Data-Driven Approaches for Electromagnetic Analysis of Traction Electrical Motors: A Proposal for a Benchmark Problem

Data-Driven Approaches for Electromagnetic Analysis of Traction Electrical Motors: a Proposal for a Benchmark Problem

Luigi Solimene *, Simone Ferrari *, Costanza Anerdi *, Fabio Freschi *, Luca Giaccone *,
Gianmarco Lorenti *, Francesco Lucchini†, Riccardo Torchio†
Piergiorgio Alotto †, Gianmario Pellegrino *, Maurizio Repetto *

*Politecnico di Torino, Department of Energy "Galileo Ferraris", Torino, Italy, 10129,
name.surname@polito.it
† Department of Industrial Engineering, University of Padova, Padova, Italy, 35131
name.surname@unipd.it

***Abstract*—The rising demand for high-performance traction electrical motors in electric vehicles necessitates efficient design solutions. This paper describes a preliminary approach to data-driven modelling of the electromagnetic design problem of traction electrical motors. The initial step consists of generating the dataset for training the data-driven models, which is performed using an FEA-based procedure. Subsequently, different data-driven approaches are compared to surrogate the output of the electromagnetic analysis, identified in the computation of the torque and the torque ripple of a specific electrical motor structure.**

I. Introduction

The need for improved performance in traction electrical motors continues to rise alongside the increasing spread in size and performance of electric vehicles. The design of traction electrical motors has to deal with strict requirements on efficiency, maximum torque and speed, and cost reduction. To effectively tackle this design task, a multi-physical approach is often required, taking into account the interaction between the electromagnetic, thermal, and structural domains. However, the multi-physical analysis for the motor design problem makes extensive use of Finite Element Analysis (FEA) codes, yielding non-negligible computational costs. Moreover, since the design goals are often contrasting, as for the torque maximization versus the power loss and overtemperature reduction, a multi-objective optimization approach is also required [1].

A *pre-design* approach is often useful to size a tentative motor meeting the design requests. As usual, a balance between the accuracy of results and the time needed to get them is requested, being in this phase the scale more in favour of quicker answers allowing to explore more diverse solutions and assess them comparatively.

To this aim, in recent years, the possibility of solving the design problem of electrical motors using data-driven approaches has emerged, reducing the overall computational cost of the design task [2]–[5]. In this paper, we propose a preliminary approach to the data-driven modelling of the electromagnetic design problem of traction electrical motors. The first task consists of the dataset generation for the training of the data-driven models, which is performed by means of a FEA-based procedure. Then, different data-driven approaches are compared to surrogate the output of the electromagnetic analysis, identified in the computation of the torque and the torque ripple of a specific electrical motor structure.

The procedure requested to generate a geometric description of a *motor family* must implement the construction of a motor geometry starting from a set of geometric/electrical degrees of freedom (dof), that is a *decoding* procedure going from the *genotypic* set of variables to the *fenotypic* geometry of the motor. Once the geometry and the supply conditions of the structure have been created, its analysis can be performed by FEA, and thus, a set of results (torque, torque ripple, etc.) can be obtained. This procedure has been implemented in the `SyR-e` [6] collection of scripts in the `Matlab` environment. As this procedure is available, it can be used to extend its analysis capabilities to other physical domains, like the structural, vibrational, and thermal ones.

Starting from these considerations, this work is a first step in the direction of defining a benchmark problem that can allow other researchers to exploit datasets to test their methodologies and procedures for a data-driven analysis of electrical motors.

II. Case Study

The paper focuses on a V-shaped Internal Permanent Magnet (IPM) motor structure since its geometry is described uniquely by well-defined rules, as are its material characteristics. A *family of motors* can be generated by considering the external dimensions of the motor (stator outer radius and axial length) as fixed as well as the power supply conditions and circuit data.

Table I gives the main specifications of the motor, while Figure 1, which describes the two-dimensional cross-section adopted for the electromagnetic simulation, highlights the seven design variables that uniquely define a geometrical configuration.

TABLE I
MOTOR STRUCTURE SPECIFICATIONS

Pole pairs	3
Number of slots	54
Number of barrier	1
Stator diameter	225 (mm)
Stack length	132 (mm)
Number of turns	21
Peak Thermal loading	171 (kW/m^2)

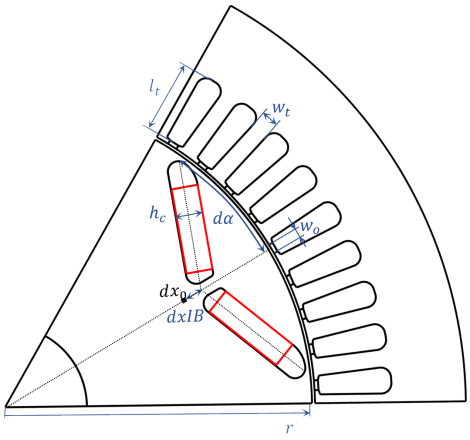


Fig. 1. Two-dimensional cross-section of the considered V-shaped IPM motor structure. The geometrical design variables are here described.

In particular, the motor geometry is described through:

- Rotor radius r
- Barrier width h_c
- Barrier shift $dxIB$
- Barrier position $d\alpha$
- Tooth length l_t
- Tooth width w_t
- Slot opening w_o.

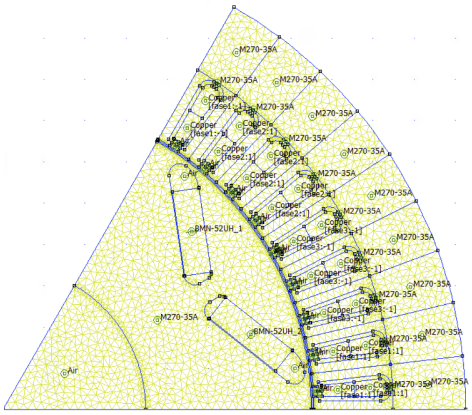
More details on the definition of the motor geometry are available in [7]. In addition to the geometrical variables, a further variable for the motor configuration definition is the current phase angle γ, taken from the q axis. The geometry construction on the basis of the described design variables is performed through the open-source motor design platform `SyR-e` [6]. The variable ranges are reported in Table II. In order to obtain a uniform distribution of the variables in the prescribed range, a hierarchical points sampling ensuring a uniform exploration of the dof space (Sobol strategy [8]) is adopted.

The electromagnetic analysis is performed by means of the `FEMM` software [9]. The following hypotheses are considered:

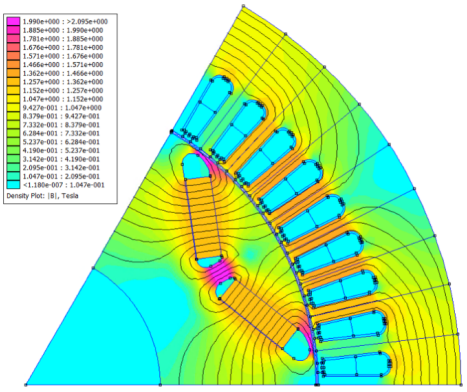
- The nonlinear magnetostatic analysis is based on the two-dimensional mesh, considering the material nonlinearities at a given reference temperature.
- In order to evaluate the torque ripple, several relative positions between the rotor and the stator are considered by magnetostatic analysis (snapshots). In particular, 30 magnetostatic simulations are performed for each motor configuration for a good resolution of the torque ripple behaviour.

TABLE II
VARIABLES SWEEP LOWER AND UPPER RANGES

Rotor radius r	[60 78] (mm)
Barrier width h_c	[0.3 0.7] (pu)
Barrier shift $dxIB$	[-4 6] (mm)
Barrier position $d\alpha$	[0.65 0.85] (pu)
Tooth length l_t	[15 22.5] (mm)
Tooth width w_t	[3.8 6.3] (mm)
Slot opening w_o	[0.1 0.4] (pu)
Current phase angle γ	[30 60] (°)



(a)



(b)

Fig. 2. (a) Triangular mesh for the electromagnetic simulation of the motor by means of FEA. (b) The output of the FEA, showing the flux lines and magnetic flux density colour map.

Figure 2 shows an example of the triangular mesh adopted for the electromagnetic finite element analysis and the output of the simulation in terms of the magnetic flux density map.

The required outcomes of the simulations are the torque T and the peak-to-peak torque ripple TR. Following this procedure, 5000 motor configurations are evaluated, providing a suitable dataset for the training of different data-driven surrogate models. The execution of the finite element analyses for the computation of the dataset outputs has required about 10 hours on a virtual machine running on a server, with 16 cores and 128 GB of RAM.

III. Dataset Analysis

Although within the suitable ranges, the combination of the design variables can lead to unfeasible design configurations. The definition of unfeasible configurations can be expressed in terms of minimum admissible magnets and copper volumes in order to achieve non-zero torque values. Thus, a preliminary filtering of the dataset can be performed in order to eliminate outlier configurations. However, on the considered dataset, defined by the variable ranges of Table II, all the design configurations resulted as feasible. The TR output is defined as a relative percentage value referred to T.

A preliminary dataset analysis can be performed by evaluating the correlation between inputs and outputs. Figure 3 describes the values of the Pearson correlation coefficient between the input variables and the outputs, represented by the torque and torque ripple, in addition to the mass of the winding copper (M_{cu}) and of the rotor magnets (M_{mag}). The Pearson correlation coefficients measure the strength of the linear covariance relationship between two variables. It is well known that the relationship between the input design variables and the outputs of the electromagnetic analysis is non-linear, and thus, for the considered case, the Pearson coefficients are not strictly representative. However, they provide a first guess of the effect of the design variables on the performance of the motor.

Two outcomes of the parameter set, copper and magnet masses, directly obtained from the motor geometry, are added to the table. Even if they are not the direct goals of the surrogate models, they highlight the effect of the geometrical variables on the motor configuration. It is apparent from the analysis of the correlation coefficients how the rotor parameters r, h_c, $dxIB$, and $d\alpha$ influence the magnet dimensions, while the stator parameters w_{t}, and l_{t}, together with the rotor radius r, mainly influences the winding dimensions. On the other hand, the analysis of the correlation coefficients for the T and the TR highlights how, for the torque, different design variables as r, w_{t}, $dxIB$, and γ have a remarkable impact on the performance, while for the torque ripple, the correlation with the input variables is less evident. This aspect reflects the possibilities of how more challenging the torque ripple modelling with respect to the torque model will be.

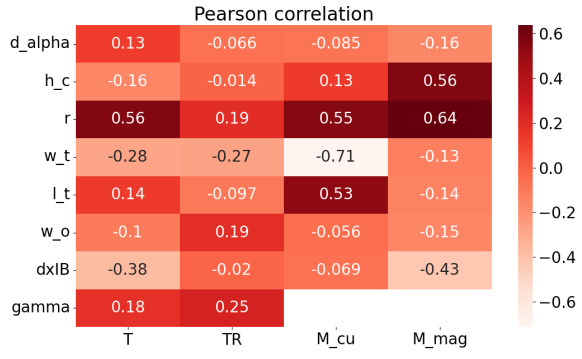


Fig. 3. Pearson correlation parameters between design variables and design outputs. In the design outputs, the copper and magnet masses are also considered. The correlation between the current phase angle and the copper and magnet masses is neglected.

IV. Data-driven surrogate models

In this section, the performances of two widely adopted data-driven regression models are compared. In particular, the tested methodologies are Feed-forward Neural Network (FNN) and Support Vector Regression (SVR). FNN and SVR models are trained to surrogate each of the two outputs.

The dataset for the training of the data-driven models is pre-processed by a random shuffling of data subsequently separated into training, validation, and testing sets (65%, 15%, 20%). Then, the inputs are normalised wrt. the mean and standard deviation of the training dataset, to obtain compatible dimensions for all the input variables.

A. Feed-forward neural network

The FNN model is implemented in Python by adopting the Sequential Network structure in Keras [10]. The neural network structures for the torque and the torque ripple surrogate models have been preliminary determined using a random search optimization to select the number of layers and the number of neurons per layer, leading to the structures defined in Tables III and IV. The number of trainable parameters is 2465 for the torque surrogate model and 761 for the torque ripple one.

TABLE III
FNN structure for torque

	Number of neurons	Activation function
Layer 1	32	Tanh
Layer 2	64	Tanh
Layer 3 (Hidden)	1	Relu

TABLE IV
FNN structure for torque ripple

	Number of neurons	Activation function
Layer 1	24	Tanh
Layer 2	16	Tanh
Layer 3	8	Tanh
Layer 4 (Hidden)	1	Relu

The models are trained on the training set by adopting the Adam algorithm to minimize the loss function, defined as the mean squared error evaluated on the validation set. The maximum number of epochs of the training phase is set at 5000, and a stop criterion is introduced after 200 iterations without improvements on the validation data to reduce overfitting on training samples.

Figures 4, 5 and 6 describe the performances of the trained FNN model for the torque prediction. In particular, Figure 4 shows the behaviour of the loss value as a function of the training iterations, computed on the training and on the validation data. Figure 5 describes the performances on the

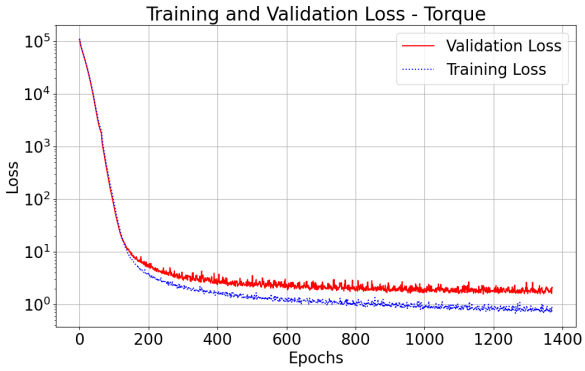


Fig. 4. Behaviour of the loss value as a function of the training iterations, for the torque FNN model.

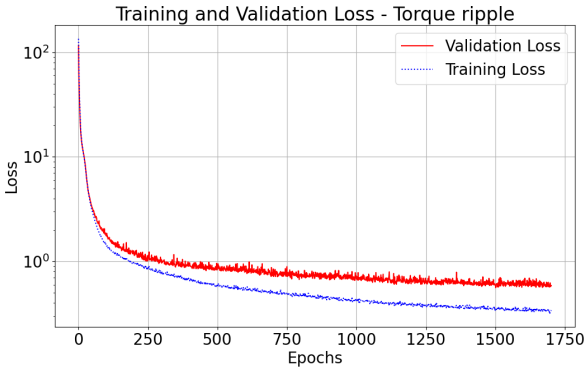


Fig. 7. Behaviour of the loss value as a function of the training iterations, for the torque ripple FNN model.

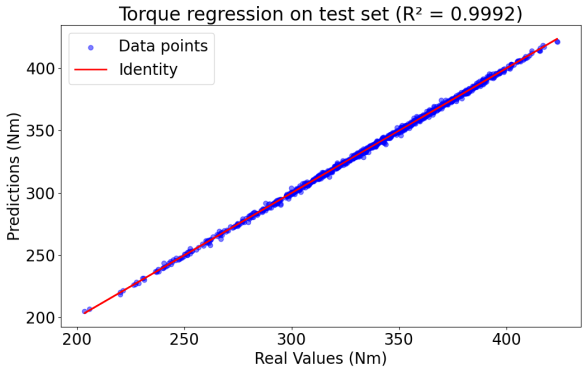


Fig. 5. Correlation between real and predicted torque values, evaluated with the FNN model on the test set.

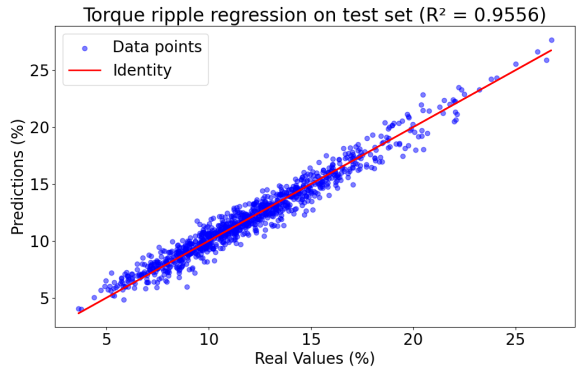


Fig. 8. Correlation between real and predicted torque ripple values, evaluated with the FNN model on the test set.

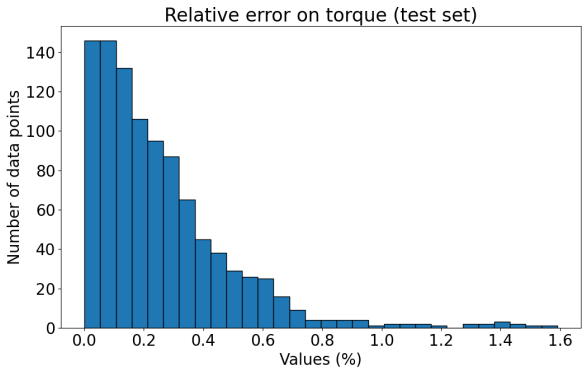


Fig. 6. Relative error between the real and the predicted torque values, evaluated with the FNN model on the test set.

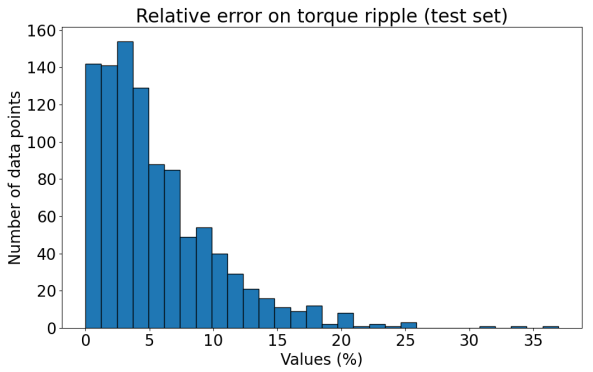


Fig. 9. Relative error between the real and the predicted torque ripple values, evaluated with the FNN model on the test set.

regression of the torque values for the test set, and Figure 6 shows the distribution of the relative error on the torque prediction on the test set. The same results are reported in Figures 7, 8 and 9 for the FNN model for the torque ripple.

B. Support vector regression

The second tested machine learning method for the surrogate modelling of torque and torque ripple of traction electrical motors is the Support Vector Regression (SVR). We used the Python implementation available in the Scikit-learn library to construct the SVR model [11], [12]. The Radial Basis Function (RBF) kernel is adopted to effectively deal with the non-linear behaviour of the design outputs as a function of the input motor design variables.

The hyperparameters of the models are determined through a grid search optimization, aiming to minimize the mean

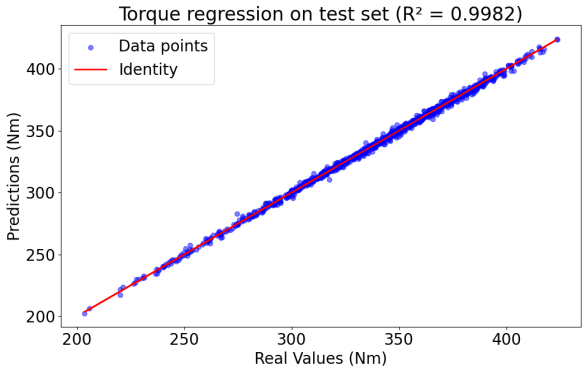


Fig. 10. Correlation between real and predicted torque values, evaluated with the SVR model on the test set.

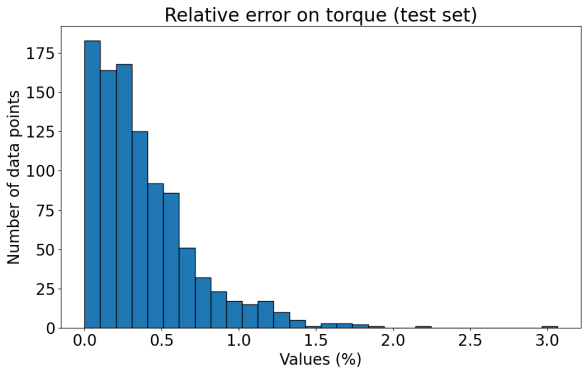


Fig. 11. Relative error between the real and the predicted torque values, evaluated with the SVR model on the test set.

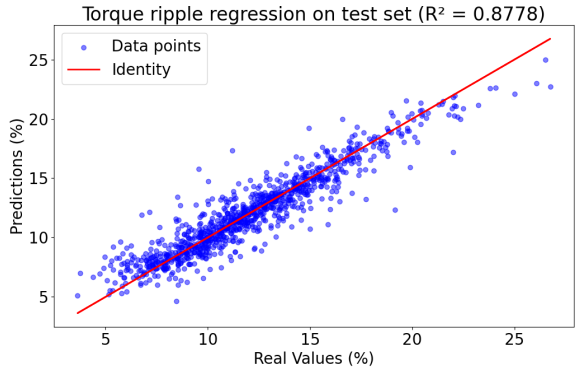


Fig. 12. Correlation between real and predicted torque ripple values, evaluated with the SVR model on the test set.

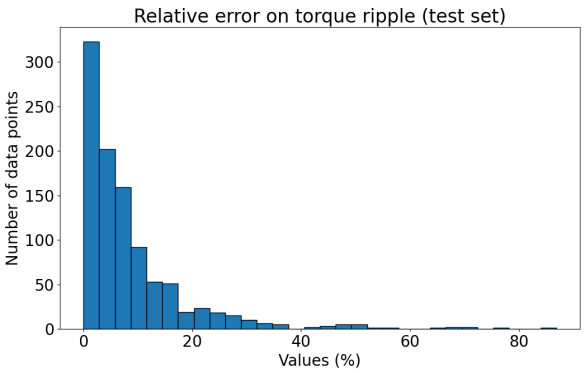


Fig. 13. Relative error between the real and the predicted torque ripple values, evaluated with the SVR model on the test set.

squared error on the validation data. The hyperparameters of the selected models are reported in Table V. The number of trained parameters that define the support vectors for torque and the torque ripple regression are 3172 and 3159, respectively.

TABLE V
SVR MODELS HYPERPARAMETERS

	C	ϵ	γ
Torque model	500	0.02	0.12
Torque ripple model	185	0.014	0.1

Figure 10 describes the performances trained SVR model on the regression of the torque values for the test set, and Figure 11 shows the distribution of the relative error on the torque prediction on the test set. The same results are reported in Figures 12 and 13 for the SVR model for the torque ripple.

V. Discussion and Future Developments

A. Accuracy of the surrogate models

Both regression methodologies show excellent performance in predicting the torque value for a given motor geometry, with extremely reduced errors, as reported in Table VI. While the errors of the torque prediction for the FNN are reduced, the SVR model can rely on a shorter training time.

The prediction of the torque ripple on the basis of the geometrical inputs is a more challenging task. However, acceptable performances are reached from FNN, while the SVR model shows poor accuracy, achieving the error values on the prediction of the torque ripple described in Table VII.

Regarding the computation time required to evaluate the design goals over 5000 data points, the results reported in

TABLE VI
MEAN, 95TH PERCENTILE, AND MAX VALUE OF THE RELATIVE ERROR ON TORQUE PREDICTION

	Mean (%)	**95 perc (%)**	**Max (%)**
FNN	0.26	0.67	1.59
SVR	0.39	1.075	3.07

TABLE VII
MEAN, 95TH PERCENTILE, AND MAX VALUE OF THE RELATIVE ERROR ON TORQUE RIPPLE PREDICTION

	Mean (%)	**95 perc (%)**	**Max (%)**
FNN	5.59	14.98	36.94
SVR	8.60	27.42	86.88

Table VIII are achieved. The computations are performed on an Intel Core i7-10510U CPU @ 1.80GHz, 8 cores, 16 GB machine.

TABLE VIII
COMPUTATION TIME OF THE SURROGATE MODELS OVER 5000 SAMPLES

FNN - T	FNN - TR	SVR - T	SVR - TR
0.3711 (s)	0.3711 (s)	2.7887 (s)	3.1153 (s)

B. Use of the models in multi-objective optimization

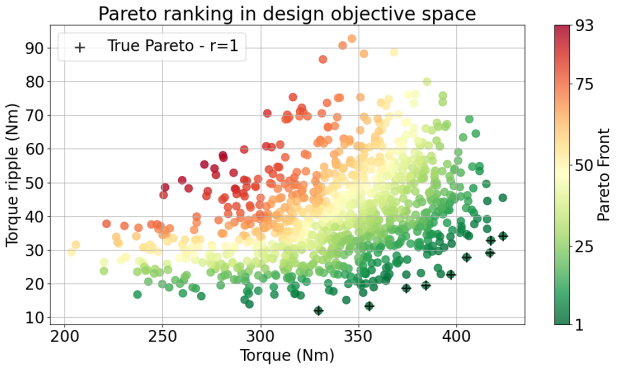
Ultimately, the aim of surrogate modelling of the electromagnetic analysis for traction electrical motors is to adopt the models in the multi-objective optimization of the design problem to reduce the computational effort compared to conventional finite element analyses. While the advantage of surrogate models in the time reduction of electromagnetic simulation is apparent, the accuracy of the proposed results has to be assessed.

In order to evaluate the ability of the tested surrogate model to estimate the optimal configurations of the motor design problem, a comparison of the Pareto Front (PF) determined on the test dataset, aiming to maximize the torque and minimize the related torque ripple, is performed. The points belonging to the test set (1000 points over the whole 5000 belonging to the dataset) are used to estimate the PF. Out of 1000 points, 9 of them are *non-dominated* and describe the PF. As they are computed by FEA, these configurations are assumed to represent the true PF. The motor performance is then computed by the surrogate data-driven models on the same points, and *non-dominated* configurations are assessed for the outputs of each surrogate model. For the FNN and the SVR outputs, the second rank of the PF is also considered, that is the set of non-dominated solutions obtained by removing the ones belonging to the first Pareto rank.

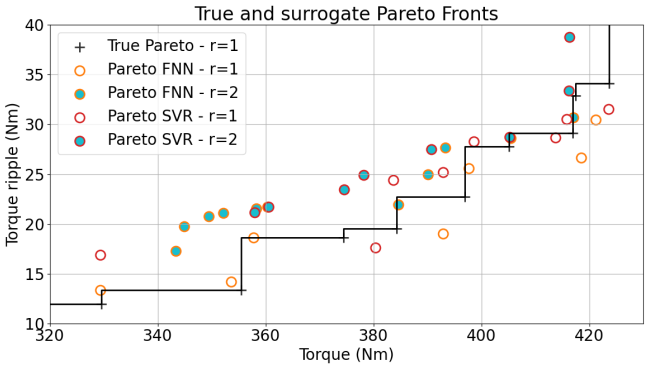
Figure 14a shows the data points of the test set, reported in the space of the two design goals, highlighting the PF determined with the FEA. The colour map indicates the Pareto ranking. Figure 14b, on the other hand, represents the comparison of the true PF with the ones determined on the basis of the prediction of the two surrogate models, considering both first and second rank.

By predicting torque and torque ripple with the FNN and SVR models, 8 and 12 non-dominated configurations are identified in the first rank, respectively. However, only 5 FNN-evaluated and 6 SVR-evaluated design configurations also belong to the Pareto set determined with FEA. As a consequence, limiting the search for the best designs to the non-dominated set of configurations determined with the surrogate models could exclude some interesting designs. Thus, since the goal of the surrogate models is to provide a fast overview of several design configurations to further investigate the most promising with conventional analyses (FEA), the range of relevant solutions can be broadened.

If we also consider the second rank, the surrogate models can identify most configurations of the real Pareto set. In particular, with the FNN model, 8 optimal configurations are identified over 9, while with the SVR model, 7 optimal configurations are identified.



(a)



(b)

Fig. 14. (a) Test data in the design objectives space. The colour indicates a data point belonging to a specific Pareto rank. The true Pareto front obtained by FEA is highlighted. (b) Comparison of the true Pareto front with the ones obtained with the surrogate models, considering both first and second rank.

The evaluation of these preliminary results reveals the convenience of adopting surrogate modelling strategies for electromagnetic analysis. In fact, the surrogate models can be effectively used in the framework of a multi-objective optimization algorithm, allowing for a deep exploration of the design space, given the reduced computational time required for the electromagnetic output evaluations. The conventional FEA approach can be used to assess the effective performances of the sole optimal and sub-optimal configurations determined through surrogate models.

Future developments will be related to the definition of open datasets on which other researchers could test their procedures.

REFERENCES

[1] F. Cupertino, G. Pellegrino, and C. Gerada, "Design of Synchronous Reluctance Motors With Multiobjective Optimization Algorithms," *IEEE Transactions on Industry Applications*, vol. 50, no. 6, pp. 3617–3627, Nov. 2014.

[2] M. H. Mohammadi, V. Ghorbanian, and D. A. Lowther, "A Data-Driven Approach For Design Knowledge Extraction of Synchronous Reluctance Machines Using Multi-Physical Analysis," in *2018 XIII International*

Conference on Electrical Machines (ICEM), Sep. 2018, pp. 479–485, iSSN: 2381-4802.

[3] J. Gu, W. Hua, W. Yu, Z. Zhang, and H. Zhang, "Surrogate Model-Based Multiobjective Optimization of High-Speed PM Synchronous Machine: Construction and Comparison," *IEEE Transactions on Transportation Electrification*, vol. 9, no. 1, pp. 678–688, Mar. 2023.

[4] M. Cheng, X. Zhao, M. Dhimish, W. Qiu, and S. Niu, "A Review of Data-driven Surrogate Models for Design Optimization of Electric Motors," *IEEE Transactions on Transportation Electrification*, pp. 1–1, 2024.

[5] V. Parekh, D. Flore, and S. Schöps, "Performance Analysis of Electrical Machines Using a Hybrid Data- and Physics-Driven Model," *IEEE Transactions on Energy Conversion*, vol. 38, no. 1, pp. 530–539, Mar. 2023.

[6] "SyR-e: Synchronous Reluctance - evolution." [Online]. Available: https://github.com/SyR-e

[7] M. Gamba, G. Pellegrino, E. Armando, and S. Ferrari, "Synchronous reluctance motor with concentrated windings for IE4 efficiency," in *2017 IEEE Energy Conversion Congress and Exposition (ECCE)*, Oct. 2017, pp. 3905–3912.

[8] I. M. Sobol', "On the distribution of points in a cube and the approximate evaluation of integrals," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 4, pp. 86–112, Jan. 1967.

[9] "FEMM: Finite Element Method Magnetics." [Online]. Available: https://www.femm.info

[10] "Keras: Deep Learning for humans." [Online]. Available: https://keras.io/

[11] "sklearn.svm.SVR — scikit-learn 1.3.2 documentation." [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html

[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.