

MODELLING AND ANOMALY DETECTION FOR DIAMOND-II BEAM LOSS MONITORS USING MACHINE LEARNING TECHNIQUES

C. Lehmann^{*1,2}, L. Bobb², P. N. Burrows¹

¹John Adams Institute, University of Oxford, UK

²Diamond Light Source, Oxfordshire, UK

Abstract

The losses measured by Beam Loss Monitors (BLMs) at synchrotron light source facilities offer useful but indirect insight into the state of the beam. Frequent changes to insertion devices, beam current, temperature, humidity, and other factors make the line between normal operation and anomalous behaviour unclear. Data was collected passively during user beam on the storage ring of Diamond Light Source over several months, as well as actively during dedicated machine shifts. A variety of machine learning techniques were employed to model the behaviour in this data, enabling anomaly detection and aiding machine operations.

INTRODUCTION

New Libera BLMs [1] have been installed and tested at Diamond Light Source in preparation for the upcoming Diamond-II upgrade [1, 2]. The new BLMs are scintillator-based and capable of measuring losses in both the fast regime (125 MHz, approaching bunch-by-bunch) and slow regimes (5 Hz). Similar installations exist at both Soleil [3] and ESRF [4]. The BLM data is commonly used to identify high loss locations and narrow down potential causes of faults [5]. With the new BLM hardware, capabilities are expanded with potential to apply machine learning (ML) and anomaly detection techniques to analyse both the fast and slow loss regimes. This ML-assisted tool can assist operators to quickly identify anomalous storage ring injections and abnormal machine states during user beam, prompting corrective action that could prevent machine faults.

The effectiveness of the anomaly detection techniques depend on the accuracy with which normal beam losses can be modelled. Unlike colliders such as the LHC [6] which have quite stable beam conditions, synchrotron light sources have a constantly changing beam state primarily due to beamlines adjusting the gap size or phase of their Insertion Devices (IDs). Feed-forward tables [7] and feedback systems [8] are used to minimise these effects but are imperfect, resulting in dynamically changing beam losses and lifetime. By modelling these losses using machine learning techniques, anomaly detection algorithms can more confidently detect abnormalities that may have severe consequences for accelerator operations. This paper explores the development of these techniques specifically for losses in the slow regime.

* corey.lehmann@physics.ox.ac.uk

EXPERIMENTAL SETUP

Diamond Light Source is a synchrotron light source facility which maintains a beam of 3 GeV electrons in the Storage Ring to provide synchrotron radiation for users at the beamlines. The Diamond-II upgrade is planned to begin December 2027 [2], and includes installation of over 100 Beam Loss Detectors (BLDs) and a corresponding 25 BLMs from I-Tech throughout the facility [1].

In preparation for this upgrade, 11 BLDs across 4 BLMs have been deployed in the Diamond storage ring for testing. The locations of these BLDs are shown in Fig. 1. All BLDs have been calibrated using LED methods similar to those used at Soleil [3].

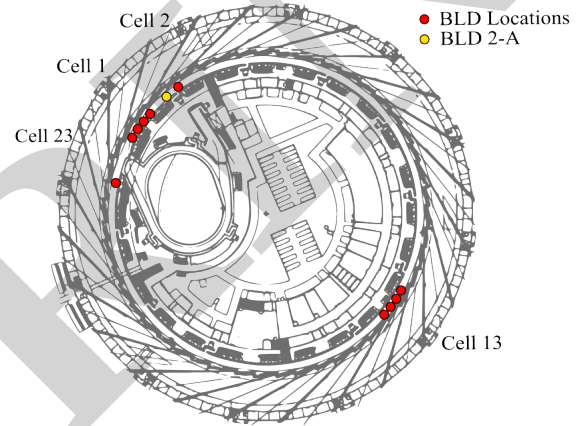


Figure 1: Positions of BLDs for initial testing in the Diamond Light Source storage ring, modified from [9]. BLD 2-A (the first BLD in DDBA Cell 2) is shown in yellow.

Each BLM offers several digitised data streams for each BLD. Testing has shown the best signals are the counts per second (cps), triggered by a static threshold on the differential data, using the 50 Ω (Fast) input impedance. Average cps for each BLD are calculated at a 5 Hz rate and stored in the Diamond Archiver [10].

DATASET PREPARATION

The list of PV categories used as input parameters to predict the BLD losses is shown in Table 1. Data was collected over a 7 month period between 05/07/2025 and 28/02/2026. Data with the machine in abnormal conditions such as low current, machine top-up or during Machine Studies were removed. The total after data cleaning amounts to over 1.9 million samples at 5 second intervals, equivalent to 110 days of user beam for each of the 73 input and 11 output parameters.

Table 1: PV Categories used as Input Variables, and their Quantities

PV Category	Number of PVs
ID Gaps	28
ID Phases	22
Beam State (eg Current, Emittance)	18
Feedback States	7
Machine Tune	6
Temperature, Humidity & Pressure	3
Beam Lifetime	2

For training and evaluation of the model, the 110 days of data were split using a cyclic partition in a ratio of 2:1:1 for training, validation and test datasets. Models use the training data to learn, the evaluation data to stop training early preventing over-fitting, and the test data only for the final evaluation [11].

MODEL EVALUATION

Model accuracy was evaluated using the Pearson residual, defined as:

$$r_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}}, \quad (1)$$

where r_i^P is the residual for BLD datastream i , y_i is the predicted count rate, and $\hat{\mu}_i$ is the measured count rate [12]. The varying sensitivity of the BLDs to anomalies leads to a distribution of residuals across the BLDs spanning an order of magnitude for many of the models. To incentivise the training to optimise more evenly across all data streams, the geometric mean of the residuals was used as the loss metric [11].

MODEL COMPARISON

A variety of model architectures and some variations were trained and evaluated over the datasets, with the results of the best model from each category shown in Table 2.

Table 2: Comparison of the Best Models for a Variety of Model Architectures using the Pearson Residual [11]

Model Category	Best Model	Train	Test
Random Forest [11]	Default	3.68	8.03
XGBoost [13]	Default	3.83	7.77
Neural Network	1x512	1.52	4.39
Partially Linear NN [14]	1x1024	1.47	4.14
Bayesian NN [15]	Default	8.64	11.92
Gaussian Process [16]	Default	3.49	6.98

The 1x1024 Partially Linear Neural Network (NN) [14] is shown to achieve the best score on the Test datasets with a Pearson residual of 4.14, and will be described in depth. This two-stage model uses a combination of a simple linear neural network as a base, with a more complex neural network used to correct the residual errors. This ensures the model defaults to the linear relationships when predicting data outside of its

trained state-space, where a standard neural network could predict more extreme values due to being unconstrained. The 1x1024 notation indicates 1 layers of 1024 nodes in the residual neural network. A dropout of 0.3 was used to add regularisation [11]. An example of this model on a day of test data is shown in Fig. 2.

Within the Partially Linear NN category, the optimal dimensions were found using a grid search iterating over number of layers from 0 to 6, and number of nodes per layer in powers of 2 from 16 to 1024. Similar hyper-parameter searches were conducted for the other architectures, although to a lesser extent.

ANOMALY DETECTION

With the Partially Linear NN 1x1024 chosen as the best current model based on its low Pearson residual in the test dataset, this can now be used for anomaly detection. A Pearson residual of 5 was chosen as the threshold for anomaly detection. This anomaly metric can be visualised over the test dataset in Fig. 3, representing how local and global beam loss anomalies can be detected for operators during user beam. Local anomalies can indicate an additional source of losses nearby, while global anomalies can indicate abnormal operating conditions which affect the entire storage ring.

Since the dataset is unlabelled, the correctness of these results is difficult to quantify especially since the nature of this method means the cause of the anomaly is unknown. As an alternative, a set of known anomalies was created during a machine shift to measure the sensitivity and confidence of these methods.

Anomalies of interest are ones which have occurred during operation before, can be easily missed by an operator, will have consequences in beam degradation or potential beam loss, and have a signature expected to be detectable by the BLMs. The disturbances investigated for demonstration are the x Pinger Magnet being left on, the y Pinger Magnet being left on, and the Multi-Bunch Feedback (MBF) with extra excitation from the Numerically Controlled Oscillator (NCO) being left on. Figure 4 shows each of these disturbances as measured by BLD 2-A, with detected anomalies shown in red. All data was taken with the vertical emittance feedback and orbit feedback off. For the MBF measurements, NCO1 was set to the betatron tune frequency with standard settings.

CONCLUSION

The use of BLMs as a beam diagnostic tool was investigated. A set of machine learning models were trained and evaluated over the test dataset, with the best model being the Partially Linear NN 1x1024 with a Pearson residual of 4.14 over the test dataset allowing for confident anomaly detection. Using this model, some recreated anomalies from the x Pinger magnets, y Pinger magnets and MBF were able to be correctly detected with good confidence.

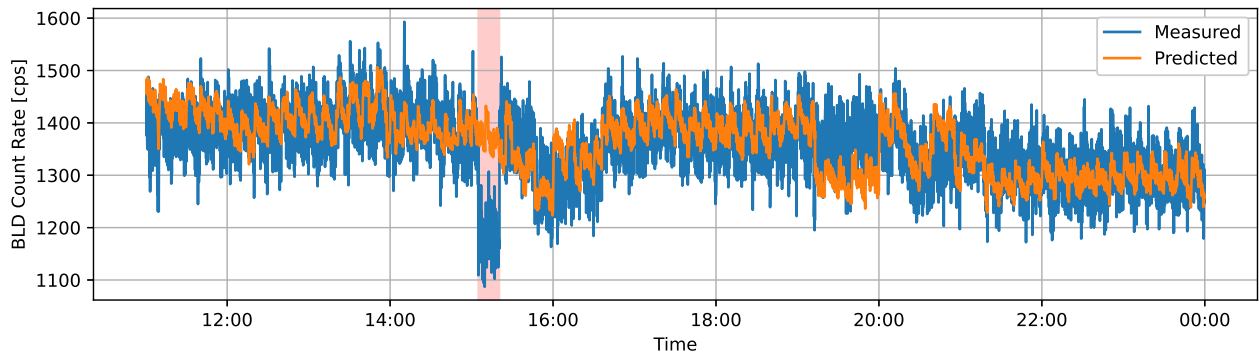


Figure 2: Example of a 13-hour period of Test data for BLD 2-A on 28/01/26, comparing the true measured values (Blue) with the values predicted by the 1x1024 Partially Linear NN model (Orange). The Red shaded areas indicate automatically detected anomalies, based on a Pearson residual threshold of 5. The plotted data has a Pearson Residual of 1.187.

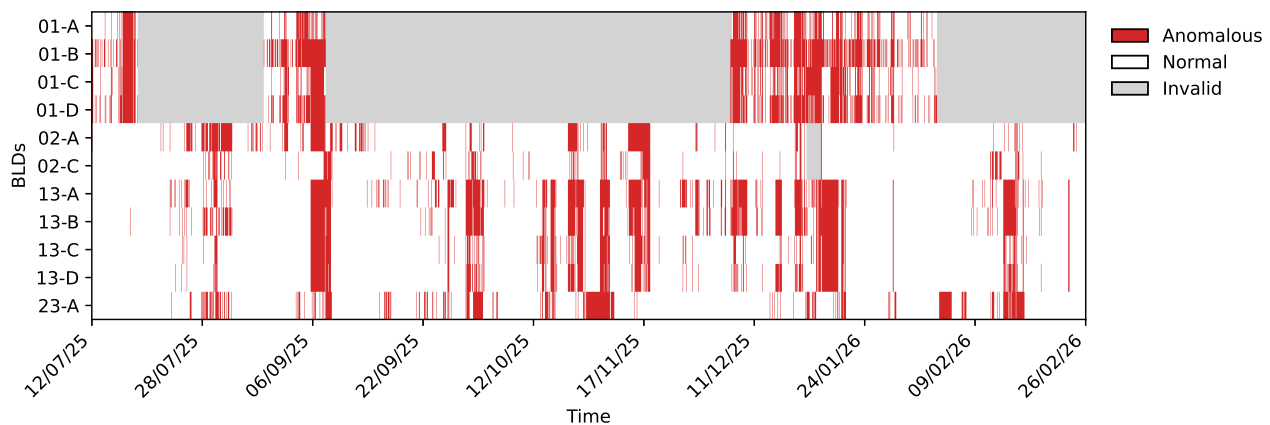


Figure 3: Visualisation of anomalies detected over 6 months from the Test dataset. The anomaly threshold for the Pearson residual was set to 5. Anomalies can be further classified as either global (simultaneously detected by most BLDs around the storage ring) or local (detected by a single BLD or cell-based grouping).

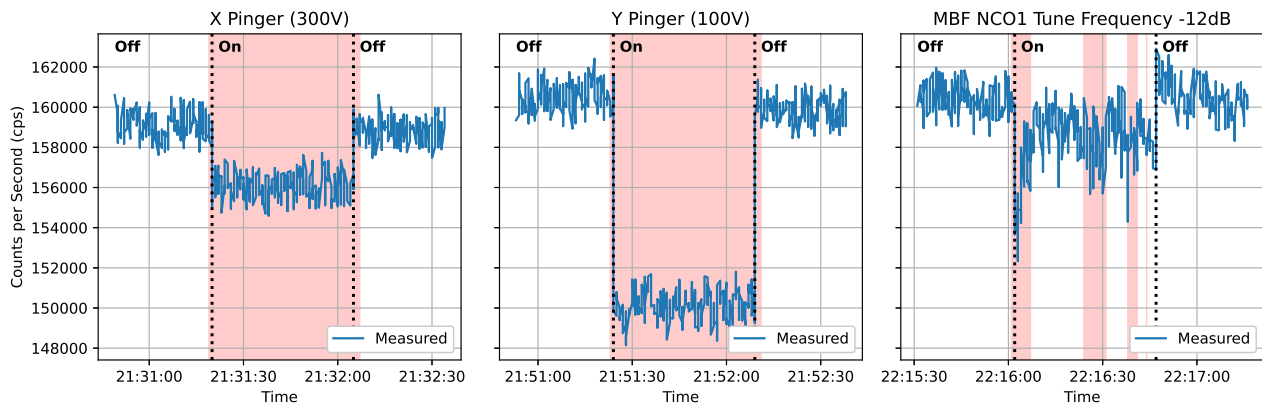


Figure 4: Examples of raw data from BLD 1-A showing the effects on the loss rate of turning on and off various disturbances. The shaded red sections show where the anomaly detection indicates an anomaly for a correctly predicting model.

REFERENCES

- [1] Instrumentation Technologies, Libera BLM, 2024, <https://www.i-tech.si/products/libera-blm/>
- [2] “Diamond-II technical design report”, Aug. 2022. <https://www.diamond.ac.uk/Home/News/LatestNews/2022/14-10-22.html>
- [3] N. Hubert, M. El Ajjouri, and D. Pédeau, “New Beam Loss Monitor System at SOLEIL”, in *Proc. IBIC’19*, Malmö, Sweden, Sep. 2019, pp. 118–121. [doi:10.18429/JACoW-IBIC2019-MOPP017](https://doi.org/10.18429/JACoW-IBIC2019-MOPP017)
- [4] L. Torino and K. B. Scheidt, “New Beam Loss Detector System for EBS-ESRF”, in *Proc. IBIC’18*, Shanghai, China, Sep.

- 2018, pp. 346–352.
doi:10.18429/JACoW-IBIC2018-WE0B01
- [5] K. Wittenburg, “Beam loss monitors”, 2018. <https://bib-pubdb1.desy.de/record/440873/>
- [6] The Large Hadron Collider, <https://home.cern/science/accelerators/large-hadron-collider/>
- [7] I. P. S. Martin *et al.*, “Active Optics Stabilisation Measures at the Diamond Storage Ring”, in *Proc. IPAC'14*, Dresden, Germany, Jun. 2014, pp. 1760–1762.
doi:10.18429/JACoW-IPAC2014-TUPRI082
- [8] I. S. Uzun, R. Bartolini, G. Rehm, J. A. Dobbing, M. T. Heron, and J. Rowland, “Initial Design of the Fast Orbit Feedback System for DIAMOND Light Source”, in *Proc. ICALEPCS'05*, Geneva, Switzerland, Oct. 2005, paper P3_030.
- [9] R. P. Walker, “The diamond light source”, 2004. <https://cas.web.cern.ch/sites/default/files/lectures/warrington-2004/walker.pdf>
- [10] M. G. Abbott, G. Rehm, and I. Uzun, “A New Fast Data Logger and Viewer at Diamond: the FA Archiver”, in *Proc. ICALEPCS'11*, Grenoble, France, Oct. 2011, paper THCHMUST03, pp. 1244–1247. <https://jacow.org/icaleps2011/papers/THCHMUST03.pdf>
- [11] K. P. Murphy, *Probabilistic Machine Learning: An Introduction*. MIT Press, 2022. <http://probml.github.io/book1>
- [12] A. C. Cameron and P. K. Trivedi, *Regression Analysis of Count Data*. Cambridge University Press, 2013.
- [13] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System”, in *Proc. KDD'16*, San Francisco, CA, USA, Aug. 2016, pp. 785–794.
doi:10.1145/2939672.2939785
- [14] T. A. Severini and J. G. Staniswalis, “Quasi-likelihood estimation in semiparametric models”, *J. Am. Stat. Assoc.*, vol. 89, no. 426, pp. 501–511, 1994.
doi:10.1080/01621459.1994.10476774
- [15] Y. Gal and Z. Ghahramani, “Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning”, in *Proc. ICML'16*, New York, NY, USA, Jun. 2016, pp. 1050–1059. <https://proceedings.mlr.press/v48/gal16.html>
- [16] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.