ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOPT008



# SUPERVISED MACHINE LEARNING FOR LOCAL COUPLING SOURCES DETECTION IN THE LHC *

F. Soubelet†, CERN, 1211 Meyrin, Switzerland and University of Liverpool, Liverpool, UK
T.H.B. Persson, R. Tomás, CERN, 1211 Meyrin, Switzerland
O. Apsimon, C.P. Welsch, University of Liverpool, Liverpool, UK

*Abstract*

Local interaction region (IR) linear coupling in the LHC has been shown to have a negative impact on beam size and luminosity, making its accurate correction for Run 3 and beyond a necessity. In view of determining corrections, supervised machine learning has been applied to the detection of linear coupling sources, showing promising results in simulations. An evaluation of different applied models is given, followed by the presentation of further possible application concepts for linear coupling corrections using machine learning.

## INTRODUCTION

Machine learning techniques have found their application in a wide range of particle accelerator control tasks in the past [1–7]. The precise knowledge of a coupling source's location and relative strength is valuable for further correction. While one can look at sudden jumps in the coupling Resonance Driving Terms (RDTs) $f_{1001}$ and $f_{1010}$ to deduce the presence of sources in between BPMs, such a method does not accurately pinpoint the location of a source, nor can it be used in locations with little instrumentation or unfavorable conditions. In the LHC, this approach is applicable for any source located in the arcs, but due to unfavorable phase advances between Beam Position Monitors (BPMs) in the Insertion Regions (IRs) and the necessity to have no strong sources between measuring BPMs, it is inappropriate for the detection of sources located within these regions. In this work, we explore the possibility of using supervised machine learning to detect betatron coupling sources in the LHC IRs.

## SUPERVISED LEARNING FOR IR COUPLING SOURCES DETECTION

In order to perform a prediction of betatron coupling sources' locations, one first needs to compute the $f_{1001}$ and $f_{1010}$ RDTs. The strength and variations of the coupling RDTs throughout the machine is then used to estimate the location of coupling sources.

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 solutions. Such a regression model requires a large data set in order to be able to generalize and produce reliable results. As from the real machine the location of sources is unknown, no real-world data is available for the training of machine learning models.

### *Data Set Generation*

In order to create a training set, simulations are performed where random rotations around the s-axis are introduced into the MAD-X [8] quadrupoles, generating a skew quadrupolar component at the affected element and thus a source of coupling. It has to be noted though, that in the simulation of the training set we use tilt components to quadrupoles, given by the DPSI variable in MAD-X, which ignores the other potential sources such as feed-down from higher order magnets. While simulating the data, the introduced tilt components (DPSI values) are the input of the simulations and the produced coupling RDTs generated from the perturbed optics functions are the output. The data was generated for Beam 1 and 2 of the LHC, for the 2018 optics settings with $\beta^* = 30$ cm and using collision tunes ($Q_x = 0.31, Q_y = 0.32$). Figure 1 shows the reconstructed coupling RDTs for a given sample, calculated through the **C** matrix [9].

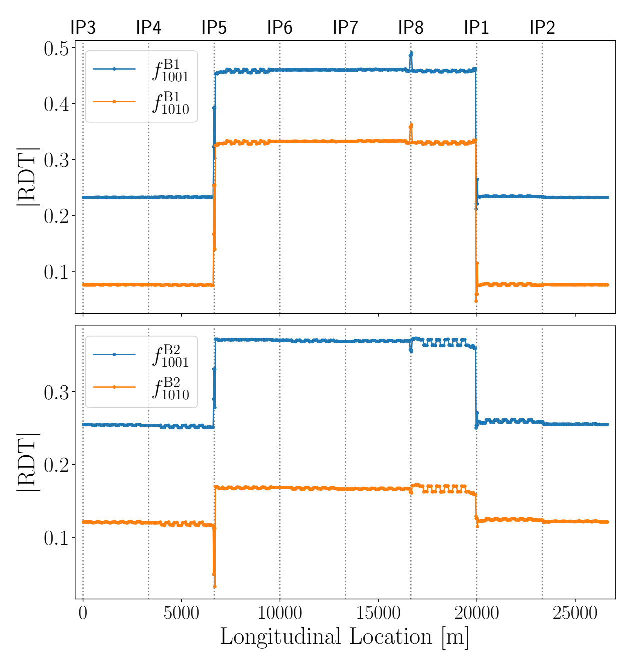


Figure 1: Coupling RDTs reconstructed for the LHC Beam 1 (top) and 2 (bottom) from the optics perturbed by the introduction of tilt errors to IR quadrupoles. Here a truncated Gaussian distribution with a standard deviation of 1 mrad was applied.

* This research is supported by the LIV.DAT Center for Doctoral Training, STFC and the European Organization for Nuclear Research.
† felix.soubelet@cern.ch

13th Int. Particle Acc. Conf. IPAC2022, Bangkok, Thailand JACoW Publishing
ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOPT008

The measurements are simulated with the MAD-X code, and each simulated sample is done by applying the following steps:

1. A truncated Gaussian distribution of tilt errors (DPSI) is applied to IR[1258] quadrupoles on Beam 1.
2. Quadrupoles located outside of the IRs are excluded as these sources can be well corrected by other means.
3. The coupling RDTs $f_{1001}$ and $f_{1010}$ are calculated at each BPM from TWISS functions for Beam 1.
4. The DPSI values for triplets are exported and applied to Beam 2 as these are common magnets.
5. A truncated Gaussian distribution is applied to the remaining IR[1258] quadrupoles in Beam 2.
6. Coupling RDTs for Beam 2 are calculated as done for Beam 1 and those two results are concatenated.

The standard deviation of the applied tilt errors was aligned with expected values from the element alignment precision in the LHC, after discussions with the alignment group.

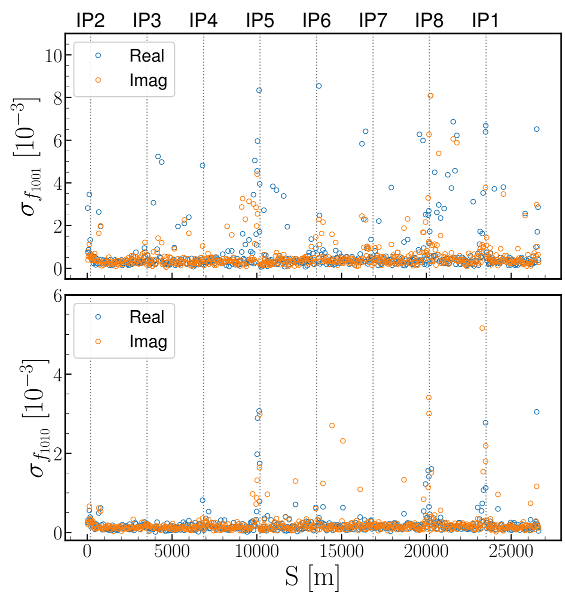


Figure 2: Standard deviation of the coupling RDTs at BPMs for Beam 1, from a batch of measurements taken on April 3, 2018. These data points were later divided into IR BPMs and arc BPMs to determine applied noise levels.

To train the model we flip this relation such that the introduced tilt errors have to be found based on given coupling RDTs computed from the perturbed optics. Therefore, the coupling RDTs reconstructed at each BPM are considered as model input (features). A vector containing the estimated DPSI value attributed to each affected quadrupole is the desired output of the trained model. A simulation data set of 50 000 samples was divided into train and test sets (75% and 25% respectively). Each sample pair consists in 4424 inputs (real and imaginary parts of each coupling RDT for each BPM for each beam) and 160 outputs (DPSI value at each affected IR quadrupole).

### *Training and Model Evaluation*

Various scoring techniques exist in the domain of model evaluation. In this study, models have been evaluated based on their $R^2$ scores (coefficient of determination) as well as the normalized mean absolute error between the true values and the model outputs. Both train and test sets were submitted to the addition of Gaussian noise on the reconstructed RDTs to simulate the uncertainty of the reconstructed RDTs.

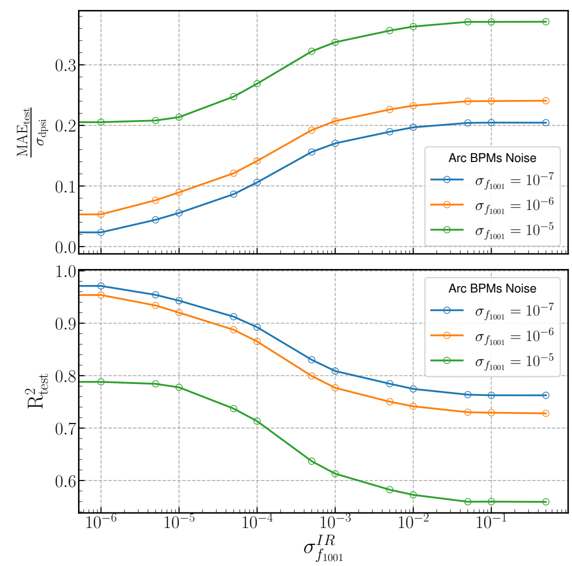


Figure 3: Normalized mean absolute error (top) and $R^2$ scores (bottom) of a Ridge model on various noised data sets. The $\sigma$ values indicated correspond to the standard deviation of the Gaussian noise distributions added to the coupling RDTs data. Each curve represents a noise level applied to arc BPMs data, while each point on these curves corresponds to a noise level applied to the inner BPMs.

The standard deviation of the added noise was determined by a statistical analysis of several measurements from the LHC Run 2. Figure 2 shows the standard deviation of coupling RDTs across Beam 1 BPMs for a batch of measurements taken on April 3, 2018. After analyzing several such batches, the following noise levels were determined:

- Coupling RDTs at arc BPMs were noised with standard deviation ranging from 0 to $10^{-5}$ absolute error.
- Inner BPMs (number 1 to 6 from IP) were noised with standard deviation ranging from 0 to $10^{-2}$.

A new data set was created for each combination of the noise levels mentioned above. Figure 3 shows the test performance of a Ridge Regression model [10] on noised data sets depending on the level of noise added to different BPMs, where the impact of noising the reconstructed coupling RDTs

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOPT008



is noticeable. Here the Mean Absolute Error (MAE) - the sum of absolute errors divided by the sample size - was normalized to the standard deviation of the applied tilts, $\sigma_{\mathrm{DPSI}} = 10^{-4}$ rad.

## RESULTS

Several models suited for regression tasks were tested, and a minimal amount of hyper-parameter tuning was performed. A simple least squares linear regression [11] showed very good results on clean data, however the introduction of noise in the data sets made its performance drop drastically, down to unusable accuracy. A decision tree regressor [12] and a random forest regressor [13]showed poor performance on all data sets. A Ridge regressor model, a linear regressor with $l^2$ regularization, showcased good performance on both clean and relatively low-noised data sets.

Results showing the best $R^2$ scores obtained by each model on both clean and noised test data sets, averaged over 1000 simulations, are presented in Table 1 where the Ridge regressor clearly outperforms its counterparts. In this table, the standard deviations of the applied noise were $\sigma = 10^{-4}$ for IR BPMs and $\sigma = 10^{-5}$ for arc BPMs.

Table 1: $R^2$ Scores of Different Tested Models

| **Model** | Clean Data | Noised Data |
|---|---|---|
| Ridge Regressor | 0.9911 | 0.8934 |
| Linear Regression | 0.9913 | 0.5638 |
| Decision Tree Regressor | 0.1385 | -0.0018 |
| Random Forest Regressor | 0.0175 | -0.0009 |

Figure 4 shows the Ridge model's predictions on a sample from the same noised data set, where a good agreement between predicted and assigned values can be observed. Performance significantly degrades with the addition of noise. Figure 5 shows the predictions and deviations of the Ridge

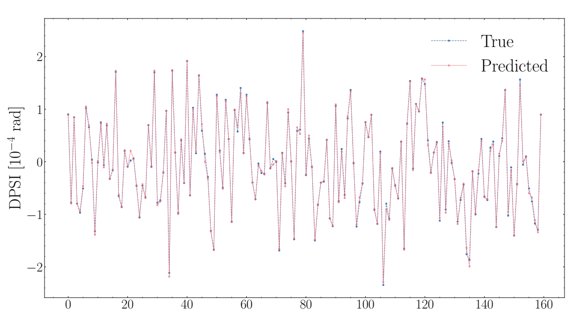


Figure 4: Assigned and predicted values DPSI with a Ridge model. Here the magnet names have been switched for numbers in order to improve the plot's clarity.

model on a noised data set ($\sigma_{\mathrm{DPSI,IRs}} = \sigma_{\mathrm{Arcs,IRs}} = 10^{-5}$, where one can notice the deviations significantly lower that the attributed errors.

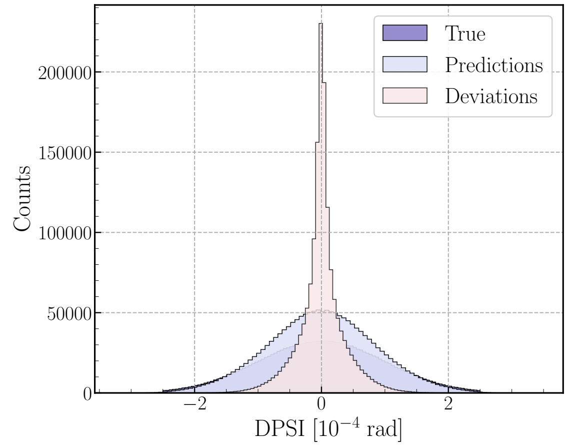


Figure 5: Histograms of the true applied DPSI values, the values predicted by the Ridge model, and the deviations from the predictions to the true values.

## FUTURE PROSPECTS

Assigning more computing resources to the determination of model parameters through hyper-parameter tuning is a first step towards improving model performance, but will not circumvent certain models' shortcomings. Another avenue of improvements would be to create a pipeline where first a denoising step is applied on the coupling RDTs data using an auto-encoder neural network [14] before feeding the results to the above prediction models. Convolutional Neural Networks (CNNs) have in the past been successfully used with impressive success on regression tasks - such as in high energy physics [15] and recently in optics measurements studies [16] - and would be a promising tool that could yield better prediction accuracy. These could give the potential to also include measurements with different machine configurations breaking the degeneracy of the IRs, providing additional insights on the local errors.

## CONCLUSIONS

We have shown that specific machine learning models are capable of predicting the IR quadrupole tilts in the LHC by assigning a representative value to specific magnets, some even when faced with noised data sets. A Ridge regressor shows the best performance among the tested models, including data sets with small amounts of noise. While usability in operation would require better accuracy on data sets with higher noise, this is an important first step towards the application of ML techniques to local coupling corrections in particle accelerators. Potential improvements such as using previously successful but more complex models and workflows have been identified which could allow to improve the performance of models discussed in this study.

## ACKNOWLEDGEMENTS

This work was supported by the STFC Liverpool Centre for Doctoral Training on Data Intensive Science (LIV.DAT) under grant agreement ST/P006752/1.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-WEPOPT008

## REFERENCES


[1] E. Fol, F. S. Carlier, J. M. C. de Portugal, A. García-Tabares, and R. Tomás, "Machine Learning Methods for Optics Measurements and Corrections at LHC," in *Proc. IPAC'18*, Vancouver, Canada, Apr.-May 2018, pp. 1967–1970, doi:10.18429/JACoW-IPAC2018-WEPAF062

[2] A. L. Edelen, S. G. Biedron, B. E. Chase, D. Edstrom, S. V. Milton, and P. Stabile, "Neural networks for modeling and control of particle accelerators," *IEEE Trans. Nucl. Sci.*, vol. 63, no. 2, pp. 878–897, 2016, doi:10.1109/TNS.2016.2543203

[3] E. Bozoki and A. Friedman, "Neural networks and orbit control in accelerators," 1994, https://cds.cern.ch/record/922258

[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, pp. 2837–2839, https://jacow.org/IPAC2012/papers/WEPPP057.pdf

[5] Y. Kijima, M. Mizota, K. Yoshida, and K. Suzuki, "A Beam Diagnostic System for Accelerator using Neural Networks," 1992, https://cds.cern.ch/record/920609

[6] E. Fol, R. Tomás, J. Coello de Portugal, and G. Franchetti, "Detection of faulty beam position monitors using unsupervised learning," *Phys. Rev. Accel. Beams*, vol. 23, p. 102 805, 10 2020, doi:10.1103/PhysRevAccelBeams.23.102805

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

[8] H. Grote and F. C. Iselin, *The MAD-X Program*, 2022, https://cern.ch/madx/releases/last-rel/madxuguide.pdf

[9] R. Calaga, R. Tomás, and A. Franchi, "Betatron Coupling: Merging Hamiltonian and Matrix Approaches," *Phys. Rev. ST Accel. Beams*, vol. 8, p. 034 001, 3 2005, doi:10.1103/PhysRevSTAB.8.034001

[10] R. M. Rifkin and R. A. Lippert, "Notes on Regularized Least-Squares," Computer Science and Artificial Intelligence Laboratory, Tech. Rep. MIT-CSAIL-TR-2007-025, CBCL-268, 2007.

[11] T. L. Lai and H. Robbins, "Strong consistency of least-squares estimates in regression models," 7, vol. 74, 1977, pp. 2667–2669, doi:10.1073/pnas.74.7.2667

[12] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification And Regression Trees*. Routledge, 2017, doi:10.1201/9781315139470

[13] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi:10.1023/a:1010933404324

[14] E. Fol and R. T. García, "Denoising of Optics Measurements Using Autoencoder Neural Networks," in *Proc. IPAC'21*, Campinas, Brazil, May 2021, pp. 3915–3918, doi:10.18429/JACoW-IPAC2021-THPAB068

[15] A. Aurisano *et al.*, "A convolutional neural network neutrino event classifier," *J. Instrum.*, 2016, doi:10.1088/1748-0221/11/09/p09001

[16] E. Fol, J. M. C. de Portugal, G. Franchetti, and R. Tomás, "Optics Corrections Using Machine Learning in the LHC," in *Proc. IPAC'19*, Melbourne, Australia, May 2019, pp. 3990–3993, doi:10.18429/JACoW-IPAC2019-THPRB077