

Article

A Physics-Informed Neural Network for PMSM Temperature Estimation Under Sparse Sampling Conditions

Linxin Yu ¹, Jianye Liang ^{1,*}, Jing Ou ², Mengran Ji ³ and Hongwei Gao ¹

¹ School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang 110159, China; 13664168266@163.com (L.Y.); 18353187231@163.com (H.G.)

² School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China; jing.ou@hit.edu.cn

³ Shenyang Chemical Co., Ltd., Shenyang 110143, China; jimengran@sinochem.com

* Correspondence: ljy17332189105@163.com

Abstract

Permanent magnet synchronous motors (PMSMs) are widely used in new energy vehicles, electric drive systems, and industrial servo applications. Excessive permanent magnet temperature may lead to magnetic performance degradation or even irreversible demagnetization; therefore, accurate estimation of permanent magnet temperature is of considerable importance. However, existing data-driven methods generally rely heavily on high-frequency measurements, and their prediction accuracy tends to deteriorate under low-frequency sampling conditions. Moreover, purely data-driven models lack explicit physical constraints, which limits their interpretability and generalization capability. To address these issues, this study proposes a physics-informed long short-term memory model for permanent magnet temperature prediction. A physics-based loss function is formulated using the PMSM d–q-axis voltage balance equations, while the d- and q-axis inductances are treated as trainable parameters during network optimization. This design enables the temperature prediction task and the electromagnetic constraints to be optimized jointly. Multi-operating-condition experiments are conducted using a publicly available electric motor temperature dataset, and the proposed model is compared with CNN, GRU, MLP-PINN and TNN models. In addition, experiments involving different downsampling ratios, errors in the high-temperature region, parameter sensitivity, physical parameter identification, and input-feature effects are performed to comprehensively evaluate the proposed model. The results show that the PINN-LSTM model achieves the best overall prediction performance, with an MAE of 1.6048 °C, an RMSE of 2.1890 °C, and an R^2 of 0.9861, outperforming all comparison models. The model also maintains high prediction accuracy in the high-temperature region, with an MAE of 1.363 °C and an RMSE of 1.896 °C. Furthermore, the parameters learned by the model can effectively reconstruct the variation trends of the d- and q-axis voltages under the test operating conditions. Sensitivity analysis of the temperature coefficients further demonstrates that the model is robust to deviations in key physical parameters. These results indicate that the proposed method can achieve accurate and robust permanent magnet temperature prediction under low-frequency sampling conditions, providing an effective solution for motor thermal-state monitoring and health management.

Keywords: permanent magnet synchronous motor; physics-informed neural network; sparse sampling; long short-term memory; temperature prediction

Academic Editor: Brian D. Fath

Received: 26 June 2026

Revised: 27 July 2026

Accepted: 5 August 2026

Published: 6 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

1.1. Research Background

Permanent magnet synchronous motors have been widely used in electric vehicles and industrial drive systems because of their high efficiency and high power density [1,2]. However, temperature rise during motor operation can accelerate winding insulation degradation, cause permanent magnet demagnetization, and shorten the service life of the motor [3]. Accurate real-time temperature estimation and monitoring have therefore become important research topics. The internal thermal behavior of a PMSM is jointly affected by electromagnetic losses, heat dissipation conditions, and complex operating conditions, resulting in strong nonlinear and time-varying characteristics [4].

In practical online monitoring applications, the temperatures of critical internal motor components are often difficult to measure directly and continuously because of constraints on sensor installation, communication and storage resources, and controller computational capacity [5]. Temperature observer models are therefore required for online estimation. As the sampling interval increases or the amount of effective measurement data decreases, local dynamic information in the time series is progressively lost, making it more difficult for the model to adequately characterize the thermal delay, heat accumulation, and long-term dependencies inherent in motor thermal processes. Although purely data-driven models can learn the nonlinear mapping between motor operating variables and internal temperatures, their estimation performance depends strongly on the quantity, coverage, and operating-condition representativeness of the training data [6]. Physics-based models, such as lumped-parameter thermal networks (LPTNs), provide explicit physical interpretations of thermal paths and temperature evolution. However, their prediction accuracy depends on equivalent-node partitioning, thermal parameter identification, and model simplification assumptions, especially when detailed motor structures and thermal parameters are unavailable [7]. Therefore, integrating the temporal representation capability of neural networks with physical prior knowledge, such as heat-transfer principles, is expected to improve temperature estimation accuracy, data efficiency, and physical interpretability under limited-data or sparse-sampling conditions [8].

1.2. Related Work

Existing motor temperature monitoring techniques have broadly evolved from physics-based modeling to data-driven approaches. Thermal models based on finite element analysis (FEA) and computational fluid dynamics (CFD) construct coupled electromagnetic–thermal systems to obtain high-accuracy loss and temperature distributions under different operating conditions [9]. Multiphysics coupling models have also been applied to the analysis of transient thermal behavior in high-speed motors and motors with complex cooling structures [10]. Despite their high physical fidelity, these methods are computationally intensive and time-consuming, making them difficult to satisfy the requirements of online estimation and control.

To balance computational efficiency and prediction accuracy, lumped-parameter thermal networks simplify the complex heat-transfer process inside electric motors into interconnected thermal resistances and capacitances, enabling rapid temperature estimation through nodal energy-balance equations [11]. Global parameter identification can improve the consistency and stability of low-order LPTNs, while multi-node structures, loss compensation, and parameter optimization further enhance their capability under transient conditions [12]. However, the accuracy of LPTNs depends on reliable thermal parameters, node partitioning, and boundary conditions, which may limit their adaptability when detailed motor information is unavailable. Despite these limitations, LPTNs remain valuable due to their physical interpretability and low computational cost. Recently,

thermal neural networks (TNNs) have further combined LPTN-based physical structures with neural networks to improve thermal modeling capability [13]. Nevertheless, these methods provide valuable solutions for temperature estimation by incorporating thermal knowledge into data-driven models. Further exploration of different physical information sources may provide additional possibilities for improving PMSM temperature estimation under sparse sampling conditions.

Sensorless temperature-estimation methods based on temperature-sensitive electrical parameters have also attracted attention. By exploiting the temperature dependence of resistance, back electromotive force, or inductance, these methods estimate internal temperatures without additional sensors. High-frequency signal injection has been applied to PMSM condition monitoring in electric vehicles [14], whereas improved high-frequency voltage injection and PWM-based flux-linkage estimation have been applied to permanent magnet temperature estimation [15,16]. However, their accuracy may be affected by measurement noise, inverter nonlinearities, and parameter uncertainties.

Data-driven methods have recently been introduced into motor temperature prediction because of their nonlinear mapping and feature-learning capabilities. Deep neural networks, residual models, and thermal-network-assisted learning methods can extract thermal characteristics from operating data and achieve accurate predictions under variable conditions [17]. Transformer models have also been used for joint temperature prediction of multiple PMSM components [18]. For rotor permanent magnets, real-time models based on voltage, current, speed, and cooling conditions provide a feasible solution for online temperature monitoring [19]. Their practical applicability may be limited when the initial rotor temperature is unavailable, the sampling interval is increased, or the training data do not adequately cover the relevant operating conditions [20].

Physics-informed neural networks provide a promising approach to addressing the limited physical consistency of purely data-driven models. Early PINNs generally employed fully connected feedforward networks and constructed governing-equation residuals through automatic differentiation, enabling the simultaneous minimization of data-fitting errors and physical constraint violations during training [21,22]. However, conventional feedforward PINNs have limited capability in representing time-series systems characterized by pronounced historical dependence and dynamic delay. More recently, researchers have attempted to combine physical constraints with recurrent neural networks and attention mechanisms to improve the modeling of long-term dynamic processes and complex interactions among variables [23,24]. For PMSM temperature estimation, motor thermal evolution exhibits a pronounced temporal accumulation effect, while voltage, current, rotational speed, and temperature are explicitly coupled through the electromagnetic behavior of the motor. Therefore, integrating the temporal feature extraction capability of LSTM networks with the physical residual constraints of PINNs is expected to improve prediction accuracy, dynamic tracking capability, and physical interpretability under sparse sampling conditions.

Overall, physics-based models offer relatively high accuracy but have limited real-time capability. Electrical-parameter-based methods are straightforward to implement but are sensitive to measurement noise and parameter disturbances. Data-driven methods provide strong nonlinear approximation capability but generally require large amounts of training data and lack sufficient physical consistency. For PMSM permanent magnet temperature estimation under low-frequency sparse sampling conditions, few existing studies have jointly integrated the voltage balance mechanism, the temperature dependence of permanent magnet flux linkage, and LSTM-based temporal modeling within a unified framework. Therefore, developing a PINN-LSTM model that combines temporal feature extraction with physical constraints is important for improving permanent magnet temperature estimation accuracy, dynamic tracking capability, and physical

interpretability. The methodological differences between existing approaches and the proposed model are further summarized in Table 1.

Table 1. Methodological Comparison of PMSM Temperature Estimation Methods.

Method Category	Representative Methods	Characteristics/Limitations	Differences in This Study
Data-driven temporal models	CNN, GRU, LSTM, Transformer	Strong temporal learning ability but limited physical interpretability and no explicit motor-physics constraint.	LSTM temporal modeling is combined with an explicit d-q voltage-balance residual.
Thermal-model-based and thermal-neural methods	LPTN, TNN, LPTN-informed LSTM	Physical priors mainly come from heat-transfer networks or thermal-state propagation, requiring thermal-node design or equivalent thermal parameters.	No complete thermal network is required; electromagnetic voltage-balance consistency is introduced instead.
Electrical-parameter and voltage-model-based methods	Rs-based methods, PM flux observers, voltage-model estimators	Physically related to temperature-sensitive electrical parameters but usually used as standalone observers with limited deep temporal learning.	The voltage-balance relationship is embedded into an LSTM estimator with bounded equivalent parameters R_{20} , L_d , L_q , and ψ_{20} .

1.3. Contributions of This Study

To address the difficulty of accurately estimating the rotor permanent magnet temperature of a PMSM under low-frequency sparse sampling conditions, this study proposes a temperature estimation framework that integrates a long short-term memory network with physics-informed learning. The LSTM network serves as the temporal feature extraction backbone and captures long-term dependencies between permanent magnet temperature and motor operating variables, including voltage, current, rotational speed, torque, and stator temperature. Meanwhile, the quasi-steady-state voltage balance equations are incorporated into the network training process. The temperature dependence of permanent magnet flux linkage, the temperature correction of stator resistance, and the d-q-axis voltage constraints are used to establish a physical relationship between measurable electrical variables and permanent magnet temperature.

To improve the stability of physics-informed training, a staged training strategy consisting of pretraining, parameter calibration, and joint fine-tuning is further developed. This strategy enables the model to balance data-fitting accuracy and physical consistency. To evaluate the proposed method, low-frequency sparse sampling scenarios are constructed using a publicly available PMSM temperature dataset. The proposed model is compared with conventional machine learning models, deep learning models, and physics-based thermal models to assess its prediction accuracy, dynamic tracking capability, and physical interpretability under multiple operating conditions.

The novelty of this study does not lie in the LSTM backbone itself, but in the proposed voltage-balance-constrained physics-informed recurrent estimation framework. Existing data-driven temporal models can learn temperature dynamics from historical operating data, but they usually lack explicit physical constraints. LPTN-, TNN-, and LPTN-informed recurrent methods introduce physical information mainly from the heat-transfer or thermal-network perspective. In contrast, this study embeds the PMSM quasi-steady-state d-q-axis voltage-balance relationship into LSTM-based PM temperature estimation. The predicted PM temperature is coupled with the physical residual through the temperature-dependent PM flux linkage, while R_{20} , L_d , L_q , and ψ_{20} are optimized as bounded equivalent electromagnetic parameters. Therefore, the proposed method combines

recurrent temporal modeling, electromagnetic voltage-balance consistency, and bounded equivalent-parameter learning for low-frequency PM temperature estimation

Figure 1 illustrates the overall framework of the proposed permanent magnet temperature estimation method, which consists of three main stages: data acquisition, model training, and practical application. First, motor operating information, including voltage, current, rotational speed, and temperature, is collected to construct the training dataset. The collected data are then used to train the PINN-LSTM model. After training, the model estimates the permanent magnet temperature from measurable operating variables, providing support for motor thermal-state monitoring and thermal management applications.

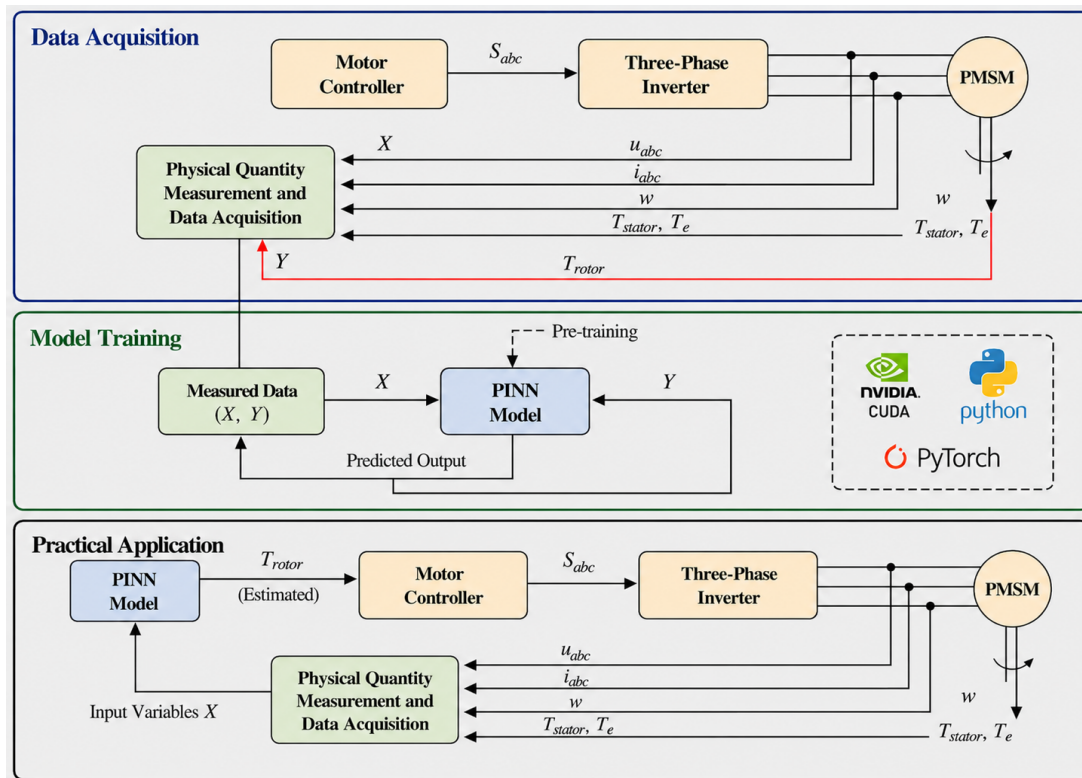


Figure 1. Overview of the process: data collection, model training, and practical application scenario.

2. Materials and Methods

This section describes the implementation of the proposed framework for PMSM permanent magnet temperature estimation. First, the source of the experimental data, the sparse-sampling procedure, and the feature normalization method are introduced. Subsequently, the architecture of the hybrid model, the formulation of the physical residuals, and the staged training strategy are presented. Finally, the experimental design, baseline models, and evaluation metrics are described to systematically assess the temperature estimation performance of the proposed method under low-frequency sparse sampling conditions.

2.1. Dataset Description

The proposed model was evaluated using a publicly available PMSM temperature dataset released by the LEA Laboratory at Paderborn University, Germany [25]. The dataset contains voltage, current, rotational speed, torque, and temperature measurements

collected during motor operation at a sampling frequency of 2 Hz and covers a wide range of load and speed conditions.

Because motor thermal processes exhibit pronounced thermal inertia, temperature variations generally occur much more slowly than changes in electrical variables. To obtain accurate and reliable temperature measurements, the relatively long duration of the operating profiles enables the dataset to capture both transient thermal evolution and long-term heat accumulation under diverse operating conditions. This procedure reduces the influence of short-term thermal transients and improves the reliability of the recorded temperature data. The main parameters of the PMSM used in the experiments are summarized in Table 2.

Table 2. Test motor parameters [26].

Type	Parameter	Unit
Power rating (nom./max.)	22/52	kW
Voltage rating	177	V
Current rating (nom./max.)	110/283	A
Torque rating (nom./max.)	110/250	N·m
Pole pair number	8	/
Sampling Duration	185	h

The dataset contains not only key electromagnetic variables, including d–q-axis currents, voltages, and rotational speed, but also synchronized temperature measurements from critical motor components, including the stator winding, stator tooth, stator yoke, and rotor permanent magnet. It should be noted that the temperature reported for each component represents the average of measurements obtained at multiple locations within the corresponding component. Owing to its broad coverage of operating conditions and clearly defined relationships among the recorded variables, this dataset has been widely used as a benchmark for electric motor thermal modeling and exhibits strong engineering relevance.

In this study, the rotor permanent magnet temperature is selected as the prediction target. Because the permanent magnets are embedded in the rotor and rotate at high speed during motor operation, their temperature is difficult to measure directly over extended periods in practical applications. Estimating the permanent magnet temperature from measurable operating variables is therefore of considerable engineering significance.

Among the selected variables, voltage, current, rotational speed, and torque characterize the electromagnetic loading condition of the motor. Coolant temperature reflects the external heat dissipation condition, whereas the temperatures of the stator winding, stator tooth, and stator yoke describe the internal thermal state of the motor. Together, these variables provide multisource operational information for permanent magnet temperature estimation.

2.2. Data Preprocessing and Construction of Low-Frequency Sampling Scenarios

The dataset comprises 69 operating profiles covering different torque and speed trajectories, with approximately 1.33 million data samples in total. Each sample contains information on the ambient condition, cooling condition, stator temperatures, and rotor temperature. The original data were recorded at a sampling frequency of 2 Hz. Considering the large thermal inertia of PMSMs, their thermal time constants are generally on the order of minutes and are substantially longer than the response times of electrical states. To reduce temporal redundancy, lower memory consumption, improve computational efficiency, and retain the long-term heat accumulation characteristics, the original dataset was downsampled by a factor of 10, resulting in a sampling interval of $\Delta t = 5$ s.

To prevent temporal data leakage between the training and test sets, the dataset was partitioned according to complete Profile ID sequences rather than by randomly shuffling individual time points. The training set was used to learn the model parameters. The validation set was used for hyperparameter selection, adjustment of the physical loss weight, and early-stopping decisions, whereas the test set was used exclusively for final performance evaluation.

Specifically, the training set contains 41 operating profiles, the validation set contains 14 profiles, and the test set contains 14 profiles, corresponding approximately to a ratio of 6:2:2. The dataset partition is illustrated in Figure 2. To estimate the rotor permanent magnet temperature during motor operation, input features capable of characterizing the main operating conditions of the motor were selected, including the stator currents i_d and i_q , the stator voltages u_d and u_q , the motor rotational speed, torque, and the temperatures of different stator components such as stator tooth temperature T_{st} , stator winding temperature T_{sw} , stator yoke temperature T_{sy} and coolant temperature T_{cool} . The prediction target is the permanent magnet temperature T_{pm} , which is directly related to the operational safety and reliability of the motor.

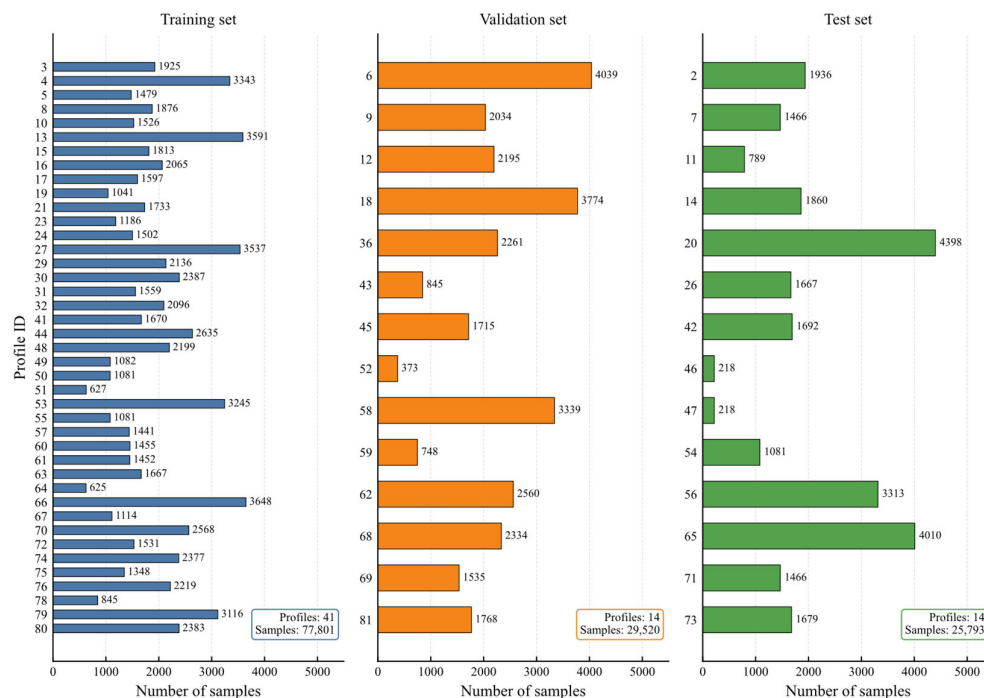


Figure 2. Partitioning of the PMSM temperature dataset into training, validation, and test sets based on complete operating profiles.

Because the motor operating variables have different physical units and markedly different numerical ranges, a Z-score normalization method based on the training-set statistics was applied to both the input features and the target variable.

$$x^* = \frac{x - \mu_{train}}{\sigma_{train}} \quad (1)$$

where x and x^* denote the variable before and after standardization, respectively, while μ_{train} and σ_{train} represent the mean and standard deviation of the corresponding variable calculated from the training set. Through this transformation, variables with different physical units and numerical ranges are mapped to comparable scales, thereby

reducing the dominance of large-magnitude variables during gradient updates and improving the stability and convergence efficiency of model training.

2.3. LSTM-Based Temperature Estimation Network

Permanent magnet temperature exhibits pronounced long-term temporal dependence. The temperature at the current time is influenced not only by the instantaneous current, rotational speed, torque, and cooling conditions, but also by the historical load and heat dissipation states over an extended period. Therefore, an LSTM network is employed as the backbone of the proposed temperature estimation model to extract long-term thermal dynamic features from multivariate time-series data [27].

As shown in Figure 3, the LSTM controls the retention, updating, and output of historical information through the forget, input, and output gates, thereby alleviating the vanishing-gradient problem encountered by conventional recurrent neural networks in long-sequence modeling. Here, f_t denotes the forget gate, h_t represents the hidden state, σ and $\tanh$ denote the sigmoid and hyperbolic tangent activation functions, respectively, and W and b represent the learnable weight matrices and bias vectors.

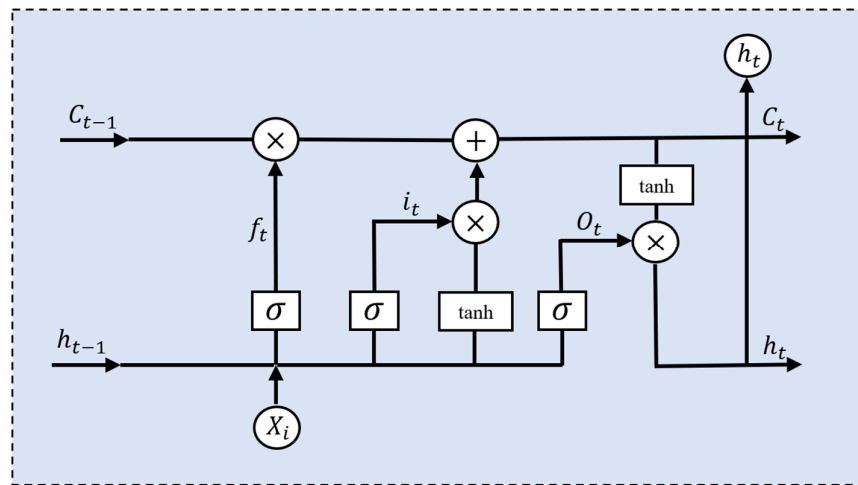


Figure 3. Basic architecture of the LSTM network.

$$f_t = \sigma(W_f h_{t-1} + W_f x_t + b_f) \quad (2)$$

$$i_t = \sigma(W_i h_{t-1} + W_i x_t + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_c h_{t-1} + W_c x_t + b_c) \quad (4)$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (5)$$

$$o_t = \sigma(W_o h_{t-1} + W_o x_t + b_o) \quad (6)$$

$$h_t = o_t \tanh(C_t) \quad (7)$$

i_t denotes the input gate, the input gate determines which new information is written into the memory cell, while the candidate state $\tilde{C}_t$ provides the new information used to update the cell state. C_t denotes the updated cell state, which combines the effects of the forget and input gates. o_t represents the output gate, and h_t is the hidden-state

output at time step t , representing the temporal features extracted by the LSTM from the current input and historical memory.

In the LSTM model, the flow of information is jointly controlled by C_t , i_t , f_t , and o_t . During network operation, the input gate writes the current motor operating-state information into the memory cell. The forget gate dynamically determines how much historical information should be retained according to its relevance to the current temperature. The memory cell continuously accumulates thermal information contained in historical operating states, thereby characterizing the processes of heat generation, transfer, and dissipation within the motor. Finally, the output gate extracts the temporal features most closely related to the permanent magnet temperature and produces the prediction output at the current time step. Through these mechanisms, the LSTM can effectively capture the long-term dependencies between permanent magnet temperature and historical load, rotational speed, and heat dissipation conditions, thereby providing an initial temperature estimate for the subsequent physics-informed constraint module.

The hidden state at the final time step of the LSTM is passed through a fully connected layer to obtain the predicted permanent magnet temperature.

Although a standalone LSTM can learn the nonlinear mapping between the input variables and permanent magnet temperature, its predictions are entirely data-driven and lack explicit physical constraints. Under low-frequency sampling, changing operating conditions, or limited-data scenarios, the model may therefore produce prediction errors that are inconsistent with the electromagnetic behavior of the motor. To address this limitation, physics-informed learning is further incorporated into the LSTM-based temperature estimation model, resulting in the proposed PINN-LSTM hybrid framework.

2.4. Formulation of the Physics-Informed Voltage Constraint and Loss Function

To improve the physical consistency of the LSTM-based temperature prediction model under low-frequency sampling conditions, the quasi-steady-state d-q-axis voltage balance equations of the PMSM are incorporated into the network training process. Unlike conventional thermal-network models, the proposed method does not require detailed motor geometry, thermal resistance-capacitance parameters, or complex heat-transfer boundary conditions. Instead, physical constraints are constructed from the electromagnetic perspective using measurable variables available in the dataset, including the d-q-axis voltages, currents, and rotational speed.

The central idea of the proposed physical constraint is that variations in permanent magnet temperature affect the permanent magnet flux linkage, which directly contributes to the q-axis voltage balance equation. Therefore, the permanent magnet temperature predicted by the neural network is not treated as an isolated output. Instead, it is further introduced into the voltage reconstruction process through a temperature-dependent flux linkage model. When the predicted temperature deviates from a physically reasonable thermal state, the corresponding equivalent flux linkage also deviates, resulting in an increased voltage reconstruction residual. Based on this relationship, the voltage residual can serve as an indirect physical constraint that guides the network toward producing permanent magnet temperature estimates consistent with the electromagnetic behavior of the motor.

Because this study focuses on low-frequency sampled data, the current derivative terms are susceptible to the effects of the enlarged sampling interval and measurement noise after downsampling. Therefore, quasi-steady-state voltage balance equations are adopted to construct the physical residuals. Specifically, the magneto-thermal characteristics of the permanent magnet material are first used to establish a mapping between the

predicted permanent magnet temperature $\hat{T}_{pm}(t)$ and the permanent magnet flux linkage ψ_{PM} :

$$\psi_{PM}(t) = \psi_{20} \cdot [1 + \alpha_{PM}(\hat{T}_{pm}(t) - T_{ref})] \quad (8)$$

where ψ_{20} denotes the reference permanent magnet flux linkage at 20 °C to be identified, α_{PM} is the reversible temperature coefficient of the permanent magnet material, set to -0.0012 K^{-1} , and T_{ref} is the reference temperature 20 °C. The temperature-dependent flux linkage ψ_{PM} serves as the key link between the neural network output and the physical equations.

Subsequently, the measured stator winding temperature T_{sw} is used to correct the stator resistance in real time, thereby reducing the influence of temperature-induced stator resistance variation on the voltage reconstruction results. R_{20} denotes the equivalent stator resistance at the reference temperature of 20 °C to be identified during model training, α_{Cu} is the temperature coefficient of copper.

$$R_s(t) = R_{20} \cdot [1 + \alpha_{Cu}(T_{sw}(t) - T_{ref})] \quad (9)$$

Based on the corrected parameters, a theoretical d-q-axis voltage model is constructed. For the motor represented in the dataset used in this study, the quasi-steady-state voltage balance equations are formulated as follows:

$$\begin{cases} \hat{u}_d(t) = R_s(t)i_d(t) - \omega_e(t)L_q i_q(t) \\ \hat{u}_q(t) = R_s(t)i_q(t) + \omega_e(t)L_d i_d(t) + \omega_e(t)\psi_{PM}(t) \end{cases} \quad (10)$$

where $\omega_e = \omega_m \cdot P$, ω_e denotes the electrical angular velocity, ω_m is the mechanical angular velocity, P is the number of pole pairs. These equations constitute the physical constraints of the proposed model. The unknown physical parameters are defined as trainable variables within the network. During backpropagation, the gradients of the physical loss with respect to L_d , L_q , and the other physical parameters are calculated, enabling parameter identification while minimizing the voltage residuals. This formulation accounts for the temperature dependence of the stator resistance and establishes a coupling relationship between the permanent magnet temperature and the voltage balance equations.

On this basis, the total loss function consists of a supervised data loss and a physical residual loss, as follows:

$$L_{total} = L_{data} + \lambda L_{phy} \quad (11)$$

$$L_{phy} = \frac{1}{N_t} \sum_t (\|u_d - \hat{u}_d\|^2 + \|u_q - \hat{u}_q\|^2) \quad (12)$$

The supervised data loss is used to quantify the error between the predicted and measured permanent magnet temperatures, whereas the physical residual loss constrains the electromagnetic state corresponding to the predicted temperature to satisfy the quasi-steady-state voltage balance equations. The physical loss weight is introduced to balance temperature prediction accuracy and electromagnetic physical consistency.

In addition, the publicly available dataset does not provide complete and accurate manufacturer-specified motor parameters. Directly adopting fixed nominal values may therefore cause the physical residuals to be inconsistent with the actual motor operating states. Accordingly, the stator resistance, d-q-axis inductances, and permanent magnet flux linkage at 20 °C are treated as trainable equivalent electromagnetic parameters and optimized jointly with the neural network parameters. Because the actual manufacturer-specified values are unavailable, the identified parameters are interpreted only as

equivalent parameters for improving voltage balance consistency, rather than as exact physical parameters of the motor. This treatment enhances the adaptability of the model under different operating conditions and low-frequency sampling scenarios.

2.5. Justification of the Quasi-Steady-State Voltage Equations

The quasi-steady-state d–q-axis voltage equations are adopted to construct the physical constraints, while the current derivative terms, $L_d \frac{di_d}{dt}$ and $L_q \frac{di_q}{dt}$, in the complete voltage equations are neglected. This simplification is based on the fact that the current derivative terms mainly reflect fast electromagnetic transients, whereas this study focuses on permanent magnet temperature variations at a sampling interval of 5 s. Because these two processes occur on markedly different time scales, the electrical state of the motor can be reasonably approximated as quasi-steady-state for temperature estimation at the time scale.

To verify the validity of this simplification and ensure consistency with the sparse-sampling setting, the contribution of the neglected inductive voltage terms was recalculated using the downsampled 5 s data.

As shown in Figure 4a, the previous analysis based on the original 2 Hz data shows that the 95th-percentile contribution ratios of the d- and q-axis inductive voltage terms are 0.003113% and 0.005674%, respectively. Figure 4b presents the results recalculated directly from the downsampled 5 s data used in the sparse-sampling experiments. In this case, the 95th-percentile magnitudes of the d- and q-axis inductive voltage terms are 0.001079 V and 0.001989 V, respectively, and their corresponding contribution ratios are only 0.001944% and 0.003667%. Even in the intervals with the largest contribution ratios, the d- and q-axis inductive voltage terms account for only 0.055100% and 0.194500%, respectively, both of which remain far below the 1% reference threshold. These results indicate that the current derivative terms have only a negligible influence on the voltage balance relationship under the actual 5 s sampling condition used in this study. Therefore, neglecting these terms and using the quasi-steady-state d–q-axis voltage equations to construct soft physical constraints is considered reasonable.

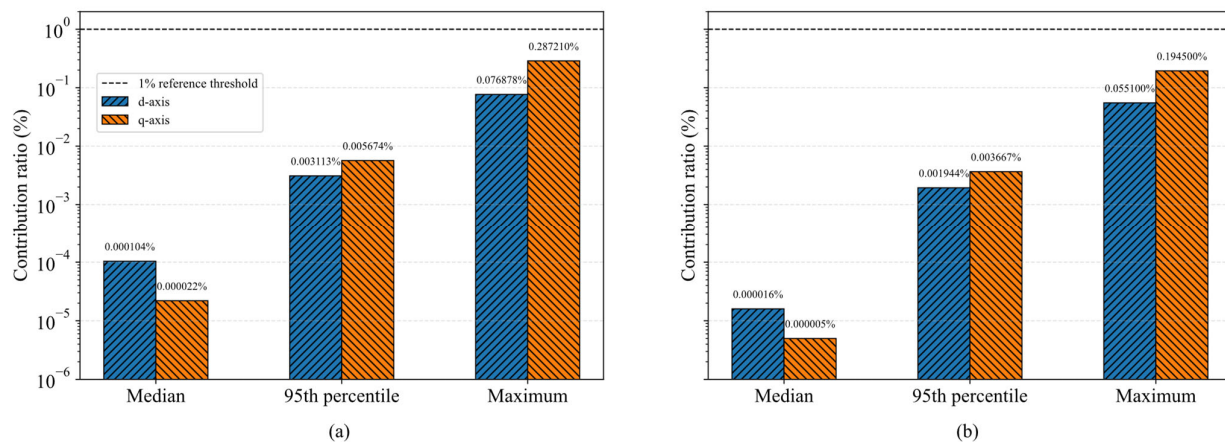


Figure 4. Contribution ratios of the neglected inductive voltage terms: (a) results from the original 2 Hz data using 5 s RMS windows; (b) results recalculated from the downsampled 5 s data. The dashed line indicates the 1% reference threshold.

2.6. Three-Stage Training Strategy

To address the gradient conflicts and sensitivity to parameter initialization encountered during PINN training, a three-stage training strategy is adopted. First, the LSTM network is pretrained in a purely data-driven manner to establish baseline prediction accuracy and provide reasonable initial network parameters. Subsequently, the network weights are frozen, and only the key motor parameters are updated using the physical residual loss to reduce systematic deviations. Finally, joint fine-tuning is performed using a decaying physical loss weight, enabling the model to accurately capture high-order non-linear relationships in the data while remaining dynamically constrained by the underlying physical laws.

Through this design, the physical relationship among permanent magnet temperature, permanent magnet flux linkage, and voltage residuals links the indirectly measurable rotor permanent magnet temperature to measurable variables, including voltage, current, and rotational speed. Consequently, physics-constrained permanent magnet temperature estimation can be achieved under low-frequency sampling conditions. The overall architecture of the proposed framework is illustrated in Figure 5.

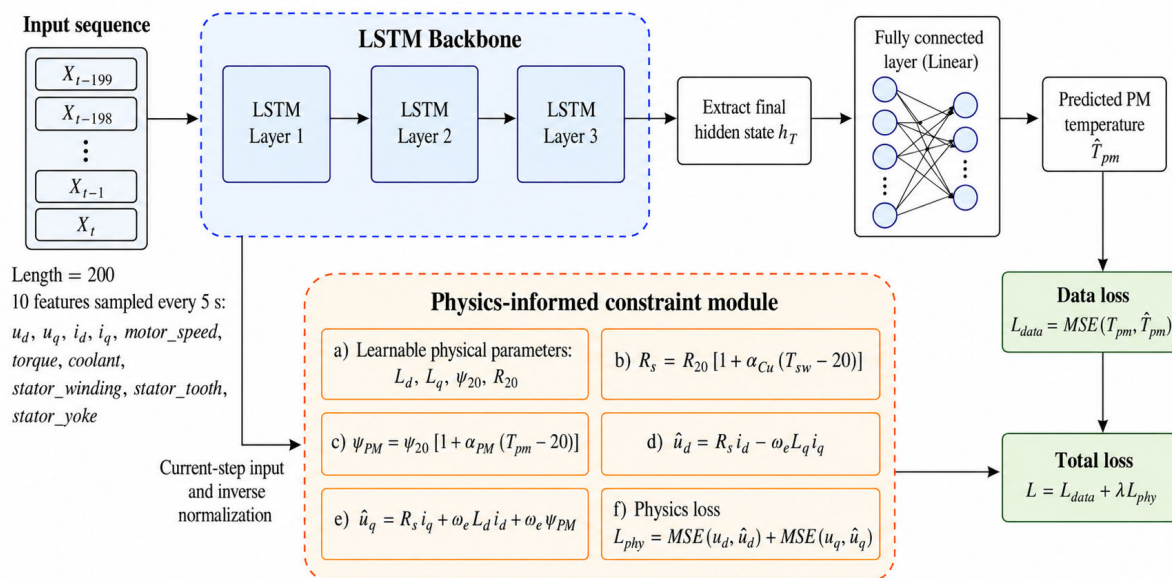


Figure 5. Overall architecture of the proposed PINN-LSTM temperature estimation model.

2.7. Experimental Design

To evaluate the effectiveness of the proposed PINN-LSTM model, a series of experiments was conducted from the perspectives of overall prediction performance, comparison with a physics-based model, ablation analysis, and adaptability to low-frequency sampling conditions.

First, using data downsampled by a factor of 10, the proposed PINN-LSTM model was compared with the CNN, GRU, and MLP-PINN models. The CNN was used to represent a local convolutional feature extraction approach, the GRU was selected as a representative recurrent time-series model, and the MLP-PINN was employed to examine the effect of physical constraints in a non-sequential network structure.

Second, an LPTN thermal model was independently constructed as a conventional physics-based baseline. This comparison was used to investigate the differences in applicability between the thermal-network approach and the proposed electromagnetically constrained neural network under the publicly available dataset.

Subsequently, an ablation study was performed by comparing the baseline LSTM, the PINN-LSTM with fixed physical parameters, the PINN-LSTM with trainable physical parameters, and the complete PINN-LSTM model. These experiments were designed to quantify the contributions of the voltage-balance physical residual, equivalent electromagnetic parameter learning, and the complete training strategy.

In addition, four downsampling factors of 5, 10, 15, and 20 were considered to evaluate the applicability of the proposed model under different low-frequency sampling conditions. Error analyses were also conducted in the high-temperature subset and in regions with frequent operating-condition changes to assess prediction reliability under thermally critical conditions and rapidly changing operating states.

2.8. Evaluation Metrics

To quantitatively evaluate the prediction accuracy, three standard statistical metrics were adopted: the root mean square error RMSE, mean absolute error MAE, and coefficient of determination R^2 . Their mathematical definitions are given as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{z}_i - z_i| \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{z}_i - z_i)^2} \quad (14)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (\hat{z}_i - \bar{z})^2}{\sum_{i=1}^N (z_i - \bar{z})^2} \quad (15)$$

where N denotes the total number of samples, and z_i , $\hat{z}_i$, and $\bar{z}$ represent the measured temperature, predicted temperature, and mean measured temperature, respectively.

MAE and RMSE are used to quantify the magnitude of the prediction errors, whereas R^2 evaluates the model's ability to reproduce the variation trend of the measured temperature.

3. Results and Discussion

This section presents the experimental evaluation and analysis of the proposed PINN-LSTM model for permanent magnet temperature estimation. First, under the same dataset partition and using data downsampled by a factor of 10, the overall performance of the proposed method is compared with that of the CNN, GRU, and MLP-PINN models. Subsequently, the conventional LPTN thermal network is separately discussed as a physics-based baseline. An ablation study, to evaluate the contributions of the voltage-balance physical residual, equivalent electromagnetic parameter learning, and the complete training strategy, is then discussed. Finally, experiments involving different downsampling factors, error analysis in the high-temperature region, and the effects of different input features, to assess the reliability of the proposed model under low-frequency sampling conditions and practical operating scenarios, are described.

The experiments were conducted on a 64-bit Windows operating system equipped with an AMD Ryzen 7 7700 CPU and 32 GB of memory. The software environment consisted of Python 3.12.11 and PyTorch 2.7.0, and model training was accelerated using an NVIDIA GeForce RTX 5060 GPU.

3.1. Comparison of Prediction Performance Across Models

To evaluate the overall effectiveness of the proposed PINN-LSTM model, a comparative experiment was first conducted using data downsampled by a factor of 10. The comparison models included CNN, GRU, MLP-PINN, and the proposed PINN-LSTM.

The model hyperparameters were determined through a stepwise tuning procedure using the validation set. The performance of the different networks was compared under the same dataset partition, input features, sequence length, and evaluation metrics, with the validation-set RMSE used as the primary criterion for hyperparameter selection. The CNN, GRU, and MLP-PINN models were also appropriately tuned on the same validation set according to their respective architectural characteristics to ensure a fair comparison.

For the proposed model, the number of LSTM hidden units and recurrent layers was set to 256 and 3, respectively. The batch size was 512, the initial learning rate was 1×10^{-4} , and the dropout rate was 0.2. During the joint fine-tuning stage, the learning rate was reduced to 1×10^{-5} . A staged annealing strategy was adopted for the physical loss weight to coordinate the optimization of the supervised data loss and the physical residual loss. During the first 20 epochs of joint fine-tuning, the physical loss weight was set to 1×10^{-4} to strengthen the constraint imposed by the voltage balance equations on the model output. From epochs 21 to 50, the weight was reduced to 1×10^{-5} , allowing the model to further minimize the temperature prediction error while maintaining physical consistency. Subsequently, the physical loss weight was reduced to 0, and the model was further refined using only the supervised data loss. This strategy prevents excessively large physical residuals from interfering with data feature learning during the early stage of training and reduces the adverse effect of overly strong physical constraints on prediction accuracy during the later stage. The hyperparameters of the compared models are summarized in Table 3.

Table 3. Hyperparameter settings of the compared models.

Hyperparameter	CNN	GRU	MLP-PINN	PINN-LSTM
Sequence length	200	200	200	200
Batch size	64	512	512	512
Maximum epochs	100	100	100/10/80	100/10/80
Network configuration	1D CNN	Three-layer GRU	Three-layer MLP	Three-layer LSTM
Hidden units	-	256	256	256
Kernel Size	(64, 3)	-	-	-
Activation function	ReLU	Sigmoid/Tanh	ReLU	Sigmoid/Tanh
Dropout rate	0.10	0.20	0.10	0.20
Optimizer	Adam	Adam	Adam	Adam
learning rate	9×10^{-4}	1×10^{-4}	1×10^{-4}	1×10^{-4}
Loss function	MSE	MSE	MSE + physical loss	MSE + physical loss

Among the comparison models, the CNN represents a local convolutional feature extraction approach, while the GRU serves as a representative time-series model. The MLP-PINN is used to evaluate the effect of introducing voltage-balance physical constraints into a non-sequential network structure. The PINN-LSTM is employed to assess the overall performance achieved by combining long-term temporal modeling with electromagnetic physical constraints.

To improve the statistical rigor of the comparison, each neural-network-based model was independently trained using three random seeds, and the results are reported as

mean $\pm$ standard deviation. As shown in Table 4, all comparison models can estimate the permanent magnet temperature to a certain extent. The MLP-PINN achieved an average MAE of 2.4146 ± 0.1988 °C, an RMSE of 3.2297 ± 0.2555 °C, and an R^2 of 0.9697 ± 0.0047 . Although physical information was incorporated into the feed-forward network, its lack of explicit temporal modeling limits its ability to characterize the thermal dynamics of motor temperature. The CNN achieved an average MAE of 2.0597 ± 0.1882 °C, an RMSE of 2.6895 ± 0.2195 °C, and an R^2 of 0.9790 ± 0.0035 by extracting local temperature variation features through convolutional operations. However, its ability to capture long-term thermal inertia and time-delay characteristics remains limited. The GRU further improved the prediction performance by using a gated recurrent structure to model historical state information, achieving an average MAE of 1.7885 ± 0.0552 °C, an RMSE of 2.4641 ± 0.0480 °C, and an R^2 of 0.9824 ± 0.0007 . This result indicates that recurrent neural networks are more suitable for capturing the long-term dependencies inherent in motor temperature variations. Compared with these baseline models, the proposed PINN-LSTM achieved the best overall performance, with an average MAE of 1.6621 ± 0.0680 °C, an RMSE of 2.3101 ± 0.1463 °C, and an R^2 of 0.9845 ± 0.0020 . Compared with the strongest data-driven baseline, namely the GRU model, the proposed method reduced the MAE and RMSE by approximately 7.1% and 6.2%, respectively. These results demonstrate that the performance improvement of the proposed method is stable across different random initializations and is not caused by a single favorable training run. The prediction curves of different models for a representative run are compared in Figure 6.

Table 4. Comparison of prediction performance among different models.

Model	MAE	MAPE	MSE	RMSE	R^2
MLP-PINN	2.4146 ± 0.1988	4.2717 ± 0.3519	10.4747 ± 1.6135	3.2297 ± 0.2555	0.9697 ± 0.0047
GRU	1.7885 ± 0.0552	3.2869 ± 0.1179	6.0732 ± 0.2356	2.4641 ± 0.0480	0.9824 ± 0.0007
CNN	2.0597 ± 0.1882	3.7698 ± 0.4239	7.2657 ± 1.2015	2.6895 ± 0.2195	0.9790 ± 0.0035
PINN-LSTM	1.6621 ± 0.0680	3.1010 ± 0.0982	5.3509 ± 0.6842	2.3101 ± 0.1463	0.9845 ± 0.0020

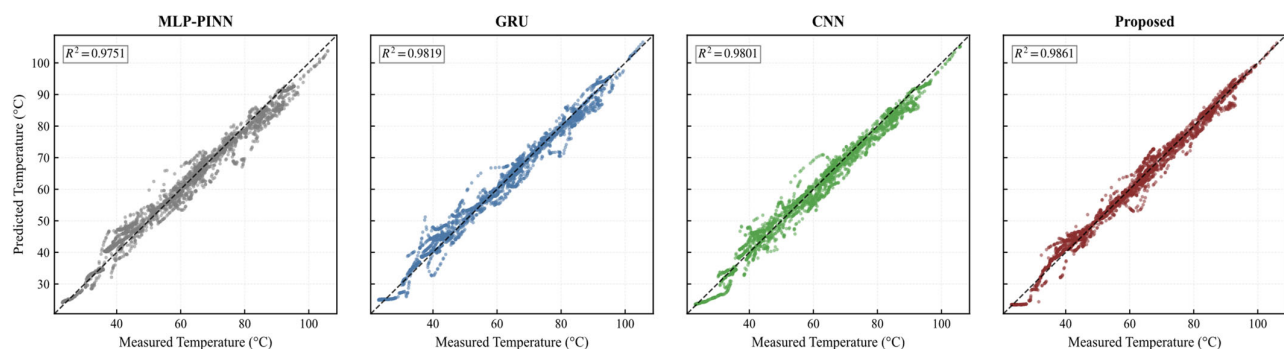


Figure 6. Comparison of prediction results of different models on the test set.

Although the improvement in R^2 appears relatively limited, this is mainly because all models can capture the overall temperature variation trend, resulting in high R^2 values. However, R^2 mainly reflects the global trend-fitting capability and is less sensitive to local absolute temperature deviations. For permanent magnet temperature estimation, MAE and RMSE are more direct indicators because they describe the actual temperature prediction error in degrees Celsius. To improve the statistical rigor of the comparison, each neural-network-based model was independently trained using three random seeds, and the results are reported as mean $\pm$ standard deviation. As shown in Table 4, the proposed PINN-LSTM achieves the lowest absolute prediction errors, with an average MAE of

1.6621 ± 0.0680 °C and an average RMSE of 2.3101 ± 0.1463 °C. Compared with the GRU, CNN, and MLP-PINN models, the average RMSE is reduced from 2.4641 °C, 2.6895 °C, and 3.2297 °C to 2.3101 °C, corresponding to reductions of approximately 6.2%, 14.1%, and 28.5%, respectively. The average MAE is also reduced by approximately 7.1%, 19.3%, and 31.2%, respectively. These results indicate that the proposed method not only maintains comparable trend-fitting capability, but also reduces the actual temperature estimation error. The additional complexity of the proposed PINN-LSTM is therefore justified by both prediction accuracy and physical interpretability. The LSTM structure captures long-term thermal inertia and time-delay characteristics, while the voltage-balance-based physical residual constrains the prediction process from the electromagnetic perspective. In contrast, the MLP-PINN lacks explicit temporal modeling, the CNN mainly extracts local temporal features, and the GRU does not explicitly impose electromagnetic physical consistency. Therefore, the proposed model provides a more balanced framework by combining long-term temporal modeling and voltage-balance physical guidance under sparse sampling conditions.

The predicted temperature curves are presented in Figure 7. The blue dashed vertical lines and the corresponding numbers denote the transitions between different operating conditions. As shown by the prediction curves, both the CNN and GRU models can generally track the overall variation trend of the permanent magnet temperature. However, noticeable deviations remain near temperature peaks and during intervals with rapid operating-condition changes. In comparison, the prediction curve of the PINN-LSTM model is closer to the measured temperature curve. This indicates that the voltage-balance constraint can further suppress physically unreasonable outputs generated by the data-driven model, thereby improving prediction stability under complex operating conditions.

The MLP-PINN model also incorporates the voltage-balance physical constraint but does not include the temporal memory mechanism of the LSTM. The comparison between the MLP-PINN and PINN-LSTM further demonstrates that, although the voltage-balance constraint can improve the physical consistency of the model, permanent magnet temperature estimation still depends on historical operating states over an extended time horizon. Therefore, the LSTM structure is essential for extracting long-term thermal dynamic features.

To further compare the dynamic tracking capability of different models under multi-operating-condition variations, two representative local intervals were selected from the overall test trajectories for enlarged analysis, as shown in Figure 8.

In the first interval, the permanent magnet temperature undergoes multiple step changes as well as repeated rises and drops. The MLP-PINN and CNN models exhibit relatively obvious deviations and time lags near abrupt temperature changes and peak values. The GRU shows improved tracking performance, whereas the prediction curve of the PINN-LSTM is closest to the measured temperature, enabling more accurate tracking of both the overall temperature trend and local turning points.

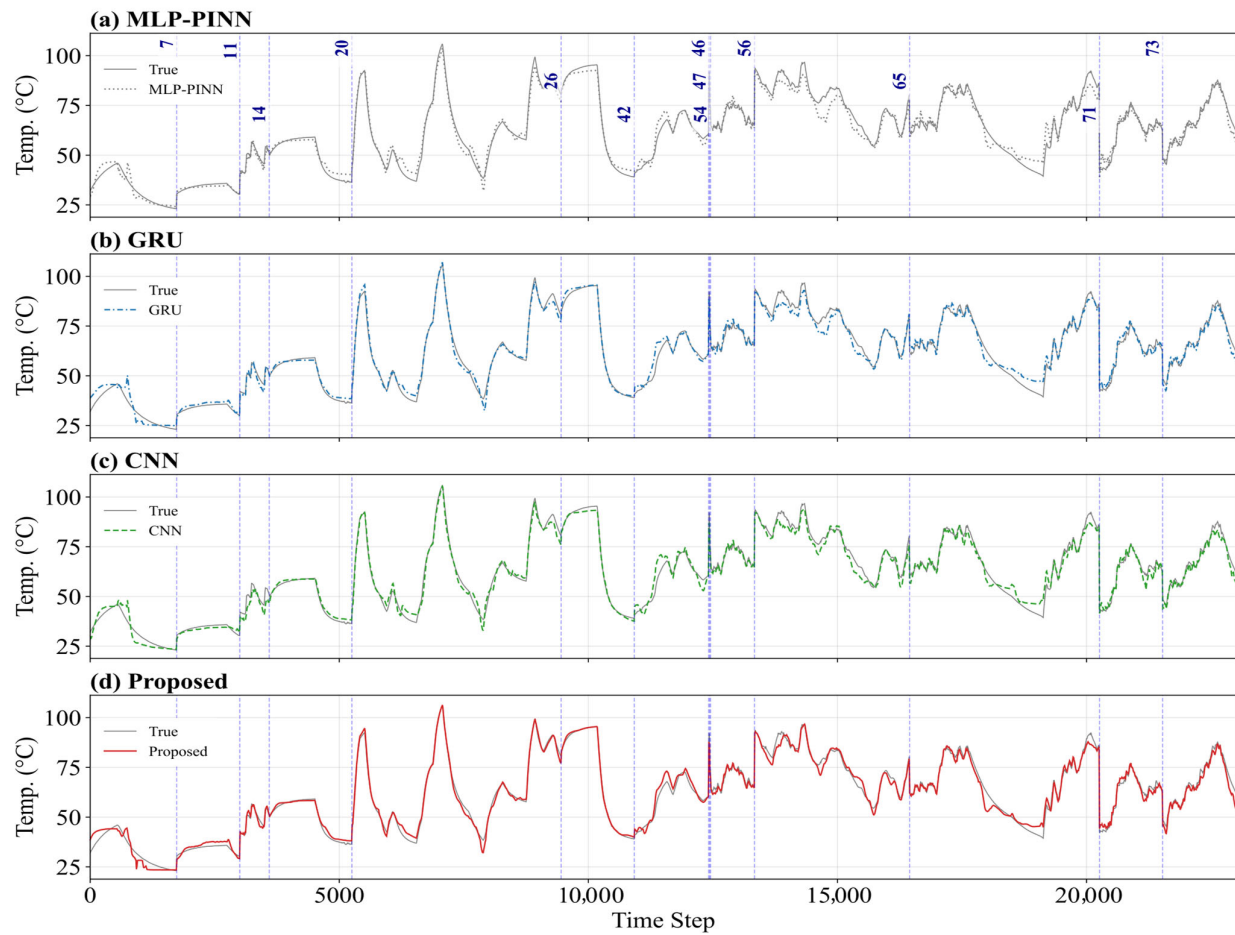


Figure 7. Comparison of predicted temperature trajectories of different models on the test set.

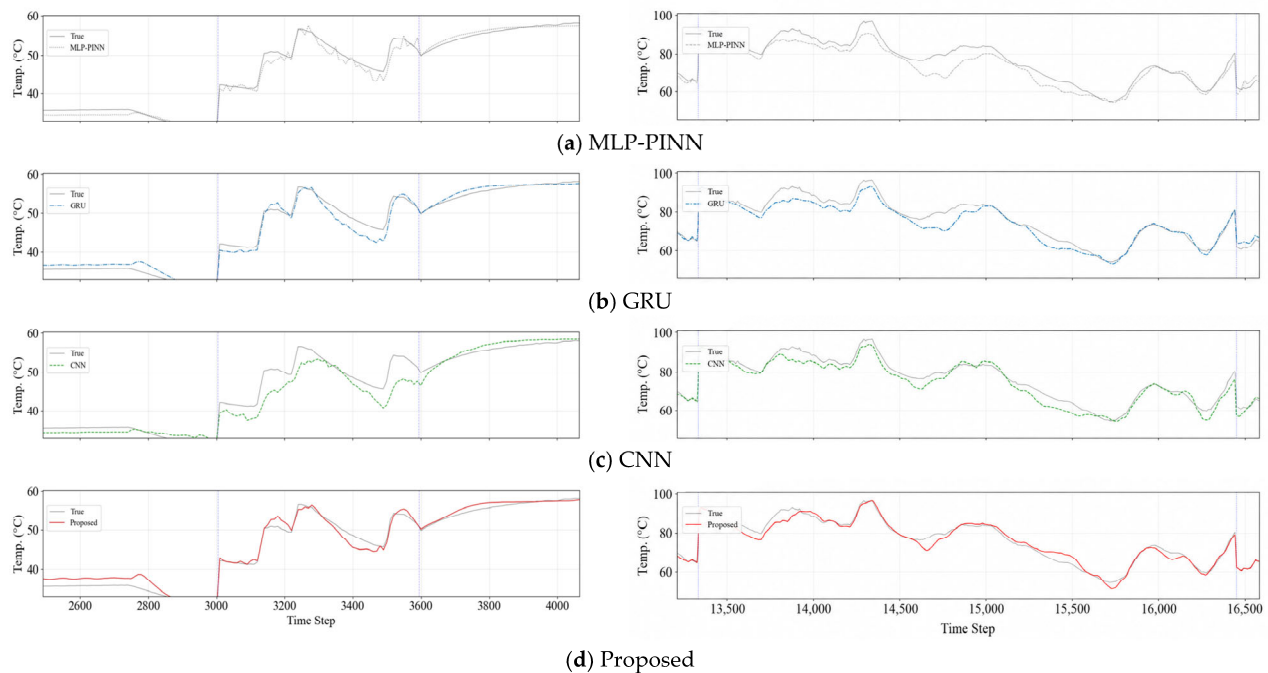


Figure 8. Comparison of model adaptability to varying operating conditions.

The second interval involves more complex variations, including continuous rises and drops in the high-temperature region and multiple peaks and valleys. Compared with the other models, the PINN-LSTM shows smaller deviations during temperature rise, temperature fall, and near extreme values, and its prediction curve is overall smoother. These results indicate that the proposed model has better dynamic tracking capability and prediction stability under complex multi-operating-condition variations.

3.2. Comparative Analysis with a Conventional LPTN Thermal Model

In addition to the data-driven models, an LPTN thermal model was constructed as a conventional physics-based model baseline. The LPTN describes heat transfer within the motor using thermal resistances, thermal capacitances, and heat-source terms. It offers relatively good physical interpretability and low computational complexity. However, its prediction performance depends strongly on the thermal-network topology, the accuracy of thermal resistance and capacitance parameter identification, and the specification of boundary heat-dissipation conditions.

The LPTN baseline developed in this study adopts a four-node structure, in which the stator winding, stator tooth, stator yoke, and permanent magnet are selected as the primary thermal nodes and denoted by T_{sw} , T_{st} , T_{sy} , and T_{pm} , respectively. Here, T_{pm} represents the permanent magnet temperature. The coolant temperature T_{cool} is treated as a thermal boundary condition. The thermal nodes are connected through equivalent thermal resistances, while the thermal capacitance of each node is used to characterize the delayed temperature response caused by thermal inertia. According to the thermal energy balance, the temperature dynamics of an arbitrary thermal node can be expressed as follows:

$$C_i \frac{dT_i}{dt} = P_i + \sum_{j \in N_i} \frac{T_j - T_i}{R_{ij}} \quad (16)$$

where T_i denotes the temperature of the i -th thermal node, C_i is its equivalent thermal capacitance, P_i represents the heat generation at node i , R_{ij} is the equivalent thermal resistance between node i and its adjacent node j , and N_i denotes the set of neighboring nodes connected to node i . The motor losses are treated as the primary heat sources and are defined as follows [28]:

$$\begin{cases} P_{cu} = mR_s (i_d^2 + i_q^2) = k_{cu} (i_d^2 + i_q^2) \\ P_{Fe} = C_{Fe} f^{1.3} B^2 G = k_{Fe} f^{1.3} \\ P_{pm} = \frac{C_{pm} B_m^2 f^2 L_a^2 L_m^3 h_m}{12 \rho_m (L_a + L_m)} = k_{pm} f^2 \end{cases} \quad (17)$$

where P_{cu} , P_{Fe} and P_{pm} denote the winding copper loss, stator iron loss, and permanent magnet eddy-current loss, respectively. Here, m is the number of motor phases; R_s is the stator phase resistance; i_d and i_q are the d- and q-axis currents, respectively; f is the electrical frequency; B and B_m are the magnetic flux density amplitudes in the stator core and permanent magnets, respectively; G is the mass of the stator core; d , l , and ρ denote the conductor diameter, conductor length, and electrical resistivity, respectively; L_a , L_m and h_m are geometric parameters associated with the permanent magnets; and ρ_m is the electrical resistivity of the permanent magnet material. The parameters k_{cu} , k_{Fe} and k_{pm} are equivalent loss coefficients.

The above equations indicate that the winding copper loss is primarily related to the squared current term $i_d^2 + i_q^2$, whereas the stator iron loss and permanent magnet eddy-current loss are mainly associated with the electrical frequency or rotational speed. Since the electrical frequency f is proportional to the mechanical rotational speed ω , these losses

can be further transformed into equivalent heat-source inputs expressed in terms of current and rotational speed. Considering that the publicly available dataset does not provide complete motor geometry, material properties, magnetic flux density distributions, or boundary heat-transfer conditions, the principal physical relationships among the measurable variables are retained, and the total equivalent heat source is expressed as follows:

$$P_{eq} = k_1(i_d^2 + i_q^2) + k_2\omega^{1.3} + k_3\omega^2 \quad (18)$$

where ω denotes the motor rotational speed, and k_1, k_2, k_3 are equivalent loss coefficients to be identified. In this expression, $k_1(i_d^2 + i_q^2)$ represents current-dependent losses, $k_2\omega^{1.3}$ characterizes the stator iron loss, and $k_3\omega^2$ represents speed-squared-dependent losses, such as the permanent magnet eddy-current loss. Through this equivalent formulation, the principal loss mechanisms can be incorporated into the LPTN model without requiring detailed geometric or material parameters, while the equivalent coefficients are identified using the training set.

Based on the four-node LPTN structure developed in this study, as illustrated in Figure 9, the thermal balance equations for the individual nodes can be written as follows:

$$\left\{ \begin{array}{l} C_{sw} \frac{dT_{sw}}{dt} = P_{cu} + \frac{T_{sy} - T_{sw}}{R_{sy-sw}} + \frac{T_{st} - T_{sw}}{R_{sw-st}} \\ C_{st} \frac{dT_{st}}{dt} = \beta_1 P_{Fe} + \frac{T_{sw} - T_{st}}{R_{sw-st}} + \frac{T_{sy} - T_{st}}{R_{sy-st}} + \frac{T_{pm} - T_{st}}{R_{st-pm}} \\ C_{sy} \frac{dT_{sy}}{dt} = \beta_2 P_{Fe} + \frac{T_{cool} - T_{sy}}{R_{cool-sy}} + \frac{T_{sw} - T_{sy}}{R_{sy-sw}} + \frac{T_{st} - T_{sy}}{R_{st-sy}} \\ C_{pm} \frac{dT_{pm}}{dt} = P_{pm} + \frac{T_{st} - T_{pm}}{R_{st-pm}} \end{array} \right. \quad (19)$$

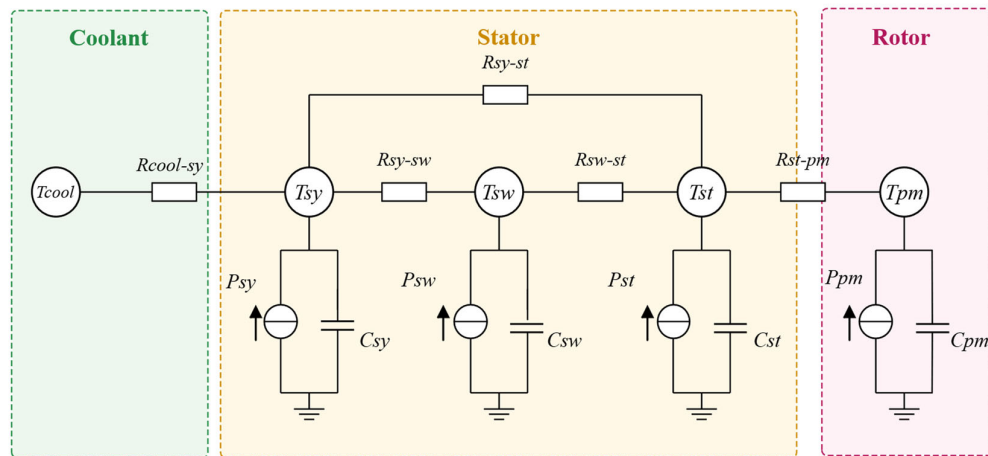


Figure 9. Structure of the four-node lumped-parameter thermal network.

For the low-frequency sampled data with a sampling interval of 5 s used in this study, the continuous-time thermal balance equations were discretized using the forward Euler method:

$$T_i(k+1) = T_i(k) + \frac{\Delta t}{C_i} \left[P_i(k) + \sum_{j \in N_i} \frac{T_j(k) - T_i(k)}{R_{ij}} \right] \quad (20)$$

$\Delta t = 5$ s, consistent with the resampling interval adopted in this study. The model performs thermal-state propagation over continuous sequences of 200 time steps, corresponding to a thermal dynamic duration of 1000 s. The equivalent thermal resistances, thermal capacitances, and heat-generation coefficients in the LPTN model were treated as identifiable equivalent parameters because the public dataset does not provide complete motor geometry, material properties, or boundary heat-transfer coefficients. These parameters were calibrated on the training set by minimizing the MSE between the predicted and measured temperatures of the available thermal nodes. The validation set was used to select the best parameter set, and the test set was used only for final evaluation.

We further introduce TNN-PM-VN as a comparison model. This model follows an LPTN-inspired recursive thermal-state neural network design. It treats the PM temperature as the internal thermal state to be estimated [29]. It uses the stator winding, stator tooth, stator yoke, and coolant temperatures as measurable external thermal nodes. It also includes a fixed virtual node to compensate for unmodeled heat-transfer paths and residual thermal dynamics.

At each sampling instant, the model takes the current PM temperature state, external thermal-node temperatures, virtual-node temperature, and operating variables as inputs. The operating variables include d-q axis voltages, d-q axis currents, motor speed, and torque. The π -Net learns the equivalent heat-generation term. The γ -Net learns the equivalent thermal conductance between the PM node and the external thermal nodes. A positively constrained parameter κ represents the inverse equivalent thermal capacitance. These components recursively update the PM temperature at the next time step. During training, the model uses the measured PM temperature at the start of each window as the initial state.

It then recursively predicts the remaining PM temperature sequence within the window. The loss function is the MSE between the predicted and measured temperature sequences. This model, together with the conventional LPTN, serves as a thermal-model-based baseline. It helps evaluate the proposed PINN-LSTM against thermal networks and their neural extensions. The prediction results are presented in Table 5.

Table 5. Comparison of thermal-model-based baselines and the proposed method.

Model	MAE	RMSE	R ²
LPTN	3.9379	6.4495	0.88159
TNN-PM-VN	1.8103	2.2883	0.98483
Proposed	1.6048	2.1890	0.98613

As shown in Table 5, the conventional four-node LPTN model shows the largest prediction error under the 5 s sparse-sampling condition, with an MAE of 3.9379 °C, an RMSE of 6.4495 °C, and an R² of 0.88159. This result indicates that the simplified low-order LPTN can capture the general temperature trend, but it struggles to reproduce the dynamic PM temperature response accurately. The main reason is that the LPTN depends on a predefined thermal-network topology and equivalent parameters, such as thermal resistances, thermal capacitances, and loss coefficients. However, the public dataset does not provide detailed motor geometry, material thermal properties, magnetic flux density distribution, or boundary heat-transfer conditions. Therefore, the identified low-order equivalent parameters can only approximate the actual heat-transfer process, which leads to larger errors during complex dynamic operation and long-term recursive prediction. The TNN-PM-VN model improves the LPTN result by introducing neural approximation into the LPTN-inspired thermal-state update. It achieves an MAE of 1.8103 °C, an RMSE of 2.2883 °C, and an R² of 0.98483.

Although the prediction curves of TNN-PM-VN and the proposed model appear visually close in Figure 10, the residual curves and quantitative metrics show that the proposed model provides lower absolute prediction errors over the full test interval. Compared with LPTN, the proposed model reduces the MAE and RMSE by approximately 59.2% and 66.1%, respectively. Compared with TNN-PM-VN, the proposed model further reduces the MAE and RMSE by approximately 11.3% and 4.3%, respectively. These results demonstrate that the proposed PINN-LSTM provides better dynamic tracking capability and physical consistency than both the conventional LPTN and its neural thermal extension under sparse sampling conditions.

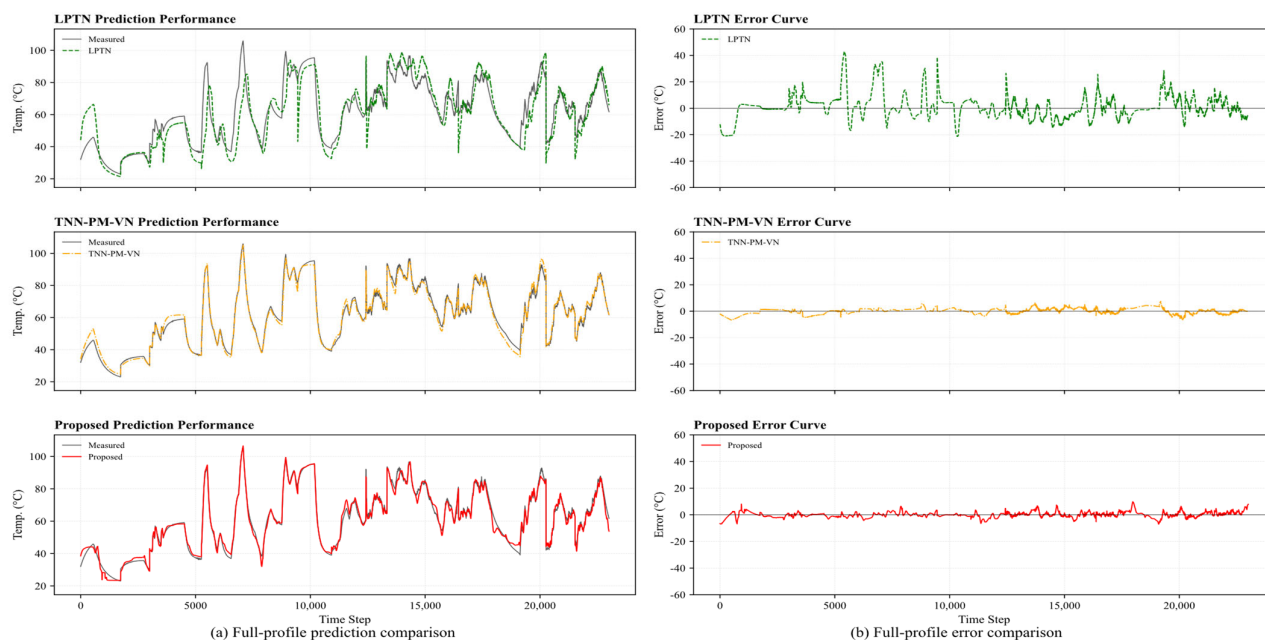


Figure 10. Comparison of prediction trajectories and errors between the LPTN and PINN-LSTM models.

To provide a clearer visual comparison among all the investigated models, the prediction and error curves of LPTN, TNN-PM-VN, MLP-PINN, CNN, GRU, and the proposed PINN-LSTM are plotted together in Figure 11. Unlike the previous separate figures, this combined figure allows the prediction trends and residual errors of all models to be compared under the same time axis and coordinate scale. Figure 11a shows the measured and predicted PM temperature curves, while Figure 11b shows the corresponding prediction errors.

As shown in Figure 11, all models can follow the overall PM temperature trend, which explains their relatively close R^2 values. However, the residual curves reveal clearer differences during operating-condition transitions. The conventional LPTN shows large error fluctuations under dynamic changes, indicating that the low-order thermal network has limited ability to describe complex transient thermal behavior. TNN-PM-VN reduces these fluctuations through LPTN-inspired neural thermal-state modeling, but noticeable deviations still appear in several transient regions. In contrast, the proposed PINN-LSTM maintains smaller and more stable residual errors during condition changes, showing better robustness to dynamic operating variations. This improvement is attributed to the combination of LSTM-based long-term temporal modeling and voltage-balance-based physical constraints.

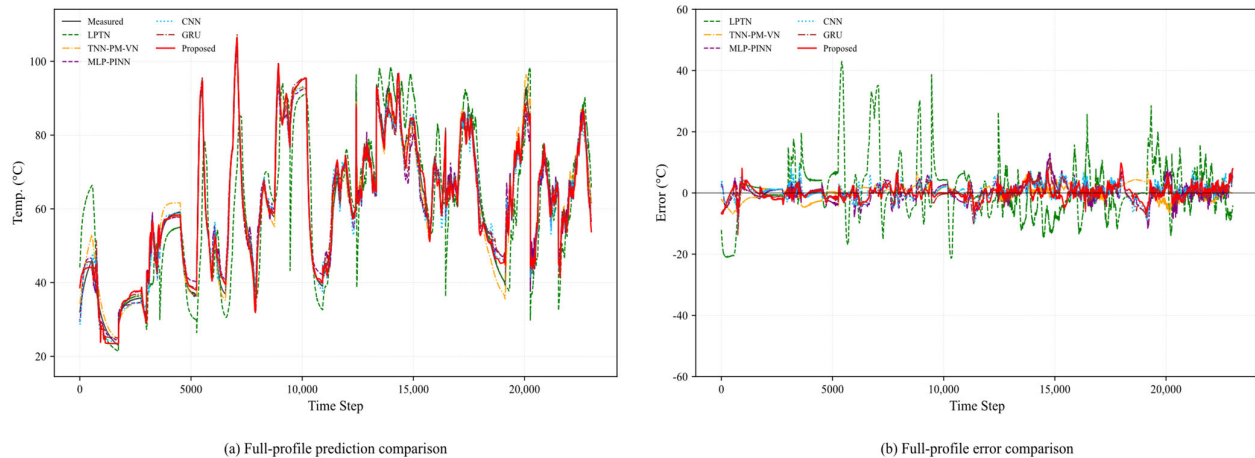


Figure 11. Full-profile prediction and error comparison of all models: (a) measured and predicted PM temperature curves; (b) prediction error curves.

In contrast, temporal data-driven models achieve lower RMSE and MAE values than the purely physics-based LPTN model, indicating that neural networks are more effective in characterizing complex nonlinear thermal dynamics from operating data. In particular, CNN and GRU improve the prediction accuracy by extracting local temporal features and historical state information, respectively. However, these conventional data-driven models still rely mainly on statistical correlations in the training data and lack explicit physical constraints. As a result, their residual errors remain larger during rapid temperature changes and operating-condition transitions. The proposed PINN-LSTM further improves the prediction performance by combining LSTM-based long-term temporal modeling with voltage-balance-based physical constraints. The LSTM captures thermal inertia and time-delay characteristics, while the electromagnetic residual guides the prediction process toward physically consistent solutions. Therefore, the results demonstrate that both temporal modeling and physical constraints are important for accurate and stable PM temperature estimation under sparse sampling conditions.

3.3. Ablation Study

To further investigate the sources of the performance improvements achieved by the PINN-LSTM model, several ablation experiments were conducted, as summarized in Table 6. The evaluated configurations include a standalone LSTM, a PINN-LSTM with fixed physical parameters, a PINN-LSTM with learnable physical parameters, and the complete PINN-LSTM model.

Table 6. Results of the ablation study.

Experiment	Voltage Constraint	Parameter Identification	Training Strategy	MAE	RMSE	R^2
A1	×	None	Single-stage data-driven training	2.0027	2.7857	0.9775
A2	√	Fixed physical parameters	Single-stage joint training	1.7037	2.3976	0.9834
A3	√	Learnable equivalent parameters	Single-stage joint training	1.6765	2.3777	0.9836
A4	√	Learnable equivalent parameters	Three-stage training with weight annealing	1.6048	2.1890	0.9861

The ablation results show that the standalone LSTM achieved an MAE of 2.0027 °C, an RMSE of 2.7857 °C, and an R^2 of 0.9775. After introducing the voltage-balance physical residual, the MAE and RMSE of the PINN-LSTM with fixed physical parameters decreased to 1.7037 °C and 2.3976 °C, respectively, while the R^2 increased to 0.9834. These results indicate that the voltage-balance physical residual effectively constrains the model predictions and improves the accuracy of permanent magnet temperature estimation.

When the electromagnetic parameters involved in the physical residual were made learnable, the resulting PINN-LSTM achieved an MAE of 1.6765 °C, an RMSE of 2.3777 °C, and an R^2 of 0.9836, further improving upon the fixed-parameter model. This result demonstrates that, when accurate manufacturer parameters are unavailable in a public dataset, adaptive identification of equivalent electromagnetic parameters enhances the compatibility of the physical residual with the actual operating data.

The complete PINN-LSTM model, which combines the voltage-balance physical residual, learnable equivalent electromagnetic parameters, and the three-stage training strategy, achieved the best prediction performance, with an MAE of 1.6048 °C, an RMSE of 2.1890 °C, and an R^2 of 0.9861. Compared with the standalone LSTM, the complete model reduced the RMSE by approximately 21.4%. These results demonstrate that the proposed method effectively integrates data-driven temporal modeling with electromagnetic physical constraints, thereby further improving the accuracy and stability of permanent magnet temperature estimation.

3.4. Sensitivity Analysis of the Physical Loss Weight

The physical loss weight λ determines the relative contribution of the voltage-balance residual to the total training loss. To evaluate its influence and justify the adopted training strategy, a sensitivity analysis was conducted by comparing several fixed λ values with the proposed annealing schedule. The fixed-weight settings included $\lambda = 0$, 1×10^{-4} , 1×10^{-5} . The adopted schedule used $\lambda = 1 \times 10^{-4}$ in the early joint fine-tuning stage, $\lambda = 1 \times 10^{-5}$ in the middle stage, and $\lambda = 0$ in the final refinement stage. The other training settings were kept unchanged.

As shown in Table 7, the fixed-weight cases exhibit similar prediction performance. When λ is set to 0, 1×10^{-4} , and 1×10^{-5} , the RMSE varies only from 2.2127 °C to 2.2217 °C, and R^2 varies from 0.98571 to 0.98582. This result indicates that the model is not highly sensitive to a specific fixed λ value. However, the fixed physical-loss weights do not clearly outperform the $\lambda = 0$ case, suggesting that a constant physical constraint is not the most effective way to use the voltage-balance residual.

Table 7. Sensitivity analysis of the physical loss weight λ .

Strategy	λ Setting	MAE	MAPE	MSE	RMSE	R^2
S0	0	1.6232	3.1282	4.8962	2.2127	0.98582
S1	1×10^{-4}	1.6228	3.1025	4.9358	2.2217	0.98571
S2	1×10^{-5}	1.6223	3.1229	4.9196	2.2180	0.98576
Adopted	$1 \times 10^{-4} \rightarrow 1 \times 10^{-5} \rightarrow 0$	1.6048	3.0479	4.7917	2.1890	0.98613

The adopted annealing strategy achieves the best overall performance, with an MAE of 1.6048 °C, an RMSE of 2.1890 °C, and an R^2 of 0.98613. This shows that the physical loss is more effective when it is used as a staged guidance term. The early and middle stages use the voltage-balance residual to improve electromagnetic physical consistency, whereas the final 10 epochs sets λ to 0 to avoid over-constraining the temperature prediction with the approximate quasi-steady-state voltage equation. Therefore, the adopted strategy provides a better balance between physical consistency and supervised prediction accuracy than fixed-weight training.

3.5. Prediction Performance Under Different Downsampling Factors

To evaluate the applicability of the proposed model under low-frequency sampling conditions, experiments were conducted using four downsampling factors: 5×, 10×, 15×, and 20×. These downsampling factors correspond to different sampling intervals. To ensure a fair comparison across experiments, the historical input duration under each downsampling setting was kept consistent with that of the baseline 10× downsampling case.

As shown in Table 8, the prediction performance of the models varies across different downsampling factors. For the CNN model, the MAE and RMSE generally increase as the downsampling factor increases, indicating that local convolutional features alone are insufficient to fully characterize long-term thermal dynamics under low-frequency sampling conditions. The GRU model achieved relatively good performance under 5× downsampling, with an MAE of 1.6720 °C and an RMSE of 2.3635 °C. However, its prediction errors gradually increased with the sampling interval, and the RMSE reached 2.8265 °C under 20× downsampling. Although the MLP-PINN incorporates physical constraints, its overall prediction errors remain relatively high because it lacks an effective temporal recurrence structure.

Table 8. Comparison of model prediction performance under different downsampling factors.

Model	Metric	5×	10×	15×	20×
CNN	MAE	1.9275	1.9870	2.1363	2.2344
	RMSE	2.5458	2.6198	2.8797	3.0383
	R ²	0.9812	0.9801	0.9760	0.9733
GRU	MAE	1.6720	1.8070	1.8752	1.9994
	RMSE	2.3635	2.5002	2.5549	2.8265
	R ²	0.9838	0.9819	0.9811	0.9769
MLP-PINN	MAE	2.1965	2.1869	2.4687	2.4993
	RMSE	2.8988	2.9354	3.3528	3.4056
	R ²	0.9757	0.9750	0.9675	0.9664
PINN-LSTM	MAE	1.7889	1.6048	1.8468	1.8710
	RMSE	2.4563	2.1890	2.5060	2.4856
	R ²	0.9825	0.9861	0.9818	0.9821

In comparison, the PINN-LSTM achieved the best or nearly the best prediction performance under the 10×, 15×, and 20× downsampling conditions. The model performed best under 10× downsampling, achieving an MAE of 1.6048 °C, an RMSE of 2.1890 °C, and an R² of 0.9861. Even under 20× downsampling, the RMSE of the PINN-LSTM remained at 2.4856 °C, which was lower than those of the other comparison models. These results indicate that combining the long-term temporal modeling capability of the LSTM with the voltage-balance physical constraint can effectively mitigate the loss of information caused by low-frequency sampling, thereby improving the stability and accuracy of permanent magnet temperature estimation.

3.6. Error Analysis in the High-Temperature Region

For motor temperature monitoring, overall metrics such as MAE, RMSE, and R² can reflect the average prediction performance but cannot fully characterize model reliability in thermally critical regions. In practical engineering applications, particular attention is paid to whether a model can maintain low prediction errors under high-temperature conditions, because excessive permanent magnet temperature may degrade magnetic properties and even increase the risk of demagnetization.

To evaluate the prediction capability of the models under thermally critical conditions, the samples corresponding to the highest 25% of permanent magnet temperatures in the test set were defined as the high-temperature region, with a threshold of approximately 74.95 °C. As shown in Figure 12, the PINN-LSTM achieved an MAE of 1.363 °C, an RMSE of 1.8958 °C, and an R^2 of 0.9125 in this region, outperforming all comparison models. Its R^2 was considerably higher than those of the CNN, GRU, and MLP-PINN models, which were 0.8460, 0.8250, and 0.6839, respectively. This indicates that the PINN-LSTM can more accurately track permanent magnet temperature variations in the high-temperature region. Moreover, compared with the CNN, GRU, and MLP-PINN models, the PINN-LSTM reduced the RMSE by approximately 24.6%, 29.3%, and 47.4%, respectively, demonstrating higher prediction accuracy and stability in the high-temperature subset.

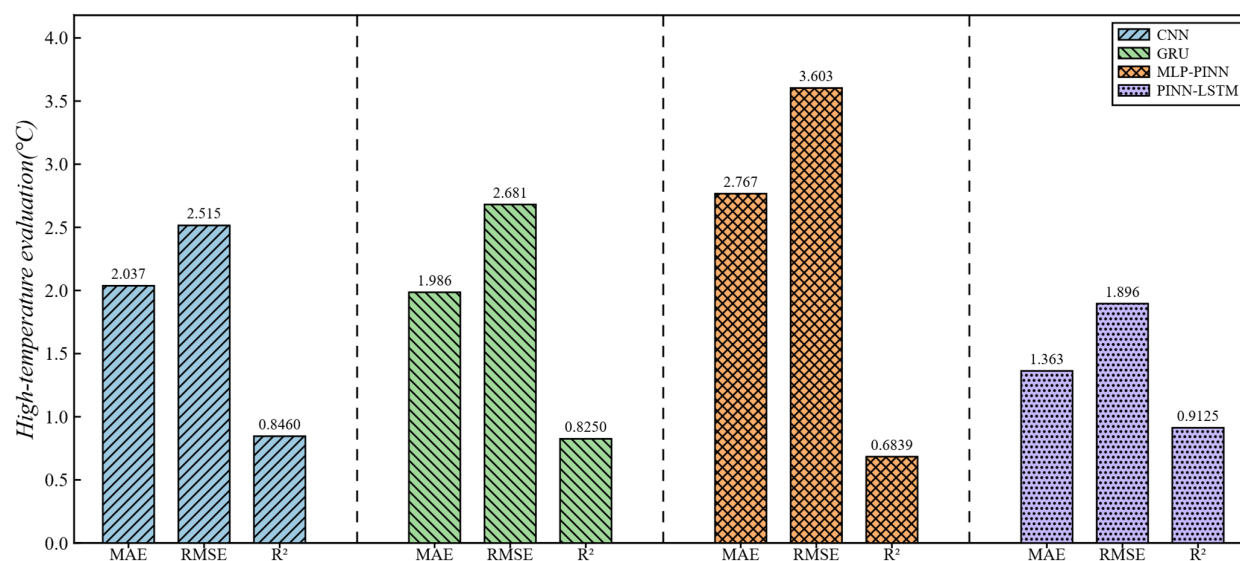


Figure 12. Comparison of prediction errors in the high-temperature region.

These results show that the PINN-LSTM not only achieves low prediction errors over the entire test set but also maintains a clear advantage in the high-temperature region. In addition, its error is substantially lower than those of the other models, indicating better error-control capability in thermally critical regions. This finding suggests that the voltage-balance physical constraint contributes to improving prediction reliability under critical thermal conditions, thereby enhancing the potential applicability of the proposed model in motor health management and safety monitoring.

3.7. Validation of the Identified Physical Parameters

In the proposed PINN-LSTM model, the voltage-balance residual is coupled with the neural-network prediction through several trainable equivalent electromagnetic parameters. To avoid physically unrealistic values during optimization, physical bounds were imposed on the reference stator resistance R_{20} , the d-axis inductance L_d , the q-axis inductance L_q , and the reference PM flux linkage ψ_{20} . It should be noted that $R_s(t)$ was not treated as an independent trainable parameter. Instead, R_{20} was optimized as a bounded equivalent parameter, and the instantaneous stator resistance $R_s(t)$ was calculated from R_{20} using the measured stator winding temperature.

To further evaluate the stability and uncertainty of the learned equivalent parameters, the proposed model was trained using three independent random seeds. The mean value,

standard deviation, and coefficient of variation of each parameter are summarized in Table 9.

Table 9. Physical bounds and uncertainty analysis of the learned equivalent electromagnetic parameters.

Parameter	Initial Value	Physical Bound	Identified Value, Mean $\pm$ std.	CV (%)	Unit
R_{20}	0.0400	[0.0100, 0.1000]	0.03532 ± 0.00073	2.07	Ω
ψ_{20}	0.0800	[0.0400, 0.1200]	0.05865 ± 0.00062	1.05	Wb
L_d	1.0000	[0.1000, 2.0000]	0.25070 ± 0.00412	1.64	mH
L_q	1.5000	[0.1000, 2.0000]	0.37162 ± 0.00281	0.76	mH

As shown in Table 8, all learned equivalent electromagnetic parameters remain within the imposed physical bounds and do not reach the upper or lower limits. The learned L_q is larger than L_d , which is consistent with the saliency characteristic of an interior PMSM. In addition, the coefficients of variation of R_{20} , ψ_{20} , L_d , and L_q are 2.07%, 1.05%, 1.64%, and 0.76%, respectively, indicating that the equivalent-parameter learning process is stable across three random seeds.

However, these parameters should be interpreted as bounded equivalent electromagnetic parameters rather than exact manufacturer-specified motor parameters. Since the true values of R_{20} , L_d , L_q , and ψ_{20} are unavailable in the public dataset, they cannot be directly validated against the actual motor parameters. Moreover, due to the coupling in the voltage-balance equations, different parameter combinations may produce similar voltage residuals. Therefore, the proposed parameter learning should be regarded as physics-guided equivalent-parameter calibration for improving voltage-balance consistency, rather than unique online identification of the true motor parameters.

Because the publicly available dataset does not provide the nominal inductance parameters of the tested motor, the relative errors between the identified inductances and their true values cannot be calculated directly. Based on this consideration, the voltage-reconstruction analysis is used only as a physics-consistency check, rather than as proof of exact parameter identification. Therefore, an indirect validation was performed using test operating profiles that were not involved in parameter calibration. Specifically, the identified L_d and L_q were substituted into the quasi-steady-state d–q-axis voltage equations. The d- and q-axis voltages were then reconstructed using the measured currents, rotational speed, and other physical parameters from the test set. Finally, the reconstructed voltages $\hat{u}_d$ and $\hat{u}_q$ were compared with the measured voltages u_d and u_q in the dataset, and the voltage-reconstruction consistency was qualitatively examined by comparing the reconstructed and measured voltage trajectories.

As shown in Figure 13, for the two selected high-temperature operating profiles, the d- and q-axis voltages reconstructed using the identified parameters can generally track the measured voltage variations well. In Profile 20, the reconstructed results accurately capture the step change in voltage from a low level to a stable operating state, and the two curves remain highly consistent after the steady-state condition is reached. During the continuous dynamic operation of Profile 56, the reconstructed voltages also follow the main fluctuation trends of the measured voltages. In particular, the reconstruction consistency of the d-axis voltage is relatively high. Although some deviations are observed in the q-axis voltage over certain intervals, the overall variation trend and magnitude range remain in good agreement. These results indicate that the identified parameter set can adequately describe the coupling relationships among voltage, current, and rotational speed under different high-temperature test conditions, thereby providing intuitive support for its physical consistency.

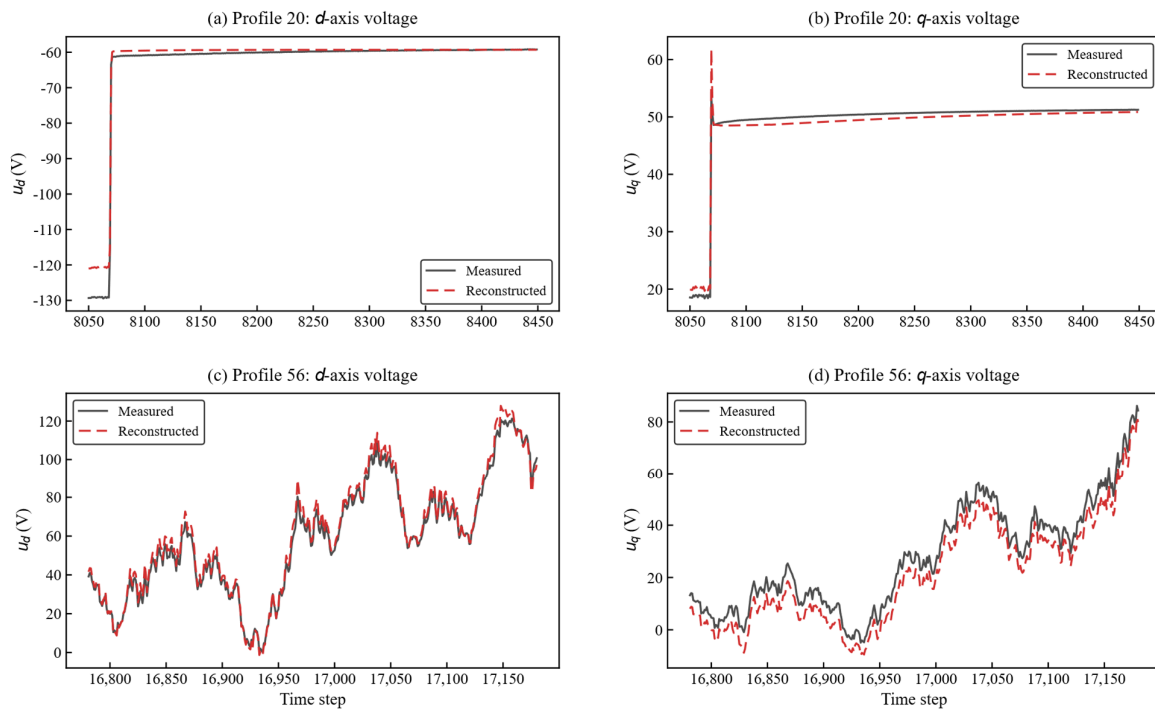


Figure 13. Reconstructed d- and q-axis voltages based on the identified parameters.

After model convergence, the identified equivalent inductances were $L_d = 0.249$ mH and $L_q = 0.373$ mH, respectively, showing the salient-pole characteristic of an interior permanent magnet synchronous motor, namely $L_q > L_d$. In terms of numerical magnitude, both inductances are below 1 mH and do not exhibit any obvious abnormality. It should be noted that the voltage reconstruction only verifies the consistency of this set of equivalent parameters with the voltage balance relationship reflected in the test data, and cannot be directly interpreted as a highly accurate measurement of the true motor inductance parameters.

3.8. Sensitivity Analysis of Temperature Coefficients

When constructing the physical constraint based on the voltage-balance equations, the permanent magnet flux-linkage temperature coefficient α_{PM} in Equation (8) and the stator winding resistance temperature coefficient α_{Cu} in Equation (9) are key physical parameters describing the temperature-dependent evolution of the electromagnetic parameters. In practical applications, the temperature characteristics of NdFeB permanent magnets may vary with material grade or aging, while the effective resistance temperature coefficient of the copper windings may also be affected by the skin effect and measurement errors. To evaluate the robustness of the proposed PINN method against uncertainties in these physical parameters, a sensitivity analysis of the two key temperature coefficients was conducted in this section.

The network architecture and hyperparameter settings were kept unchanged, while artificial perturbations were introduced into α_{pm} and α_{cu} in the physical loss function. The perturbation levels were set to $\pm 10\%$ and $\pm 20\%$ of their nominal values. The nominal value of α_{pm} was set to -0.0012 K $^{-1}$, as adopted in this study, while α_{cu} was set to the standard temperature coefficient of copper, 0.00393 K $^{-1}$. Both α_{PM} and α_{Cu} were simultaneously perturbed by the same relative percentage. The resulting changes in permanent magnet temperature estimation performance on the test set were then evaluated, as summarized in Table 10.

Table 10. Effects of different temperature coefficients on model performance.

Perturbation Level	MAE	RMSE	R ²
+20%	1.6027	2.2031	0.98595
+10%	1.6479	2.2437	0.98542
0	1.6048	2.1890	0.98613
−10%	1.6048	2.2145	0.98580
−20%	1.6095	2.2039	0.98594

Table 10 summarizes the model performance when the relevant physical parameters were perturbed within $\pm 20\%$ of their nominal values. The results show that all evaluation metrics remained highly stable. The RMSE varied only slightly, from 2.19 °C to 2.24 °C, while R^2 consistently remained above 0.985. These results indicate that prior uncertainties in the physical parameters did not significantly affect the temperature estimation accuracy.

This robustness can be attributed to the soft-constraint coupling mechanism. The physical equations provide supplementary electromagnetic consistency constraints on the model, thereby reducing the likelihood of physically unreasonable predictions under sparse sampling conditions. Meanwhile, the LSTM-based data-driven branch exhibits strong adaptive compensation capability and can partially absorb the residual deviations caused by parameter perturbations. Consequently, the proposed model retains physical interpretability while remaining flexible in the presence of parameter uncertainties encountered in practical engineering applications, resulting in greater robustness than conventional physics-based models with fixed parameters.

3.9. Effects of Different Temperature Input Combinations on Prediction Performance

To investigate the contributions of different temperature measurement locations to permanent magnet temperature estimation and assess the feasibility of reducing the number of temperature sensors, PINN-LSTM models were constructed using different combinations of temperature inputs while keeping the network architecture and training parameters unchanged. The importance of each temperature measurement location was evaluated by analyzing the variations in prediction accuracy across the different input combinations.

The stator winding temperature, coolant temperature, stator tooth temperature, and stator yoke temperature were selected as candidate thermal features, and eight input combinations were designed. All models retained the operating variables, including voltage, current, rotational speed, and torque, while only the temperature input features were varied. MAE, RMSE, and R^2 were adopted as the evaluation metrics, and the experimental results are presented in Table 11.

Table 11. Evaluation metrics of the model under different temperature input combinations.

Temperature Input Combination	MAE (°C)	RMSE (°C)	R ²
T_{sw}	3.4667	4.6465	0.9375
$T_{sw} + T_{cool}$	3.3092	5.5462	0.9109
$T_{sw} + T_{st}$	3.7356	5.3331	0.9176
$T_{sw} + T_{sy}$	3.7095	5.4112	0.9152
$T_{sw} + T_{st} + T_{sy}$	2.0382	2.6365	0.9799
$T_{sw} + T_{cool} + T_{sy}$	2.4035	3.2684	0.9691
$T_{sw} + T_{cool} + T_{st}$	1.6695	2.2942	0.9848
$T_{sw} + T_{cool} + T_{sy} + T_{st}$	1.6048	2.1890	0.9861

The four temperature inputs correspond to the stator winding, stator yoke, stator tooth, and coolant temperatures. Compared with the model using temperatures from all four locations, removing any one of the four temperature inputs reduced the prediction accuracy to varying degrees. The influence of each temperature measurement location on model performance is closely related to the motor structure and internal heat-transfer mechanisms.

Among these features, the stator tooth temperature can effectively reflect the thermal state near the air gap and the variation in stator–rotor thermal coupling, making it a key input for accurate permanent magnet temperature estimation. The coolant temperature characterizes the thermal boundary condition of the motor and therefore plays an important role in the long-term temperature-rise process. In comparison, the additional information provided by the stator yoke temperature is relatively limited.

When the stator winding, coolant, and stator tooth temperatures are used as inputs, the model achieves prediction performance close to that of the full-feature model. Therefore, in practical applications, these three temperature signals can be preferentially retained to reduce the complexity of the temperature measurement system while maintaining satisfactory prediction accuracy.

4. Conclusions

To address the difficulty of directly measuring permanent magnet temperature in PMSMs and the degradation of prediction accuracy under low-frequency sampling conditions, this study proposed a PINN-LSTM temperature estimation method. The proposed model combines LSTM-based temporal feature extraction with a physics-informed voltage-balance constraint. By incorporating the PMSM d-q-axis voltage equations into the training process, the model jointly considers data-driven temperature fitting and electromagnetic physical consistency.

The rationality of the quasi-steady-state voltage constraint was verified by evaluating the neglected current derivative terms. The results based on both the original 2 Hz data and the downsampled 5 s data show that the contribution ratios of the inductive voltage terms remain far below the 1% reference threshold. This supports the use of the quasi-steady-state d-q-axis voltage equations as soft physical constraints under the sparse-sampling condition considered in this study.

Experiments on a public PMSM thermal dataset demonstrate that the proposed PINN-LSTM achieves better overall prediction performance than CNN, GRU, MLP-PINN, LPTN, and TNN-PM-VN models. In a representative run, the proposed model achieved an MAE of 1.6048 °C, an RMSE of 2.1890 °C, and an R^2 of 0.9861. The results under different downsampling factors, high-temperature regions, and operating-condition transitions further show that the proposed method has good prediction accuracy and stability under sparse sampling conditions.

Despite these promising results, this study still has limitations. The experiments were conducted on a single public PMSM dataset from one motor/testbench, and the proposed method has not yet been validated on different motor types, cooling structures, or independent experimental platforms. Therefore, the cross-motor and cross-platform generalization capability of the model requires further investigation. In addition, although GRU was included as a representative sequential baseline in this study, more advanced time-series models, such as Temporal Fusion Transformer, Informer, Autoformer, and other Transformer-based long-sequence architectures, were not fully investigated. Future work will incorporate multiphysics simulation data and experimental data from different motors to further evaluate model transferability, improve robustness under unknown operating conditions, and enhance the engineering applicability of the proposed method. Moreover, Transformer-family models will be introduced as extended baselines or

integrated with physics-informed constraints to further evaluate the potential of advanced sequence modeling for PMSM temperature estimation.

Author Contributions: Conceptualization, J.L., L.Y. and H.G.; methodology, J.L.; software, J.L.; validation, J.L., L.Y. and M.J.; formal analysis, J.L.; investigation, J.L. and L.Y.; resources, J.O., M.J. and H.G.; data curation, J.L.; writing—original draft preparation, J.L.; writing—review and editing, L.Y., J.O., M.J. and H.G.; visualization, J.L.; supervision, J.O. and H.G.; project administration, H.G.; funding acquisition, J.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number [52207041]; Liaoning Provincial Science and Technology Plan Joint Program (Key Research and Development Program Project), 2025JH2/101800399.

Data Availability Statement: The dataset used in this study is publicly available from the Electric Motor Temperature Dataset. The implementation code and processed data-splitting scripts can be made available from the corresponding author upon reasonable request.

Conflicts of Interest: Author Mengran Ji was employed by the company Shenyang Chemical Co., Ltd., Shenyang, China. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Zhu, Z.-Q.; Liang, D. Perspective of Thermal Analysis and Management for Permanent Magnet Machines, with Particular Reference to Hotspot Temperatures. *Energies* **2022**, *15*, 8189.
2. Gronwald, P.-O.; Kern, T.A. Traction Motor Cooling Systems: A Literature Review and Comparative Study. *IEEE Trans. Transp. Electr.* **2021**, *7*, 2892–2913.
3. He, L.; Feng, Y.; Zhang, Y.; Tong, B. Methods for temperature estimation and monitoring of permanent magnet: A technology review and future trends. *J. Braz. Soc. Mech. Sci. Eng.* **2024**, *46*, 174.
4. Su, Y.; Zhao, J.; An, G.; Jin, W.; Li, S.; Nie, Y.; Xu, G. An Overview of Recent Advances in the Online Temperature Estimation of PMSM in Electric Vehicle Applications. *Electronics* **2026**, *15*, 1249.
5. Xu, X.; Zhu, T.; Zhao, Z.; Liu, W. Loss calculation and temperature prediction method for high-speed amorphous alloy PMSM considering the coupling effect of multi-physical field factors. *J. Magn. Magn. Mater.* **2025**, *614*, 172689.
6. Cao, L.; Fan, X.; Li, D.; Kong, W.; Qu, R.; Liu, Z. Improved LPTN-based online temperature prediction of permanent magnet machines by global parameter identification. *IEEE Trans. Ind. Electron.* **2022**, *70*, 8830–8841.
7. Li, J.; Zhao, S.; Yang, X.; Cao, J.; Gao, M. Transient Temperature Evaluation of PMSM by the Combination of LPTN and Segmented Iron Losses Model with Compensation Terms. *IEEE Trans. Transp. Electr.* **2025**, *11*, 10113–10124.
8. Hughes, R.; Haidinger, T.; Pei, X.; Vagg, C. Real-time temperature prediction of electric machines using machine learning with physically informed features. *Energy AI* **2023**, *14*, 100288.
9. Jin, L.; Mao, Y.; Wang, X.; Lu, L.; Wang, Z. A Model-Based and Data-Driven Integrated Temperature Estimation Method for PMSM. *IEEE Trans. Power Electron.* **2024**, *39*, 8553–8561.
10. Tang, P.; Zhao, Z.; Li, H. Short-Term Prediction Method of Transient Temperature Field Variation for PMSM in Electric Drive Gearbox Using Spatial-Temporal Relational Graph Convolutional Thermal Neural Network. *IEEE Trans. Ind. Electron.* **2024**, *71*, 7839–7852.
11. Liu, Z.; Kong, W.; Fan, X.; Li, Z.; Peng, K.; Qu, R. Hybrid thermal modeling with LPTN-informed neural network for multinode temperature estimation in PMSM. *IEEE Trans. Power Electron.* **2024**, *39*, 10897–10909.
12. Jing, H.; Chen, Z.; Wang, X.; Wang, X.; Ge, L.; Fang, G.; Xiao, D. Gradient boosting decision tree for rotor temperature estimation in permanent magnet synchronous motors. *IEEE Trans. Power Electron.* **2023**, *38*, 10617–10622.
13. Kirchgässner, W.; Wallscheid, O.; Böcker, J. Thermal neural networks: Lumped-parameter thermal modeling with state-space machine learning. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105537.
14. Reigosa, D.D.; Fernandez, D.; Yoshida, H.; Kato, T.; Briz, F. Permanent-Magnet Temperature Estimation in PMSMs Using Pulsating High-Frequency Current Injection. *IEEE Trans. Ind. Appl.* **2015**, *51*, 3159–3168.

15. Feng, G.; Lai, C.; Iyer, K.L.V.; Kar, N.C. Improved high-frequency voltage injection based permanent magnet temperature estimation for PMSM condition monitoring for EV applications. *IEEE Trans. Veh. Technol.* **2018**, *67*, 216–225.
16. Xiao, S.; Griffo, A. PWM-based flux linkage and rotor temperature estimations for permanent magnet synchronous machines. *IEEE Trans. Power Electron.* **2020**, *35*, 6061–6069.
17. Guo, H.; Ding, Q.; Song, Y.; Tang, H.; Wang, L.; Zhao, J. Predicting temperature of permanent magnet synchronous motor based on deep neural network. *Energies* **2020**, *13*, 4782.
18. Jung, H.S.; Oh, D.Y. Estimation of stator and magnet temperatures of IPMSM from active and reactive energies at medium and high speeds. *IEEE Trans. Transp. Electr.* **2023**, *9*, 2983–2993.
19. Hu, J.; Sun, Z.; Xin, Y.; Jia, M. Real-time prediction of rotor temperature with PMSMs for electric vehicles. *Int. J. Therm. Sci.* **2025**, *208*, 109516.
20. Feng, G.; Lai, C.; Li, W.; Li, Z.; Kar, N.C. Efficient permanent magnet temperature modeling and estimation for dual three-phase PMSM considering inverter nonlinearity. *IEEE Trans. Power Electron.* **2020**, *35*, 7328–7340.
21. Jing, H.; Chen, Z.; Ge, L.; Wang, X.; Xiao, D. Attention-based Deep Learning Method for Rotor Temperature Estimation in Permanent Magnet Synchronous Motors. *IEEE Trans. Power Electron.* **2026**, *41*, 3481–3492.
22. Chen, Y.; Xiao, Y.; Ji, C.; Su, B. An Online Torque Prediction Method of PMSM Using PINNs. *IEEE Access* **2025**, *13*, 196796–196808.
23. Sheng, Z.; Wang, D.; Fu, J.; Hu, J. A computationally efficient spatial online temperature prediction method for PM machines. *IEEE Trans. Ind. Electron.* **2022**, *69*, 10904–10914.
24. Yang, Z.; Dong, H.; Man, J.; Jia, L.; Qin, Y.; Bi, J. Online deep learning for high-speed train traction motor temperature prediction. *IEEE Trans. Transp. Electr.* **2024**, *10*, 608–622.
25. Kirchgässner, W.; Wallscheid, O.; Böcker, J. Estimating electric motor temperatures with deep residual machine learning. *IEEE Trans. Power Electron.* **2020**, *36*, 7480–7488.
26. Kirchgässner, W.; Wallscheid, O.; Böcker, J. Thermal Neural Networks: Lumped-Parameter Thermal Modeling with State-Space Machine Learning. *arXiv* **2021**. Available online: <https://arxiv.org/pdf/2103.16323v2> (accessed on 25 June 2026).
27. Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 420.
28. Wang, H.; Zhu, B.; Feng, Y.; Gao, Z.; Diao, L. A Hybrid Physical and Data-Driven Framework for Temperature Evaluation of Permanent Magnets in PMSMs. *IEEE Trans. Electr. Electron. Eng.* **2026**, *21*, 755–764.
29. Ramones, A.I.; Shi, B.; Engels, M.; Eisenbarth, M.; Andert, J. Estimating the real-time temperature of a 48V permanent magnet synchronous motor using a thermal neural network. *Energy Convers. Manag. X* **2025**, *27*, 101140.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.