

# Application of machine learning techniques at the CERN Large Hadron Collider

---

**F.F. Van der Veken<sup>\*a,b</sup>, G. Azzopardi<sup>a,b</sup>, F. Blanc<sup>c</sup>, L. Coyle<sup>a,c</sup>, E. Fol<sup>a,d</sup>,  
M. Giovannozzi<sup>a</sup>, T. Pieloni<sup>a,c</sup>, S. Redaelli<sup>a</sup>, L. Rivkin<sup>c,e</sup>, B. Salvachua<sup>a</sup>,  
M. Schenk<sup>a,c</sup>, R. Tomás Garcia<sup>a</sup>, and G. Valentino<sup>a,b</sup>**

<sup>a</sup>*CERN, Beams Department  
1211, Meyrin, Switzerland*

<sup>b</sup>*University of Malta  
MSD2080, Msida, Malta*

<sup>c</sup>*Ecole Polytechnique Federale Lausanne,  
1015, Lausanne, Switzerland*

<sup>d</sup>*Johann-Wolfgang-Goethe University  
60323, Frankfurt am Main, Germany*

<sup>e</sup>*Paul Scherrer Institut  
5232, Villigen, Switzerland*

*E-mail:* [frederik.van.der.veken@cern.ch](mailto:frederik.van.der.veken@cern.ch)

Machine learning techniques have been used extensively in several domains of Science and Engineering for decades. These powerful tools have been applied also to the domain of high-energy physics, in the analysis of the data from particle collisions, for years already. Accelerator physics, however, has not started exploiting machine learning until very recently. Several activities are flourishing in this domain, in view of providing new insights to beam dynamics in circular accelerators, in different laboratories worldwide. This is, for instance, the case for the CERN Large Hadron Collider, where since a few years exploratory studies are being carried out. A broad range of topics have been addressed, such as anomaly detection of beam position monitors, analysis of optimal correction tools for linear optics, optimisation of the collimation system, lifetime and performance optimisation, and detection of hidden correlations in the huge data set of beam dynamics observables collected during the LHC Run 2. Furthermore, very recently, machine learning techniques are being scrutinised for the advanced analysis of numerical simulations data, in view of improving our models of dynamic aperture evolution.

*European Physical Society Conference on High Energy Physics - EPS-HEP2019 -  
10-17 July, 2019  
Ghent, Belgium*

---

<sup>\*</sup>Speaker.

## 1. Introduction

In recent years, Machine Learning (ML) techniques have found their way into the field of accelerator physics. The first attempts to use these techniques in beam diagnostics and beam control systems already date from a few decades ago [1, 2], but only recently some sizeable progress has been made. At the CERN Large Hadron collider (LHC), several ML applications are under study and in full development, four of which are discussed here.

In Section 2 it is detailed how supervised learning (SL) has reduced the time needed to align the LHC collimation system. In Section 3 we use unsupervised learning (UL) to clean optics measurements, and supervised ML to correct optics functions in the machine. In Section 4 it is shown how UL can be used in tracking simulations to detect outliers efficiently. Finally, in Section 5 we use supervised ML to create a model that directly relates measured beam lifetime to certain machine settings.

## 2. Alignment of the LHC Collimators

The LHC relies on a collimation system to absorb unavoidable beam losses before they reach the superconducting magnets and eventually lead to a quench [3]. The collimation system makes use of 100 collimators, whereby each of them is made of two parallel absorbing blocks positioned symmetrically around the beam. The position of each collimator's jaws must respect a hierarchy, with the settings determined following a beam-based alignment (BBA) established in [4]. This procedure moves collimator jaws separately towards the beam halo, whilst monitoring the measured beam-loss signal provided by a dedicated Beam Loss Monitoring (BLM) device positioned immediately downstream. A collimator is said to be aligned when both jaws are centred around the beam after touching the beam halo.

Collimators must be aligned each year during beam commissioning, to ensure the correct setup for the LHC prior to achieving nominal operation conditions. Their settings are monitored along the year, and different collimator setups are required when machine parameters are changed.

The alignment of a jaw relies on the ability to classify between alignment spikes and non-alignment spikes in the BLM signal. A collimator must continuously move towards the beam, ignoring any non-alignment spikes, until a clear alignment spike is observed. An alignment spike consists of a steady-state signal before the spike, the loss spike itself, the temporal decay of losses, and a steady-state signal after the spike. The steady-state is a result of the continuous scraping of halo particles when the jaw positions are fixed. The deeper a jaw cuts into the beam halo the more the steady-state signal increases. Any other spikes which do not follow this pattern are classified as non-alignment spikes. Such non-alignment spikes arise due to other factors, e.g. beam instabilities or mechanical vibrations of the opposite jaw [5].

The process of spike recognition was automated by casting it as a classification problem, as explained in [5]. ML models were trained to distinguish between the two spike patterns in the BLM losses. A dataset was assembled from previous alignment campaigns in 2016 and 2018. Fourteen manually-engineered features were extracted from this dataset and were analysed. The five most important features (1 feature for spike height, 3 features for exponential decay, 1 feature for collimator position) were used to train six machine learning models for comparison, namely

Logistic Regression (LR), Neural Network (NN), Support Vector Machine (SVN), Decision Tree (DT), Random Forest (RF), Gradient Boost (GB). Each model was analysed in-depth, optimised using hyper-parameters, and thoroughly tested on unseen data.

The machine learning model was then incorporated into the BBA software [6], together with the necessary threshold-selection algorithm [7] and cross-talk analysis [8]. This new alignment software was successfully used for all alignments throughout 2018, and decreased the alignment time of 79 collimators at injection by 71.4% compared to the semi-automatic alignment in 2017, namely from 2.8 hours to 50 minutes [9, 10].

### 3. Optics Measurements and Corrections

SL methods are used to explore the opportunities to build regression models to reconstruct magnet errors from optics perturbations caused by these errors. UL is needed in order to detect faulty Beam Position Monitor (BPM) signals, which produce nonphysical outliers in the optical functions computed from BPM data.

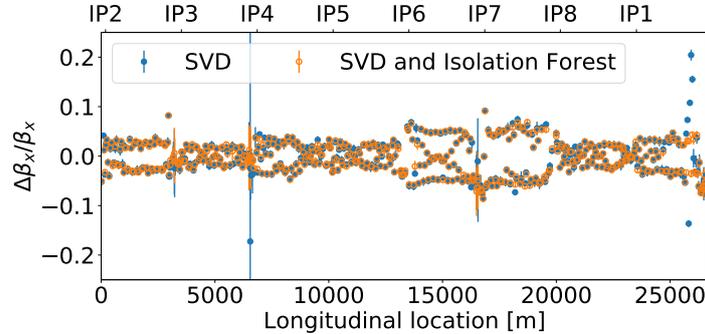
Currently, optics corrections in the LHC are performed in two steps, i.e. local, using the Segment-by-Segment technique [11], and global corrections using the Response Matrix technique [12]. Several regression models have been trained on MAD-X simulations [13], producing input-output pairs for supervised training. Each pair contains the errors in the circuits (quadrupoles powered in series to be trimmed) as input, and the horizontal and vertical phase advances of the simulated optics, perturbed by the given errors in the circuits, as output. After training, models are tested on simulated optics perturbed with single quadrupoles to examine the method performance on more realistic data. The comparison of  $\beta$ -beating corrections achieved by different regression models is given in Table 1. The results clearly demonstrate the ability of ML-based regression models to compute efficiently the optics corrections (see [14] and [15] for more detail).

$\beta$ -beating %	peak	rms
Uncorrected	$32 \pm 10$	$11 \pm 3$
Response Matrix	$11 \pm 5$	$3 \pm 2$
Orthogonal Matching Pursuit	$11 \pm 2$	$3.5 \pm 0.8$
Convolutional Neural Network	$11 \pm 2$	$3.2 \pm 0.5$
Ridge regression	$10 \pm 2$	$2.9 \pm 0.8$
Linear regression	$9 \pm 2$	$2.6 \pm 1.7$

**Table 1:** Comparison of  $\beta$ -beating averaged over 100 simulations considering the standard deviation as uncertainty. The optics measurements are simulated using the  $\beta^* = 40$  cm optics, 2016 for Beam 1.

The application of UL in optics measurements is motivated by the appearance of nonphysical spikes in the optics functions computed from BPM turn-by-turn readings. The presented cleaning technique uses harmonic analysis of the BPM signal (tune, amplitude, and noise-to-amplitude ratio) as input, in order to detect faulty BPMs as anomalies, before the optical functions are obtained from these data. Several clustering methods have been compared; the Isolation Forest (IF) algorithm [16] achieves the best results. The comparison of the  $\beta$ -beating computed from data cleaned with previously-existing techniques [17], and data additionally cleaned using the IF algorithm, is shown

in Fig. 1. This method recently became a standard part of optics measurements at the LHC, and has been successfully used during beam commissioning and machine development for different optics settings in 2018 (see [18] for more detail).



**Figure 1:** The comparison between beta-beating computed before and after IF cleaning demonstrates that anomaly detection significantly reduces the number of nonphysical spikes. The optics is computed for Beam 2 with  $\beta^* = 50$  cm.

#### 4. Dynamic Aperture Studies

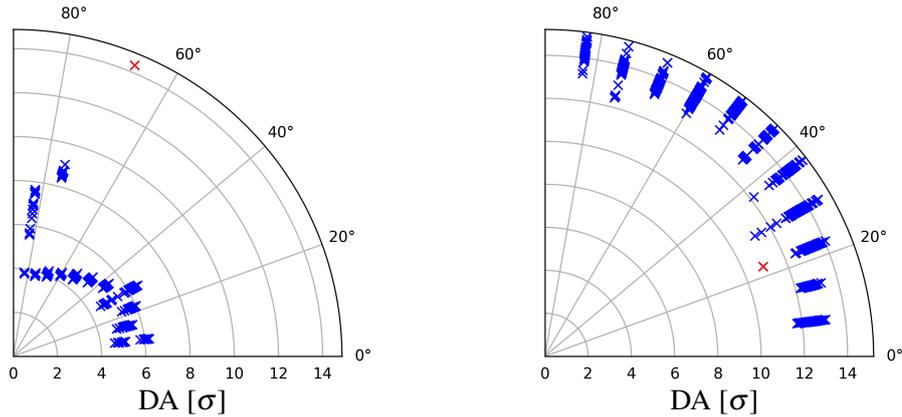
One of the important tools in the study of beam dynamics is the concept of Dynamic Aperture (DA), which represents the extent of the smallest connected volume in phase space that is stable for a given amount of time. It can be estimated from tracking simulations, where it is calculated for different random realisations of the magnetic errors (the so-called seeds) and over a given set of initial conditions uniformly distributed in polar co-ordinates in physical space.

It is not uncommon that for a given seed and angle the DA differs a lot from the value obtained for the other seeds, thus generating an outlier. It might be due to a numerical error or file corruption, in which case a remedy should be found, or be the result of a particular sensitivity of the seed to the underlying physics (e.g. due to internal cancellations or closeness to resonances). In both cases, it could be argued that it is not representative of a real-life scenario and it should be removed from the following analyses.

One has to make sure to distinguish a set of outliers from a justifiable split of a set of points into a certain number of clusters. First, for each angle the DA values for the different seeds are rescaled between the minimum and maximum values. These points are then clustered with help of the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [19]. For those points that are not assigned to any cluster after applying the algorithm, the Local Outlier Factor (LOF) is calculated to quantify their outlier strength. Finally, a minimum threshold on the distance between an outlier and the nearest cluster is imposed, and we only recognise a DA value as an outlier when it is a minimum or a maximum for that angle.

The algorithm has been tuned to avoid false negatives as much as possible. The flagged outliers are reviewed manually anyway and false positives are thus removed. On the left side of Fig. 2, we see an example of a point that is clearly an outlier, and correctly identified as such by the algorithm. On the right side, we see an example of a point that is an outlier for this specific angle,

however, when comparing it to its neighbouring angles this might not necessarily be true anymore. If one or both minima of the neighbouring angles are points from the same seed and have a value similar to the outlier, it might be a false positive instead and could be kept. On the other hand, the governing dynamics that defines the DA can be very different from angle to angle, so even when the neighbouring points are similar the point might still be an outlier, e.g. due to resonance structures. It is clear that particular care needs to be taken in these cases before drawing any conclusions, and additional investigations might be advisable.



**Figure 2:** Example of a DA computation with a true outlier (left) and a false positive (right).

For most studies, we observe that outliers are far more present at high angles and that certain seeds contain far more outliers than others. The analyses of the distribution of outliers potentially gives us more insight into the sensitivity of the underlying physics and investigations are still ongoing.

A final investigation of the usage of ML techniques in DA studies concerns the time evolution of DA. It is known from theory that the DA shrinks over time, following well-defined scaling laws [20–23]. These laws can be used to extrapolate a CPU-intensive simulation, done for a relatively short tracking time, to realistic and hence much larger timescales. So far this has led to very promising results, however, the fitting procedure is rather sensitive to several details. We conducted some very preliminary research into using a Recurrent Neural Network (RNN) to extrapolate the DA. This is of course an overly simplistic brute-force approach as it contains no information on the theoretical knowledge about the actual scaling law. The results indicate that the RNN is not really able to make an accurate prediction. We only used a basic implementation however, so improvements can probably be made. On the other hand, one very interesting approach, which is currently being investigated, is not to completely discard the theoretical knowledge, but to use ML techniques to improve the above-described fitting procedure, e.g. by finding a set of optimal fitting weights for the deterministic models.

## 5. Beam Lifetime Optimisation

The LHC is a complex machine with numerous intertwined systems, each potentially impacting the dynamics and stability of the beams. As such, building a rigorous model of particle losses

occurring in the LHC is a very daunting task, but it would offer valuable insight into the inner workings of the machine, to push further its performance. The main goal is to develop a system capable of determining the optimal set of operational knobs, so as to maximise the beams' intensity lifetimes, given a specific machine configuration. This system could then assist operators in the control room in the decision-making process.

The approach taken is to make use of the swaths of LHC data acquired through the several instrumentation systems to build a data-driven surrogate model of the beams' lifetime. This could then be coupled with an optimisation algorithm to determine the optimal operational knobs.

This problem was treated with a SL framework. The output of the model is the beam lifetime, and the inputs are the operational knobs of the machine, i.e. tunes, chromaticities and magnet currents. The data cover an entire operational year, but in order to simplify the input/output relationship, a small section of the complete machine cycle, corresponding to the injection energy, has been used, although this can be easily extended.

Several SL models were trained and compared, and the best performance was achieved with a Gradient Boosted Decision Tree model [24], which was paired with an off-the-shelf simplex optimiser [25] to extract the optimal machine configuration from the trained lifetime response.

The distribution of the input data was far from ideal, which is to be expected, as the operators, rightfully, do not explore the input parameter space during physics runs. This hindrance could potentially severely hamper the predictive power of the surrogate model.

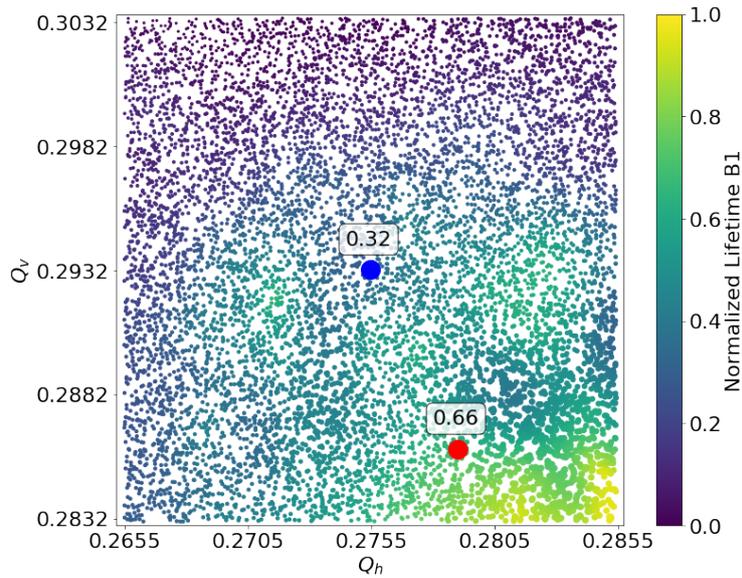
The parameter space was further explored with the help of a dedicated machine experiment, in which multiple random tune scans were performed over varying machine configurations. The data collected during this study are used to benchmark and supplement the current model. We observed a number of instabilities that increased the beams' emittances, thus reducing the collider's performance. Such instabilities are not taken into account by the model and are an obvious weakness of this setup. Nonetheless, ignoring this blind spot and restricting ourselves to the, albeit naive, lifetime optimisation problem, the model does agree with the lifetime optimal regions of the vertical  $Q_v$  vs horizontal  $Q_h$  tune diagrams, see Fig. 3. The model is capable of moving towards the optimal regions, although falls short of the maximum. To treat this problem rigorously, the model should take into account the emergence of instabilities in a multi-objective optimisation framework.

## 6. Conclusions

A few applications of ML at the CERN LHC, using a variety of algorithms based on both supervised and unsupervised learning, have been presented and discussed in this paper.

UL has been applied to detect outliers in optics-function measurements, to identify faulty BPMs, and to detect outliers in DA simulation results, to spot numerical errors or remove unlikely realisations of the magnetic lattice. In both cases, the algorithms work as expected and will be fine-tuned further during their usage. A centralised SL approach, in which every user automatically contributes by using the manual verification of flagged outliers as training for the Neural Network, is being considered for DA simulations.

SL has been applied to automatise the alignment of collimators during LHC beam commissioning, to correct optics functions, and to predict the beam lifetime based on the LHC operational settings. For the collimators' alignment, the time gain as compared to manual alignments has been



**Figure 3:** Beam lifetime as a function of the LHC working point as measured for Beam 1. Red dot: Lifetime-optimised working point, as determined by the model, which is close, but not equal, to the measured maximum.

impressive enough to make the ML implementation the default one. A continued cross-talk analysis will allow to do more alignments in parallel. Concerning the optics corrections, already the basic NN implementation produced very interesting results. A larger dataset is being generated and more error sources and non-linearities are being added in order to create a more general model.

The beam lifetime ML model showed satisfactory results, once the parameter space was uncorrelated by a dedicated LHC machine experiment. Of course, this is but a first implementation, which will be used as a basis to build upon to further ameliorate the framework.

## References

- [1] E. Bozoki and A. Friedman, *Neural networks and orbit control in accelerators*, in *Proc. 4th European Particle Accelerator Conf. (EPAC'94)*, London, England, (1994) 1589.
- [2] E. Meier, Y. E. Tan and G. LeBlanc, *Orbit Correction Studies using Neural Networks*, in *Proc. 3rd Int. Particle Accelerator Conf. (IPAC'12)*, New Orleans (LA), USA, (2012) 2837.
- [3] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole et al., *LHC Design Report*, CERN Yellow Reports: Monographs. CERN, Geneva, 2004, [10.5170/CERN-2004-003-V-1](https://arxiv.org/abs/10.5170/CERN-2004-003-V-1).
- [4] R. W. Aßmann et al., *Expected performance and beam-based optimization of the LHC collimation system*, in *Proc. 9th European Particle Accelerator Conf. (EPAC'04)*, Lucerne, Switzerland, (2004) 1825.
- [5] G. Azzopardi et al., *Automatic spike detection in beam loss signals for LHC collimator alignment*, *Nucl. Instrum. Methods Phys. Res.* **A934** (2019) 10.
- [6] G. Azzopardi et al., *Software Architecture for Automatic LHC Collimator Alignment using Machine Learning*, in *Proc. 17th Int. Conf. on Accelerator and Large Experimental Physics Control Systems (ICALEPCS'19)*, New York (NY), USA, (2019) .

- [7] G. Azzopardi et al., *Automatic Beam Loss Threshold Selection for LHC Collimator Alignment*, in *Proc. 17th Int. Conf. on Accelerator and Large Experimental Physics Control Systems (ICALPCS'19)*, New York (NY), USA, (2019) .
- [8] G. Azzopardi, B. Salvachua and G. Valentino, *Data-driven cross-talk modeling of beam losses in LHC collimators*, *Phys. Rev. Accel. Beams* **22** (2019) 083002.
- [9] G. Azzopardi, A. Muscat, S. Redaelli, B. Salvachua and G. Valentino, *Operational results of LHC collimator alignment using machine learning*, in *Proc. 10th Int. Particle Accelerator Conf. (IPAC'19)*, Melbourne, Australia, (2019) 1208.
- [10] G. Azzopardi et al., *Operational results on the fully automatic LHC collimator alignment*, *Phys. Rev. Accel. Beams* **22** (2019) 093001.
- [11] M. Aiba et al., *First  $\beta$ -beating measurement and optics analysis for the CERN Large Hadron Collider*, *Phys. Rev. ST Accel. Beams* **12** (2009) 081002.
- [12] T. Persson et al., *LHC optics commissioning: A journey towards 1% optics control*, *Phys. Rev. Accel. Beams* **20** (2017) 061002.
- [13] L. Deniau et al., "MAD - Methodical Accelerator Design." Available at <http://madx.web.cern.ch/madx/>.
- [14] E. Fol, J. M. Coello de Portugal, G. Franchetti and R. Tomás, *Optics corrections using machine learning in the LHC*, in *Proc. 10th Int. Particle Accelerator Conf. (IPAC'19)*, Melbourne, Australia, (2019) 3990.
- [15] E. Fol, J. M. Coello de Portugal, G. Franchetti and R. Tomás, *Application of Machine Learning to Beam Diagnostics*, in *Proc. 39th Int. Free Electron Laser Conf. (FEL'19)*, Hamburg, Germany, (2019) .
- [16] F. T. Liu, K. M. Ting and Z.-H. Zhou, *Isolation forest*, in *Proc. 8th IEEE Int. Conf. on Data Mining (ICDM'08)*, IEEE Computer Society, (2008) 413.
- [17] R. Calaga and R. Tomás, *Statistical analysis of RHIC beam position monitors performance*, *Phys. Rev. Accel. Beams* **7** (2004) 042801.
- [18] E. Fol, J. M. Coello de Portugal and R. Tomás, *Unsupervised Machine Learning for Detection of Faulty Beam Position Monitors*, in *Proc. 10th Int. Particle Accelerator Conf. (IPAC'19)*, Melbourne, Australia, (2019) 2668.
- [19] M. Ester, H.-P. Kriegel, J. Sander and X. Xu, *A density-based algorithm for discovering clusters in large spatial databases with noise*, in *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining (KDD'96)*, Portland (OR), USA, (1996) 226, <https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf>.
- [20] M. Giovannozzi and F. F. Van der Veken, *Description of the luminosity evolution for the CERN LHC including dynamic aperture effects. Part I: the model*, *Nucl. Instrum. Methods Phys. Res.* **A905** (2018) 171 [1806.03058].
- [21] M. Giovannozzi and F. F. Van der Veken, *Description of the luminosity evolution for the CERN LHC including dynamic aperture effects. Part II: application to Run 1 data*, *Nucl. Instrum. Methods Phys. Res.* **A908** (2018) 1 [1806.03059].
- [22] F. F. Van der Veken and M. Giovannozzi, *Scaling Laws for the Time Dependence of Luminosity in Hadron Circular Accelerators based on Simple Models of Dynamic Aperture Evolution*, in *Proc. 61st ICFA Advanced Beam Dynamics Workshop on High-Intensity and High-Brightness Hadron Beams (HB'18)*, Daejeon, Korea, (2018) 260.

- [23] A. Bazzani, M. Giovannozzi, E. H. Maclean, C. E. Montanari, F. F. Van der Veken and W. Van Goethem, *Advances on the modeling of the time evolution of dynamic aperture of hadron circular accelerators*, *Phys. Rev. Accel. Beams* **22** (2019) 104003 [1909.09516].
- [24] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma et al., *Lightgbm: A highly efficient gradient boosting decision tree*, in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan et al., eds., pp. 3146–3154, Curran Associates, Inc., (2017),  
<http://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree.pdf>.
- [25] J. A. Nelder and R. Mead, *A Simplex Method for Function Minimization*, *The Computer Journal* **7** (1965) 308  
[<http://oup.prod.sis.lan/comjnl/article-pdf/7/4/308/1013182/7-4-308.pdf>].