in MATLAB file format, comprising 28 parameter matrices of dimension 280 (bunches) × 1033 (macropulses), along with a corresponding 1 × 1033 vector of laser energy values.

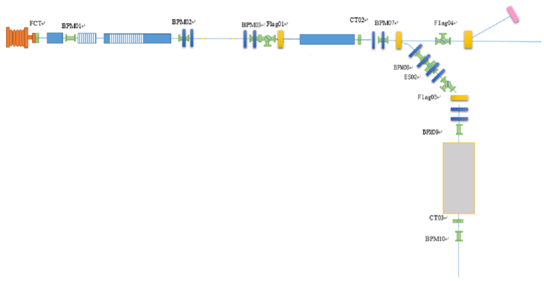

Figure 2: The beam measurement system in FELiChEM.

To prepare the data for machine learning modeling, the following preprocessing steps were performed: First, all parameters for each bunch were directly used as features, and the raw data were restructured into an input matrix X of dimension 1033 (samples) × 7840 (features) and a target vector y of dimension 1033. Next, to mitigate numerical biases caused by differences in the units of various physical parameters (e.g., phase and charge), features were grouped by type (Phase, Sum, X, Y) and standardized using Z-score normalization. The target variable, FEL energy, was globally standardized. The processed dataset was randomly split into training and independent test sets in a 9:1 ratio to ensure unbiased model evaluation.

Neural Network Model Architecture

To model the complex nonlinear relationship between BPM multi-parameters and laser power, we designed a deep fully connected neural network (FELPowerPredictor) as illustrated in Fig. 3. The network architecture processes the 7840-dimensional input feature vector through three progressively compressed hidden layers (2560 → 640 → 32 neurons). Each layer implements the transformation[11]:

$$\mathbf{h}^{(l)} = \sigma\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right), \tag{1}$$

where $h^{(l)}$ represents the activation value of this layer, and $W^{(l)}$ and $b^{(l)}$ are the weight matrix and bias vector of that layer, respectively. σ is the nonlinear activation function, and in this paper, we selected the Rectified Linear Unit (ReLU), i.e., σ(z) = max(0, z).Between each hidden layer, we incorporate Batch Normalization to accelerate convergence and Dropout (rate=0.2) to prevent overfitting. The final output layer performs a linear transformation to generate the continuous laser power prediction. This hierarchical design enables the network to progressively extract relevant features while maintaining computational efficiency.

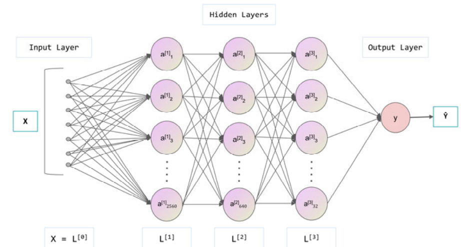

Figure 3: FELPowerPredictor neural network.

The model training aims to minimize the Mean Squared Error (MSE) between the predicted power and the actual power. The loss function L is defined as:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 , \tag{2}$$

where **N** is the number of training samples. We adopted the Adaptive Moment Estimation (Adam) optimizer to update the network parameters, with the learning rate set to 0.001. The training was conducted under the PyTorch framework, and the TensorBoard tool was used to monitor the loss changes during the training process. The model was fitted on the training set, and its performance was ultimately evaluated on an independent test set to ensure its generalization capability. The model with the lowest validation loss was saved for subsequent analysis and application.

Model Deconstruction Analysis

To identify key parameters influencing laser power output, we employed permutation importance analysis to deconstruct the trained predictive model[12]. This method quantifies feature importance by randomly shuffling each feature's values in the test set and measuring the corresponding decrease in model performance (MSE). A larger performance degradation indicates a more important feature.

In practice, we first established baseline performance on the original test set. Each feature was then shuffled multiple times (5 iterations), with the importance score calculated as the difference between baseline and average permuted performance. Positive scores indicate meaningful feature contributions.

By analyzing the top-ranked features in terms of importance, we can map the machine learning results back to physical reality, precisely identifying critical BPM probes, sensitive parameter types, and key bunch regions. This analysis goes beyond "black-box" prediction and provides clear optimization directions and quantitative basis for beam tuning, such as prioritizing the calibration of phase measurements for specific BPM probes or enhancing stability control in critical bunch regions. The model-agnostic nature and intuitiveness of the permutation importance method ensure the reliability and interpretability of the analytical results.

ISBN: 978-3-95450-262-2 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2025-MOPCO02

ANALYSIS AND DISCUSSION

This study conducted 3,000 training epochs using an RTX 4090 GPU, with the model achieving a coefficient of determination (R^2) of 0.990 on the training set and 0.843 on the test set. Figure 4 shows the scatter plot of predicted versus actual values and the curve of predicted/actual values versus sample indices from the training set, demonstrating nearly perfect fitting performance. The data points are tightly distributed on both sides of the y=x reference line, and the predicted values almost completely coincide with the actual values in the time series curve.

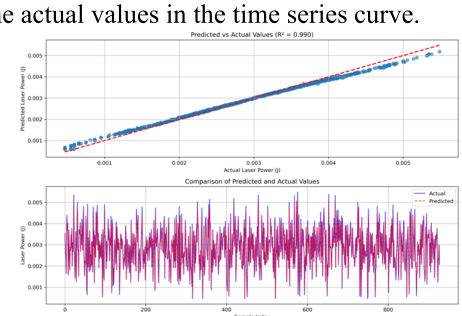

Figure 4: Prediction performance on training set.

Figure 5 presents the results on the test set, where the scatter plot shows that most data points are distributed near the y=x reference line, though some outliers exist. The time series curve indicates that the model effectively captures the overall variation trend of laser power, though some deviations or delays in prediction occur at certain abrupt change points.

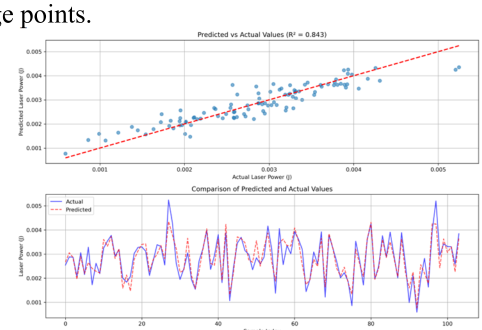

Figure 5: Prediction results on test set.

The excellent performance on the training set demonstrates the powerful fitting capability of the neural network model, while the performance gap on the test set indicates some degree of overfitting. Analysis suggests that the primary reason for the limited predictive performance is the relatively small training dataset (1,033 macro-pulse samples). For learning complex nonlinear relationships involving 7,840 features, more training data would help the model better capture the precise mapping between input features and output power, reduce overfitting risk, and improve generalization capability.

Feature analysis based on the permutation importance method revealed key factors affecting laser power. The BPM importance bar chart in Fig. 6 shows that the importance of BPM9 and BPM10 is significantly higher than other probes. This finding strongly aligns with physical intuition: as the last beam position monitor upstream of the undulator, BPM9's measurements directly reflect the electron beam state before entering the undulator; while as the first monitor downstream of the undulator, BPM10 captures the electron beam changes after interaction. The combined information from these two probes effectively characterizes the key physical state of the FEL gain process.

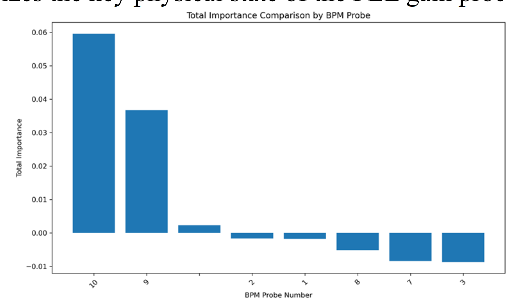

Figure 6: BPM importance.

Regarding bunch importance, the bunch range importance bar chart in Fig. 7 indicates that the head bunches (No. 0-50) have the most significant impact on laser power. However, these bunches show considerable transverse distribution deviation, falling outside the beam core range. It remains unclear whether this phenomenon is related to their importance.

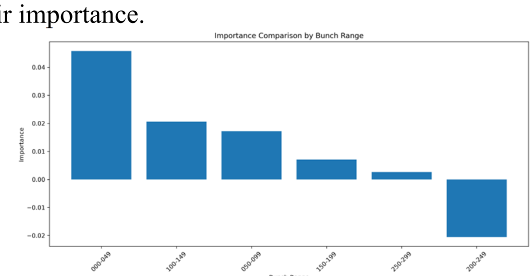

Figure 7: Bunch range importance.

The current dataset reveals significant differences in the distribution of parameter importance across various BPM probes. Specifically, transverse position parameters demonstrate the highest importance in BPM10, while the sum signal dominates in importance among most other probes. This differential feature distribution is likely associated with the current maturity level of the prediction model. Although the model trained on limited data volume has achieved a coefficient of determination (R^2) of 0.843 overall, demonstrating good predictive capability, it may still lack sophistication in fine-grained learning of feature correlations, potentially affected by data volume constraints and overfitting risks. We anticipate that with the accumulation and incorporation of more experimental data into training, the model will not only achieve further improvements in predictive accuracy but, more importantly, will develop more consistent and stable evaluations of parameter importance distribution patterns—either by eliminating the currently observed importance differences or by providing clearer and more reasonable physical explanations for these differences.

ISBN: 978-3-95450-262-2 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2025-MOPCO02

REFERENCES

[1] H.T. Li *et al.,* "Design of FELiChEM, the first infrared free-electron laser user facility in China," *Chin. Phys C*, vol. 41, no. 1, p. 018102, Jan. 2017.
grl=43143;;249:70446:27424234;435

[2] H. T. Li and Q. K. Jia, "Commissioning and First Lasing of the FELiChEM: A New IR and THz FEL Oscillator in China", in *Proc. FEL'19*, Hamburg, Germany, Aug. 2019, pp. 15-19.
grl=4314;75<2MDFrZ0IHO534<0PRD3:

[3] X. Yang, Y. B. Leng, and Y. M. Zhou, "Real-time optimizatio n of bunch-by-bunch 3D information extraction software HO TCAP". *Nucl. Tech.*, vol. 47(2), p. 020102, 2024
grl=43144;;<2m135860654<153571kmv17:1353435

[4] Y.M. Deng *et al.*, "Current status and latest developments in electron storage ring bunch-by-bunch beam diagnostic techn iques", *Nucl. Tech.*, vol. 47(10), p 100201, 2024.
grl=43144;;<2m135860654<153571kmv17:1433534

[5] Y.-M. Deng, Y.-B. Leng, X.-Y. Xu, J. Chen, and Y.-M. Zhou, "Ultrahigh spatiotemporal resolution beam signal reconstruction with bunch phase compensation," *Nucl. Sci. Tech.*, vol. 35, no. 5, May 2024.
grl=431433:2v7469803570347770|

[6] X.-Y. Xu et al., "HOTCAP: a new software package for high-speed oscilloscope-based three-dimensional bunch charge and position measurement," *Nucl. Sci. Tech.*, vol. 32, no. 11, Nov. 2021. grl=431433:2v746980354033<990}

[7] X. Yang et al., "Bunch-resolved 3D beam position measurement system and its application in FELiChEM", in *Proc. Proc. IBIC'2024*, Beijing, Sep. 2024, pp. 23-26.
grl=4314;75<2MDFrZ0LELF53570WXFL4

[8] H. Wang, Y. Leng, and Y. Zhou, "Technique for Extracting Initial Parameters of Longitudinal Phase Space of Freshly Injected Bunches in Storage Rings, and Its Applications," *Instrum. Exp. Tech.,* vol. 9, no. 3, p. 17, Jul. 2025.
grl=43166<32lqvwuxphqwv<36334:

[9] X. Yang, H.-S. Wang, Y.-M. Zhou, and Y.-B. Leng, "Determining beam transverse absolute position by triangulation of multi-electrode signal phase differences," *Nucl. Sci. Tech.*, vol. 35, no. 8, Jul. 2024.
grl=431433:2v7469803570347<;0|

[10] H.-S. Wang, X. Yang, Y.-B. Leng, Y.-M. Zhou, and J.-G. Wang, "Bunch-length measurement at a bunch-by-bunch rate based on time–frequency-domain joint analysis techniques and its application," *Nucl. Sci. Tech.*, vol. 35, no. 4, Apr. 2024.
grl=431433:2v7469803570347760}

[11] Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors" *Nature*, vol. 323.6088, pp. 533-536, 1986.
grl=431436;2656866d3

[12] L. Breiman, "Random Forests," *ACM Trans. Probab. Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
grl=43143562d=4343<66737657