

# USING SURROGATE MODELS TO ASSIST ACCELERATOR TUNING AT ISIS

A. A. Saoulis\*, K. R. L Baker, H. V. Cavanagh, R. E. Williamson,  
ISIS Neutron and Muon Source, Rutherford Appleton Laboratory, U.K.  
S. Basak, J. Cha, J. Thiyyagalingam,  
Scientific Machine Learning, Rutherford Appleton Laboratory, U.K.

## Abstract

High intensity hadron accelerator performance is often dominated by the need to minimise and control beam losses. Operator efforts to tune the machine during live operation are often restricted to local parameter space searches, while existing physics-based simulations are generally too computationally expensive to aid tuning in real-time. To this end, Machine Learning-based surrogate models can be trained on data produced by physics-based simulations, and serve to produce fast, accurate predictions of key beam properties, such as beam phase and bunch shape over time. These models can be used as a virtual diagnostic tool to explore the parameter space of the accelerator in real-time, without making changes on the live machine. At the ISIS Neutron and Muon source, major beam losses in the synchrotron are caused by injection and longitudinal trapping processes, as well as high intensity effects. This paper describes the training and inference performance of a neural network surrogate model of the longitudinal beam dynamics in the ISIS synchrotron, from injection at 70 MeV to 800 MeV extraction, and evaluates the model's ability to assist accelerator tuning.

## INTRODUCTION

Machine learning (ML) has emerged as a valuable tool across many sub-disciplines within accelerator physics. Recent work has demonstrated its potential in control [1–3], tuning and optimisation [4–7], and virtual diagnostics [8–10]. Several of these recent advances rely on ML-based surrogate models [11, 12], which can offer an accurate substitute for traditional physics-based simulations with several orders of magnitude reduction in computation time. This paper will focus on training a parameter-to-image convolutional neural network (CNN) that takes the initial beam parameters and machine settings as input to reconstruct simulation-generated images representing key beam properties in the ISIS synchrotron.

At the ISIS Neutron and Muon source up to  $3 \times 10^{13}$  protons per pulse are accelerated from 70 MeV to 800 MeV by a 50 Hz rapid cycling synchrotron [13]. The majority of beam losses at the facility occur in the synchrotron, due to injection and longitudinal trapping processes as well as high intensity effects. Since operators must rely on feedback from the live machine during tuning, they are restricted to making small, incremental changes to ensure important operational constraints such as low loss are always satisfied.

This approach can be time-consuming for operators, and such a restricted parameter space search is likely to yield sub-optimal local minima for a given optimisation problem.

A method for searching the parameter space without such constraints is clearly desirable, and physics-based simulations can be leveraged to address this. However, since these simulations are generally too computationally expensive to aid tuning in real-time, fast-executing surrogate models may offer an alternative that allows for real-time use in the control room to aid machine tuning.

## Modelling the ISIS Synchrotron

The accelerator physics group at ISIS has developed a C++ turn-by-turn physics-based simulation of the longitudinal dynamics of the ISIS synchrotron [14]. The simulation takes in a wide range of inputs that define properties of the injected beam as well as how the RF settings vary over the 10 ms injection cycle, and outputs several bunch properties at each turn, including longitudinal bunch charge distribution and emittances.

One key measurement that is frequently used as a diagnostic during tuning is an image representing the evolution of the longitudinal charge density of a bunch in the synchrotron over time, see Fig. 1. Each row represents the longitudinal charge distribution of the bunch at a particular turn. This is referred to as a “waterfall plot” by ISIS operators, and can be used to quickly diagnose the state of the synchrotron by identifying common unstable modes such as oscillations in the mean phase and root mean square bunch length. The CNN models explored in this paper were trained to reconstruct these waterfall plots from simulation data.

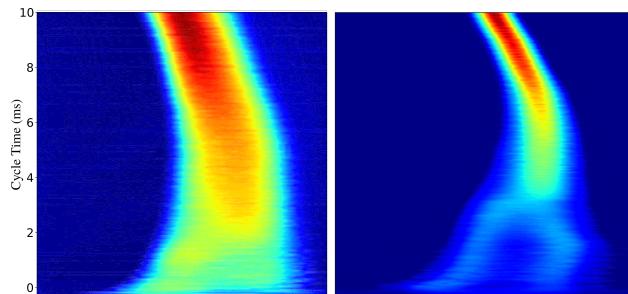


Figure 1: Left: Waterfall plot acquisition from a single beam position monitor in the ISIS synchrotron. Right: Simulation-generated waterfall plot. The bunch charge distribution evolves from injection at the bottom of the plot to extraction at the top over the 10 ms acceleration cycle.

\* alex.saoulis@stfc.ac.uk

## DATASET GENERATION

The physics-based simulation takes in a set of 5 scalar parameters, representing important beam properties including injection energy mismatch, intensity, and space-charge effects, as well as three time series as input. The time series define how the 1RF gap volts, the 2RF gap volts, and the 2RF phase difference vary over the 10 ms acceleration cycle. Each time series was interpolated at a fixed set of 17 points across the 10 ms cycle, leading to  $3 \times 17$  time series inputs. This gave a total of 56 inputs to be fed into the surrogate model. The scalar parameters were randomly sampled within reasonable ranges defined by the accelerator physics experts. Sampling of the time series inputs required more care, and proved to be of key importance in producing a useful surrogate model.

The last three years of historical machine settings were analysed, over which the machine settings were generally stable. The mean and ranges of the machine settings were calculated, and the sampling strategy was adapted to ensure that the entire range was covered over many samples.

A scheme of adding the mean historical settings to a combination of element-wise Gaussian noise and continuous random variations was adopted, see Fig. 2. The continuous random variations introduced local correlations in the time series, preventing the large losses and instabilities that come with mismatched RF buckets in the synchrotron, while avoiding global correlations that would distort the model's learning process. The random noise helped to mitigate the problem of the surrogate model over-relying on local correlations in its inputs to make predictions.

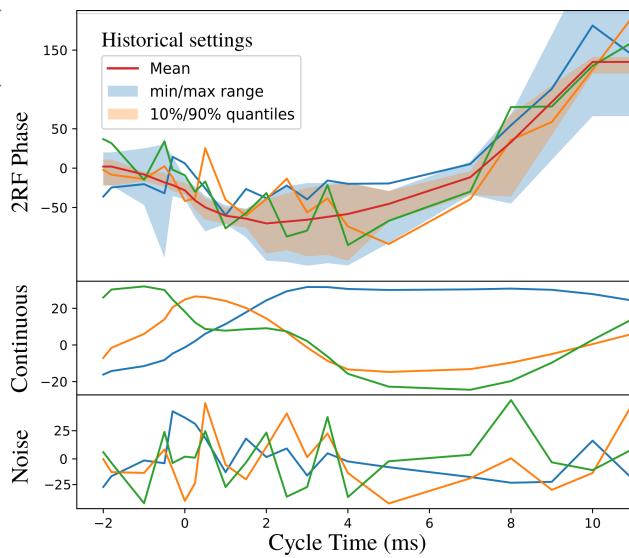


Figure 2: Three example synthetic time series inputs (solid orange, blue, and green lines), overlaid onto the historical machine settings statistics.

This final sampling strategy was used to generate 15000 input-image pairs, generated by running the physics simulations in an embarrassingly parallel fashion across 100 CPUs. The turn-by-turn longitudinal charge density was

downsampled and interpolated to generate a  $200 \times 100$  pixel image representing the waterfall plot output of each simulation. 11000 points were used for training the model, and the remaining 4000 points were split equally into validation and test sets, used for hyperparameter tuning and model evaluation respectively.

## MODEL TRAINING

Initially, a simple parameter-to-image CNN architecture was experimented with and underwent hyperparameter tuning. The model was trained to minimise the  $L_2$  pixel-wise error between the "ground truth" simulation-generated waterfall plot and the surrogate model generated reconstruction. All model training and evaluation was done using TensorFlow [15] and Keras [16], using the default Adam optimiser. The final architecture employed, given in Fig. 3, achieved low reconstruction loss and successfully reproduced most key features in the waterfall plots.

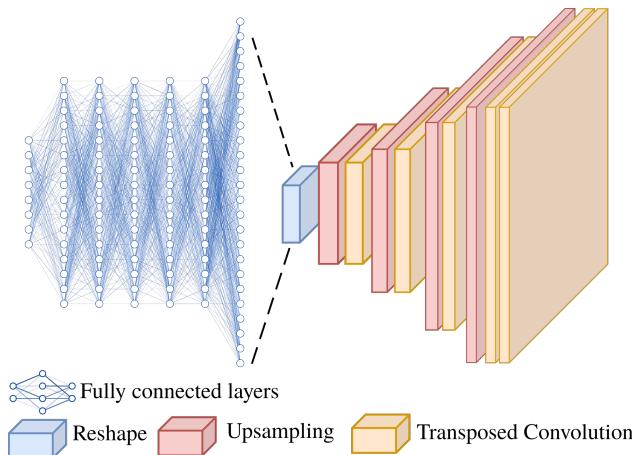


Figure 3: The base model consisted of 5 fully-connected layers with 256 nodes each, followed by 4 blocks of upsampling, batch normalisation, and transposed convolutional layers. ReLU activations used everywhere.

One issue that became apparent, however, was the model's inability to reproduce high frequency oscillations in the mean position of the bunch over time. These oscillations can be caused by a mismatch between the incoming beam and the RF bucket, and may be indicative of growing instabilities in the bunch. Accurately reproducing these oscillations in the surrogate model was therefore important from an operational point of view. Many different standard loss functions and architectures were explored to achieve this.

Some success was found using a GAN architecture (by appending a discriminator to the base model) that adversarially learnt features in the waterfall plots [20]. Finally, a custom loss term that explicitly compared the amplitude of oscillation over time between the ground truth waterfall plot and the reconstruction was added to the loss function. This greatly improved reproduction of these oscillations in the ML-generated images. Results are shown in Table 1, with some example reconstructions given in Fig. 4.

Table 1: Performance metrics of the investigated architectures and loss functions on the hold-out test set. None of the modifications significantly improved  $L_2$  pixel loss. Only the custom loss model and the GAN architecture, which was trained adversarially to learn features beyond pixels, significantly improved reproduction of mean phase oscillations.

| Performance Metric                    | Base Model | $L_1$ Pixel Loss [17] | ResBlock Decoder [18] | cVAE [19] | GAN [20]     | Custom Loss  |
|---------------------------------------|------------|-----------------------|-----------------------|-----------|--------------|--------------|
| $L_2$ Pixel Loss ( $\times 10^{-5}$ ) | 6.36       | 5.91                  | 6.07                  | 6.03      | 8.67         | 6.25         |
| Oscillation Loss                      | 0.128      | 0.110                 | 0.107                 | 0.132     | <b>0.075</b> | <b>0.063</b> |

Figure 4: Two example simulation-generated waterfall plots alongside the reconstructions of the different models. Mean phase oscillations are identified by the high frequency horizontal spikes along the plots, which many models fail to reproduce.

All models that were only trained on a pixel-wise loss function failed to reproduce the oscillations consistently, irrespective of the complexity of the architecture employed. Predicting the exact turn numbers (i.e. the exact row on the waterfall plot) of the peaks and troughs of the oscillations is an extremely difficult problem without relying on the numerical integration of a physics-based model; the surrogate model likely learned to smooth over the oscillations since it could not predict the exact pixels. An oscillation-dependent loss term provided a training signal that taught the surrogate model to mimic the amplitude of the mean phase oscillations, rather than reproduce them pixel-wise. This behaviour was sufficient as operators are primarily interested in qualitative features of the oscillations, such as amplitude and rough position. This approach could be extended to other features that a simple pixel-wise loss function fails to fully capture.

## EXPERIMENTAL DEPLOYMENT

An attempt to evaluate the performance of a zero-shot transfer from simulation to live operation was made in the last ISIS operating cycle. A GUI was developed to allow for simple interfacing between the control system and the model. Unfortunately, due to operational issues at ISIS stemming from several recent upgrades, the synchrotron RF settings were undergoing large, frequent changes in the run-up to deployment. It was therefore not possible to train a model over the live RF settings in time for deployment. Once machine settings at ISIS stabilise, it is hoped that a proper evaluation of the surrogate model can be undertaken.

## CONCLUSION

Once an efficient strategy for sampling the large input space of this problem was developed, it was straightforward to develop an ML-based surrogate with a simple architecture that could accurately reconstruct most important features in the target waterfall plots. In addition, hand-crafted loss terms provide an avenue to reconstructing qualitative features that are important to operators where pixel-wise loss does not suffice. The rapid-executing surrogate shows potential to aid machine tuning by providing operators with real-time feedback during parameter space searches.

The difficulties encountered in evaluating the model on the live machine highlight a serious shortcoming of ML-based surrogate models: such surrogate models are only useful insofar as stable operating conditions can be guaranteed.

## ACKNOWLEDGEMENTS

The authors are grateful to the Scientific Machine Learning group (SciML) in the Scientific Computing Department (SCD) at STFC, and the Synchrotron Group at ISIS, for the invaluable guidance and technical expertise provided throughout this project. Computing resources were provided by SciML and SCD through the PEARL and in-house OpenStack cloud clusters.

## REFERENCES

- [1] A. L. Edelen, S. G. Biedron, B. E. Chase, D. Edstrom, S. V. Milton, and P. Stabile, “Neural networks for modeling and control of particle accelerators,” *IEEE Transactions*

- [1] Content from this work may be used under the terms of the CC BY 4.0 licence (© 2022). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI
- on Nuclear Science, vol. 63, no. 2, pp. 878–897, 2016. doi: 10.1109/TNS.2016.2543203
- [2] A. Scheinker, “Adaptive machine learning for robust diagnostics and control of time-varying particle accelerator components and beams,” *Information*, vol. 12, no. 4, 2021. doi: 10.3390/info12040161
- [3] A. Scheinker, F. Cropp, S. Paiagua, and D. Filippetto, “An adaptive approach to machine learning for compact particle accelerators,” *Scientific reports*, vol. 11, no. 1, pp. 1–11, 2021. doi: 10.1038/s41598-021-98785-0
- [4] G. Azzopardi, A. Muscat, S. Redaelli, B. Salvachua, and G. Valentino, “Operational Results of LHC Collimator Alignment Using Machine Learning,” in *Proc. IPAC’19*, Melbourne, Australia, May 2019, pp. 1208–1211. doi: 10.18429/JACoW-IPAC2019-TUZZPLM1
- [5] J. Duris *et al.*, “Bayesian optimization of a free-electron laser,” *Phys. Rev. Lett.*, vol. 124, p. 124801, 2020. doi: 10.1103/PhysRevLett.124.124801
- [6] A. Hanuka *et al.*, “Physics model-informed gaussian process for online optimization of particle accelerators,” *Phys. Rev. Accel. Beams*, vol. 24, p. 072 802, 7 2021. doi:10.1103/PhysRevAccelBeams.24.072802
- [7] R. Roussel, A. Hanuka, and A. Edelen, “Multiobjective bayesian optimization for online accelerator tuning,” *Phys. Rev. Accel. Beams*, vol. 24, p. 062 801, 6 2021. doi:10.1103/PhysRevAccelBeams.24.062801
- [8] L. Gupta, A. Edelen, N. Neveu, A. Mishra, C. Mayes, and Y.-K. Kim, “Improving surrogate model accuracy for the LCLS-II injector frontend using convolutional neural networks and transfer learning,” *Machine Learning: Science and Technology*, vol. 2, no. 4, p. 045 025, 2021. doi:10.1088/2632-2153/ac27ff
- [9] 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, p. 112 802, 2018. doi:10.1103/PhysRevAccelBeams.21.112802
- [10] O. Convery, L. Smith, Y. Gal, and A. Hanuka, “Uncertainty quantification for virtual diagnostic of particle accelerators,” *Phys. Rev. Accel. Beams*, vol. 24, p. 074 602, 2021. doi: 10.1103/PhysRevAccelBeams.24.074602
- [11] A. Edelen, N. Neveu, M. Frey, Y. Huber, C. Mayes, and A. Adelmann, “Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems,” *Phys. Rev. Accel. Beams*, vol. 23, p. 044 601, 2020. doi:10.1103/PhysRevAccelBeams.23.044601
- [12] A. E. Pollard, D. J. Dunning, and M. Maheshwari, “Learning to Lase: Machine Learning Prediction of FEL Beam Properties,” Shanghai, China, Oct. 2021, presented at ICALEPS’21, Shanghai, China, Oct. 2021. paper WEPV020, unpublished.
- [13] D. Findlay, J. Thomason, S. Fletcher, R. de Laune, and E. Cooper, *A Practical Guide to the ISIS Neutron and Muon Source*, Accessed: 01-06-2022. <https://www.isis.stfc.ac.uk/Pages/A%20Practical%20Guide%20to%20the%20ISIS%20Neutron%20and%20Muon%20Source.pdf>
- [14] R. E. Williamson, B. G. Pine, and C. M. Warsop, “Longitudinal Dynamics Studies for ISIS Upgrades,” in *Proc. PAC’09*, Vancouver, Canada, May 2009, pp. 4411–4413. <https://jacow.org/PAC2009/papers/FR5PFP046.pdf>
- [15] Martín Abadi *et al.*, *TensorFlow: Large-scale machine learning on heterogeneous systems*, Software available from tensorflow.org, 2015, <http://tensorflow.org/>
- [16] F. Chollet *et al.* “Keras,” 2015. <https://github.com/fchollet/keras>
- [17] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, “Loss functions for image restoration with neural networks,” *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 47–57, 2017. doi:10.1109/TCI.2016.2644865
- [18] L. Cai, H. Gao, and S. Ji, “Multi-stage variational auto-encoders for coarse-to-fine image generation,” in *Proceedings of the 2019 SIAM International Conference on Data Mining (SDM)*, pp. 630–638. doi:10.1137/1.9781611975673.71
- [19] K. Sohn, H. Lee, and X. Yan, “Learning structured output representation using deep conditional generative models,” in *Advances in Neural Information Processing Systems*, vol. 28, 2015. <https://proceedings.neurips.cc/paper/2015/file/8d55a249e6baa5c06772297520da2051-Paper.pdf>
- [20] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, “Autoencoding beyond pixels using a learned similarity metric,” in *Proceedings of The 33rd International Conference on Machine Learning*, vol. 48, 2016, pp. 1558–1566. <https://proceedings.mlr.press/v48/larsen16.html>