*Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering) in press*
www.jzus.zju.edu.cn; www.springer.com/journal/11582
E-mail: jzus_a@zju.edu.cn

**Research Article**

https://doi.org/10.1631/jzus.A2400527

# Predicting the temperature of CRTS III ballastless tracks in cold regions based on a TCN-Track model

Jie LIANG[1,2,4], Shijie DENG[1,2]✉, Juanjuan REN[1,2], Wenlong YE[1,2,4], Kaiyao ZHANG[5], Dacheng LI[3], Ronghe ZHANG[3]

[1]*School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, China;*
[2]*State Key Laboratory of Rail Transit Vehicle System, Southwest Jiaotong University, Chengdu 610031, China*
[3]*Railway Line and Station Yard Design Department, China Railway Design Corporation, Tianjin 300308, China*
[4]*Engineering Research Center of Railway Industry of Operation Safety Assurance, Chengdu 610031, China*
[5]*Shanghai Shentong Metro Construction Group Co., Ltd, shanghai 201103, China*

**Abstract:** The uneven distribution of the temperature field in the track structure, caused by various meteorological factors such as extremely low temperatures and snowfall, leads to significant temperature loads and is the primary cause of damage to China Railway Track System (CRTS) III ballastless tracks in cold regions during service. In this study, to predict the temperature of the track structure accurately, we analyzed meteorological data collected from the Shenyang area and identified the factors that had most effect on the track temperature field. We propose a Temporal Convolutional Network-based (TCN-based) temperature field prediction model for ballastless tracks (TCN-Track), which enhances the ability to extract and fuse local and global features from complex long-term meteorological data. The results indicate that the proposed TCN-Track model performs well in predicting track temperature fields from meteorological data, with a Mean Absolute Error (MAE) ranging from 0.26 to 0.39, a Root Mean Square Error (RMSE) ranging from 0.32 to 0.50, and Correlation Coefficient (R) values ranging from 0.888 to 0.985. Compared with a Long Short-Term Memory (LSTM) model, the MAE of the TCN-Track model was reduced by 89.17% and the RMSE by 88.51%. This method offers a new solution for accurately predicting the temperature field of ballastless tracks in cold regions, aiding in predicting and preventing track damage caused by low temperatures.

**Key words:** Cold regions; CRTS III ballastless tracks; Temperature prediction; Meteorological variables; Time prediction model.

Unedited

## 1 Introduction

China's climate is complex and diverse, with cold regions covering nearly 50% of the country's land area (Ren et al., 2021). Typical cold regions include northeast China, the Qinghai-Tibet Plateau, and Xinjiang. Currently, the total operating mileage of ballastless tracks in these cold regions exceeds 8000 km, encompassing high-speed railway lines such as the Harbin-Dalian, Harbin-Qiqihar, and Shenyang-Dandong lines (Ren et al., 2022). Developed by integrating the advantages of other ballastless tracks structures, the China Railway Track System (CRTS) III ballastless tracks is used extensively in high-speed railways in cold regions, supporting speeds of 350 km/h or more (Yu et al., 2019; Zhao et al., 2022; Wang et al., 2022). The track structure comprises steel rails, fasteners, a prefabricated track slab, self-compacting concrete layers, a geotextile isolating layer, and a base plate (Fig. 1). During its service life, the CRTS III ballastless tracks in cold regions faces prolonged exposure to cold environments, resulting in issues such as cracking due to concrete shrinkage stress concentration, changes in the geometric state of the ballastless tracks structure, interlayer displacement and separation caused by upward mapping, and track surface deformation (Xie et al., 2023; Chen et al., 2024; Sun et al., 2023). The main cause of these issues is the uneven temperature distribution within the track structure due to meteorological factors, such as extremely low temperatures and snowfall, leading

✉ Shijie DENG, dkjust@my.swjtu.edu.cn

Received Nov. 19, 2024; Revision accepted Apr. 17, 2025; Crosschecked

© Zhejiang University Press 2025

to significant temperature loads. Conventional methods for monitoring the temperature-field status of track structures are costly and insufficient for providing low-temperature warnings. Therefore, it is essential to conduct temperature prediction research on CRTS III ballastless tracks in cold regions, based on meteorological factors, to ensure their durability and reliability in severely cold environments.

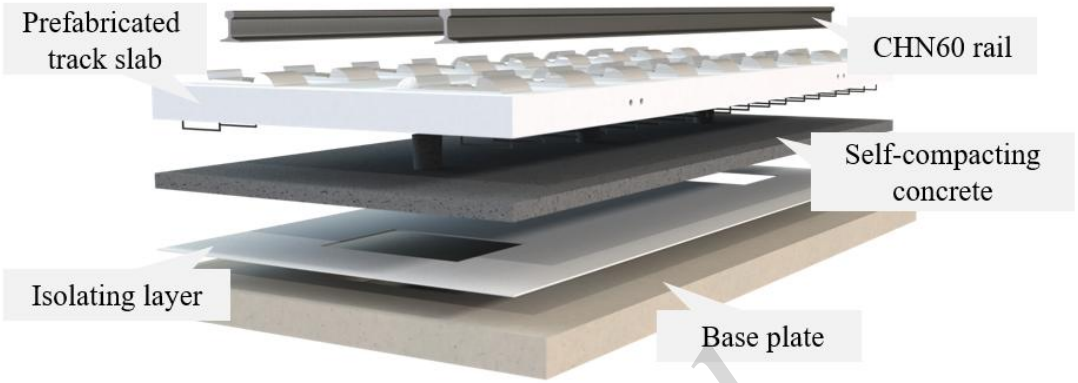



**Fig. 1 CRTS III ballastless tracks structure**

Scholars, both domestically and internationally, have conducted studies on temperature field predictions for concrete structures. Results have shown that the temperature of concrete structures is influenced by environmental conditions and material properties (Berwanger and Synko, 1976; Mirambell and Aguado, 1990; Wang et al., 2010). Analytical solutions derived through mathematical methods, such as Laplace transform and separation of variables, can quickly predict the short-term temperature distribution of concrete structures. Based on this research, Wang et al. (2023) used Bayesian hypothesis testing to establish a temperature prediction model for track structures. Ou et al. (2015; 2014) derived an analytical solution for the temperature field of ballastless tracks structures based on solid heat transfer principles and verified its accuracy with measured data. Using only a single meteorological factor can result in unstable prediction accuracy. To improve accuracy, Yang et al. (2015; 2015) studied the surface temperature of a CRTS I twin-block ballastless tracks in Chengdu through an on-site physical model test and proposed an empirical formula for the maximum surface temperature of the track bed slab in winter. However, because this method considers only two meteorological factors, there is still room for improvement. To address this, Liu et al. (2020) examined the effects of meteorological factors such as temperature, solar radiation, and wind speed on the temperature distribution of an asphalt support layer (ASL) using on-site data and numerical simulation, revealing the characteristics of the temperature distribution within the ASL. Chen (2020) furthered this research by establishing a temperature field prediction model suitable for both sunny and cloudy weather, demonstrating that comprehensive consideration of multiple meteorological factors can more accurately predict ballastless tracks temperature. Similarly, Hong et al. (2019) included the sun's position in the energy balance equation to predict the ballastless tracks temperature, underscoring the importance of considering multiple meteorological factors. Most temperature prediction models for track structures use empirical formulas and statistical analyses. While they can roughly predict temperature field trends in different regions, these methods have significant limitations in addressing the complex nonlinear relationship between meteorological data and track temperature, involving numerous thermodynamic parameters, simplifications, and assumptions. In complex environments and extreme weather conditions, prediction errors can be large.

Machine learning models can automatically capture and learn complex nonlinear relationships without explicit mathematical expressions, effectively identifying complex patterns and potential factors influencing temperature changes (Ye et al., 2024). This capability makes machine learning highly effective for temperature prediction. Kanjanun et al. (2021) and Hong et al. (2021) used meteorological parameters such as temperature, wind speed, and solar radiation as input features for machine learning models to predict ballastless tracks temperatures,

Unedited

significantly improving prediction accuracy through powerful pattern recognition. Zhou et al. (2024) analyzed the impact of different meteorological factors on temperature prediction for CRTS II ballastless tracks using artificial neural network (ANN), convolutional neural network (CNN), and long short-term memory (LSTM) models, comparing their predictive performance under various conditions. Liu et al. (2018) and Ngamkhanong et al. (2022) studied the temperature field inside a slab track using back propagation (BP) neural network and ANN models, and established temperature prediction models for track structures. These studies underscore the enormous potential of machine learning to predict ballastless tracks temperatures.

Although there have been advances in predicting the temperature field of track structures, there are still limitations in cold regions: (1) Many models consider the impact of only current meteorological data on the temperature field of track structures, ignoring the cumulative effect of past meteorological data, such as temperature, solar radiation, and snowfall. (2) Current mainstream temperature prediction models, such as BP and ANN, predict temperature only at a specific moment and do not meet engineering requirements. (3) The ballastless tracks is a multi-layered structural system, and predicting the temperature of the track surface alone cannot grasp the overall mechanical behavior. To address these deficiencies, in this study we aimed to understand the temperature state of the multi-layered structure of ballastless tracks in cold regions and implement early warnings for the ballastless tracks structure's temperature field, ensuring the safe operation of high-speed railways. We introduce a temperature prediction method for ballastless tracks in cold regions, using a temporal convolutional network-track (TCN-Track) model. TCN-Track fully accounts for the cumulative effect of meteorological data, with its robust capability to extract and integrate both local and global features of such data. By training and testing the TCN-Track model with data collected in the Shenyang area, we have achieved accurate predictions of the future temperature states of the multi-layered ballastless tracks structure. This advance offers a reliable method for early warning systems concerning temperature stress in ballastless tracks structures, thereby enhancing the safety and operational efficiency of high-speed railways in cold regions.

## 2 Meteorological and track temperature data collection

During the service life of ballastless tracks, they undergo constant alternating changes in external meteorological factors (Ren et al., 2024). Factors, such as temperature, solar radiation, and snowfall, induce temperature changes in the track structure through various heat transfer methods (Shrivastava et al., 2023; Arasu et al., 2022; Anjali et al., 2019). A long-term monitoring test plan is required to identify the key meteorological factors for predicting the temperature field of the track structure.

### 2.1 Test plan for monitoring the temperature field of a ballastless tracks

To facilitate long-term monitoring of temperature and associated meteorological parameters at key track structure locations in cold regions, it is crucial to select a test site that embodies a typical cold climate. According to the Chinese Design Code for Heating, Ventilation and Air Conditioning of Civil Buildings (GB50736-2012), regions with a coldest-month average temperature of ≤-10 °C are classified as "severe cold zones". After a comprehensive review of historical weather data and on-site investigations, Shenyang was identified as an ideal location. In Shenyang, the average temperature during the coldest month ranges from -20 to -10 °C, with annual minimum temperatures dropping to between -35 and -20 °C. This combination of factors confirmed that Shenyang is highly representative of the severe cold conditions needed for this study.

To avoid boundary effects on the scaled model, a 1:1 full-scale model of the CRTS III ballastless tracks was constructed at the test site (Fig. 2). As a unit structure, the CRTS III ballastless tracks is significantly influenced by temperature gradients, particularly by vertical temperature variations (Zhou et al., 2023; Liu et al., 2023). To ensure that temperature measurements were representative, appropriate testing locations needed to be selected. Consequently, five measuring points were established along the vertical direction at the center of the track structure (Z0, Z1, Z2, Z3 and Z4). Temperature sensors included both probe and surface-mounted types, and data were collected using a 48-channel integrated automated W2000 low-power temperature

Unedited

acquisition system. The probe type, a stable Pt1000 platinum thermistor temperature sensor, monitored temperatures within the track structure within a range of from −50 to 200 ℃. Surface-mounted temperature sensors monitored surface temperatures within the same range, with a measurement accuracy of ±0.20 ℃. Long-term meteorological data, including temperature, solar radiation, and wind speed, were collected using a Davis 6162C automated weather station. The collected meteorological data and track temperature field data provided a robust foundation for subsequent track temperature field prediction models.

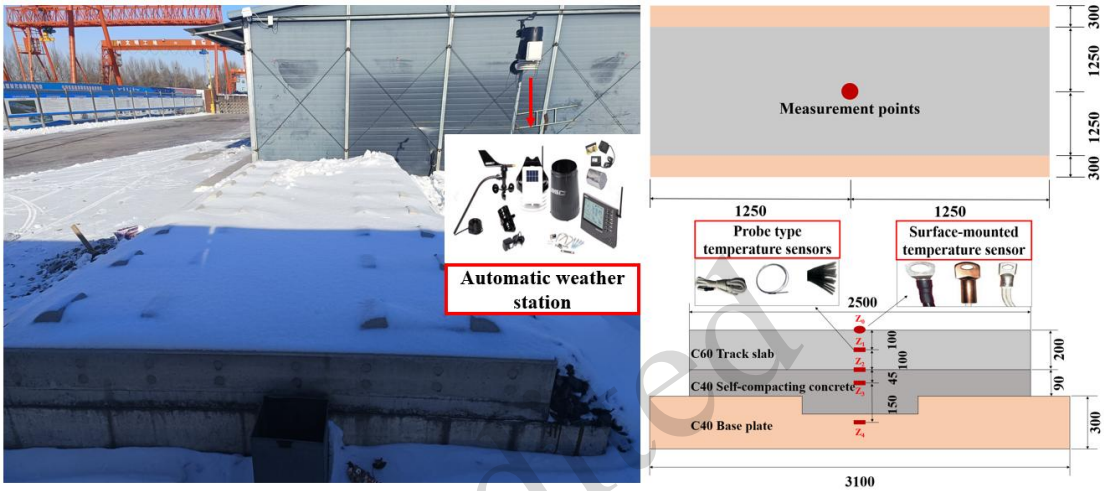



**Fig. 2 Layout of sensors for full-scale testing**

## 2.1 Test plan for monitoring the temperature field of a ballastless tracks

The monitoring period extended from November 13, 2022, to February 6, 2023, with data measured every half hour, resulting in 20,440 meteorological data points and 20,440 track structure temperature data points. This method preserves the temperature variation pattern while eliminating certain measurement errors (Xu et al., 2017). Periodic temperature changes directly affected the track structure temperature. Fig. 3 shows the temperature and track temperature data, revealing a consistent overall trend. A strong correlation between temperature and plate surface temperature is evident. Spatially, as the depth increased, the influence of temperature on track structure temperature weakened, showing nonlinear characteristics. Temporally, the occurrence of the lowest track structure temperature lagged behind the temperature, mainly because of the variable thermal conductivity coefficients of each ballastless tracks layer in cold regions. Typically, the track structure exhibits external heat and internal heat during daytime warming, and external cold and internal heat during nighttime cooling. The maximum temperature difference between the surface temperature of the track plate and the temperature [AS2.1][jl2.2]can reach 16.8 ℃ . The surface temperature of the track is often higher than the atmospheric temperature[AS3.1][jl3.2], which is attributable to the track's larger thermal capacity, which absorbs more heat, and its high thermal inertia, resulting in a lower cooling rate (Esper et al., 2024).

Uncorrected

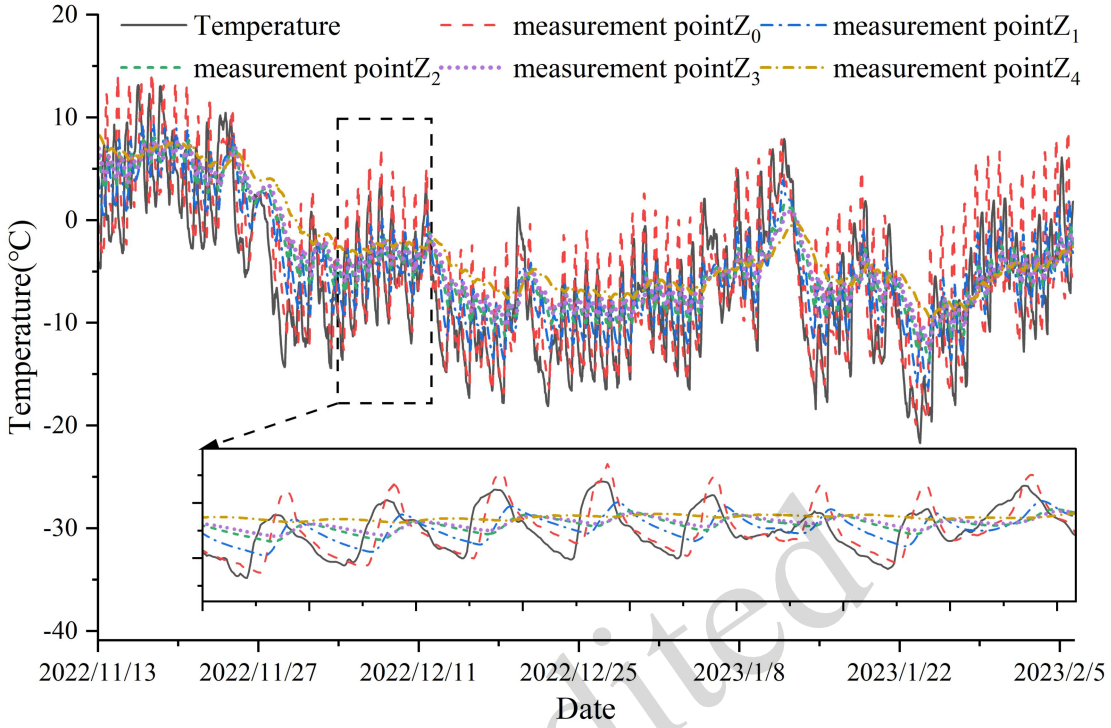



**Fig. 3 Track temperature and atmospheric temperature data**

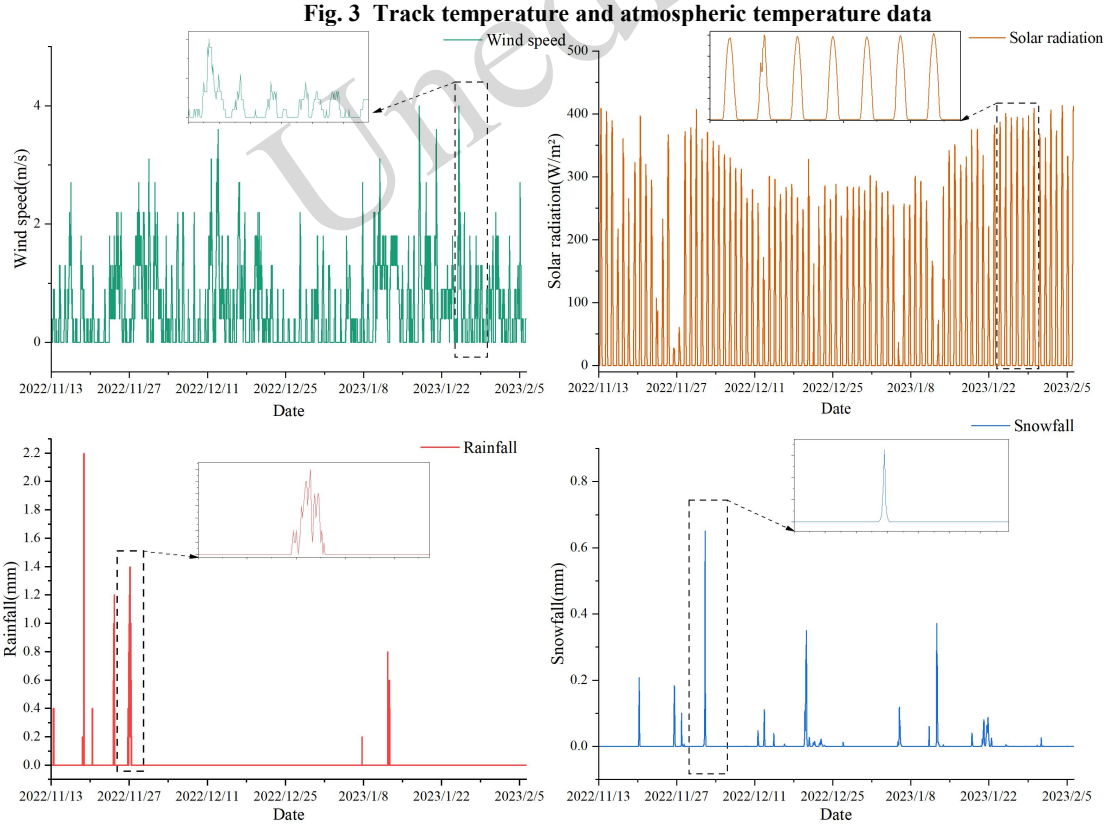



**Fig. 4 Meteorological data**

Fig. 4 presents the results identifying the main meteorological factors affecting the temperature of the track structure. Solar radiation is the main source of heat for buildings and their surroundings. In the monitoring area, the daytime solar radiation intensity was about 300 W/m² and peaked at 413 W/m². The

Unedited

daily solar radiation reached its maximum around 12:00, with less radiation in the morning and evening, concentrated mainly between 8:00 and 16:00. A higher solar radiation intensity results in more heat being absorbed by the track structure, leading to higher temperatures, which is consistent with the literature (Lou et al., 2023; Zhang et al., 2022). The graph also shows an average wind speed of 0.5 m/s, with a maximum of 4 m/s. Wind speed determines the convective heat transfer coefficient at the boundary of the track structure, which affects the heat transfer rate. However, the low wind speeds during the monitoring stage affected mainly the track plate surface. High rainfall typically causes erosion and damage to structures. However, during this monitoring period of low rainfall, its impact on the track structure temperature was the main concern. The maximum rainfall recorded was 2.2 mm, which caused rapid temperature changes in the track structure on rainy days. The maximum snowfall during this period was 0.6 mm. Snow coverage acts as insulation, reducing the heat exchange between the track structure and the environment. The high reflectivity of the snow's white surface further decreased the solar radiation heat received by the track structure.

## 3 Track temperature prediction model

In this section we introduce a temperature prediction method for ballastless tracks in cold regions based on meteorological factors. This method is based on time prediction models such as TCN and LSTM and incorporates a multi-head attention mechanism, fully utilizing the advantages of each model, such as TCN's parallel processing ability and long sequence dependency capture ability, LSTM's memory and ability to handle nonlinear relationships, and multi-head attention mechanism's global feature extraction ability, with the aiming of improving the accuracy, robustness, and generalization ability of prediction. By combining the advantages of these models, it is possible to predict the temperature changes in ballastless tracks more effectively in cold regions.

### 3.1 Network structure

Considering the accuracy and computational cost of temperature prediction, time-series models such as TCN and LSTM were optimized, leading to the development of the TCN-Track model for predicting track temperature. The overall network structure is illustrated in Fig. 5. The ballastless tracks is directly exposed to the atmospheric environment, where heat transfer occurs through three main mechanisms—conduction, radiation, and convection—as governed by the basic principles of heat transfer. Meteorological factors, including temperature, solar radiation, wind speed, rainfall, and snowfall, fundamentally influence the track's thermal behavior: temperature reflects the current thermal state; solar radiation is the main heat source warming the track; wind speed modulates convective heat exchange between the track and its surroundings; rainfall alters surface moisture and the effective thermal conductivity; and snowfall, due to its distinct thermal properties, forms an insulating layer that absorbs heat during melting.

Inconsistent data ranges and proportions between different feature values may lead to issues such as reduced training speed and decreased model performance. Therefore, it is necessary first to normalize the input meteorological data to accelerate model convergence. Subsequently, the data are fed into a TCN module consisting of multiple layers, where each layer contains dilated causal convolution, ReLU activation function, weight normalization (Weight Norm), and a Dropout mechanism to ensure stable propagation of information flow while capturing features at different time scales. After each layer of TCN output, the feature transformation layer (CWR) adjusts the number of channels through 1x1 convolution to match the input requirements of the LSTM encoding module. LSTM models long-term dependencies through memory units and gating mechanisms, while GRU uses reset and update gates to improve computational efficiency. The two work in parallel to enhance the expressive power of temporal patterns. However, because LSTM and GRU rely mainly on sequence information, it is difficult to comprehensively model global dependencies. Therefore, a multi head attention mechanism is introduced after their output, which calculates the correlation between different time steps in parallel through multiple attention heads, enabling the model to focus dynamically on key time steps and enhance feature expression through weighting and enhancement. Finally, the attention outputs of each layer are mapped to the target dimension through fully connected layers, completing the final prediction task.

Unedited

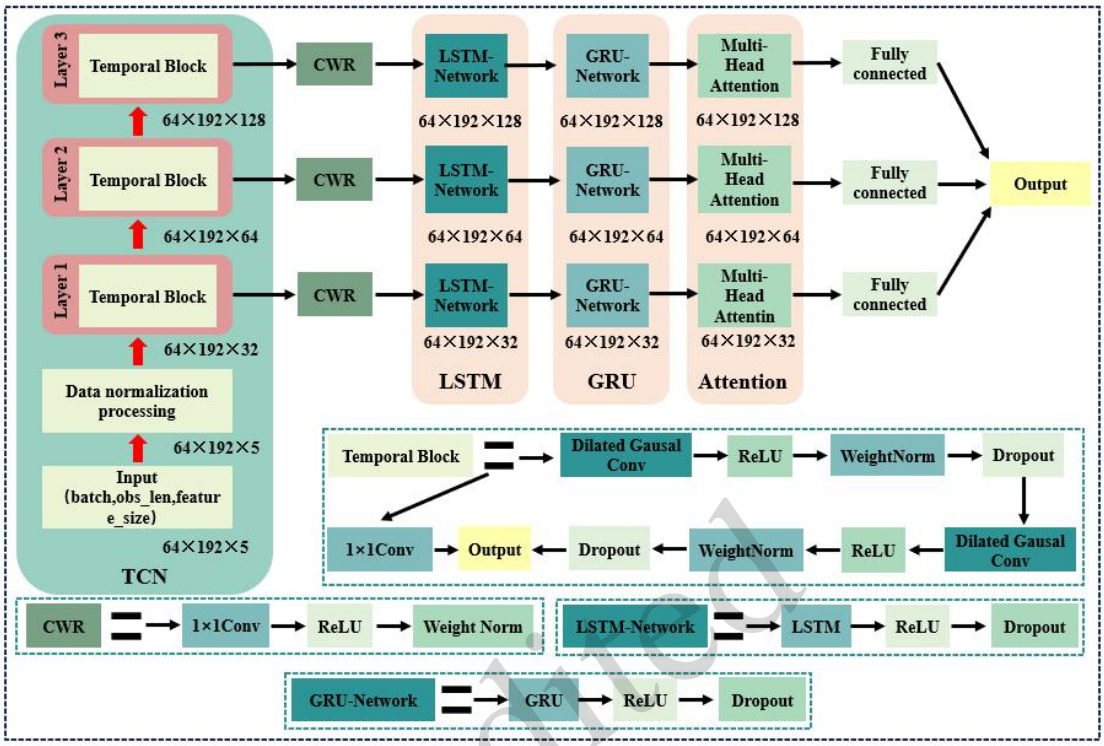



**Fig. 5 Structure of the TCN-Track model**

### 3.2 Temporal Convolutional Network

TCN is a variant of CNN (Bai et al., 2018; Hewage et al., 2020) that uses causal convolution, dilated convolution, and unbiased padding specifically for processing temporal data. The TCN structure is illustrated in Fig. 6. In TCN, causal convolution is used to ensure that the output at time step $t$ is computed based solely on data from time step $t$ and earlier, thereby preventing any influence from data at time step $t$+1 or later. This approach is grounded in the fundamental principles of CNN. The core concept of CNN is to apply convolutional kernels in a sliding-window fashion over the input data, whereby local regions are weighted and summed to extract salient features. Mathematically, this convolution operation is represented in Eq. (1), where each output $y_t$ is obtained by weighting and summing the inputs from the current time step and the preceding $t$-1 time steps. The weight-sharing mechanism inherent in CNN significantly reduces the number of parameters and imparts translational invariance, enabling the network to recognize similar local patterns regardless of their position within the input sequence. Additionally, CNN uses a local receptive field design, which confines the focus to small segments of the input data, thereby effectively capturing local correlations. As these local features are propagated through multiple convolutional layers, they are gradually integrated into higher-level abstract representations. This hierarchical feature extraction not only enhances the model's ability to represent complex patterns but also facilitates parallel computation, thereby greatly improving computational efficiency.

Zero-padding was used to fill the ends of the data, preventing discrepancies in input and output sizes caused by the coverage range of the convolution kernel. This method ensures temporal alignment and preserves edge information. The number of zeros filled in each layer was calculated using Eq. (2). Mainstream gated recurrent neural networks often require computing data at time step $t$−1 and before calculating data at time step $t$, leading to inefficiency. Although traditional CNNs can perform efficient parallel computing, they focus mainly on local data features, making it challenging to extract key features from longer time-series data. Dilation convolution addresses these issues by setting the dilation rate to skip specific step sizes during convolution operations while maintaining the convolution kernel size. This

Uncorrected

approach expands the receptive field of the TCN without increasing the number of parameters, as shown in Eq. (3). Capturing dependencies in longer time-series typically requires larger convolutional kernels and more layers to expand the receptive field, which can increase the risk of overfitting and degrade model performance. To counter this, the TCN network incorporates a residual module to prevent gradient vanishing and uses Dropout regularization to reduce the risk of overfitting.

$$y_t = \sum_{i=0}^{n-1} k_i \cdot x_{t-i} \tag{1}$$

$$P = (n-1) \cdot d \tag{2}$$

$$y_t = \sum_{i=0}^{n-1} k_i \cdot x_{t-d \cdot i} \tag{3}$$

where $y_t$ denotes output at time step $t$, $k_i$ denotes the $i$-th convolutional kernel, $n$ denotes the convolution kernel size, $x_t$ denotes the input sequence at time step $t$, $P$ denotes the number of zeros filled in, and $d$ denotes the expansion rate.

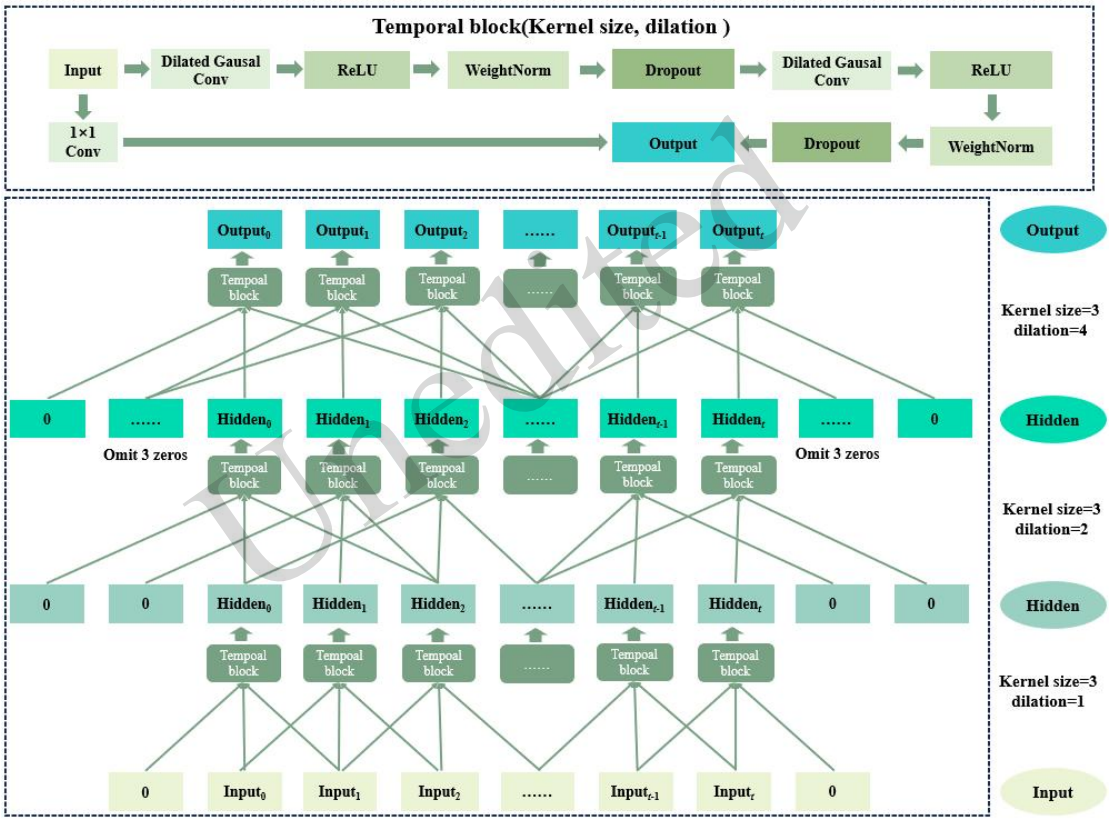



**Fig. 6 Structure of the TCN model.**

### 3.3 Gated Recurrent Neural Networks

The LSTM is a variant of the recurrent neural network (RNN) that incorporates a forget gate, input gate, and output gate to enhance its functionality (Liu et al., 2018; Ghimire et al., 2023). LSTM retains the RNN's capability to process temporal data while better capturing long-term dependencies, avoiding issues of gradient explosion or vanishing. As shown in Fig. 7 and Eq. (4), LSTM's information flow is controlled mainly by three memory factors: the forget memory factor $f_t$, the input memory factor $i_t$, and the output memory factor $O_t$. These factors were generated by applying a Sigmoid function to $X_t$ and $h_{t-1}$, differing only in their weight parameters. Specifically, $f_t$ determines which information from the previous time step will be retained, $i_t$ determines whether candidate memory cells generated through tanh operations will update the historical information, and dictates which information will be output. The values of these memory factors ranged from 0 to 1, with 0 indicating complete abandonment and 1

indicating full retention.

$$
\begin{aligned}
f_t &= \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \\
i_t &= \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \\
\tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, X_t] + b_C) \\
C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \\
O_t &= \sigma(W_O \cdot [h_{t-1}, X_t] + b_O) \\
h_t &= O_t \cdot \tanh(C_t)
\end{aligned} \tag{4}
$$

where $f_t$ denotes the forgetting memory factor, ranging from 0 to 1, $i_t$ denotes the input memory factor at time step $t$, ranging from 0 to 1, $\tilde{C}_t$ denotes candidate memory cells at time step $t$, $C_t$ denotes memory cells at time step $t$, $O_t$ denotes the output memory factor at time step $t$, ranging from 0 to 1, $h_{t-1}$ denotes the output at time step $t-1$, $W_f$, $W_i$, $W_C$, $W_O$ denotes the weight, $b_f$, $b_i$, $b_C$, $b_O$ denote bias terms, $X_t$ denotes input at time step $t$, $h_t$ denotes output at time step $t$, and $\sigma$ denotes the sigmoid function.

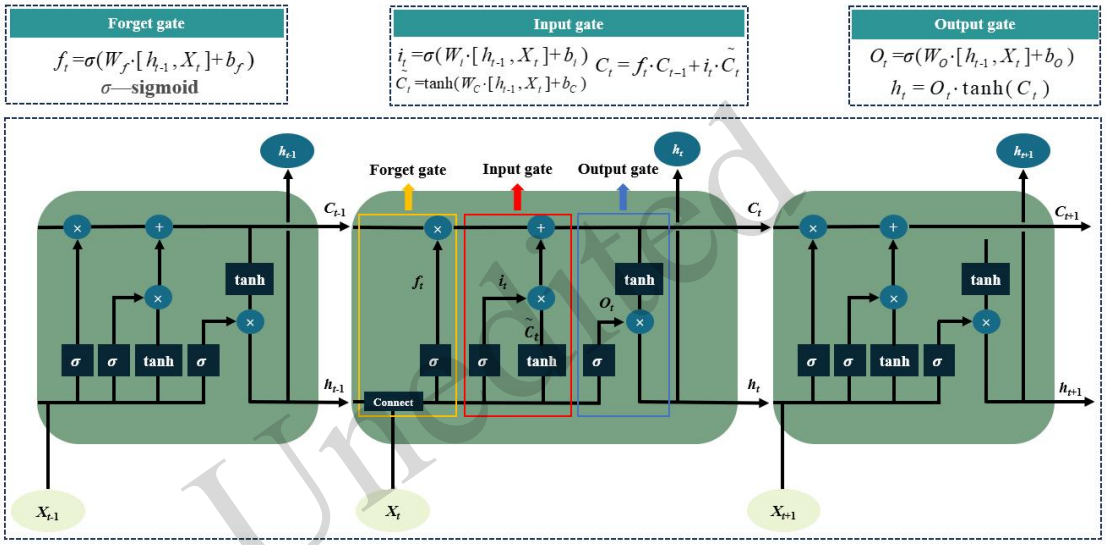



**Fig. 7 Structure of the LSTM model.**

Gated recurrent units (GRU) is a variant of RNN, with a network structure as shown in Fig. 8. Similar to LSTM, GRU uses a gating mechanism to control information flow, but it has fewer parameters, a simpler structure, and higher computational efficiency. GRU's gates are the reset gate and the update gate. The forget gate and input gate in LSTM are combined into an update gate, and GRU lacks independent memory cells. The calculation formula is shown in Eq. (5). Based on Fig. 8 and Eq. (5), the reset gate determines which input information combines with the previous output, while the update gate controls the input information.

$$
\begin{aligned}
Z_t &= \sigma(W_Z \cdot [h_{t-1}, X_t] + b_Z) \\
R_t &= \sigma(W_R \cdot [h_{t-1}, X_t] + b_R) \\
\tilde{H}_t &= \tanh(W_H \cdot [R_t \cdot h_{t-1}, X_t] + b_H) \\
H_t &= \tilde{H}_t \cdot Z_t + H_{t-1} \cdot (1 - Z_t)
\end{aligned} \tag{5}
$$

where $Z_t$ denotes the update gate weight coefficient at time step $t$, $R_t$ denotes the reset gate weight coefficient at time step $t$, $\tilde{H}_t$ denotes the candidate state at time step $t$, $H_t$ denotes the output at time step $t$, $W_Z$, $W_R$, $W_H$ denote the weight, and $b_Z$, $b_R$, $b_H$ denote the bias term.

Unedited

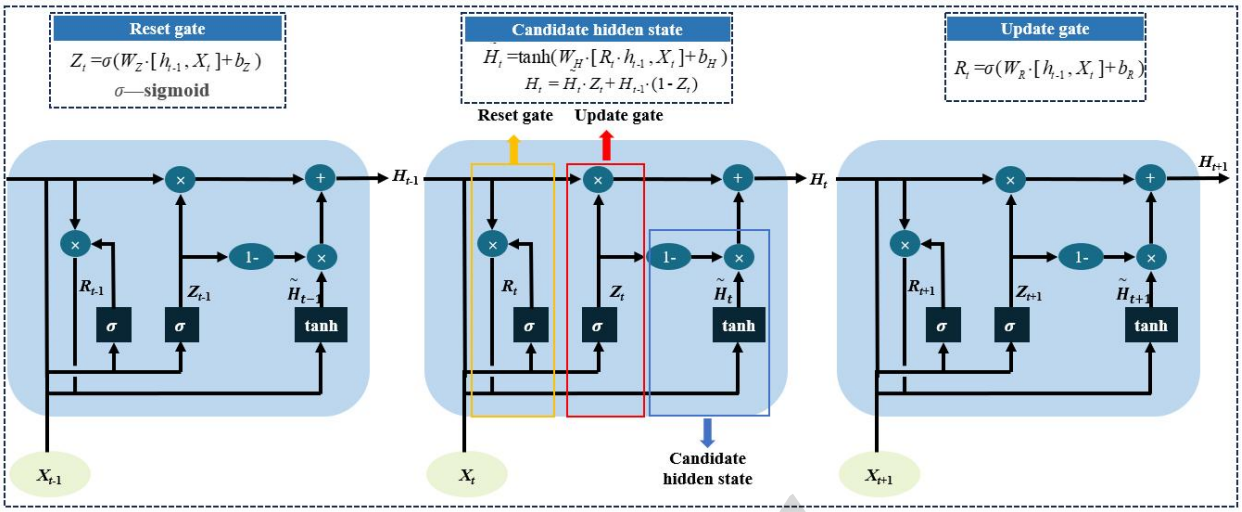



**Fig. 8 Structure of the GRU model.**

### 3.4 Multi-Head Attention

The attention mechanism assigns different weights to the input sequence, allowing the model to focus on significant parts while reducing the effect of less important parts, greatly enhancing the model's global feature extraction ability. The multi-head attention mechanism comprises multiple independent self-attention mechanisms, which better capture potential relationships between features from various perspectives. Its network structure is depicted in Fig. 9.

The input matrix is first transformed linearly to produce three matrices: queries ($Q$), keys ($K$), and values ($V$). These matrices capture different aspects of the input features by projecting them into distinct subspaces. To enable the model to attend to various aspects of the input from different representation subspaces, $Q$, $K$, and $V$ are then split along their last dimensions into h smaller matrices, yielding h separate sets of ($Q_i$, $K_i$, $V_i$). for $i$ = 1, 2, …, h. For each head, the self-attention mechanism computes an attention score matrix using the scaled dot-product operation (Eq. (6)). After processing through each head independently, the outputs of all h heads are concatenated along the feature dimension. This concatenated matrix is then passed through a final linear transformation, which integrates the diverse information captured by the different heads into a unified representation.

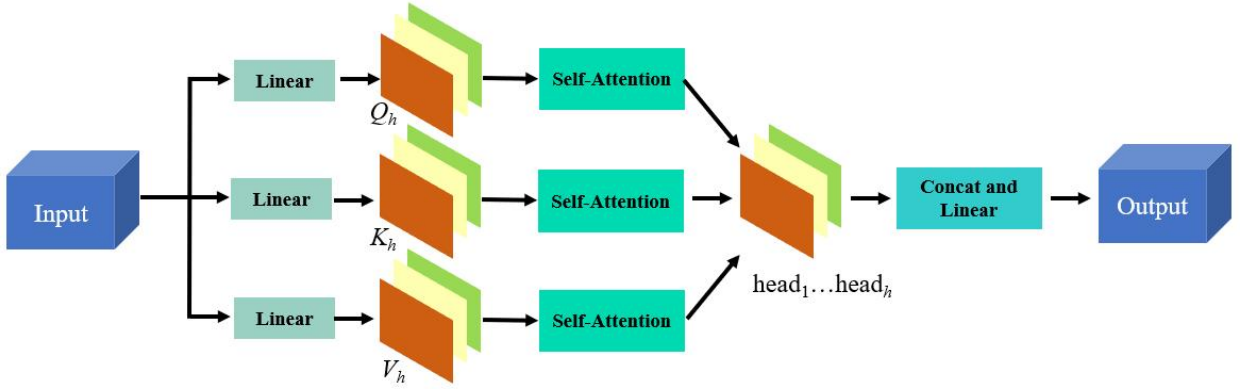



**Fig. 9 Structure of the Multi-Head Attention model.**

$$head_h=\text{Attention}(Q_h,K_h,V_h)=\text{softmax}\left(\frac{Q_hK_h^T}{\sqrt{d_k/h}}\right)V_h \quad (6)$$

where $Q_h$ denotes the $h$-th query matrix, $K_h$ denotes the $h$-th key matrix, $V_h$ denotes the $h$-th value matrix, $d_k$ denotes the dimension of the key matrix, and $head_h$ denotes the output of the $h$-th head.

Unedited

## 4 Prediction results and analysis of the temperature field of the ballastless tracks

In this section we introduce the evaluation metrics for the track temperature prediction model, the model parameter settings, and a comparison of the model prediction results.

### 4.1 Evaluation Metrics

We evaluated the model's performance using three metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Correlation Coefficient (R). The RMSE, which is sensitive to larger errors, reflects the model's prediction accuracy in extreme weather conditions in cold regions but is easily affected by outliers. The calculation formula is shown in Eq. (7). The MAE treats each error equally, providing a simple and intuitive measure. The calculation formula is shown in Eq. (8). R evaluates the linear correlation between predicted and true values. The calculation formula is shown in Eq. (9). These metrics comprehensively reflect the performance of the track temperature field prediction model.

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2} \tag{7}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \tag{8}$$

$$\mathrm{R}=\frac{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)\left(\hat{y}_i-\bar{\hat{y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2\sum_{i=1}^{n}\left(\hat{y}_i-\bar{\hat{y}}\right)^2}} \tag{9}$$

where $y_i$ denotes the *i*-th predicted value, $\hat{y}_i$ denotes the *i*-th true value, n denotes the total number of samples, $\bar{y}$ denotes the mean of predicted values, and $\bar{\hat{y}}$ denotes the mean of the true value.

### 4.2 Model parameters

The TCN-Track model code, developed using the PyTorch deep learning framework and Python, runs on an i5-12490F CPU, NVIDIA GeForce RTX 4060 Ti GPU, and Windows 10 operating system. The input feature matrix size is 64 × 192 × 5, where 64 represents the batch size, 192 is the time window size, and 5 is the number of features. This setup predicts the track temperature for a particular day based on meteorological data from the previous four days. The initial learning rate is set to 0.001, and the optimizer used is AdamW, an Adam optimization algorithm variant that achieves regularization effects without additional weight decay parameters. The main activation functions are ReLU and tanh. Relevant literature (Lea et al., 2017; Bai et al., 2018) indicates that TCN shows minimal performance improvement with increasing layers and carries a risk of overfitting. Thus, to balance prediction accuracy and computational time, TCN was set to three layers. Dividing the dataset into training and validation sets effectively prevents overfitting and ensures model robustness. In this study, 80% of the data were used for training and the remaining 20% for validation.

Selecting the appropriate loss function is crucial for improving model performance. Common loss functions for temperature prediction include MSE, MAE, MAPE, and Huber Loss (Sadhukhan et al., 2022; Yang et al., 2022). MSE is sensitive to outliers, while MAE is robust and less affected by large errors. MAPE's high sensitivity to values close to zero makes it unsuitable for temperature prediction of ballastless tracks in cold regions. Huber Loss combines the advantages of MSE and MAE, using MAE to accelerate loss reduction when errors are large and MSE to maintain sensitivity when errors are small. Thus, we selected Huber Loss as the loss function, with the calculation formula shown in Eq. (10), where $\delta$ is the hyperparameter set manually.

$$\mathrm{Huber\ Loss}\left(y,\ \hat{y}\right)=\begin{cases}\frac{1}{2}\left(y-\hat{y}\right)^2 & \left|y-\hat{y}\right|\le\delta\\ \delta\left|y-\hat{y}\right|-\frac{1}{2}\delta^2 & \text{otherwise}\end{cases} \tag{10}$$

### 4.3 Prediction results

To investigate the impact of different epochs on the performance of track temperature field prediction models, epochs were set to 100, 150, 200, 250, and 300. The optimal results are highlighted in bold in Table 1. The results show that measuring points $Z_0$, $Z_1$, and $Z_2$ achieved the best prediction outcomes at 250 epochs, while points $Z_3$ and $Z_4$ reached optimal results at 150 and 200 epochs, respectively. Further increases in epochs lead to minimal changes or even deterioration in prediction results, indicating a risk of overfitting.

As the depth of the track structure increases, temperature changes become more pronounced, enabling the model to capture these variations more accurately. Consequently, deeper track structures

Unedited

yield more precise temperature predictions and require fewer epochs. At 250 epochs, the prediction results for measuring points $Z_3$ and $Z_4$ deteriorate slightly but remain within an acceptable range. Therefore, in this study we set the epochs to 250 for the subsequent tests.

**Table 1 Prediction results of different epochs**

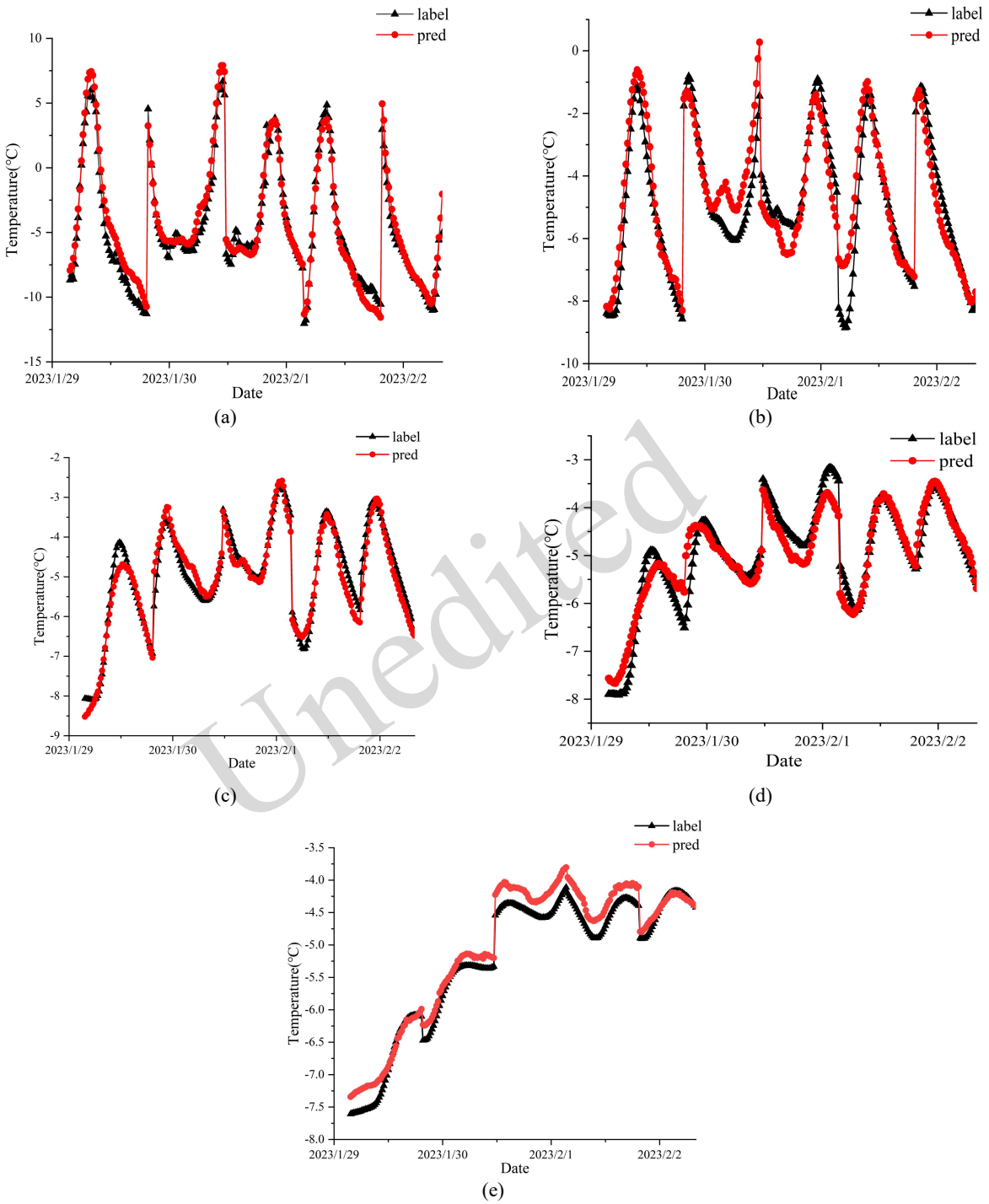
| Measurement point | Metrics | Epochs | | | | |
|---|---|---|---|---|---|---|
| | | 100 | 150 | 200 | 250 | 300 |
| $Z_0$ | RMSE | 6.13 | 2.73 | 0.84 | **0.77** | 0.81 |
| | MAE | 5.69 | 2.08 | 0.69 | **0.62** | 0.62 |
| | R | 0.795 | 0.836 | 0.944 | **0.980** | 0.974 |
| $Z_1$ | RMSE | 5.32 | 0.89 | 0.69 | **0.64** | 0.64 |
| | MAE | 5.25 | 0.80 | 0.64 | **0.53** | 0.58 |
| | R | 0.807 | 0.939 | 0.944 | **0.961** | 0.942 |
| $Z_2$ | RMSE | 3.89 | 1.18 | 0.62 | **0.62** | 0.79 |
| | MAE | 3.27 | 0.93 | 0.49 | **0.49** | 0.60 |
| | R | 0.811 | 0.898 | 0.923 | **0.934** | 0.900 |
| $Z_3$ | RMSE | 0.69 | 0.62 | **0.51** | 0.51 | 0.66 |
| | MAE | 0.56 | 0.49 | **0.39** | 0.42 | 0.53 |
| | R | 0.872 | 0.881 | **0.888** | 0.884 | 0.875 |
| $Z_4$ | RMSE | 0.36 | **0.32** | 0.39 | 0.43 | 0.86 |
| | MAE | 0.36 | **0.32** | 0.36 | 0.36 | 0.80 |
| | R | 0.829 | **0.839** | 0.834 | 0.830 | 0.807 |

The size of the convolutional kernel directly affects the model's receptive field, feature capture ability, and computational efficiency. A kernel that is too small may cause the model to overly focus on local features, hindering its ability to capture global information. Conversely, a kernel that is too large can significantly increase the parameter count and the risk of overfitting. Therefore, determining the optimal convolution kernel size is crucial. As shown in Table 2, the overall performance is optimal when the convolution kernel size is 9 and further increasing the kernel size decreases the model's performance. The test data encompassed five days from January 29, 2023, to February 2, 2023. Fig. 10 displays the prediction results for measurement points Z0, Z1, Z2, Z3, and Z4, respectively, showing good overall prediction accuracy for the track structure temperature. For subsequent model comparisons, epochs were set to 250, and the kernel size to 9.

Unedited

**Table 2 Prediction results of different convolutional kernel sizes**

| Measurement point | Metrics | Kernel size | | | | |
|---|---|---|---|---|---|---|
| | | 3 | 6 | 9 | 12 | 15 |
| $Z_0$ | MAE | 0.62 | 0.49 | **0.39** | 0.48 | 0.46 |
| | RMSE | 0.77 | 0.61 | **0.50** | 0.59 | 0.57 |
| | R | 0.965 | 0.978 | **0.985** | 0.984 | 0.984 |
| $Z_1$ | MAE | 0.53 | 0.45 | **0.38** | 0.48 | 0.45 |
| | RMSE | 0.64 | 0.55 | **0.47** | 0.58 | 0.53 |
| | R | 0.961 | 0.968 | 0.974 | 0.981 | **0.982** |
| $Z_2$ | MAE | 0.49 | 0.46 | **0.37** | 0.39 | 0.39 |
| | RMSE | 0.62 | 0.56 | **0.45** | 0.47 | 0.46 |
| | R | 0.934 | 0.910 | 0.931 | 0.926 | **0.955** |
| $Z_3$ | MAE | 0.42 | 0.38 | **0.34** | 0.39 | 0.39 |
| | RMSE | 0.51 | 0.47 | **0.42** | 0.48 | 0.48 |
| | R | 0.884 | **0.918** | 0.895 | 0.839 | 0.878 |
| $Z_4$ | MAE | 0.36 | 0.30 | **0.26** | 0.25 | 0.30 |
| | RMSE | 0.43 | 0.34 | **0.32** | 0.31 | 0.36 |
| | R | 0.830 | 0.837 | **0.888** | 0.861 | 0.850 |




Unedited

**Fig. 10 Temperature prediction results of different measuring points: (a) Measurement point position Z0; (b) Measurement point position Z1; (c) Measurement point position Z2; (d) Measurement point position Z3; (e) Measurement point position Z4**

To demonstrate the superiority of the TCN-Track model, it was compared with six other time prediction models and evaluated comprehensively using RMSE, MAE, and R as indicators. The results are shown in Table 3. For the GRU, LSTM, and GRU-LSTM models, epochs were set to 400, and the convolution kernel size was 3. For the TCN, TCN-LSTM, and TCN-GRU models, epochs were set to 400 with a convolution kernel size of 9. The correlation coefficient of the model in this study is lower only

than that of the TCN-LSTM model at measurement point $Z_2$, but it still reaches 0.931. The indicators for other measurement points are significantly better than those of other models. Compared to other models, the TCN-Track model showed a decrease in MAE of 3.61% to 89.17%, a decrease in RMSE of 6.67% to 88.51%, and an increase in R of 1.45% to 29.57%.

**Table 3 Model comparison**

| Algorithm | Metrics | Measurement point | | | | |
|---|---|---|---|---|---|---|
| | | $Z_0$ | $Z_1$ | $Z_2$ | $Z_3$ | $Z_4$ |
| GRU | MAE | 3.51 | 2.46 | 1.48 | 1.08 | 1.00 |
| | RMSE | 4.12 | 2.92 | 1.74 | 1.32 | 1.08 |
| | R | 0.77 | 0.851 | 0.86 | 0.722 | 0.767 |
| TCN | MAE | 2.49 | 1.81 | 1.50 | 1.34 | 1.31 |
| | RMSE | 3.20 | 2.20 | 1.69 | 1.49 | 1.35 |
| | R | 0.782 | 0.686 | 0.872 | 0.700 | 0.704 |
| LSTM | MAE | 3.60 | 1.56 | 1.29 | 1.59 | 1.15 |
| | RMSE | 4.35 | 1.96 | 1.56 | 1.79 | 1.31 |
| | R | 0.833 | 0.732 | 0.821 | 0.710 | 0.704 |
| TCN-LSTM | MAE | 1.80 | 1.12 | 0.90 | 0.35 | 0.28 |
| | RMSE | 2.14 | 1.34 | 1.00 | 0.45 | 0.35 |
| | R | 0.922 | 0.914 | **0.962** | 0.826 | 0.866 |
| TCN-GRU | MAE | 1.44 | 1.36 | 0.91 | 0.67 | 0.58 |
| | RMSE | 1.82 | 1.71 | 1.10 | 0.82 | 0.65 |
| | R | 0.938 | 0.937 | 0.904 | 0.882 | 0.821 |
| GRU-LSTM | MAE | 2.29 | 1.71 | 1.32 | 1.49 | 0.73 |
| | RMSE | 2.91 | 2.05 | 1.53 | 1.64 | 0.81 |
| | R | 0.901 | 0.911 | 0.908 | 0.832 | 0.798 |
| TCN-Track | MAE | **0.39** | **0.38** | **0.37** | **0.34** | **0.26** |
| | RMSE | **0.50** | **0.47** | **0.45** | **0.42** | **0.32** |
| | R | **0.985** | **0.974** | 0.931 | **0.895** | **0.888** |

## 5 Conclusions

In this article we present an in-depth investigation into the prediction of track structure temperature fields. We introduce a method for predicting temperature fields for CRTS III ballastless tracks in cold regions, based on on-site measurement data from the Shenyang area. We offer innovative ideas and techniques for predicting the temperature of ballastless tracks on high-speed railways in cold regions. The specific contributions are as follows:

(1) The on-site measured data from the Shenyang area were systematically analyzed across time and space dimensions, exploring the effects of temperature, solar radiation, wind speed, rainfall, and snowfall on the track temperature field.

(2) The TCN-Track model incorporates the cumulative effects of meteorological factors and improves feature extraction by merging TCN, LSTM, and GRU models with a multi-head attention mechanism, achieving precise temperature predictions for multi-layer track structures. The model predicts the daily temperature for each layer of the ballastless tracks structure with an MAE between 0.26 and 0.39, an RMSE between 0.32 and 0.50, and an R between 0.888 and 0.985.

(3) Compared to models like TCN, LSTM, and GRU for predicting track temperature fields, the TCN-Track model reduced the MAE by 3.61% to 89.17%, reduced the RMSE by 6.67% to 88.51%, and increased R by 1.45% to 29.57%.

## Acknowledgments

We gratefully acknowledge support from the National Natural Science Foundation of China (No. 52278461, 52308467, NO. 52425213). The results and opinions presented are those of the authors and do not necessarily reflect those of the sponsoring agencies.

## Author contributions

Jie LIANG, Wenlong YE, Kaiyao ZHANG designed the research and processed the corresponding data. Jie LIANG wrote the first draft of the manuscript. Shijie DENG and Juanjuan REN provided financial support. Dacheng LI helped to organize the manuscript. Ronghe ZHANG revised and edited the final version.

## Conflict of interest

Jie LIANG, Shijie DENG, Juanjuan REN, Wenlong YE, Kaiyao ZHANG, Dacheng LI, Ronghe ZHANG declare that

Unedited

they have no conflict of interest.

## References


Anjali T., Chandini K., Anoop K., et al., 2019. Temperature prediction using machine learning approaches. 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), p.1264–1268.

Arasu A I, Modani M, Vadlamani N R, 2022. Application of machine learning techniques in temperature forecast. 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA), p.513–518.

Bai S, Kolter J Z, Koltun V, 2018. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271, 10.

Berwanger C, Symko Y, 1976. Finite-element solutions for thermal stresses in steel-concrete composite bridges: Stresses produced by steady-state temperatures are computed. Analyses using the finite-element method are compared to the experimental results. *Experimental Mechanics*, 16: 168-175.

Chen W, Li S, Wang W, et al., 2024. Analysis on crack propagation of CRTS III slab ballastless track under temperature loads and freeze–thaw deterioration. *Theoretical and Applied Fracture Mechanics*, 129: 104206.

Chen W, Li S, Wang W, et al, 2024. Analysis on crack propagation of CRTS III slab ballastless track under temperature loads and freeze–thaw deterioration. Theoretical and Applied Fracture Mechanics, 129: 104206.

Chong W, Tramontini R, Pivoto Specht L, 2009. Application of the Laplace transform and its numerical inversion to temperature profile of a two-layer pavement under site conditions. *Numerical Heat Transfer, Part A, Applications*, 5(11): 1004-1018.

Esper J, Torbenson M, Büntgen U, 2024. 2023 summer warmth unparalleled over the past 2,000 years. *Nature*, 1-2.

Ghimire S, Nguyen-Huy T, AL-Musaylh M S, et al., 2023. Integrated Multi-Head Self-Attention Transformer model for electricity demand prediction incorporating local climate variables. *Energy and AI*, 14: 100302.

Hewage P, Behera A, Trovati M, et al., 2020. Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station. *Soft Computing*, 24: 16453-16482.

Hong S, Park C, Cho S, 2021. A rail-temperature-prediction model based on machine learning: warning of train-speed restrictions using weather forecasting. *Sensors*, 21(13): 4606.

Hong S U, Kim H U, Lim N H, et al., 2019. A rail-temperature-prediction model considering meteorological conditions and the position of the sun. *International Journal of Precision Engineering and Manufacturing*, 20: 337-346.

Kanjanun K, Bin Y, Shuang'ao Y, et al., 2021. GRNN prediction model for temperature-induced deformation of CRTS II unballasted slab track. *Applied Science and Engineering Progress*, 15(4):5662.

Lea C, Flynn M D, Vidal R, et al., 2017. Temporal convolutional networks for action segmentation and detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p.156–165.

Liu H M, Lu H Y, He Y L, et al., 2018. Neural network based temperature field mapping model for CRTS II type ballastless track. IOP Conference Series: Materials Science and Engineering, p.012042.

Liu J, Zhang T, Han G, et al., 2018. TD-LSTM: Temporal dependence-based LSTM networks for marine temperature prediction. *Sensors*, 18(11): 3797.

Liu R, Song L, Cao J, et al., 2023. Early warping deformation of high-speed railway CRTS III track slabs. *Engineering Structures*, 294: 116651.

Liu S, Chen X, Yang J, et al., 2020. Numerical study and in-situ measurement of temperature features of asphalt supporting layer in slab track system. *Construction and Building Materials*, 233: 117343.

Lou P, Shi T, 2023. Thermal arching and interfacial damage evolution of CRTS-II slab track under solar radiation in alpine and plateau regions. *Alexandria Engineering Journal*, 74: 301-315.

Ministry of Housing and Urban-Rural Development of the People's Republic of China, 2012. Design code for heating ventilation and air conditioning of civil buildings, GB 50736-2012. National Standards of the People's Republic of China.

Mirambell E, Aguado A, 1990. Temperature and stress distributions in concrete box girder bridges. *Journal of Structural Engineering*, 116(9): 2388-2409.

Ngamkhanong C, Kaewunruen S, 2022. Prediction of thermal-induced buckling failures of ballasted railway tracks using artificial neural network (ANN). *International Journal of Structural Stability and Dynamics*, 22(05): 2250049.

Ou Z, Lu S, Qun C, 2015. Simplified calculation of temperature field in high-speed railway ballastless track structure. *Journal of Zhejiang University (Engineering Science)*, 49(3): 482-487.

Ren J, Deng S, Zhang K, et al., 2021. Design theories and maintenance technologies of slab tracks for high-speed railways in China: a review. *Transportation Safety and Environment*, 3(4): tdab024.

Ren J, Zhang S, Xu X, et al., 2022. Progress of research on service performance evolution and improvement technology of ballastless track in cold regions. *China Railway*, 2022(08): 76-87.

Sadhukhan B, Chakraborty S, Mukherjee S, 2022. Investigating the relationship between earthquake occurrences and climate change using RNN-based deep learning approach. *Arabian Journal of Geosciences*, 15(1): 31.

Shrivastava V K, Shrivastava A, Sharma N, et al., 2023. Deep learning model for temperature prediction: an empirical


Unedited

study. *Modeling Earth Systems and Environment*, 9(2): 2067-2080.

Sun Z, Lou P, Huang X, 2023. Temperature gradient of ballastless track in large daily temperature difference region and its influence on dynamic responses of vehicle-track-bridge system. *Alexandria Engineering Journal*, 85: 114-131.

Wang D, 2010. Analytical solutions for temperature profile prediction in multi-layered pavement systems. PhD Thesis, University of Illinois at Urbana-Champaign, USA.

Wang J, Gao L, Zhao W, et al., 2022. Evolution mechanism of interlayer fatigue properties of CRTS III slab track. *Construction and Building Materials*, 360: 129459.

Wang S, Gao J, Lin C, et al., 2023. Condition assessment of high-speed railway track structure based on sparse Bayesian extreme learning machine and Bayesian hypothesis testing. *International Journal of Rail Transportation*, 11(3): 364-388.

Xu B, Dan H C, Li L, 2017. Temperature prediction model of asphalt pavement in cold regions based on an improved BP neural network. *Applied Thermal Engineering*, 120: 568-582.

Yang R, Wan Z, Liu X, et al., 2015. Temperature Field Test of CRTS I twin-Block Ballastless Track in Winter. *Journal of Southwest Jiaotong University*, 50(3): 454-460.

Yang Y, Zhou H, Wu J, et al., 2022. Robustified extreme learning machine regression with applications in outlier-blended wind-speed forecasting. *Applied Soft Computing*, 122: 108814.

Ye W, Ren J, Li C, et al., 2024. Intelligent detection of surface defects in high-speed railway ballastless track based on self-attention and transfer learning. *Structural Control and Health Monitoring*, 2024(1): 2967927.

Yu Z, Xie Y, Tian X, 2019. Research on mechanical performance of CRTS III plate-type ballastless track structure under temperature load based on probability statistics. *Advances in Civil Engineering*, 2019(1): 2975274.

Zhang Q, Dai G, Tang Y, 2022. Thermal analysis and prediction methods for temperature distribution of slab track using meteorological data. *Sensors*, 22(17): 6345.

Zhao L, Che Y, 2022. Numerical simulation of temperature field of CRTS III ballastless track with Tarmac in Harbin area. *Highlights in Science, Engineering and Technology*, 28: 360-366.

Zhou L, Zhang T, Luo Y, 2023. Insights of the vehicle-track-girder system dynamic response changes caused by the thermal deformation of CRTS III ballastless track in the natural environment. *Construction and Building Materials*, 400: 132745.

Zhou R, He M, Xu H, et al., 2024. Comparing various neural network methods for temperature prediction of CRTS II slab track on transition sections. *Intelligent Transportation Infrastructure*, liae004.

Zumin O, Fujian L, 2014. Analysis and prediction of the temperature field based on in-situ measured temperature for CRTS-II ballastless track. *Energy Procedia*, 61: 1290-1293.

Ren J, Qu C, Liang J, et al., 2024. Experimental study on temperature field characteristics of CRTS III prefabricated slab track in cold regions. *Archives of Civil and Mechanical Engineering*, 24(3).

## 中文概要

**题　目：基于 TCN-Track 模型的寒区 CRTS III 无砟轨道温度场预测研究**

**作　者：**梁捷[1,2,4]，邓世杰[1,2]，任娟娟[1,2]，叶文龙[1,2,4]，章恺尧[5]，李大成[3]，张荣鹤[3]

**机　构：**[1]西南交通大学，土木工程学院，中国成都，610031；[2]西南交通大学，轨道交通运载系统全国重点实验室，中国成都，610031；[3]中国铁路设计集团有限公司，中国天津，300308；[4]运营安全保障铁路行业工程研究中心，中国成都，610031；[5]上海申通地铁建设集团有限公司，中国上海，201103

**目　的：**针对寒区 CRTS III 无砟轨道因气象要素时空异质性导致的温度场非均匀分布问题，本文旨在通过分析沈阳地区气象数据，量化极端低温与降雪等气候因子对轨道热力耦合效应的影响机制，进而开发基于 TCN-Track 模型的温度场预测方法。该研究通过融合多尺度气象特征与热传导机理，构建具备动态预测能力的智能模型，为寒冷环境下高铁轨道的预防性维护和热应力调控提供理论依据与技术支撑。

**创新点：**1. 系统分析极端低温、降雪等气象因子对轨道多层结构温度分布的累积影响特性，为温度场建模提供科学依据；2. 突破传统模型仅基于瞬时气象数据的局限，实现了长期气象信息的深层特征提取与序列依赖建模；3 建立了轨道结构多层温度场预测体系，不仅覆盖轨道表面温度，同时揭示内部结构温度变化规律，全面支撑轨道力学性能演变分析。

**方　法：**1. 通过实测气象数据分析，识别寒冷地区无砟轨道温度场分布的主控因素，并揭示极端低温、太阳辐射与降雪等历史累积效应对温度演变规律的影响（图 3 和 4）；2. 基于时间卷积网络（TCN）理论，提出 TCN-Track 预测模型，构建局部与全局气象特征融合的温度场时序映射关系（图 5）；3. 通过超参数分析，

Unedited

确定模型训练轮数与卷积核尺寸对预测性能的影响规律，优化得到最优组合参数（表1、表2）4. 通过模型训练与对比实验，验证TCN-Track在温度场预测精度上优于传统LSTM方法，显著降低了预测误差（表3，图10）。

**结　论：** 1. 对沈阳地区实测数据进行了系统的时空分析，探究了气温、太阳辐射、风速、降雨和降雪等气象因素对轨道结构温度场的影响规律；2. 提出了一种TCN-Track模型，融合了TCN、LSTM和GRU网络，并引入多头注意力机制，有效整合气象因素的累积效应并提升特征提取能力，实现了对无砟轨道多层结构温度场的精准预测。模型在预测每日轨道结构各层温度时，取得了MAE在0.26至0.39之间、RMSE在0.32至0.50之间、相关系数R在0.888至0.985之间的优异结果；3. 与TCN、LSTM和GRU等单一模型相比，TCN-Track模型在轨道温度场预测任务中将MAE降低了3.61%至89.17%，RMSE降低了6.67%至88.51%，相关系数R提高了1.45%至29.57%。

**关键词：** 寒区；CRTS III型板式无砟轨道；温度预测；气象变量；时序预测模型

Unedited