

# MACHINE LEARNING APPLIED AT THE LHC FOR BEAM LOSS PATTERN CLASSIFICATION

G. Valentino\*, University of Malta, Msida, Malta  
 B. Salvachua, CERN, Geneva, Switzerland

## Abstract

Beam losses at the LHC are constantly monitored because they can heavily impact the performance of the machine. One of the highest risks is to quench the LHC superconducting magnets in the presence of losses leading to a long machine downtime to recover cryogenic conditions. Smaller losses are more likely to occur and have an impact on the machine performance, reducing the luminosity production or reducing the lifetime of accelerator systems due to radiation effects, such as magnets. Understanding the characteristics of the beam loss, such as the beam and the plane, is crucial to correct them. Regularly during the year, dedicated loss map measurements are performed to validate the beam halo cleaning of the collimation system. These loss maps have the particular advantage that they are performed in well controlled conditions and can therefore be used by a machine learning algorithm to classify the type of losses during the LHC machine cycle. This study shows the result of the beam loss classification and its retrospective application to beam loss data from the 2017 run.

## INTRODUCTION

A multi-stage collimation system [1] protects the LHC from normal and abnormal beam losses. To monitor beam losses, the LHC is equipped with a Beam Loss Monitoring (BLM) system [2], with around 3600 ionization chambers distributed around the ~27 km length of the accelerator. The collimator jaw positions for various stages in the LHC machine cycle are determined through a beam-based alignment procedure [3], and are qualified for regular LHC operation by creating losses in the horizontal (H) or vertical (V) planes of one beam, and observing the resulting loss map [4], a snapshot of the 1 Hz BLM signals versus the longitudinal position in the ring. The losses are created by exciting the beam using the transverse damper [5]. Examples of loss maps at flat top (energy: 6500 GeV) in each of the two beams and two planes in each beam are shown in Fig. 1. In the plots, the black, red and blue bars represent readings from BLMs at collimators, normal conducting and superconducting magnets respectively.

The collimation hierarchy is correctly set up when the highest losses in the ring occur at the IR7 primary collimators (TCP), followed by the secondary collimators (TCSG) and absorbers (TCLA). The collimation system cleaning inefficiency is the ratio of the leakage of protons to the IR7 dispersion suppressor superconducting magnets to the losses at the primary collimator. Determining the characteristics of the time-varying beam losses during the LHC machine cycle,

\* gianluca.valentino@um.edu.mt

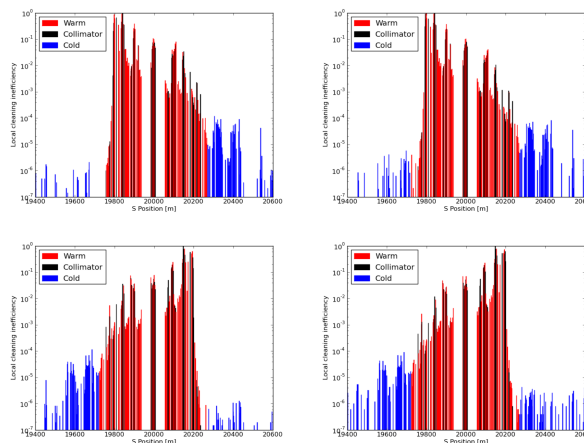


Figure 1: Examples of loss maps (zoom into IR7) during excitations in the B1H (top left), B1V (top right), B2H (bottom left) and B2V (bottom right) planes.

such as the beam and plane, is necessary to understand the impact of the losses on the machine performance, luminosity production and lifetime of accelerator components. The aim is therefore to be able to classify between the four types of loss planes: Beam 1 (B1) horizontal, B1 vertical, Beam 2 (B2) horizontal and B2 vertical.

An existing beam loss decomposition method [6] based on Singular Value Decomposition (SVD) is already able to determine the individual contributions of the four loss planes, and works for both off-momentum and betatron losses. It uses a calibration factor, obtained through dedicated collimator scraping measurements, to convert BLM readings in Gy/s to proton/s. A subset of only six BLMs per beam at horizontal and vertical collimators, determined through experience with measurements, is used. This vector is then decomposed as a linear combination of the individual B1H/B1V/B2H/B2V contributions.

Machine learning techniques have been applied to a wide variety of use cases to understand patterns in data, automate processes and anomaly detection. In the latter case, some work has already been performed to detect minor changes in the loss map hierarchy [7, 8] in LHC. In this paper, we use Principal Component Analysis to determine a different set of features which may be used to discern between beam loss planes when compared to the SVD technique, as well as train a classifier on loss map datasets using different features, applying the model to beam loss data from the LHC machine cycle.

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2018). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

This is a preprint — the final version is published with IOP

## DATASET PREPARATION

The data from around 3600 BLMs were extracted each time the transverse damper blow-up was running during the 2017 LHC proton physics run. This resulted in a total of 5893 loss maps, which were then narrowed down based on the following criteria. Firstly, it was ensured that the intensity loss in each loss map was more than  $1 \times 10^8$  p/s to have sufficient resolution in the BLM signals in IR 7. Secondly, only loss maps in which the collimator positions were identical were considered. Finally, visual checks of the loss map plots were made to ensure a correct hierarchy was in place. The final breakdown of the loss maps split by beam / plane and energy is shown in Table 1. As the beam characteristics and IR7 collimator settings are different between injection energy and top energy (which includes loss maps at flat top, and with squeezed non-colliding and colliding beams), separate models were trained for the respective cases.

Table 1: Final Set of Loss Maps Split by Beam / Plane and Energy (injection = 450 GeV, top energy = 6500 GeV)

Beam and Plane	Injection	Flat Top
B1H	84	496
B1V	132	599
B2H	127	383
B2V	123	129

## FEATURE SELECTION

Three feature sets were considered. The first set consisted of all the BLMs in the IR7 longitudinal position range 19400 - 20600 m (261 BLMs). The second set consisted of only the 12 BLMs (6 per beam, located at primary and secondary collimators in IR3 and IR7) used in the existing SVD technique. Finally, the third set of features was obtained by using the scikit-learn [9] implementation of Principal Component Analysis (PCA) to convert the set of 261 BLMs in IR7 into an equivalent number of linearly uncorrelated principal components. The cumulative explained variance ratio, shown in Fig. 2, indicates that the first eight principal components can already be used to explain  $\sim 100\%$  of the data, and therefore would be sufficient to use as features to discriminate between loss map planes. The first three principal components are shown in Fig. 3, in which the distinction between loss map planes can be appreciated visually from the clusters of loss maps formed.

For each feature set, the original BLM readings for each loss map were normalized in two different ways. In the first normalization procedure, all the readings were normalized to a BLM situated in the middle of IR7 (TCSG.A4R7.B1). In the second procedure, all the BLMs allocated to detecting losses in a particular beam were normalized to one of the first BLMs in the beam direction in IR7 (TCP.A6L7.B1 and TCP.A6R7.B2 when considering all IR7 BLMs, and TCP.C6L7.B1 and TCP.C6R7.B2 when considering only the BLMs used in the existing decomposition method). In reality however, the BLMs will detect electromagnetic showers which propagate in the tunnel originating from either beam.

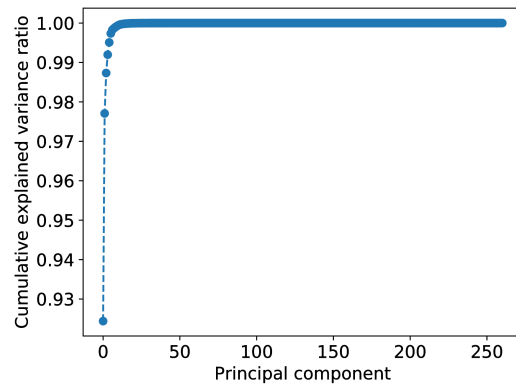


Figure 2: Cumulative explained variance ratio versus principal component number (sorted by highest contribution).

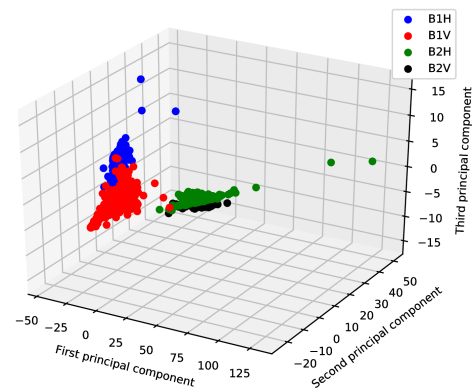


Figure 3: First three principal components for injection loss maps.

## MODEL TRAINING

Two separate models were trained for the loss maps at injection energy and at top energy. Each of the four loss map datasets (B1H, B1V, B2H, B2V) were split in a 75-25 ratio into training and testing datasets respectively. The allocation of a particular loss map to the training or testing datasets was done randomly. The scikit-learn implementation of the Gradient Boosting Classifier (GBC) [10] was trained on each of the three feature sets described previously. The models were then used to predict labels for the as yet unseen testing dataset, and the predicted labels were compared to the original labels to measure the success rate. Grid search cross-validation was used to determine the best parameters for GBC, which were found as shown in Table 2.

Table 2: Optimal GBC Parameters from Cross-Validation

Parameter	Description	Value
n_estimators	Number of boosting stages to perform	1000
max_depth	Max. depth of individual regression estimators	2
min_samples_split	Min. # samples required to split internal node	2

This is a preprint — the final version is published with IOP

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2018). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

## RESULTS

### Classification of Loss Maps

The final success rates for the models trained for loss maps at injection and at top energy were calculated by averaging the prediction performance on the testing dataset over five tries due to the random allocation of loss maps to the datasets. The results with the initial BLM readings separated by beam and normalized to separate BLMs are shown in Fig. 4. For this type of normalization, the PCA feature set performs better. On the other hand, the success rates are ~100% when the BLM readings are normalized to the same BLM in each loss map. This illustrates that the normalization procedure has a greater bearing on the final success rate than the selected feature set.

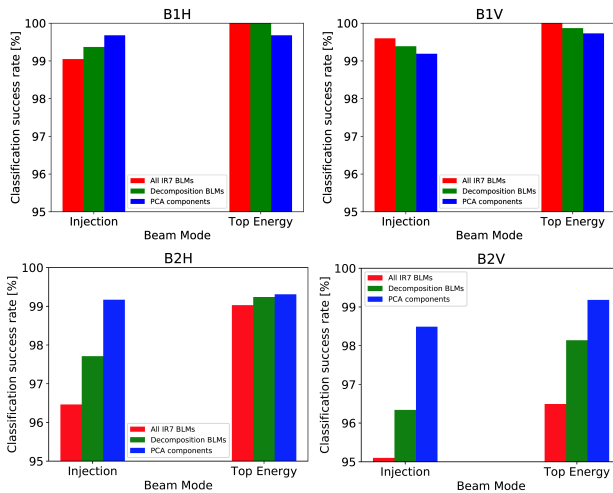


Figure 4: Classification success rates with initial BLM readings separated by beam and normalized to separate BLMs.

### Classification of Losses during the Cycle

The machine learning models are trained on loss maps performed in controlled conditions, in which the beam and plane are known. These models were then applied to losses during the beam squeeze in the LHC machine cycle. An example from Fill 6266 is shown in Fig. 5, from which it can be seen that the SVD technique predicts that the dominant contribution should initially be B1H for most of the squeeze up to ~1000 s, when it is then superseded by B1V and B2V.

The output from the GBC classifier is a class prediction probability in the range (0, 1) where a probability of less than 0.5 would indicate one class e.g. B1H, and a probability of more than 0.5 would indicate another class e.g. B1V. The class prediction probability applied to the same losses in the squeeze of Fill 6266 is shown in Fig. 6. First, the beam (B1 or B2) is predicted by using the machine learning model trained with the PCA features obtained after normalizing the initial BLM readings to the same BLM. Then, given that for most of the squeeze a loss in B1 is predicted, the machine learning model trained with the same features obtained after the initial BLM readings are separated by beam and normalized to

separate BLMs, and the initial dominance of B1H followed by a switch to B1V at ~1000 s is observed. Therefore, a good comparison is noted between the SVD technique and the machine learning model.

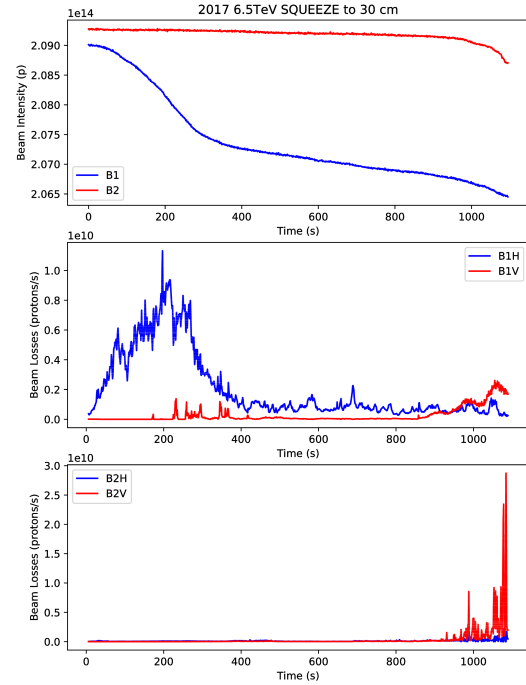


Figure 5: Intensity loss and contributions of losses in the different beams and planes predicted by the SVD technique.

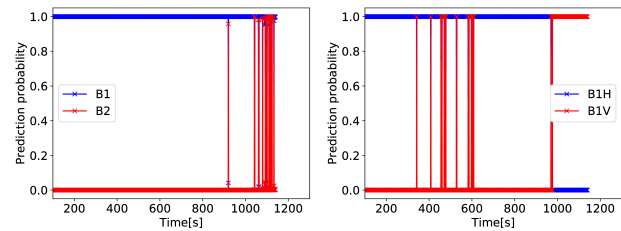


Figure 6: Contributions of losses in the different beams and planes predicted by GBC (left - initial BLM readings normalized to the same BLM to predict the beam; right - initial BLM readings separated by beam and normalized to separate BLMs to predict plane for B1).

## CONCLUSIONS

Determining the characteristics of beam losses in the LHC is an important step towards understanding their impact on the machine performance and long-term effects on accelerator components. Machine learning techniques were used to determine suitable features and train a Gradient Boosting Classifier model to classify between different types of loss planes in the LHC. The trained models were then used to classify beam losses during the machine cycle.

## ACKNOWLEDGEMENTS

The support from LHC collimation project team in measurements and analysis of the beam loss maps is appreciated.

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2018). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

This is a preprint — the final version is published with IOP

## REFERENCES

- [1] S. Redaelli *et al.*, “LHC Collimation for the Run II and Beyond”, presented at HB 2016, Malmo, Sweden, 2016.
- [2] E. B. Holzer *et al.*, “Beam loss monitoring system for the LHC”, Proc. of IEEE NSS, Puerto Rico, USA, pp. 1052-1056, 2005.
- [3] G. Valentino *et al.*, “Semiautomatic beam-based LHC collimator alignment”, Phys. Rev. ST Accel. Beams 15, 051002 (2012).
- [4] V. Moens *et al.*, “Comparison of LHC beam loss maps using the transverse damper blow up and tune resonance crossing methods”, in *Proc. 4th Int. Particle Accelerator Conf. (IPAC'13)*, Shanghai, China, May 2013, paper MOPWO050, pp. 1008-1010.
- [5] W. Hofle *et al.*, “Controlled transverse blow-up of high-energy proton beams for aperture measurements and loss maps”, in *Proc. 3rd Int. Particle Accelerator Conf. (IPAC'12)*, New Orleans, USA, May 2012, paper THPPR039, pp. 4059-4061.
- [6] B. Salvachua *et al.*, “Decomposition of beam losses at LHC”, in *Proc. 8th Int. Particle Accelerator Conf. (IPAC'17)*, Copenhagen, Denmark, May 2017, paper MOPAB009, pp. 88-91.
- [7] P. Theodoropoulos, M.Sc. Thesis, Univ. College London, London, England, CERN-THESIS-2015-128, 2015.
- [8] G. Valentino *et al.*, “Anomaly detection for beam loss maps in the Large Hadron Collider”, J. Phys.: Conf. Ser. 874, 012002 (2017).
- [9] F. Pedregosa *et al.*, Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.
- [10] J. H. Friedman, “Greedy function approximation: A gradient boosting machine”. Ann. Statist. 29 (2001), no. 5, pp. 1189-1232.