FRONTIERS IN ENGINEERING | SCIENCE | MANAGEMENT
MNS-UET, MULTAN

FRONTIERS IN ENGINEERING, SCIENCE & MANAGEMENT

https://journal.mnsuet.edu.pk/

Research Paper

Explainable AI for Modeling Magnetic Flux Distribution in MCCB Actuators: A Configuration-Aware Analysis

Mehmet Onur YAĞIR [a, *], **Sevda GÜL** [b]

[a] Sakarya University Faculty of Engineering Department of Mechanical Engineering 54050, Sakarya, Turkey.

[b] Department of Electronics and Automation, Adapazarı Vocational School, Sakarya University, 54050 Serdivan, Turkey.

ARTICLE INFORMATION

Received	10 July 2026
Received in revised form	07 Aug 2026
Accepted	11 Aug 2026
Available Online	15 Sep 2026

*CORRESPONDING AUTHOR

Mehmet Onur YAĞIR

Sakarya University Faculty of Engineering Department of Mechanical Engineering 54050, Sakarya, Turkey.

Email mehmetyagir@sakarya.edu.tr

ABSTRACT

Background: This study presents an Explainable Artificial Intelligence (XAI)–based analysis of the three-dimensional magnetic flux density components (B_x, B_y, B_z) generated in magnetic actuators used in molded case circuit breakers (MCCBs). The analyses were conducted on a dataset obtained from electromagnetic simulations performed in CST Studio Suite, comprising a total of 1,186,066 samples across 27 different actuator configurations. The dataset includes 16 input features (Full Mode), covering design parameters, spatial coordinates, and physical field quantities.

Methods: To investigate how the model benefits from different levels of information, three feature modes were compared: (i) Full Mode, which includes all variables; (ii) Design–Spatial Mode, which incorporates design parameters and spatial coordinates; and (iii) Design-Only Mode, which relies solely on design parameters. To prevent data leakage, the evaluation was carried out using a configuration-based group splitting strategy, and model stability was assessed through repeated experiments.

Results: The results show that the Full Mode, as the most physically informative representation, achieves the highest accuracy, particularly for the B_y and B_z components. The Design–Spatial Mode, despite its simpler feature structure, largely preserves this high performance. In contrast, the Design-Only Mode exhibits a significant drop in performance. These findings highlight that spatial information is critical for the point-wise prediction of local magnetic flux density components, whereas relying solely on design parameters is insufficient. Furthermore, feature importance analysis and SHapley Additive exPlanations (SHAP)-based explanations indicate that the model's decisions are grounded in physically meaningful variables. Quantitatively, the Full Mode attains R^2 = 0.9152, 0.9910, and 0.9941 for B_x, B_y, and B_z, respectively, while the more compact Design–Spatial Mode remains highly competitive (R^2 = 0.9846, 0.9916, and 0.9658), and the Design-Only Mode drops sharply (R^2 = 0.0194, 0.1015, and 0.0063 for B_x, B_y, and B_z respectively), confirming that spatial information — rather than design parameters alone — is the key driver of point-wise prediction accuracy.

Conclusion: Overall, the findings demonstrate that, in data-driven modeling of electromagnetic systems, not only predictive accuracy but also the level of information and interpretability play a crucial role.

Keywords: Explainable Artificial Intelligence (XAI), Magnetic Flux Prediction, MCCB, Magnetic Actuator, Feature Importance, Spatial Features, Surrogate Modeling, Random Forest

1. Introduction

Molded case circuit breakers (MCCBs) are widely used low-voltage protection devices in industrial and commercial power distribution systems (Q. Wang et al., 2023). The reliable operation of these devices is directly governed by the behavior of their internal magnetic actuator mechanisms, which convert electromagnetic energy into controlled mechanical motion. The performance of magnetic actuators is determined by complex interactions among excitation conditions, material properties, geometric structure, and the spatial distribution of the magnetic field. Therefore, variations in parameters such as coil winding turns, permanent magnet strength, and material selection can lead to significant changes in the magnetic flux density distribution (Cho et al., 2022).

In the analysis of such systems, high-fidelity electromagnetic simulation tools such as CST Studio Suite enable detailed investigation of field distributions. In the literature, studies aimed at understanding the behavior of MCCB systems are largely based on physics-driven simulation approaches. Research on arc interruption performance and current-carrying behavior in low-voltage circuit breakers demonstrates that the electromagnetic characteristics of these systems can be modeled in detail (Cho et al., 2022; Q. Wang et al., 2023); comparably detailed electromagnetic-response simulation methodologies have also been applied to other safety-critical electrical systems, such as coupled cable and structural response analysis under lightning strikes (Chen et al., 2025). Similarly, studies examining the influence of external transverse magnetic fields on arc evolution and interruption behavior in circuit breakers have revealed the sensitivity of system behavior to physical parameters (S. Liu et al., 2024). Related work on transient protection and current-limiting behavior in modern power-electronic-dominated grids, including grid-forming converter current-limiting algorithms and transient stability analysis (Qoria et al., 2020), further illustrates the continued relevance of protective-device modeling across evolving power system architectures. More recent studies continue to explore different physical aspects of MCCB systems through approaches such as comparative arc interruption analysis (Zhang et al., 2025) and circuit-based air arc modeling.

While these studies provide valuable insights into the physical behavior of electromagnetic systems, they are largely limited to deterministic modeling and accuracy-oriented analyses. Moreover, most existing work primarily focuses on prediction accuracy, without sufficiently addressing which input variables the model relies on and what physical relationships it actually learns. However, in engineering applications, achieving high accuracy alone is not sufficient; understanding the decision-making mechanism of the model is equally critical for ensuring reliability and interpretability. In this context, explainable artificial intelligence (XAI) approaches enable the analysis of complex model behavior, offering more transparent and trustworthy solutions for engineering problems (Barredo Arrieta et al., 2020; Tjoa and Guan, 2021). Furthermore, it has been emphasized in the literature that the success of artificial intelligence models in electromagnetic systems depends not only on the choice of model but also strongly on data quality and the adequacy of feature representation (Ali et al., 2023; Zha et al., 2025). This highlights the need for a systematic investigation of how different feature groups contribute to model performance.

In this study, the objective is not only to predict the three-dimensional magnetic flux density components in MCCB magnetic actuators, but also to determine under which levels of information these components can be reliably estimated. To this end, the dataset is analyzed under three different feature modes: (i) Full Mode, which includes all variables; (ii) Design–Spatial Mode, which incorporates only design parameters and spatial information; and (iii) Design-Only Mode, which relies solely on design parameters. This framework allows for a systematic evaluation of the role of spatial information in magnetic field prediction, while the dependency of the model on specific variables is assessed using explainable artificial intelligence techniques. This study goes beyond conventional accuracy-driven approaches in data-driven modeling of electromagnetic systems by contributing to a more physically interpretable understanding of model behavior.

To more clearly position this study relative to recent work, we note that explainable machine learning has increasingly been applied to electromagnetic-field-related prediction tasks outside the power-device domain: SHAP- and LIME-based explanations have been used to interpret ensemble-model predictions of ambient electric-field strength in urban environments across multiple European cities (Givisis et al., 2025; Kiouvrekis et al., 2025), with tree-based ensembles (Random Forest, XGBoost, LightGBM) consistently identified as strong predictors whose feature-importance rankings differ from those of simpler models. In parallel, physics-informed machine learning has emerged as a complementary paradigm for embedding governing electromagnetic equations directly into data-driven models, improving generalization and physical consistency in applications ranging from eddy-current couplers to artificial electromagnetic materials (Deng et al., 2025; Karniadakis et al., 2021; J. Wang et al., 2025; Wu et al., 2024). Compared to this body of work, the present study differs in that it does not embed governing equations into the learning algorithm (a physics-informed approach) or explain a black-box ensemble comparison (a multi-model XAI approach); instead, it uses a single tree-ensemble model analyzed using post-hoc explanations to systematically isolate how much of the predictive signal is attributable to spatial information versus design parameters alone -a question that, to our knowledge, has not been directly addressed for MCCB magnetic actuators.

2. Materials and Methods

2.1. Dataset

The dataset used in this study was generated through numerical simulations performed in CST Studio Suite, designed to represent the electromagnetic behavior of the magnetic actuator system in molded case circuit breakers (MCCBs) (Figure 1(a)). The dataset comprises a total of 27 different actuator configurations, created to analyze the impact of varying design parameters on system performance. Detailed descriptions of the parameters used in these configurations are provided in Table 1, while a subset of these parameters is visually illustrated in Figure 1(b).

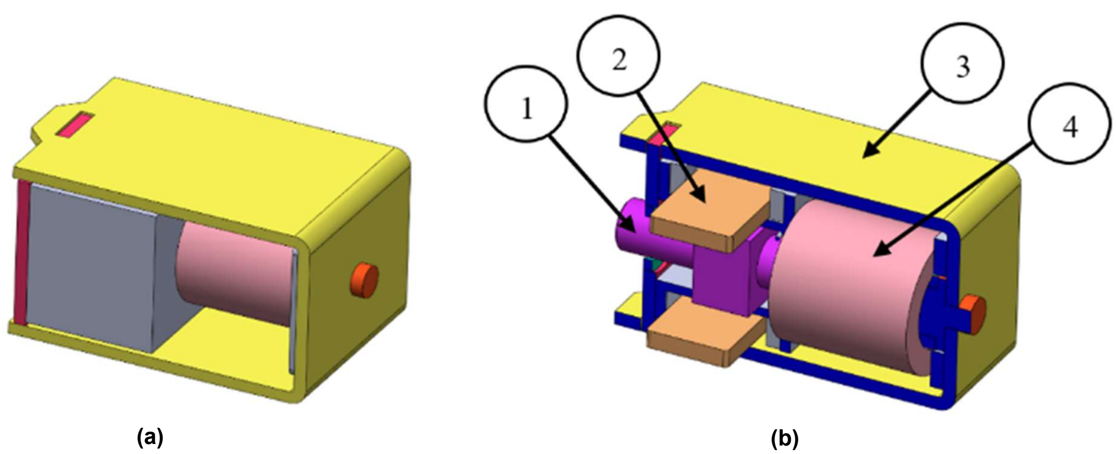

Fig. 1 - (a) Magnetic actuator assembly; (b) Cross-sectional view of the actuator design and its main components: (1) shaft, (2) magnet, (3) chassis, and (4) coil.

Table 1 - Parametric design space of the actuator configurations

Parameter	Values	Description
Magnet level (Gauss)	800, 900, 1000	Magnetic flux density level of the permanent magnet
Coil turns	800, 900, 1000	Number of coil windings
Top cover material	Brass, Alpaca, St 37	Material of the top cover
Total combinations	(3 × 3 × 3 = 27)	All possible parameter combinations

The 27 configurations correspond to a deterministic full-factorial design of experiments (3 × 3 × 3), rather than a random or Latin-hypercube sample; every combination of the three categorical/discrete design parameters in Table 1 was simulated. This design ensures complete, systematic coverage of the discrete design space defined by these parameters, and each configuration was additionally sampled at a dense grid of spatial locations to characterize the local field distribution, yielding the 1,186,066-sample dataset described below.

2.2. CST Simulation Results Overview Across Configurations

Figure 2 graphically summarizes the main characteristics of the CST-generated dataset, including the complete factorial design space formed by three magnet flux-density levels, three coil-turn levels, and three top-cover materials, as well as the configuration-wise sample counts, spatial sampling domain, and distributions of the B_x, B_y, and $B_z$components. The component distributions exhibit pronounced concentrations near zero together with long-tailed field values, reflecting the spatially heterogeneous nature of the magnetic field. To complement this graphical overview, Table 2 reports the spatial mean and maximum absolute values of B_x, B_y, and B_z separately for all 27 actuator configurations. The mean resultant flux magnitude was calculated as $|B| = \sqrt{B_x^2 + B_y^2 + B_z^2}$. Increasing the coil turns from 800 to 1000 consistently increased the mean resultant flux magnitude across all nine magnet-level/top-cover combinations. The increase ranged from 17.6% to 18.7% in eight combinations, confirming that the number of coil turns is a major determinant of the spatially averaged magnetic field. A notably larger increase of 39.9% was observed for the 800-G magnet level and brass top-cover combination because its 800-turn coil configuration exhibited an unusually low mean resultant flux magnitude (3.115 mV·s/m²) compared with the corresponding Alpaca and St 37 steel configurations (3.674 and 3.675 mV·s/m², respectively). This value was verified directly against the raw CST export and retained without modification. Nevertheless, a solver-side verification, particularly a mesh-convergence analysis for this configuration, is recommended before the observed difference is interpreted as a confirmed physical effect. For nearly all other fixed magnet/coil combinations, changing the top-cover material altered the mean resultant flux magnitude by only approximately 0.1%–0.4%, indicating its limited influence on the spatially averaged field. The component-wise maxima reveal distinct parameter-dependent effects: the maximum $|B_x|$increases primarily with magnet level, whereas the maximum $|B_y|$and $|B_z|$increase predominantly with coil turns. In contrast, the top-cover material produces negligible differences in the reported component-wise maxima.

Table 2. CST-simulated flux-density component summary showing the spatial mean and maximum magnitudes of $|B_x|$, $|B_y|$, and $|B_z|$(mV·s/m²) for each of the 27 actuator configurations.

Configuration	Magnet parameter	Coil parameter	Top-cover material	Top-cover μr	Mean $\lvert B_x\rvert$	Max $\lvert B_x\rvert$	Mean $\lvert B_y\rvert$	Max $\lvert B_y\rvert$	Mean $\lvert B_z\rvert$	Max $\lvert B_z\rvert$
Model 1	800	800	Brass	1	1.003	73.897	2.186	87.513	0.599	60.187
Model 2		900			1.471	73.898	2.723	99.769	0.709	67.710
Model 3		1000			1.556	81.269	2.990	110.856	0.757	75.234
Model 4	800	800	Alpaca	30	1.387	73.897	2.457	88.681	0.661	60.188
Model 5		900			1.472	73.898	2.724	99.769	0.710	67.712
Model 6		1000			1.558	81.268	2.991	110.856	0.758	75.235
Model 7	800	800	St 37	1000	1.387	73.897	2.457	88.681	0.662	60.188
Model 8		900			1.473	73.898	2.724	99.769	0.711	67.712
Model 9		1000			1.559	81.268	2.992	110.856	0.760	75.235
Model 10	900	800	Brass	1	1.474	83.133	2.497	88.681	0.693	60.187
Model 11		900			1.560	83.134	2.763	99.769	0.742	67.710
Model 12		1000			1.645	83.135	3.030	110.856	0.790	75.234
Model 13	900	800	Alpaca	30	1.476	83.133	2.498	88.681	0.694	60.188
Model 14		900			1.556	83.134	2.752	99.769	0.743	67.712
Model 15		1000			1.647	83.135	3.032	110.856	0.791	75.235
Model 16	900	800	St 37	1000	1.476	83.133	2.498	88.681	0.695	60.188
Model 17		900			1.562	83.134	2.765	99.769	0.744	67.712
Model 18		1000			1.647	83.135	3.032	110.856	0.793	75.235
Model 19	1000	800	Brass	1	1.563	92.369	2.537	88.681	0.726	60.187
Model 20		900			1.648	92.370	2.804	99.768	0.775	67.710
Model 21		1000			1.734	92.371	3.071	110.856	0.823	75.234
Model 22	1000	800	Alpaca	30	1.564	92.369	2.538	88.681	0.728	60.188
Model 23		900			1.650	92.370	2.805	99.768	0.776	67.712
Model 24		1000			1.736	92.371	3.072	110.856	0.825	75.235
Model 25	1000	800	St 37	1000	1.565	92.369	2.539	88.681	0.729	60.188
Model 26		900			1.650	92.370	2.805	99.768	0.777	67.712
Model 27		1000			1.736	92.371	3.072	110.856	0.826	75.235

Figure 3 and Figure 4 present representative CST-simulated magnetic flux density maps (side cross-section and top view, respectively) for the actuator.

In the side cross-section shown in Figure 3, the magnetic flux density is distributed symmetrically along the coil axis. The highest flux density occurs in the air gap between the moving core and the stationary magnetic circuit, while the flux intensity decreases gradually with distance from this region. The flux lines are largely closed within the magnetic circuit, and leakage flux to the surrounding region remains limited.

In the top view (Figure 4), the magnetic flux density is symmetrically distributed around the coil center, with flux vectors concentrating in the air gap and oriented regularly from the center toward the outer regions. The field intensity reaches its maximum in the central region and decreases with increasing radial distance. Taken together, both views indicate that the magnetic field is effectively concentrated in the target region, the magnetic circuit operates in a balanced manner, and the resulting flux distribution is favorable for electromagnetic force generation.

Table 3 compares CST-simulated (true) and Random Forest-predicted mean flux density, converted to physical units (Vs/m²), for all three feature modes (Full, Design–Spatial, Design-Only) and all three output components (B_x, B_y, B_z), restricted to the four configurations held out from training (genuine out-of-sample predictions). Prediction error generally increases as the feature set is reduced, consistent with the R² pattern in Table 5: median relative error across the 12 held-out cases is 0.11% for the Full Mode, 4.87% for the Design–Spatial Mode, and 19.70% for the Design-Only Mode. One held-out configuration (Magnet = 1000 G, Coil = 800, Top-cover = Brass) is a clear exception: for this configuration, all three modes show much larger relative errors, and for the Full and Design-Spatial Modes specifically the B_x and B_z

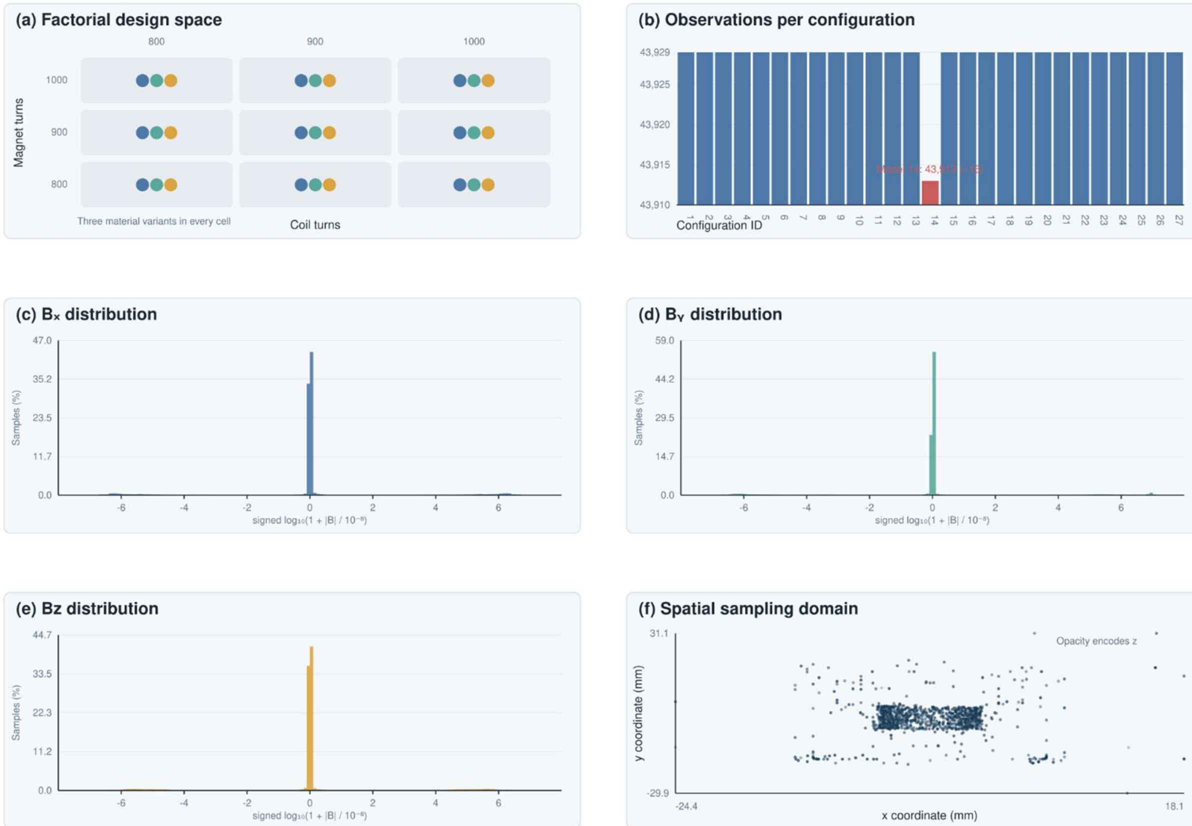

Figure 2. Graphical characteristics of the simulation dataset. (a) Full-factorial design space comprising three magnet levels (800, 900, and 1000 G), three coil-turn levels (800, 900, and 1000 turns), and three top-cover materials. (b) Number of spatial observations in each configuration. (c–e) Distributions of the B_x, B_y, and B_z components represented using a signed logarithmic transformation. (f) Spatial sampling domain in the $x-y$ plane, with opacity indicating the z −coordinate.

components (whose true magnitude is very close to zero for this configuration) show very large percentage errors despite small absolute deviations - a known limitation of percentage-based error metrics near zero. The Design-Only Mode additionally shows one severe miss for this configuration (B_y: 98.1% relative error), which is consistent with its comparatively low R² in Table 5 and Table 7. We report this configuration's results in full rather than removing it and identify it as a priority case for the uncertainty-quantification future work discussed in the Conclusion.

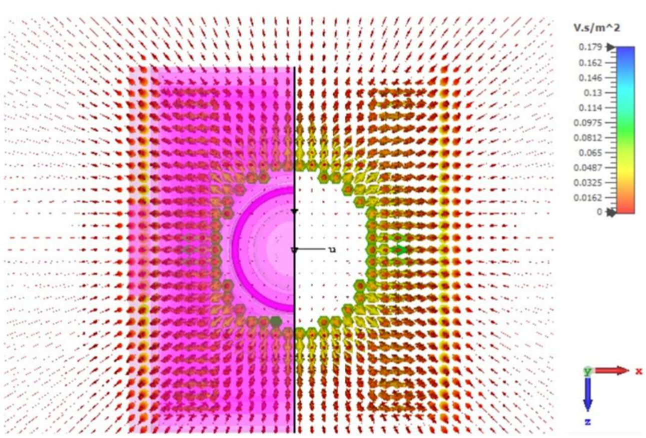

Figure 3. Side cross-sectional view of the CST-simulated magnetic flux density distribution (Vs/m²), showing vector orientation around the moving core and air gap.

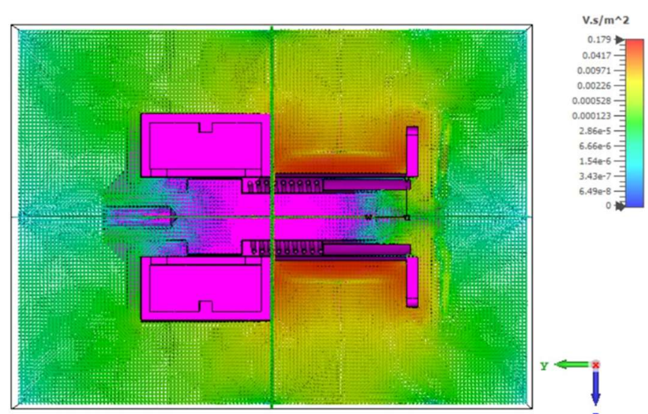

Figure 4. Top view of the CST-simulated magnetic flux density distribution (Vs/m²) around the coil and actuator assembly.

The dataset obtained from the simulations consists of a total of 1,186,066 samples. Each sample represents the behavior of the magnetic field at a specific point in space and is composed of input and output variables. The input features used in the Full Mode comprise a total of 16 variables, which are categorized into three main groups:

a) **Design parameters:** magnet level, number of coil turns, and material properties

Table 3. Configuration-level comparison of CST-simulated vs. AI-predicted mean flux density (physical units), for all three feature modes and all three output components, restricted to the four held-out (genuine out-of-sample) test configurations.

Mode	Magnet	Coil	Material	Comp.	True (Vs/m²)	Pred (Vs/m²)	Err %
				B_x	0.000011	0.000011	0.70
Full	900	1000	Brass	B_y	0.001243	0.001243	0.00
				B_z	0.000003	0.000003	0.11
				B_x	0.000012	0.000012	0.26
Full	1000	800	Alpaca	B_y	0.000990	0.000990	0.00
				B_z	0.000001	0.000001	1.37
				B_x	0.000011	0.000052	388.97
Full	1000	800	Brass	B_y	0.000991	0.001100	10.94
				B_z	0.000001	0.000050	3563.83
				B_x	0.000011	0.000011	0.01
Full	1000	900	Brass	B_y	0.001116	0.001116	0.01
				B_z	0.000002	0.000002	0.40
				B_x	0.000011	0.000011	2.17
Design-Spatial	900	1000	Brass	B_y	0.001243	0.001230	1.01
				B_z	0.000003	0.000002	7.66
				B_x	0.000012	0.000015	26.03
Design-Spatial	1000	800	Alpaca	B_y	0.000990	0.001008	1.81
				B_z	0.000001	0.000001	11.20
				B_x	0.000011	0.000013	23.59
Design-Spatial	1000	800	Brass	B_y	0.000991	0.001258	26.95
				B_z	0.000001	0.000001	27.81
				B_x	0.000011	0.000014	25.80
Design-Spatial	1000	900	Brass	B_y	0.001116	0.001120	0.33
				B_z	0.000002	0.000002	4.87
				B_x	0.000011	0.000008	28.21
Design-Only	900	1000	Brass	B_y	0.001243	0.001218	2.02
				B_z	0.000003	0.000003	19.70
				B_x	0.000012	0.000016	40.68
Design-Only	1000	800	Alpaca	B_y	0.000990	0.000992	0.20
				B_z	0.000001	0.000002	38.79
				B_x	0.000011	0.000012	15.49
Design-Only	1000	800	Brass	B_y	0.000991	0.000019	98.07
				B_z	0.000001	0.000001	5.21
				B_x	0.000011	0.000013	12.02
Design-Only	1000	900	Brass	B_y	0.001116	0.001181	5.80
				B_z	0.000002	0.000002	23.43

b) **Spatial information:** x, y, z coordinates of the sampled points in the magnetic field

c) **Physical quantities:** magnetic field intensity components (H_x, H_y, H_z), energy density, and derived field-related quantities

The target output variables to be predicted by the model are the three-dimensional magnetic flux density components, namely B_x, B_y, and B_z. Standard preprocessing steps were applied to the dataset prior to model training. Continuous variables were normalized to reduce the impact of scale differences on the model, while categorical design parameters were transformed into numerical representations using one-hot encoding to enable direct processing by the learning algorithms. These steps helped mitigate potential biases arising from feature scale disparities during training.

A column-wise data-quality check confirmed no missing values across the full dataset (0 missing entries in every input and output column, verified programmatically). For completeness, the exact input feature sets used in each mode are as follows. Full Mode (16 features): magnet level (Gauss), coil turns, housing/top-cover/shaft relative permeability, spatial coordinates (x, y, z), magnetic field intensity components (H_x, H_y, H_z), effective field cross-sectional area, energy density, real (measured) permeability, and one-hot encoded top-cover material indicators. Design–Spatial Mode (11 features) excludes the field-intensity components, area, and energy density, retaining design parameters and spatial coordinates. Design-Only Mode (8 features) further excludes the spatial coordinates, retaining only component-level design parameters and material indicators.

This data structure provides a comprehensive representation for data-driven modeling of electromagnetic systems by jointly incorporating design parameters and spatial information. However, the multi-configuration nature of the dataset requires a specialized data splitting strategy to ensure a reliable assessment of the model's generalization capability. In particular, random sample-based splitting may lead to data leakage, as samples originating from the same actuator configuration could appear in both training and test sets. To address this issue, a configuration-based group splitting strategy was adopted instead of a conventional random train/test split.

2.3. Data Splitting Strategy

In this study, a dedicated data splitting strategy was employed to ensure that the model performance reflects true generalization capability. Since the dataset contains a large number of spatial samples derived from the same physical actuator configurations, conventional random sample-based train/test splitting can lead to data leakage. In such cases, the model may encounter samples in the test phase that are highly similar to those seen during training, resulting in overly optimistic performance estimates. To address this issue, a configuration-based group splitting strategy was adopted. The dataset was grouped based on 27 distinct actuator configurations, each representing a unique combination of design parameters. Samples belonging to the same configuration were strictly assigned to a single subset, ensuring that no spatial samples from the same physical configuration were shared between the training and test sets. In each experiment, 4 out of the 27 configurations were reserved for the test set, while the remaining 23 configurations were used for training. This setup ensures that the model is evaluated on entirely unseen actuator configurations during testing. To assess model stability, the splitting procedure was not limited to a single partition. Instead, five repeated experiments were conducted using different combinations of configurations. This allowed for analyzing the consistency of model performance across previously unseen configurations. Overall, this approach enables the evaluation of whether the model learns physically meaningful relationships rather than merely exploiting data-specific similarities.

2.4. Feature Modes

To investigate how the model utilizes different types of information and to assess the impact of varying information levels on prediction performance, the dataset was evaluated under three distinct feature modes.

a) Full Mode: In this mode, all input variables were utilized. This includes design parameters, spatial coordinates, and physical field quantities, providing the richest possible representation of information accessible to the model. Therefore, the Full Mode is considered the reference scenario, where local magnetic behavior can be modeled with the highest level of physical information. It should be noted that the Full Mode's field-intensity inputs (H_x, H_y, H_z, in A/m) are themselves outputs of the same CST simulation used to generate the target flux-density values. The Full Mode therefore functions as an upper-bound reference and ablation baseline for quantifying how much predictive power is attributable to field-derived information, rather than as a standalone deployable surrogate that would allow users to bypass CST simulation entirely. In contrast, the Design–Spatial and Design-Only Modes rely only on design parameters and/or spatial coordinates known prior to simulation and therefore represent the practically deployable configurations. The real permeability retained in the Design–Spatial Mode represents a material property specified prior to simulation and was therefore treated as a design input rather than a field-derived feature.

b) Design–Spatial Mode: In this mode, only design parameters and spatial coordinates were used, while simulation-derived physical quantities were excluded from the input set. These excluded variables include magnetic field intensity components (H_x, H_y, H_z), energy density, and other derived field-related features. This setup allows for evaluating the extent to which the model can predict magnetic flux density components using only design and positional information, without relying on field-derived physical proxy variables.

c) Design-Only Mode: In this mode, only component-level design parameters were used, with both spatial coordinates and physical field quantities completely excluded. As such, this mode does not aim to provide a complete solution for point-wise prediction of local fields; rather, it serves as an ablation scenario to examine how model performance changes when spatial and physical context is removed.

Through these three feature modes, the study systematically analyzes which level of information is critical for predicting magnetic flux density components. In particular, the contribution of spatial information and the additional benefit provided by physical field quantities are explicitly compared.

2.5. Random Forest

In this study, a Random Forest regression model was employed to analyze the magnetic flux density components (B_x, B_y, B_z). Random Forest is an ensemble-based method that combines multiple decision trees and is widely preferred due to its ability to model nonlinear relationships and its robustness against overfitting (Breiman, 2001). The final prediction is obtained by averaging the outputs of Tindividual decision trees:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x) \tag{1}$$

where $f_t(x)$denotes the output of the t-th decision tree.

The model was trained separately for each output component *(B_x, B_y, B_z)*, and identical hyperparameters were used across all feature modes. To ensure reproducibility, the number of trees was fixed at *n* estimators = 250, the maximum tree depth at max depth = 18, and the random seed at random state = 42. This setup allows performance differences to be attributed to the level of information provided by the feature modes, rather than variations in model complexity. It is important to note that the Random Forest model was not employed to propose a novel prediction architecture, but rather to provide a reliable and stable learning framework for interpretability-driven analysis. Moreover, its inherent structure is well-suited for feature effect interpretation, making it an appropriate choice for explainability studies. In this context, model outputs were further analyzed using SHAP and permutation importance methods. The objective of this study is not to identify the best-performing predictive algorithm, but to investigate model explainability and the role of information level (feature modes) on prediction behavior. Random Forest was selected as a nonlinear tree-ensemble backbone because it provides stable, well-established performance on large tabular datasets (Breiman, 2001) and supports efficient TreeSHAP-based post-hoc interpretation (Lundberg et al., 2020; Lundberg and Lee, 2017), not because it is intrinsically transparent; the explainability in this study is obtained through post-hoc analysis (SHAP and permutation importance) applied to the trained ensemble, rather than from the model architecture itself. A dedicated multi-model benchmarking study comparing Random Forest against alternative ensemble learners (e.g., XGBoost, LightGBM) on predictive accuracy is left for future work, as it is outside the explainability-focused scope of the present manuscript. We note that this approach differs from the position that intrinsically interpretable models should be preferred over post-hoc explanation of black-box models entirely (Rudin, 2019); post-hoc SHAP-based explanation of a high-performing tree ensemble remains a widely used and practical engineering approach for extracting physical insight, even where a fully transparent model is not adopted.

2.6. Evaluation Metrics

In this study, model performance was evaluated using the coefficient of determination (R^2), mean absolute error (MAE), and root mean squared error (RMSE) as the primary metrics. These measures provide complementary perspectives by assessing both prediction accuracy and error magnitude. The coefficient of determination (R^2) indicates the proportion of variance in the target variable that is explained by the model. Values of R^2closer to 1 indicate a higher level of agreement between predicted and actual values. It is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \hat{y}\right)^2} \tag{2}$$

where y_i denotes the true value, $\hat{y}_i$ the predicted value, and $\bar{y}$ the mean of the target variable.

The mean absolute error (MAE) represents the average of the absolute differences between predicted and actual values, providing a direct and easily interpretable measure of error magnitude:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y_i - \hat{y}_i\right| \tag{3}$$

The root mean squared error (RMSE), on the other hand, penalizes larger errors more heavily and is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{4}$$

Together, these three metrics enable a comprehensive evaluation of both the overall model fit and the error characteristics.

In addition, symmetric mean absolute percentage error (SMAPE) and relative absolute error (RAE) were computed as supplementary metrics. However, it is important to note that SMAPE can become unstable in regions where the target values approach zero. Therefore, the primary interpretation of the results is based on R^2, MAE, and RMSE. The MAE and RMSE values in Table 5 are computed on the standardized (z-scored) target variables used directly for model training and evaluation, consistent with common practice for multi-output regression with targets of very different physical scales. Converting to physical units (using the population mean and standard deviation of each raw target: B_x: mean=1.21×10^{-5}, std=6.252×10^{-3}; B_y: mean=1.126×10^{-3}, std=1.287×10^{-2}; B_z: mean=2.06×10^{-6}, std=3.601×10^{-3} Vs/m²), the Full Mode MAE corresponds to approximately 1.14×10^{-4} Vs/m² (B_x), 1.17×10^{-4} Vs/m² (B_y), and 3.72×10^{-5} Vs/m² (B_z) — all small relative to the physical range of each component. R^2 and RAE are invariant to this affine (mean-centering plus scaling) transformation and are therefore unaffected by this clarification. SMAPE is not invariant to mean-centering, since it depends on the absolute distance of each value from zero rather than on relative differences alone; the SMAPE values in Table 5 should accordingly be read as computed on the standardized target scale, where zero corresponds to the population mean of each component rather than to a physically zero flux density.

2.7. Computational Cost and Data Quality Checks

To address computational-cost reporting, wall-clock training and prediction times were measured for each feature mode on a workstation with a 20-logical-core Intel processor (scikit-learn 1.7.2, Python 3.10.16). Results are summarized in Table 4: training time decreases substantially as the feature set is reduced (from 139.5 s per target in the Full Mode to 70.8 s in the Design–Spatial Mode and 19.0 s in the Design-Only Mode, for ~1.01 million training samples), while prediction throughput exceeds 5.9×10^5 samples/s in all modes and reaches 9.6×10^5 samples/s in the Design-Only Mode, confirming that inference cost is negligible relative to training cost across all configurations.

Table 4. Mean per-target training and prediction time across the three feature modes (n = 1,010,350 training samples; n = 175,716 test samples per mode).

Feature Mode	Train time (s/target, mean)	Predict time (s, mean)	Predict time (ms/sample)	Throughput (samples/s)
Full	139.50	0.294	0.00167	597,719
Design–Spatial	70.81	0.289	0.00164	610,372
Design-Only	18.95	0.184	0.00104	959,347

A data-quality screening was performed directly on the raw, un-standardized CST simulation output (1,186,066 samples across all 27 configurations, in physical units), rather than on the model-input feature space, to ensure the reported statistics reflect true physical quantities. No missing values and no fully duplicate rows were found (0 duplicate rows, 0.0000%). Design parameters take their intended engineering values: magnet level and coil turns each span 800–1000 (mean 900, std 81.65, reflecting the three tested levels), and the top-cover relative permeability spans 1 (Brass) to 1000 (St 37), with Alpaca at 30; the body and shaft permeability are fixed at 1000. Spatial coordinates span approximately –24.4 to +18.1 mm (H_x), –29.9 to +31.1 mm (H_y), and –10.0 to +35.0 mm (H_z). The magnetic field intensity components (H_x, H_y, H_z, in A/m) and the target flux-density components (B_x, B_y, B_z, in Vs/m²) are strongly peaked and long-tailed: for example, B_x [A/m] ranges from −81,202 to +81,351 A/m with an interquartile range of only ~3.6×10^{-5} A/m, and B_y [Vs/m²] ranges from −0.0766 to +0.1109 Vs/m² with a mean of just 0.0011 Vs/m². This confirms, using real physical units, that most spatial sample points lie far from the region of concentrated magnetic flux near the actuator core — where field values are effectively near-zero — while a small fraction of points close to the magnet/coil assembly exhibit much larger field magnitudes. Consistently, an IQR-based outlier screen (recomputed on this raw data) flags approximately 35–36% of samples in the field-intensity and flux-density columns as statistical outliers under this rule; this matches the earlier screen performed on the standardized model-input features almost exactly, confirming that the finding is a genuine property of the physical field distribution rather than an artifact of preprocessing. No rows were removed on this basis.

3. Results and Discussion

When the results obtained under the three feature modes are evaluated collectively, it is evident that the predictability of the magnetic flux density components strongly depends on the level of information provided to the model. In particular, the performance differences observed between the Full Mode, Design–Spatial Mode, and Design-Only Mode clearly demonstrate the critical role of spatial information and field-derived physical quantities in modeling local magnetic behavior.

Table 5. Comparative prediction performance across the three feature modes.

Mode	Output	R²	MAE	RMSE	SMAPE (%)	RAE
Full Mode	B_x	0.9152	0.0183	0.3031	2.55	0.0707
	B_y	0.9910	0.0091	0.0926	2.57	0.0318
	B_z	0.9941	0.0103	0.0780	41.80	0.0492
Design–Spatial Mode	B_x	0.9846	0.0260	0.1293	92.51	0.1005
	B_y	0.9916	0.0110	0.0896	3.53	0.0384
	B_z	0.9658	0.0237	0.1872	27.01	0.1126
Design-Only Mode	B_x	0.0194	0.2548	1.0303	198.41	0.9867
	B_y	0.1015	0.2762	0.9272	139.76	0.9674
	B_z	0.0063	0.2080	1.0095	113.02	0.9902

Table 5 reports results from a single representative train/test configuration split, used for the detailed error analysis, SHAP interpretation, and per-configuration comparison presented throughout this section. Run-to-run variability across five independently repeated splits is reported separately in Table 7; this variability is small for B_y and B_z (std ≤ 0.007) but more substantial for B_x in the Full Mode (mean R^2 = 0.877, std = 0.133 across repeats, compared with R^2 = 0.9152 for the specific split shown in Table 5). Readers seeking a robustness-aware estimate of expected performance, rather than a single-split value, should reference Table 7 alongside Table 5. The results obtained under the Full Mode show that when all input variables are utilized, the model achieves particularly high accuracy for the B_y and B_z components. The corresponding R^2 values for B_x, B_y, and B_z are 0.9152, 0.9910, and 0.9941, respectively. These findings indicate that the model is able to construct a strong representation when provided with the richest level of physical information. However, the relatively lower accuracy observed for the B_x component compared to the other two suggests that it may involve more complex spatial interactions. Therefore, the Full Mode is not considered the most practical engineering solution, but rather a reference upper-bound scenario representing the highest level of available physical information.

In the Design–Spatial Mode, despite the exclusion of field-derived physical quantities from the input set, the model continues to achieve high performance for the $B_x$and $B_y$components. The corresponding R^2values are 0.9846, 0.9916, and 0.9658, respectively. Notably, for the $B_x$component, the Design–Spatial Mode yields a higher R^2value than the Full Mode. While the $B_y$component is predicted with similarly high accuracy in both modes, a noticeable performance drop is observed for $B_z$in the Design–Spatial Mode. This suggests that the $B_z$component is more sensitive to additional field-related physical information. Overall, the Design–Spatial Mode offers the most balanced solution, combining high predictive performance with a simpler and more interpretable feature representation from an engineering perspective. A particularly noteworthy finding of this study is that the Design–Spatial Mode outperforms the Full Mode for the $B_x$component. Although this may initially appear counterintuitive, it highlights that the inclusion of additional physical variables does not necessarily improve model performance. In the Full Mode, magnetic field intensity components and derived physical variables, while highly correlated with the target, may introduce redundancy and noise, thereby complicating the learning process—especially for the $B_x$component. In contrast, the Design–Spatial Mode provides a more compact feature space, allowing the model to focus more directly on spatial patterns. This likely contributes to improved generalization performance, particularly for components that are more sensitive to spatial variation, such as B_x. This finding emphasizes that more information does not al-

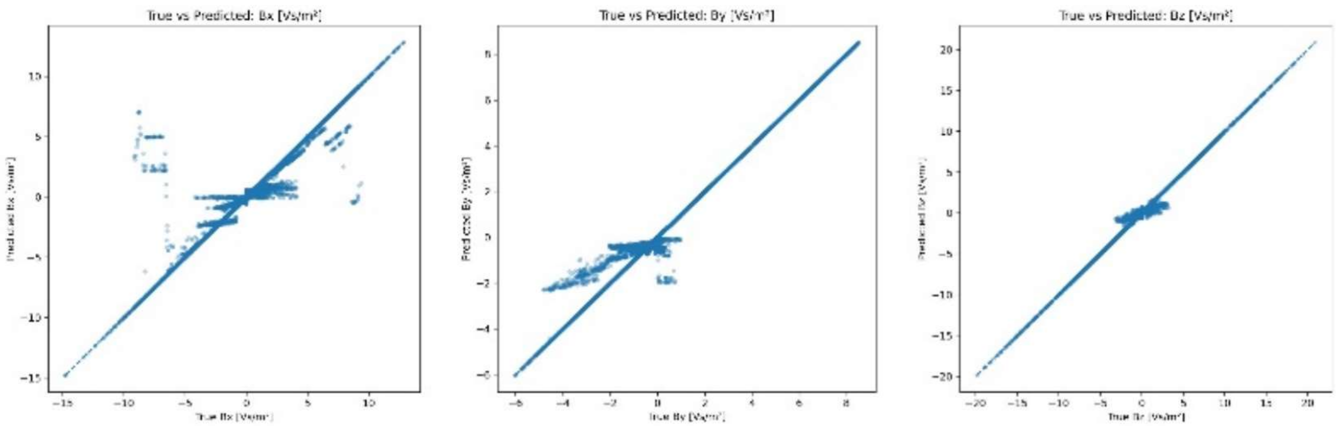

Figure 5. Scatter plots of true versus predicted values for the B_x, B_y and B_z components under the Full Mode.

ways lead to better model performance, and that appropriate feature selection plays a critical role.

The exact numerical values underlying Figure 6 are reported in Table 5 (R², MAE, RMSE, SMAPE, and RAE for each mode and output component), so that precise figures are available without relying on visual estimation from the chart. In contrast, the Design-Only Mode exhibits a substantial degradation in performance across all output components. The R^2 values for B_x, B_y, and $B_z$are 0.0194, 0.1015, and 0.0063, respectively. At first glance, these results may suggest that design parameters alone are insufficient. However, the nature of the problem must also be considered. Since local magnetic flux density components are inherently dependent on specific spatial locations, removing spatial coordinates makes it fundamentally difficult for the model to reconstruct point-wise field behavior. Therefore, the Design-Only Mode should not be interpreted as a direct prediction scenario, but rather as an ablation setting that demonstrates how model performance collapses when spatial and physical context is removed, clearly highlighting the indispensable role of spatial information.

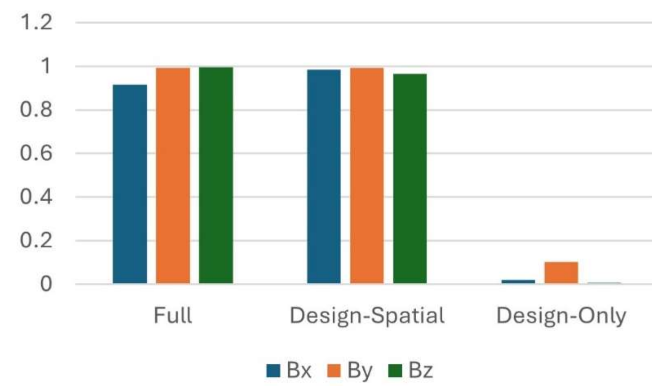

Figure 6. Comparison of R^2values across the three feature modes.

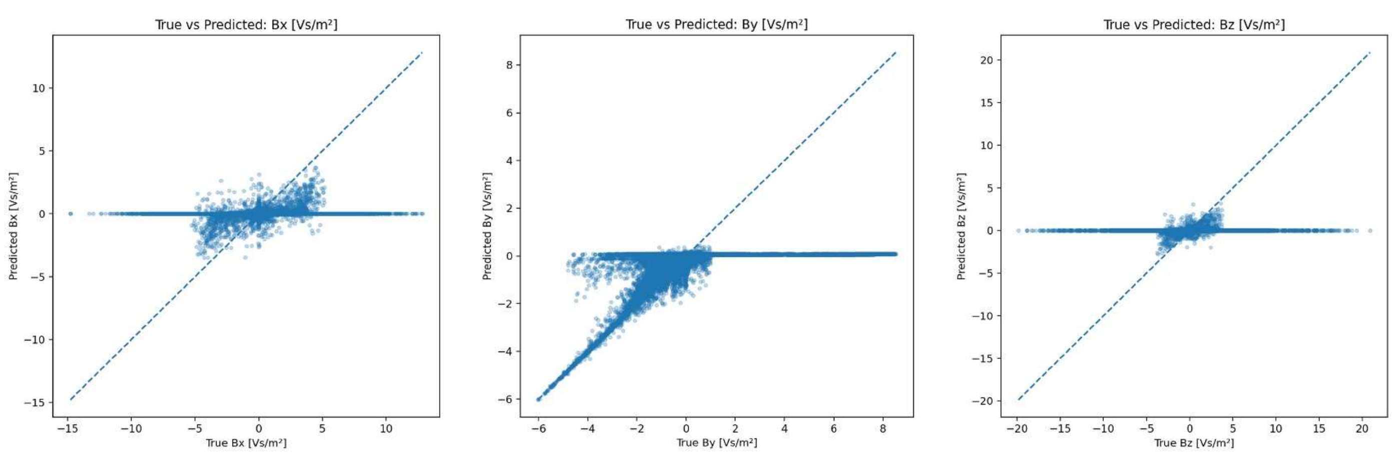

Figure 7. Comparison of true and predicted values for the Design-Only Mode.

Global feature importance analysis further supports these findings. In the Full Mode, the most dominant variables are the spatial y-coordinate and the magnetic field intensity components H_x, H_z, and H_y. These are followed by relative permeability, derived field quantities, and energy density. This indicates that the model relies on both spatial location and physically meaningful field-related variables when interpreting local magnetic behavior. In the Design–Spatial Mode, as shown in Figure 8, the most influential features are the spatial coordinates x, z, and y. Relative permeability follows these variables, while the coil parameter contributes at a lower but still meaningful level. This suggests that, when field-derived physical quantities are removed, the model primarily relies on spatial patterns to make predictions, reinforcing the dominance of spatial encoding in local field estimation.

SHAP analysis provides deeper insight by revealing not only which variables are important, but also the direction and magnitude of their influence on the model output. The SHAP summary plots obtained for the Full Mode show that magnetic field intensity components and spatial coordinates exert strong directional effects, particularly for the $B_y$and $B_z$outputs. This finding further supports the physically rich representation provided by the Full Mode. Quantitatively, mean |SHAP| values in the Full Mode confirm this directional pattern in a physically consistent way: for B_x, the dominant contributor is H_x [A/m] (mean |SHAP| = 0.200), consistent with B_x being governed by the horizontal field-intensity component; for B_y, H_y [A/m] dominates (0.284), matching the axis-aligned relationship between B_y and H_y; and for B_z, H_z [A/m] dominates (0.211), matching the same axis-aligned relationship for the vertical component. In all three components, real (measured) permeability is consistently the second most influential variable (mean |SHAP| ≈ 0.04–0.05), reflecting its role in modulating flux density through material behavior, while design parameters such as coil turns and magnet level contribute negligibly (mean |SHAP| below 0.001) once the field-derived quantities are available. For the Design–Spatial Mode, the SHAP summary plots indicate that spatial coordinates play a central role in the model's decision-making process. In particular, the positional variables (x, y, and z) and relative permeability exhibit clear and consistent effects on the predictions.

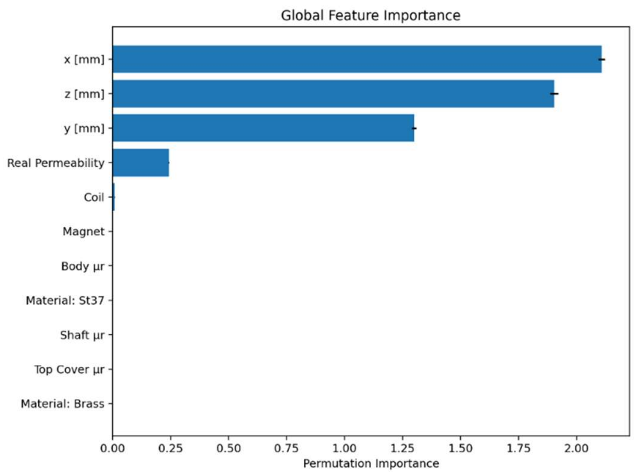

Figure 8. Global feature importance for the Design–Spatial Mode.

Table 6. Permutation feature importance (mean ± std) for the Design–Spatial Mode, corresponding to Figure 8.

Feature	Importance (mean)	Std
x [mm]	2.108	0.014
z [mm]	1.903	0.017
y [mm]	1.300	0.010
Real permeability	0.242	0.001
Coil	0.0092	0.0003
Magnet	0.0011	0.00004

This demonstrates that, even in the absence of field-derived physical quantities, the model is capable of capturing meaningful relationships through spatial patterns, highlighting the structural role of spatial encoding in governing model behavior. Consistent with the permutation-importance ranking reported in Table 6, The permutation importance values for the Design–Spatial Mode confirm that the spatial coordinates x [mm] (2.108) and z [mm] (1.903) exert the strongest influence, followed by y [mm] (1.300) and real permeability (0.242), whereas the coil and magnet parameters make substantially smaller contributions (≤0.009). This pattern is physically expected: with field-intensity quantities removed, the model must reconstruct local flux density primarily from spatial position, and permeability remains informative because it directly scales the material's magnetic response. In contrast, the SHAP outputs obtained for the Design-Only Mode are limited in interpretability due to the significantly lower prediction performance, reflecting the constraints imposed by insufficient information rather than providing strong physical insights.

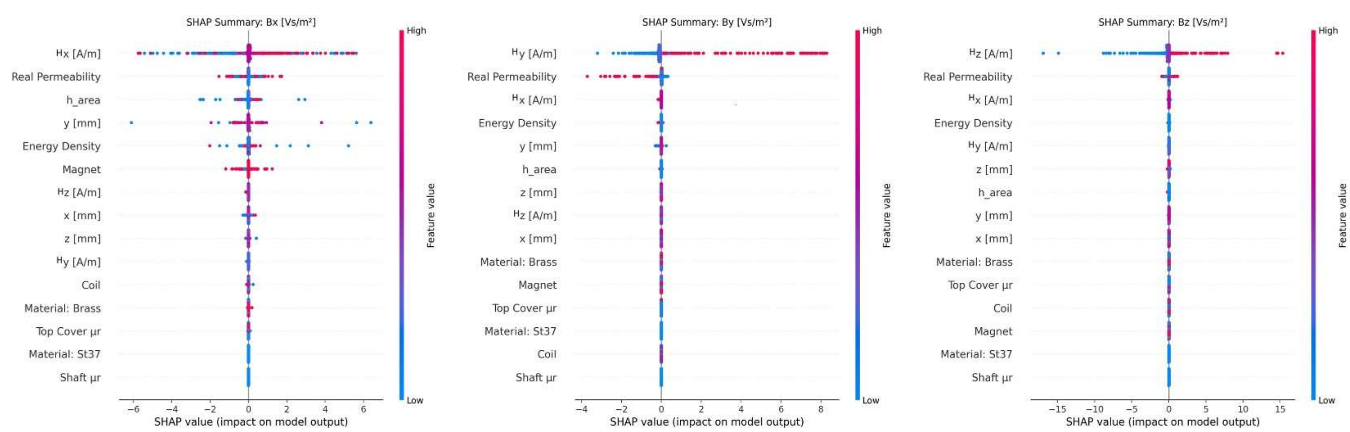

Figure 9. SHAP summary plots for the Full Mode.

Model stability was further evaluated through repeated configuration-based experiments, as summarized in Table 7. In the Full Mode, the mean R^2values across repetitions were 0.8766, 0.9924, and 0.9974 for B_x, B_y, and B_z, respectively. In the Design–Spatial Mode, the corresponding values were 0.9361, 0.9944, and 0.9649. In contrast, the Design-Only Mode yielded substantially lower mean R^2values of 0.0184, 0.0792, and 0.0066. These results indicate that the Design–Spatial Mode provides a more balanced and stable performance, particularly for the $B_x$and $B_y$components, while the Full Mode remains more advantageous for predicting B_z. Notably, the relatively high standard deviation observed for the $B_x$component in the Full Mode (±0.133) suggests that model performance is sensitive to different test configurations. In contrast, the Design–Spatial Mode exhibits a significantly more consistent performance profile, with a lower standard deviation of ±0.048 for B_x, highlighting its superior robustness across unseen configurations. To explore whether the performance differences between feature modes are statistically meaningful, paired two-sided t-tests were conducted across the five repeated, identically-partitioned experiments for each output component, with a Holm-Bonferroni correction applied across the resulting nine pairwise comparisons (three mode pairs × three components) to control the family-wise error rate. Full Mode and Design–Spatial Mode were not significantly different for B_x or B_y after correction (B_x: t = -1.31, p_{Holm} = 0.52; B_y: t = -1.07, p_{Holm} = 0.52), while both significantly outperformed the Design-Only Mode for all three magnetic flux density components (six component-wise comparisons in total: two mode pairs × three components; p_{Holm}< 0.001 in each). For B_z, the Full Mode significantly outperformed the Design–Spatial Mode (t = 26.62, p_{holm}< 0.001), consistent with B_z benefiting more from field-derived physical quantities. These exploratory results are consistent with the Design–Spatial Mode being a practically competitive alternative to the Full Mode for

B_x and B_y, while the Design-Only Mode is consistently and substantially inferior across all components; a stronger statistical design — for example grouped k-fold cross-validation across more of the 27 configurations, or a per-configuration held-out analysis — would be needed to confirm these patterns with greater confidence, and we note this as a direction for future work alongside uncertainty quantification. Given the limited number of paired repetitions (n=5), these statistical findings should be interpreted as exploratory.

Table 7. Stability analysis results (R^2 mean ± std).

Mode	B_x	B_y	B_z
Full Mode	0.877 ± 0.133	0.992 ± 0.007	0.997 ± 0.003
Design–Spatial Mode	0.936 ± 0.048	0.994 ± 0.003	0.965 ± 0.003
Design-Only Mode	0.018 ± 0.004	0.079 ± 0.066	0.007 ± 0.001

High SMAPE values observed in certain modes also warrant attention. In particular, elevated SMAPE values are observed for $B_x$in the Design–Spatial Mode, for $B_z$in the Full Mode, and across all outputs in the Design-Only Mode. This behavior is primarily due to the presence of target values that are close to zero, where percentage-based error metrics tend to become unstable. Therefore, the primary interpretation of the results is based on R^2, MAE, and RMSE. Overall, while the Full Mode represents the most physically rich scenario, the Design–Spatial Mode emerges as the most balanced solution, offering high predictive performance alongside a cleaner and more interpretable feature representation. The Design-Only Mode, on the other hand, serves not as a final predictive solution but as an ablation scenario, illustrating how performance degrades when spatial and physical context is removed. The key finding of this study is that spatial information is critical for the reliable prediction of local magnetic flux density components, and that representations based solely on design parameters are insufficient for accurate point-wise field estimation.

Conclusions

In this study, a mode-based analysis of magnetic flux density components (B_x, B_y, B_z) in MCCB magnetic actuators was conducted using an explainable machine learning approach. Three distinct feature modes were defined and systematically compared under a configuration-based data splitting strategy. The findings demonstrate that model performance depends not only on the chosen learning algorithm, but more critically on the nature and level of information provided to the model. The Full Mode, offering the most physically comprehensive representation, achieved high accuracy, particularly for the $B_y$and $B_z$components. The Design–Spatial Mode, despite excluding field-derived physical quantities, largely preserved its performance and produced balanced and robust results, especially for the $B_x$and $B_y$components. In contrast, the significant performance degradation observed in the Design-Only Mode indicates that representations based solely on design parameters are insufficient for reconstructing local magnetic field behavior.

The key outcome of this study is that spatial information plays a critical role in the reliable prediction of local magnetic flux density components. Explainability analyses further support this conclusion, revealing that model decisions are primarily driven by physically meaningful variables, including spatial coordinates, magnetic field intensity components, and relative permeability. In this context, the Design–Spatial Mode emerges as the most balanced solution, offering both high predictive performance and a more interpretable and practically meaningful engineering representation.

This work can be extended in several directions. First, incorporating uncertainty quantification (e.g., prediction intervals via quantile regression or bootstrap resampling) would enable the evaluation of prediction confidence; this is identified as a planned direction for future work rather than part of the present scope, which is instead focused on explainability under varying information levels. Second, adopting group-based cross-validation and alternative machine learning models could provide a more comprehensive assessment of generalization capability. Third, integrating explainability outputs into inverse design and optimization frameworks could facilitate the identification of design parameters that achieve targeted magnetic field behavior. Finally, validating the proposed approach using experimental data would strengthen the applicability of simulation-driven learning in real-world systems.

Beyond uncertainty quantification, several further directions are worth pursuing. Physics-informed machine learning, which embeds Maxwell's-equation-derived constraints directly into the learning process, could improve generalization beyond the sampled design space and has shown promise in related electromagnetic-field modeling tasks (Deng et al., 2025; Karniadakis et al., 2021; J. Wang et al., 2025; Wu et al., 2024). Transfer learning across actuator geometries or magnetic-material families could reduce the CST simulation burden for new configurations by reusing knowledge from the present dataset, as demonstrated for electromagnetic metamaterial design (Peng et al., 2024). Finally, coupling the present surrogate model with multi-objective optimization algorithms would allow the framework to move from prediction toward inverse design building on recent surrogate-assisted multi-objective optimization frameworks developed for electrical machines (L. Liu et al., 2025). Overall, the results indicate that the model captures not only data-driven correlations but also physically meaningful relationships, reinforcing the reliability of the proposed approach for electromagnetic system modeling.

Author Contributions (Mandatory)

Mehmet Onur YAĞIR: Conceptualization, visualization, manuscript writing, analysis

Sevda GÜL: Manuscript writing, methodology; data organization; analysis, visualization

Acknowledgements

No support was received from any organization or individual.

Conflicts of Interest Statement

The authors declare that they have no conflicts of interest regarding this study.

Funding Information

The corresponding author funded this study without any external funding support.

AI Statement

No artificial intelligence (AI) or AI-assisted tools were used in the preparation, writing, analysis, or editing of this manuscript.

References

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N. and Herrera, F. (2023), "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence", Information Fusion, Vol. 99, pp. 101805.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R. and Herrera, F. (2020), "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI", Information Fusion, Vol. 58, pp. 82-115.

Breiman, L. (2001), "Random Forests", Machine Learning, Vol. 45 No. 1, pp. 5-32.

Chen, H., Li, X., Zhou, C., Tan, Y. and Shen, Y. (2025), "Simulation Study on Electromagnetic Response and Cable Coupling Characteristics of eVTOL Under Lightning Environment", *Electronics*, pp. 3661.

Cho, Y.-M., Park, H.-J., Lee, J.-J. and Lee, K.-A. (2022), "Analysis of Characteristics of Low Voltage Circuit Breaker by External Magnetic Field", *Energies*, pp. 8156.
Deng, Y., Fan, K., Jin, B., Malof, J. and Padilla, W. J. (2025), "Physics-informed learning in artificial electromagnetic materials", Applied Physics Reviews, Vol. 12 No. 1, pp. 011331.
Givisis, I., Kalatzis, D., Christakis, C. and Kiouvrekis, Y. (2025), "Comparing Explainable AI Models: SHAP, LIME, and Their Role in Electric Field Strength Prediction over Urban Areas", *Electronics*, pp. 4766.
Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S. and Yang, L. (2021), "Physics-informed machine learning", Nature Reviews Physics, Vol. 3 No. 6, pp. 422-440.
Kiouvrekis, Y., Givisis, I., Panagiotakopoulos, T., Tsilikas, I., Ploussi, A., Spyratou, E. and Efstathopoulos, E. P. (2025), "A Comparative Analysis of Explainable Artificial Intelligence Models for Electric Field Strength Prediction over Eight European Cities", *Sensors*, pp. 53.
Liu, L., Li, Z., Kang, H., Xiao, Y., Sun, L., Zhao, H., Zhu, Z. Q. and Ma, Y. (2025), "Review of surrogate model assisted multi-objective design optimization of electrical machines: New opportunities and challenges", Renewable and Sustainable Energy Reviews, Vol. 215, pp. 115609.
Liu, S., Jin, J., Chen, J., Yu, M., Liu, Z. and Geng, Y. (2024), "Evolution and interruption characteristics of DC vacuum arc under Halbach transverse magnetic fields", International Journal of Electrical Power & Energy Systems, Vol. 157, pp. 109844.
Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S.-I. (2020), "From local explanations to global understanding with explainable AI for trees", Nature Machine Intelligence, Vol. 2 No. 1, pp. 56-67.
Lundberg, S. M. and Lee, S.-I. (2017), "A unified approach to interpreting model predictions", in *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Curran Associates Inc., Long Beach, California, USA, pp. 4768–4777.
Peng, R., Ren, S., Malof, J. and Padilla, W. J. (2024), "Transfer learning for metamaterial design and simulation", Vol. 13 No. 13, pp. 2323-2334.
Qoria, T., Gruson, F., Colas, F., Kestelyn, X. and Guillaud, X. (2020), "Current limiting algorithms and transient stability analysis of grid-forming VSCs", Electric Power Systems Research, Vol. 189, pp. 106726.
Rudin, C. (2019), "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead", Nature Machine Intelligence, Vol. 1 No. 5, pp. 206-215.
Tjoa, E. and Guan, C. (2021), "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI", IEEE Transactions on Neural Networks and Learning Systems, Vol. 32 No. 11, pp. 4793-4813.
Wang, J., Wang, D., Wang, S., Li, W. and Jiang, Y. (2025), "Dimensionless Physics-Informed Neural Network for Electromagnetic Field Modelling of Permanent Magnet Eddy Current Coupler", IET Electric Power Applications, Vol. 19 No. 1, pp. e70084.
Peng, R., Ren, S., Malof, J., & Padilla, W. J. (2024). Transfer learning for metamaterial design and simulation. Nanophotonics, 13(13), 2323–2334.
Wang, Q., Ge, S., Cao, W., Yu, S. and Liao, Z. (2023), "Study of Arc Interruption Characteristics under Rated Current in Low Voltage Circuit Breakers", *Energies*, pp. 4114.
Wu, Y., Sicard, B. and Gadsden, S. A. (2024), "Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring", Expert Systems with Applications, Vol. 255, pp. 124678.
Zha, D., Bhat, Z. P., Lai, K.-H., Yang, F., Jiang, Z., Zhong, S. and Hu, X. (2025), "Data-centric Artificial Intelligence: A Survey", Vol. 57 No. 5 ACM Computing Surveys., pp. Article 129.
Zhang, B., Zhang, J., Cheng, Y., Chen, Q. and Zhang, Q. (2025), "Simulation and Experimental Study of Arc Model in a Low-Voltage Distribution Network", *Energies*, pp. 420.