OPTICS CORRECTIONS USING MACHINE LEARNING IN THE LHC

E. Fol[*], J. Coello de Portugal, G. Franchetti[†1], R. Tomás, CERN, 1211 Geneva 23, Switzerland

[1]also at Johann-Wolfgang Goethe University, 60438 Frankfurt am Main, Germany

[†] GSI Helmholtzzentrum für Schwerionenforschung, 64291, Darmstadt, Germany

Abstract

Optics corrections in the LHC are based on a response matrix between available correctors and observables. Supervised learning has been applied to optics correction in the LHC demonstrating promising results on simulations and demonstrating the ability to reach acceptably low β-beating. A comparison of different algorithms to the traditional response matrix approach is given, and it is followed by the presentation of further possible concepts to obtain optics corrections using machine learning (ML).

INTRODUCTION

ML techniques have found their application in a wide range of accelerator control tasks [1, 2]. Attempts to build beam diagnostics and beam control systems using ML have been made already in the past decades [3–5], including orbit corrections using dipole field strengths predicted from the orbit deviations using neural networks [3].

In the LHC, global optics corrections are performed trimming quadrupolar fields aiming to reduce the difference between the measured and design optics functions around the ring. The strengths of the quadrupole circuits required to achieve the desired low β-beating are obtained using the Response Matrix (RM) approach [6]. The identification of local error sources and finding the correction is based on the Segment-by-Segment technique [7] and is used mostly around the interaction points (IPs). In this work we focus on global optics correction only, the local corrections in the triplets region are usually performed prior to global correction computation.

SUPERVISED LEARNING APPLIED TO OPTICS CORRECTIONS

In order to compute the corrections, the measured data have to be compared with the ideal optics design. The deviations from ideal optics have to be compensated by computed corrections [6, 8, 9]. In terms of machine learning, this task can be defined as a regression problem that can be solved by training a model using measurements and corresponding corrections. Such a regression model requires a large dataset in order to be able to generalize and produce reliable results. As corrections are only performed few times per year, training on corrections from the past is not possible, since not enough data is available. Another approach for data acquisition is to simulate optics perturbation with known errors in the MAD-X variables that correspond to physical circuits in the machine. To correct the perturbed optics the circuit

strengths predicted by the regression model just have to be applied with the opposite sign.

Dataset Generation

In order to create a training set, random errors are introduced into the MAD-X-variables that represent physical circuits. It has to be noted though, that in the simulation of the training set we use the strength of the circuits and each circuit represents in general the powering of a set of several quadrupoles. While simulating the data, the circuit errors are the input of the simulations and the produced optics functions perturbed by the introduced errors are the output. The data is generated for Beam 1 for the 2016 optics settings, with $\beta^* = 40$ cm and using injection tunes.

To train the model we flip this relation, such that the circuit errors have to be found based on given optics perturbation. The optics measurements in the LHC are mainly concerned by the phase advance measured between neighbouring beam position monitors (BPM). Therefore, the phase advance deviations from the nominal optics measured at each BPM are considered as model input (features). A correction knob, which is a list containing the correction applied to each variable, is the desired output of the trained model. A simulation
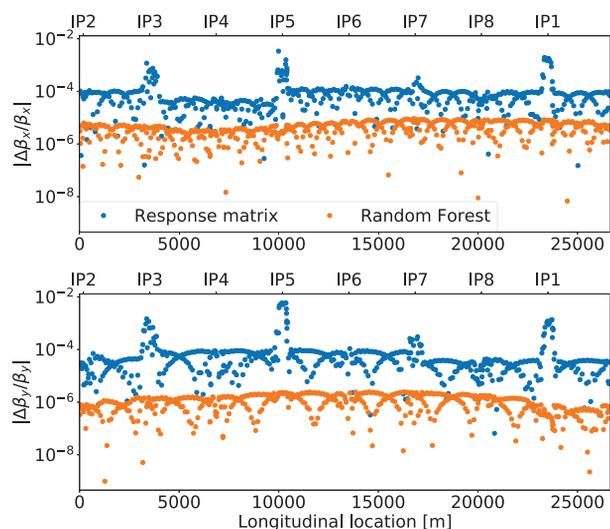

Figure 1: Simulated beta-beating after applying corrections computed with linear response matrix and predicted by Random Forest regressor on a measurement simulated with circuit errors and no noise. The mean absolute error between the introduced errors and computed corrections is 2×10^{-6} for response matrix and 3×10^{-8} for Random Forest.

---

* elena.fol@cern.ch

MC6: Beam Instrumentation, Controls, Feedback and Operational Aspects

T33 Online Modeling and Software Tools

data set of 100 000 samples was divided into train and test set (70% and 30% respectively), each sample pair consists of 1106 inputs (number of BPMs in both planes) and 190 outputs (correction variables).

Training and Model Evaluation

Several scoring techniques exist in the domain of model evaluation. In previous studies [10, 11] the comparison between ML models has been done based on the mean absolute error between true errors in the circuits and model output. The optics measurements in the test set were given perturbations by circuits only (not individual magnets) and no noise. The ML model trained using Random Forest (RF) regressor [12] achieved remarkably good results on correcting the β-beating as shown in Fig. 1. However, with the introduction of noise and more realistic simulation conditions, the time required to train a RF increased significantly from the order of tens of minutes to several hours, making this model unfeasible for further application for now. The reason for the increased training time has to be studied in the future.

In this study we aim to predict corrector values from simulated optics perturbed by individual quadrupoles which reflects the real state of optics perturbations in the LHC. The test dataset generated with circuit errors is therefore used only to control the fitting of the models during the training. To assert the ability of the model to predict the correctors values from the optics perturbed by single magnets, the measurements are simulated under following conditions:

- Gaussian distribution of quadrupolar errors with $3 \times 10^{-6} \mathrm{m}^{-2}$.

- Quadrupoles in the IP triplets and skew quadrupoles are excluded, since these errors are assumed to have been corrected with local correction techniques. These error sources can be considered in a next study.

- Phase advance noise is $10^{-3}\ 2\pi$ in a BPM at $\beta = 171$ m and scaled with the inverse square root of the β at the rest of BPMs.

The regression model for correction prediction can be trained using different ML techniques. These techniques usually require careful optimization and hyper-parameter tuning. In this study we use non-optimized models to acquire preliminary results on ability of ML models to correct the optics. The model has to be able to find an approximation of the function that describes the relation between input and output by learning from the data. We start with a ordinary least squares linear regression model. With new incoming training data samples, the model updates the fit by minimizing the sum of the squares of the residuals [13]. As next step we increase the complexity of the model by using Ridge regression, which applies regularization given by the L2-norm [14] in order to penalize the regression coefficients to achieve more stable prediction. Another method considered for correction prediction is Orthogonal Matching Pursuit (OMP) which is based on the K-SVD algorithm
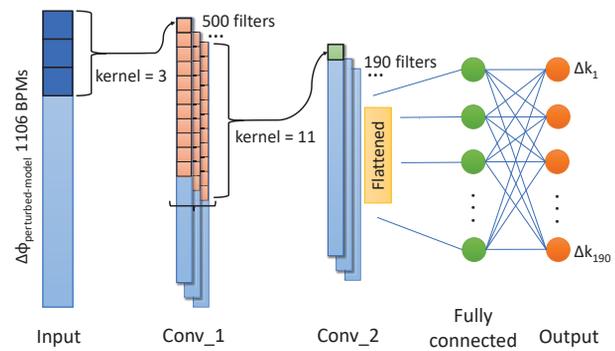

Figure 2: The composition of CNN implemented to predict 190 correctors values to correct the optics perturbed by individual quadrupole errors.

using Batch Orthogonal Matching Pursuit [15]. The models are implemented using scikit-learn python library [16].

Convolutional Neural Network

Regression tasks can be also solved using more complex ML techniques such as neural networks. In this study we use a special kind of neural networks called Convolutional Neural Network (CNN) [17] which has found a wide application in image-processing tasks demonstrating impressive results [18] among others in high-energy physics [19]. The main advantage of CNN is the ability to capture the spatial dependencies through the application of learned filters. In other words, the network can be trained to understand and extract spatially correlated features. This fact makes the CNN specially appealing for the application on correction prediction, as the determination of many optical parameters depends on the relationship between neighboring BPMs [20–22]. Opposite to most of the existing applications, here we apply a CNN on non-image based data, but on 1-dimensional vectors of phase-advance errors. The Keras python library [23] with TensorFlow backend [24] has been used in order to build and train the CNN model. For the first preliminary attempt presented in this study, we chose a simple architecture shown in Fig. 2. The network uses ReLu activation function [25] and the weights were initialized using random uniform distribution.

The dataset generation using MAD-X took several hours and it is the main concern regarding the time needed to prepare a ML model. The training itself takes less than a minute in case of regression models and around one hour for a non-optimized CNN.

RESULTS

As described in the previous section, we evaluate the models on simulated measurements perturbed by single-quadrupole errors. The optics correction results using different ML models as well as the correction obtained by the traditional response matrix approach are presented in Table 1. Response matrix was performed using default strength delta values for the LHC of $2 \times 10^{-5} \mathrm{m}^{-2}$. Comparing the peak and rms β-beating after applying the obtained corrections,
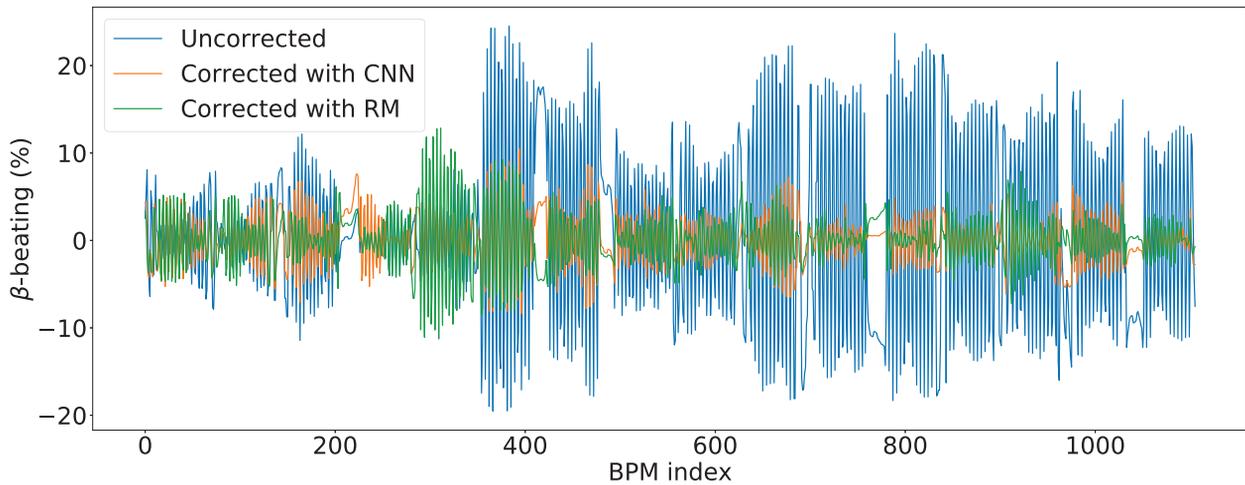
Figure 3: Simulated β-beating after applying corrections computed with linear RM and CNN trained on errors in the circuits, reducing rms β-beating from 9.5% to 3.2% and 3.0%, respectively. The measurement is simulated using the optics with $\beta^* = 40$ cm and with the introduction of betatronic phase noise.

we conclude that all methods achieve similar performance. Response matrix shows the largest peak β-beating and linear regression produces smallest rms β-beating. Figure 3 shows an example of β-beating correction using CNN demonstrating the ability of the method to achieve the correction comparable with traditional response matrix method. To be noted is that no tuning or optimization techniques have been applied to the ML models and hence, further improvements of results are expected. Considering possible model tuning, CNN has a greater potential for improvements due to a bigger model parameter space and various architecture options and hence, this method is preferred to the others for future studies.

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

β-beating %	peak	rms
Uncorrected	32±10	11±3
Response Matrix	11±5	3±2
OMP	11±2	3.5±0.8
CNN	11±2	3.2±0.5
Ridge regression	10±2	2.9±0.8
Linear regression	9±2	2.6±1.7

FUTURE PLANS

Observing the great potential of CNN to obtain sufficient optics corrections, the simple preliminary model used in the presented research should be further improved and studied. CNN architecture allows to extract the intermediate representation of the data, which can be extremely useful for the understanding of the training process and possibly find new relations or observables. Recently, iterative optics corrections using analytic response matrix approach have been implemented for the LHC [26]. CNN corrections can

be applied in a similar manner such that the corrections produced by an iteration are used to compute the expected optics which is then passed as the input for the next iteration.

Another possible solution for ML-based optics correction could be *Reinforcement Learning* [27]. This concept is based on environment-agent interaction. The agent takes an action on the environment, and the environment reacts producing a reward, which is used by the agent to learn how to improve its actions. For optics correction, minimization of β-beating can be used as the task to be solved. The task of the agent is to find optimal corrector values. As environment we can use MAD-X in order to compute the new state (β-beating) as response to the action (correctors values) taken by agent model (e.g. neural network). Once the model has been trained, it should be able to find optimal corrections to achieve as low β-beating as possible. Opposite to presented supervised learning approach, no input-output pairs are needed to train this model.

CONCLUSIONS

We have shown that every ML model applied in this study is capable of predicting corrector variables values needed do achieve β-beating comparable to traditional response matrix approach even without any optimization or parameter tuning. Linear regression performs better than response matrix, probably due to the benefit of extracting an average linear response over the training population instead of only using the unperturbed model. This is an important first step towards application of ML techniques to optics corrections in particle accelerators. All supervised ML models achieved similar results, but considering the large space for improvements of CNN through tuning of the network parameters and possibility to study the intermediate information representation between the layers, CNN appears to have the most potential to be applied to optics corrections and be useful to improve optics quality.

REFERENCES

[1] E. Fol, J. M. Coello de Portugal, and R. Tomas, "Application of Machine Learning to Beam Diagnostics", presented at the IBIC'18, Shanghai, China, Sep. 2018, paper TUOA02.

[2] A.L. Edelen *et al.*, "Neural Networks for Modeling and Control of Particle Accelerators", *IEEE Trans. Nucl. Sci.*, vol. 20, pp. 878-897, 2016. `doi:10.1109/TNS.2016.2543203`

[3] E. Bozoki and A. Friedman, "Neural Networks and Orbit Control in Accelerators", in *Proc. 4th European Particle Accelerator Conf. (EPAC'94)*, London, UK, Jun.-Jul. 1994, pp. 1589–1592.

[4] E. Meier, G. LeBlanc, and Y. E. Tan, "Orbit Correction Studies using Neural Networks", in *Proc. IPAC'12*, New Orleans, LA, USA, May 2012, paper WEPPP057, pp. 2837–2839.

[5] Y. Kijima, M. Mizota, K. Suzuki, and K. Yoshida, "A Beam Diagnostic System for Accelerator using Neural Networks", in *Proc. 3rd European Particle Accelerator Conf. (EPAC'92)*, Berlin, Germany, Mar. 1992, pp. 1155–1158.

[6] T. Persson *et al.*, "LHC optics commissioning: A journey towards 1% optics control", *Phys. Rev. Accel. Beams* vol. 20, p. 061002, 2017. `doi:10.1103/PhysRevAccelBeams.20.061002`

[7] M. Aiba *et al.*, "First β-beating measurement and optics analysis for the CERN Large Hadron Collider" , *Phys. Rev. Accel. Beams* vol. 12, p. 081002, 2009. `doi:10.1103/PhysRevSTAB.12.081002`

[8] R. Tomás *et al.*, "Record low β-beating in the LHC", *Phys. Rev. ST Accel. Beams* vol. 15, p. 091001, 2012. `doi:10.1103/PhysRevSTAB.15.091001`

[9] R. Tomás, M. Aiba, A. Franchi, and U. Iriso, "Review of linear optics measurement and correction for charged particle accelerators", *Phys. Rev. Accel. Beams* vol. 20, p. 054801, 2017. `doi:10.1103/PhysRevAccelBeams.20.054801`

[10] E. Fol, "Evaluation of Machine Learning Methods for LHC Optics Measurements and Corrections Software", Master thesis, CERN-THESIS-2017-336.

[11] E. Fol, F. S. Carlier, J. M. Coello de Portugal, A. Garcia-Tabares, and R. Tomás, "Machine Learning Methods for Optics Measurements and Corrections at LHC", in *Proc. 9th Int. Particle Accelerator Conf. (IPAC'18)*, Vancouver, Canada, Apr.-May 2018, pp. 1967–1970. `doi:10.18429/JACoW-IPAC2018-WEPAF062`

[12] L. Breiman, "Random forests", *Mach. Learn.*, vol. 45, no.1, pp. 5-32, Oct. 2001. `doi:10.1023/A:1010933404324`

[13] T. L. Lai and H. Robbins, "Strong consistency of least-squares estimates in regression models", in *Journal of Multivariate Analysis*, vol. 23, no. 1, pp. 77–92, 1987. `doi:10.1016/0047-259X(87)90179-5`

[14] R.M. Rifkin and R.A. Lippert, "Notes on Regularized Least Squares", Tech. Rep. MIT-CSAIL-TR-2007-025, CBCL-268, 2007.

[15] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient Implementation of the K-SVD Algorithm using Batch Orthogonal Matching Pursuit", CS Technion report CS-2008-08 2008.

[16] F. Pedregosa *et al.*, Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825–2830, 2011.

[17] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series", in *Handbook of brain theory and neural networks*, M. A. Arbib, Cambridge, MA, USA: MIT Press, 1998, pp. 255–258.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "ImageNet classification with deep convolutional neural networks", in *Proc. of the 25th International Conference on Neural Information Processing Systems (NIPS'12)*, Nevada, USA, Dec. 2012, pp. 1097–1105.

[19] A. Aurisano *et al.*, "A convolutional neural network neutrino event classifier", *Journal of Instrumentation*, vol. 11, no. 9, p. 09001, 2016. `doi:10.1088/1748-0221/11/09/P09001`

[20] A. Langner and R. Tomás, "Optics measurement algorithms and error analysis for the proton energy frontier", *Phys. Rev. ST Accel. Beams* vol. 18, p. 031002, 2015. `doi:10.1103/PhysRevSTAB.18.031002`

[21] A. Langner, G. Benedetti, M. Carlà, U. Iriso, Z. Martí, J. Coello de Portugal, and R. Tomás, "Utilizing the N beam position monitor method for turn-by-turn optics measurements", *Phys. Rev. Accel. Beams* vol. 19, p. 092803, 2016. `doi:10.1103/PhysRevAccelBeams.19.092803`

[22] A. Wegscheider, A. Langner, R. Tomás and A. Franchi, "Analytical N beam position monitor method", *Phys. Rev. Accel. Beams* vol. 20, p. 111002, 2017. `doi:10.1103/PhysRevAccelBeams.20.111002`

[23] F. Chollet *et al.*, Keras, `https://keras.io` (2015).

[24] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning", in *Proc. 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI'16)*, Savannah, GA, USA, Nov. 2016, pp. 265–283.

[25] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks", in *Proc. of the 14th international conference on artificial intelligence and statistics*, Lauderdale, FL, USA, Apr. 2011, pp. 315–323.

[26] J. W. Dilly *et al.*, "An Updated Global Optics Correction Scheme", CERN-ACC-NOTE-2018-0056.

[27] R.S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MA, USA: MIT Press, 1998.