International Journal of Heat and Technology
Vol. 43, No. 6, December, 2025, pp. 2323-2338
Journal homepage: http://iieta.org/journals/ijht

Thermo–Electromagnetic Co-Simulation of Electric Motor Cooling Systems Using Particle Swarm Optimization

Zihao Shen

School of Civil Engineering, Yancheng Institute of Technology, Yancheng 224061, China

Corresponding Author Email: 13921826526@139.com

https://doi.org/10.18280/ijht.430630

Received: 5 October 2025
Revised: 2 November 2025
Accepted: 19 November 2025
Available online: 31 December 2025

Keywords:
electric motor cooling systems, thermo–electromagnetic coupling, PSO, surrogate model, co-simulation, adaptive optimization

ABSTRACT

The increasing efficiency and miniaturization of electric motors have made thermo–electromagnetic coupling a critical challenge in cooling system design. Conventional sequential simulation–optimization approaches suffer from high computational cost and difficulties in achieving high accuracy and efficiency. Moreover, existing particle swarm optimization (PSO) variants fail to adequately accommodate the physical characteristics of thermo–electromagnetic coupling. To address these limitations, an adaptive chaotic PSO framework integrating a dual-path surrogate model and thermal-state feedback was proposed, breaking the conventional serial paradigm and establishing a parallel, interactive simulation–optimization closed loop, thereby enabling a deep integration of the accuracy of physical simulation with the efficiency of intelligent optimization. The results demonstrated that the dual-path surrogate model achieved significantly higher predictive accuracy than a single model, with the average root mean square error (RMSE) reduced by 50.8% and further decreased to 1.75% after calibration through the interactive validation closed loop. Compared with purely high-fidelity simulation-based optimization, the proposed method reduced computational time by 91.1% and decreased the number of convergence iterations by 47.1% relative to conventional PSO. It achieved reductions of 17.9% in the maximum motor temperature, 33.9% in cooling pumping power, and 23.5% in cooling system volume, while maintaining a high-fidelity validation error of only 2.1%. Ablation studies further confirmed that the three core modules–namely, the dual-path surrogate model, thermal-state feedback, and the interactive validation closed loop–are all indispensable. This study establishes a new paradigm for the efficient collaborative design of electric motor cooling systems and provides significant implications for enhancing motor power density and operational reliability.

1. INTRODUCTION

Driven by the rapid development of new energy vehicles and intelligent industrial manufacturing, electric motors have been advancing toward higher efficiency and increased compactness [1, 2], resulting in continuously rising power density and increasingly pronounced thermo–electromagnetic coupling effects. The elevation of winding temperature induces an increase in electrical resistivity [3], which in turn intensifies electromagnetic losses and forms a positive thermo–electromagnetic feedback loop [4, 5], substantially elevating the risk of thermal failure. Consequently, an efficient thermal management system has become a fundamental prerequisite for ensuring operational reliability and service life of electric motors [6]. As the primary carrier of thermal management, the design quality of the cooling system directly determines motor power density, operational efficiency, and durability [7, 8]. Conventional experience-based design approaches or optimization methods relying on a single physical field are unable to simultaneously satisfy multi-objective requirements such as temperature regulation, energy efficiency, and structural compactness [9-11], and therefore no longer meet the demands of modern electric motor design.

Thermo–electromagnetic co-simulation techniques provide an effective means for accurately characterizing coupling effects. High-fidelity simulations are capable of precisely resolving internal loss distributions and temperature field characteristics within electric motors; however, their prohibitive computational cost severely limits applicability in multi-parameter, multi-objective optimization scenarios requiring extensive iterations. Intelligent optimization algorithms—particularly PSO—have been widely adopted in motor design optimization due to their simple structure, rapid convergence, and ease of engineering implementation [12, 13]. Nevertheless, existing PSO variants fail to sufficiently accommodate the complex physical characteristics inherent to thermo–electromagnetic coupling. As a result, optimization processes lack explicit physical guidance, are prone to convergence toward designs exhibiting local thermal concentration, and struggle to achieve a coordinated improvement in both accuracy and computational efficiency [14, 15]. Accordingly, the development of a synergistic design methodology that integrates the physical fidelity of simulation with the efficiency of intelligent optimization has emerged as a critical pathway for addressing thermo–electromagnetic co-optimization challenges in electric motor cooling systems.

In the field of thermo–electromagnetic co-simulation of electric motors, extensive research efforts have been undertaken. With respect to high-fidelity simulation, multiphysics coupling techniques based on the finite element method (FEM) and finite volume method (FVM)—including electromagnetic–fluid–thermal coupled simulations—have become the mainstream technical approach for accurately analyzing thermo–electromagnetic coupling phenomena [16, 17]. However, the inherent limitation of low computational efficiency has remained largely unresolved. To balance accuracy and efficiency, surrogate model–assisted simulation techniques have been rapidly developed. Models such as artificial neural networks (ANNs), Gaussian process regression (GPR), and response surface methods have been widely employed for fast prediction of motor losses or temperature fields [18, 19]. Nevertheless, most existing studies rely on a single surrogate model architecture, which is insufficient to simultaneously accommodate the dual-path coupling characteristics of electrical losses and thermal temperature fields. As a result, predictive accuracy and generalization capability are limited, rendering such approaches inadequate for high-precision optimization requirements.

The application of PSO and its variants in electric motor design optimization has continued to expand. Conventional PSO has been successfully applied to single-objective optimization problems, such as winding parameters and cooling structure dimensions. To alleviate premature convergence, strategies including chaotic mapping and adaptive parameter regulation have been incorporated, leading to enhanced population diversity and improved convergence behavior [20]. However, most existing PSO variants primarily focus on mathematical or algorithmic enhancements [21, 22], while insufficient attention has been paid to the underlying physical mechanisms of thermo–electromagnetic coupling. Consequently, optimization processes lack sensitivity and responsiveness to temperature field states, often steering solutions toward designs with localized thermal concentration and failing to achieve a globally optimal balance between thermal and electromagnetic performance.

Regarding simulation–optimization co-design, existing studies predominantly adopt a serial paradigm of "simulation followed by optimization," in which simulation data are first generated to train surrogate models, and optimization is subsequently performed based on these models. Under this paradigm, surrogate model accuracy remains static and cannot adapt to the dynamically evolving design space during optimization, leading to accuracy degradation in later optimization stages. Moreover, the separation between simulation and optimization prevents simulation data from dynamically feeding back into model refinement, thereby hindering coordinated improvements in accuracy and efficiency. Although intermittent simulation-based calibration mechanisms have been explored in some studies, a dynamic interactive closed-loop framework has not yet been established. As a result, the fundamental imbalance between accuracy and efficiency remains unresolved, and the lack of a dynamic interaction–driven co-design framework constitutes a critical research gap.

In response to the core limitations identified in existing studies—including insufficient surrogate model adaptability, the absence of physical guidance in optimization processes, and weak coordination between simulation and optimization—the primary objective of this study is to develop a thermo–electromagnetic co-design methodology for electric motor cooling systems that integrates physical simulation with intelligent optimization. Through this integration, simultaneous improvements in optimization efficiency, predictive accuracy, and engineering applicability are targeted.

To achieve this objective, the principal innovations are summarized below. First, a dual-path surrogate model architecture is proposed, in which dedicated surrogate submodels are constructed for electrical loss prediction and thermal temperature field prediction, respectively. Coupled parameters are introduced to enable bidirectional data interaction between the two paths, allowing accurate representation of the dual-path characteristics inherent to thermo–electromagnetic coupling and significantly enhancing both predictive accuracy and generalization capability. Second, a thermal-state feedback–driven adaptive PSO strategy is developed, in which temperature field uniformity metrics are embedded into the parameter regulation mechanism. As a result, the optimization process is endowed with real-time awareness and responsiveness to physical field states, providing explicit physical guidance and effectively mitigating designs prone to localized thermal concentration. Third, a dynamic optimization–high-fidelity simulation interactive verification loop is established. Through periodic simulation–based calibration of optimal solutions during optimization, surrogate model bias is dynamically corrected, while newly generated high-fidelity simulation data are continuously incorporated into the training sample database, enabling the co-evolution of surrogate model accuracy and optimization reliability.

The remainder of this study is organized below. Section 2 presents the fundamental mechanisms of thermo–electromagnetic coupling in electric motors, the basis of PSO, and the core theories of the surrogate model, followed by the mathematical modeling of the thermo–electromagnetic co-optimization problem. Section 3 details the proposed adaptive chaotic PSO–based thermo–electromagnetic co-simulation framework, including dual-path surrogate model construction, adaptive optimizer design, and implementation of the dynamic interactive closed-loop process. Section 4 conducts comparative experimental validation using a 200 kW permanent magnet synchronous drive motor with a liquid-cooled system as the case study, with emphasis on optimization efficiency, accuracy, and engineering feasibility. Section 5 provides an in-depth discussion of the physical implications of the results, the engineering applicability and limitations of the proposed method, and potential directions for future research. Section 6 concludes with a summary of the main findings and clarifies the academic contributions and engineering significance of the study.

2. THEORETICAL BACKGROUND AND PROBLEM FORMULATION

2.1 Thermo–electromagnetic coupling mechanism of electric motors

The essence of thermo–electromagnetic coupling in electric motors lies in the mutual interaction between electrical losses and the temperature field. Electrical losses constitute the primary heat sources driving the evolution of the temperature field, while the resulting temperature rise, in turn, alters the electrical properties of motor materials, thereby modifying the spatial distribution of losses and forming a closed-loop coupled effect. Electrical losses mainly consist of copper losses and iron losses, both of which exhibit pronounced temperature dependence. Copper losses are governed by Joule's law, with consideration of the temperature dependence of winding resistivity. The copper loss can be expressed as $P_{cu}=\int J^2 \rho(T)dV$, where J denotes the stator current density, $\rho(T)$ represents the temperature-dependent electrical resistivity, and V is the winding volume. In addition, $\rho(T)=\rho_0[1+\alpha(T\text{-}T_0)]$, where ρ_0 is the resistivity at the reference temperature T_0, and α is the temperature coefficient of resistance. Iron losses are composed of hysteresis losses, eddy current losses, and additional losses, and are commonly evaluated using a modified Steinmetz formulation $P_{fe}=k_{\mathrm{h}}fB^n+k_ef^2B^2+k_af^{1.5}B^{1.5}$, where k_h, k_e, and k_a denote the hysteresis, eddy current, and additional loss coefficients, respectively; f is the magnetic field frequency; and B is the magnetic flux density amplitude. Temperature alters magnetic permeability, which subsequently affects the effective values of these loss coefficients.

Heat transfer within the motor is realized through conduction, convection, and radiation, among which the design of the cooling system directly governs convective heat transfer efficiency and thus represents the core of thermal management. Heat conduction in solid domains follows Fourier's law, and the corresponding governing equation is expressed as:

$$\rho_c c_c \frac{\partial T}{\partial t} = \nabla \cdot (k_c \nabla T) + q_v \tag{1}$$

where, ρ_c, c_c, and k_c denote the density, specific heat capacity, and thermal conductivity of the solid material, respectively, while q_v represents the volumetric heat generation rate. Fluid flow and convective heat transfer within the cooling channels are governed by the continuity equation, the Navier–Stokes equations, and the energy equation. The energy equation is expressed as:

$$\rho_f c_f \left(\frac{\partial T_f}{\partial t} + \vec{u} \cdot \nabla T_f\right) = \nabla \cdot \left(k_f \nabla T_f\right) \tag{2}$$

where, ρ_f, c_f, and k_f are the density, specific heat capacity, and thermal conductivity of the coolant, respectively, and u denotes the fluid velocity vector. The coupling between convective heat transfer and solid heat conduction is realized through thermal flux continuity conditions at the solid–fluid interfaces.

The core of the thermo–electromagnetic coupling feedback loop lies in the loss–temperature positive feedback induced by the temperature dependence of electrical resistivity. As winding temperature increases, electrical resistivity rises, leading to increased copper losses; the additional heat generation further elevates temperature. If the heat dissipation capacity of the cooling system is insufficient, thermal accumulation may occur, potentially resulting in thermal failure. This coupled feedback mechanism can be mathematically described through the temperature-dependent loss relationship:

$$P_{\mathrm{loss}}(T) = P_{\mathrm{cu}}(T) + P_{fe}(T) \tag{3}$$

where, k_{conv} denotes the convective heat transfer coefficient, which is determined by the structural parameters of the cooling system and the properties of the fluid. The temperature coefficient of resistance α and the convective heat transfer coefficient k_{conv} serve as key parameters characterizing coupling intensity. The former determines the sensitivity of electrical losses to temperature variations, whereas the latter governs the effectiveness with which generated heat can be dissipated through the cooling system. Together, these parameters dictate the steady-state behavior and transient response characteristics of the thermo–electromagnetic coupled system.

2.2 Fundamentals of PSO

PSO performs optimization searches by emulating the collective foraging behavior of bird flocks. The fundamental mechanism lies in the iterative updating of particle positions and velocities, through which the swarm progressively converges toward regions containing optimal solutions. In the classical PSO framework, each particle represents a candidate solution within the design space, and its velocity and position are updated according to the following formulation:

$$\vec{v}_{i,d}^{\,t+1}=w\vec{v}_{i,d}^{\,t}+c_1r_1(\vec{p}_{i,d}^{\,t}\text{-}\vec{x}_{i,d}^{\,t})+c_2r_2(\vec{p}_{g,d}^{\,t}\text{-}\vec{x}_{i,d}^{\,t}),\vec{x}_{i,d}^{\,t+1}=\vec{x}_{i,d}^{\,t}+\vec{v}_{i,d}^{\,t+1} \tag{4}$$

where, $\vec{v}_{i,d}^{\,t}$ and $\vec{x}_{i,d}^{\,t}$ denote the velocity and position of the i-th particle in the d-th dimension at iteration t, respectively. The inertia weight w controls the degree to which historical velocity information is retained, thereby balancing global exploration and local exploitation. The parameters c_1 and c_2 represent the individual and social learning factors, respectively, governing particle attraction toward its personal best position $\vec{p}_{i,d}^{\,t}$ and the global best position $\vec{p}_{g,d}^{\,t}$. The variables r_1 and r_2 are uniformly distributed random numbers in the interval [0,1], introduced to enhance search diversity.

In multi-objective particle swarm optimization (MOPSO), the primary objective is the identification of a Pareto-optimal solution set. The central concept underpinning this process is Pareto dominance: a solution $\vec{x}_a$ is said to dominate another solution $\vec{x}_b$ if $\vec{x}_a$ is no worse than $\vec{x}_b$ with respect to all objective functions and is strictly better in at least one objective. A Pareto-optimal solution is defined as a solution that is not dominated by any other solution, and the collection of all such solutions forms the Pareto front. The solution strategy in MOPSO relies on maintaining an external archive to store non-dominated solutions. This archive is continuously updated during the iterative process, enabling the eventual extraction of a well-distributed Pareto front. The construction of the fitness function is critical to multi-objective optimization performance. Common approaches include the weighted-sum method and the ε-dominance method. The weighted-sum method transforms a multi-objective problem into a single-objective formulation by assigning predefined weights to each objective, making it suitable for scenarios with clearly defined preference structures. In contrast, the ε-dominance method relaxes strict dominance conditions by introducing an ε-threshold, effectively preserving diversity along the Pareto front and mitigating solution clustering, which is particularly advantageous for complex multi-objective optimization problems.

2.3 Core theory of the surrogate model

The surrogate model is constructed by fitting data obtained from high-fidelity simulations or experiments to establish simplified mappings between inputs and outputs, thereby enabling rapid prediction of complex physical processes. The fundamental objective is to achieve a balance between accuracy and computational efficiency. In this study, neural networks and GPR are adopted as the primary surrogate model, as each exhibits distinct and complementary advantages. Neural networks are well suited for capturing highly nonlinear and complex mapping relationships, whereas GPR provides explicit quantification of predictive uncertainty, which is advantageous for accuracy assessment. Deep neural networks (DNNs) consist of input, hidden, and output layers, and training is performed by minimizing the discrepancy between predicted and reference values through backpropagation. Nonlinear mapping capability is introduced through activation functions, enabling adaptation to the prediction of high-dimensional and nonlinear problems, such as electrical losses. Convolutional neural networks (CNNs) extract spatial features of input data via convolutional and pooling layers, making them particularly effective for predicting outputs with spatial distribution characteristics, such as temperature fields.

GPR is formulated as a probabilistic model based on Gaussian processes, in which similarity between samples is defined through kernel functions, thereby constructing a probabilistic distribution of the outputs. Commonly used kernels include the squared exponential kernel, the Matérn kernel, and the rational quadratic kernel. Among these, the Matérn kernel exhibits superior robustness to noisy data and is therefore more suitable for fitting engineering simulation data. Training of GPR is performed by maximizing the marginal likelihood to determine optimal kernel function parameters. During prediction, both the mean and variance are obtained, with the variance serving as a direct quantitative measure of predictive uncertainty and providing an intuitive basis for surrogate model accuracy evaluation. Regardless of whether neural networks or GPR are employed, training performance is inherently dependent on the availability of sufficiently representative sample data. The distribution characteristics of the samples directly determine model generalization capability. Consequently, appropriate design of experiments (DoE) strategies are required to ensure adequate coverage of the design space.

Accuracy assessment of the surrogate model is critical to ensuring their reliability. Commonly adopted evaluation metrics include the coefficient of determination R^2, RMSE, and mean absolute error (MAE). R^2 is used to quantify the goodness of fit between predicted and reference data and is defined as:

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{5}$$

where, y_i denotes the reference (true) value, $\hat{y}_i$ represents the predicted value, and $\bar{y}$ is the mean of the reference values. A value of R^2 closer to unity indicates superior model fitting performance. The RMSE characterizes the square root of the mean squared prediction error and is sensitive to outliers, whereas the MAE reflects the average magnitude of prediction errors and exhibits greater robustness to outliers. Joint utilization of these three metrics enables a comprehensive evaluation of surrogate model accuracy, satisfying the precision requirements of iterative optimization. The RMSE and MAE are computed as:

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{6}$$

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{7}$$

2.4 Thermo–electromagnetic co-optimization modeling of the electric motor cooling system

The core of the thermo–electromagnetic co-optimization problem for electric motor cooling systems lies in the coordinated optimization of electrical and thermal design parameters under thermo–electromagnetic coupling constraints, with the objective of achieving an optimal balance among multiple performance criteria. The design variables are required to span the entire thermo–electromagnetic interaction chain and are classified, according to physical attributes, into electrical path variables, thermal path variables, and coupling interface variables. Electrical path variables include the stator current density J, winding turn number N, and air-gap magnetic flux density B_g. These parameters directly determine the intensity of electromagnetic loss generation and are constrained by rated motor power and torque requirements. For example, $J \in [2.5, 5.0]$ A/mm^2 and $N \in [80, 120]$, ensuring a balance between loss mitigation and power performance. Thermal path variables encompass the cooling channel diameter d, coolant flow velocity v, fin spacing s, and fin height h, which directly govern convective heat transfer efficiency. The admissible ranges of these variables are constrained by manufacturing feasibility and installation space limitations, e.g., $d \in [8,15]$ mm and $s \in [5,12]$ mm, avoiding excessive fabrication difficulty and sharp increases in flow resistance associated with overly small dimensions. The coupling interface variables are centered on the real-time winding temperature T_w, which serves as the critical intermediary linking electrical losses to the temperature field. Its dynamic evolution constitutes the fundamental feedback mechanism underlying thermo–electromagnetic coupling.

The optimization objectives focus on thermal management performance, energy efficiency, and structural compactness of the cooling system, resulting in three mutually constrained objective functions. The first objective aims to minimize the maximum motor temperature T_{max} to ensure operational reliability, expressed as $\min f_1(x) = T_{max}(x)$, subject to the temperature safety constraint $T_{max} \leq 155$°C. The second objective seeks to minimize the cooling system pumping power P_{pump}, thereby reducing auxiliary energy consumption. The pumping power is quantified based on the relationship between channel pressure drop and volumetric flow rate, given by:

$$\min f_2(x)=P_{pump}(x)=\Delta P(x)\cdot Q(x)/\eta_{pump} \tag{8}$$

where, ΔP denotes the total pressure drop across the cooling channels, Q represents the coolant volumetric flow rate, and η_{pump} is the pump efficiency. The third objective aims to minimize the cooling system volume V_{cool} in order to enhance motor power density. This objective is evaluated based on the integrated geometric parameters of the cooling channels and fins, expressed as:

$$\min f_3\,(\mathrm{x})=V_{cool}(\mathrm{x})=f(d,s,\mathrm{h},L) \tag{9}$$

where, L denotes the cooling channel length. The constraint set is further refined into physical constraints, engineering constraints, and mathematical constraints. Physical constraints include an upper limit on cooling channel pressure drop ($\Delta P \leq 200$ kPa) and torque compliance under rated operating conditions. Engineering constraints account for manufacturing process limitations and geometric boundaries imposed by installation space. Mathematical constraints explicitly define the admissible ranges of all design variables, thereby establishing a well-bounded design space.

Based on the aforementioned design variables, objective functions, and constraints, the problem can be formulated as a standard multi-objective optimization model, expressed in a general form as:

$$\begin{gathered} \min f(x)=[f_1(x),f_2(x),f_3(x)]^T \\ \text{s.t. } g_k(\mathrm{x})\leq 0 \ (k=1,2,\ldots,m) \\ h_l(\mathrm{x})=0 \ (l=1,2,\ldots,n) \ x\in X=[x_{elec},x_{therm}]^T \end{gathered} \tag{10}$$

where, x denotes the design variable vector, composed of electrical path variables $x_{elec} = [J, N, B_g]^T$ and thermal path variables $x_{therm} = [d, v, s, h]^T$. The functions $g_k(x)$ represent inequality constraints, encompassing temperature limits, pressure drop restrictions, and manufacturing constraints, while $h_l(x)$ denote equality constraints, primarily corresponding to electromagnetic performance matching requirements under rated operating conditions. X defines the feasible design domain. This formulation provides a comprehensive description of the multi-objective optimization problem under thermo–electromagnetic coupling, thereby establishing a rigorous mathematical foundation for the subsequent design and solution of the adaptive chaotic PSO algorithm.

3. THERMO–ELECTROMAGNETIC CO-SIMULATION FRAMEWORK BASED ON ADAPTIVE CHAOTIC PSO

3.1 Overall framework architecture

The proposed thermo–electromagnetic co-simulation framework based on adaptive chaotic PSO is designed to overcome the efficiency bottleneck inherent in traditional serial "simulation–optimization" workflows by establishing a parallel, interactive, and mutually driven system that deeply integrates the physical fidelity of simulation with the efficiency of intelligent optimization. Through the coordinated operation of five core modules, the framework achieves thermo–electromagnetic co-optimization. These modules follow a hierarchical and closed-loop logic of problem formulation → efficient prediction → intelligent optimization → accuracy calibration → decision output, progressing sequentially while forming a closed loop. First, coupled problem definition and solution space mapping provide the foundational boundaries of the framework by explicitly specifying the mathematical representations of design variables, objective functions, and constraints, as well as the admissible solution space. Second, the dual-path surrogate model, serving as the core intermediate layer, enables rapid and accurate prediction of electrical losses and thermal temperature fields, thereby providing efficient evaluation support for iterative optimization. Third, the adaptive chaotic particle swarm optimizer performs multi-objective search based on surrogate model outputs, while a thermal-state feedback mechanism ensures explicit physical guidance throughout the optimization process. Fourth, the optimization–simulation interactive verification closed loop dynamically calibrates surrogate model accuracy by feeding high-fidelity simulation results back into model updates, thereby enhancing predictive reliability. Finally, a multi-criteria decision-making (MCDM) system selects engineering-preferred solutions from the Pareto-optimal set. An overview of the thermo–electromagnetic co-simulation framework based on adaptive chaotic PSO is illustrated in Figure 1.

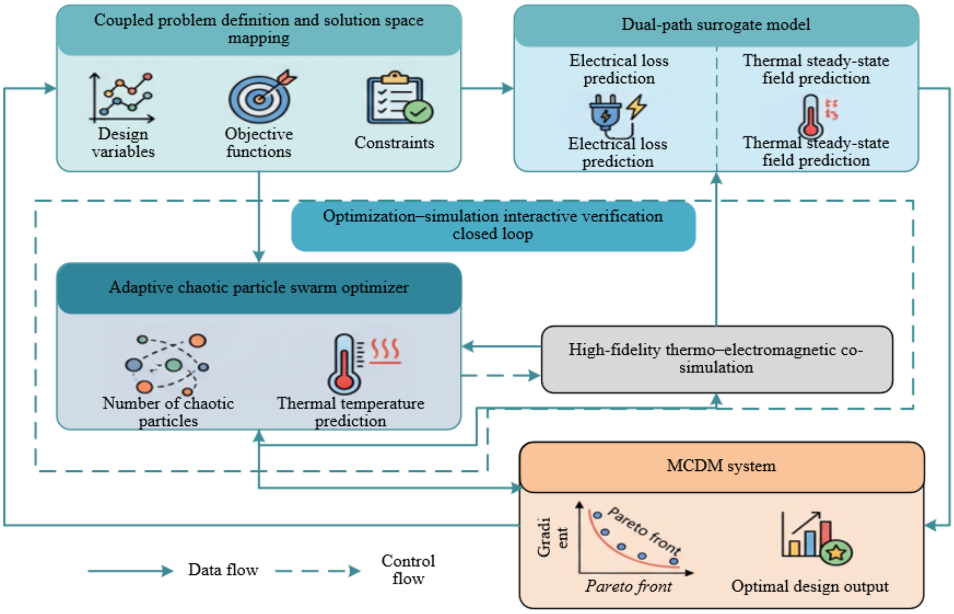


Figure 1. Thermo–electromagnetic co-simulation framework based on adaptive chaotic PSO

The data flow and control flow within the framework jointly constitute a mutually driven closed-loop chain. Design variable boundaries and objective function representations generated by the coupled problem definition module are used to guide sample acquisition and training of the dual-path surrogate model. The surrogate-predicted electrical losses and temperature fields are then supplied to the optimizer as fitness evaluation inputs, driving iterative position updates of the particle swarm. During optimization, selected candidate optimal solutions are fed into the high-fidelity thermo–electromagnetic co-simulation module for accuracy calibration. The resulting high-precision data are subsequently utilized to update surrogate model parameters and, in parallel, to enrich the training sample database, thereby enhancing model generalization capability. The updated surrogate model is then fed back to the optimizer, enabling the coordinated improvement of optimization accuracy and efficiency. Upon completion of the optimization iterations, the generated Pareto-optimal solution set is passed to the MCDM system, from which an engineering-feasible optimal design is selected. The decision outcome may further be used to retrospectively assess the rationality of the coupled problem definition, thereby establishing a full-process closed-loop regulation mechanism.

3.2 Dual-path surrogate model construction

The core design principle of the dual-path surrogate model is to accommodate the intrinsic dual-path characteristics of thermo–electromagnetic coupling. This is achieved by independently constructing an electrical surrogate model and a thermal surrogate model, while enabling bidirectional data interaction through coupling parameters. In this manner, the nonlinear mapping relationships between electrical losses and thermal temperature fields are accurately captured, while high predictive efficiency is maintained. A schematic illustration of the proposed architecture is provided in Figure 2. Sample acquisition constitutes the foundation of surrogate model construction. Latin hypercube sampling (LHS) is employed to achieve uniform coverage of the full design variable space, thereby ensuring sample representativeness and diversity and preventing degradation of model generalization caused by insufficient local-space information. All sample data are generated through high-fidelity thermo–electromagnetic co-simulation, with the full-dimensional design variables serving as inputs. The corresponding outputs include loss distributions of motor components, temperature fields of critical components, and pressure drop across the cooling system. Data preprocessing is conducted sequentially, including normalization, outlier elimination, and correlation analysis. Normalization is applied to eliminate dimensional discrepancies among variables and to enhance training efficiency. Outlier removal is performed using the three-sigma (3σ) criterion to ensure data quality. Correlation analysis is subsequently employed to identify input variables exhibiting strong relevance to the output responses, thereby effectively reducing model complexity and improving generalization capability.

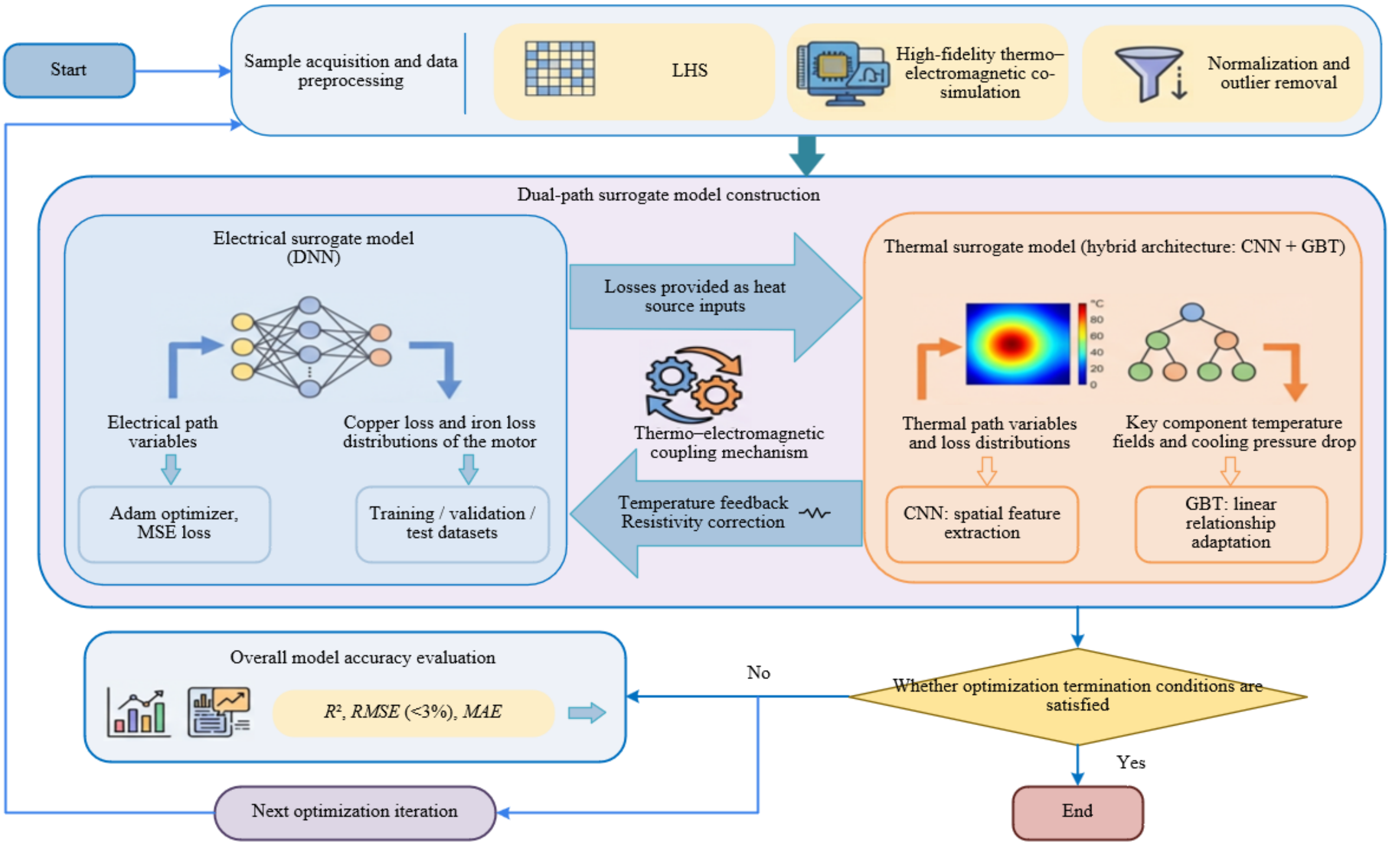


Figure 2. Schematic diagram of the dual-path surrogate model

The electrical surrogate model is implemented using a DNN architecture to accommodate the strong nonlinear relationships between electrical losses and electrical path variables. The model inputs consist of electrical path variables, while the outputs correspond to the distributions of copper losses and iron losses within the motor. The network architecture is designed such that the number of input-layer neurons matches the dimensionality of the electrical variables, three fully connected hidden layers are employed, and the number of output-layer neurons corresponds to the loss categories. The rectified linear unit (ReLU) activation function is adopted to enhance nonlinear fitting capability. Model training is performed using the Adam optimizer to minimize the mean squared error (MSE) loss function. The dataset is partitioned into training, validation, and test subsets with a ratio of 7:2:1, which are used for parameter learning, hyperparameter tuning, and generalization assessment, respectively. The thermal surrogate model adopts a hybrid architecture combining a CNN and a gradient boosting tree (GBT). The inputs include thermal path variables and the loss distributions predicted by the electrical surrogate model, while the outputs comprise the maximum temperature, average temperature, temperature difference of key components, and the cooling system pressure drop. Within this hybrid structure, the CNN is responsible for capturing the spatial distribution characteristics of the temperature field, whereas the GBT is employed to accommodate the predominantly linear relationships between pressure drop and the input variables. The training strategy remains consistent with that of the electrical surrogate model, with particular emphasis placed on validating the predictive accuracy of temperature field distributions to ensure reliable thermal state assessment.

The dual-path surrogate model achieves close alignment with the thermo–electromagnetic coupling mechanism through a bidirectional coupling mechanism. The loss distributions predicted by the electrical surrogate model are provided as thermal source inputs to the thermal surrogate model, serving as the primary driving factors for temperature field prediction. Conversely, the winding temperature feedback coefficients predicted by the thermal surrogate model are fed back into the electrical surrogate model to correct the temperature-dependent resistivity calculation, thereby accurately representing the coupled feedback loop of "temperature–resistivity–loss." Overall model accuracy is evaluated using R^2, RMSE, and MAE. It is required that the RMSE of key output parameters does not exceed 3%, ensuring that the surrogate model provides both efficient and accurate performance evaluation for subsequent optimization iterations and achieves a balanced trade-off between optimization efficiency and precision.

3.3 Adaptive chaotic PSO optimizer driven by thermal state feedback

Parameter self-adaptation is achieved through thermal-state feedback, enabling a balanced enhancement of global exploration and local exploitation capabilities. Particle encoding is implemented using a real-valued representation, whereby the position vector of each particle directly corresponds to a complete cooling system design scheme, expressed as $x_i = [x_{i,elec}, x_{i,therm}]^T$, where $x_{i,elec}$ denotes the subset of electrical path variables and $x_{i,therm}$ represents the subset of thermal path variables. The particle dimensionality is consistent with the total number of design variables. Population initialization is performed using a Tent chaotic mapping to generate a uniformly distributed initial swarm, thereby mitigating premature convergence caused by uneven initial distributions. The Tent map is defined as:

$$z_{k+1}=\begin{cases}2z_k & 0\leq z_k<0.5\\ 2(1-z_k) & 0.5\leq z_k\leq 1\end{cases} \tag{11}$$

where, z_k denotes the chaotic variable at the k-th iteration. The initialization procedure consists of first generating the chaotic sequence z_k, followed by a linear mapping into the feasible domain of the design variables $[x_{min}, x_{max}]$, given by $x = x_{min} + z_k(x_{max}-x_{min})$.

A chaotic perturbation mechanism is introduced to balance global exploration and local exploitation during the iterative process. The perturbation is triggered either at later stages of iteration or when population diversity falls below a predefined threshold. Population diversity is quantified using the standard deviation of fitness values, defined as:

$$D=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_i-\bar{f}\right)^2} \tag{12}$$

where, N is the population size, f_i denotes the fitness value of the i-th particle, and $\bar{f}$ represents the mean fitness of the population. When it drops below the diversity threshold D_{th}, the perturbation mechanism is activated. The perturbation strategy applies a small-amplitude chaotic disturbance to the particle's best positions, expressed as:

$$p_{i,new}=p_{i,old}+\delta(z_k-0.5)(x_{max}-x_{min}) \tag{13}$$

where, $p_{i,old}$ denotes the historical best position of the particle, and δ=0.08 is the perturbation coefficient, selected to balance disturbance intensity and convergence stability. The principal innovation lies in the thermal-state–feedback–driven adaptive parameter adjustment mechanism, in which the thermal state is quantified using a temperature field uniformity index, defined as:

$$\theta=\frac{T_{max}-T_{avg}}{T_{avg}} \tag{14}$$

where, T_{max} is the maximum motor temperature, and T_{avg} is the average temperature. A larger value of θ indicates a more non-uniform temperature field and a higher risk of thermal concentration.

Adaptive parameter regulation is driven by θ, through which the inertia weight w, cognitive learning factor c_1, and social learning factor c_2 are dynamically updated according to:

$$\begin{gathered}w=w_{\max}-\frac{w_{max}-w_{min}}{1+exp\left(-k(\theta-\theta_{th})\right)},\\ c_1=c_{1,base}+k_1\theta, c_2=c_{2,base}-k_2\theta\end{gathered} \tag{15}$$

where, w_{max} = 0.9 and w_{min} = 0.4 define the upper and lower bounds of the inertia weight, respectively; θ_{th} = 0.15 is the thermal-state threshold; k = 5 is the decay coefficient; $c_{1,base}$ = 1.5 and $c_{2,base}$ = 1.5 are the baseline learning factors; and k_1 = 0.8 and k_2 = 0.8 are adaptive coefficients. When $\theta > \theta_{th}$, both w and c_1 are increased while c_2 is reduced, thereby encouraging global exploration to escape regions associated with thermal concentration. Conversely, when $\theta \leq \theta_{th}$, w and c_1 are decreased and c_2 is increased, promoting refined local exploitation. The multi-objective optimization procedure is conducted using an ε-dominance strategy to identify Pareto-front solutions. An external archive is employed to store non-dominated solutions, while crowding-distance sorting is applied to preserve solution diversity. Through this mechanism, convergence toward a uniformly distributed Pareto front is ensured.

3.4 Optimization–high-fidelity simulation interactive verification closed loop

The core design logic of the optimization–high-fidelity simulation interactive verification closed loop is to dynamically calibrate surrogate model accuracy, thereby establishing a virtuous cycle in which optimization drives model learning, and improved models enhance optimization accuracy. This mechanism addresses the degradation of predictive accuracy that may occur during long-term iterative optimization and enables coordinated improvements in both optimization efficiency and precision. A schematic illustration is provided in Figure 3. Although the surrogate model substantially reduces computational cost, prediction bias may increase when optimization iterations concentrate on regions near optimal solutions, where sample information is often sparse. In contrast, high-fidelity simulation is capable of accurately resolving the underlying thermo–electromagnetic coupling physics. Dynamic interaction between these two components effectively compensates for the surrogate model's accuracy deficiencies. The triggering mechanism of the closed loop adopts a dual-criterion strategy, combining periodic triggering and accuracy-based triggering, to ensure both timeliness and necessity of calibration. Periodic triggering is configured to execute once every five optimization iterations, thereby balancing calibration frequency against computational overhead. Accuracy-based triggering is determined by the surrogate model prediction error evaluated on solutions stored in the external archive, quantified using the RMSE:

$$RMSE_{val}=\sqrt{\frac{1}{M}\sum_{j=1}^{M}(y_{j,sim}\text{-}y_{j,pre})^{2}} \tag{16}$$

where, M denotes the number of archive solutions selected for validation, $y_{j,sim}$ represents the reference values obtained from high-fidelity simulation, and $y_{j,pre}$ denotes the corresponding surrogate model predictions. When $RMSE_{val} > 5\%$, calibration is forcibly triggered to prevent accuracy deterioration from adversely influencing the optimization trajectory.

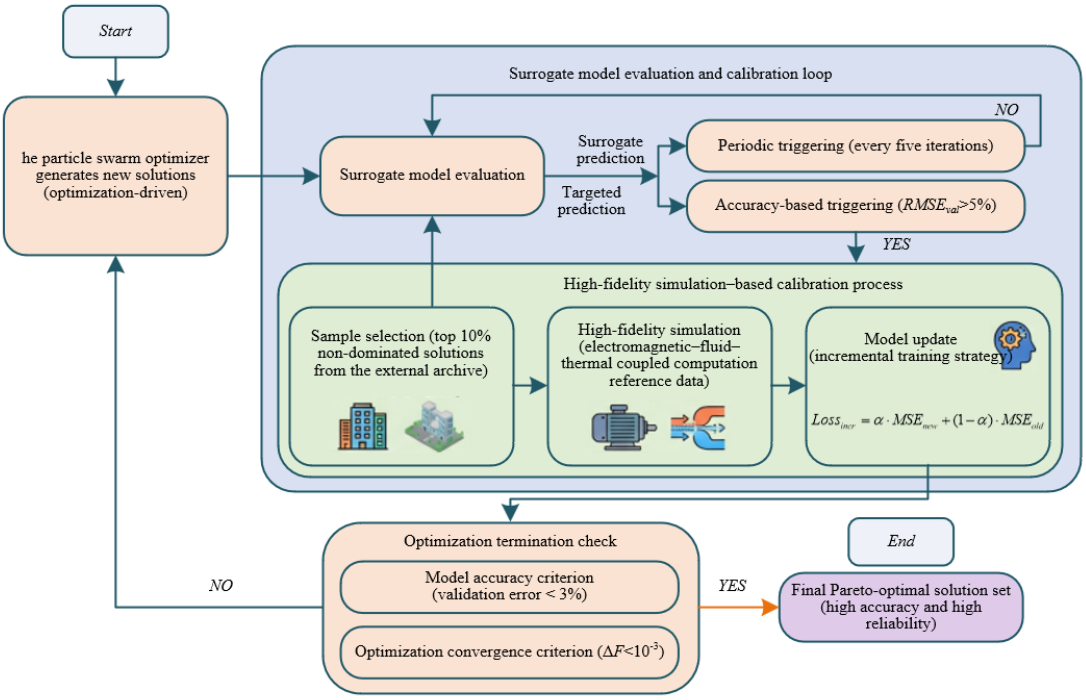


Figure 3. Schematic of the optimization–high-fidelity simulation interactive verification closed loop

The calibration procedure strictly follows a progressive sequence of sample selection → high-fidelity simulation → model update, thereby ensuring both calibration efficiency and effectiveness. During sample selection, the top 10% of non-dominated solutions from the external archive are chosen as calibration samples. These solutions represent the current optimal design region; calibrating against them enables targeted improvement of model accuracy in critical regions. High-fidelity simulation is performed using electromagnetic–fluid–thermal coupled simulations. The selected samples are supplied to the high-fidelity models to compute accurate benchmark data for model calibration, including motor loss distributions, temperature fields of key components, and cooling system pressure drop. Model updating is conducted using an incremental training strategy, whereby complete retraining of the surrogate model is avoided. Instead, newly acquired high-fidelity samples are appended to the training dataset, and model parameters are fine-tuned to achieve accuracy correction. The loss function for incremental training is defined as:

$$Loss_{incr}=\alpha\cdot MSE_{new}+(1\text{-}\alpha)\cdot MSE_{old} \tag{17}$$

where, MSE_{new} denotes the MSE associated with newly added samples, MSE_{old} represents the MSE over the original training dataset, and $\alpha = 0.6$ is a weighting coefficient selected to balance calibration toward new samples while preserving the existing generalization capability of the model.

Termination of the closed loop is required to satisfy the dual criteria of model accuracy and optimization convergence, thereby ensuring the completeness and reliability of the optimization process. The model accuracy criterion requires that the validation error of the surrogate model remain below 3% for two consecutive calibration cycles, indicating that predictive accuracy in the vicinity of the optimal solution region has stabilized and meets the required precision. The optimization convergence criterion is determined based on the variation of solutions stored in the external archive. The variation is defined as the mean Euclidean distance between non-dominated solutions in the archive across two consecutive generations and is expressed as:

$$\Delta F=\frac{1}{M}\sum_{j=1}^{M}\sqrt{\sum_{d=1}^{D}(f_{j,d}^{t+1}-f_{j,d}^{t})^{2}} \tag{18}$$

where, D denotes the dimensionality of the objective function space, and $f_{j,dt}$ represents the value of the d-th objective for the j-th archived solution at iteration t. When $f_{j,dt} < 10^{-3}$, the optimization process is considered to have converged. Once both criteria are satisfied, the interactive verification closed loop is terminated, and the optimizer proceeds to complete the remaining iterations based on the updated high-accuracy surrogate model. A stable Pareto-optimal solution set is then obtained as the final output. Through the dynamic complementarity between high-fidelity simulation and the surrogate model, the closed-loop framework preserves the efficiency of the surrogate model while ensuring the reliability of the optimization results, thereby providing essential support for achieving a balanced trade-off between accuracy and efficiency in thermo–electromagnetic co-optimization.

3.5 MCDM system

The primary objective of the MCDM system is to integrate practical engineering requirements and to select a unique, feasible optimal design from the Pareto-optimal solution set produced by the optimization process, thereby ensuring effective linkage between multi-objective optimization outcomes and engineering application. Because solutions on the Pareto front are mutually non-dominated, optimality cannot be determined based on any single objective alone; consequently, a decision mechanism incorporating engineering preferences is required. In this study, the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is adopted to construct the decision model. This method ranks candidate solutions by evaluating their relative closeness to the ideal solution, offering advantages of conceptual simplicity, computational efficiency, and suitability for multi-objective, multi-attribute decision scenarios. As a result, the influence of engineering preferences on decision outcomes can be objectively reflected.

Reasonable allocation of decision weights is critical to ensuring the scientific validity of the decision process. The Analytic Hierarchy Process (AHP) is employed to determine the weights of individual objective functions, enabling quantitative representation of engineering preferences. Within AHP, a hierarchical structure is constructed, and pairwise comparisons of objective importance are performed based on expert judgment, forming a judgment matrix $A = (a_{ij})_{n\times n}$, where n denotes the number of objective functions and a_{ij} represents the relative importance of the i-th objective with respect to the j-th objective. The maximum eigenvalue λ_{max} of the judgment matrix and its corresponding eigenvector are computed, and the normalized eigenvector yields the weight vector $w = [w_1, w_2, w_3]^T$. To ensure consistency of the weight assignment, a consistency check is performed using the consistency ratio $CR = CI/RI$, where $CI = (\lambda_{max}-n) / (n-1)$ is the consistency index and RI is the random consistency index. When CR < 0.1, the judgment matrix is considered to satisfy the consistency requirement. For different application scenarios, decision weights can be dynamically adjusted to reflect specific engineering priorities. For example, in new energy vehicle drive motor applications, higher weights are typically assigned to cooling power consumption and maximum motor temperature, whereas in industrial motor applications, greater emphasis is placed on cooling system volume and reliability.

The TOPSIS decision-making procedure strictly follows a progressive sequence of normalization → weighting → ideal-solution determination → closeness-based ranking. The first step involves standardization of objective functions. Since all objectives considered in this study are of the minimization type, range normalization is employed to eliminate dimensional inconsistencies, expressed as:

$$z_{ij}=\frac{x_{j,max}-x_{ij}}{x_{j,max}-x_{j,min}} \tag{19}$$

where, x_{ij} denotes the value of the j-th objective function for the i-th Pareto solution, $x_{j,max}$ and $x_{j,min}$ represent the maximum and minimum values of the j-th objective, respectively, and z_{ij} is the normalized objective value. In the second step, a weighted decision matrix is constructed by combining the normalized matrix with the weight vector, i.e., $Z = z_{ij}\cdot w_j$, thereby explicitly reflecting the relative importance of each objective. The third step involves computation of the positive and negative ideal solutions. The positive ideal solution is defined as $Z^+ = [\,max\,(z_{i1}w_1), max\,(z_{i2}w_2), max\,(z_{i3}w_3)]$, representing the optimal weighted values of each objective, whereas the negative ideal solution is defined as $Z^- = [\,min\,(z_1w_1), min\,(z_2w_2), min\,(z_3w_3)]$, corresponding to the least favorable weighted values. In the fourth step, the Euclidean distances between each solution and the ideal solutions, as well as the corresponding closeness coefficients, are computed. The distance metric is defined as:

$$d_i^+=\sqrt{\sum_{j=1}^{3}(Z_{ij}-Z_j^+)^2},d_i^-=\sqrt{\sum_{j=1}^{3}(Z_{ij}-Z_j^-)^2} \tag{20}$$

The closeness coefficient is defined as $C_i = d_i^+ / (d_i^+ + d_i^-)$, with C_i in the range of [0, 1]. A larger value of C_i indicates that the corresponding solution is closer to the positive ideal solution. Finally, all Pareto-optimal solutions are ranked according to C_i, and the solution with the highest closeness coefficient is selected as the recommended optimal design. The corresponding design parameters and predicted performance indicators are then reported, providing a direct basis for engineering decision-making and validation.

4. EXPERIMENTAL VALIDATION AND RESULTS ANALYSIS

4.1 Experimental setup

The experimental study was conducted on a 200 kW permanent magnet synchronous drive motor for new energy vehicle applications. The principal specifications are as follows: a rated power of 200 kW, a rated speed of 10,000 r/min, windings fabricated using Class F insulated copper conductors, a stator core composed of 50W470 silicon steel laminations, and a rated voltage of 350 V. The cooling system adopts a liquid-cooled jacket configuration, with initial design parameters specified as a cooling channel diameter of 10 mm, 12 cooling channels, a fin spacing of 8 mm, a fin height of 15 mm, and a coolant consisting of a 50% ethylene glycol–water solution operating at an initial flow velocity of 2 m/s. The simulation platform is composed of three main components: a high-fidelity co-simulation module, a surrogate model and optimization algorithm module, and a hardware support module. High-fidelity co-simulation was implemented through coupled simulations using ANSYS Maxwell and ANSYS Fluent, with data exchange across electromagnetic–fluid–thermal interfaces realized via ANSYS Workbench, enabling accurate computation of electrical losses and temperature fields. The surrogate model and optimization algorithm were implemented in Python, with neural networks constructed using the TensorFlow framework and GPR realized through the Scikit-learn library. The adaptive chaotic PSO algorithm was developed using customized code. Hardware support was provided by an Intel Xeon Gold 6330 CPU (28 cores, 56 threads) and an NVIDIA A100 GPU (80 GB memory), ensuring efficient execution of large-scale simulation and optimization tasks.

The experimental parameters were configured below. For optimization, the population size is set to 50, the maximum number of iterations to 100, the chaotic perturbation threshold to 0.05, the thermal-state evaluation threshold to 0.15, and the interactive verification interval to K = 5 iterations. For surrogate modeling, the electrical surrogate model employs a DNN with three hidden layers, a learning rate of 0.001, and 500 training epochs. In the thermal surrogate model, the CNN consists of two convolutional layers and two pooling layers, while the GBT is configured with 100 decision trees; training parameters are consistent with those of the DNN. The evaluation metric system encompasses three categories of core indicators. Optimization efficiency is assessed using the number of convergence iterations and the total computational time. Optimization accuracy is evaluated based on improvements in objective function values, the RMSE of surrogate model predictions, and high-fidelity validation error. Solution diversity is quantified using the crowding distance of the Pareto front, calculated as the mean Euclidean distance between adjacent solutions along the front, where larger values indicate a more uniformly distributed Pareto front.

4.2 Comparative experimental design

To systematically verify the superiority of the proposed approach, four categories of comparative experiments were designed. The comparison dimensions encompass surrogate model architecture, optimization algorithm enhancement strategies, and simulation–optimization coordination modes, while pure high-fidelity simulation–based optimization is adopted as the accuracy benchmark to ensure comprehensive and targeted validation. Comparative Method 1 employs conventional PSO combined with a single DNN surrogate model under a serial simulation–optimization paradigm. The principal distinctions lie in the replacement of the dual-path architecture with a single surrogate model, the absence of a thermal-state feedback mechanism, and the decoupled serial execution of optimization and simulation. This configuration is used to evaluate the contributions of the dual-path surrogate model and the interactive verification closed loop to optimization performance. Comparative Method 2 adopts chaotic PSO with a dual-path surrogate model under a serial paradigm. While the dual-path surrogate architecture is retained, the thermal-state–feedback–driven adaptive parameter regulation mechanism is omitted. This setup is designed to assess the impact of thermal-state feedback on optimization guidance and convergence behavior. Comparative Method 3 corresponds to the proposed method, integrating adaptive chaotic PSO, the dual-path surrogate model, and the interactive verification closed loop, thereby incorporating all core innovations. The baseline comparison group consists of pure high-fidelity simulation–based optimization, in which multi-objective optimization is performed directly using coupled ANSYS Maxwell–Fluent simulations without surrogate model acceleration. This benchmark is employed to quantify surrogate-induced accuracy loss and to evaluate the accuracy recovery capability of the proposed framework. All comparative experiments are conducted using the same research object, identical initial design parameters, and optimization objectives, ensuring uniform experimental conditions. The comparative analysis focuses on three key dimensions: optimization efficiency, optimization accuracy, and Pareto front distribution characteristics, thereby enabling a systematic evaluation of the comprehensive advantages of the proposed approach.

4.3 Experimental results and analysis

4.3.1 Surrogate model accuracy validation

Surrogate model accuracy constitutes the foundation for reliable optimization. In this subsection, the predictive performance of the dual-path surrogate model is evaluated through comparison with a single DNN surrogate model, while the dynamic error evolution under the interactive verification closed loop is further analyzed to validate the effectiveness of both the model architecture and the closed-loop mechanism. The experimental results are summarized in Table 1.

Table 1. Comparison of surrogate model prediction accuracy

Model Type	Metric	Copper Loss Prediction	Temperature Field Prediction	Pressure Drop Prediction	Average RMSE
Single DNN surrogate model	R^2	0.921	0.896	0.903	-
	RMSE (%)	4.82	6.35	5.17	5.45
Dual-path surrogate model (without the interactive loop)	R^2	0.968	0.957	0.962	-
	RMSE (%)	2.35	3.12	2.58	2.68
Dual-path surrogate model (with the interactive loop)	R^2	0.983	0.976	0.979	-
	RMSE (%)	1.52	2.03	1.71	1.75
Accuracy improvement via the interactive loop (%)	RMSE reduction	35.3	35.0	33.7	34.7

As indicated in Table 1, the dual-path surrogate model demonstrates substantially higher predictive accuracy than the single DNN surrogate model. The average RMSE is reduced by 50.8% relative to the single-model approach, with the temperature field prediction achieving an R^2 value of 0.957 and an RMSE of 3.12%. This improvement is attributed to the strong adaptability of the dual-path architecture to the intrinsic dual-path characteristics of thermo–electromagnetic coupling: the electrical surrogate model is dedicated to capturing the nonlinear relationships between losses and electrical variables, while the thermal surrogate model, based on a hybrid architecture, accurately resolves spatial temperature field distributions and the predominantly linear relationship associated with pressure drop. In contrast, a single surrogate model struggles to simultaneously accommodate these heterogeneous mapping characteristics. Further enhancement is achieved through the interactive verification closed loop, which reduces the average RMSE to 1.75%, corresponding to an additional reduction of 34.7%. It is observed that surrogate model errors consistently decrease upon closed-loop activation, particularly during later optimization stages when the search concentrates on the optimal solution region. Incremental training with newly generated high-fidelity simulation samples effectively compensates for prediction bias caused by sparse sample coverage in this region, thereby ensuring robust accuracy support for optimization iterations.

4.3.2 Optimization efficiency analysis

Optimization efficiency was quantified in terms of computational time and iteration count to convergence, with performance differences among the three comparative methods and the baseline group evaluated. The experimental results are summarized in Table 2. The convergence criterion is defined as a change in the Pareto front smaller than 10^{-3} for five consecutive generations.

Table 2. Comparison of optimization efficiency

Method	Computational Time (min)	Convergence Iterations	Time Reduction vs. Baseline (%)	Iteration Reduction vs. Method 1 (%)
Baseline: pure high-fidelity simulation optimization	1480	92	-	-
Method 1: conventional PSO + single surrogate model + serial mode	320	85	78.4	-
Method 2: chaotic PSO + dual-path surrogate model + serial mode	215	68	85.4	20.0
Proposed method: adaptive chaotic PSO + dual-path surrogate model + interactive closed loop	132	45	91.1	47.1

As shown in Table 2, the proposed method exhibits the highest optimization efficiency. The total computational time is reduced to 132 min, corresponding to a 91.1% reduction relative to the baseline group, and reductions of 58.8% and 38.6% relative to Methods 1 and 2, respectively. The number of iterations to convergence is reduced to 45, representing decreases of 47.1% and 33.8% compared with Methods 1 and 2, respectively. The efficiency gains are primarily attributed to the synergistic effects of chaotic perturbation and thermal-state feedback. Chaotic initialization enhances population diversity and mitigates ineffective searches during early iterations, while chaotic perturbations introduced in later stages balance global exploration and local exploitation, thereby reducing iteration waste associated with premature convergence. The thermal-state feedback mechanism, driven by a temperature-field uniformity index, dynamically adjusts optimization parameters: when thermal concentration is detected, the inertia weight and cognitive learning factor are increased, thereby guiding particles to rapidly escape regions of thermal concentration and preventing repeated iterations within unfavorable areas of the design space; when the temperature field becomes more uniform, parameter settings shift toward intensified local exploitation, accelerating convergence to the optimal region. In addition, the high-efficiency predictions of the dual-path surrogate model, together with the precise calibration provided by the interactive closed loop, further achieve a balance between efficiency and accuracy by preventing ineffective iterations caused by insufficient surrogate accuracy.

4.3.3 Optimization performance analysis

Optimization performance was evaluated from four perspectives: Pareto front distribution, objective function improvement, flow field and temperature field characteristics, and high-fidelity validation accuracy. The experimental results are summarized in Table 3. Pareto uniformity was quantified using the crowding distance, where a larger value indicates a more uniform distribution. Objective improvement ratios were calculated relative to the initial design.

As shown in Table 3, the crowding distance achieved by the proposed method reaches 0.61, which is close to that of the baseline group (0.63) and substantially higher than those obtained by Methods 1 and 2. This result indicates a more uniformly distributed Pareto front, thereby providing a richer set of candidate solutions for MCDM. In terms of objective improvement, the proposed method reduces T_{max} by 17.9%, P_{pump} by 33.9%, and V_{cool} by 23.5% relative to the initial design. These improvements are slightly inferior to those of the baseline but significantly outperform both comparative methods. Method 1 exhibits limited T_{max} reduction (9.5%) due to insufficient accuracy of the single surrogate model, which introduces bias in the optimization direction. Method 2 lacks thermal-state feedback and therefore fails to effectively avoid thermal concentration, resulting in a T_{max} improvement that is 4.2 percentage points lower than that of the proposed method. High-fidelity validation indicates that the optimal solution obtained by the proposed method exhibits a validation error of only 2.1%, satisfying the predefined 3% accuracy requirement. This outcome demonstrates that the interactive verification closed loop effectively compensates for surrogate-induced accuracy loss and ensures the reliability of the optimization results.

According to the analysis of the temperature and flow fields for the optimal design, pronounced thermal concentration is observed at the winding end regions in the initial design, with T_{max} reaching 168°C. After optimization using the proposed method, the cooling channel diameter is increased to 12 mm, fin spacing is reduced to 6 mm, and coolant velocity is optimized to 2.5 m/s, resulting in a more uniform flow distribution. Consequently, the convective heat transfer coefficient is increased by 28%, the winding end temperatures are significantly reduced, and T_{max} decreases to 138°C. The temperature-field uniformity index θ is reduced from 0.21 to 0.09, indicating effective mitigation of thermal concentration.

4.3.4 Ablation study

To verify the necessity of the three core modules—namely the dual-path surrogate model, thermal-state feedback, and the interactive verification closed loop—an ablation study was conducted by sequentially removing each module to construct three ablated variants. Performance differences between these variants and the complete proposed method were analyzed. The experimental results are summarized in Table 4.

Table 3. Comparison of optimization performance

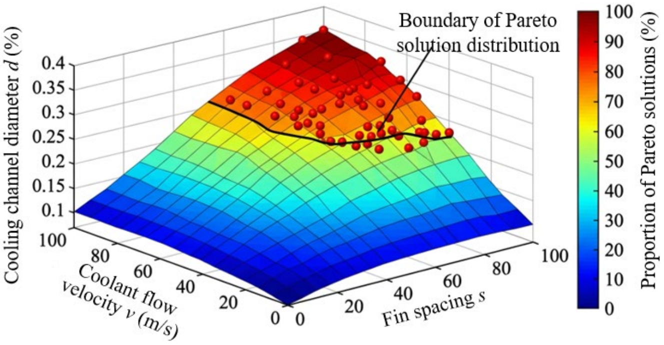
Method	Pareto Crowding Distance	Maximum Motor Temperature T_{max} (°C)	T_{max} Improvement (%)	Pumping Power P_{pump} (W)	P_{pump} Improvement (%)	Cooling System Volume V_{cool} (L)	V_{cool} Improvement (%)	High-Fidelity Validation Error (%)
Initial design	-	168	-	620	-	8.5	-	-
Method 1	0.32	152	9.5	510	17.7	7.8	8.2	6.8
Method 2	0.45	145	13.7	465	25.0	7.2	15.3	4.2
Proposed method	0.61	138	17.9	410	33.9	6.5	23.5	2.1
Baseline: pure high-fidelity simulation optimization	0.63	136	19.0	402	35.2	6.4	24.7	0.8

Table 4. Results of the ablation study

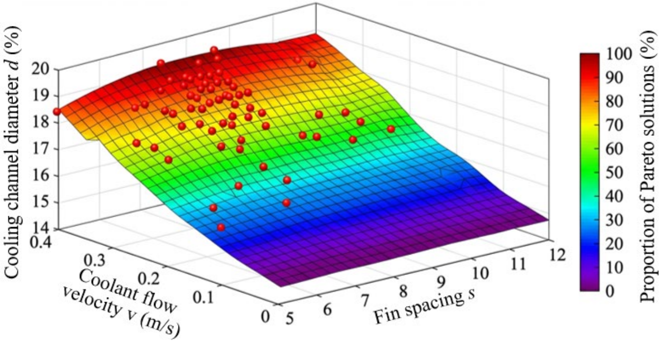
Experimental Group	Core module Configuration	Computational Time (min)	Convergence Iterations	T_{max} Improvement (%)	P_{pump} Improvement (%)	High-Fidelity Validation Error (%)	Overall Performance Degradation (%)[1]
Complete proposed method	Dual-path + thermal-state feedback + interactive closed loop	132	45	17.9	33.9	2.1	-
Ablation Group 1	Single surrogate model + thermal-state feedback + interactive closed loop	185	62	12.3	24.5	4.8	31.8
Ablation Group 2	Dual-path + without thermal-state feedback + interactive closed loop	163	75	13.1	26.8	2.3	25.1
Ablation Group 3	Dual-path + thermal-state feedback + without interactive closed loop	128	51	14.5	29.2	5.3	19.0

Note: [1] Overall performance degradation is calculated as the averaged percentage decrease across optimization effectiveness, efficiency, and accuracy, with equal weighting.

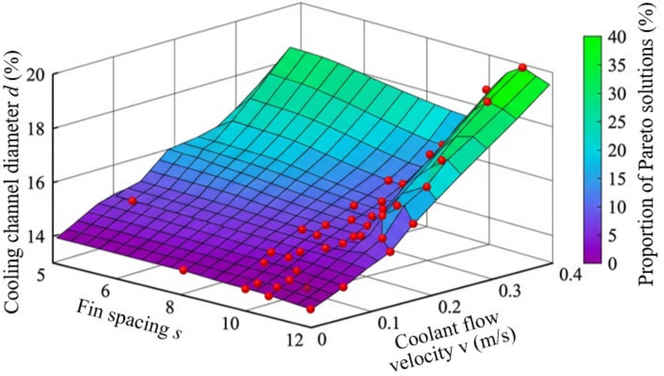
As indicated in Table 4, all three core modules play essential roles in the overall performance of the proposed framework. When the dual-path surrogate model is removed (Ablation Group 1), the overall performance degradation reaches 31.8%, computational time increases by 40.2%, the T_{max} improvement decreases by 31.8%, and the validation error rises to 4.8%, demonstrating that the dual-path architecture constitutes the foundation for both predictive accuracy and optimization efficiency. When the thermal-state feedback mechanism is removed (Ablation Group 2), the number of convergence iterations increases by 66.7%, and the T_{max} improvement decreases by 26.8%, indicating that thermal-state feedback effectively guides the optimization trajectory, reduces ineffective iterations, and enhances both convergence efficiency and optimization effectiveness. When the interactive verification closed loop is removed (Ablation Group 3), the high-fidelity validation error increases to 5.3%, accompanied by a pronounced reduction in objective improvement, confirming that the closed-loop mechanism ensures optimization reliability through dynamic calibration of surrogate model accuracy. Overall, the synergistic integration of the three modules enables the proposed method to achieve simultaneous improvements in efficiency, accuracy, and optimization effectiveness. The ablation results demonstrate that none of the modules can be omitted without incurring substantial performance degradation, thereby underscoring their collective indispensability.



(a) Distribution of the proportion of Pareto solutions satisfying the constraint $T_{\max} \leq 140°\mathrm{C}$



(b) Distribution of the proportion of Pareto solutions satisfying the constraint $P_{\mathrm{pump}} \leq 450$ W



(c) Distribution of the proportion of Pareto solutions satisfying the constraint $V_{cool} \leq 7$ L

Figure 4. Three-dimensional distribution of the proportion of Pareto solutions satisfying multiple performance constraints as a function of key design variables of the motor cooling system

To quantitatively characterize how cooling channel diameter, fin spacing, and coolant flow velocity influence the distribution of Pareto solutions that satisfy different thermo–electromagnetic performance constraints in multi-objective optimization—and thereby to provide intuitive decision support for coordinated cooling system parameter design—this analysis was conducted. In Figure 4(a), a high proportion of Pareto solutions approaching 100% is observed when the cooling channel diameter is relatively large and the fin spacing is relatively small, indicating that this parameter combination effectively satisfies the temperature control constraint. Figure 4(b) shows that regions with moderate coolant flow velocity and appropriate fin spacing yield a higher proportion of Pareto solutions, confirming that parameter settings within this range can achieve a favorable balance between pumping power consumption and heat transfer efficiency. Figure 4(c) reveals a pronounced performance trade-off, where the proportion of Pareto solutions exceeds 30% only within localized parameter combinations, indicating a strong mutual constraint between cooling system compactness and thermo–electromagnetic performance. These results not only clarify the influence of key design variables on different performance constraints but also visually demonstrate the distribution characteristics of Pareto solutions in multi-objective optimization. Collectively, the findings substantiate that the proposed adaptive chaotic PSO framework can effectively explore high-quality parameter regions under multiple constraints in thermo–electromagnetic co-optimization of motor cooling systems, thereby providing quantitative support for parameter trade-off decisions in engineering design.

Overall, the experimental results demonstrate that the dual-path surrogate model, by aligning with the intrinsic dual-path characteristics of thermo–electromagnetic coupling, significantly enhances prediction accuracy; the thermal-state–feedback–driven adaptive chaotic PSO, through dynamic parameter regulation, strengthens the physical guidance of the optimization process and markedly improves convergence efficiency; and the optimization–high-fidelity simulation interactive verification closed loop effectively calibrates surrogate model accuracy, ensuring the reliability of optimization outcomes. While achieving a 91.1% reduction in computational time, the proposed method realizes coordinated optimization of maximum motor temperature, pumping power, and cooling system volume, with performance approaching that of pure high-fidelity simulation–based optimization. These results establish an efficient and reliable new solution for thermo–electromagnetic co-optimization of electric motor cooling systems.

5. DISCUSSION

The proposed method demonstrates pronounced advantages in both efficiency and accuracy over comparative approaches, primarily attributable to the precise alignment of three core modules with the intrinsic characteristics of thermo–electromagnetic coupled optimization. The thermal-state feedback mechanism incorporates physical field information into the optimization process via the temperature field uniformity index θ. Therefore, parameter adaptation is no longer governed solely by algorithmic heuristics but is endowed with explicit thermal management guidance. This design effectively mitigates the stagnation of solutions in regions of thermal concentration that commonly arises from indiscriminate search strategies in conventional optimization. The experimentally observed reduction of θ from 0.21 to 0.09 exhibits a strong correlation with the improvement in T_{max}, thereby validating the rationality of employing this index as the basis for adaptive regulation. The dual-path surrogate model architecture accurately matches the dual-path nature of thermo–electromagnetic coupling. The electrical surrogate model leverages a DNN to accommodate the strong nonlinearity inherent in loss prediction, while the thermal surrogate model employs a CNN to capture the spatial characteristics of temperature fields. Relative to a single DNN surrogate, the average RMSE is reduced by 50.8%. Compared with GPR, which incurs high computational cost, and support vector machines, which often exhibit limited generalization capability, the DNN + CNN dual-path configuration achieves an effective balance between accuracy and efficiency, thereby providing reliable and rapid performance evaluation for iterative optimization. The optimization–high-fidelity simulation interactive verification closed loop further resolves the degradation of surrogate accuracy caused by insufficient sampling in the vicinity of optimal solutions through a dynamic calibration mechanism. As a result, the final validation error is controlled within 2.1%, enabling coordinated attainment of computational efficiency and predictive accuracy.

The proposed method exhibits substantial engineering applicability. With a 91.1% reduction in computational time relative to pure high-fidelity simulation-based optimization—representing an improvement exceeding one order of magnitude—the framework can be directly integrated into engineering design workflows for motor cooling systems, substantially shortening development cycles and reducing research and development costs. In application scenarios with stringent requirements on thermal management and power density, such as new energy vehicle drive motors and high-efficiency industrial motors, the method enables rapid generation of multi-objective balanced optimal designs, thereby providing direct support for engineering decision-making. Nevertheless, several limitations remain. First, initial sample acquisition relies on high-fidelity simulations; although LHS maximizes sample representativeness, constructing an initial dataset that adequately covers the full design space still incurs non-negligible computational expense. Second, the threshold value of the temperature-field uniformity index θ is derived from sample statistics specific to a given motor type; motors with different power ratings or structural configurations require re-calibration of this threshold, which weakens methodological generality. Third, the present study focuses on steady-state thermo–electromagnetic coupling, whereas real motor operation involves transient conditions such as start-up and acceleration. The dynamic coupling between losses and temperature under transient regimes has not yet been addressed, thereby limiting applicability under complex operating conditions.

Compared with existing studies, the proposed method exhibits three principal advantages. First, relative to conventional PSO variants, prior work has predominantly relied on mathematical adjustments of inertia weights and learning factors to improve convergence, with limited consideration of the underlying thermo–electromagnetic coupling physics. As a result, optimization trajectories may deviate from engineering requirements. By introducing thermal-state feedback, physical guidance is imparted into the optimization process, effectively addressing this limitation; experimentally, convergence efficiency is improved by 47.1% relative to conventional PSO. Second, compared with existing surrogate-assisted optimization approaches, most studies adopt a single surrogate architecture and lack a dynamic interaction mechanism between simulation and optimization, making it difficult to balance accuracy and efficiency. Through the coordinated design of a dual-path surrogate model and an interactive verification closed loop, the average prediction RMSE is reduced to 1.75%, while computational efficiency is markedly enhanced. Third, in contrast to the serial simulation–optimization paradigm, in which simulation and optimization are decoupled and simulation data cannot dynamically inform model refinement, the proposed parallel interactive closed-loop framework establishes a virtuous cycle in which optimization drives model learning and improved models, in turn, enhance optimization accuracy. This mechanism substantially reduces computational cost and yields a pronounced efficiency advantage.

Future research may extend the applicability and comprehensiveness of the method along four directions. (i) Extension to transient thermo–electromagnetic co-optimization, through the development of dynamic surrogate models capable of capturing time-varying loss and temperature characteristics, thereby accommodating transient operating conditions. (ii) Incorporation of active learning strategies, whereby samples with the highest information gain—identified via uncertainty quantification—are selected for high-fidelity simulation, optimizing the initial sample acquisition process and further reducing the computational cost of sample construction. (iii) Integration of thermo–electromagnetic–structural multiphysics coupling, by incorporating structural stress and vibration constraints into the optimization objectives, thereby enhancing design completeness and avoiding structural reliability issues arising from purely thermo–electromagnetic optimization. (iv) Establishment of a motor cooling system experimental test bench, enabling experimental validation of optimized designs, comparison between simulation predictions and measured data, and further model calibration to strengthen engineering credibility.

6. CONCLUSIONS

To address the fundamental challenges of imbalanced accuracy and efficiency and the lack of physical guidance in thermo–electromagnetic co-optimization of motor cooling systems, an adaptive chaotic PSO–based thermo–electromagnetic co-simulation optimization framework was proposed in this study. A parallel simulation–optimization interactive closed-loop paradigm was established. To accommodate the dual-path characteristics inherent to thermo–electromagnetic coupling, a dual-path surrogate model architecture was devised, enabling efficient and accurate prediction of electrical losses and thermal temperature fields, respectively. A thermal-state–feedback–driven adaptive PSO strategy was further proposed, in which a temperature-field uniformity index is embedded into dynamic parameter regulation, thereby imparting explicit physical guidance to the optimization process and enhancing both global exploration and local exploitation in complex design spaces. Comprehensive experiments conducted on a 200 kW permanent magnet synchronous drive motor systematically validated the proposed framework in terms of optimization efficiency, predictive accuracy, and engineering applicability.

The dual-path surrogate model demonstrates markedly higher predictive accuracy than a single-surrogate counterpart, with the average RMSE reduced by 50.8% relative to a single DNN model. The optimization–high-fidelity simulation interactive verification closed loop further calibrates surrogate accuracy, constraining the average RMSE within 1.75% and providing robust support for iterative optimization. The synergistic effects of thermal-state feedback and chaotic perturbation substantially enhance PSO performance, yielding a 47.1% improvement in convergence efficiency relative to conventional PSO while effectively avoiding unfavorable designs associated with local thermal concentration;

correspondingly, the temperature-field uniformity index is reduced from 0.21 to 0.09. Compared with traditional serial simulation–optimization workflows, computational efficiency is dramatically improved, with computational time reduced by 91.1%. Simultaneously, coordinated multi-objective optimization is achieved, including a 17.9% reduction in maximum motor temperature, a 33.9% reduction in cooling system pumping power, and a 23.5% reduction in cooling system volume, with overall performance approaching that of pure high-fidelity simulation–based optimization.

From an academic perspective, a new paradigm of deep synergy between physical simulation and intelligent optimization is established. By leveraging a dual-path surrogate model tailored to the characteristics of thermo–electromagnetic coupling, introducing thermal-state feedback to impart explicit physical guidance to the optimization process, and employing an interactive closed loop to ensure coordinated improvement of accuracy and efficiency, the inherent limitations of traditional serial workflows are effectively overcome. This framework provides both theoretical support and a reusable methodological system for the efficient optimization of thermo–electromagnetic coupled systems. From an engineering perspective, substantial practical value is demonstrated. The proposed approach can be directly integrated into the engineering design workflow of motor cooling systems, significantly shortening design cycles and reducing development costs. Its efficient and coordinated optimization capability facilitates the advancement of electric motors toward higher efficiency and increased compactness, thereby offering critical technical support for motor thermal management design in application domains such as new energy vehicles and industrial intelligent manufacturing.

REFERENCES

[1] Gör, H., Dalcalı, A. (2024). Design and optimization of a high-performance multi-barrier IPMS motor for an electric scooter and bicycle. Royal Society Open Science, 11(3): 231650. https://doi.org/10.1098/rsos.231650

[2] Gómez, J.R., Sousa, V., Eras, J.J.C., Gutiérrez, A.S., Viego, P.R., Quispe, E.C., De León, G. (2022). Assessment criteria of the feasibility of replacement standard efficiency electric motors with high-efficiency motors. Energy, 239: 121877. https://doi.org/10.1016/j.energy.2021.121877

[3] Waltman, D.J., Superczynski, M.J. (1995). High-temperature superconducting magnet motor demonstration. IEEE Transactions on Applied Superconductivity, 5(4): 3532-3535. https://doi.org/10.1109/77.482146

[4] Varshney, H., Jain, H., Tiwari, R. (2023). Thermal-electric modeling: a new approach for evaluating the impact of conservation voltage reduction on cooling equipment. Buildings, 13(5): 1287. https://doi.org/10.3390/buildings13051287

[5] Biedinger, J.M., Vilain, J.P. (1998). Design of low-voltage variable-frequency electrical motors. I: Analysis of the magnetic-thermal-electric coupled problem. European Physical Journal-Applied Physics, 3(1): 101-118.

[6] Wang, J., Zhang, S., Huang, H., Li, H., Yin, S., Yuan, W., Tang, Y. (2025). UTVC based thermal management of air-cored linear synchronous motor for high-precision CNC machine tool. Thermal Science and Engineering Progress, 63: 103722. https://doi.org/10.1016/j.tsep.2025.103722

[7] Zhang, K., Qi, Y., Le, B., Feng, Z. (2024). Design and experimental study of water cooling system for external rotor permanent magnet synchronous motor. IEEE Access, 12: 88034-88047. https://doi.org/10.1109/ACCESS.2024.3418017

[8] Lim, D.H., Lee, M.Y., Lee, H.S., Kim, S.C. (2014). Performance evaluation of an in-wheel motor cooling system in an electric vehicle/hybrid electric vehicle. Energies, 7(2): 961-971. https://doi.org/10.3390/en7020961

[9] Shen, Q., Zhu, H., Zhuo, J., An, P., Guo, D. (2025). Enhanced cooling strategy of YASA axial flux permanent magnet motor for electric automobile application. IEEE Transactions on Transportation Electrification, 11(5): 11065-11074. https://doi.org/10.1109/TTE.2025.3570739

[10] Konovalov, D., Tolstorebrov, I., Kobalava, H., Lamb, J.J., Eikevik, T.M. (2023). Experimental investigation of a low-temperature three-circuit cooling system for an electric motor under varying loads. Energies, 16(24): 8019. https://doi.org/10.3390/en16248019

[11] Lee, K.H., Cha, H.R., Kim, Y.B. (2016). Development of an interior permanent magnet motor through rotor cooling for electric vehicles. Applied Thermal Engineering, 95: 348-356. https://doi.org/10.1016/j.applthermaleng.2015.11.022

[12] Ma, C., Qu, L. (2015). Multiobjective optimization of switched reluctance motors based on design of experiments and particle swarm optimization. IEEE Transactions on Energy Conversion, 30(3): 1144-1153. https://doi.org/10.1109/TEC.2015.2411677

[13] Boukhalfa, G., Belkacem, S., Chikhi, A., Benaggoune, S. (2019). Genetic algorithm and particle swarm optimization tuned fuzzy PID controller on direct torque control of dual star induction motor. Journal of Central South University, 26(7): 1886-1896. https://doi.org/10.1007/s11771-019-4142-3

[14] Polat, M., Yildiz, A. (2023). GA and PSO based approaches for fast optimization of external rotor switched reluctance motor design parameters. Advances in Electrical and Computer Engineering, 23(2): 11-18. https://doi.org/10.4316/AECE.2023.02002

[15] Allaoua, B. (2012). Intelligent PID DC motor speed control alteration parameters using particle swarm optimization. In Artificial Intelligence Resources in Control and Automation Engineering, pp. 3-14. https://doi.org/10.2174/97816080512671120101

[16] Nisar, F., Rojek, J., Nosewicz, S., Kaszyca, K., Chmielewski, M. (2025). Coupled thermo-electric discrete element model for spark plasma sintering. Powder Technology, 458: 120957. https://doi.org/10.1016/j.powtec.2025.120957

[17] Chen, J., Hua, X., Zhang, X. (2017). Two-dimensional numerical simulation of thermo-electric coupling model in semiconductor bridge ignition system. International Journal of Heat and Mass Transfer, 113: 195-202. https://doi.org/10.1016/j.ijheatmasstransfer.2017.05.052

[18] Qiao, L., Liu, A., Ren, D., Meng, F., Wang, Y., Liu, Y. (2025). Cooling system design and temperature field analysis for high-speed switched reluctance motor

considering multi-field coupling of electromagnetic-thermal-flow. IEEE Access, 13: 77196-77209. https://doi.org/10.1109/ACCESS.2025.3562070

[19] Xu, Z., Xu, Y., Yang, S., Wang, Y., Wang, Y. (2026). A novel electromagnetic-thermal-fluid bidirectional coupled model considering multiple factors for temperature rise and loss calculation of high current density motor. International Journal of Thermal Sciences, 220: 110330. https://doi.org/10.1016/j.ijthermalsci.2025.110330

[20] Yan, Z., Deng, C., Li, B., Zhou, J. (2014). Novel particle swarm optimization and its application in calibrating the underwater transponder coordinates. Mathematical Problems in Engineering, 2014(1): 672412. https://doi.org/10.1155/2014/672412

[21] Niknam, T., Farsani, E.A., Nayeripour, M., Firouzi, B.B. (2011). Hybrid fuzzy adaptive particle swarm optimization and differential evolution algorithm for distribution feeder reconfiguration. Electric Power Components and Systems, 39(2): 158-175. https://doi.org/10.1080/15325008.2010.526990

[22] Zhang, Z. (2010). Quantum-behaved particle swarm optimization algorithm for economic load dispatch of power system. Expert Systems with Applications, 37(2): 1800-1803. https://doi.org/10.1016/j.eswa.2009.07.042