ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-TUPS63

OVERVIEW OF MACHINE LEARNING BASED BEAM SIZE CONTROL DURING USER OPERATION AT THE ADVANCED LIGHT SOURCE*

Thorsten Hellert†, Tynan Ford, Simon C. Leemann, Hiroshi Nishimura, Marco Venturini
Lawrence Berkeley National Laboratory, Berkeley, CA, USA
Andrea Pollastro, IMPALab, DIETI, University of Naples Federico II, Naples, Italy

Abstract

Past research at the Advanced Light Source (ALS) provided a proof-of-principle demonstration that machine learning (ML) methods could be effectively employed to compensate for the significant perturbations to the transverse electron beam size induced by user-controlled adjustments of the insertion devices. However, incorporating these methods into the ALS' daily operations has faced notable challenges. The complexity of the system's operational requirements and the significant upkeep demands have restricted their sustained application during user operation. In this paper we summarize the development of a more robust ML-based algorithm that utilizes a novel online fine-tuning approach and its systematic integration into the day-to-day machine operations.

INTRODUCTION

The performance of storage ring light sources is critically reliant on the stability of the radiation output in terms of source position/angle and intensity. On shorter timescales, radiation intensity stability is impacted by the electron beam's transverse size response to changes in the insertion device (ID) parameters during operation. Typically, in 3rd generation light source horizontal beam size remains largely constant across different ID settings, while vertical size is sensitive to skew quadrupole-field errors from IDs. To correct these, storage rings like the Advanced Light Source use quadrupole and skew quadrupole correctors in a feed-forward configuration [1]. Corrections based on beam measurements create lookup tables specifying necessary lattice corrections for each ID, which combine via linear superposition. However, these corrections are compromised by short-term drifts during measurements and long-term drifts from external factors, reducing their effectiveness.

Although feedback systems can compensate for these drifts based on real-time beam size monitoring, noise limitations and insufficient closed-loop bandwidth impede their effectiveness. Recent studies suggest augmenting lookup tables with a neural network trained on beam size, ID parameters, and dispersion wave data to improve control [2], yet sustained implementation has been challenging.

This paper summarizes the findings and advancements detailed in [3], focusing on the online fine-tuning mechanism and the performance during user operation. An illustration of the effectiveness of the method is given in Fig. 1 showing data during about a week of ALS user operation after its implementation in the fall of 2023. As the users continually adjust the ID parameter setpoints to accommodate their experiments (the traces in the two top graphs), the measured rms vertical beam size is seen to remain stabilized within a band that is very close to the estimated ~0.3 µm rms noise floor (red trace in the bottom graph). For comparison, the plot also shows the inferred beam size (blue trace) that would have been observed with the ML-based feed-forward (FF) system turned off.

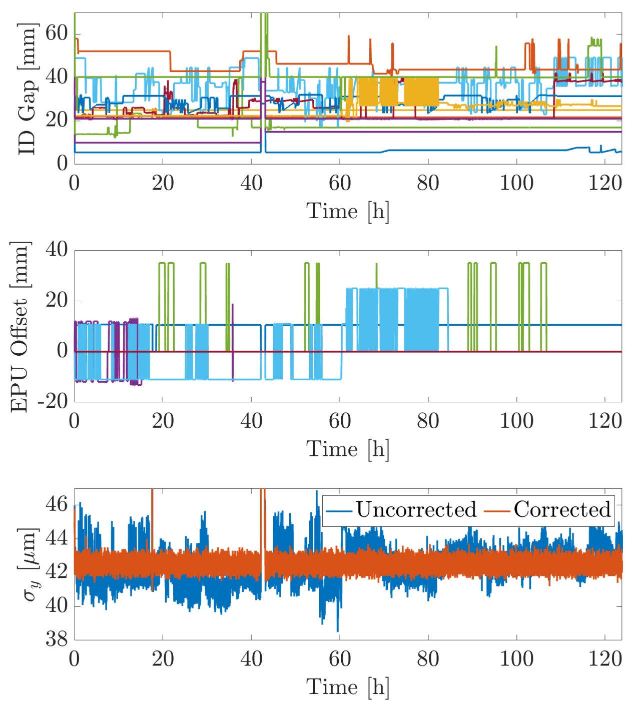

Figure 1: Operational performance of the ML-based ID FF system during a user run starting on November 7, 2023. Shown are the vertical ID gaps (top), the elliptical polarized undulator (EPU) phase or longitudinal offsets (center), and the vertical electron beam size (bottom) as measured at ALS diagnostic beamline 3.1 (red) and as inferred (blue) if no correction had been applied. One beam outage occurred at hour 42 during that 5 day window; notably, the beam size control algorithm dis- and re-engaged automatically without human intervention (see Section for a detailed discussion).

ONLINE FINE-TUNING

As detailed in Ref. [3], the task of training a base model capable of independently and effectively predicting beam size

* This work is supported by the U.S. Department of Energy (DOE) under Contract No. DE-AC02-05CH11231.
† thellert@lbl.gov

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-TUPS63

presents a formidable challenge. The scarcity of available machine time for acquiring sufficient training data necessitates the adoption of an alternative approach. This method involves the continuous adaptation of the model during user operations.

During the proof of principle studies [2], such adaptation has been applied through online retraining by integrating the original training data with a randomly downsampled segment from the ongoing user run, subsequently continuing the training of the active model. However, with hundreds of thousands of samples in the data set, retraining demanded approximately 15 min to complete on a CPU, which significantly limits the model's reactivity to changes in the ID configuration space.

However, these challenges are widely recognized in the Deep Learning community. Training a neural network (NN) from scratch requires substantial computational power, memory resources, and large training datasets [4]. Fine-tuning [5, 6] is a widely used solution for addressing these challenges. The core of this approach lies in the realization that rapid and efficient adaptation to new data—achieved through the adjustment of NN weights—can greatly enhance the model's predictive accuracy while circumventing the extensive data processing typically necessitated by comprehensive training phases [7].

In our approach to fine-tuning the model [3], we exclusively utilize data acquired during the current user run which is stored in a first-in-first-out buffer, opting not to incorporate the original training data. To safeguard against any potential runaway scenarios, each training cycle commences with the original base model. This strategy effectively *anchors* the model, ensuring it remains closely aligned with our dedicated training dataset, thereby maintaining stability and reliability in the model's predictive performance.

An ostensible difficulty is that to fine tune the NN during operation, one would need to know the uncorrected beam size data, whereas only measurements of the beam size after correction are available, since the NN FF system is always active. One method to overcome this difficulty is to derive the presumed uncorrected beam size σ_{y0} from knowledge of the measured corrected value σ_y and the current vertical dispersion wave (see Ref. [3]).

It is worth highlighting that the online fine-tuning method employed in our study essentially functions as a form of feedback (FB). By adjusting the amount of data in the fine-tuning buffer, we control the noise level in the data, and by tuning the fine-tuning hyperparameters, we modulate the model's responsiveness to new data. The balance between FF and FB elements in such an elegant way is a key feature of our ID compensation algorithm which contributes significantly to its robustness.

NN FF SYSTEM PERFORMANCE

In this section we evaluate the capabilities of the NN-based ID FF algorithm to stabilize the vertical beam size at the ALS. As detailed in Ref. [2], scanning transmission x-ray microscopy (STXM) [8] beamlines are very sensitive to variations of the transverse photon distribution and the quality of their experiments can be significantly impacted by such fluctuations.

Our measurements have established a linear relationship between variations in the vertical beam size at the diagnostic beamline 3.1 and subsequent changes in intensity observed in STXM scans taken at beamline 5.3.2.2 (consistent with resulting vertical beam size changes being driven by perturbations of the vertical dispersion wave, the dominating contribution to the vertical emittance). Specifically, we observed that a 10 % change in the vertical beam size resulted in approximately a 9 % intensity change in the STXM scan, which was common during user operations over the duration of a STXM measurement.

Performance During User Operation

At the time of writing, the NN-based FF system had been continuously operational for two months. The performance over about one week was showcased in the introduction in FIG 1 and is typical. The vertical beam size stability has been remarkably consistent, to within an average of 0.32 µm rms per user run (or 0.75%), closely approaching the measurement noise floor at 0.3 µm rms.

This can be seen in more detail in FIG 2: data points to the right of the shaded area, where the two-month period with operating NN FF system is segmented into 7 uninterrupted user operation intervals. For each interval we report the rms beam size fluctuations before correction (blue), and as corrected using the base model without fine-tuning (red), and finally as measured with correction by the fine-tuned system (crosses). The blue and red data points are inferred quantities; specifically the blue data were obtained by sub-

Figure 2: The data on the right of the shaded area is with the NN FF system fully deployed: crosses represent the vertical beam size rms fluctuations as corrected by the fine-tuned NN model and measured. They are compared to the uncorrected (blue circles) and partially corrected (red circles) beam size fluctuations, the latter representing the correction made by the NN FF system without fine-tuning; the values for both of these data sets are inferred estimates (see body text). Each data point is a time average over about one week of user operation. The data points on the left of the shaded area represent a backward-test of the NN model based on archived data.

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-TUPS63

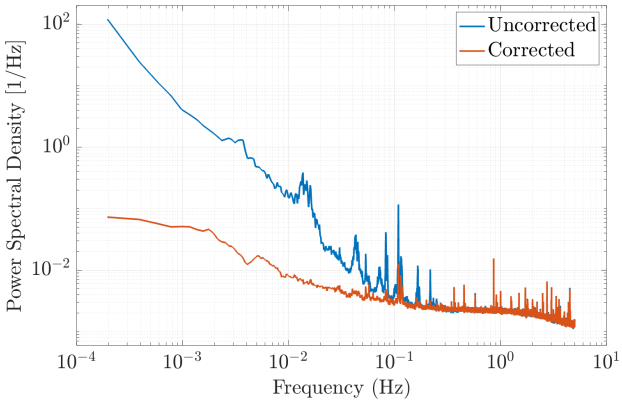

Figure 3: Frequency spectrum of the vertical beam size during 760 h of user operation with fully functional NN FF system (red), and the inferred beam size without correction (blue). The features at around 0.1 Hz are attributed to EPUs, the spikes above 1 Hz are associated to beam injection transients during top off.

tracting the contribution due to the vertical dispersion wave adjustments from the measurement of the stabilized vertical beam size. Barring the small hysteresis effects not accounted for in our simplified model (see Ref. [3]), we believe the uncorrected beam size so calculated should be a fairly accurate estimate of the actual beam size that would have been observed without correction. Note that in the figure the data points are time averages during the operation period (about a week).

The red data points preceding the shaded area are the result of a study in which the NN model trained on June 2024 was retroactively applied to archived data of beam and ID parameters from the preceding year, which is not suitable for training a model [3]. The backward-test results indicate that the uncorrected and corrected data points remain consistent over time, suggesting minimal impact from machine drifts on the neural network model's accuracy and implying that frequent model refreshes may be unnecessary and accuracy could be enhanced by accumulating extended user operation data.

Additional insight can be gained by data analysis in the frequency domain. In FIG 3, a discrete Fourier transformation of 760 h of user operation data with the NN FF system on confirms the system's effectiveness over a broad frequency range. Spikes at around 0.1 Hz, linked to EPU phase switching, indicate areas for improvement with better training data. The loss of correction effectiveness at lower frequencies is also influenced by EPU phase switching, which, occurring over extended periods, can cause low-frequency data modulation. However, the total integrated spectral power in this range is small, minimally impacting overall performance and emphasizes the robust performance of the NN-based FF system across the evaluated time frame.

Recovery After Beam Outage

During the two-month operation, the facility experienced 12 beam outages, each followed by recovery. Intervals between events ranged from minutes to hours. In each instance, the NN-based FF disengaged at the trip and autonomously re-engaged before user operation resumed, without manual intervention.

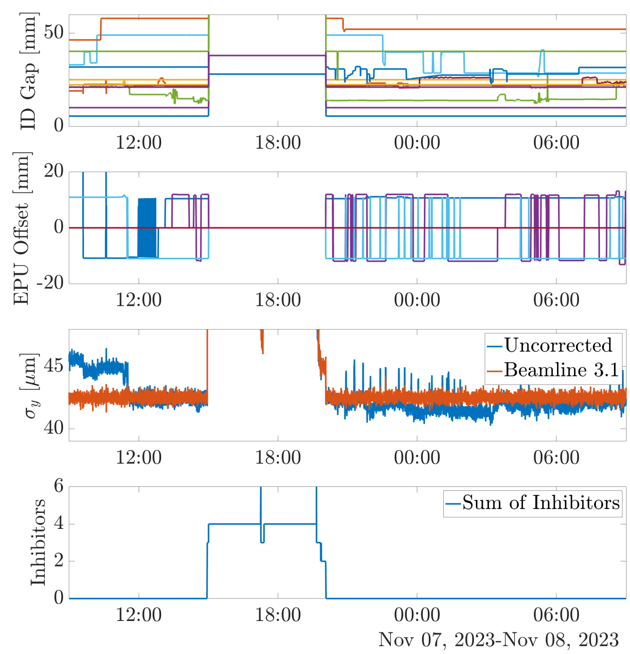

Figure 4: Example of beam dump with subsequent restart of the NN-based FF without human intervention. The top two plots show ID gaps and EPU offsets, respectively. The 3rd plot shows the vertical beam size and the lower plot the sum of the currently active inhibitor PVs.

An example of a beam outage event, caused by an RF power trip, followed by a machine refill and closing of the NN-based FF loop without human intervention, is shown in FIG 4. The beam was lost at 14:57, immediately triggering three of the six inhibitor PVs designed to prevent the NN-based FF from acting on the skew quadrupoles under conditions that are not operationally safe and reliable. At 17:15, during the process of reloading the lattice, the skew quadrupole power supplies exhibited transient conditions, as indicated by the activation of all six inhibitor PVs. Following this, the machine was refilled. However, a subsequent RF fault caused another beam loss. The machine was successfully filled at 19:56, which was then followed by the closure of the ID gaps. From 20:03 onwards, all conditions for closing the FF loop were met, and skew quadrupole corrections were one again applied. This is evidenced by the vertical beam size returning to its target value of 42.5 µm.

REFERENCES

[1] C. Steier, S. Marks, S. Prestemon, D. Robin, R. D. Schlueter, and A. Wolski, "Study of Row Phase Dependent Skew Quadrupole Fields in Apple-II type EPUs at the ALS", in *Proc.*

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-TUPS63

EPAC'04, Lucerne, Switzerland, Jul. 2004, paper MOPKF071, pp. 479–481.

[2] S. C. Leemann *et al.*, "C. Demonstration of Machine Learning-Based Model-Independent Stabilization of Source Properties in Synchrotron Light Sources", *Phys. Rev. Lett.*, vol. 123, p. 194801, 2019.

[3] T. Hellert *et al.*, "Application of deep learning methods for beam size control during user operation at the Advanced Light Source", *Phys. Rev. Accel. Beams*, to be published.

[4] R. Patel and A. Chaware, "Transfer learning with fine-tuned MobileNetV2 for diabetic retinopathy", in *Proc. 2020 International Conference For Emerging Technology (INCET)*, pp. 1–4, 2020.

[5] T. Goodfellow, Y. Bengio,and A. Courville, *Deep learning*, MIT press, 2016.

[6] C. Bishop and H. Bishop, *Deep Learning: Foundations and Concepts*, Springer Cham, 2023. doi:10.1007/978-3-031-45468-4

[7] Y. Guo *et al.*, "Spottune: transfer learning through adaptive fine-tuning", in *Proc. IEEE/CVF Conference On Computer Vision And Pattern Recognition*. pp. 4805–4814, 2019.

[8] T. Feggeler, A. Levitan, M. Marcus, H. Ohldag, and D. Shapiro, "Scanning transmission X-ray microscopy at the Advanced Light Source", *J. Electron. Spectrosc. Relat. Phenom.*, vol. 267, p. 147381, 2023.