MACHINE LEARNING TECHNIQUES FOR DESIGN OF COMPLEX ACCELERATOR MAGNETS*

S. Gresty[1†], T. Gallagher[1], A. Wolski[1], University of Liverpool, Liverpool, United Kingdom
B. Muratori[1], ASTeC STFC Daresbury Laboratory, Daresbury, Warrington, United Kingdom
[1]also at Cockroft Institute, Daresbury, United Kingdom

Abstract

The design of multipole and other magnets for accelerators is typically an iterative process in which the magnet geometry is optimised for the required beam dynamics properties. Modelling the field for a given geometry can be computationally expensive, so exploring the parameter space can be a time-consuming procedure. The task is particularly challenging when complex field properties are needed (for example, in magnets with several multipole components or with longitudinal field variation). Combined function magnets with several multipole components are particularly useful in accelerators with tight spatial constraints such as an X-ray Free Electron Laser (XFEL). Surrogate models using neural networks can provide a way of rapidly generating possible magnet geometries for given field or beam dynamics requirements. In this contribution, we discuss how machine learning tools may be used to improve the efficiency of the design process for complex accelerator magnets, and present results from a case study based on a combined function magnet for the beam spreader in a future XFEL.

INTRODUCTION

Accelerator magnets generally need to meet demanding field quality specifications if the accelerator is to achieve its performance goals; as a result, design studies for a new accelerator project will usually include a significant amount of magnet modelling work. The required beam properties, specified during the accelerator design process, are used to determine the field properties and field quality needed in each magnet or family of magnets. Modelling of the magnets is then commonly performed using computer codes based on finite element analysis (FEA), for example [1, 2] or boundary integral methods [3]: the design work proceeds iteratively until the desired field quality is achieved. Although numerical computer codes are capable of good accuracy in modelling the field from a given magnet design, calculation of the field is often computationally expensive, and if many design iterations are needed, achieving the required specifications can take a considerable amount of time.

In this paper, we present a technique to assist with accelerator magnet design based on the use of a machine learning (ML) surrogate model to predict the geometry of a magnet given a set of required multipole components. Surrogate models have been shown to be useful for optimisation problems in accelerator research, including for separate components (e.g. [4–6]) and for systems of many components (e.g. [7–11]). Machine learning has also been used to model hysteresis in magnets to assist accelerator tuning [12], but so far there has been little work on surrogate models for magnet design.

One previous study in this area [13] has demonstrated the ability to use machine learning for multipole magnet design in the case of "idealised" magnets in which the material in the core of the magnet is assumed to have infinite permeability and no field saturation limit. In this case, field calculations are simplified since the pole face lies on a surface of constant scalar potential. In these proceedings, we show that a similar machine learning approach can be applied to normal-conducting electromagnets with material in the core having realistic magnetic properties: we show an example based on permeability and field saturation limits typical of iron commonly used in accelerator magnets. The technique aims to increase the efficiency of magnet design work, especially in cases where field specifications may be complex (involving, for example, several multipole components) and where different parameter spaces need to be explored.

Although the use of machine learning tools means that some computational time is needed to generate training data for the surrogate model, generation of the training data is more readily automated than the iterative design process: once a surrogate model is established, it can be used to explore a wide range of design parameters without the need to recompute the field (using a numerical field solver) with each small change in the design. With further development of the technique, surrogate models could be trained to estimate magnet designs directly from beam dynamics specifications (e.g. in terms of required transfer functions). In principle, the techniques may be extended to other magnet technologies, e.g. permanent magnets.

The use of machine learning tools to improve the efficiency of magnet design could prove especially useful for magnets that are not simply conventional multipoles with essentially just a single multipole field component. Magnets with a single component are relatively straightforward, and design work for a magnet for a specific application can often be informed by existing devices. However, modern accelerators are increasingly using complex magnets, including magnets with several multipole field components and with strengths that vary along the length of a magnet: one example is the use of "longitudinal gradient dipoles" to help reduce the beam emittance in ultra-low emittance light sources [14–16]. Beam spreaders in high-power X-ray Free

* This work is supported by Science and Technology Facilities Council, UK, through a grant awarded to The Cockcroft Institute
† sophie.gresty@gmail.com

Electron Lasers also present challenges for magnet design, especially in the vicinity of beam separation points where strict control of the beam optics must be maintained in two closely-spaced beamlines [17–19]: this leads to congested areas where combining separate magnets into a single device can ease the engineering requirements. Here, we consider the example of a magnet combining quadrupole and sextupole field components in a design for the beam spreader in a future UK XFEL [20]. As well as easing space constraints, combining discrete magnets into a single device can help to limit the crosstalk between closely positioned components.

IDEALISED AND REAL MULTIPOLE MAGNETS

Neglecting end effects, so that there is no dependence of the field on longitudinal position in the magnet, the potential ϕ in a multipole magnet can be written:

$$\phi(r,\theta) = \sum_{n=-N}^{N} a_n r^{|n|} e^{in\theta}, \tag{1}$$

where a_n is the n^{th} multipole component, with $n = 1$ the dipole component, $n = 2$ the quadrupole component etc., and r and θ are the radial and azimuthal coordinates in a cylindrical polar coordinate system. For normal multipoles, $a_{-n} = -a_n$, and for skew multipoles $a_{-n} = a_n$, with a_n real in each case. In principle, the limits $\pm N$ in the summation in (1) extend to $\pm\infty$. The magnetic field components are:

$$\mathbf{B}_{\text{skew}} + i\mathbf{B}_{\text{norm}} = \nabla\phi, \tag{2}$$

where $\mathbf{B}_{\text{norm}}$ and $\mathbf{B}_{\text{skew}}$ are superpositions of normal and skew multipole field components, respectively.

In an idealised (theoretical) case, where the material in the core of a normal-conducting electromagnetic multipole has infinite permeability and there is no field saturation limit, the pole faces lie on surfaces of constant magnetic scalar potential (equipotential surfaces). In that case, the shape of the pole faces needed to generate a specified set of multipole components can be determined simply by constructing the appropriate equipotential surface. The radial coordinate r can be found as a function of θ, such that $\phi(r,\theta) = \pm\phi_{\text{ref}}$, where ϕ_{ref} is a fixed "reference" potential, and the sign alternates between adjacent poles. Assuming that the multipole field strengths are specified, the aperture of the magnet is fixed by the value chosen for ϕ_{ref}.

In practice, a number of effects lead to changes in the geometry of the pole faces needed to generate a given set of multipole components, so that the pole faces are not simply equipotential surfaces. The magnetic properties of the iron in the core of the magnet are particularly important: the permeability of the iron may be large, but will certainly not be infinite; and the field will approach saturation as the field strength increases. The finite extent of the poles is also an important factor. The magnet design process requires careful optimisation of the geometry of the pole faces to achieve the specified strengths for the required multipole components, while minimising the strengths of other components. Typically, the field with a given geometry is computed using a numerical field solver; small adjustments are then made to the geometry to try to improve the field quality (the match between the multipole strengths predicted by the numerical code and the specified or required multipole field strengths). This process must be repeated a number of times until a satisfactory field quality is achieved.

It is worth commenting that given the field in a multipole magnet, the multipole components can be determined from the field on the surface of a cylinder of given radius with axis along the axis of the magnet. Choosing a large "reference radius" r_{ref} for the cylinder improves the accuracy of the calculation of the multipole strengths; but the radius must be within the aperture of the magnet. In principle, either the radial or the azimuthal component of the field can be used to calculate the multipole coefficients. Given the expression for the azimuthal component (for example):

$$B_\theta = \frac{1}{r}\frac{\partial\phi}{\partial\theta} = \sum_{n=-N}^{N} ina_n r^{|n|-1} e^{in\theta}, \tag{3}$$

it can be seen that the multipole coefficients a_n can be found by applying a Fourier transform:

$$a_n = -\frac{i}{2\pi n r_{\text{ref}}^{|n|-1}} \int_0^{2\pi} B_\theta e^{-in\theta}\, d\theta. \tag{4}$$

APPLICATION OF MACHINE LEARNING TO MAGNET DESIGN

Modelling the field in a magnet with a given geometry can be computationally expensive. This can reduce the efficiency of the magnet design process, especially in the case of complex field requirements where it may be desirable to explore a range of options with (for example) different magnet lengths, apertures, and pole widths. The availability of a technique allowing rapid computation of the required pole face geometry for given multipole field components (i.e. solving the "inverse" problem to computing the field components from a given geometry) for magnets with realistic material properties could help increase the efficiency of the design process.

One way of addressing the problem is to construct a surrogate model, based on a neural network trained to relate the multipole components in a magnet with given geometry and material properties (permeability and saturation limit) to the multipole components in a magnet with the same geometry, but with idealised properties (infinite permeability and no saturation limits). Given a set of required specific multipole field components, the corresponding idealised components can immediately be obtained from the surrogate model, and the necessary pole face geometry is then readily obtained as an equipotential surface corresponding to the idealised multipole components. In principle, this technique can be applied to a wide variety of cases.

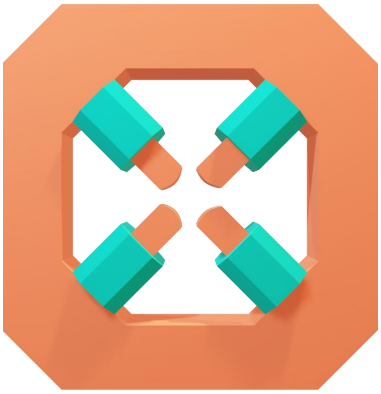


Figure 1: Example multipole magnet model for combined function quadrupole-sextupole. The iron yoke and poles are shown in orange and the coils are shown in teal.

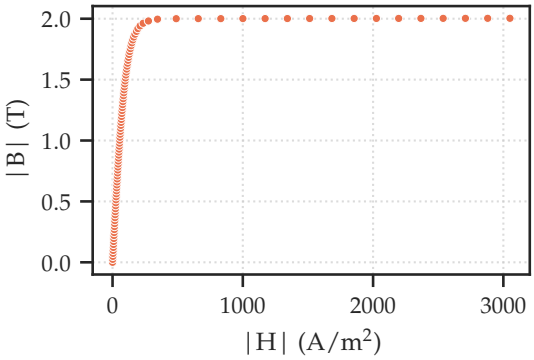


Figure 2: B-H curve for the iron used to construct the yoke and poles for magnet models in Radia.

UK-XFEL CASE STUDY

To provide training data for the surrogate model, the magnetics code Radia [3] was used to simulate over twenty-five thousand cases with varying multipole components within a target range based on the UK XFEL combined quadrupole-sextupole parameters. Each case was constructed by generating (randomly) a set of idealised multipole field components, i.e. a set of coefficients $a_{n,\text{ideal}}$ from $n = 2$ (quadrupole) up to $n = 4$ (octupole). An equipotential surface was computed for these field components and used to specify the geometry of the pole faces in the Radia model. After solving the field, the "true" coefficients $a_{n,\text{true}}$ were computed using (4). The coefficients $a_{n,\text{true}}$ are used as the input to a neural network, which provides the coefficients $a_{n,\text{ideal}}$ (from which the required pole face geometry can readily be found) as output. The aperture, pole width and magnetic properties of the iron in the core are kept constant across all cases, although in principle, these quantities can be varied and included as additional variables in the surrogate model. We have not so far implemented the potential additional degrees of freedom, but plan to do so in a future step.

A typical magnet model is shown in Fig. 1. The magnetic properties of the iron in the magnet (B-H curve) are shown in Fig. 2.

The neural network relating the multipole coefficients $a_{n,\text{true}}$ (inputs) to $a_{n,\text{ideal}}$ (outputs) contains a single hidden layer, with 88 nodes; a nonlinear activation function (leaky relu) is used to allow for modelling the nonlinear relationship between the input and output values (arising from the magnetic properties of the iron in the magnet). The training hyperparameters are given in Table 1. The architecture and hyperparameters were optimised using a Bayesian algorithm in Optuna [21]. Training data were scaled to have zero mean and unit standard deviation, to optimise training.

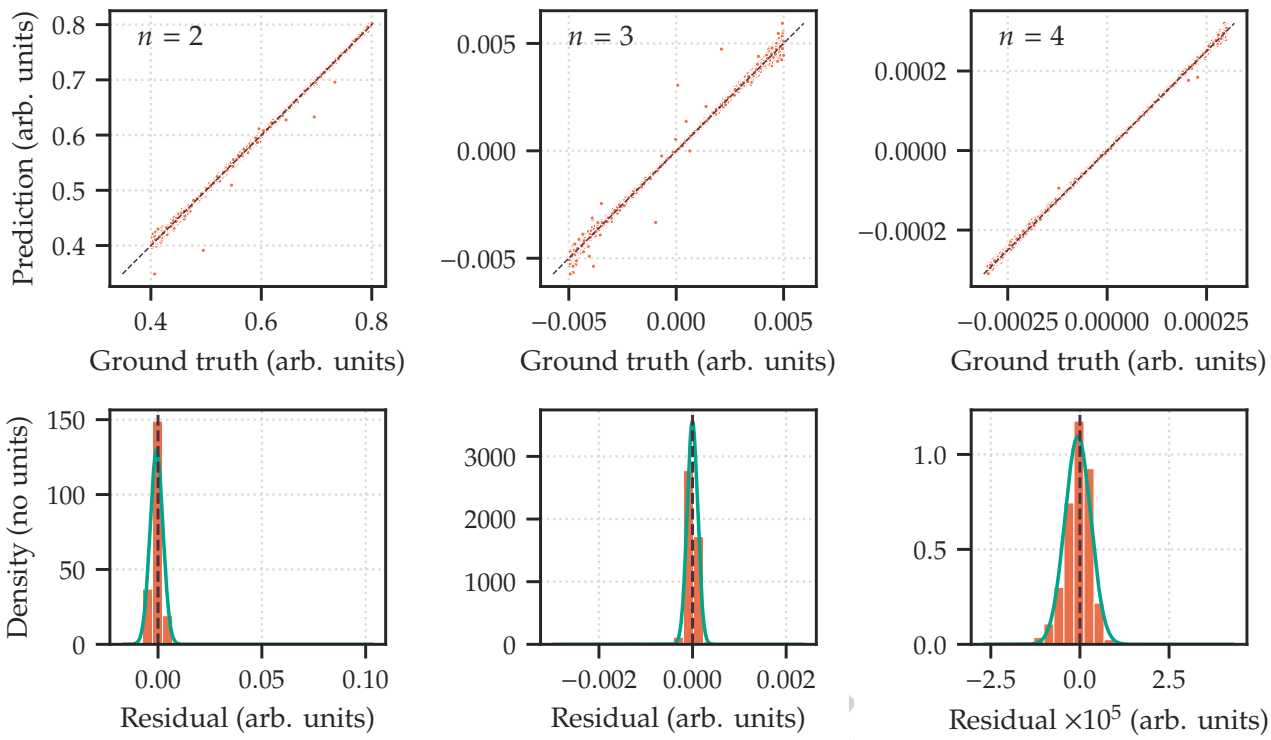
The upper row of plots in Fig. 3 shows results from training the model: values of the $a_{n,\text{ideal}}$ coefficients predicted by the model are plotted against the ground truth values for each multipole component for a set of unseen validation data. Distributions of the residuals (difference between predicted and ground truth values) are shown in the lower row of plots in the figure. Dashed lines on each plot show the ideal correlation (zero residual). Note that the values of $a_{n,\text{ideal}}$ (ground truth and predicted) shown in Fig. 3 are used simply to specify an equipotential surface that defines the geometry of the pole faces in a Radia magnet model; the true multipole components (the coefficients $a_{n,\text{true}}$ provided as input to the machine learning model) depend on the (nonlinear) magnetic properties of the iron and the current in the coils in the magnet model.

The accuracy of the predictions from the ML model is indicated by the mean of the absolute values of the residuals shown in Fig. 3 (lower row of plots): the mean absolute residual values are 0.14%, 0.10%, and 3.11% for the quadrupole, sextupole, and octupole components, respectively.

Table 2 shows a comparison between the target (required) values for the multipole components in a combined quadrupole-sextupole, and the field components computed in Radia for two different pole face geometries and for three different nominal lengths. In the ML cases, the required multipole coefficients $a_{n,\text{true}}$ are given as inputs to the trained machine learning model, and the pole face geometry is constructed from the coefficients $a_{n,\text{ideal}}$ obtained as outputs (by using these coefficients to define equipotential surfaces). In the idealised cases in Table 2, the required multipole coefficients $a_{n,\text{true}}$ are used directly to construct the pole faces as equipotential surfaces. Note that in the idealised cases, the current in the magnet model is adjusted so that the quadrupole gradient in the model matches the required

Table 1: Training hyperparameters for the neural network relating the true multipole components in a magnet with nonlinear magnetic properties to the components in an idealised magnet with the same pole face geometry.

Learning rate	0.000694
Batch size	32
Dropout rate	0.00516
Optimiser	adam
L2 regulariser	0.0000944
Loss function	Huber (δ = 8.69)



(a) Quadrupole (b) Sextupole (c) Octupole

Figure 3: Upper row: comparison of predicted multipole components with ground truth values for the neural network validation data. Lower row: distributions of residuals in the validation data, where the residual is the difference between the predicted value of the multipole component and the ground truth.

Table 2: Integrated field gradients for multipole components ($n = 1$ for dipole, $n = 2$ quadrupole, etc.) in a combined function multipole magnet. The idealised case has pole face geometry based on an equipotential surface. In the ML case the geometry is predicted by a machine learning model. For both cases, the field is computed (using `Radia`) with realistic magnetic properties for the iron core. The same ML model is used for each of the different lengths.

		Integrated field gradient (Tm^{2-n})					
		$L = 0.5$ m		$L = 0.455$ m		$L = 0.333$ m	
Multipole order, n	**Target**	**Idealised**	**ML**	**Idealised**	**ML**	**Idealised**	**ML**
1	0	0.01	0.00	0.01	0.00	0.01	0.00
2	11.71	11.71	11.83	11.71	11.64	11.71	11.73
3	−91.33	−260.17	−90.69	−255.37	−90.56	−200.20	−94.08
4	0	3277.44	294.65	3848.51	294.81	3820.87	−254.45

gradient. In the ML case, no separate scaling is needed, since the scaling is "built into" the machine learning model (the training data are obtained by solving a model with given current in the coils). Figure 4 shows a comparison between the pole face geometries in the idealised and the ML cases for 0.5 m magnet length.

Since it is the integrated multipole field that is significant for the beam dynamics in an accelerator, the specifications for the UK XFEL combined quadrupole-sextupole magnet are given in terms of the integrated field strengths rather than the strengths of the multipole components themselves. In the magnet design work, the target strengths for the field components will scale in inverse proportion to the length of the magnet. Due to the nonlinear properties of the iron, each field component may have a different scaling with the current in the coil, for fixed pole face geometry. This means that different lengths of magnet will require different pole face geometries to achieve multipole field components in a fixed ratio. Use of the machine learning (surrogate) model, which is trained with data that is independent of magnet length, allows rapid exploration of the effects of changes in magnet length, without the need to manually adjust the pole face geometry (with some iteration) with each change. Table 2 shows results for three different lengths of magnet (0.5 m, 0.455 m and 0.333 m); the integrated target field is the same in each case. The same machine learning model is used to generate the pole face geometry for each length.

Although the quadrupole components in the idealised cases matches the required value (because the current is tuned to meet this criterion) the sextupole components are

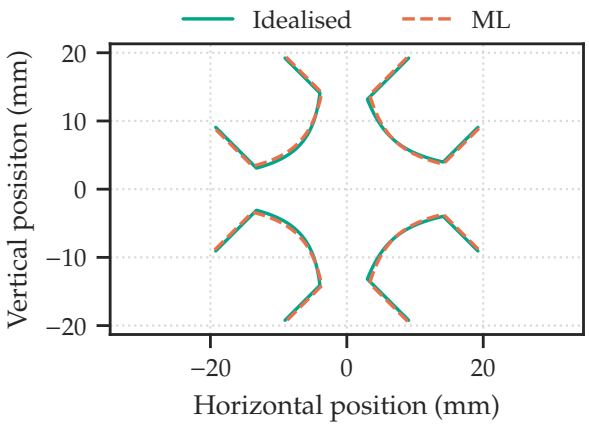


Figure 4: Pole face geometries for an idealised magnet (solid line) and a magnet with realistic magnetic properties (dashed line) for the UK XFEL combined quadrupole-sextupole.

between two and three times larger than the required value. There is also a significant octupole component. In the ML cases, no adjustment of the current is made, since the coil current is essentially "built into" the machine learning model. Because the model is not perfectly accurate, the quadrupole component in each of the ML cases does not match exactly the target value; this could be corrected by making some small adjustment to the current. The sextupole components in the ML cases are close to the target values, and although the octupole components are not zero, they are an order of magnitude smaller than the components in the idealised cases. The machine learning model provides a close approximation to the required pole face geometry for each of the three magnet lengths in Table 2; final optimisation can be performed manually.

SUMMARY AND CONCLUSIONS

The results from the surrogate model demonstrate proof of principle for the application of machine learning techniques to the design of magnets in accelerators. The model is trained to relate a set of specified multipole components to a set representing (in the form of an equipotential surface) the pole face geometry required to achieve the specified components in a device using an iron core with given magnetic properties. This provides a good first approximation for the geometry, which can then be further refined using standard magnetic modelling codes. Although some computational time is needed to generate training data, this initial step can easily be automated, and the trained model can then be used for rapid exploration of a range of design variations. This can be useful when designing complex magnets requiring the superposition of several multipole components: we have reported the specific example of a combined quadrupole-sextupole magnet for the beam spreader of the UK XFEL.

It should be emphasised that the machine learning technique that we have described is not intended as a complete replacement for the use of conventional magnet modelling tools (such as computer codes based on FEA or boundary integral methods). Although the ML technique can produce a good approximation to an appropriate magnet design, and further improvements in performance may be possible (e.g. through use of more advanced architectures), some final optimisation of the design will almost certainly be needed, based on expert use of a magnet modelling code.

So far, we have assumed fixed magnetic properties (permeability and saturation) for the iron core of the magnet, and fixed values of some parameters defining the geometry, such as the aperture and nominal pole width. In principle, the technique can be extended to include variations in these quantities, as well as in the multipole components. This will be investigated in future studies.

The work up to now has focused on 2D multipoles, neglecting variations in multipole strengths along the length of a magnet. However, 3D multipoles are of significant practical interest, for example in the use of longitudinal gradient dipoles in ultra-low emittance electron storage rings. In principle, the technique we have described can be extended to 3D magnetic fields: this will allow the inclusion also of end effects (fringe fields). 3D fields were considered for idealised magnets in [13]. Data sets for the 3D case (i.e. 3D multipole coefficients) are significantly larger than for the 2D case, especially when a detailed description of the longitudinal variation is needed. Although there will be greater demands in terms of computational resources (in particular, in generating training data) there should not be any fundamental obstacles, and we aim to explore the generalisation of our technique to 3D fields in future work.

Finally, there may be significant benefits in using machine learning models to determine the pole face geometry needed to meet specific beam dynamics requirements directly from an appropriate description of the beam dynamics (e.g. in the form of a transfer map). Although this problem will introduce additional complexity, there should again be no fundamental obstacle, and is something that we also hope to investigate in future.

ACKNOWLEDGEMENTS

We would like to thank A. Potter (Diamond Light Source) for specifications and information on the combined function quadrupole-sextupole magnet for the UK XFEL beam spreader. We would also like to thank A. Pollard (ASTeC) for computational resources.

REFERENCES

[1] CST Studio Suite, https://www.3ds.com/products/simulia/cst-studio-suite

[2] Comsol Multiphysics, https://www.comsol.com/

[3] Radia, https://www.esrf.fr/home/Accelerators/instrumentation--equipment/Software/Radia.html

[4] D. Koser, J. M. Conrad, L. H. Waites, D. Winklehner, A. Adelmann, M. Frey and S. Mayani, "RFQ beam dynamics optimization using machine learning", in *Proc. IPAC'21*, Campinas, SP, Brazil, May 2021, pp. 3100–3102. doi:10.18429/JACoW-IPAC2021-WEPAB203

[5] D. Koser, L. Waites, D. Winklehner, M. Frey, A. Adelmann, and J. Conrad, "Input beam matching and beam dynamics design optimizations of the IsoDAR RFQ using statistical and machine learning techniques", *Frontiers in Physics*, vol. 10, 2022. doi:10.3389/fphy.2022.875889

[6] J. Villarreal, D. Winklehner, D. Koser and J. M. Conrad, "Neural networks as effective surrogate models of radio-frequency quadrupole particle accelerator simulations", Jun. 2024, arXiv:2210.11451. doi:10.48550/arXiv.2210.11451

[7] A. Edelen, N. Neveu, M. Frey, Y. Huber, C. Mayes and A. Adelmann, "Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems", *Phys. Rev. Accel. Beams*, vol. 23, no. 4, p. 044601, 2020. doi:10.1103/PhysRevAccelBeams.23.044601

[8] 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 Transactions on Nuclear Science*, vol. 63, no. 2, pp. 878–897, 2016. doi:10.1109/TNS.2016.2543203

[9] K. Sun, X. Chen, X. Zhao, X. Qi, Z. Wang, and Y. He, "Surrogate model of particle accelerators using encoder-decoder neural networks with physical regularization", *International Journal of Modern Physics A*, vol. 38, no. 26–27, p. 2350145, 2023. doi:10.1142/S0217751X23501452

[10] F. M. Velotti, B. Goddard, V. Kain, R. Ramjiawan, G. Z. D. Porta, and S. Hirlaender, "Towards automatic setup of 18 MeV electron beamline using machine learning", *Machine Learning: Science and Technology*, vol. 4, no. 2, 2023. doi:10.1088/2632-2153/acce21

[11] S. Chauhan, A. Edelen, C. Emma and S. Gessner, "Modeling and optimization of the FACET-II injector with machine learning algorithms", in *Proc. IPAC'24*, Nashville, TN, USA, May 2024, pp. 913–916. doi:10.18429/JACoW-IPAC2024-MOPS79

[12] R. Roussel, A. Edelen, D. Ratner, K. Dubey, J. P. Gonzalez-Aguilera, Y. K. Kim and N. Kuklev, "Differentiable Preisach modeling for characterization and optimization of particle accelerator systems with hysteresis", *Phys. Rev. Lett.*, vol. 128, no. 20, p. 204801, 2022. doi:10.1103/PhysRevLett.128.204801

[13] T. Gallagher, A. Wolski, and B. D. Muratori, "A machine learning approach to shaping magnetic fringe fields for beam dynamics control", *J. Phys.: Conf. Ser.*, vol. 2687, p. 062031, 2024. doi:10.1088/1742-6596/2687/6/062031

[14] J. Guo and T. O. Raubenheimer, "Low Emittance e-/e+ Storage Ring using Bending Magnets with Longitudinal Gradient", in *Proc. EPAC'02*, Paris, France, Jun. 2002, paper WEPLE001, pp. 1136–1138.

[15] C.-X. Wang, "Minimum emittance in storage rings with uniform or nonuniform dipoles", *Physical Review Special Topics - Accelerators and Beams*, vol. 12, no. 6, Jun. 2009. doi:10.1103/physrevstab.12.061001

[16] V. S. Kashikhin *et al.*, "Longitudinal Gradient Dipole Magnet Prototype for APS at ANL", *IEEE Trans. Appl. Supercond.*, vol. 26, no. 4, pp. 1–5, Jun. 2016. doi:10.1109/tasc.2016.2521896

[17] K. Tono, T. Hara, M. Yabashi, and H. Tanaka, "Multiple-beamline operation of SACLA", *J. Synchrotron Radiat.*, vol. 26, no. 2, pp. 595–602, Feb. 2019. doi:10.1107/s1600577519001607

[18] F. Obier, W. Decking, M. Hüning, and J. Wortmann, "Fast kicker system for European XFEL beam distribution," in *Proc. 39th Int. Free Electron Laser Conf. (FEL'19)*, Hamburg, Germany, Aug. 2019, paper WEP013. doi:10.18429/JACoW-FEL2019-WEP013

[19] S. U. De Silva *et al.*, "A compact beam spreader using RF deflecting cavities for the LCLS-II," in *Proc. 5th Int. Particle Accelerator Conf. (IPAC'14)*, Dresden, Germany, Jun. 2014, pp. 2666–2668. doi:10.18429/JACoW-IPAC2014-WEPRI075

[20] UK XFEL, https://www.xfel.ac.uk/

[21] Optuna, https://optuna.org/