VIRTUAL BEAM POSITION MONITOR FOR PARTICLE DRIVEN PLASMA WAKEFIELD ACCELERATORS

J. Wolfenden[1]*, C. Welsch[1], University of Liverpool, Liverpool, United Kingdom
E. Belli[2], C. Pakuza, M. Turner
European Organization for Nuclear Research (CERN), Geneva, Switzerland
[1]Cockcroft Institute, Daresbury, United Kingdom
[2]University of Oxford, Oxford, United Kingdom

Abstract

Beam parameters at the interaction point (IP) of a particle driven wakefield accelerator (PWFA) are carefully controlled to optimise the output energy and beam quality of the accelerated witness beam. However, regardless of the parameters achieved, misalignments of beams at a PWFA IP have been shown to produce significant detrimental effects to the witness beam output energy and quality. Online monitoring of pointing stability, or beam jitter, at the IP is therefore essential to discern between beam quality optimisation needs and misalignment issues. However, it is not possible to use standard beam position monitors (BPMs) at these locations. Presented in this contribution is the demonstration of a proposed virtual beam position monitor (vBPM), applied to the IP of the AWAKE experiment at CERN that uses proton bunches as the driver. Upstream BPMs on the proton transfer line are used to infer the transverse position at the IP. A simulation study has been conducted to train a convolutional neural network (CNN) model to perform this inference. Dominant BPMs are highlighted, pointing towards areas of future optimisation to improve the proton beam jitter. An application to recent data collected at AWAKE is also presented, demonstrating the online reconstruction potential for future AWAKE runs.

INTRODUCTION

The alignment of a beam to an interaction point (IP) is often critical to the performance of the interaction. A clear example is the injection point in particle driven wakefield acceleration (PWFA), where incoming beam misalignment can significantly affect the acceleration process and the accelerated witness beam that is produced. This can include a drop in output energy [1], witness beam emittance growth [2], and hosing experienced by the driver [3]. Alignment is required both spatially and temporally, although this contribution focuses solely on transverse spatial alignment.

When the PWFA output beam is sub-optimal, transverse spatial misalignment is only one of several possible contributing factors, since the accelerated beam quality also depends on injection phase, beam loading, plasma-density evolution, beam matching and incoming beam conditions [4, 5]. However, it is often not possible to monitor transverse alignment directly at the PWFA injection point, since beam position monitors (BPMs) and screens cannot be placed there without interfering with operation. Therefore, this contribution develops a new non-invasive method to measure beam-IP alignment: a virtual beam position monitor (vBPM).

* joseph.wolfenden@cockcroft.ac.uk

VBPM MODEL DESIGN

A vBPM is a new virtual diagnostic (VD) [6, 7], whose inputs are upstream BPM readings, which are used to infer the transverse arrival position of the beam at the IP, operating completely non-invasively. A test case of the proton beam driver at the Advanced Wakefield Experiment (AWAKE) at CERN was selected to efficiently develop this model and methodology. At AWAKE, a 400 GeV SPS proton bunch propagates through a rubidium plasma, where it undergoes seeded self-modulation into a train of microbunches [8, 9]. These microbunches resonantly drive wakefields that have been used to accelerate externally injected witness electrons to GeV energies [5, 10]. The transport line delivering the 400 GeV proton beam from the SPS to AWAKE is by far the largest source of arrival position jitter at the injection point, with uncertainties, such as magnet power supplies, playing a large factor. In this initial proof of concept, all BPM readings along the transport line were included, enabling later permutation analysis to identify which BPMs have the largest effect on the output.

The model architecture chosen was a 1D convolutional neural network (CNN), as shown in Fig. 1. This was selected because it preserves the ordered sequence of the upstream BPM readings along the beamline, allowing the convolutional layers to learn local correlations between neighbouring BPM signals. The model takes the 18 pairs of upstream BPM readings, and passes these through three residual convolution blocks, which systematically increase the number of feature channels from 2 to 128. A residual convolutional block differs from a standard convolutional block by using a skip connection to combine the transformed block input with the output of the convolutional layers. This was used to help preserve information from the original BPM signals while allowing the network to learn deeper, higher-level feature transformations, reducing the risk of useful positional information being degraded through successive convolutional layers. The feature representation is then pooled and then passed through fully connected layers to infer the downstream transverse position, given by the output (X, Y) coordinates.

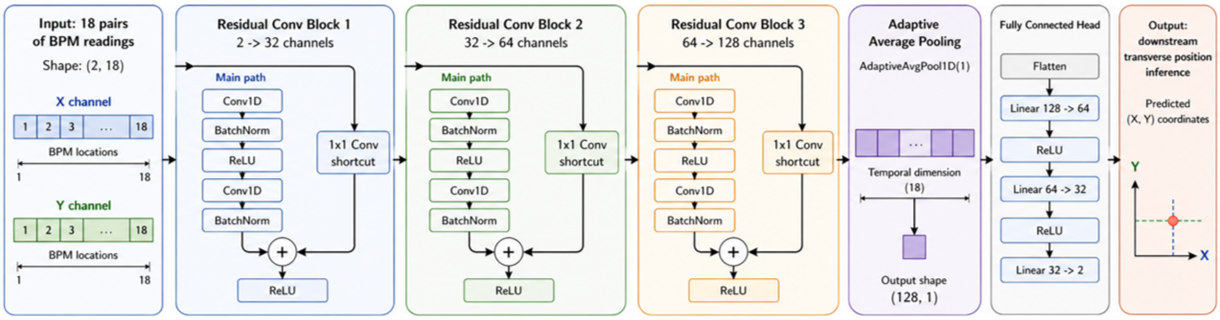

Figure 1: A schematic diagram representing the architecture, input, and output of the vBPM model.

DATA GENERATION AND TRAINING

The model was trained using data generated from MAD-X simulations of the proton transfer line. Trajectory jitter was introduced using thin kickers placed at eight locations along the beamline. These locations were selected to span the full transfer line, sample distinct optics regions and major lattice features, exploit BPM clusters where correlated trajectory information may be present, reduce degeneracy in the downstream position prediction, and allow horizontal and vertical jitter to be introduced independently.

For each simulated trajectory, the kicker strengths were sampled independently from zero-mean Gaussian distributions. The initial particle position and angle were also randomly sampled. The standard deviations used were 2 μrad for the kicker strengths, 100 μm for the initial transverse position, and 10 μrad for the initial angle. These values were chosen by inspecting the resulting downstream position distributions and selecting a range that extended beyond the experimentally observed variation, with the intention of supporting later transfer learning studies using experiment.

A total of 50,000 simulated trajectories were generated and split into training, validation and evaluation sets using a 70/15/15 split. An ensemble of eight models was then trained using the same dataset but different random initialisations. The final inferred downstream position was calculated as the mean ensemble prediction, while the standard error across the ensemble was used to provide an uncertainty estimate in both the horizontal and vertical coordinates.

To assess the robustness of the inferred downstream position to measurement uncertainty in the BPM inputs, additional versions of the ensemble were trained with noise applied to the input BPM readings. The injected noise was implemented as a random percentage perturbation to the BPM signals, with noise levels varied from 0% to 30% in 5% increments. For each noise level, a new ensemble of eight models was trained using the same underlying simulated trajectories. This allowed the effect of BPM input noise on the downstream position inference to be studied, and also provided a comparison between models trained with and without noise augmentation.

Each model in the ensemble was trained using the Adam optimiser with a learning rate of 10^{-3} and a mean squared error loss function. Training was monitored using validation loss, with learning rate reduced by a factor of 0.5 if validation loss plateaued, and early stopping applied when no further validation improvement was observed. Model checkpoints were saved during training, and the model state with the lowest validation loss was used for evaluation.

SIMULATION-BASED PERFORMANCE

The performance of the trained ensembles was first evaluated using the simulated evaluation data. In the absence of injected BPM noise, the predicted and true downstream positions showed close agreement, confirming that the CNN was able to learn the mapping from the upstream BPM readings to the downstream transverse position under ideal input conditions. The effect of BPM input noise was then studied by comparing ensembles trained with different levels of injected noise, the results of this are presented in Fig. 2. The mean absolute error in the prediction remained broadly unchanged up to 10% evaluation noise. Above this level, the ensemble performances began to separate, with a slightly larger spread observed at higher noise levels. Interestingly, the best overall performance was obtained from the ensemble trained with 15% injected noise, suggesting that moderate noise augmentation may improve generalisation by reducing sensitivity to small perturbations in the BPM inputs. Overall, regardless of the noise augmentation, the 1D CNN remained robust to injected BPM noise, suggesting that the model is making use of correlated features across the BPM sequence rather than simply relying on the absolute raw BPM values.

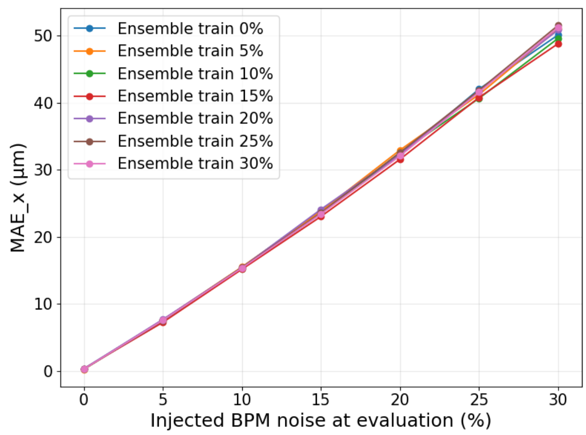

Figure 2: The effect of noise augmentation in evaluation for ensembles trained with different noise augmented data.

APPLICATION TO EXPERIMENT

The trained model was then applied to experimental proton transfer line data. The dataset consisted of approximately 250 beam shots, for which the beam was transported through the injection region. Two profile screens, located approximately 1 m upstream and 13 m downstream of the injection point, were used to reconstruct the proton trajectory and interpolate the transverse beam position at the injection point. The screen-derived positions were corrected into the BPM coordinate system by estimating the average offset between the screen and nearby BPM readings for each shot.

When the simulation-trained ensemble was applied directly to the corrected experimental data, the predictions showed poor agreement with the measured injection-point positions. Inspection of the upstream BPM data indicated a clear offset between the simulated and experimental BPM distributions, motivating the use of a simple transfer-learning approach. An affine calibration layer was inserted between the input BPM readings and the first residual convolutional block. The weights of the simulation-trained CNN were frozen, and only this calibration layer was trained using the experimental data, with the same 70/15/15 train, validation and evaluation split. The layer applied an independent linear scale and offset to each BPM input channel, giving 18 pairs of calibration parameters for each transverse plane, while preserving the mapping learnt from the simulation data.

The effect of this calibration step is shown in Fig. 3. After training the affine calibration layer, the agreement between the CNN prediction and the interpolated screen-based position improved substantially. The calibrated model achieved an accuracy of approximately 20 μm, demonstrating that the simulation-trained CNN could be transferred to experimental data using only a small number of trainable calibration parameters and a small number of experiment data points.

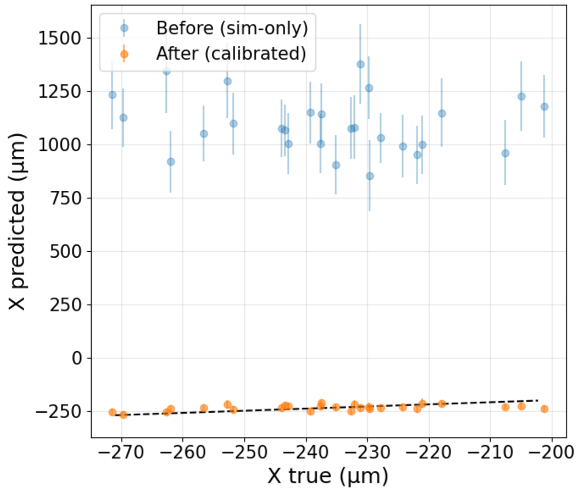

Figure 3: A comparison of model performance on experiment data with and without the affine calibration layer.

The calibrated CNN was also compared with a simple two-BPM extrapolation, using the final two BPMs upstream of the injection point. As shown in Fig. 4, the CNN showed significantly better agreement with the screen-derived injection position, while the two-BPM extrapolation showed substantially larger scatter. This indicates that the CNN is making use of information distributed across the upstream BPM sequence, rather than relying only on the final local trajectory measurement.

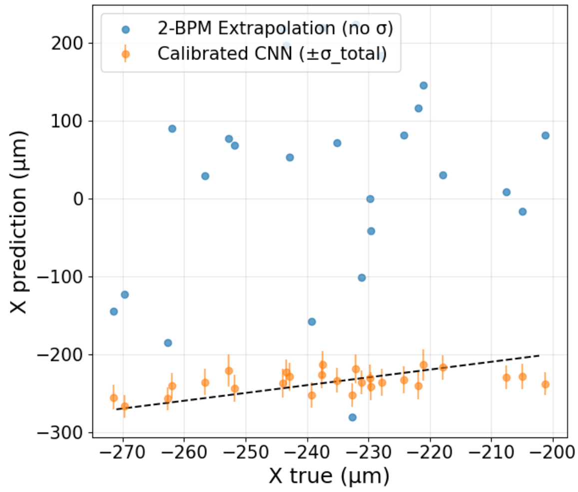

Figure 4: Comparison of model performance against a simple two-BPM extrapolation for experiment evaluation data.

CONCLUSION

This work demonstrates a proof-of-concept virtual diagnostic for inferring the downstream transverse proton position from upstream BPM readings using a 1D CNN. After transfer from MAD-X simulation data to experimental data using a simple affine calibration layer, the model achieved an accuracy of approximately 20 μm and outperformed a simple two-BPM extrapolation. This suggests that the CNN is able to exploit information distributed across the upstream BPM sequence, rather than relying only on the final local trajectory measurement.

Further work will focus on assessing the ensemble spread as an uncertainty estimate, testing the method over a wider operational range, and improving the overall accuracy towards the $< 10\ \mu$m level. Additional studies will include BPM-dependent noise modelling, permutation and ablation tests to identify the most important inputs, ensemble convergence studies, hyperparameter optimisation, and application to other beams, future runs, and comparable facilities.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon Europe Research and Innovation programme via the PACRI project under Grant Agreement No 101188004.

REFERENCES

[1] V. Lee *et al.*, "Precision alignment and tolerance of a plasma wakefield accelerator in a laser-ionized plasma source", *Phys. Rev. Accel. Beams*, vol. 29, no. 4, p. 041001, Apr. 2026. doi:10.1103/tgby-pksc

[2] L. Hildebrand *et al.*, “Beam realignment with emittance preservation in a plasma wakefield-accelerator stage”, *Phys. Rev. Lett.*, vol. 135, no. 19, p. 195002, Nov. 2025. doi:10.1103/cy5x-b81v

[3] M. Turner *et al.*, “Indirect self-modulation instability measurement concept for the awake proton beam”, *Nucl. Instrum. Methods Phys. Res. A*, vol. 829, pp. 314–317, Sep. 2016. doi:10.1016/j.nima.2016.01.060

[4] S. Schröder *et al.*, “High-resolution sampling of beam-driven plasma wakefields”, *Nat. Commun.*, vol. 11, no. 1, pp. 5984–, Nov. 2020. doi:10.1038/s41467-020-19811-9

[5] E. Gschwendtner *et al.*, “The AWAKE Run 2 Programme and Beyond”, *Symmetry*, vol. 14, no. 8, p. 1680, Aug. 2022. doi:10.3390/SYM14081680

[6] 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*, vol. 12, no. 2, p. 61, Jan. 2021. doi:10.3390/INFO12020061

[7] O. Convery, L. Smith, Y. Gal, and A. Hanuka, “Uncertainty quantification for virtual diagnostic of particle accelerators”, *Phys. Rev. Accel. Beams*, vol. 24, no. 7, p. 074602, Jul. 2021. doi:10.1103/PhysRevAccelBeams.24.074602

[8] E. Adli *et al.*, “Experimental observation of proton bunch modulation in a plasma at varying plasma densities”, *Phys. Rev. Lett.*, vol. 122, no. 5, p. 054802, Feb. 2019. doi:10.1103/PhysRevLett.122.054802

[9] M. Turner *et al.*, “Experimental observation of plasma wakefield growth driven by the seeded self-modulation of a proton bunch”, *Phys. Rev. Lett.*, vol. 122, no. 5, p. 054801, Feb. 2019. doi:10.1103/PhysRevLett.122.054801

[10] E. Adli *et al.*, “Acceleration of electrons in the plasma wakefield of a proton bunch”, *Nature*, vol. 561, no. 7723, pp. 363–367, Sep. 2018. doi:10.1038/s41586-018-0485-4