UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

Vinicius Antonio Meneguitte Alves

A cost-efficient data-driven framework for rail temperature prediction

Juiz de Fora
2026

Vinicius Antonio Meneguitte Alves

A cost-efficient data-driven framework for rail temperature prediction

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Civil. Área de concentração: Estruturas e materiais

Orientador(a): Prof. Dr. Alexandre Abrahão Cury

Coorientador(a): Prof. Dr. Flávio de Souza Barbosa

Juiz de Fora

2026

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Alves, Vinicius.
A cost-efficient data-driven framework for rail temperature prediction / Vinicius Antonio Meneguitte Alves. – 2026.
46 f. : il.

Orientador(a): Alexandre Abrahão Cury
Coorientador(a): Flávio de Souza Barbosa
Dissertação (Engenharia Civil) – Universidade Federal de Juiz de Fora, Faculdade de Engenharia. Programa de Pós-Graduação em Engenharia Civil, 2026.

1. Machine Learning. 2. Railways. 3. Temperature predction. I. Cury, Alexandre, orient. II. Barbosa, Flávio, coorient. III. Título.

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

VINICIUS ANTONIO MENEGUITTE ALVES

Título: "A cost-efficient data-driven framework for rail temperature prediction"

Dissertação apresentada ao Programa de Pós Graduação em Engenharia Civil da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Civil. Área de concentração: Estruturas e Materiais

Aprovada em 22/05/2026

BANCA EXAMINADORA

Prof. Dr. Alexandre Abrahão Cury - Orientador e Presidente da Banca
Universidade Federal de Juiz de Fora

Prof. Dr. Flávio de Souza Barbosa - Coorientador
Universidade Federal de Juiz de Fora

Dra. Rafaelle Piazzaroli Finotti Amaral - Membro titular interno
Universidade Federal de Juiz de Fora

Dr. Rafael de Oliveira Teloli - Membro titular externo
Supmicrotech - ENSMM, Institut FEMTO-ST

Prof. Dr. Gustavo Henrique Nalon - Membro titular interno
Universidade Federal de Juiz de Fora

Juiz de Fora, 25/05/2026.

Documento assinado eletronicamente por **Alexandre Abrahao Cury**, **Professor(a)**, em 25/05/2026, às 16:06, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rafael de Oliveira Teloli**, **Usuário Externo**, em 26/05/2026, às 03:37, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rafaelle Piazzaroli Finotti Amaral**, **Usuário Externo**, em 26/05/2026, às 12:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Gustavo Henrique Nalon**, **Professor(a)**, em 26/05/2026, às 18:37, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Flavio de Souza Barbosa**, **Professor(a)**, em 28/05/2026, às 23:06, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **3000497** e o código CRC **8A4ABC41**.

Dedico este trabalho aos meus pais.

AGRADECIMENTOS

A conclusão deste mestrado é o reflexo de uma jornada tão desafiadora quanto gratificante, a qual não teria sido possível sem o apoio fundamental de pessoas que caminharam ao meu lado.

Aos meus pais, expresso minha mais profunda gratidão. Pela parceria incondicional, pela paciência nos momentos de ausência e por serem o alicerce que me permitiu chegar até aqui.

Aos meus dois irmãos, agradeço pelas conversas, pelas dicas valiosas e, acima de tudo, pela segurança e pelo incentivo que sempre me transmitiram.

Aos meus orientadores, Alexandre Cury e Flávio Barbosa, agradeço pela orientação empática, pela confiança depositada em meu trabalho e por todo o apoio compartilhado ao longo destes anos de pesquisa. Seus ensinamentos foram fundamentais para o meu amadurecimento acadêmico e profissional.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), agradeço pelo suporte financeiro por meio do Projeto de Desenvolvimento de estratégias SHM para ferrovias.

Por fim, aos meus amigos, meu muito obrigado por tornarem os dias árduos mais leves e por transformarem o percurso em uma experiência muito mais calma.

RESUMO

Altas temperaturas nos trilhos representam riscos significativos à integridade estrutural das vias férreas, aumentando a probabilidade de flambagem. Este trabalho propõe um modelo utilizando apenas a temperatura ambiente e a umidade relativa como variáveis de entrada, grandezas rotineiramente coletadas por estações meteorológicas convencionais, eliminando a dependência de redes complexas de sensores. Além disso, compara cinco algoritmos de aprendizado de máquina: Redes Neurais Artificiais (ANN), Regressão por Vetores de Suporte (SVR), Floresta Aleatória (RF), redes Long Short-Term Memory (LSTM) e XGBoost. Duas estratégias de previsão foram implementadas: um modelo global de longo prazo, treinado com dois anos de dados históricos, e um modelo de janela deslizante de curto prazo, com horizonte de até três horas. Os modelos foram treinados com dados ferroviários brasileiros de 2022–2023 e validados em um conjunto independente de 2024. A otimização de hiperparâmetros foi realizada via Optuna em estrutura de árvore. Os resultados indicam que o RF obteve a melhor precisão no horizonte de uma hora ($R^2 = 0{,}94$; erro médio absoluto = 1,19 °C), enquanto a ANN se destacou na previsão de longo prazo ($R^2 = 0{,}87$; erro médio absoluto = 2,2 °C). Assim, entende-se que a estratégia proposta demonstra acurácia competitiva com conjunto de entradas reduzido, oferecendo uma solução prática e escalável para redes ferroviárias com infraestrutura de monitoramento limitada.

Palavras-chave: Ferrovias; Aprendizado de máquina; Previsão de temperatura; Monitoramento estrutural; Flambagem de trilhos.

ABSTRACT

High rail temperatures pose significant risks to track integrity, increasing the likelihood of buckling. This study proposes a cost-effective, data-driven framework for rail temperature prediction using only ambient temperature and relative humidity as inputs, variables routinely collected by standard meteorological stations, thereby eliminating dependence on complex sensor arrays. Five machine learning algorithms were benchmarked: Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest (RF), Long Short-Term Memory networks (LSTM), and XGBoost. Two complementary prediction strategies were evaluated: a long-term global model trained on a two-year historical dataset, and a short-term sliding-window model predicting up to three hours ahead. Models were trained on Brazilian railway data from 2022–2023 and validated on an independent 2024 dataset. Hyperparameters were optimized via Optuna using a Tree-structured Parzen Estimator (TPE). Results demonstrate that ANN delivered the lowest long-term error (mean absolute error = 2.0978 °C; $R^2 = 0.8571$), closely matched by XGBoost, which obtained the highest R^2 (0.8652), while RF achieved the best short-term precision at the one-hour horizon ($R^2 = 0.9395$; mean absolute error = 1.1944 °C). Despite the deliberately limited input set, the proposed framework achieves competitive accuracy, offering a practical solution for railway networks operating under limited monitoring infrastructure.

Keywords: rail temperature; track buckling; railway safety; time series forecasting; feature reduction; structural health monitoring.

List of Figures

List of Tables

Contents

1 GENERAL INTRODUCTION

The Brazilian railway sector plays a strategic role in the large-scale transport of cargo, especially minerals and agricultural products, representing approximately 20% of the national freight transport matrix (ILOS - Instituto de Logística e Supply Chain, 2024). Currently, Brazil has about 30,000 kilometers of railways, and despite being recognized as the most sustainable, safe, and efficient mode for transporting large volumes, the country still faces limitations related to modernization, infrastructure maintenance, and network expansion (EPL - Empresa de Planejamento e Logística S.A., 2021). The National Logistics Plan (EPL - Empresa de Planejamento e Logística S.A., 2021), published in 2021, sets a goal to increase the railway share to 36% by 2035, highlighting the need for investments in infrastructure, innovation, and monitoring technologies.

Within the logistical scope, rail temperature prediction also allows for more efficient planning of the railway network. By identifying sections that will be out of service on certain days, it is possible to reroute traffic in advance, minimizing the operational costs of a stationary locomotive.

In this context, this study proposes a comparative performance evaluation of various artificial intelligence models for rail temperature prediction using a reduced and cost-effective set of input features. This analysis is fundamental for identifying the most efficient and accurate approach, considering factors such as the ability to capture non-linear patterns and temporal data dependencies. This research is justified by the vital importance of anticipating railway track temperatures that could severely compromise train operations, thereby avoiding accidents and network delays.

This work addresses a significant gap in the literature, as there is a scarcity of similar studies conducted within the Brazilian context. While current international research often relies on predictive systems requiring a vast array of complex meteorological and operational input variables, this high dimensionality hinders practical deployment. In Brazil, railway transport companies face substantial challenges in acquiring and continuously monitoring such extensive datasets. Therefore, by deliberately restricting the input features to readily available data, the proposed models offer a highly applicable and cost-effective solution tailored to the operational reality of the national railway sector.

1.1 OBJECTIVE

The primary objective of this study is to evaluate different machine learning models for predicting rail temperature, using historical rail temperature data and ambient temperature as input variables.

1.2 SPECIFIC OBJECTIVES

The specific objectives of this study are to apply algorithms widely used for prediction and time series analysis, such as Artificial Neural Networks (ANN), Support Vector Machines (SVR), Random Forest (RF), Long Short-Term Memory networks (LSTM) and Extreme Gradient Boosting (XGBoost), comparing them through appropriate evaluation metrics and evaluating by a long-term and a short-term temperature prediction strategies.

1.3 TEXT STRUCTURE

This work is organized as follows: the general introduction presents the research topic along with the general and specific objectives. Chapter 2 consists of the paper to be submitted to the journal *Railway Engineering Science* (Q1), which incorporates the literature review within it. Finally, Chapter 3 provides the concluding remarks and offers recommendations for future research.

1.4 REFERENCES

EPL. Plano Nacional de Logística 2035. Brasília, DF, 2021. Disponível em: <https://portal.epl.gov.br/plano-nacional-de-logistica-2035>.

ILOS. Matriz de transporte de carga (//prime.ilos.com.br/>

2 ARTICLE - A COST-EFFICIENT DATA-DRIVEN FRAMEWORK FOR RAIL TEMPERATURE PREDICTION

Article submitted to the scientific journal Railway Engineering Science.

Abstract. High rail temperatures pose significant risks to track integrity, increasing the likelihood of buckling. This study proposes a cost-effective, data-driven framework for rail temperature prediction using only ambient temperature and relative humidity as inputs, variables routinely collected by standard meteorological stations, thereby eliminating dependence on complex sensor arrays. Five machine learning algorithms were benchmarked: Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest (RF), Long Short-Term Memory networks (LSTM), and XGBoost. Two complementary prediction strategies were evaluated: a long-term global model trained on a two-year historical dataset, and a short-term sliding-window model predicting up to three hours ahead. Models were trained on Brazilian railway data from 2022–2023 and validated on an independent 2024 dataset. Hyperparameters were optimized via Optuna using a Tree-structured Parzen Estimator (TPE). Results demonstrate that ANN delivered the lowest long-term error (mean absolute error = 2.0978 °C; R^2 = 0.8571), closely matched by XGBoost, which obtained the highest R^2 (0.8652), while RF achieved the best short-term precision at the one-hour horizon (R^2 = 0.9395; mean absolute error = 1.1944 °C). Despite the deliberately limited input set, the proposed framework achieves competitive accuracy, offering a practical solution for railway networks operating under limited monitoring infrastructure.

Keywords: rail temperature; track buckling; railway safety; time series forecasting; feature reduction; structural health monitoring.

2.1 INTRODUCTION

Railway infrastructure represents a critical component of the global transportation system, with networks spanning diverse climatic environments ranging from temperate European corridors to subtropical East Asian high-speed lines and tropical South American freight routes. Across all these contexts, thermal effects on railway tracks represent a persistent operational and safety challenge (Palin et al., 2021). In Europe, extreme summer heat waves have been documented to cause service disruptions, speed restrictions, and increased maintenance demands (Ferranti et al., 2018). In East Asia, high-speed railways must manage large seasonal and diurnal temperature variations that affect track geometry and structural integrity (Zhou et al.; Han et al., 2024, 2020). Globally, the risk of track buckling under thermal stress is a recognized threat demanding robust monitoring and prediction systems (Agustin, Wu e Ngamkhanong, 2025).

One of the main challenges associated with railway operations is related to climatic

conditions, especially in tropical regions (Jie et al., 2025) such as Brazil, as high ambient temperatures and large daily thermal variations directly affect the integrity of the rails, which can reach temperatures well above ambient levels. This phenomenon intensifies the effects of thermal expansion, resulting in risks of track buckling and fracture (Ferranti et al., 2018). Such occurrences, in addition to compromising safe operations, can cause serious accidents, interrupt freight transport, and generate high corrective maintenance costs for concessionaires (Palin et al., 2021).

In this context, the prediction of rail temperature assumes a central role in terms of safety and logistics (Agustin, Wu e Ngamkhanong, 2025). Accurate forecasting allows operators to anticipate critical thermal conditions and to adopt preventive measures, such as speed reductions in specific sections, load redistribution, and scheduled maintenance, thereby mitigating failure risks and extending the lifespan of the infrastructure.

More broadly, computational models have been increasingly adopted to assess the structural behavior and reliability of railway infrastructure, whether calibrated with experimental measurements (Vidal, Cury e Barbosa, 2026) or trained on numerical simulation data (Melo et al., 2025). Within this wider trend, artificial intelligence-based models have emerged as a particularly promising strategy for rail temperature prediction, offering notable gains in precision, efficiency, and sustainability for ensuring the operational resilience of modern railway networks (Agustin, Wu e Ngamkhanong, 2025).

Despite growing international research on rail temperature forecasting, most existing frameworks rely on complex, multi-sensor meteorological inputs that are difficult to maintain in regions with limited infrastructure, leaving a significant operational gap that simpler, data-efficient approaches could fill. To address this gap, the primary contribution of this study is the proposal of a machine learning-based prediction model that requires only ambient temperature and relative humidity, two variables routinely measured by standard meteorological stations.

By deliberately restricting the input feature set, the proposed framework prioritizes operational feasibility and cost-effectiveness over absolute predictive performance. This design choice is particularly relevant for railway networks in regions where installing and maintaining an extensive array of specialized sensors (solar radiation, wind speed or cloud cover) is logistically or economically prohibitive. Rather than aiming to outperform the most complex models reported in the literature, this work demonstrates that a carefully selected subset of machine learning algorithms can achieve useful accuracy with reduced data requirements, thereby allowing broader adoption of proactive thermal monitoring in resource-constrained contexts.

To achieve this goal, five machine learning algorithms widely used for prediction and time series analysis in Structural Health Monitoring are employed (Júnior et al., 2026): Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest (RF),

Long Short-Term Memory networks (LSTM), and Extreme Gradient Boosting (XGBoost). All models are trained and validated using a real-world dataset of rail temperatures collected directly from an active Brazilian railway, ensuring that the evaluation reflects actual field conditions. Their performance is compared through four metrics and assessed within two complementary prediction strategies:

- A long-term approach that utilizes the full historical dataset (two years of data) for model training to generate long-term forecasts as required.
- A short-term approach that uses the preceding 6 hours of data to predict the subsequent 3-hour interval.

From an operational standpoint, this forecasting capability provides significant utility for railway management. By utilizing extended weather forecasts, operators can generate a comprehensive rail buckling risk calendar for the entire upcoming week. This proactive insight allows maintenance teams to identify critical risk windows well in advance, optimizing scheduling and enhancing track safety protocols before critical thermal thresholds are reached.

Conversely, the short-term approach offers distinct advantages in terms of immediate precision. By utilizing recent historical lags, short-term models can capture sudden microclimatic shifts and account for thermal inertia inherent to the rail infrastructure. This integration of immediate past data can reduce prediction error for the upcoming hours, making it a valuable tool for traffic management, automated emergency alerts, and decision-making.

2.2 BACKGROUND AND RELATED WORK

Historically, rail temperature prediction methodologies are grouped into three primary approaches: empirical equation-based models, multivariate regression models, and thermal analysis models.

The first approach estimates rail temperature as a linear function of a single parameter. While these models are computationally efficient and easy to implement, they generally yield lower predictive accuracy compared to more complex methods (Hunt, 1994).

The second approach predicts rail temperature using multiple meteorological variables, such as ambient air temperature, wind speed, and cloud cover. These models tend to outperform single-parameter equations. For instance, Wu et al. (Wu et al., 2012) employed a multivariate regression with 24 features, achieving an R^2 of 0.9630 and an RMSE of 2.560°C. However, a significant limitation of this approach is its lack of generalizability across different climatic regions.

The third approach models the thermodynamic interactions between the rail and the surrounding environment. This method offers superior generalizability but requires precise knowledge of specific rail properties (Ou e Li, 2014), such as reflectance and emissivity. For example, Han et al. (Han et al., 2020) conducted thermofluidic analysis using numerical simulations based on a solar radiation model to predict rail temperatures and mitigate summer buckling events in Korea.

Although the literature presents several established approaches for rail temperature forecasting (Hong et al.; Bruzek et al., 2019, 2014), more recent studies have explored autoregressive statistical methods, such as that proposed by Pham et al. (Pham et al., 2025), with a particular focus on machine learning techniques.

Among these machine learning approaches, Hong et al. (Hong, Park e Cho, 2021) compared XGBoost, SVR, RF, polynomial regression of order 2 (PR2) and ANN algorithms. These models utilized several meteorological variables as input features obtained from weather forecasting services. The XGBoost algorithm exhibited the best performance in capturing thermal fluctuations and was used to forecast temperatures up to 64 hours ahead at 3-hour intervals.

The study by Liu et al. (Liu et al., 2023) investigated the thermal characteristics of rails on a bridge in China using a genetic algorithm-optimized back propagation (GA-BP) neural network. The model was fed with meteorological input features, specifically air temperature and solar radiation. Their research specifically focused on predicting the minimum and maximum temperatures of the day.

Zhou et al.(Zhou et al., 2024) investigated temperature prediction on a high-speed railway line in China by comparing various neural network architectures, ANN, CNN and LSTM. These models utilize meteorological variables such as air temperature, solar radiation, and wind speed. Their study also incorporates the track's own temperature as an input feature, which is a different approach from other studies and brings an advantage in its comparative metrics. The study demonstrated high precision in capturing seasonal and diurnal thermal variations to forecast the temperature 10 minutes in advance.

Pham et al. (Pham et al., 2025) proposed a monitoring and prediction system for the Seoul subway in South Korea using a Bayesian Long Short-Term Memory (BLSTM) model. The model is structured into specific time windows to forecast the temperature 3 hours in advance.

Cho et al. (Cho et al., 2026) proposed a data-driven approach using regression and machine learning models, such as RANSAC, RF, and ANN, to estimate rail temperatures on a bridge in South Korea. The models were developed using air temperature, solar irradiation, and land surface temperature as primary input features to identify high-risk buckling conditions. Their findings revealed that RF and Linear Regression performed best, accurately capturing thermal peaks that can exceed 35°C during the summer. The

methodology provides a reliable tool for predicting rail temperatures in locations where direct sensor installation is technically challenging.

Table 1 summarizes the general information and methodological characteristics of the machine learning models identified in the literature, highlighting a wide array of complex input variables, such as solar radiation, wind speed, solar azimuth, rainfall, wind speed, cloud cover, etc.

It is worth noting that the body of machine learning work on rail temperature prediction has so far been concentrated in East Asia, particularly China and South Korea, as reflected in Table 1. In other regions, rail temperature forecasting has been addressed predominantly through physically based or statistical models rather than machine learning. In the United Kingdom, Chapman et al. (Chapman et al., 2008) proposed an energy-balance model of rail surface temperature driven by standard meteorological variables. In the United States, Zhang and Lee (Zhang e Lee, 2008) developed a heat-transfer model fed with real-time weather data, an approach subsequently consolidated into operational forecasting systems. In Australia, Wu et al. (Wu et al., 2012) formulated the multivariate regression equation discussed above, while in Europe, Piloto et al. (Piloto et al., 2022) validated a lumped-thermal model against field measurements in Portugal. This geographic and methodological distribution indicates that data-driven approaches remain comparatively unexplored outside East Asia, reinforcing the relevance of evaluating such methods under the climatic and operational conditions of other regions, such as the Brazilian network considered in this study.

As previously stated, the models proposed in this study utilize only two input parameters: ambient temperature and relative humidity. By reducing the need for specific meteorological sensors, this approach simplifies the operational requirements for railway companies. Consequently, it offers a more cost-effective and rapid implementation path, allowing for the initiation of thermal control and monitoring protocols with lower initial investment and less technical complexity.

Table 1 – Related rail temperature prediction models

Model	**Ref**	**Variables (Inputs)**	**Location**
Bayesian LSTM	(Pham et al., 2025)	Air temp., solar radiation, wind speed, solar azimuth, and altitude.	South Korea
ANN (Case 1)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step	China
ANN (Case 2)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, wind	China
ANN (Case 3)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, solar radiation	China

Continues on the next page

Related rail temperature prediction models (continued)

Model	Ref	Variables (Inputs)	Location
ANN (Case 4)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, humidity	China
ANN (Case 5)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, wind, humidity, solar radiation	China
CNN (Case 1)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step	China
CNN (Case 2)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, wind	China
CNN (Case 3)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, solar radiation	China
CNN (Case 4)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, humidity	China
CNN (Case 5)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, wind, humidity, solar radiation	China
LSTM (Case 1)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step	China
LSTM (Case 2)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, wind	China
LSTM (Case 3)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, solar radiation	China
LSTM (Case 4)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, humidity	China
LSTM (Case 5)	(Zhou et al., 2024)	Ambient temp., rail temp. at previous step, wind, humidity, solar radiation	China
BP	(Liu et al., 2023)	Air temp., solar radiation intensity, wind speed, precipitation, humidity	China
GA-BP	(Liu et al., 2023)	Air temp., solar radiation intensity, wind speed, precipitation, humidity	China
PR2	(Hong, Park e Cho, 2021)	Air temp., rainfall, wind speed, cloud cover, relative humidity, azimuth, altitude, and solar irradiance	South Korea
SVM	(Hong, Park e Cho, 2021)	Air temp., rainfall, wind speed, cloud cover, relative humidity, azimuth, altitude, and solar irradiance	South Korea
ANN	(Hong, Park e Cho, 2021)	Air temp., rainfall, wind speed, cloud cover, relative humidity, azimuth, altitude, and solar irradiance	South Korea

Continues on the next page

Related rail temperature prediction models (continued)

Model	Ref	Variables (Inputs)	Location
RF	(Hong, Park e Cho, 2021)	Air temp., rainfall, wind speed, cloud cover, relative humidity, azimuth, altitude, and solar irradiance	South Korea
XGBoost 1	(Hong, Park e Cho, 2021)	Air temp., rainfall, wind speed, cloud cover, relative humidity, azimuth, altitude, and solar irradiance	South Korea
XGBoost 2	(Hong, Park e Cho, 2021)	Air temp., rainfall, wind speed, cloud cover and relative humidity	South Korea
RF	(Cho et al., 2026)	Air temp., solar irradiation, and land surface temperature	South Korea
SVM	(Cho et al., 2026)	Air temp., solar irradiation, and land surface temperature	South Korea
ANN	(Cho et al., 2026)	Air temp., solar irradiation, and land surface temperature	South Korea

Source: Organized by the author (2026).

2.3 DATASET AND PREPROCESSING

2.3.1 INPUT VARIABLE SELECTION

The choice of ambient temperature and relative humidity as the sole meteorological inputs is not merely a pragmatic simplification, it is physically grounded. Rail surface temperature is primarily driven by the net energy balance at the rail–atmosphere interface, which is dominated by convective heat exchange with the surrounding air and by longwave radiative emission, both of which are strongly modulated by ambient temperature. Relative humidity, in turn, influences the latent heat flux and affects the effective sky emissivity, thereby modulating the nocturnal cooling rate of the rail. Although solar irradiance is the dominant forcing during peak heating hours, its effect is already encoded in the temporal features (hour of day and day of year) included in the model, which capture the diurnal and seasonal cycles of solar incidence at the study location.

The selection of a limited set of input variables is a strategic priority for both temperature prediction strategies. For short-term operational deployments, a reduced number of sensors—specifically temperature and humidity—lowers the financial and logistical barriers for companies to implement monitoring systems across extensive geographic networks compared to multi-sensor stations. In the long-term context, the dependency on forecasted inputs introduces a cumulative risk of error propagation. As the number of input variables increases, the model becomes increasingly sensitive to the inaccuracies of external meteorological forecasts. Furthermore, obtaining reliable, long-range predictions for complex variables such as solar radiation, cloud cover, and wind speed is often impractical or computationally expensive, potentially compromising the model's utility in real-world forecasting scenarios.

2.3.2 DATA DESCRIPTION

The dataset employed in this study comprises rail temperature, ambient temperature and relative humidity readings. A dataset of rail temperatures was obtained via a thermometer installed on a southeastern railway network in Brazil. In parallel, ambient temperature and relative humidity data from the same city were sourced from INMET, the National Meteorological Institute.

Rail temperature measurements were recorded at 15-minute intervals, providing high temporal resolution. Conversely, ambient temperature and relative humidity data were provided on an hourly basis. This discrepancy in time scales presented a challenge for comparative analysis, requiring a data alignment process prior to the application of predictive modeling techniques.

To synchronize the time series, a preprocessing step was performed: the four rail temperature readings recorded every 15 minutes within each hour were averaged to produce a single hourly value, aligning them with the frequency of the meteorological data. Table 2 illustrates this procedure for a representative three-hour excerpt; the same averaging was applied to the entire dataset.

Table 2 – Temperature and Humidity Monitoring

Time		**Rail Temp. (ºC)**	**Average Rail Temp. (ºC)**	**Ambient Temp. (ºC)**	**Humidity (%)**
13:00	13:00	51.8	52.1	25.0	62.0
13:15		52.0			
13:30		51.8			
13:45		52.7			
14:00	14:00	53.6	53.2	26.4	57.3
14:15		53.1			
14:30		52.8			
14:45		53.4			
15:00	15:00	53.1	51.7	27.5	51.4
15:15		52.4			
15:30		51.0			
15:45		50.1			

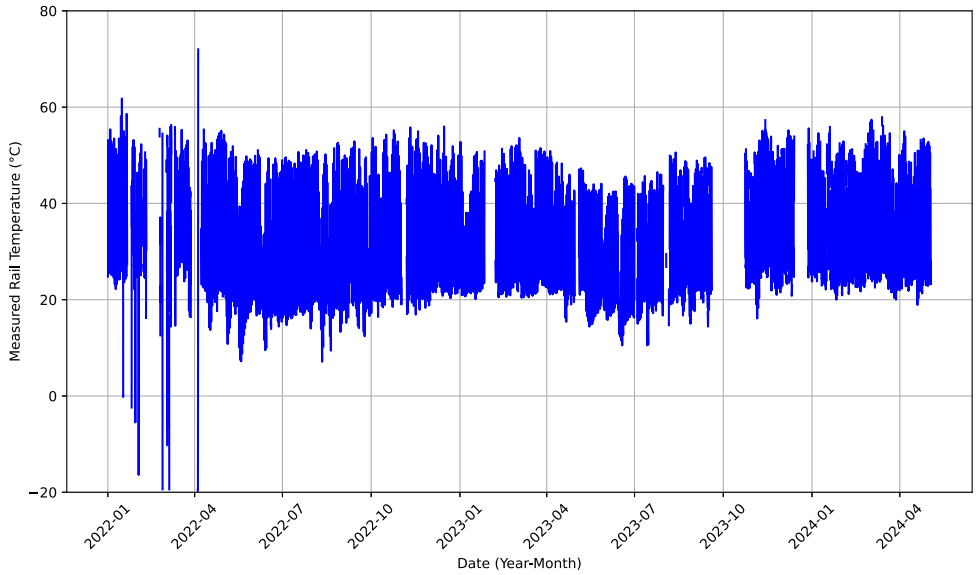
Source: Prepared by the author (2026).

The data span is from January 1, 2022, to March 5, 2024, offering a comprehensive overview of local weather conditions throughout this interval. To ensure precise temporal alignment, these variables are cross-referenced using day and hour timestamps. Specifically, the temporal features are encoded such that the days of the year range from 1 to 365, while the hourly cycle is represented on a scale from 0 to 23. This structured indexing allows the models to account for both seasonal variations and daily thermal fluctuations.

Figure 1 presents the average rail temperature calculated at hourly intervals throughout the analyzed period, while Figures 2 and 3 illustrate, respectively, the evolution of ambient temperature and relative humidity over the same timeframe. As can be observed in all three figures, the time series exhibit intermittent gaps, which are attributed to data acquisition failures

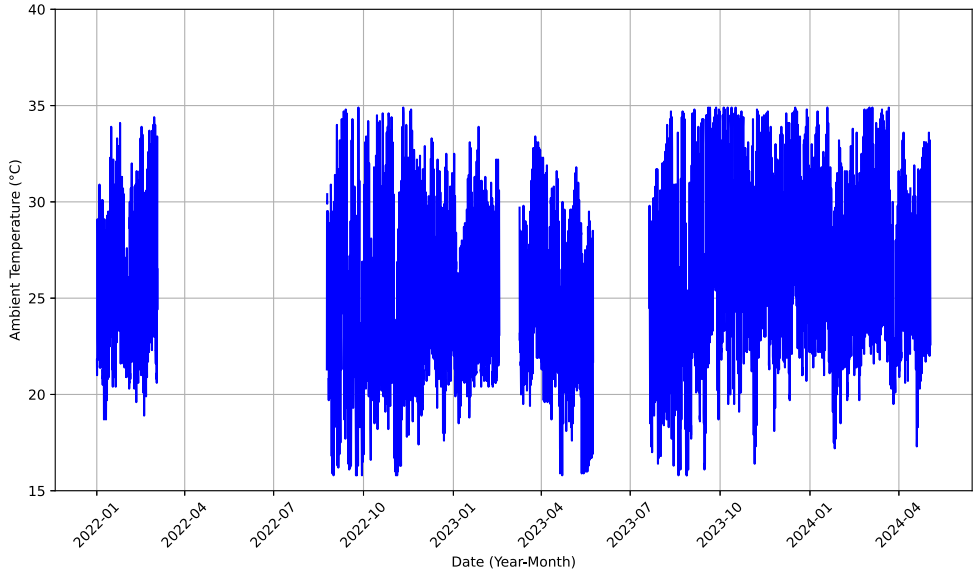
detected in both the rail temperature sensor and the meteorological records provided by INMET. These missing intervals were identified during the preprocessing stage and handled as discussed in detail in Section **2.3.2**.

Figure 1 – Rail temperature data.



Source: Prepared by the author (2026).

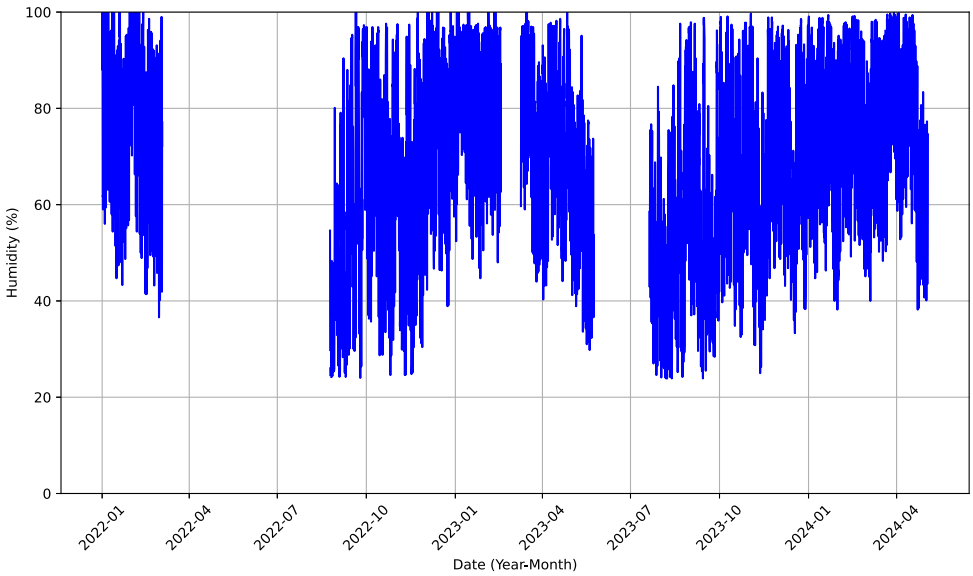
Figure 2 – Ambient temperature data.



Source: Prepared by the author (2026).

Fig. 4 presents the temporal variation of rail and ambient temperatures alongside relative humidity during September 2022, a representative period selected to facilitate detailed visualization of the daily thermal cycles. The figure reveals a clear diurnal pattern in all three

Figure 3 – Relative humidity data.



Source: Prepared by the author (2026).

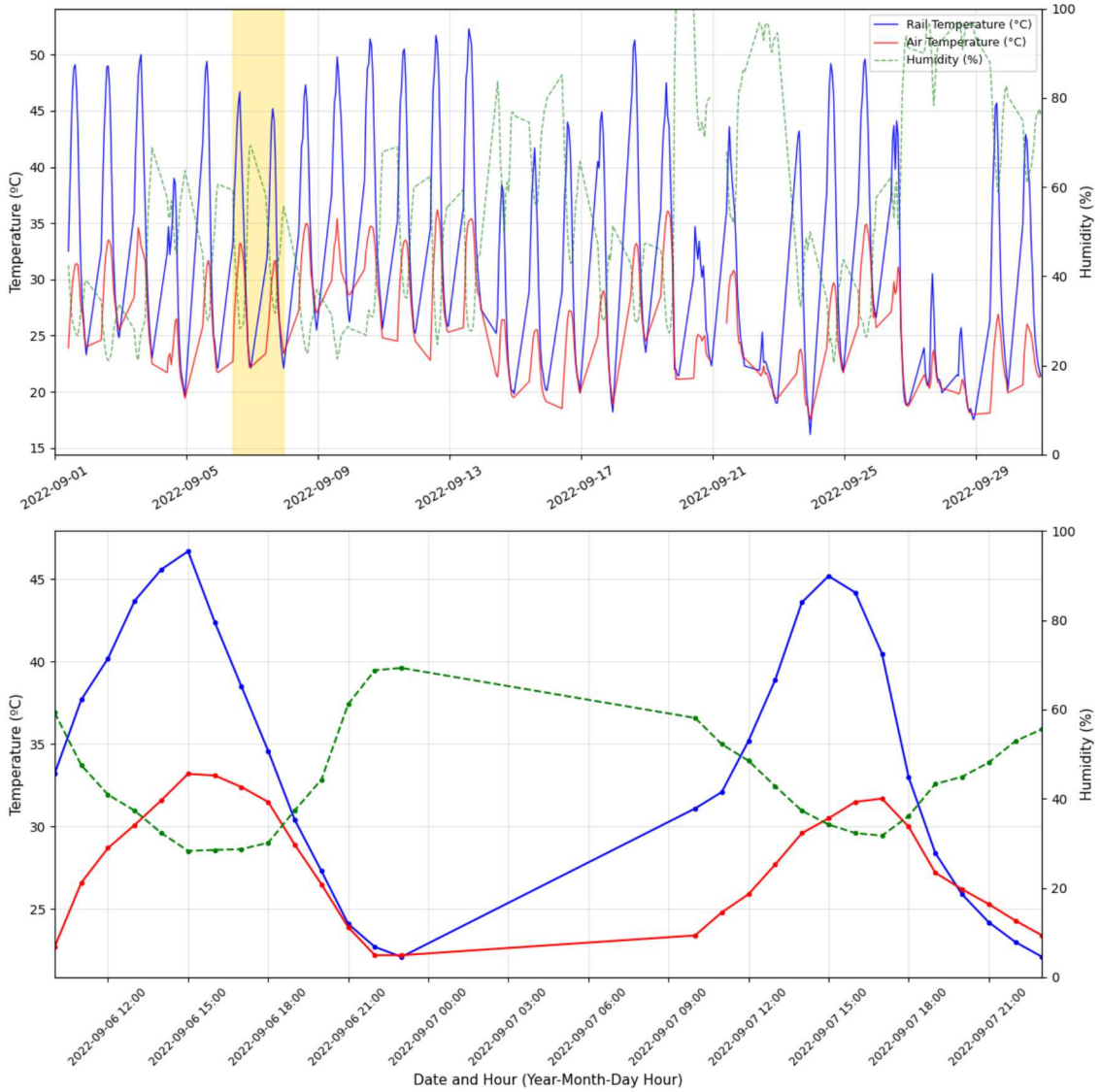
variables: rail and ambient temperatures rise progressively throughout the day, reaching their respective peaks in the early afternoon, while relative humidity follows an inverse trend, dropping to its lowest values during peak heating hours. Rail temperatures consistently exceed ambient values throughout the entire period, with differences that can surpass 20^{o}C during daytime hours, occasionally reaching values above 50^o C, whereas ambient temperature rarely surpasses 35^o C. This pronounced thermal differential reflects both the high absorption capacity of the steel rail when exposed to direct solar radiation and its low thermal reflectivity. During nighttime and periods of elevated humidity, this gap narrows considerably, as convective cooling and increased atmospheric moisture contribute to reducing the thermal differential between the rail and its surrounding environment, reinforcing the negative correlation between temperature and humidity observed throughout the entire month.

The predictive framework developed in this study is built upon a multivariate dataset, where the input features consist of ambient temperature, relative humidity, and temporal markers, the day of the year and the hour of the day. These variables provide the necessary context to capture both daily thermal cycles and broader seasonal trends. The target output of the models is the rail temperature, which is estimated as a continuous value through time series regression.

To capture the cyclical nature and continuity of the temporal variables, a trigonometric transformation was applied to the hour and day of the year timestamps. In their original numerical format, such as values ranging from 0 to 23 for hours or 1 to 365 for the day of the year, these variables exhibit an artificial discontinuity at boundary periods. Example, the transition from 23 : 00 to 00 : 00 or from December 31 to January 1.

This discontinuity can mislead machine learning models into treating temporally adjacent values as distant in the feature space. To overcome this, the timestamps were decomposed into

Figure 4 – Temperature data on the rail, ambient temperature, and humidity for September 2022.



Source: Prepared by the author (2026).

sine and cosine components, mapping the cyclical variables onto a continuous unit circle. This transformation preserves the temporal continuity and maintains a consistent distance between all time steps, as defined in equations (2.1) and (2.2):

$$t_h = \sin\left(\frac{2\pi \cdot t}{T_h}\right) \tag{2.1}$$

$$t_y = \cos\left(\frac{2\pi \cdot t}{T_y}\right) \tag{2.2}$$

where t denotes the timestamp value; $T_h = 24$ and $T_y = 365$ are the periods associated with the daily and annual cycles, respectively; and t_h and t_y are the corresponding cyclic encodings that

capture the periodic nature of hour-of-day and day-of-year variations.

2.3.3 DATA CLEANING AND FEATURE ENGINEERING

Figures 1, 2 and 3 reveal the occurrence of some anomalies, primarily related to the absence of records during certain intervals. Since the data were obtained from an external source, they are subject to operational failures or temporary unavailability, resulting in gaps within the analyzed dataset.

Given this scenario, it is essential to adopt missing data handling strategies to ensure the consistency and completeness of the information used, thereby avoiding bias in the analysis. As observed in Figures 2 and 3, there are extensive continuous periods completely lacking data. Over such large time gaps, any attempt at mathematical approximation or imputation would be highly unreliable and fail to reflect the true environmental conditions. Therefore, the exclusion of observations with missing variables was chosen in this study, thus ensuring the integrity and coherence of the database for subsequent analyses.

Distortions in the measurements are observable in Figure 1. An anomaly occurs between January and April 2022, where extreme temperatures outside the expected pattern are recorded. Visual inspection revealed physically implausible values, such as abrupt spikes exceeding 80°C or negative readings during daytime, which could not be attributed to actual thermal events. These discrepancies were likely caused by sensor drift, inadequate calibration, or logging errors.

Thus, for outlier detection, the Interquartile Range (IQR) statistical method was adopted. In this study, a multiplier of $1.0 \times$ IQR was utilized. The lower bound is defined by Equation 2.3 and the upper bound by Equation 2.4 where Q_1 corresponds to the lower quartile of the data set, Q_3 to the upper quartile and the IQR being calculated as the difference between the third quartile (Q_3) and the first quartile (Q_1).

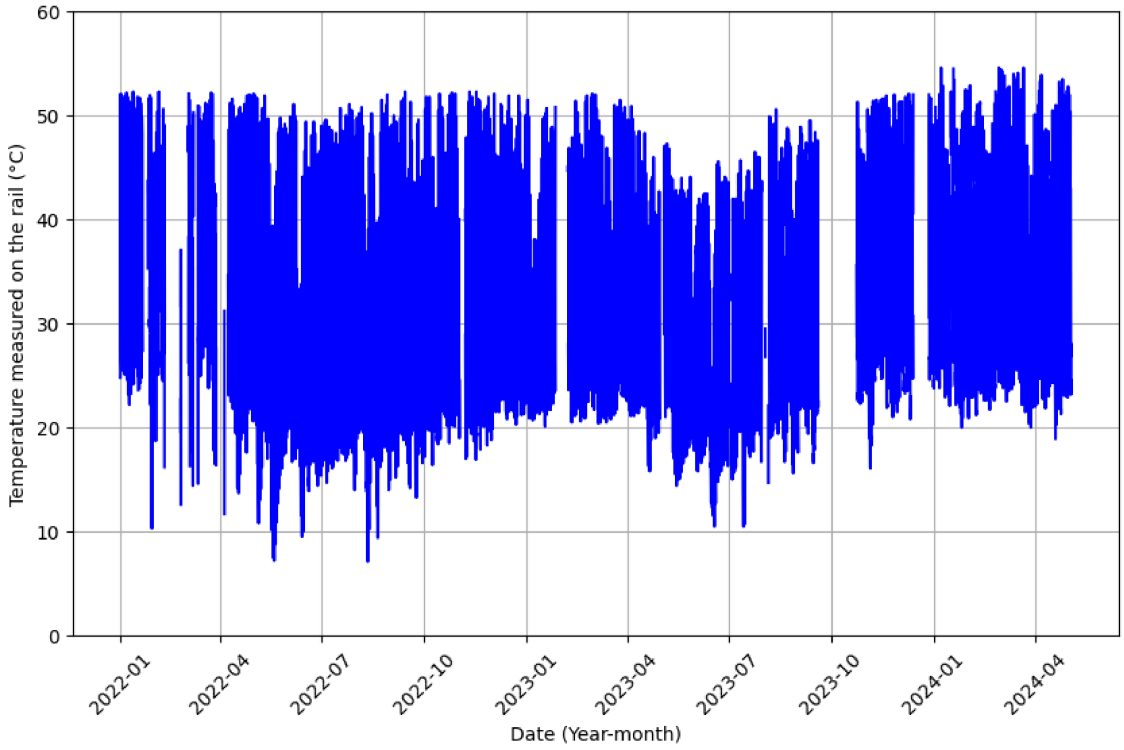
$$\text{Lower bound} = Q_1 - (1 \cdot IQR) \tag{2.3}$$

$$\text{Upper bound} = Q_3 + (1 \cdot IQR) \tag{2.4}$$

This stricter threshold was chosen because a preliminary analysis showed that the standard 1.5 multiplier still retained a number of non-physical values that would distort model training. By establishing these tolerance thresholds, it was possible to systematically isolate and remove discrepant values that would otherwise skew the analysis. The resulting dataset, Figure 5, retains the natural thermal variability, including summer peaks above 50°C, while excluding only those measurements that are inconsistent with the expected physical behavior of the rail.

For all the developed models, the data collected from January 1, 2022, to December 31, 2023, were allocated for training, while the observations spanning from January 1 to March 5, 2024 (i.e. the austral summer, a window with high-temperature regimes, most relevant to buckling) were reserved exclusively for testing. Consequently, the results presented in this study solely reflect the models' performance on data from the year 2024. This allows the evaluation of their generalization capabilities when generating forecasts within a temporal interval distinct from the one used for learning.

Figure 5 – Rail temperature data after outlier removal.



Source: Prepared by the author (2026).

2.4 METHODOLOGY

Two predictive strategies were adopted, short-term and long-term forecasting. The short-term forecasting strategy is designed to capture the thermal inertia of the rails, requiring a sliding window of the previous 6 hours of historical data to project temperatures for a multi-step horizon of +1h, +2h, and +3h.

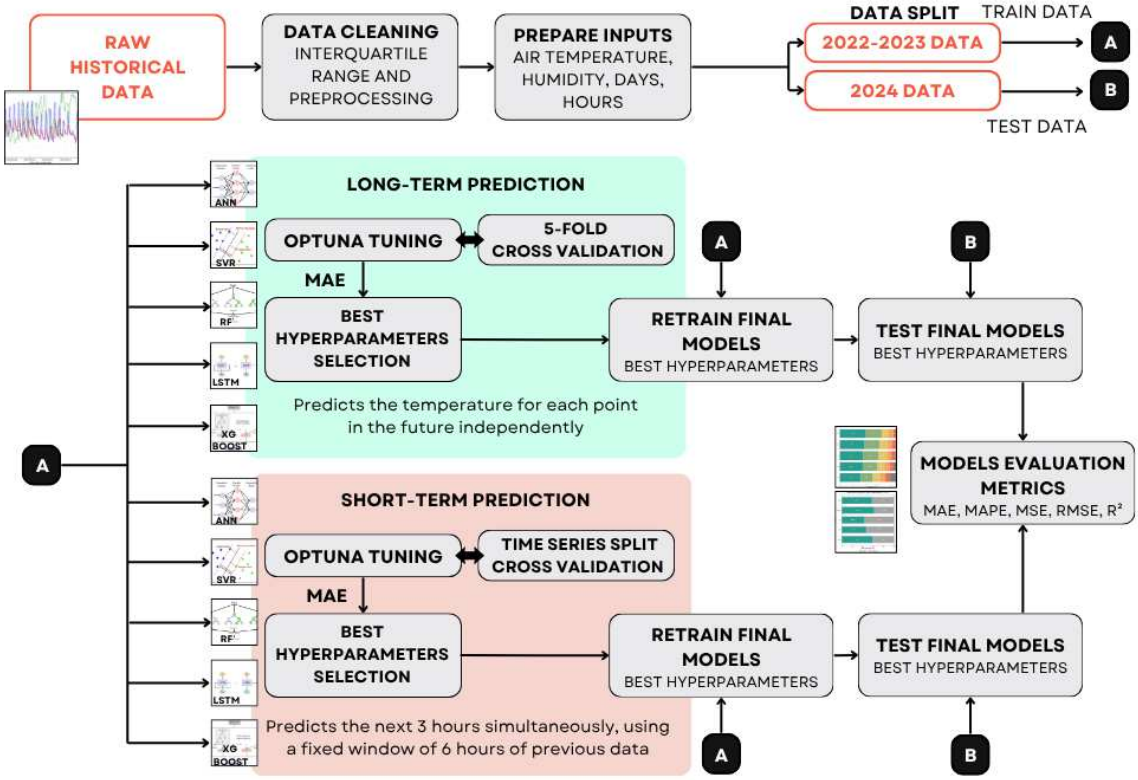
In contrast, the long-term approach is developed to provide a generalized mapping of rail temperatures at specific future intervals. In practice, the model generates a sequence of hourly rail temperature predictions corresponding directly to the availability of external weather forecasts. For every future hour targeted, the user must provide a set of forecasted exogenous inputs, specifically ambient temperature and relative humidity. The model then processes each hourly dataset to predict the rail temperature for that specific hour.

2.4.1 PROPOSED FRAMEWORK

The methodological framework, illustrated in Figure 6, encompasses the complete workflow from raw historical data to model evaluation. The process initiates with data cleaning and input preparation, followed by a data split into training (2022-2023) and testing (2024) sets. Two parallel prediction strategies were implemented: long-term prediction, which forecasts temperature for each future point independently using 5-fold cross-validation, and short-term prediction, which predicts the next 3 hours simultaneously using a 6-hour sliding window with

time series split validation. Both strategies employed multiple algorithms (ANN, SVR, RF, LSTM and XGBoost), with Optuna hyperparameter tuning and final model evaluation through metrics including MAE, MAPE, MSE, RMSE, and R^2.

Figure 6 – Methodological flowchart summarizing the data processing and modeling approach.



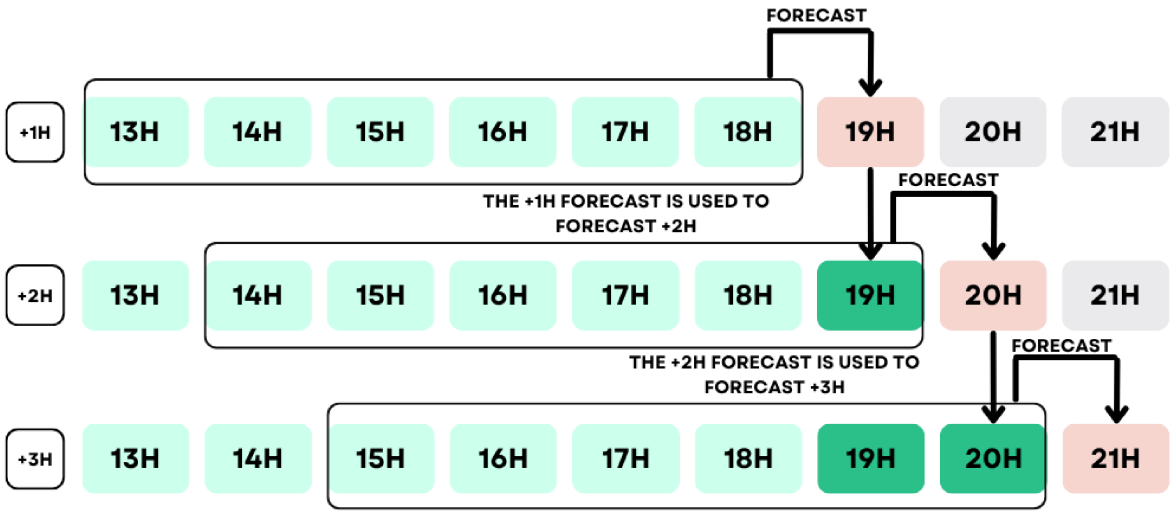
Source: Prepared by the author (2026).

Specifically to short-term strategy, as seen in Figure 7, to predict the temperature at $t+1$, the model utilizes the preceding 6 hours of actual historical observations. For the $t+2$ prediction, the input window shifts to input 5 hours of actual historical data along with the previously predicted value at $t+1$ acting as the most recent lag. This recursive feedback loop is extended to the $t+3$ step, where the model relies on 4 hours of actual historical data and the two previously generated predictions ($t+1$ and $t+2$), systematically updating the sliding window to capture real-time physical dynamics.

2.4.2 MODEL TRAINING AND HYPERPARAMETER TUNING

The training, validation, and test sets were composed of the variables: ambient temperature, rail temperature, relative humidity, day and hour of the reading. During the training and

Figure 7 – Short-term Sliding Window.



Source: Prepared by the author (2026).

validation phases, all these variables were utilized, with the rail temperature strictly defined as the target variable. In the testing phase, only the ambient temperature, relative humidity, day and hour were provided as input features, leaving the model with the task of estimating the rail temperature.

To ensure statistical robustness and the generalization capacity of the models, a cross-validation strategy was applied during the hyperparameter optimization phase. For short-term models, characterized by their strong sequential and autoregressive dependence, the standard implementation of the TimeSeriesSplit from the scikit-learn library was adopted. This technique specifically addresses the fundamental limitation of traditional k-fold methods when applied to time series data: the violation of temporal dependency and the subsequent risk of data leakage. By avoiding the random shuffling of observations, this method imposes a strict chronological constraint, ensuring that the training set invariably precedes the validation set in each successive fold. For long-term models, a 5-fold method was utilized. This approach enabled the evaluation of the model's consistency across various climatic regimes distributed throughout the historical series, ensuring that the learning of physical temperature patterns was not biased by any single, specific period.

Although LSTM networks are natively architected to utilize sequential dependencies through historical lags, a deliberate methodological choice was made to omit temporal sequences in the long-term forecasting scenario. This strategy was implemented to ensure a standardized baseline, allowing all models, including traditional machine learning algorithms like RF and XGBoost, to be evaluated under identical input conditions. Consequently, when removed of its sequential context, the LSTM underperformed relative to the tree-based ensembles. Without the capacity to exploit recurrent temporal dynamics, the network operates essentially as a feedforward regressor, a regime where robust non-linear estimators such as RF and XGBoost

inherently exhibit superior performance.

When it comes to hyperparameter tuning, traditional techniques such as grid search and random search have been widely employed in the literature. Although useful, these methods present significant limitations, including high computational costs and low efficiency in high-dimensional search spaces (Bergstra e Bengio, 2012). Modern hyperparameter optimization methods address these shortcomings by enabling a more intelligent and adaptive exploration of the search space, thereby reducing training time and improving predictive performance (Shahriari et al., 2016).

Therefore, automated hyperparameter tuning for each model was conducted using the Optuna framework, an independent and open-source optimization library, and the Tree-structured Parzen Estimator (TPE) sampling algorithm to efficiently navigate the search space. The objective function was integrated with the previously described cross-validation strategy. During each trial, the model was instantiated and subjected to the complete training and testing cycle across the five data splits, with performance scored as the arithmetic mean of the MAE across all folds. The algorithm then sought to minimize this global average, ensuring that the selected hyperparameters provided the most generalizable configuration rather than one optimized for a single data partition. Once all trials were complete, Optuna identified the hyperparameter combination yielding the lowest MAE. The final model was then retrained on the full training set (2022–2023) using these optimal parameters and applied to the independent test dataset (2024) to generate predictions and compute the final evaluation metrics. All simulations were performed on a desktop PC equipped with an Intel Core Ultra 9 285K CPU, 64 GB of RAM, a 1 TB SSD, and an NVIDIA GeForce RTX 5070 GPU.

2.4.3 MACHINE LEARNING REGRESSION MODELS

This study benchmarks five machine learning algorithms: Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest (RF), Long Short-Term Memory networks (LSTM), and eXtreme Gradient Boosting (XGBoost). The following subsections briefly describe each algorithm and outline its architecture.

To ensure the full reproducibility of this research and facilitate future studies, the entire computational framework was developed in Python (version 3.10) utilizing widely adopted, open-source libraries. Data preprocessing, standard scaling, and traditional machine learning algorithms—specifically Support Vector Regression (SVR) and Random Forest (RF)—were implemented using the scikit-learn library. The gradient boosting models were built utilizing the official xgboost package, while the deep learning architectures, including the multilayer Artificial Neural Network (ANN) and the Long Short-Term Memory (LSTM) network, were constructed, trained, and optimized using the PyTorch framework. The automated search for hyperparameters was orchestrated via the Optuna library.

2.4.3.1 **Artificial neural networks**

Artificial Neural Networks (ANNs) are bio-inspired computational models structured as layers of interconnected neurons that perform parallel, adaptive processing (Basheer e Hajmeer; Gurney, 2000, 2018). Their primary appeal lies in the ability to approximate complex non-linear functions without requiring explicit knowledge of the underlying physical relationships. In the feedforward architecture adopted here, information propagates in a single direction from input to output through one or more hidden layers, with each neuron applying a weighted linear combination of its inputs followed by a non-linear activation function (Chen et al., 2017).

Table 3 presents the best hyperparameters and architecture of the ANN model.

Table 3 – Best hyperparameters - ANN

Hyperparameter	**ANN Long Term**	**ANN Short Term**
hidden_layers	3	4
neurons_per_layer	195	145
learning_rate	0.00395686	0.00032744
batch_size	64	64
epochs	386	426
dropout_rate	0.101793	0.275978
weight_decay	0.00000198	0.00013436
activation_function	relu	selu

Source: Prepared by the author (2026).

2.4.3.2 **Support Vector Regression**

Support Vector Regression (SVR) is the regression variant of Support Vector Machines (SVM), a class of supervised learning algorithms grounded in statistical learning theory (Murphy; Smola e Schölkopf, 2012, 2004). SVR maps input features into a high-dimensional space via a kernel function and seeks a regression hyperplane that minimizes prediction error subject to a tolerance margin; errors within this margin are ignored, while those exceeding it are penalized symmetrically, making SVR robust to outliers (Awad e Khanna, 2015).

Table 4 presents the best hyperparameters and architecture of the SVR model.

Table 4 – Best hyperparameters - SVR

Hyperparameter	**SVR Long Term**	**SVR Short Term**
kernel	rbf	rbf
C	0.328434	0.589018
epsilon	0.000690	0.018065
gamma	float	auto

Source: Prepared by the author (2026).

2.4.3.3 **Random Forest**

Random Forest (RF) is an ensemble method that aggregates predictions from a large number of independently trained decision trees, each constructed on a bootstrap sample of the training data with a random subset of features considered at each split (Breiman, 2001). This double randomization substantially reduces variance compared to a single tree and confers strong resistance to overfitting, even in relatively high-dimensional settings (Hastie, Tibshirani e Friedman, 2009).

Table 5 presents the best hyperparameters and architecture of the RF model.

Table 5 – Best hyperparameters - Random Forest

Hyperparameter	**RF Long Term**	**RF Short Term**
n_estimators	547	298
max_depth	21	23
min_samples_leaf	14	10
min_samples_split	7	11
criterion	mse	absolute_error

Source: Prepared by the author (2026).

2.4.3.4 **Long Short-Term Memory**

Long Short-Term Memory (LSTM) networks are a specialized class of Recurrent Neural Networks designed to capture temporal dependencies across both short and extended time horizons (Hochreiter e Schmidhuber, 1997). Their gated memory-cell architecture (comprising input, forget, and output gates) selectively retains or discards information across time steps, overcoming the vanishing-gradient problem that limits standard RNNs (Al-Selwi et al.; Pawar, Jalem e Tiwari, 2024, 2019).

Table 6 presents the best hyperparameters and architecture of the LSTM model.

Table 6 – Best hyperparameters - LSTM

Hyperparameter	**LSTM Long Term**	**LSTM Short Term**
lstm_layers	2	1
lstm_units	436	36
learning_rate	0.001479	0.00130684
batch_size	64	128
epochs	214	266
dropout_rate	0.355742	0.241065
weight_decay	0.00000411	0.00066178

Source: Prepared by the author (2026).

2.4.3.5 **Extreme Gradient Boosting**

Extreme Gradient Boosting (XGBoost) is an optimized, regularized implementation of the gradient boosting framework that constructs an ensemble of decision trees sequentially, with each tree correcting the residual errors of its predecessors (Lu et al., 2020). Its sparsity-aware split-finding algorithm and block-structured data storage enable efficient training on large datasets while maintaining strong generalization performance (Chemura et al., 2020).

Table 7 presents the best hyperparameters and architecture of the XGBoost model.

Table 7 – Best hyperparameters - XGBoost

Hyperparameter	**XGBoost Long Term**	**XGBoost Short Term**
n_estimators	2310	917
learning_rate	0.0202128	0.00557550
max_depth	3	3
subsample	0.7186169	0.837761
colsample_bytree	0.9588725	0.929141
reg_alpha	0.0291913	9.557877
reg_lambda	0.0020609	2.448901

Source: Prepared by the author (2026).

2.4.4 PERFORMANCE METRICS

Selecting appropriate metrics is essential for evaluating the quality of regression model predictions. The metrics employed in this study are the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2). Each metric captures a different aspect of model behavior: MAE and MAPE are highly interpretable and express error in intuitive units, MSE and RMSE penalize large errors more heavily, and R^2 reflects the model's overall explanatory power. Analyzing these metrics jointly therefore provides a more complete and balanced picture of predictive performance.

MAPE measures the average prediction error in percentage terms, making it particularly useful for interpreting results relative to the actual values (Hyndman e Athanasopoulos, 2018), and is calculated using Eq. (2.5):

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{2.5}$$

MAE computes the average absolute difference between predicted and observed values, as defined in Eq. (2.6). Because it does not square the errors, MAE is less sensitive to outliers and provides a direct measure of average error magnitude (James et al., 2013).

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2.6}$$

MSE computes the average of the squared errors, thereby penalizing larger deviations more severely, as shown in Eq. (2.7). Its mathematical convenience makes it particularly common in model optimization (James et al., 2013).

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2.7}$$

RMSE is the square root of the MSE, expressing error in the same units as the original data, as given by Eq. (2.8). Like MSE, it penalizes large errors heavily, but its interpretation is more intuitive (James et al., 2013).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2.8}$$

The R^2 represents the proportion of the variance in the dependent variable that is predictable from the independent variables, as defined in Eq. (2.9). It serves as a key indicator of the model's fit, where a value closer to 1 implies a stronger correlation between the observed and predicted data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2.9}$$

In all equations, n is the number of observations, y_i is the actual value for observation i, and $\hat{y}_i$ is the corresponding value predicted by the model.

2.5 RESULTS

2.5.1 LONG-TERM PREDICTION RESULTS

A linear regression model (baseline) was initially tested to verify whether a simple model would be sufficient to solve the problem. As shown in Table 8, the baseline model achieved an MAE of 3.5341 and an R^2 of 0.7711, significantly underperforming in comparison to the machine learning models studied.

The performance metrics for the long-term prediction strategy, detailed in Table 8, indicate that the machine learning models maintained robust accuracy, even as the temporal horizon increased. The ANN model achieved the lowest MAE (2.0978) and the lowest MAPE, whereas XGBoost attained the highest coefficient of determination (R^2) of 0.8652, together with the lowest MSE and RMSE. ANN and XGBoost therefore demonstrated competitive and highly similar overall performances, consolidating themselves as the most effective options for the long-term approach.

The SVR and LSTM models exhibited inferior performance in this category. The SVR scored an MAE of 2.5964 and an R^2 of 0.8294, while the LSTM obtained the highest error indices in the table, with an MAE of 2.7204 and an R^2 of 0.8050. These data suggest that, for the analyzed dataset, ANN, XGBoost and RF are better-suited to capture long-term temporal patterns using ambient temperature and relative humidity as input data.

Table 8 – Long-Term Prediction Metrics including Baseline

Model	**MAE**	**MAPE (%)**	**MSE**	**RMSE**	**R^2**
ANN	**2.0978**	**6.5913**	**11.9057**	**3.4505**	**0.8571**
LSTM	2.7204	8.8101	16.2488	4.0310	0.8050
RF	2.4220	7.0848	12.0923	3.4774	0.8549
SVR	2.5964	7.6064	14.2193	3.7708	0.8294
XGBoost	2.2589	6.8045	11.2358	3.3520	0.8652
Baseline	3.5341	11.2700	19.0770	4.3680	0.7711

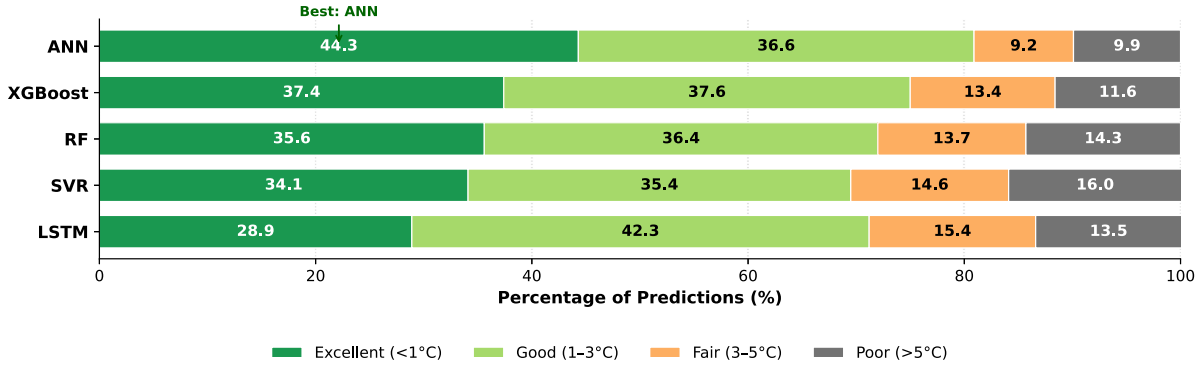
Source: Prepared by the author (2026).

Figure 8 presents the absolute error distribution for each predictive model in the long-term approach, categorized into four quality levels. A key finding is that all five models concentrate the vast majority of their predictions within the two most favorable classes (Excellent and Good, corresponding to errors below 3°C). Specifically, the ANN model achieves the highest combined proportion in this range, with 80.9% of its predictions falling below the 3°C threshold, followed by XGBoost (75.0%), RF (72.0%), SVR (69.5%), and LSTM (71.2%). This indicates that, regardless of the algorithm chosen, approximately seven out of ten predictions deviate from the measured rail temperature by less than 3°C, which is a encouraging result given that only ambient temperature and relative humidity are used as input features. When the analysis focuses on the most precise class (errors below 1°C), the ANN model stands out with 44.3%, substantially ahead of XGBoost (37.4%), RF (35.6%), SVR (34.1%), and LSTM (28.9%). Conversely, regarding predictions classified as Poor (errors exceeding 5°C), the ANN again presents the best result with the lowest incidence (9.9%), while SVR exhibits the highest proportion of critical errors (16.0%). It is also noteworthy that the LSTM, despite presenting the lowest Excellent rate, achieves the largest Good share (42.3%), suggesting that its errors tend to remain moderate rather than extreme. Overall, these results confirm the ANN as the most suitable model for the long-term prediction strategy, offering the best balance between high precision and low incidence of critical deviations.

Fig. 9 presents the comparative performance of the models in two distinct scenarios: situations where the real temperature exceeds the critical threshold of 45°C, and situations with temperatures below this value. It is observed that the models' accuracy is significantly higher when the actual temperature is below 45°C, with correct prediction rates (right) ranging between 93.7% and 98.9%. However, in the challenging scenario where actual temperatures exceed 45°C, the models' effectiveness declines. Nevertheless, ANN and LSTM stand out positively in this regard, achieving correct prediction rates of 87.6% and 71.0%, respectively. This indicates that, although LSTM yields slightly higher MAE than the ANN, it possesses a good capacity to correctly identify extreme heat peaks in long-term forecasts.

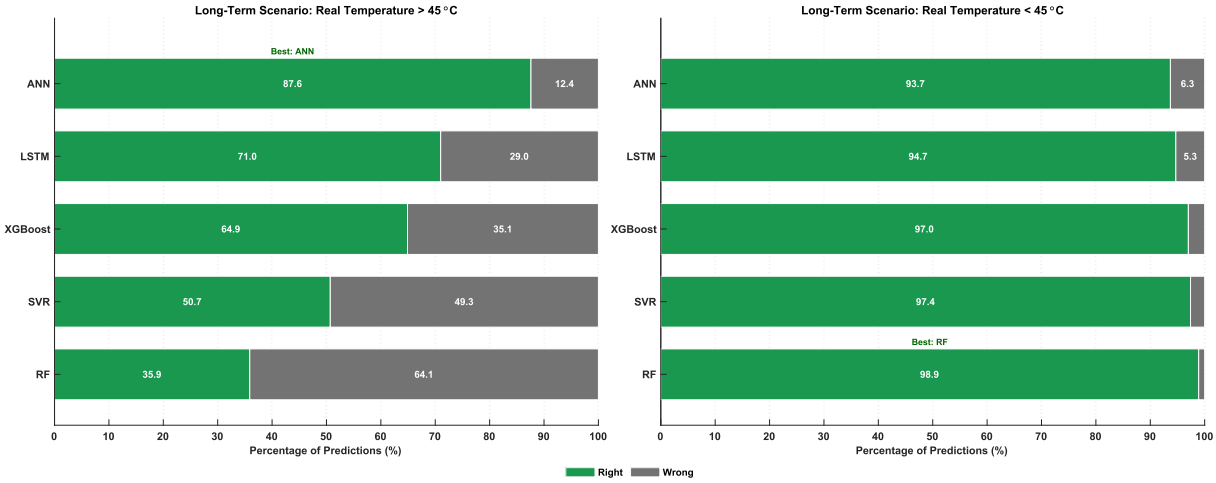
Figure 10 presents the results obtained for the long-term prediction of rail temperature. It can be observed that the XGBoost, SVR, RF, LSTM, and ANN models show a consistent trend in tracking the actual temperature, successfully capturing the daily temperature fluctuations throughout February 2024. On days when the rail temperature remained low, the predictions

Figure 8 – Absolute error ratio for long-term prediction model.



Source: Prepared by the author (2026).

Figure 9 – Real Temperature Scenarios Higher and Lower than 45ºC for long-term prediction model.



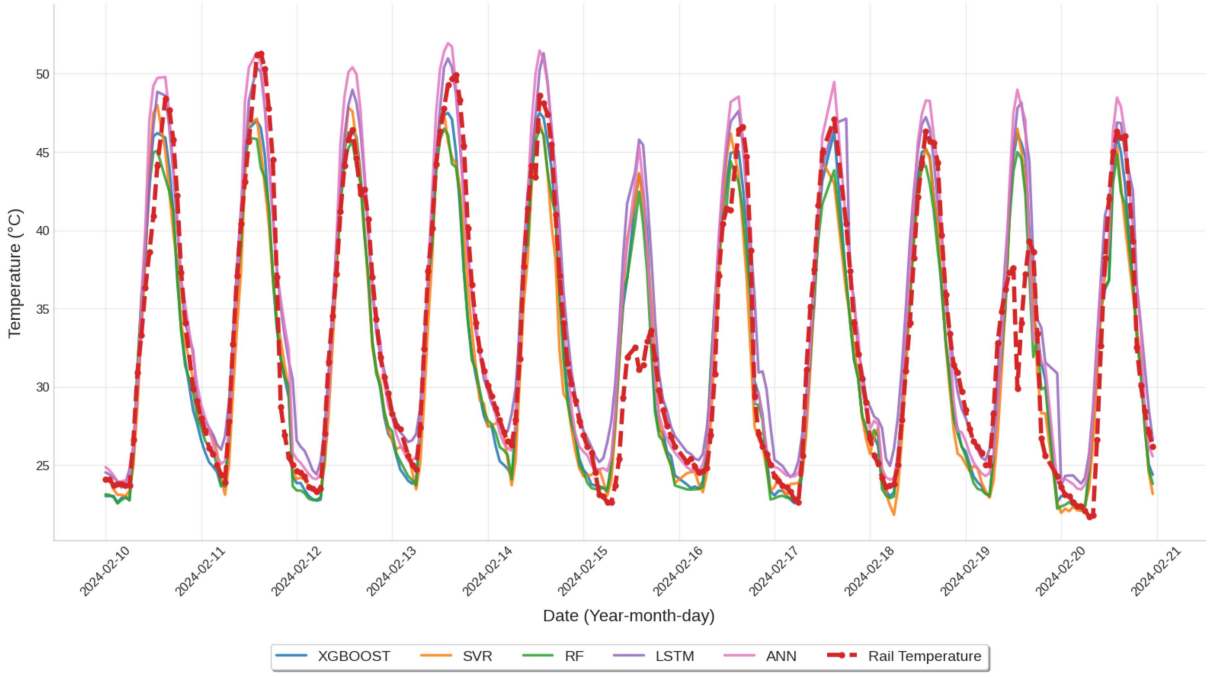
Source: Prepared by the author (2026).

were slightly higher, although they did not exceed 45ºC. This does not necessarily represent a problem for the company's management, as it did not result in false positives for rail issues such as buckling. Furthermore, it is noteworthy that, despite the inherent complexity of the long-term horizon, the visual metrics indicate a significant adherence of the models regarding the thermal peaks observed in the rail.

2.5.2 SHORT-TERM PREDICTION RESULTS

The performance metrics obtained from the short-term prediction process are presented in Table 9. The data were compiled to evaluate the accuracy of various statistical and machine learning models across a three-hour horizon.

Figure 10 – Rail temperature predictions - Long Term.



Source: Prepared by the author (2026).

Table 9 – Short-Term Prediction Metrics

Model	**Step**	**MAE**	**MAPE (%)**	**MSE**	**RMSE**	**R^2**
ANN	1	1.3469	3.9065	5.1744	2.2747	0.9379
XG Boost	1	1.3318	3.8030	5.3335	2.3094	0.9360
LSTM	1	1.2800	3.7345	5.0887	2.2558	0.9390
RF	**1**	**1.1944**	**3.4558**	**5.0412**	**2.2453**	**0.9395**
SVR	1	1.2282	3.6184	5.2380	2.2887	0.9372
ANN	2	1.9280	5.6038	10.0448	3.1694	0.8795
XG Boost	2	2.0819	6.0356	10.6298	3.2603	0.8725
LSTM	2	2.0395	5.9637	10.7175	3.2738	0.8714
RF	**2**	**1.8960**	**5.5335**	**10.1370**	**3.1839**	**0.8784**
SVR	2	2.0062	5.9478	11.0188	3.3195	0.8678
ANN	3	2.5295	7.3050	15.2437	3.9043	0.8171
XG Boost	3	2.6327	7.6888	15.3859	3.9225	0.8154
LSTM	3	2.6015	7.5614	15.7516	3.9688	0.8110
RF	**3**	**2.4178**	**7.0724**	**14.6071**	**3.8219**	**0.8247**
SVR	3	2.5155	7.4927	15.6769	3.9594	0.8119

Source: Prepared by the author (2026).

The analysis of the performance metrics for short-term prediction reveals that the RF model exhibited the highest overall accuracy among all tested models, consistently delivering the best results across the three time horizons. At the +1h horizon (Step 1), the RF achieved the lowest MAE of 1.1944 and the highest coefficient of determination (R^2) of 0.9395, demonstrating

a superior ability to map short-term thermal fluctuations compared to deeper architectures like ANN and LSTM.

The LSTM and SVR models demonstrated the least satisfactory performance for this temporal scale at the furthest horizon. At the +3h horizon (Step 3), the LSTM recorded the lowest R^2 (0.8110), while the MAE reached 2.6015. The LSTM's performance still struggles to outperform simpler ensemble methods in the short term. The thermal inertia of the rail may be sufficiently captured by the local patterns that RF explores through its decision trees. Despite the optimization, recurrent models like LSTM often require larger datasets or more complex feature engineering to significantly surpass the efficiency of shallow learners in very short horizons.

A clear decrease in performance is observed as the forecasting horizon extends from +1h to +3h. For instance, the RF model's MAE increased from 1.1944 at Step 1 to 2.4178 at Step 3, while its R^2 decreased from 0.9395 to 0.8247. This trend is consistent across all models, reflecting the natural increase in uncertainty as the prediction moves further from the last known observation. Despite this error propagation, the ANN remains a solid alternative, maintaining a competitive R^2 of 0.8171 at the +3h horizon, positioning itself alongside RF as the most reliable models for short-term thermal monitoring.

As depicted in Fig. 11, the absolute error distribution for short-term horizons reinforces the high precision of the models, with a notable performance by the RF model in the most accurate range. At the +1h horizon, the RF model concentrates 67.4% of its predictions within the error up to 1°C, surpassing the ANN and LSTM, which present 61.4% and 62.4% respectively in the same category. As the horizon extends to +3h, this precision naturally declines across all architectures, with the RF maintaining a leading 43.5% of predictions within the 1°C margin, while the ANN and SVR follow with 40.5% and 41.5% respectively.

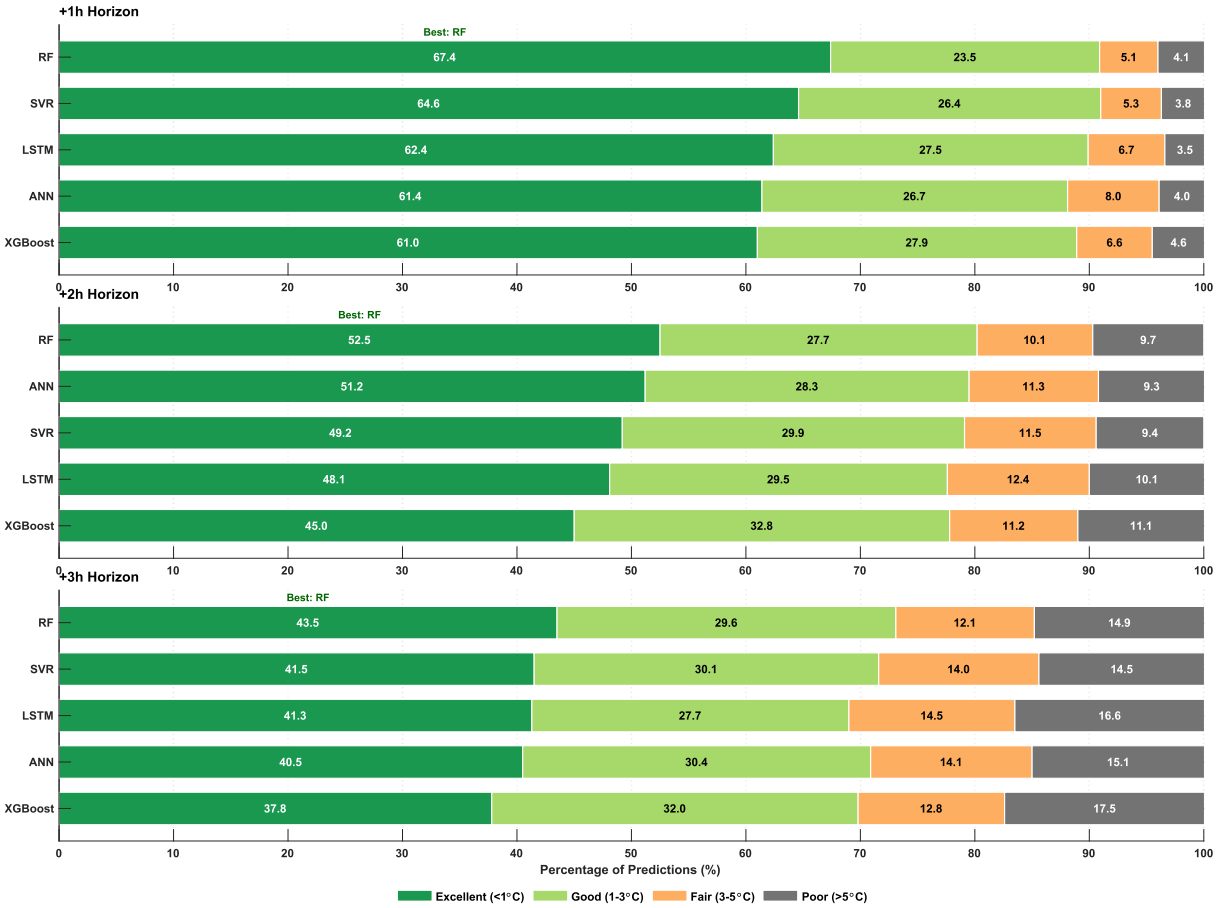
As the prediction horizon extends to +3h, precision naturally declines across all architectures due to accumulating uncertainties. Nevertheless, the models' capacity to sustain reliable predictions within the acceptable ≤ 3°C threshold remains robust. At the +3h horizon, the RF model maintains a leading 73.1% of its predictions within the 3°C margin (43.5% in the Excellent range and 29.6% in the Good range). The SVR and ANN architectures follow a similar resilient trend, retaining 71.6% and 70.9% of their errors within this green zone, respectively, while the XGBoost model exhibits a performance at 69.8%.

The operational reliability analysis, focusing on the incidence of errors ranges above 5°C, reveals a progressive increase in severity as the prediction horizon advances. At the +1h horizon, all models demonstrate high reliability, with critical error rates remaining below 5% . However, at the +3h horizon, these rates increase significantly. The XGBoost model exhibits the highest incidence of severe deviations at the furthest horizon, reaching 17.5%, followed by the LSTM with 16.6% and the ANN with 15.1%.

When comparing the intermediate error profiles, the RF and SVR models show consistent behavior, with a combined frequency of errors below 2°C exceeding 80% at the +1h horizon (82.9% for RF and 82.4% for SVR). In contrast, at the +3h horizon, the models present a more

dispersed distribution. The RF and SVR continue to show slightly better stability in critical ranges compared to XGBoost, which presents a more relevant portion of errors in the range above 5°C. This analysis confirms that while all models are highly effective for the immediate hour, the RF and ANN offer a more balanced distribution for longer short-term horizons, minimizing the frequency of severe failures.

Figure 11 – Absolute error ratio for short-term prediction model.



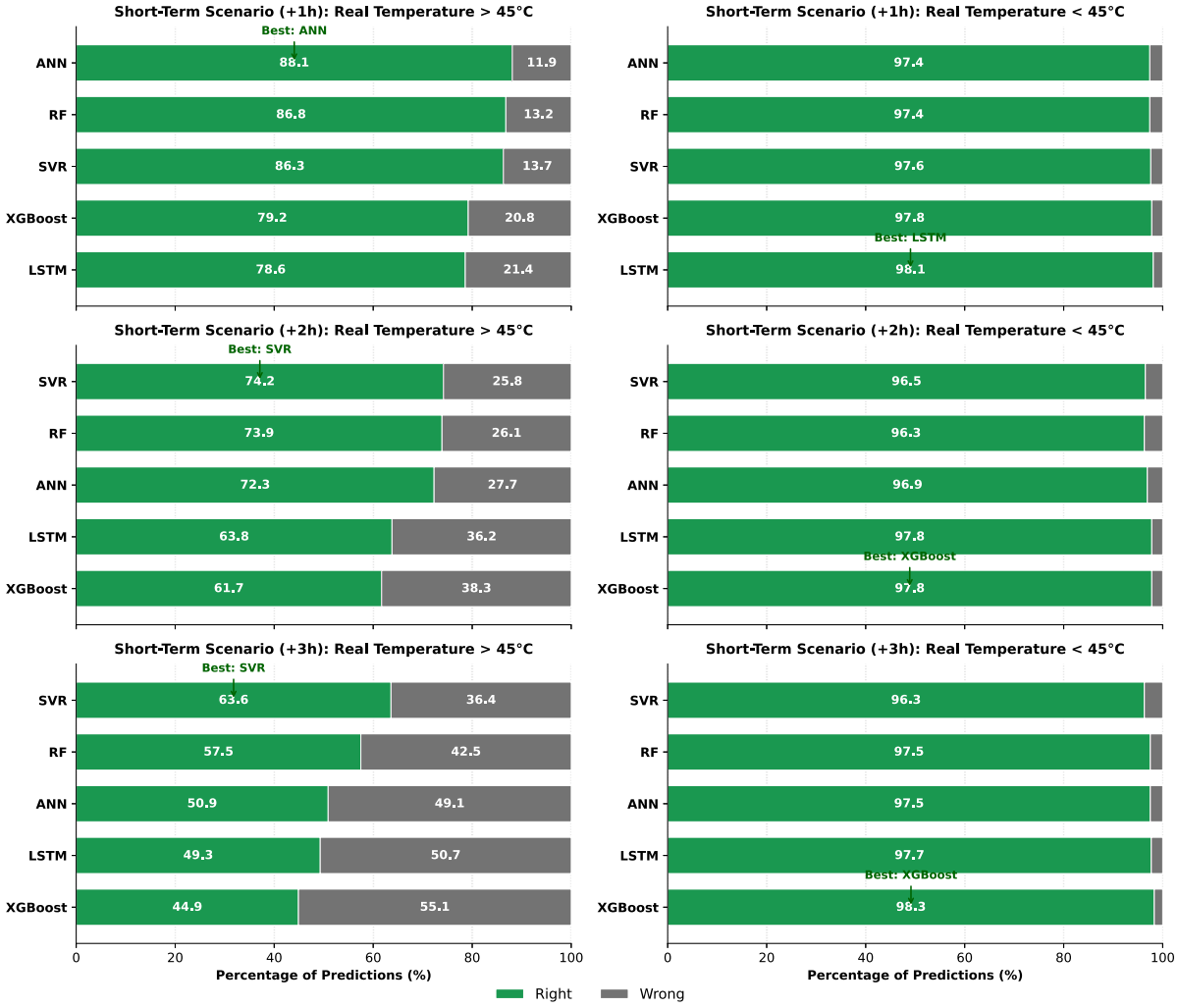
Source: Prepared by the author (2026).

The evaluation of accuracy in extreme scenarios in Figure 12 reveals distinct model behaviors when the real temperature exceeds the 45°C threshold. The ANN model, exhibiting the best overall metrics, achieved a correct prediction rate of 88.1% at the +1h horizon for the 45°C scenario. This performance is superior to that of the LSTM (78.6%) and better than XGBoost, which achieved 79.2% success rate in this critical scenario.

In scenarios where temperatures are below 45 °C, most models exhibit a near-perfect correct prediction rate, frequently exceeding 97%. In the more demanding scenario above 45 °C, however, the ANN leads at the +1h horizon with a correct prediction rate of 88.1%, followed by RF (86.8%) and SVR (86.3%), whereas XGBoost (79.2%) and LSTM (78.6%) are found behind. This pattern persists across the +2h and +3h horizons, where RF and SVR retain the highest accuracy in the extreme range. These results indicate that, for short-term forecasting,

the ANN, RF, and SVR models offer the most balanced and precise behavior under both typical and extreme conditions.

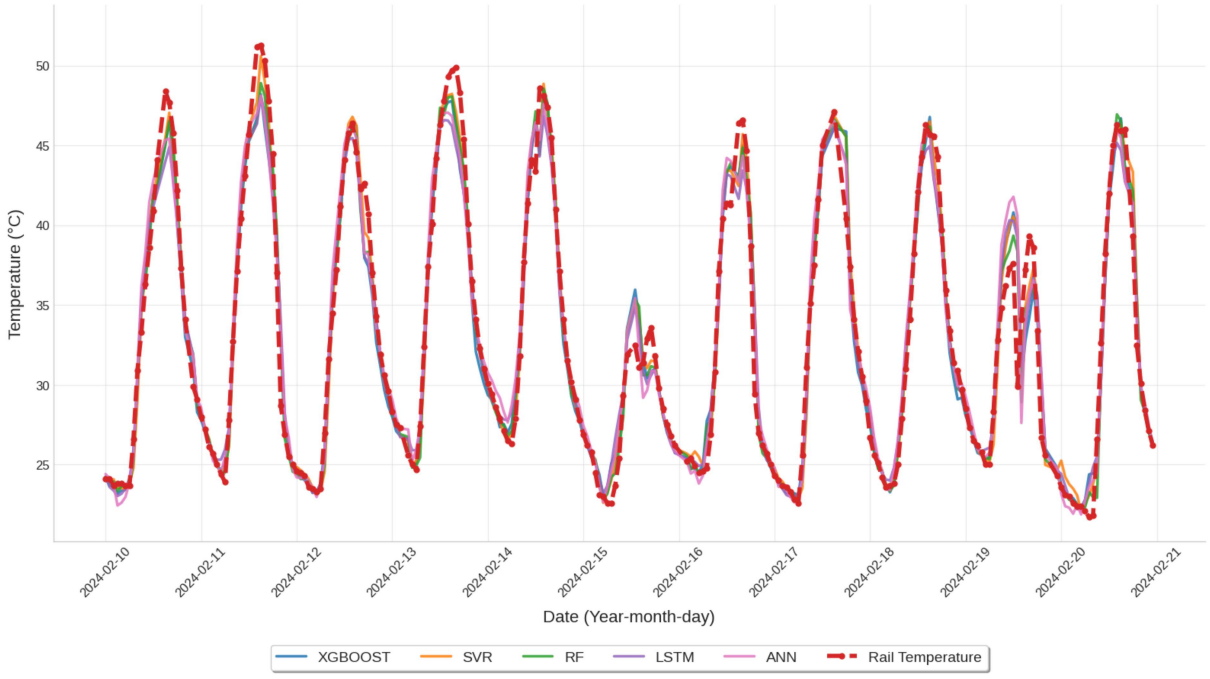
Figure 12 – Real Temperature Scenarios for short-term prediction



Source: Prepared by the author (2026).

Figure 13 illustrates the performance of the algorithms for the short-term (+1h horizon) strategy. In this scenario, the proximity between the prediction curves and the actual rail temperature becomes even more evident, demonstrating that the reduction of the forecasting window enhances the accuracy of the machine learning models. For the short-term scenario, the models were able to more closely approximate lower temperatures, effectively avoiding false positives for structural issues such as buckling.

Figure 13 – Rail temperature predictions - Short Term.



Source: Prepared by the author (2026).

2.5.3 BENCHMARKING AGAINST EXISTING MODELS

Table 10 provides a comparative literature review of rail temperature prediction models, summarizing their performance in terms of MAE, RMSE, MAPE, and R^2, along with the forecast horizon adopted in each study.

Table 10 – Performance Metrics of Rail Temperature Forecasting Models Observed in the Literature

Model	**Ref**	**MAE**	**RMSE**	**MAPE**	R^2	**Forecast**
Bayesian LSTM	(Pham et al., 2025)	0.090 ºC	2.210 ºC	-	0.930	3h
ANN (Case 1)	(Zhou et al., 2024)	3.400 ºC	4.350 ºC	-	0.720	10 min
ANN (Case 2)	(Zhou et al., 2024)	4.550 ºC	5.500 ºC	-	0.540	10 min
ANN (Case 3)	(Zhou et al., 2024)	3.200 ºC	3.950 ºC	-	0.760	10 min
ANN (Case 4)	(Zhou et al., 2024)	2.850 ºC	3.600 ºC	-	0.810	10 min
ANN (Case 5)	(Zhou et al., 2024)	2.550 ºC	3.250 ºC	-	0.840	10 min
CNN (Case 1)	(Zhou et al., 2024)	2.180 ºC	2.820 ºC	-	0.620	10 min
CNN (Case 2)	(Zhou et al., 2024)	2.050 ºC	2.680 ºC	-	0.660	10 min
CNN (Case 3)	(Zhou et al., 2024)	1.750 ºC	2.220 ºC	-	0.760	10 min
CNN (Case 4)	(Zhou et al., 2024)	1.750 ºC	2.350 ºC	-	0.740	10 min
CNN (Case 5)	(Zhou et al., 2024)	2.380 ºC	2.980 ºC	-	0.780	10 min
LSTM (Case 1)	(Zhou et al., 2024)	2.980 ºC	3.850 ºC	-	0.770	10 min
LSTM (Case 2)	(Zhou et al., 2024)	2.980 ºC	3.860 ºC	-	0.760	10 min
LSTM (Case 3)	(Zhou et al., 2024)	2.420 ºC	3.100 ºC	-	0.850	10 min
LSTM (Case 4)	(Zhou et al., 2024)	2.760 ºC	3.470 ºC	-	0.820	10 min
LSTM (Case 5)	(Zhou et al., 2024)	2.300 ºC	2.930 ºC	-	0.870	10 min
Back Propagation (BP)	(Liu et al., 2023)	1.514 ºC	2.014 ºC	5.560%	-	Daily
GA-BP	(Liu et al., 2023)	1.328 ºC	1.651 ºC	5.040%	-	Daily
Back Propagation (BP)	(Liu et al., 2023)	1.700 ºC	2.467 ºC	11.808%	-	Daily
GA-BP	(Liu et al., 2023)	1.411 ºC	2.081 ºC	8.414%	-	Daily
PR2	(Hong, Park e Cho, 2021)	0.693 ºC	2.316 ºC	-	0.971	64h (3h)
SVM	(Hong, Park e Cho, 2021)	0.068 ºC	2.135 ºC	-	0.972	64h (3h)
ANN	(Hong, Park e Cho, 2021)	0.519 ºC	1.732 ºC	-	0.984	64h (3h)
RF	(Hong, Park e Cho, 2021)	0.029 ºC	0.685 ºC	-	0.997	64h (3h)
XGBoost 1	(Hong, Park e Cho, 2021)	0.008 ºC	0.518 ºC	-	0.998	64h (3h)
XGBoost 2	(Hong, Park e Cho, 2021)	0.043 ºC	1.744 ºC	-	0.982	64h (3h)

Source: Prepared by the author (2026).

Now, the performance metrics obtained from the long-term and short-term models (Tables 8 and 9, respectively) are benchmarked against existing models found in the literature (Table 10).

It should be noted that the models reviewed in Table 10 were developed under different experimental conditions and forecasting horizons and, most importantly, with more input feature sets. The near-optimal errors reported by Hong et al. (Hong, Park e Cho, 2021) (e.g., MAE of 0.008°C for XGBoost) were obtained using up to eight meteorological variables, including solar irradiance, wind speed, and cloud cover, which are not always available in many railway networks. The following comparison therefore does not aim to directly rank models, but rather to illustrate the performance trade-off between input complexity and predictive accuracy.

2.5.3.1 Long-Term Models

This section presents a comparative analysis between the long-term predictive results obtained in the present study, Table 8, and established benchmarks from the literature, which are summarized in Table 10.

- ANN: consistently outperforms all five cases from ANN, CNN and LSTM (Zhou et al., 2024), including their best result of R^2 of 0.840, achieving an MAE of 2.0978°C and an R^2 of 0.8571. It should be emphasized that the inclusion of rail temperature as an input variable in Zhou et al. (2024) likely provided a performance advantage, given that this information was unavailable to the other models under comparison. It nonetheless remains less precise than the ANN benchmark (Hong, Park e Cho, 2021) with an MAE of 0.519°C and R^2 of 0.984.

- XGBoost: achieves an MAE of 2.2589°C and an R^2 of 0.8652, surpassing the majority of neural network configurations reported by Zhou et al. (2024). However, it does not approach the precision of XGBoost 1 (Hong, Park e Cho, 2021).

- RF: delivers robust performance with an MAE of 2.4220, outperforming the standard cases from (Zhou et al., 2024). Despite this, it does not reach the high precision recorded by (Hong, Park e Cho, 2021), consistent with the pattern observed across all models in this study.

- SVR: with an MAE of 2.5964°C, outperforms the primary ANN (Case 1) and LSTM (Case 1) configurations (Zhou et al., 2024) but falls short of the SVM benchmark (Hong, Park e Cho, 2021).

- LSTM: achieves an MAE of 2.7204°C, significantly outperforming the LSTM (Case 1) (Zhou et al., 2024) (MAE of 2.980°C) and approaching the performance of its more complex configurations.

2.5.3.2 Short-Term Models

This section presents a comparative analysis between the short-term predictive results obtained in the present study, Table 9, and established benchmarks from the literature, which are summarized in Table 10.

- ANN: outperforms all five configurations presented by Zhou et al. (2024), where even the best-performing scenario (Case 5) yielded an MAE of approximately 2.550°C and an R^2 of 0.840 for a much shorter 10-minute forecast. In contrast, the proposed ANN achieves an MAE of 1.3469°C and 2.5295°C for the +1h and +3h horizons, respectively. However, it remains less precise than the ANN reported by (Hong, Park e Cho, 2021), which achieved an MAE of 0.519°C and an R^2 of 0.984 for a 3-hour horizon, as that study utilized a broader and more complex set of meteorological inputs.

- XGBoost: achieves an MAE of 1.3318°C and an R^2 of 0.9360 for the +1h horizon. While these results demonstrate high reliability for immediate forecasts, the error indices increase to an MAE of 2.6327°C at the +3h horizon. This performance does not approach the near-zero error (MAE of 0.008°C) reported for XGBoost 1 (Hong, Park e Cho, 2021), a discrepancy that can be attributed to the significantly more parsimonious input feature set used in this work, prioritized for ease of operational implementation.

- RF: delivers the best overall performance among the proposed models, with an MAE of 1.1944°C and an R^2 of 0.9395 at the +1h horizon. This yields results in a range that is more accurate than all cases reported by Zhou et al. (2024), including their best LSTM and CNN cases. Nevertheless, its error at the +3h horizon (MAE 2.4178°C) remains above the benchmark (MAE of 0.029°C) established by Hong et al. (2021), reflecting the inherent trade-off between model simplicity and absolute precision.

- SVR: demonstrates a balanced performance across horizons, with an MAE ranging from 1.2282°C (+1h) to 2.5155°C (+3h). This performance surpasses the ANN and LSTM configurations (Zhou et al., 2024), which reported MAE values of approximately 3.400°C and 2.980°C for very short-term predictions. It falls short of the SVM (MAE of 0.068°C) from Hong et al. (2021) and the Bayesian LSTM (MAE of 0.090°C) reported by Pham et al. (2025), models which rely on more computationally intensive frameworks or additional input parameters.

- LSTM shows a performance transition from an MAE of 1.2800°C (+1h) to 2.6015°C (+3h), with the R^2 declining from 0.9390 to 0.8110. While this updated architecture is significantly more competitive than previous versions, it is still outperformed by the Bayesian LSTM with MAE of 0.090°C (Pham et al., 2025). Furthermore, most LSTM configurations in Zhou et al. (2024) yield MAE values between 2.300°C and 2.980°C for 10-minute forecasts, indicating that the proposed LSTM maintains comparable or superior precision even for the significantly longer +3h horizon.

2.6 CONCLUSIONS

The performance analysis across different temporal scales reveals that the RF model achieved the highest accuracy in short-term forecasting, particularly at the +1h horizon with an MAE of 1.1944°C and R^2 of 0.9395. In long-term forecasting, the ANN model stood out as the most accurate architecture, presenting an MAE of 2.0978°C and an R^2 of 0.8571. XGBoost

delivered competitive results across both settings, while SVR and LSTM showed varying degrees of stability depending on the temporal window. In the long-term approach, the LSTM recorded the highest error indices among all evaluated algorithms, being outperformed by SVR on every aggregate metric; its relative strength was confined to the detection of rail temperatures above 45 °C. This indicates that recurrent architectures require dedicated tuning to effectively capture thermal dependencies under the current input configuration.

Beyond general accuracy, a critical finding emerges when considering extreme-temperature scenarios. In the short-term horizon, models such as LSTM and RF demonstrated high reliability for immediate intervals; in the long-term setting, ANN and LSTM led the detection of rail temperatures above 45°C, achieving a correct prediction rate of 87.6% and 71.0% respectively. This performance trade-off is particularly relevant for railway safety systems and should guide model selection according to the primary operational objective.

A central contribution of this work is demonstrating that meaningful rail temperature predictions, with R^2 reaching 0.94 and MAE as low as 1.19°C for short-term forecasts, can be obtained from a minimal set of inputs consisting of ambient temperature and relative humidity. While the proposed models do not reach the absolute predictive accuracy of systems trained on eight or more meteorological variables, they achieve a favorable balance between performance and deployability. This reflects a deliberate design choice aligned with the operational constraints of railway infrastructure management, particularly in tropical and developing regions, where the consistent acquisition of high-fidelity environmental data from a wide array of sensors is often logistically unfeasible.

Furthermore, the utility of this streamlined input configuration extends beyond general predictive performance to the robust detection of critical thermal extremes. The proposed ANN models demonstrated remarkable efficacy in forecasting high-risk rail temperatures exceeding 45 °C. The ANN achieved an accuracy of 87.6% in long-term predictions for these extreme events, while maintaining a high accuracy of 88.1% for short-term (+1h) forecasts. These results underscore that a minimal meteorological dataset is not only sufficient for routine monitoring but is also capable of anticipating dangerous temperature peaks, thereby providing infrastructure managers with a low-cost tool that can support established safety procedures.

The demonstrated framework can therefore support operational decisions, including heat-related speed restrictions and track buckling risk assessment, even under limited data conditions.

2.6.1 Limitations and Future Work

Despite these contributions, the present study has limitations that point to relevant directions for future work. The models were trained and evaluated within a specific geographic and climatic context, and their generalization to other railway networks warrants further investigation.

Additionally, although solar irradiance is the dominant driver of rail heating during peak hours, it was not used as an explicit input; its influence is instead captured indirectly through the temporal features (hour of day and day of year). These features encode only the mean

(climatological) diurnal and seasonal solar cycle and cannot represent the day-to-day variability in solar load caused by cloud cover and clear-sky conditions. Because the most extreme rail temperatures (above 45 °C) typically occur on clear-sky days, this is a likely contributor to the reduced predictive accuracy observed for such events (Sections 2.5 and 2.6). Incorporating a direct measure of solar load—such as measured or forecast solar irradiance, or a clear-sky index—therefore represents a promising direction for improving the detection of critical thermal extremes.

Future efforts will focus on targeted strategies to enhance predictive performance within this critical thermal range, including the exploration of class-imbalance mitigation techniques, threshold-specific loss functions, and the incorporation of supplementary meteorological or infrastructure variables capable of improving model sensitivity to heat extremes.

2.7 ACKNOWLEDGMENTS

The authors would like to thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) – grants 407256/2022-9, 303550/2025-2, 402533/2023-2 and 303982/2022-5, FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) – grants APQ-00032-24 and APD-01113-25, and DNIT (Departamento Nacional de Infraestrutura de Transportes) – through the Termo de Execução Descentralizada (TED) UFJF/DNIT, for their financial support.

2.8 REFERENCES

AGUSTIN, D.; WU, Q.; NGAMKHANONG, C. A review on railway track buckling prediction methods. *Construction and Building Materials*, v. 466, p. 140295, 2025.

AL-SELWI, S. M. et al. Rnn-lstm: From applications to modeling techniques and beyond—systematic review. *Journal of King Saud University-Computer and Information Sciences*, v. 36, n. 5, p. 102068, 2024.

AWAD, M.; KHANNA, R. *Efficient Learning Machines: Theories, Concepts and Applications for Engineers and System Designers*. 1. ed. Berkeley, CA: Apress, 2015.

BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, v. 43, n. 1, p. 3–31, 2000.

BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. *The Journal of Machine Learning Research*, v. 13, n. 1, p. 281–305, 2012.

BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001.

BRUZEK, R. et al. Rail temperature prediction model and heat slow order management. In: AMERICAN SOCIETY OF MECHANICAL ENGINEERS. *Proceedings of the ASME/IEEE Joint Rail Conference*. [S.l.], 2014. p. V001T01A014.

CHAPMAN, L. et al. Modelling of rail surface temperatures: a preliminary study. *Theoretical and Applied Climatology*, v. 92, n. 1–2, p. 121–131, 2008.

CHEMURA, A. et al. The impact of land-use/land cover changes on water balance of the heterogeneous Buzi sub-catchment, Zimbabwe. *Remote Sensing Applications: Society and Environment*, v. 18, p. 100292, 2020.

CHEN, M. et al. *Machine learning for wireless networks with artificial intelligence: A tutorial on neural networks*. 2017.

CHO, M. et al. A data-driven approach for rail temperature estimation from air temperature, solar irradiation, and land surface temperature. *Sustainable Cities and Society: Advances*, v. 2, n. 2, p. 100038, 2026. ISSN 3051-052X.

EPL - Empresa de Planejamento e Logística S.A. *Plano Nacional de Logística 2035*. Brasília, DF, 2021. Disponível em: <https://portal.epl.gov.br/plano-nacional-de-logistica-2035>.

FERRANTI, E. et al. The hottest July day on the railway network: insights and thoughts for the future. *Meteorological Applications*, v. 25, n. 2, p. 195–208, 2018.

GURNEY, K. *An introduction to neural networks*. Boca Raton: CRC Press, 2018.

HAN, S. H. et al. Numerical study on the characteristics of temperature distribution in continuous welded rail by solar radiation and rail orientation. *Journal of Mechanical Science and Technology*, v. 34, n. 11, p. 4819–4829, 2020.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The elements of statistical learning: data mining, inference, and prediction*. 2. ed. New York: Springer, 2009.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural Computation*, v. 9, n. 8, p. 1735–1780, 1997.

HONG, S.; PARK, C.; CHO, S. A rail-temperature-prediction model based on machine learning: warning of train-speed restrictions using weather forecasting. *Sensors*, v. 21, n. 13, p. 4606, 2021.

HONG, S. U. et al. A rail-temperature-prediction model considering meteorological conditions and the position of the sun. *International Journal of Precision Engineering and Manufacturing*, v. 20, n. 3, p. 337–346, 2019.

HUNT, G. *An Analysis of Track Buckling Risk*. London, UK, 1994.

HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. Melbourne: OTexts, 2018.

ILOS - Instituto de Logística e Supply Chain. *Matriz de transporte de carga (% TKU)*. 2024. Disponível em: <https://prime.ilos.com.br/>.

JAMES, G. et al. *An introduction to statistical learning: with applications in R*. New York: Springer, 2013.

JIE, H. et al. Defect detection and classification of railway track system for in-service mrt in tropical regions using a contactless tbms and adaptive-dbscan. *IEEE Transactions on Intelligent Transportation Systems*, v. 26, n. 11, p. 18695–18707, 2025.

JúNIOR, L. T. D. et al. The trajectory of data-driven structural health monitoring: A review from traditional methods to deep learning and future trends for civil infrastructures. *Computer Modeling in Engineering & Sciences*, Tech Science Press, v. 146, n. 2, p. 1–10, 2026. ISSN 1526-1506. Disponível em: <http://dx.doi.org/10.32604/cmes.2026.075433>.

LIU, D. et al. Temperature field test and prediction using a GA-BP neural network for CRTS II slab tracks. *Railway Engineering Science*, v. 31, n. 4, p. 381–395, 2023.

LU, H.-p. et al. Short-term prediction of building energy consumption employing an improved extreme gradient boosting model: A case study of an intake tower. *Energy*, v. 203, p. 117756, 2020.

MELO, R. et al. Out-of-roundness wheel damage identification in railway vehicles using autoencoder models. *Applied Sciences*, MDPI AG, v. 15, n. 5, p. 2662, mar. 2025. ISSN 2076-3417. Disponível em: <http://dx.doi.org/10.3390/app15052662>.

MURPHY, K. P. *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: MIT Press, 2012.

OU, Z.; LI, F. Analysis and prediction of the temperature field based on in-situ measured temperature for CRTS-II ballastless track. *Energy Procedia*, v. 61, p. 1290–1293, 2014.

PALIN, E. J. et al. Implications of climate change for railway infrastructure. *Wiley Interdisciplinary Reviews: Climate Change*, v. 12, n. 5, p. e728, 2021.

PAWAR, K.; JALEM, R. S.; TIWARI, V. Stock market price prediction using lstm rnn. In: *Emerging Trends in Expert Applications and Security*. Germany: Springer, 2019. p. 493–503.

PHAM, K. et al. An AIoT system for real-time monitoring and forecasting of railway temperature. *Journal of Civil Structural Health Monitoring*, v. 15, n. 3, p. 915–925, 2025.

PILOTO, P. A. G. et al. Validation of a rail temperature model with experimental measurements. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, v. 236, n. 9, p. 1104–1113, 2022.

SHAHRIARI, B. et al. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, v. 104, n. 1, p. 148–175, 2016.

SMOLA, A. J.; SCHöLKOPF, B. A tutorial on support vector regression. *Statistics and Computing*, v. 14, n. 3, p. 199–222, 2004.

VIDAL, R.; CURY, A.; BARBOSA, F. Model development and efficient calibration of railway transition zones using experimental data. *Railway Engineering Science*, p. 1–20, 2026.

WU, Y. et al. Rail temperature prediction model. In: *Proceedings of the CORE 2012 Global Perspectives Conference on Railway Engineering*. Brisbane, Australia: [s.n.], 2012. Disponível em: <https://acquire.cqu.edu.au/articles/conference_contribution/Rail_temperature_prediction_model/13461110>.

ZHANG, Y.-J.; LEE, S. Modeling rail temperature with real-time weather data. In: *Transportation Research Circular E-C126: Fourth National Conference on Surface Transportation Weather and Seventh International Symposium on Snow Removal and Ice Control Technology*. Indianapolis, IN, USA: Transportation Research Board, 2008. p. 37–48.

ZHOU, R. et al. Comparing various neural network methods for temperature prediction of crts ii slab track on transition sections. *Intelligent Transportation Infrastructure*, v. 3, p. liae004, 02 2024. ISSN 2752-9991.

3 FINAL CONSIDERATIONS

This research demonstrates that accurate results can be achieved using a reduced set of input features, such as ambient temperature and humidity. This strategy balances predictive power with operational feasibility, offering a cost-effective solution for railway safety and maintenance planning in regions where complex meteorological data collection is challenging.

The performance analysis across different time scales indicates that the Random Forest (RF) model is the most effective for short-term forecasting, while Artificial Neural Networks (ANN) provide the highest accuracy for long-term predictions. Regarding extreme temperature detection, ANN an LSTM proved most reliable for long-term monitoring, whereas LSTM and RF models excelled in short-term intervals.

Several directions are identified for future development of the proposed framework. The most immediate extension is the incremental incorporation of solar irradiance data, either from on-site pyranometers or from freely available satellite-based products, to evaluate whether this single additional variable meaningfully reduces prediction error during peak heating hours without compromising the operational simplicity of the system. A second priority is the validation of the models across geographically distinct segments of the Brazilian railway network to assess model transferability across different climatic regimes. From an architectural standpoint, bidirectional LSTM networks and attention-based temporal models represent promising candidates to address the limitations observed in this study, particularly for the detection of extreme temperature events.

Finally, unifying the short-term and long-term models into a single software platform and integrating it with existing track infrastructure management systems would represent the critical step from research prototype to operational tool, enabling maintenance teams to receive threshold-based, early-warning alerts for heat-related speed restrictions and track buckling risk.