MagNet Challenge 2023: Data-Driven Model for Power Loss Estimation of Magnetic Components - Final Report

Z. Li, *Student member, IEEE*, R. Mirzadarani, *Student member, IEEE*, R. Liu, *Student member*, L. Wang, *Student member, IEEE*, T. Luo, *Student member, IEEE*, D. Lyu, *Student member, IEEE*, M. Ghaffarian Niasar, Z. Qin, *Senior member, IEEE*

***Abstract*- Traditional methods like Steinmetz's equation (SE) and its improved version (iGSE) have shown limited accuracy in calculating power loss for magnetic materials. To address this, we introduce an innovative approach that combines the Fast Fourier Transform (FFT) with a Feed-forward Neural Network (FNN), leading to a significant improvement in loss prediction accuracy. Our model, optimized through Multi-Objective Optimization (MOO), balances reduced complexity and heightened accuracy. To address data limitations, transfer learning is successfully employed for new materials with sparse data. This strategy offers the potential for improved training efficiency and broader applicability in the field.[1]**

I. Introduction

The empirical Steinmetz Equation (SE), introduced in 1890, has long served as the foundation for the standard modeling techniques for losses estimation in power magnetics. Despite various improvements such as iGSE and i2GSE, the accuracy of these curve-fitting techniques is still relatively low [1], [2]. The imprecise model will lead to a rough magnetic design. Thus, a large design margin is required to achieve an acceptable loss performance of the magnetic components.

One possible approach is using neural networks to accurately predict the loss of magnetic material, considering complicated material response mechanisms to different excitations and the factors involved, such as temperature and DC bias [3]. Different neural network structures have been proposed in [4], including FNN (Feed-forward Neural Network), which uses a four-layer FNN taking peak flux density B, duty ratio D, and fundamental frequency f to predict core loss density P_V. The disadvantage of this method is that the B waveform shape information cannot be included by only using the peak flux density value. One possible approach is to train different FNNs for different B waveforms. However, there are an unlimited number of B waveform shapes. Thus, it is impractical to categorize and then design separate FNNs for each case first.

A more effective method to represent the B waveform information, irrespective of the waveform's shape, is by applying FFT (Fast Fourier Transform). In this report, we utilize FFT to transform the waveform into harmonics. The magnitudes and frequencies of these harmonics, along with temperature data, are then input into the FNN for training. This approach limits the number of parameters and decreases the complexity; thus, the training time is considerably shorter in comparison to other more complex methods. To the best of the authors' knowledge, this is the first research to integrate FFT with FNN for estimating magnetic power loss.

To further improve the model performance, considering both accuracy and model size, a MOO (Multi-Objective Optimization) is carried out. To tackle the limited data size of the new materials, transfer learning has been taken to fine-tune the pre-trained model based on a large dataset of existing materials.

In this report, we detail a neural network modelling approach that integrates MOO and transfer learning to accurately forecast the power loss in magnetic materials under conditions of limited data availability. Accordingly, the report is organized as follows. Section II presents the employed FNN

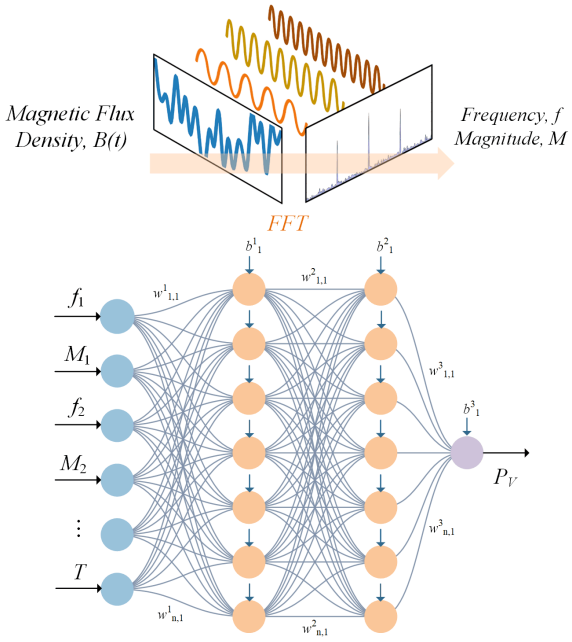

Fig. 1 FNN with FFT

Z. Li (Z.Li-20@tudelt.nl), R. Mirzadarani (R.Mirzadarani@tudelt.nl), R. Liu (Ruijun.Liu@tudelt.nl), L. Wang (L.Wang-11@tudelt.nl), T. Luo (T.Lou-1@tudelt.nl), D. Lyu (D.Lyu@tudelt.nl), M. Ghaffarian Niasar (M.GhaffarianNiasar@tudelft.nl), and Z. Qin (Z.Qin-2@tudelft.nl) are with Department of Electrical Sustainable Energy, Delft University of Technology, Delft, 2628CD, The Netherlands.

architecture, elaborates on the rationale behind the chosen input format and the number of harmonics, and conducts a comparative analysis with other sophisticated AI methodologies. Section III details two approaches for predicting power loss in new materials: normal training and transfer learning, each integrating MOO. This section further compares the developed FNNs, aiming to identify the most effective model regarding both parameter count and accuracy. Finally, section IV concludes the report.

II. Selected Neural Network and Comparison with State of the Art

The FNN structure is shown in Fig. 1. As shown in the figure, the inputs of the FNN are [f_1, M_1, f_2, M_2, …, f_n, M_n, T], where f_k and M_k represent kth harmonic frequency and magnitude, and this input structure is marked as case 0 for later comparison.

A. Selection of Input Format for FNN

Other input formats also have been considered, for example [f_1, M_1, p_1, f_2, M_2, p_2, …, f_n, M_n, p_n, T], where p_k represents the phase information of the kth harmonic, and this case is marked as case 1; and [M_1, M_2, …, M_n, f_1, T], where in this case only the fundamental harmonic frequency is included, and this is named as case 2. A comparison between these three cases has been carried out, as shown in Table 1.

Table 1
Compared Cases

Cases	Neurons in hidden layers	Total number of parameters	Avg. relative error on test set
0	(32,64,32)	4929	1.29
1	(64,64,64)	10433	2.2
2	(32,64,32)	4641	1.69

The test is carried out based on N27 material, and N27 data is split into 60%, 20%, and 20% for train, validation, and test datasets, respectively. Ten harmonics have been chosen for all three cases. The number of neurons in the hidden layers are determined based on the experience, and the activation function is ReLU. The training optimizer is set as Adam, and the model is trained is in PyTorch. An exponentially decayed learning rate strategy is implemented to yield a better model convergence, where the initial learning rate is 0.004, and the decaying rate is 90% per 150 epochs [4]. In section III.A, a hyperparameter optimization will be carried out.

According to the comparison result, there are two conclusions. First, adding harmonics phase information does not help increase loss prediction accuracy. Second, when the model is trained using only a single fundamental frequency as input, its performance is inferior compared to when it is trained with both the frequencies and magnitudes of the harmonic components paired together. This observation holds under the scenario where the network parameters are initially selected based on empirical experience and not subject to further optimization. Therefore, the inputs for the FNN are chosen to be in the format of [f_1, M_1, f_2, M_2, …, f_n, M_n, T].

An early stopping method is implemented in the training process to prevent overfitting and save computational resources. Early stopping halts training when a pre-set number of epochs pass without improvement in validation loss. It is used in the following model training process.

B. Selection of Number of Harmonics for FNN

A further question is how many numbers of harmonics are suitable to consider for model training and prediction. Three different FNN models are built for three, five, and ten harmonic inputs, respectively. A comparison between these three FNNs is shown in Table 2. The model training and testing settings are the same as the tests in Table 1.

Table 2
Comparison of Different Numbers of Harmonics as Input

No. harmonics	Neurons in hidden layers	Total number of parameters	Avg. error on test set	Max. error on test set
3	(16,32,16)	1217	1.38	7
5	(16,32,16)	1281	1.44	12
10	(32,64,32)	4929	1.29	17

The difference in neuron settings in the middle layer is due to the different number of inputs. As can be seen from the comparison results, though with the total number of parameters differences, the results of using three harmonics, five harmonics, and ten harmonics are similar. For the simplicity of the model structure, which is essential for training and future optimization, three harmonics are chosen for the input. Consequently, the input format for the FNN in this study has been established as [f_1, M_1, f_2, M_2, f_3, M_3, T].

C. Comparison with Other Neural Network Structures

With the given magnetic flux density $B(t)$, magnetic field strength $H(t)$, frequency f, temperature T, and power loss P information, different AI models can be applied to predict the power loss. Apart from the scalar-to-scalar model presented in this report, [4] also offered two other models: sequence-to-scalar model using LSTM plus FNN and sequence-to-sequence model using encoder-decoder structure with LSTM. A detailed explanation of these two methods can be found in [4].

To compare with the models mentioned above, similar models, as shown in [4] are developed. The models are trained with N27 data, and for the LSTM+FNN, the input signal $B(t)$ is sampled at 1024-time steps per cycle, while for the encoder-decoder, it is down-sampled to 102-time steps per cycle to manage the higher computational demands. Each model is trained for 10000 epochs. As a result, a comparison between the proposed model, the LSTM+FNN, and the encoder-decoder structure is shown in Table 3. Additionally, the optimized FNN is detailed in section III.A is also shown for benchmarking. The neural network training was conducted in a PC equipped with an AMD Ryzen 7 5800H processor and an NVIDIA GeForce RTX 3060 graphics card.

The encoder-decoder is the most complex model, and the power loss estimation accuracy is unsatisfactory. The reason is that the power loss is not directly generated as the neural

network's output. On the contrary, the power loss is indirectly calculated by integration.

Table 3
Comparison between the FNN, LSTM plus FNN, and Encoder-Decoder

Cases	Total number of parameters	Training time	Avg. relative error on the test set
Initial FNN	1217	~1 hour	1.38
Optimised FNN	1419	~1 hour	0.75
LSTM+FNN	2809	~20 hours	1.37
Encoder-decoder	28097	~20 hours	3.34

The LSTM+FNN model achieves the highest accuracy without considering the optimization of the three models. Because it effectively captures detailed information from the B waveforms and directly generates power loss estimations. However, the sequence-to-scalar model presents two disadvantages compared to the scalar-to-scalar model. First, it is more complex and requires longer training time, which is unsuitable for further optimization. Second, in the case of transfer learning using the data measured by other research groups, adjustments like down-sampling/up-sampling or modifying the model structure might be needed for transfer learning, especially when handling data sequences of different lengths.

In comparison, the proposed FNN model is the simplest and quickest for training. With further optimization, the model's performance can be increased significantly. Given its simplicity and high accuracy, the proposed FNN has been demonstrated to be suitable for the following development and optimization.

III. Multi-Objective Optimization, Transfer Learning and Comparison

The overall model construction and selection methods for the five new materials are presented in Fig. 2. In the normal training approach, for the five new materials, the FNN model for each material is first optimized with the MOO tool Optuna, as detailed in the next section. New materials data are split into 70%, 20%, and 10% for training, validation, and test purposes, respectively. Then, the best hyperparameter

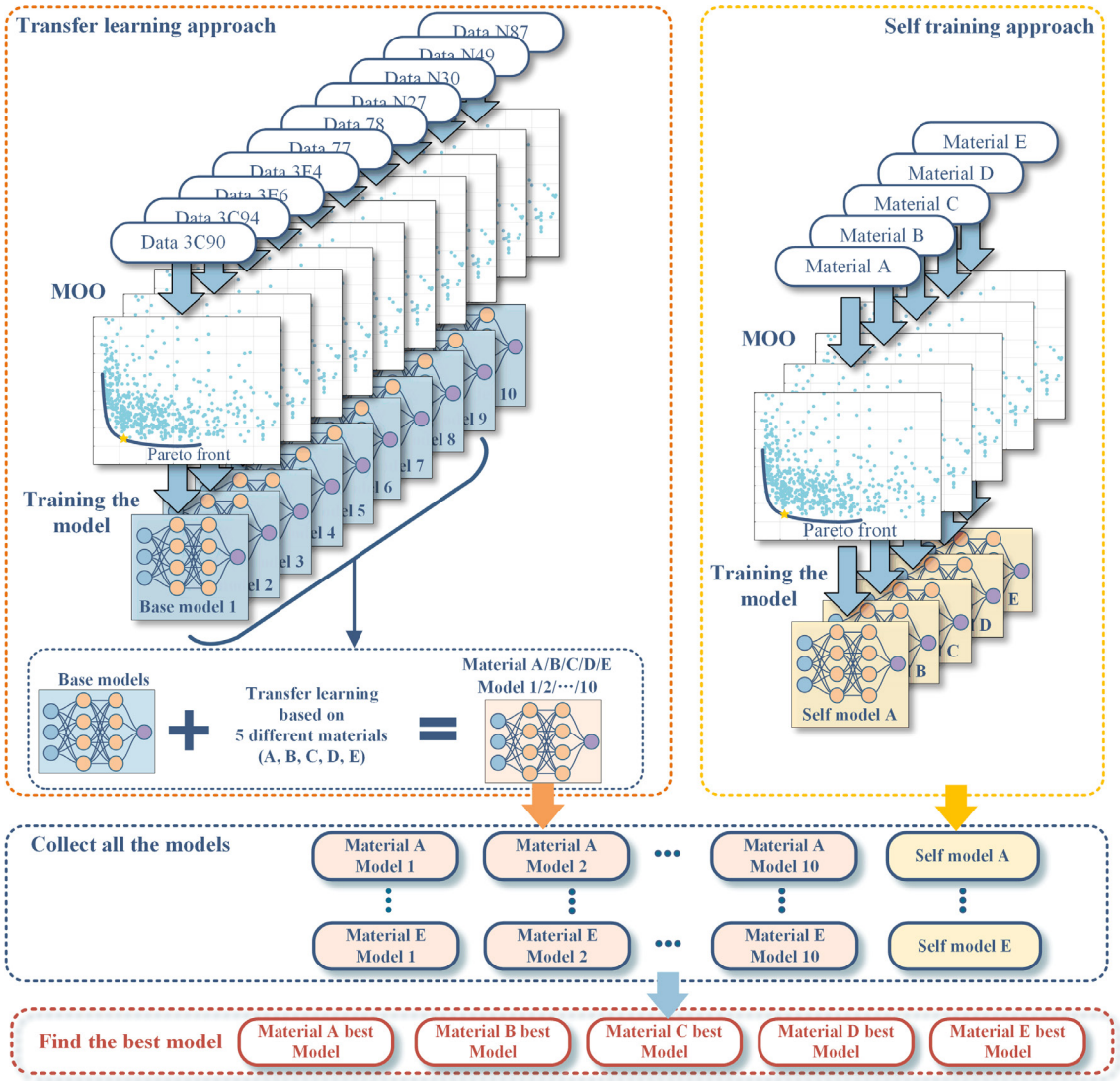

Fig. 2 Overall Model Training and Comparison Process

combination with a low number of parameters and low validation loss is selected. The test dataset performance is recorded for later comparison.

Although with MOO, the test performance is satisfactory with the normal training process, due to the low number of new material data, a better approach is by using transfer learning. This method holds significant importance in real-world scenarios, such as for designers who may lack adequate data on new materials required for normal training models. This part will be analyzed in III.B. The overall approach for transfer learning, as shown in Fig. 2, also starts with a MOO for the ten existing materials. This will increase the accuracy of the following step which uses transfer learning. Once the optimal model for each material is determined, the five new materials will be subjected to transfer learning across the ten models through fine-tuning. The ten existing materials data are split into 70%, 15%, and 15% for training, validation, and test purposes, respectively. This distribution strategy differs slightly from the previous one, as in the earlier data split, we aimed to maximize the use of the limited data available for training the new material. As a result, 50 new transfer-learned models for the five new materials will be acquired, and the test data results will be recorded. The transfer-learned and normal training results described previously are compared to select the final best model. We have also experimented with training the base model using a combined dataset of several materials (such as 3C90, N27, and 77). However, the outcomes of transfer learning with this approach did not surpass those achieved when training the base model exclusively with data from a single material (3C90).

A. Multi-Objective Optimization

In assessing the neural network model, both prediction accuracy and the number of parameters are crucial metrics. While accuracy reflects the model's effectiveness, the parameter count directly indicates its complexity and significantly impacts the time required for training and fine-tuning. To optimize the model considering both values, a MOO is needed.

MOO aims to find the optimal set of parameters, simultaneously optimizing multiple objectives, leading to trade-offs and compromises. Optuna, a popular Python library, provides a comprehensive framework for conducting MOO. Utilizing Optuna for MOO, we focus on two primary objectives: reducing the validation loss to enhance model accuracy and minimizing the number of parameters. NSGA-II (Non-dominated Sorting Genetic Algorithm II) is favored as the optimization engine for its effective balance in ranking and diversity preservation in MOO contexts. The search space is defined as follows: the number of middle layers in the FNN structure varies between 2 and 5, each containing 8 to 64 neurons. The activation functions considered include ReLU, LeakyReLU, Tanh, Sigmoid, ELU, SELU, and SiLU. Possible optimizers are Adam, SGD, RMSprop, and AdamW. The batch size options are set at 64, 128, or 256.

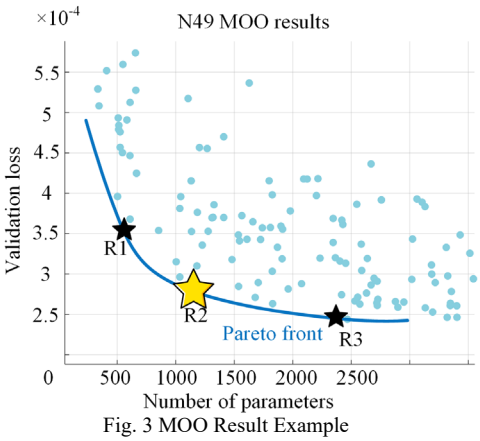

Fig. 3 MOO Result Example

For each of the five new materials, MOO involves 500 trials, each comprising 3000 epochs. In contrast, for the ten existing materials, which have larger datasets, MOO consists of 200 trials with 1000 epochs each. An example MOO result is presented in Fig. 3, where stars indicate the best hyperparameter combinations. Table 4 summarizes these combinations, and the optimizer for three cases is Adam. The N49 test dataset demonstrates that accuracy increases with the number of parameters. However, to strike a balance between the number of parameters and accuracy, R2 is selected as the optimal choice for further transfer learning due to its proximity to the knee point of the Pareto front. A similar analysis approach is applied to the other materials.

Table 4
Three Different Points on N49 MOO Pareto Front

case	Batch size	Hidden layers neurons	No. params.	Act. Func.	N49 avg.	Material E TF avg.
R1	64	(21, 17)	560	Tanh	2.01	2.75
R2	128	(29, 29)	1132	Tanh	1.9	2.49
R3	128	(38, 19, 33, 19)	2371	SiLU	1.46	2.35

B. Transfer Learning

After the MOO, ten base models trained on existing materials are ready for transfer learning. New material data will be directly input into the pre-trained base model. This process will fine-tune the neural network's parameters, adapting the model to accommodate the characteristics of the new materials.

A test is carried out to testify the MOO results and the transfer learning effectiveness. Based on Table 4 N49 trained models, the new material E data are applied for fine tuning. The transfer learning results are shown in the last column of the table. It can be similarly concluded that an increase in the number of parameters enhances the accuracy of transfer learning results. To strike an optimal balance between the number of parameters and accuracy, the knee point on the Pareto front is identified as the ideal combination of hyperparameters.

C. Normal Training Results and Transfer Learning Results Comparison

Based on the overall process shown in Fig. 2, the results of the transfer learning approach and the normal training

Table 5
Ten Existing Materials Optimized MOO Models Transfer Learning Results

material (number of parameters)	Self-training result		Transfer learning results									
			Material A		Material B		Material C		Material D		Material E	
	Average error	95% error	Average error	95% error	Average error	95% error	Average error	95% error	Average error	95% error	Average error	95% error
N27(1419)	0.75	1.97	2	6.7	0.75	1.89	1.39	3.93	3.16	10	2.4	7.25
N30(2197)	0.41	1.1	2.08	7.85	0.7	1.65	1.1	3.24	3.7	10	2.28	7
N49(1132)	1.9	5.78	2.688	8.9	0.79	2.2	1.37	3.9	5.11	19	2.49	7
3F4(2117)	1.45	3.8	2.58	7.8	0.84	2.17	1.23	3.4	5.2	15.7	2.66	10.9
3E6(1295)	0.455	1.19	2.25	6.3	0.74	1.98	1.23	3.62	4.26	10.65	2.77	8
3C90(2705)	0.83	2.35	2.23	8.1	0.76	1.97	1.2	3.78	3.86	10.9	2.38	9.24
3C94(3021)	0.75	2	3.1	10.4	0.75	2.02	1.37	4.2	5.9	22.4	2.63	7.09
N87(1525)	0.74	2	2.54	7.67	0.73	1.9	1.19	4	4.6	14.9	2.74	10.7
77 (2454)	0.77	2.1	2.2	6.37	0.92	2.48	1.21	3.54	4.77	14	2.2	6.4
78(2285)	0.72	1.89	2.19	7.1	0.69	1.9	1.21	3.72	5.23	15.38	2.69	8.17

Table 6
Comparison of the Best Transfer Learned Results and Regular Trained Results for the Five New Materials

New Material	Self-training result			Transfer learning result			
	Number of parameters	Average error	95% error	Number of parameters	Average error	95% error	Base model
A	1662	2.18	6.73	1419	2	6.7	N27
B	552	0.9	2.2	2197	0.7	1.65	N30
C	935	1.65	4.4	2197	1.1	3.24	N30
D	1857	6.3	21.8	1419	3.16	10	N27
E	1551	2.47	7	2454	2.2	6.4	77

approach for the five new materials are presented in Table 5 and Table 6. Table 6 compares the transfer learning results for the five new materials, and the best results are highlighted in green. Table 7 presents test dataset accuracy for the MOO models for the five new materials to compare with the transfer learning results.

As seen in Table 6, from an accuracy perspective, the results of transfer learning significantly outperform those of self-training, mainly due to the limited dataset available for the new materials. Although the number of parameters in transfer learning approaches is slightly higher than in self-training methods, this is not a significant concern for applying transfer learning to new materials. The rapidity of the transfer learning process, as demonstrated in our tests, enables quick adaptation to the characteristics of new materials. Consequently, despite the larger parameter size, the speed of transfer learning makes it a more practical choice. Therefore, considering both the superior accuracy and the efficiency in learning new material properties, the transfer learning models are deemed the most effective for new materials.

The number of parameters for the final model for the five new materials is summarized in Table 7 for the committee's convenience.

Table 7
Number of Parameters for New Materials

Material	A	B	C	D	E
Number of parameters	1419	2197	2197	1419	2454

IV. Conclusion

This report comprehensively presents a neural network based modelling approach for accurately predicting the power loss of magnetic materials with limited data, incorporating MOO and transfer learning techniques. The fundamental neural network employed in this study is a FNN. Through a detailed comparison of various input structures and the number of harmonics, the chosen inputs are the magnitudes and frequencies of the first, second, and third harmonics, plus temperature. Compared to other advanced methods, such as LSTM or encoder-decoder models, the proposed method achieves satisfactory accuracy with lower complexity. The model's performance is further enhanced by MOO, targeting both a reduction in the number of parameters and an increase in accuracy. Finally, a comparative analysis between the transfer learning and self-training approaches, both integrated with MOO, is conducted for new materials with limited data. The transfer learning shows a better performance with considerable lower error and marginal increase in the number of parameters.

References

[1] F. Kral, "Analysis and Implementation of Algorithms for Calculation of Iron Losses for Fractional Horsepower Electric Motors," Master's Graz University of Technology, 2017.

[2] C. N. M. Smailes, R. Fox, J. Shek, M. Abusara, G. Theotokatos, P. McKeever, "Evaluation of core loss calculation methods for highly nonsinusoidal inputs," presented at the IET International Conference on AC and DC Power Transmission, 2015.

[3] D. Serrano. e. al., "Why MagNet: Quantifying the Complexity of Modeling Power Magnetic Material Characteristics," *IEEE Transactions on Power Electronics,* vol. 38, no. 11, pp. 14292-14316, 2023, doi: 10.1109/TPEL.2023.3291084.

[4] H. Li. e. al., "How MagNet: Machine Learning Framework for Modeling Power Magnetic Material Characteristics," *IEEE Transactions on Power Electronics,* vol. 38, no. 12, pp. 15829-15853, 2023, doi: 10.1109/TPEL.2023.3309232.