ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-THPAB055

RECONSTRUCTION OF LINEAR OPTICS OBSERVABLES USING SUPERVISED LEARNING

E. Fol*, H. Garcia-Morales, R. Tomás
CERN, Geneva, Switzerland

Abstract

In the LHC, most of the optical functions can be obtained from turn-by-turn beam centroid data. However, the measurement of such observables as β^* and the dispersion function require special dedicated techniques and additional operational time. In this work, we propose an alternative approach to estimate these observables using supervised machine learning, in case the dedicated measurements are not available but turn-by-turn data are. The performance of developed estimators is demonstrated on LHC simulations. Comparison to traditional techniques for the computation of β-function will be also provided.

MOTIVATION

In order to obtain the normalised dispersion function, several beam excitations on- and off-momentum are required. Considering turn-by-turn motion in the absence of betatron coupling, the vertical dispersion can be neglected. The normalised horizontal dispersion $\frac{D_x}{\sqrt{\beta_x}}$ is used as an observable which does not depend on BPM calibration, as it is calculated as a ratio between two calibration-dependent quantities, cancelling the effect of calibration factor [1]. Reconstructing the normalised dispersion from a single transverse beam excitation, without any momentum shift, can save operational time by avoiding numerous measurement acquisitions with on- and off-momentum conditions.

The β-function at Interaction Points (IPs), also referred to as β^*, is typically computed in the LHC using the k-modulation technique [2, 3], which also produces accurate β measurements at the BPMs next to the IPs of four main experiments. Control of the beam in the IPs is crucial to ensure luminosity balance between experiments. Including β^* as a constraint in the global correction computation improves the results of existing correction techniques. This concerns the currently used response matrix approach, as well as a recently developed ML-based optics correction technique [4] to be tested in LHC Run III commissioning.

Besides providing additional input to optics correction algorithms, this approach can be advantageous also for the analysis of historical data where the measurements of such observables have not been performed.

INTRODUCTION TO SUPERVISED LEARNING

In this study, we employ the concept of supervised learning in order to build linear regression models for the prediction of normalised dispersion and β-function from the phase advance deviations from nominal design. To implement the supervised learning approach, a data set consisting of correlated input and output variables needs to be provided. Since the relation between phase advances and the optics functions to be predicted is known to be linear, a linear regression model, so called *Ridge Regression* [5] is applied. This choice allows faster training and ease the model parameter tuning in comparison to non-linear complex models, such as neural networks.

The Ridge regression model minimises the residual sum of squares between the true targets in the training data, and the targets predicted by the linear approximation. The tuning parameter α is responsible for the weights' regularisation - a special technique to control the weights' update during the training. The regression problem is formally described as the square of the Euclidean norm

$$\min_{w} ||\vec{X}w - \vec{y}||_2^2 + \alpha||w||_2^2, \quad (1)$$

where w is the matrix containing the weights of the regression model, $\vec{X}$ is the input data vector and $\vec{y}$ the vector of targets to be predicted by the model. We also use a special technique to improve the prediction quality on unseen data called *bagging*. Bagging is based on the idea of training several estimators on subsets of available training data and using the average of predictions made by single estimators as final prediction of a target value. In this study, optimal model parameters found by using cross-validation are 10 Ridge estimators combined into one model, each using 80% of training data and regularisation parameter $\alpha = 1 \times 10^{-3}$.

DATA AND MODEL GENERATION

In order to create a data set for the reconstruction of normalised dispersion and β-function, thousands of MAD-X simulations for $\beta^* = 40$ cm optics, introducing different distributions of quadrupolar gradient errors have been performed. Specifically this optics setting is used in the simulations since small β^* measurements can limit the performance of traditional measurements techniques such as k-modulation.

The required input features and output targets are extracted from the generated simulations. From a single simulation, two different data samples are produced, i.e. employing either the simulated normalised dispersion deviations from nominal model in the entire lattice as sample output or β-function at the BPMs left and right from IPs 1, 2, 5 and 8. As input, phase advance deviations from the nominal model are used in both data sets. To provide realistic data, the phase advances include the noise estimated from LHC measurements. In total, 80000 simulations are generated,

* elena.fol@cern.ch

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-THPAB055

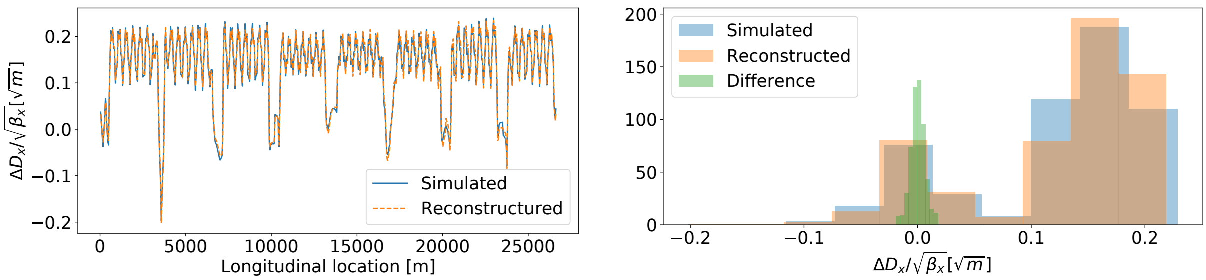

Figure 1: Reconstruction of horizontal normalised dispersion deviations in beam 1 from noisy phase advance data. The left plot illustrates the agreement between simulated and reconstructed values in one exemplary simulation, while the histogram on the right shows statistics obtained from 1000 simulations.

providing two data sets, where each sample consists of 2048 phase advances as input variables, and 32 or 1024 output targets corresponding to β-function and normalised dispersion reconstruction, respectively. In both cases, the trained regression models produce a combined reconstruction of optics functions in both beams simultaneously, including either both planes in the β-function reconstruction or horizontal plane only when reconstructing the normalised dispersion.

RECONSTRUCTION OF NORMALISED DISPERSION

The Ridge regression model is trained on 64 000 samples and tested with the remaining 16 000 samples generated as described above. We also use a separate set of 100 validation simulations in order to obtain statistically significant results on an independent data set. Figure 1 shows the comparison between normalised dispersion variation with respect to the nominal model predicted by the regression model and the corresponding actual simulated function. We provide a comparison based on a single simulation in order to demonstrate the agreement of the values at different BPM locations, as well as summarised result of reconstruction obtained from 100 validation simulation. The relative rms prediction error on the validation set is 7%.

RECONSTRUCTION OF β-FUNCTION

The current method available for the computation of β^*, k-modulation, is based on gradient modulation in the triplet magnets left and right of the IPs. By measuring the resulting tune changes, the average β-functions in the triplets can be calculated, which are then propagated towards IPs to obtain the β^*. This technique includes time consuming quadrupole current modulation and requires the cleaning of tune measurements [6] before computing the β-functions. Providing the estimates of β-function can shorten the analysis time, by reducing the steps to be performed if the traditional method is used.

The result of reconstructing the β-function at the BPMs next to the IPs is presented in Fig. 2. The Ridge regression model utilises 80 000 samples (80% and 20% for training and test respectively), each consisting of simulated noisy phase advance deviations in both beams and planes as input and β deviations in horizontal and vertical planes simulated for both beams at the BPMs left and right next to the IPs 1, 2, 5 and 8 being the output. The validation is then performed on a set of 100 LHC simulations. The resulting reconstruction error on validation set is 0.9% which is comparable to the uncertainty of k-modulation technique for the β^* measurements with $\beta^* = 40$ cm [2, 7].

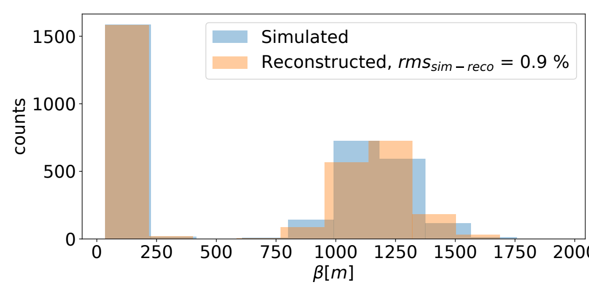

Figure 2: Comparison between simulated and reconstructed β in both planes and for both beams at the BPMs next to four main IPs for 100 simulations. RMS reconstruction error is 0.9%.

In order to verify the ability of the linear regression model to predict the β-functions around IPs from phase advances on realistic LHC data, the measurements taken in uncorrected machine in 2016 with $\beta^* = 40$ cm are used [8]. Since k-modulation measurements are not available in these data sets, the prediction of ML-model is compared to the values computed by the N-BPM method [9, 10], where the β-function is inferred from the phase advances between different BPMs combinations. Both, regression model and N-BPM method use the phase advances to reconstruct the β-function. However, the calculation of β from phase using the N-BPM method is known to be less reliable around IPs. Figure 3 shows the β-function values at the BPMs next to the four IPs, left and right, computed for both beams in horizontal and vertical planes. The agreement between the measured values computed using the N-BPM method and the reconstruction predicted by the regression model shows

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-THPAB055

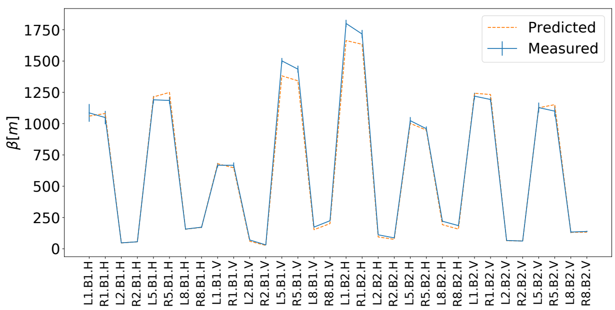

Figure 3: Agreement between measured and predicted β-function at BPMs left and right from IP1, 2, 5 and 8, including the values for both beams, horizontal and vertical planes.

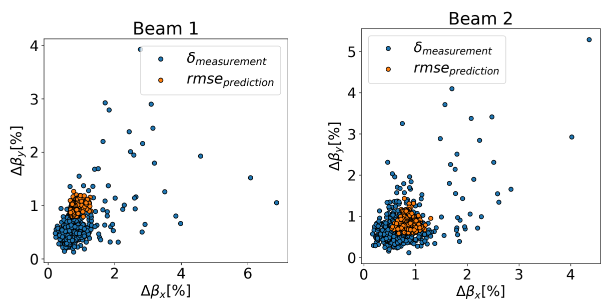

Figure 4: Comparison between the uncertainties in the computation of β-function measured in LHC commissioning in 2016 and rms error of ML-based β prediction obtained from 100 LHC simulations using $\beta^* = 40$ cm optics.

that the ML-model is capable to make accurate prediction on unseen operational data.

We also performed an additional verification of the approach to reconstruct the β-function using supervised regression models. By training a model on an extended data set, where the output targets consist of the full set of simulated β-function at all BPMs, we can compare the ML-based approach to the traditional method in terms of accuracy. The optics analysis tools, currently used for optics measurements at the LHC, also provide statistical error bars of optics functions calculated from the data obtained from several beam excitations. In Fig. 4, we provide a comparison between the calculated error bars in the measurement of β-function for $\beta^* = 40$ cm optics and rms error of ML-based reconstruction obtained from 100 LHC simulations. While the reconstruction error is identical for the majority of BPMs, independently from the plane and beam, the measurement uncertainty calculated using traditional methods can depend on the location. Moreover, malfunctioning BPMs can also impact the accuracy of optics computations. Although ML-based reconstruction is globally less accurate than currently applied methods, the reconstructed β values can be utilised in order to replace the measurements at the locations where large error bars are observed, or at the location of faulty BPMs, where the data are missing.

SUMMARY

The application of linear regression models allows to obtain estimates of the quantities whose computation otherwise requires additional beam excitation, as in the case of dispersion function or performing advanced optics analysis techniques as k-modulation in order to measure β^*. In both cases, supervised regression models estimate the target outputs from the phase advances which are available in any measurement set through harmonic analysis of turn-by-turn data [11, 12].

Table 1 summarises the scores achieved by training and testing the models on 80000 of LHC simulations used to build the data samples required for the supervised learning. The lower accuracy of normalised dispersion reconstruction can be explained by the fact that the regression model needs to update its weights during the training with respect to a much larger amount of output targets compared to the reconstruction of β-function around IPs. While the achieved scores state that there is no overfitting towards the training sets, the relative errors of prediction demonstrate that improvements are still possible. Hence, further steps of this study will be focused on boosting the prediction accuracy, potentially by the means of more sophisticated regression models.

Table 1: Explained Variance R^2, Mean Absolute Error of Prediction (MAE) and Residual Error of Prediction Relative to the Corresponding Simulated True Values, Demonstrating the Performance of Regression Models Trained to Predict $\beta_{x,y}$ Next to IPs and Normalised Dispersion $D_x/\sqrt{\beta_x}$

	$\beta_{x,y}$	$D_x/\sqrt{\beta_x}$
$R^2_{train/test}$	0.995/0.996	0.81/0.8
$MAE_{train/test}$	4.49 / 4.61 [m]	0.0058 / 0.006 [$\sqrt{\text{m}}$]
$\frac{rms_{residual}}{rms_{true}}$	0.9%	7%

The presented approach demonstrates the reconstruction of specific observables relevant for the optics analysis and corrections' computation, without performing dedicated measurements of these observables. Thus, it supports offline optics analysis by providing otherwise missing data and allows to speed up optics measurements and corrections at the LHC, saving the costly operational time.

REFERENCES

[1] R. Calaga, R. Tomas, and F. Zimmermann, "BPM Calibration Independent LHC Optics Correction", in *Proc. 22nd Particle Accelerator Conf. (PAC'07)*, Albuquerque, NM, USA, Jun. 2007, paper THPAS091, pp. 3693–3695.

[2] F. Carlier and R. Tomás, "Accuracy and feasibility of the β^* measurement for LHC and high luminosity LHC using k-modulation", *Phys. Rev. Accel. Beams*, vol. 20, p. 011005, 2017. doi:10.1103/PhysRevAccelBeams.20.011005

[3] H. Garcia-Morales *et al.*, "Phase advance constraint in K-modulation for β^* determination in the LHC", unpublished.

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-THPAB055

[4] E. Fol, R. Tomás, and G. Franchetti, “Supervised learning-based reconstruction of magnet errors in circular accelerators”, *Eur. Phys. J. Plus*, vol. 136, p. 365, 2021. doi:10.1140/epjp/s13360-021-01348-5

[5] R. M. Rifkin and R. A. Lippert, “Notes on Regularized Least Squares”, MIT Libraries, Cambridge, MA, United States, Rep. MIT-CSAIL-TR-2007-025, 2007.

[6] H. Garcia Morales, R. Tomas, and E. Fol, “Unsupervised Learning Techniques for Tune Cleaning Measurement”, presented at the 12th Int. Particle Accelerator Conf. (IPAC’21), Campinas, Brazil, May 2021, paper MOPAB184, this conference.

[7] H. G. Morales, “HL inner triplet error prediction using ML”, presented at Momentum compaction and ML, CERN, Geneva, Switzerland, Dec. 2020, unpublished.

[8] 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

[9] A. Langner *et al.*, “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

[10] 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

[11] R. Bartolini and F. Schmidt, “Sussix: A computer code for frequency analysis of non–linear betatron motion”, CERN, Geneva, Switzerland, Rep. CERNSL-Note-98-017, 1998.

[12] Harpy, https://github.com/pylhc/omc3/tree/master/omc3/harpy