SURROGATE MODELS FOR THE EUROPEAN XFEL OPERATION

B. Veglia*, J. Kaiser, S. Tomin, S. Walker,
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Abstract

Numerical beam dynamics simulation codes are essential for designing and studying particle accelerators, but their computational cost can make them unsuitable for online use and predictions during operations. The use of machine learning-based surrogate models can significantly reduce the required computational time whilst still providing an accurate prediction of the beam properties. In this paper, we present the first results on the training of surrogate models for the prediction of the longitudinal phase space (LPS) at the European XFEL. Finally, we discuss the potential application of such models in the development of a virtual diagnostic tool for use in the European XFEL control room as well as a fast estimator for the final LPS based on the user-provided compression parameters.

INTRODUCTION

The rise of X-ray free-electron lasers (XFELs) over the past decade has provided researchers with ultra-short, coherent X-rays capable of probing matter at the atomic level. At the heart of these facilities is the electron beam; information on its longitudinal phase space (LPS) is vital for achieving the desired lasing performance and developing advanced FEL modes. To monitor this property effectively, scientists rely on longitudinal diagnostics instrumentation. The European XFEL (EuXFEL), schematically shown in Fig. 1, is a 3.4 km hard X-ray FEL in the Hamburg region, Germany. It is equipped with two transverse deflecting structures (TDSs) [1] — one located in the injector diagnostics section and another downstream of the third bunch compressor (BC2) — a coherent radiation intensity spectrometer (CRISP) [2, 3], which reconstructs the current profile, and a single-plate corrugated structure (also referred to as a passive streaker) [4] installed downstream of SASE2. However, all of these diagnostics have important limitations. The TDS downstream of BC2 provides a temporal resolution of only about 10 fs, while the CRISP spectrometer does not measure the phase of the emitted coherent radiation, making the reconstruction ambiguous with respect to the beam orientation and particularly sensitive to noise at low bunch compression. Finally, the analysis of the streaked beam after after interacting with the passive streaker requires a complicated reconstruction [5] making it non-compatible with online tuning.

Ideally, high-resolution TDSs would be situated downstream of the undulator lines in order to directly inspect the LPS in order to directly tune the undulator lasing. However, this is a longer-term goal, likely several years away from completion at the time of writing, and will require considerable investment. Traditional numerical simulations would provide a high-accuracy description of beam dynamics by modeling complex collective effects [6]. However, the substantial computational effort required for these simulations precludes their deployment as live diagnostics for machine operations. To overcome these difficulties, a machine-learning-based surrogate model was developed to provide fast high-fidelity longitudinal phase space predictions.

Virtual Diagnostics

Neural networks for particle accelerator can be used as virtual diagnostics, where learned surrogate models provide information about the state of an accelerator that would otherwise be difficult or even impossible to acquire using a physical diagnostic as in our case. The provided information can give useful data about experiments and help operators tune the accelerator. Previous work has demonstrated the feasibility of using ML models as virtual diagnostics to non-destructively predict the LPS distribution of FACET-II single bunch operation (in simulation) and at LCLS (in experiment) [7, 8]. Neural networks were trained in these studies to create a mapping between non-destructive diagnostic inputs (e.g., linac and e-beam diagnostics that are available on a single shot basis) and the LPS.

In this work, we present the ongoing work into building a ML-based surrogate model for reconstructing the two-dimensional LPS at the EuXFEL from RF settings and the CRISP spectrum. In consideration of the limited amount of LPS measurements available, we decided to base the training of such model on simulated data.

METHODS AND RESULTS

Dataset Generation

In order to effectively train a surrogate model, large scale simulated data is required. For this project, start-to-end simulations were used, tracking the electron beam from the gun down to the end of the collimation section (before the switchyard) including collective effects. For each simulation, the RF parameters for the injector and the first two linacs are randomly selected from a range of values around the working points, as listed in Table 1.

The first three parameters listed in Table 1 are the first-, second-, and third-order coefficient in beam energy profile at the injector exit, which define the detailed compression scenario in the first bunch compressor. The last two parameters, the L1 and L2 chirps, control the compression strengths in the second and third chicane sections. Preliminary studies were conducted in the past [9, 10], where the RF settings where adjusted with relatively small variations based on the nominal design. The aim of this project is to extend

* bianca.veglia@desy.de

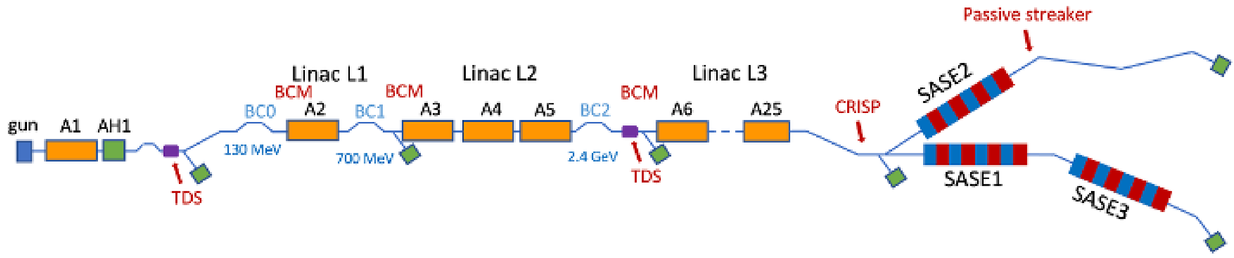

Figure 1: Schematic layout of the EuXFEL.

Table 1: The RF Parameter Input Features and the Values Ranges Used in Simulations for the Dataset Generation

Parameter (unit)	**lower bound**	**upper bound**
Injector chirp ($1/m$)	-10	-7
Injector curvature ($1/m^2$)	-100	+300
Injector cubic coefficient ($1/m^3$)	0	30000
L1 chirp ($1/m$)	-13	-5
L2 chirp ($1/m$)	-13	-5

the dataset to cover a wider range of parameters to create a robust control room tool for various experimental working points. The generation of this extensive dataset relied on start-to-end simulations, leveraging a robust and validated simulation toolkit. The details of this simulation model and its validation are presented in the following section.

Simulation Model

Ocelot [11, 12] is a Python-based toolkit for flexible FEL and storage ring multiphysics simulations. It provides a modular framework for beam dynamics simulations allowing for particle tracking, optics functions calculations, matching of Twiss parameters in a magnetic lattice by optimizing a set of variables. The Ocelot model of physical processes includes collective effects as space charge, coherent synchrotron radiation, wakefields and more.

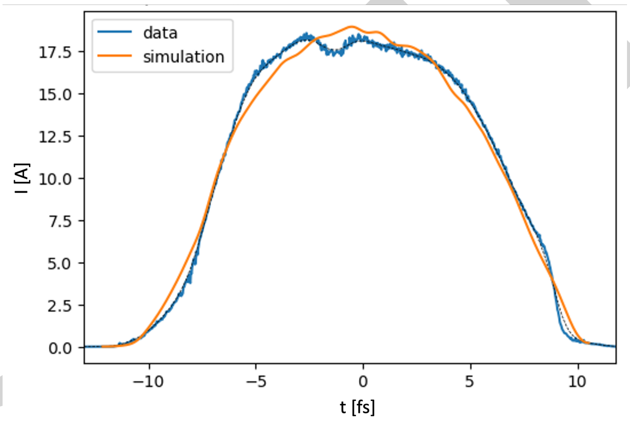

Figure 2: Comparison of simulations vs measured current profile at the injector TDS.

Our simulations and injector settings were tested against dedicated measurements in March last year shown in Fig. 2, where very good agreement was found. Following the successful validation of the simulation model, it was employed to generate the training data necessary for developing a neural network-based surrogate model.

Neural Network Architecture

We propose using a neural network machine learning model that is trained on data from a start-to-end beam dynamics simulation to combine scalar and spectral information in order to infer either the current profile or the two-dimensional LPS. To this end, we construct a model that receives two inputs: a 5D vector of RF settings and a 240-D THz formfactor sampled from 0.7 THz to 58 THz. These are then concatenated and passed through the NN.

The surrogate model is designed to reconstruct the 2D longitudinal density distribution from scalar inputs. The architecture employs a hybrid approach: a dense neural network first processes the RF settings and form factor data into a compact 512-dimensional latent space for feature representation. This representation is then expanded by a convolutional decoder, which uses upsampling layers to generate a high-resolution (300×300) image of the phase space. To ensure physical units, the network includes a secondary output head that predicts the energy and spatial scales (ranges) of the distribution. The model is trained in a supervised setup to minimise the difference between the reconstructed and the ground truth LPS images. The mean squared error loss function used here includes weight masking to selectively assign more importance to the non-zero regions of the image. An Adam [13] optimiser is used with a learning rate $\gamma = 0.001$. Figure 3 shows a few samples from the validation set, comparing ground truth and NN-predicted 2D LPS images. It can be observed that the model captures the shape of the beam distribution but fails to resolve the fine details, exhibiting some blurring in all the cases. This could be partially attributed to the inductive bias of the chosen convolutional architecture. The network's layers inherently prioritize the reconstruction of continuous, low-frequency features, which are more stable during the gradient descent process. Also the loss function might be responsible for smearing the intensity over a certain area to minimize the error. The model needs to be further optimized to reach the required level of accuracy to be deployed as a virtual diagnostics tool.

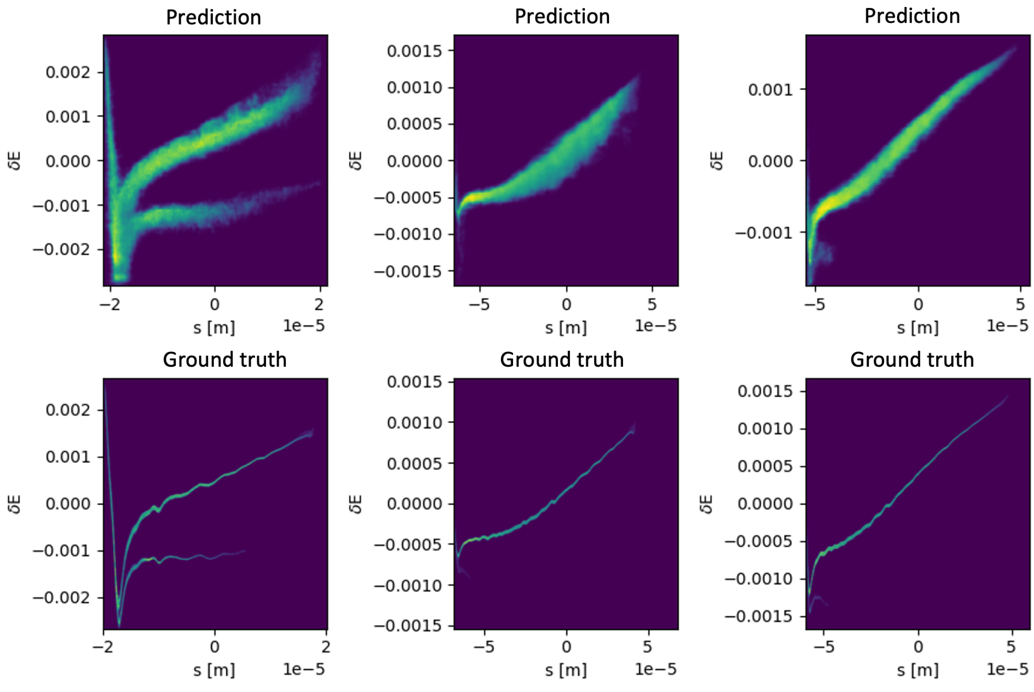

Figure 3: Three examples from the validation set comparing ground truth and reconstructed 2D LPS images.

Operational Applications

The developed model serves a dual-purpose for machine operations. As a *virtual diagnostic*, it takes real-time readbacks of RF settings and measured CRISP spectra to provide operators with an instantaneous visualization of the LPS.

Conversely, the model can also act as a *fast estimator*. By decoupling the model from live diagnostic inputs (using only the RF parameters), we can perform rapid "what-if" studies. While traditional OCELOT simulations of a single working point take several minutes, the surrogate model can provide predictions much faster. This enables the optimization of compression parameters to achieve specific peak currents or current distributions before applying settings to the actual machine.

CONCLUSIONS AND OUTLOOK

In this paper, we demonstrated the first steps toward a machine-learning-based virtual diagnostic for the longitudinal phase space at the European XFEL. By leveraging the Ocelot simulation toolkit, we generated a comprehensive dataset validated against experimental injector measurements. Our current surrogate model successfully reconstructs the overall 2D LPS geometry and density from non-destructive scalar and spectral inputs.

Although the model seams to be able to capture the beam's global compression state, further refinement is required to improve the accuracy of the predictions. To address the discrepancies observed ongoing work is testing new formulations of loss function and upsampling process. We are also investigating the use of generative adversarial networks or variational autoencoders to sharpen the reconstructed images and recover the slice details. The corrected model will then be integrated into the EuXFEL control system via a dedicated graphical user interface, enabling live comparisons between predicted and measured beam properties during facility operation.

This work lays the foundation for two primary tools intended for the EuXFEL control room. First, the virtual diagnostic tool will provide high-repetition-rate, non-destructive feedback on the beam distribution. Second, the fast estimator functionality will allow users to explore the compression parameter space in real-time, significantly accelerating the setup time for new experimental configurations.

ACKNOWLEDGEMENTS

The authors are thankful for the support through the Maxwell computational resources operated at DESY.

REFERENCES

[1] E. Volobuev *et al.*, "Transverse deflecting structure XFEL TDS INJ", *J. Phys.: Conf. Ser.*, vol. 747, p. 012083, Sep. 2016. doi:10.1088/1742-6596/747/1/012083

[2] S. Wesch, B. Schmidt, C. Behrens, H. Delsim-Hashemi, and P. Schmüser, "A multi-channel THz and infrared spectrometer for femtosecond electron bunch diagnostics by single-shot spectroscopy of coherent radiation", *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 665, pp. 40–47, 2011. doi:10.1016/j.nima.2011.11.037

[3] N. M. Lockmann, C. Gerth, B. Schmidt, and S. Wesch, "Noninvasive THz spectroscopy for bunch current profile reconstructions at MHz repetition rates", *Phys. Rev. Accel. Beams*, vol. 23, no. 11, p. 112801, Nov. 2020. doi:10.1103/PhysRevAccelBeams.23.112801

[4] S. Tomin, W. Decking, N. Golubeva, A. I. Novokshonov, T. Wohlenberg, and I. Zagorodnov, “Longitudinal Phase Space Diagnostics with Corrugated Structure at the European XFEL”, in *Proc. IPAC’22*, Bangkok, Thailand, pp. 275–278, Jul. 2022. doi:10.18429/JACoW-IPAC2022-MOPOPT020

[5] P. Dijkstal, W. Qin, and S. Tomin, “Longitudinal phase space diagnostics with a non-movable corrugated passive wakefield streaker”, *Phys. Rev. Accel. Beams*, vol. 27, no. 5, p. 050702, 2024. doi:10.1103/PhysRevAccelBeams.27.050702

[6] I. Zagorodnov, S. Tomin, Y. Chen, and F. Brinker, “Experimental validation of collective effects modeling at injector section of X-ray free-electron laser”, *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 995, p. 165111, 2021. doi:10.1016/j.nima.2021.165111

[7] C. Emma, A. Edelen, M. J. Hogan, B. O’Shea, G. White, and V. Yakimenko, “Machine learning-based longitudinal phase space prediction of particle accelerators”, *Phys. Rev. Accel. Beams*, vol. 21, no. 11, p. 112802, Nov. 2018. doi:10.1103/PhysRevAccelBeams.21.112802

[8] C. Emma, A. Edelen, A. Hanuka, B. O’Shea, and A. Scheinker, “Virtual Diagnostic Suite for Electron Beam Prediction and Control at FACET-II”, *Information (Basel)*, vol. 12, no. 2, 2021. doi:10.3390/info12020061

[9] J. Kaiser, A. Eichler, S. Tomin, and Z. Zhu, “Machine learning for combined scalar and spectral longitudinal phase space reconstruction”, in *Proc. IPAC’23*, Venice, Italy, pp. 4464–4467, Sep. 2023. doi:10.18429/JACoW-IPAC2023-THPL019

[10] S. T. Z. Zhu and J. Kaiser, “Application of machine learning in longitudinal phase space prediction at the european XFEL”, in *Proc. FEL2022*, Trieste, pp. 391–394, Aug. 2022. doi:10.18429/JACoW-FEL2022-WEP12

[11] S. Tomin, I. Agapov, M. Dohlus, and I. Zagorodnov, “OCELOT as a Framework for Beam Dynamics Simulations of X-Ray Sources”, in *Proc. IPAC’17*, Copenhagen, Denmark, 2017, pp. 2642–2645. doi:10.18429/JACoW-IPAC2017-WEPAB031

[12] I. Agapov, G. Geloni, S. Tomin, and I. Zagorodnov, “OCELOT: A software framework for synchrotron light source and FEL studies”, *Nucl. Instrum. Methods Phys. Res. A*, vol. 768, pp. 151–156, 2014. doi:10.1016/j.nima.2014.09.057

[13] D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization”, 2017, doi:10.48550/arXiv.1412.6980,