A Rail-Temperature-Prediction Model Based on Machine Learning: Warning of Train-Speed Restrictions Using Weather Forecasting

Sunguk Hong [ID], Cheoljeong Park [ID] and Seongjin Cho * [ID]

School of Mechanical Engineering, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, Korea; hsu12375@gmail.com (S.H.); pffiro@gmail.com (C.P.)
* Correspondence: scho@cnu.ac.kr; Tel.: +82-42-821-5648

Abstract: Predicting the rail temperature of a railway system is important for establishing a rail management plan against railway derailment caused by orbital buckling. The rail temperature, which is directly responsible for track buckling, is closely related to air temperature, which continuously increases due to global warming effects. Moreover, railway systems are increasingly installed with continuous welded rails (CWRs) to reduce train vibration and noise. Unfortunately, CWRs are prone to buckling. This study develops a reliable and highly accurate novel model that can predict rail temperature using a machine learning method. To predict rail temperature over the entire network with high-prediction performance, the weather effect and solar effect features are used. These features originate from the analysis of the thermal environment around the rail. Precisely, the presented model has a higher performance for predicting high rail temperature than other models. As a convenient structural health-monitoring application, the train-speed-limit alarm-map (TSLAM) was also proposed, which visually maps the predicted rail-temperature deviations over the entire network for railway safety officers. Combined with TSLAM, our rail-temperature prediction model is expected to improve track safety and train timeliness.

Keywords: intelligent transportation system (ITS); machine learning; rail temperature; buckling; XGBoost; structural health monitoring

1. Introduction

Rail temperature is important for rail safety. High rail temperature is a direct cause of buckling on railway tracks. Buckling is the result of excessive deformation in a high rail-temperature environment. When severe, it can derail the train. Although train derailment by buckling is an infrequent event, it causes catastrophic casualties of human lives and property [1–4]. Moreover, the air temperature, which is closely related to rail temperature, is continuously increasing under global warming effects, further enhancing the risk of buckling.

The recent expansion of high-speed trains has increased the demand for continuous welded rails (CWRs), which reduce vibration and noise to offer a comfortable riding experience. Unfortunately, CWRs are also vulnerable to buckling because they are welded together and lack the space to expand. Therefore, controlling the buckling and monitoring the rail temperature is imperative for railway safety management. To prevent buckling-induced train derailment, current railway companies issue train-speed-limit orders based on the rail temperature, which is monitored in real time.

Train-speed-limit orders are directly related to the track safety and train timeliness. If the orders for preventing track buckling (for example, limiting the train speed or spraying the tracks with water) are implemented without planning, the sudden adjustment of the train-operating schedule might cause bottlenecks or traffic congestion. For example, during the hottest July day on record in the UK, there were 12,800 heat-related delay minutes

and an additional 23,700 min caused by unplanned train-speed restrictions [5,6]. If the rail temperature in an area can be predicted beforehand, schedulers can preplan the speed limit of the train, adjust the intervals between trains, or cool the rail by water spraying, thereby improving train timeliness and track safety. To achieve this goal, researchers have developed rail-temperature-prediction models (RTPMs) that predict the rail temperature based on the local weather conditions around the rail.

The previous RTPMs are classified into three types: empirical equation-based models, multivariate regression models, and thermal analysis models [7–13].

RTPMs based on empirical equations simply predict the rail temperature as a linear function of air temperature. These models are easy to use because they predict the rail temperature from a single parameter, but deliver lower performance (R^2 = 0.9021, root mean square error (RMSE) = 5.866 °C) than other RTPMs [8,9].

Multivariate regression RTPMs predict the rail temperature not only from the air temperature, but also from other weather conditions such as wind speed and cloud cover. Such RTPMs generally outperform those based on empirical equations. Wu et al. developed a multivariate regression RTPM called the Bureau of Meteorology (BoM) prediction equation (1–24 h), which delivers the highest performance to date (R^2 = 0.9630, RMSE = 2.560 °C) but requires 24 features. Therefore, this model is applicable only to specific countries or environments, and is not easily generalized [10].

Finally, thermal analysis RTPMs thermodynamically model the environment around the actual rail. Since thermodynamic models consider the laws of nature, they are more generalizable than RTPMs based on statistical methods, which tend to require specific data. However, to guarantee high performance (R^2 = 0.9334, RMSE = 3.799 °C), such models require precise knowledge of the rail properties, such as the reflectance and emissivity properties [7,10–14].

The previous RTPMs are hampered by two limitations. First, their performance in predicting the rail temperature is too low for practical use. In previous RTPMs based on thermodynamic principles (R^2 = 0.9334, RMSE = 3.799 °C), the predicted rail temperature deviated by up to 15 °C from the measured rail temperature [7]. Since the main goal of RTPMs is preventing buckling on hot days, this difference is unacceptable in practical use. Furthermore, the criteria of train speed limits are shifted to a 4–5 °C change in rail temperature, seriously weakening the reliability of RTPMs [15,16].

Second, the previous RTPMs predicted the rail temperature at a single point. This approach is not meaningful because actual rails are continuously interconnected over several hundred kilometers. To develop RTPMs for practical uses, the range of the rail-temperature prediction must be broadened, for example, from small-town connectivity (local networks) to state- or country-wide connectivity (entire network).

The next generation of RTPMs for practical uses will require high performance RTPM and mapping applications, and will quickly present the rail temperatures from small networks to entire networks.

Recently, machine learning approaches such as artificial neural network (ANN), support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGBoost) have attracted great interest by developers of high-performance regression models. Machine learning allows a computer to learn the relationship between the data (the input) and the results (the output). For instance, a regression model that predicts the room temperature and daily maximum air temperature by machine learning delivers higher prediction performance than other types of methods [17,18].

Herein, we propose a machine learning-based rail temperature prediction model (RTPM) with the highest performance to date (maximum R^2 = 0.9984, RMSE = 0.518 °C) that can predict the rail temperature over an entire network. The method, called Chungnam National University RTPM (CNU RTPM), outperforms the previous RTPMs for predicting high rail temperatures (over 40 °C). The CNU RTPM performance is due to the selected features obtained by analyzing the thermal environment around the rail. With these

features, the CNU RTPM can predict the rail temperature over the entire network using weather forecast data alone.

Additionally, a structural health-monitoring application, called the train-speed-limit alarm-map (TSLAM), was developed, which shows the deviation of the predicted rail temperature over the entire network, enabling quick searching of danger regions. This application is easily combined with the CNU RTPM. TSLAM is also available worldwide because it uses the global weather forecast data, not merely a particular region's data.

2. Measurement

2.1. Measurement of Rail Temperature and Local Weather Conditions

Direct sampling of the actual rail environment during train operation is impractical for safety reasons. Rail temperatures are usually measured indirectly within a measurement station that simulates the environment of the railway installation [7,10,11,13,19]. As in previous work, we collected the rail temperature and weather data every 10 min at a measurement station from August of 2016 to May of 2017 [7,20]. The constructed measurement station consisted of a 500-mm long KS 50n rail, a data acquisition system (DAQ), a weather station, and K-type thermocouples. The station was installed at a low-traffic site of Chungnam National University (CNU). CNU is located in Daejeon in Korea (latitude: 36.36°, longitude: 127.34°). A schematic of the constructed measurement station is shown in Figure 1 and a photograph of the installed measurement system is shown in Figure A1.
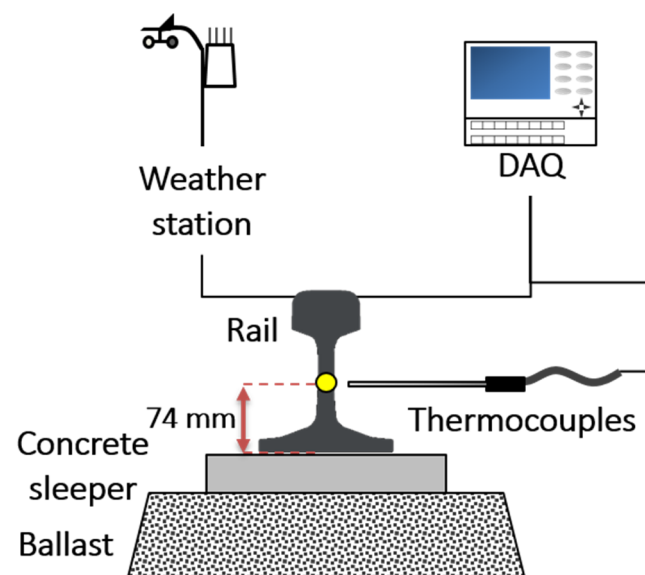

Figure 1. Schematic of the measurement system.

The rail of the measurement system was installed on a ballast-and-concrete sleeper oriented in the south–north direction, which is minimally influenced by shadow. As the actual rail is shadowed by mountains, trees, soundproof walls, and other objects, its temperature is lower than that of the test rail [21]. By minimizing the influence of shadows, a higher temperature is guaranteed in the measurement system than in the shaded area of the real system. Since the purpose of this study was to improve the safety of train operation, we deliberately constructed a conservative measurement system. A model based on conservative measurement data makes conservative predictions.

2.2. Measurement of Local Weather Conditions

Around the world, weather forecasts report the air temperature, relative humidity, rainfall amount, wind speed, and cloud cover. Using these weather factors as features, we can construct a globally available rail-temperature-prediction model. As confirmed in previous studies, the weather factors provided by weather forecasts are decisive predictors

of rail temperature. Indeed, a conventional empirical RTPM predicts the rail temperature from the air temperature alone [8,9]. The air temperature and rail temperature are highly correlated.

The weather station (Vantage Pro2, Davis, CA, USA) installed in the measurement system measures the local weather data (air temperature, relative humidity, rainfall amount, and wind speed). The measured local weather data are transmitted to the DAQ. The measured environmental data were acquired once at every 10-min intervals, because it takes about 10-min to achieve thermal equilibrium in the rail measurement system [7]. The rainfall amount is difficult to measure at a specific moment, so it was obtained by summing the values measured over 10 min. The cloud cover, which cannot be measured by the measurement system, was borrowed from the data of the Korea Meteorological Administration (KMA), located 2.4 km from the measurement station.

2.3. Measurement of Rail Temperature

To develop our novel RTPM, measuring the rail temperature in summer when rail buckling usually occurs is needed. However, in summer, the internal and surface rail temperatures can differ by as much as 7 °C, because the sun rises to a higher altitude than in winter, and the rail surface receives a large amount of solar radiation [7,19,22]. Such deviations in rail temperature constitute a noise in the predictive regression model. We overcome this problem by setting a representative point as follows.

Recently, we showed that the deformation of the KS 50N rail at 74 mm from the bottom of the rail represents the average deformation of the whole rail [20]. In this study, the rail temperature at the point of average deformation of the KS 50N rail was taken as the representative temperature of the whole rail. To measure the rail temperature at this point, K-type thermocouple probes were inserted by drilling.

3. Feature Study: Feature Selection Based on the Thermal Analysis

3.1. Modeling of the Thermal Environment around the Rail

The rail is exposed to complex conduction, convection, and radiation processes that depend on the weather factors (such as the air temperature, wind speed, and cloud) and solar factors (such as solar irradiance) (see Figure 2) [7,11,12]. RTPMs based on the thermal analysis typically predict the rail temperature by balancing the heat transfers [7,11–13]. The energy balance equation is given as

$$\dot{E}_{sun} - \left(\dot{E}_{conv} + \dot{E}_{rad} \right) = 0, \tag{1}$$

where $\dot{E}_{sun}$ is the heat flux of the global solar irradiance and $\dot{E}_{conv}$ and $\dot{E}_{rad}$ are the heat fluxes of convection and radiation, respectively.
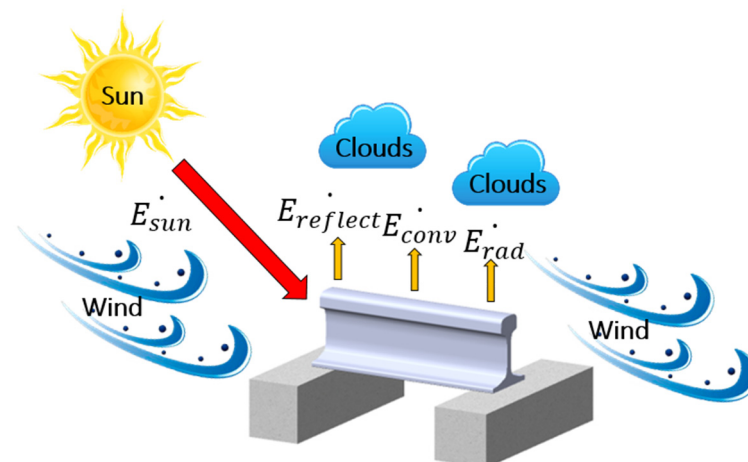

Figure 2. Thermal environment around the rail.

$\dot{E}_{sun}$ is a major determinant of rising rail temperature and must be modeled in detail [23]. When solar irradiance moves through the atmosphere, it is partially lost to absorption, reflection, and diffusion by moisture, dust, and clouds [24]. If the RTPM is built without considering the detailed solar irradiance, it may predict high rail temperatures inaccurately during the daytime. The poor performance of high rail-temperature prediction is a fatal weakness because high rail temperatures cause rail buckling, especially during the summer daytime.

$\dot{E}_{conv}$ is closely related to wind speed [7,13]. Generally, the higher the wind speed, the better the heat exchange between the atmosphere and rail surface, considering that the wind speed is high and the air and rail temperatures are similar.

$\dot{E}_{rad}$ is closely related to the air temperature above the cloud and rail surface emissivity [7]. Based on the Stefan–Boltzmann law, $\dot{E}_{rad}$ increases to the fourth power of the absolute temperature [25]. In the study of rail temperatures, $\dot{E}_{rad}$ is of less importance compared to other heat transfer mechanisms. Typically, rail temperatures range from -10 to $60\,^\circ$C. In this temperature range, $\dot{E}_{rad}$ is negligible compared to that of other heat transfer mechanisms.

Previous RTPMs based on the regression analysis focused only on the weather factors and did not properly reflect the relationship between solar irradiance and rail temperature, which is critical for rail-temperature prediction. For precise rail-temperature prediction, the rail–sun relationship (solar-effect features) was treated as important while developing the RTPM based on machine learning. Based on the thermal analysis and preliminary studies, the features were classified into weather (air temperature, rainfall amount, wind speed, cloud cover, and relative humidity) and solar effect (azimuth, altitude, and total solar irradiance (TSI)) features.

3.2. Weather Features

Most previous regression analysis-based RTPMs predicted rail temperatures considering only weather conditions [8–10]. The Hunt models predicted the rail temperature using a single feature (air temperature), whereas the BoM prediction equation used 24 features, including the temperature of the Earth's surface and pressure [8,10].

When building machine learning-based RTPMs, including many weather-features account for various weather situations, but risks the curse of dimensionality and increases the difficulty of data acquisition. To avoid these problems, the relevant features mainly affecting rail temperature must be selected from several weather factors.

Among the weather features, we selected the air temperature, rainfall amount, wind speed, cloud cover, and relative humidity. The effects of the above weather features on the rail temperature were reported in previous studies, and are summarized below.

Air temperature: A change in the air temperature will always change the rail temperature [12]. As mentioned earlier, the simplest RTPM (Hunt models) uses only the air temperature as a feature [8].

Rainfall amount: Moisture is crucial in suppressing the maximum rail temperature because of the high specific heat of water ($4.184\,\mathrm{J \cdot g^{-1} \cdot K^{-1}}$)) [13].

Wind speed: As mentioned earlier, wind speed is closely related to convection. When the wind speed increases, the difference between the air temperature and rail temperature decreases. This effect is enhanced in summer, when the difference between the rail and air temperatures is already large [11,13].

Cloud cover: Cloud cover is closely related to solar irradiance. On cloudy days with high cloud cover, the portion of lost solar irradiance increases and the ratio of diffuse horizontal irradiance (DHI) to global solar irradiance (GHI) increases accordingly [13].

Relative humidity: A change in the relative humidity slightly changes the rail temperature [13]. In Chapman's model, the relative humidity is assumed to be closely related to the minimum rail temperature [13].

The suggested weather features were measured as described in Section 2 and are easily obtainable worldwide because they are common components of global weather forecasts.

By virtue of the weather features, the CNU RTPMs built by the suggested features can predict the rail temperatures over an entire network.

### 3.3. Solar Effect Features

The sun is the main heat source of the rail and causes a high rail temperature during the daytime. Therefore, the influence of the sun must be considered in the rail temperature prediction. To predict the rail temperature over the entire network, we here define the solar effects (azimuth, altitude, and TSI) as new features that can be simply calculated in terms of time (year, month, day, hour, minute, and second) and location (latitude and longitude). By using the solar effect features in the RTPM-building, the CNU RTPM can account for the influence of solar irradiance on rail temperature. The solar effect features release the temporal and location dependencies of the RTPMs, because they are physical variables that change in time and space. Since the rail temperature has increased under global warming effects, RTPMs that depend on a specific time and location cannot predict the rail temperature in the coming decades or at other locations. In future RTPMs that predict the rail temperatures of entire networks, the solar effect features are essential.

3.3.1. Modeling of Solar Irradiance

The GHI is closely related to rail temperature, so it must be considered in the rail-temperature prediction. The measured rail temperature and GHI data during the summer of 2016 at the measurement station are compared in Appendix B. Note that the rail temperature and GHI exhibited similar temporal dynamics (Figure A2a) and were positively correlated (Figure A2b; R^2 = 0.6870). The phase difference in Figure A2a is attributable to the heat transfer delay caused by the heat capacity of the rail.

Clearly, the GHI is an important predictor of rail temperature and would improve the performance of RTPM. As this variable is not provided in the weather forecast, it was replaced by the TSI on the Earth's upper atmosphere. The TSI includes the solar power over all incident wavelengths per unit area and changes slowly as the Earth elliptically orbits the sun [25,26]. As the TSI depends on the distance between the Earth and the sun, it can be predicted accurately. The GHI can then be predicted from the TSI and the weather factors [26–29].

The TSI, DNI, and DHI are geometrically defined in Figure 3a. Upon entering the atmosphere, the TSI is scattered and absorbed by weather factors such as humidity and clouds. At noon on a clear day, approximately 25% of the TSI is scattered and absorbed [30].
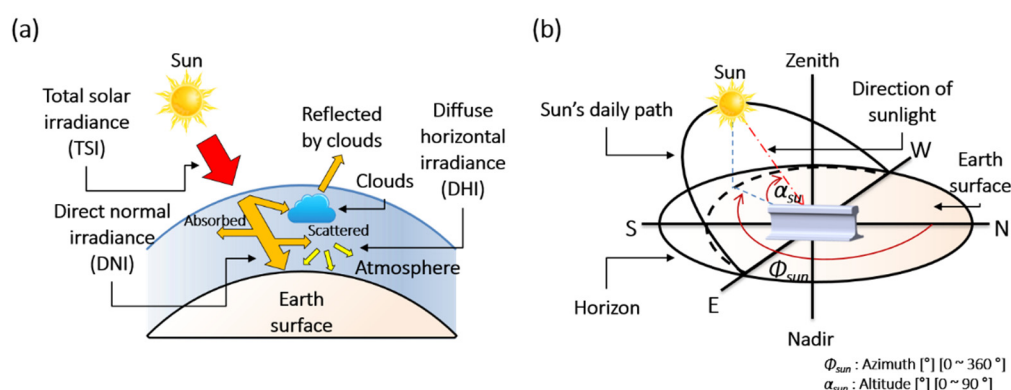
Figure 3. Geometric definitions of (**a**) total solar irradiance (TSI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI). (**b**) Azimuth and altitude in the horizontal. coordinate system.

The DNI is the solar irradiance directly arriving from the sun. The DNI is measured as the flux of the beam radiation through a plane perpendicular to the sun's direction [23].

The DHI represents the solar irradiance scattered by the atmosphere. The DHI is measured on a horizontal surface, assuming that the radiation (excluding the circumsolar radiation) enters from all points in the atmosphere.

The GHI is the total irradiance from the sun on a horizontal surface on Earth [27]. It is computed by summing the DNI and DHI.

In summary, since the TSI is the reduced GHI by weather factors, simultaneously adopting the weather factors and TSI as features in the RTPM produces the same effect as directly adopting the GHI as the feature. The TSI is calculated as described in Appendix C.

3.3.2. Direction of Solar Position

As the Earth is spherical, the TSI enters the atmosphere at an oblique angle, so the GHI (the summed DNI and DHI) varies in each region. The DNI is oriented along the direction of the sunlight, which is defined by its azimuth and altitude. The changing direction of sunlight alters the sites at which the rails receive energy directly from the sun. Therefore, even when the GHI remains constant, the rail temperature changes with azimuth and altitude [7].

In the horizontal coordinate system, the direction of the sun is expressed by its azimuth and altitude, as shown in Figure 3b [30]. The azimuth (Φ_{sun}) is the angle between the projected vector and the north (N) vector. The altitude (α_{sun}) is the angle between the sunlight vector and its projected vector on the perpendicular plane.

The azimuth and altitude are calculated in terms of time (year, month, day, hour, minute, and second) and location (latitude and longitude). The temporal terms are astrophysically meaningful for predicting the rail temperature. If the time is directly applied to the model without any post-processing, the model's performance is not guaranteed after decades or longer, and its reliability is severely degraded.

In this study, the altitude and azimuth were computed from the local latitude, longitude, and time by the Michalsky's method [31].

4. Building the CNU RTPMs

CNU RTPMs are based on machine learning and various statistical methods: extreme gradient boosting (XGBoost), support vector machine (SVM), random forest (RF), artificial neural network (ANN), and polynomial regression of order 2 (PR2). Their hyperparameters were optimized by tuning. The performances of the various CNU RTPMs were compared and the best-performing CNU RTPM was clarified. We built the models using the Python libraries: XGBoost, Scikit-learn, random forest, and Tensorflow.

4.1. Machine Learning and Statistical Methods

4.1.1. Extreme Gradient Boosting (XGBoost)

XGBoost is popularly chosen by contestants in data and machine learning competitions, owing to its higher performance than other methods [32–37]. XGBoost is suitable for regression and classification problems. At each step, the algorithm generates a weak learner and accumulates it into the total model. If the weak learner follows the gradient direction of the loss function, the learning method is called a gradient boosting machine [37].

XGBoost is an ensemble of classification and regression trees (CART). An ensemble method is a machine learning technique that combines the results of several submodels to determine the best result. Consider a dataset $D = \{(X_i, y_i) : i = 1 \cdots n, X_i \in R^m, y_i \in R\}$ with n data, m feature spaces, a target value y_i, and a predicted value $\hat{y}_i$ Let $\hat{y}_i$ be the result of an ensemble represented as follows:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(X_i), \ f_k \in \ F, \tag{2}$$

where *K* is the number of trees and f_k is a function in the functional space *F* of all CARTs.

Let L be the objective function, which consists of a training loss term and a regularization term as follows:

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k), \tag{3}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\|w\|^2. \tag{4}$$

The training loss term l measures the difference between $\hat{y}_i$ and y_i, and the regularization term Ω penalizes the complexity of the model to avoid overfitting.

The objective function is optimized by tree boosting of Equation (3). Define $\hat{y}_i^t$ as the prediction of the i-th instance at the t-th iteration. To minimize the objective function, a term f_t is added to Equation (3), giving Equation (5). Equation (5) is then simplified by the Taylor expansion to give Equation (6):

$$L^t = \sum_{i=1}^n l(y_i, \hat{y}_i^{t-1} + f_t(X_i)) + \Omega(f_t) \tag{5}$$

$$\check{L}^t = \sum_{i=1}^n [g_i f_t(X_i) + \frac{1}{2}h_i f_t^2(X_i)] + \Omega(f_t), \text{where} g_i = \partial_{\hat{y}_i^{t-1}} l(y_i, \hat{y}_i^{t-1}), h_i = \partial_{\hat{y}_i^{t-1}}^2 l(y_i, \hat{y}_i^{t-1}). \tag{6}$$

The loss reduction after the best split from a given node is given by Equation (7). Note that this function depends only on the loss function and the regularization parameter γ. Clearly, this algorithm optimizes any loss function that provides the first and second-order gradients [32].

$$L_{split} = \frac{1}{2}\left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda}\right] - \gamma. \tag{7}$$

4.1.2. Support Vector Machine (SVM)

The objective of the SVM is to find the optimal hyperplane in an N-dimensional space (where N is the number of features) that maximizes the distance between the hyperplane and the nearest data point on each side. The optimal hyperplane should distinctly classify the data points [38]. SVMs are employed in regression and classification problems, and are applicable to both linear and nonlinear data [39].

As in previous studies, our study employs an SVM for predicting the solar generation from weather forecast data. The kernel function was a radial basis function with two hyperparameters: the kernel coefficient γ and the penalty coefficient C.

4.1.3. Random Forest (RF)

An RF is a substantial modification of bagging that builds a large collection of decorrelated trees, and then averages them [40]. Bagging is one of the ensemble methods. Like SVMs, RFs are applicable to both linear and nonlinear data in regression and classification problems.

In previous studies, RFs have predicted the daily maximum air temperature from the solar radiation, albedo, latitude, longitude, and other solar-related parameters. We similarly adopt the solar effect features in temperature prediction. Here we optimize the number of variables to be split at each node (M_{try}) and the number of trees in each run (N_{tree}) as the hyperparameters.

4.1.4. Polynomial Regression of Order 2 (PR2)

PR2 is a traditional statistical method applied in regression modeling. It is suitable when the independent variable is a quadratic function of the dependent variable [38]. Additional hyperparameter settings are not required. The PR2 can be expressed as

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \epsilon_i, \tag{8}$$

where y_i is the dependent variable, x_i is the independent variable, β_i denotes the unknown parameters, ϵ_i is an error term, and i denotes a row of data.

4.1.5. ANN (Artificial Neural Network)

ANNs are powerful connectionist systems vaguely inspired by biological neural networks [41]. An ANN consists of multilayer perceptrons, and is suitable for solving various problems such as computer vision, speech recognition, and regression analyses [42].

Previously, ANNs were applied to body-temperature prediction in wearable-device studies [43]. In our study, an ANN predicts not the body (interior) temperature but the skin (surface) temperature. Specifically, we employ a feed-forward neural network with a rectified linear unit as the activation function and AdamOptimizer as the optimizer. The learning rate and batch size were 0.05 and 100, respectively. We optimize the number of nodes (one hidden layer) as the hyperparameter.

4.2. Hyperparameter Tuning by K-Fold Cross Validation

The performance of any model based on machine learning largely depends on the value(s) of the hyperparameter(s), which must be determined before running the model. In machine learning, the process of determining the hyperparameter(s) is called "tuning." [44–46]. The suitability of the value range of the hyperparameter(s) tends to depend on the user's intuition and experience.

In this study, the hyperparameters of XGBoost, SVM, and RF were optimized by K-fold cross validation, which divides the samples into training and test samples. The model was constructed using the training data and verified using the test data. K-fold cross validation simply and effectively checks whether the model has overfitted the training data.

K-fold cross validation divides the samples into k uniform groups and performs k iterations of cross validation. After cross validation based on the constructed hyperparameter values, the optimal combinations of hyperparameters are found using the GridSearchCV function in the Python library Scikit-learn. By default, GridSearchCV uses 3-fold cross validation ($k = 3$); however, applying 5-fold cross validation is conducted to more accurately optimize the hyperparameter combination.

The results of the hyperparameter tuning are provided in Appendix D.

4.3. Comparison Methods of Model Performance

In this study, 70% of the measured data were randomly selected as the training data and the remaining 30% were reserved as the test data. To account for the randomness in distributing the training and test data, the evaluation was performed five times on the test data and individual performances were averaged to obtain the final performance value. As the performances of the training and test data were not significantly different at the time of the experiment, the results of the test data are presented. The performances were described by the mean absolute error (MAE), coefficient of determination (R^2), and RMSE.

MAE is the absolute value of the average error (average difference between the original and predicted values). A lower MAE corresponds to a higher performance of the model. The R^2 measures how well the predicted values match the original values. Its value ranges from 0 (no correlation between the actual and predicted values) to 1 (perfect correlation between the two values). RMSE is another measure of the average error in the prediction. As MAE is generally used in machine learning and RMSE is conventionally used in RTPMs, both measures are computed in the present analysis. Similar to MAE, a low RMSE implies a high model performance. Conclusively, a good prediction model is characterized by low MAE, high R^2, and low RMSE. MAE, R^2, and RMSE measures are respectively expressed as follows:

$$MAE = \frac{\sum_{i=1}^{n} |y_I - \overline{y_i}|}{n}, \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \overline{y_i})^2}{\sum_{i=1}^{n} (y_i - \hat{y_i})^2} \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y_i})^2},$$

(11)

where n is the total number of test data, y_i and $\bar{y}_i$ are the measured data and the model predicted values, respectively, and $\hat{y}$ is the average output value of the test data.

The features and measurement conditions determine the reliability and versatility of the prediction model. As the model is built using the measured data, a small quantity of measured data will degrade the model reliability. Meanwhile, if the model includes an excessive number of features, it cannot easily adapt to different environments and the measurement system requires many instruments. Accordingly, models with too many features have low versatility.

5. Train-Speed-Limit Alarm-Map (TSLAM)

The actual rails are interconnected continuously for several hundred kilometers. Therefore, predicting the rail temperature in a local region is meaningless. Our TSLAM framework, developed by Python, visually presents the predicted rail temperature over the entire network. TSLAM visualizes the predicted rail temperature deviations computed by the CNU RTPMs with the selected features. Using TSLAM, train safety officers can quickly map the rail-temperature deviations and order the train company to limit the train-speed in advance.

The TSLAM can also forecast the rail temperature for up to 64 h based on the weather forecast data alone. This forecasting reduces train delays owing to the imposed train-speed limit because the train company can preadjust the train operation interval. Finally, the TSLAM improves track safety and train timeliness.

The TSLAM operates via a four-step algorithm: data acquisition, data preprocessing, rail-temperature-prediction, and data visualization (see Figure 4). The four steps of the TSLAM algorithm are detailed below.
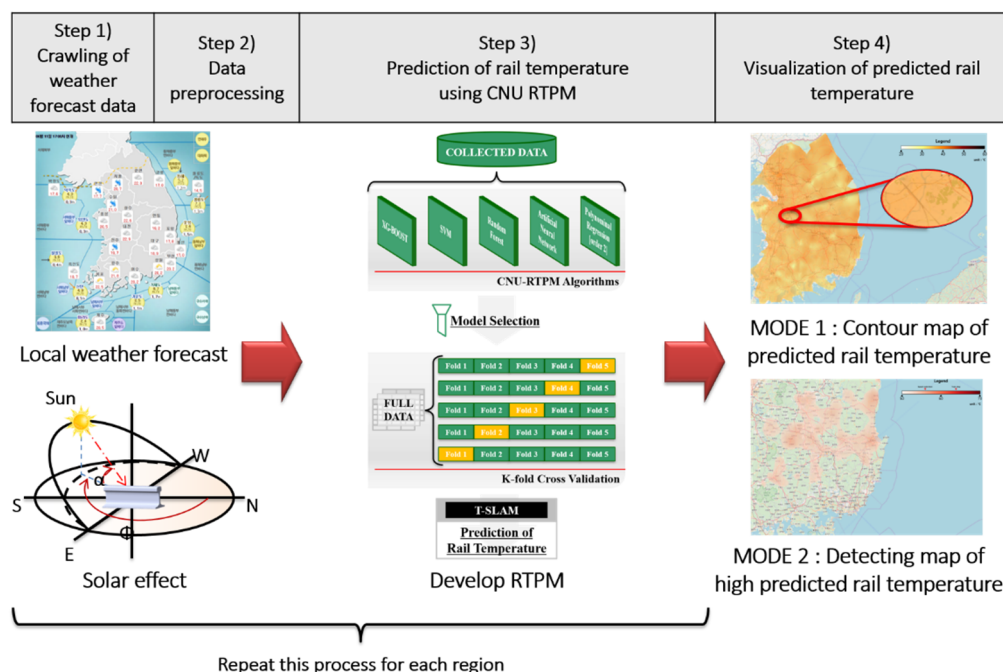

Figure 4. Framework of mapping the local rail-temperature prediction.

Step (1): Web crawling of weather forecast data: The KMA divides the Korean peninsula into (5×5)-km^2 grid squares and forecasts the weather in each area for up to 64 h at 3-h intervals. Step 1 of the TSLAM algorithm obtains the local forecast data by web crawling (defined as data extraction from a web page [47]). The obtained data comprise

the location and meteorological forecasts in each area. In this step, if the data source of the application programming interface (API) provided by the open weather map (OWM) is changed, the country and its local weather forecast data are newly obtained. Originally, we considered using the Korean local weather forecast of the OWM rather than the local weather forecast of the KMA. However, web crawling the weather forecast data of KMA, which provides a detailed local weather forecast over more regions in Korea than in the OWM, was selected.

Step (2): Data preprocessing: The KMA provides local coordinate data in its own coordinate system, which must then be converted to general-purpose latitudes and longitudes. This step is redundant if the data source is OWM's API because the OWM directly provides local latitude and longitude data. This step then computes the azimuth and altitude from the latitude and longitude at the forecast time and converts the data unit into units compatible with CNU RTPM.

Step (3): Prediction of the rail temperature using CNU RTPM: This step predicts the rail temperature from CNU RTPMs. CNU RTPM with the highest performance is then installed as the main model. Through this process, data are converted into three-dimensional data of latitudes, longitudes, and predicted rail temperatures.

Step (4): Visualization of the predicted rail temperature: This step displays the transformed three-dimensional data on a map for the user. The displayed image can be expressed in either of the following two modes, Mode 1 or Mode 2, by adjusting the range of the legend. Mode 1 is a contour map showing the predicted rail temperature in all regions of the selected country. Mode 1 compares the predicted rail-temperature deviations in each region but does not clarify the regions in which rails are dangerously deformed. Conversely, Mode 2 shows the regions of high predicted rail temperature at which the train-speed should be limited to ensure safe railway operation. These data quickly inform the train safety officer of the risky sites.

6. Results and Discussion

6.1. Data Configuration

The most important goal for building a model based on machine learning is the acquisition of numerous high-quality data. Herein, high-quality data were acquired using the measurement station. The rail temperature was determined at the point where rail deformation represented the average rail deformation. Moreover, features providing a globally usable model were selected.

Measurements were continuously collected over a 10-month period from August 2016 to May 2017, obtaining 35,252 samples. The measured data were the air temperature, relative humidity, wind speed, rainfall amount, and rail temperature. The cloud cover data, which were not measured, were instead borrowed from the data of nearby KMA. The maximum, minimum, average, and standard deviation of the dataset are included in the Appendix A. Korea's climate is characterized by a large annual range of air temperatures (-10.1–38.0 °C in the collected dataset) and by various weather events, such as monsoon rains and snowfall, which are dynamically generated in different seasons. The Korean weather data are thus considered suitable for creating globally available models.

The azimuth, altitude, and TSI were calculated by the method presented in Section 3 and added to the final samples. The resulting model features were weather features (air temperature, rainfall amount, wind speed, cloud cover, and relative humidity) and solar effect features (azimuth, altitude, and TSI).

Herein, 70% of the data (24,676 samples) were randomly selected as the training data and the remaining 30% (10,576 samples) were reserved as the test data. Figure 5 is a detailed flowchart of the methodology.
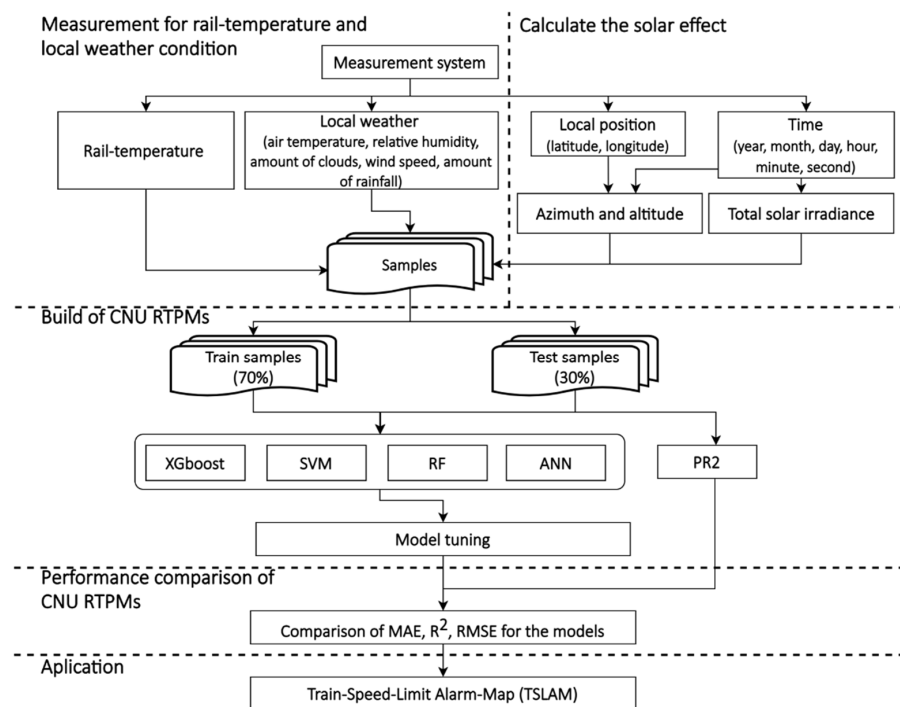
Figure 5. Flowchart of the proposed methodology.

6.2. Feature Importance

Tree-based machine learning methods such as XGBoost can calculate the importance of features. The feature importance quantitatively evaluates the effect of a particular feature on the model performance. By ranking and comparing the importance of each feature, we can remove the unnecessary features with low feature importance from the model. Such low-importance features can degrade the performance of the model.

The feature importance values computed by XGBoost are graded in Figure 6. The most important, second-most important, and third-most important features were the TSI, azimuth, and altitude, respectively. Note that these features are the solar effect features reflecting the GHI, which are obtained by simple calculations but largely affect the performance of the model. That is, the solar effect features play an important role in predicting the rail temperature.
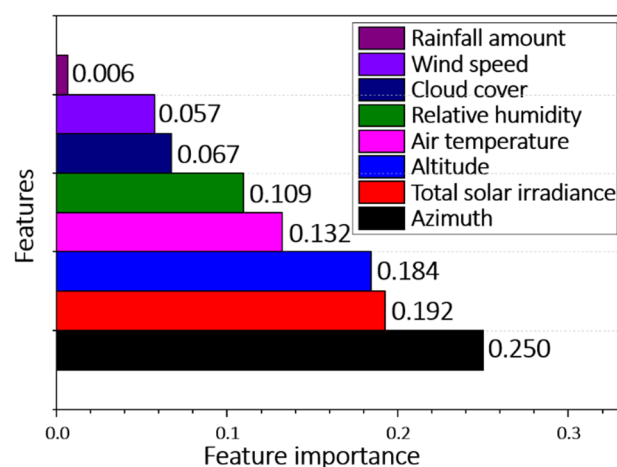

Figure 6. Graded importance of features in the rail-temperature. prediction, obtained by XGBoost.
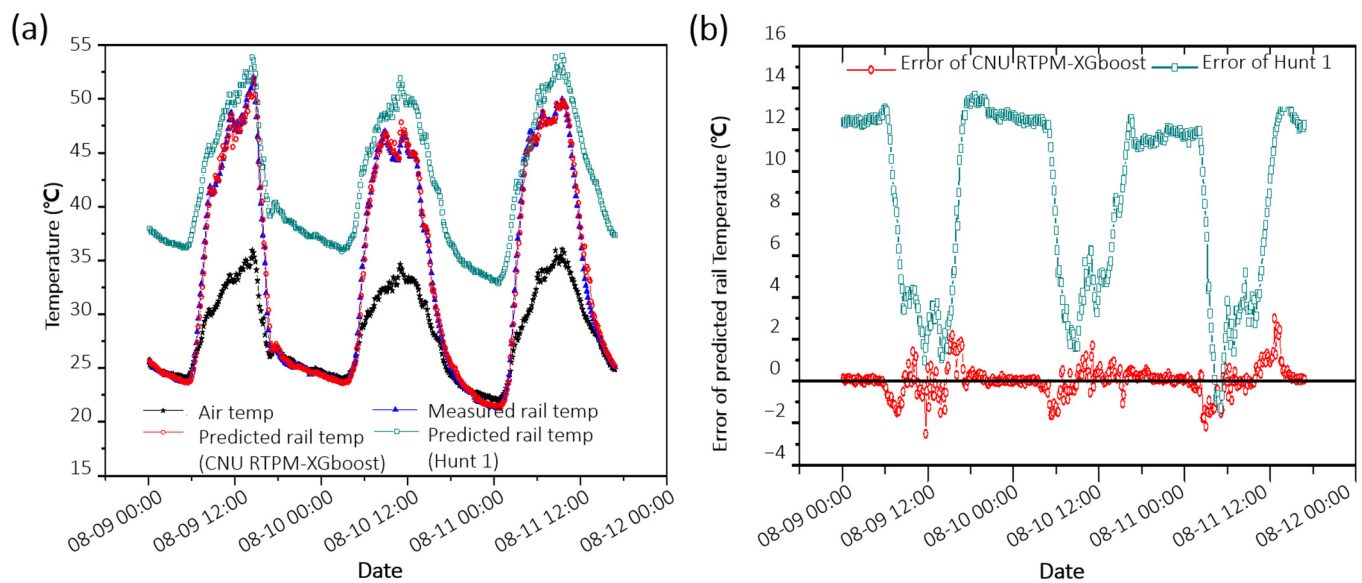
6.3. Model Performance Comparison

6.3.1. Performance at the Whole Range of Rail Temperatures

Table 1 compares the performances of the previous RTPMs (white) and the newly developed CNU RTPMs (yellow and orange). To compare the models' performances, the datasets used in the RTPMs should be similar. However, direct performance comparison was difficult because of the restricted data set used in developing previous RTPMs. Thus, the Hunt model, which uses the air temperature as the sole feature, was used to confirm the similarity of the dataset used in this study with that used in other RTPMs. Besides, the Hunt model is common because similar studies have also compared model performance using it [7,10,13]. In a study by Wu (development of the BoM prediction equation), the RMSE of the Hunt 1 model was determined as 6.952 °C, while in this study, it was determined as 5.866 °C. These results disprove the similarity between Wu's study and the measurement conditions adopted in this research, such as the rail installation environment and climate. Figure 7 shows the predicted and measured rail temperatures from 9 August to 12 August in 2016, when the rails were the hottest. The Hunt 1 model predicts the temperature of the rails to be somewhat higher, while the CNU RTPM–XGBoost predicts it with high accuracy regardless of day or night conditions.

Table 1. Comparison of prediction errors in previous RTPMs and CNU RTPMs.

Performance Range	Performance in Whole Range of Rail Temperature			Performance in High Rail Temperature Range (over 40 °C)		
Model Name	MAE (°C)	R^2	RMSE (°C)	MAE (°C)	R^2	RMSE (°C)
Hunt model 1 [8]	2.913	0.9021	5.866	2.660	0.4471	3.925
Hunt model 2 [8]	14.396	0.9021	15.117	3.661	0.4471	4.448
Munro model [1]	Unknown	0.9180	Unknown	Unknown	Unknown	Unknown
BoM Prediction Equation (1–24 h) [10]	0.136	0.9630	2.560	Unknown	Unknown	Unknown
Weather station Regression model [10]	0.659	0.8960	4.193	Unknown	Unknown	Unknown
CNU RTPM–PR2	0.693	0.9709	2.316	0.103	0.5605	2.7108
Chapman's model [13]	0.200	Unknown	2.500	Unknown	Unknown	Unknown
CNU Heat Transfer model [7]	1.537	0.9334	3.799	0.618	0.3406	5.935
CNU RTPM–SVM	0.068	0.9720	2.135	1.055	0.5465	2.802
CNU RTPM–ANN	0.519	0.9839	1.732	1.213	0.7115	2.355
CNU RTPM–RF	0.029	0.9972	0.685	0.191	0.9199	0.927
CNU RTPM–XGBoost	0.008	0.9984	0.518	0.119	0.9415	0.771
CNU RTPM–XGBoost Without a solar effect	0.043	0.9816	1.744	0.570	0.7243	1.898

Figure 7. (**a**) Comparison of the predicted rail temperature and the measured rail temperature graphs. (**b**) Error of predicted rail temperature in August 2016.

Among the previous RTPMs, the BoM prediction equation showed the highest performance (MAE = 0.136 °C, R^2 = 0.9630, RMSE = 2.560 °C) over the whole range of rail temperatures. However, the BoM equation requires 24 features, which severely reduces its versatility. Moreover, the BoM equation is constructed from data collected over one month during the Australian winter, casting doubt on its global and seasonal applicability. In contrast, the CNU RTPMs use consecutive data collected over 10 months (including summer and winter), ensuring their reliability over the four seasons. Moreover, the weather factors and solar effects are easily obtainable, ensuring the versatility of the CNU RTPMs. The weather factors are globally available through weather forecasts, and the solar effects are readily calculated at any given time. These selected features enable easy combination of CNU RTPM and TSLAM.

The highest performer among the RTPMs over the whole range of rail temperatures was CNU RTPM–XGBoost (MAE = 0.008 °C, R^2 = 0.9984, RMSE = 0.518 °C). The RMSE was approximately 2 °C lower in CNU RTPM–XGBoost than in the BoM prediction model, the previously highest performance RTPM.

Accordingly, CNU RTPM–XGBoost was selected as the main model in this study and its hyperparameters were optimized by K-fold cross validation. Adding the solar effect features improved the performance of CNU RTPM–XGBoost over that of CNU RTPM–XGBoost without solar effects, built using the weather data only.

6.3.2. Performance at the High Rail-Temperature Range (over 40 °C)

The most important practical requirement of RTPM is predicting the high rail temperatures that cause buckling. When predicting rail temperatures over 40 °C (Table 1), the CNU RTPM–RF and CNU RTPM–XGBoost delivered the highest performance among the tested models (CNU RTPM–RF: MAE = 0.191 °C, R^2 = 0.9199, RMSE = 0.927 °C; CNU RTPM–XGBoost: MAE = 0.119 °C, R^2 = 0.9415, RMSE = 0.771 °C). The ability to predict high rail temperatures was conferred by a selected algorithm and proposed features.

Since they adopt the ensemble method, tree-based machine learning algorithms (CNU RTPM–RF and CNU RTPM–XGBoost) are more suitable for high temperature prediction than other machine learning algorithms. Tree-based RTPMs are composed of various submodels constructed under specific conditions. As one of these submodels predicts high rail temperatures, tree-based RTPMs can outperform models with other architectures.

Additionally, we showed that the newly proposed solar effect features further improve the performance of predicting the high rail temperature in RTPM by comparing the performance of the CNU RTPM–XGBoost with or without the solar effect ($R^2 = 0.9415$ and 0.7243, respectively).

6.3.3. Raw Error Data of CNU RTPM

The practical applicability of CNU RTPMs was investigated on raw error data. The performances of the CNU RTPMs with the same features are compared in Figure 8. This figure shows the box plots of the errors in the CNU RTPMs, allowing a comparison of the raw error data in each CNU RTPMs.
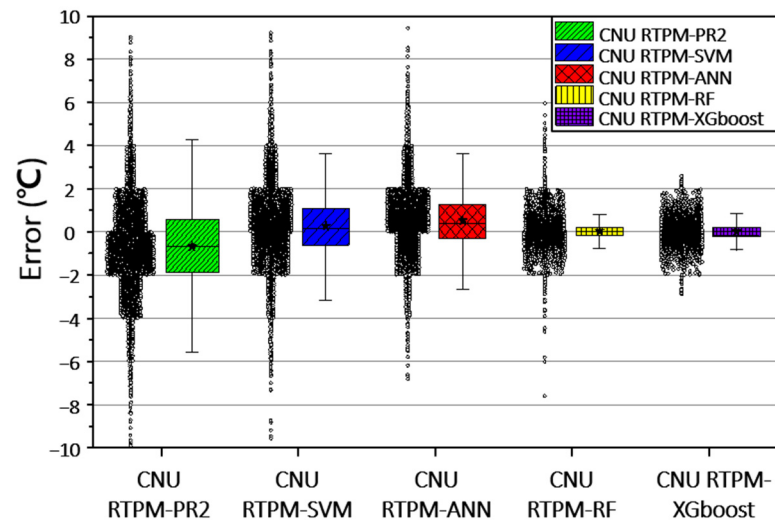
Figure 8. Boxplots showing the error ranges in the CNU RTPMs based on SVM, ANN, PR2, RF, and XGBoost.

The box plots differ in shape; the fence length (distance between the upper and lower whiskers) is shorter in the CNU RTRMs based on machine learning (SVM, ANN, RF, and XGBoost) than in the regression model (PR2). The tree-based RTPMs (RF and XGBoost) yielded the shortest fence lengths around a median of 0 °C.

The error ranged by almost 9 °C in CNU RTPM–PR2, SVM, and ANN, nearly 6 °C in CNU RTPM–RF, and nearly 3 °C in CNU RTPM–XGBoost. A narrow error range guarantees a reliable rail-temperature prediction that minimizes the risk of incorrect prediction.

As confirmed by its short fence lengths and narrow error range (3 °C), the CNU RTPM–XGBoost is the most suitable algorithm in practical situations, because the criteria of train-speed limits are shifted to a 4–5 °C change in rail temperature. This varies from country to country. In Korea, high-speed trains run at 230 km·h^{-1} when the rail temperature is 55–60 °C, 70 km·h^{-1} when the rail temperature is 60–64 °C, and are cancelled when the rail temperature exceeds 64 °C [15]. In the UK, the train speed is determined by the stress-free temperature (STF, usually equal to 27 °C in the UK). On well-maintained tracks, the speed limit is 60 mph and 20 mph at rail temperatures of STF + 37 °C and STF + 42 °C, respectively. However, under poor track conditions (inadequate ballast), the temperatures of the 60 mph and 20 mph speed limits reduce to STF + 13 °C and STF + 15 °C, respectively [21].

6.4. TSLAM with CNU RTPM–XGboost

TSLAM is a structural health-monitoring application for analyzing railway safety using CNU RTPM–XGBoost, the highest-performing CNU RTPM. Using TSLAM, railway safety managers can know the rail temperature in advance or in real-time. Based on the predicted rail temperature, railway safety managers could decide on safety measures, such as limiting the train speed and spraying water.

Figure 9a shows the graphical user interface (GUI) of TSLAM. The user selects the time,

country, and map mode through input-selection widgets and obtains the rail temperature predicted by CNU RTPM–XGBoost. The predicted rail temperature is displayed either in Mode 1 (the standard mode showing the predicted rail temperatures in all regions) or Mode 2 (the detection mode displaying the danger zones on the map). The text-display panel reports the risk points demanding rail-temperature management and the train-speed limits recommended by the Korea Railroad Operation Safety Managers Association [15] (see Appendix E). Equipped with this information, railway safety officers can either plan the lowering of the rail temperature by water spraying or adjust the train operation interval in advance.
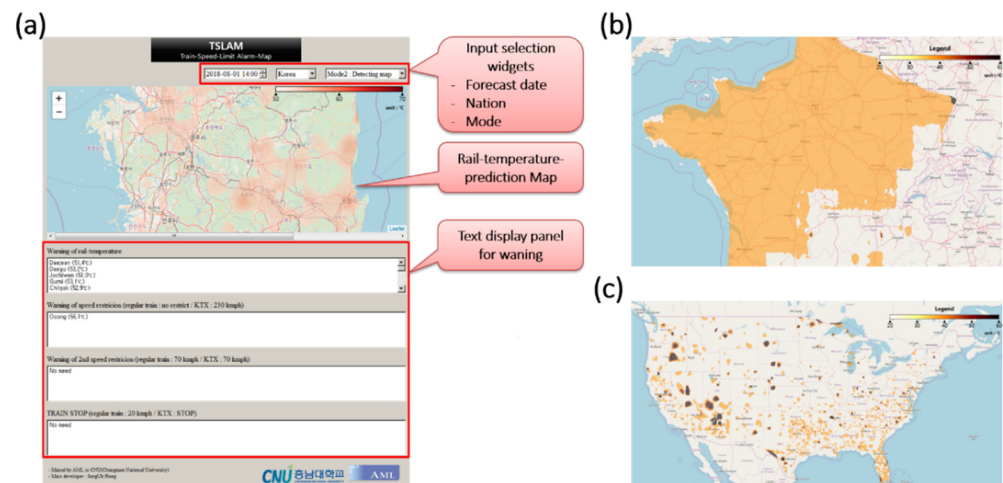
Figure 9. (**a**) GUI of TSLAM and maps of the predicted local rail temperatures in (**b**) France and (**c**) the USA.

The main advantage of TSLAM is its worldwide applicability, enabled by including the globally available features in CNU RTPM–XGBoost. A user can select a country through the GUI, and visually observe the predicted rail temperature of that country using the weather forecast provided by the OWM. In other words, users can obtain the rail temperature or warnings from devices installed with TSLAM, independent of location. As an example, Panels (b) and (c) of Figure 9 show the predicted rail-temperature maps of France and the USA, respectively, in Mode 1. In this mode, users can compare the regional deviations in the predicted rail temperatures. In Mode 2 (Figure 9a), the user must know the train driving rules of the selected country.

TSLAM can improve the efficiency of a train system controlled by a safety management system. As mentioned above, TSLAM requires the weather forecast data from the OWM, which are easily accessible. Therefore, operators of a train safety management system can acquire the information required for rerouting or applying derailment-prevention measures at any locality. With this information, they can appropriately adjust the train interval, avoid bottlenecks, and efficiently operate the rail system.

Additionally, TSLAM can support the indirect measurement of the rail temperature over the entire network. Generally, the rail temperature is directly measured using thermocouples; however, installing thermocouples over the entire network results in high installation and maintenance costs. Additionally, because of the safety issue, it is difficult to attach the thermocouples in some networks. However, TSLAM prediction showed nearly the same result as directly measuring the rail. Thus, TSLAM is expected to serve as a supporting system that provides the predicted rail temperature using real-time weather measurement data or weather forecast data.

6.5. Limitations and Directions for Improvement

The features of CNU RTPM were designed by using the thermal analysis. The aim was to improve the performance of rail-temperature prediction and to extend its network applicability. However, some relevant features were excluded because of data unavailability and restriction. The particulate matter (PM) is one of the excluded features, which can be incorporated into an improved RTPM in future work.

Recently, PM2.5 (with diameters of 2.5 μm or less) has been shown to decrease the DNI proportion and increase the DHI proportion in the GHI [48]. By definition, the DNI is related to the direction of sunlight and induces a temperature distribution on the cross- section of the rail. In contrast, the DHI is less directional and tends to smooth the temperature distribution on the rail cross-section. In regions with high PM2.5 concentration, this interplay causes a slight difference between the rail temperature predicted by CNU RTPM and the actual rail temperature.

Moreover, sleeper material might be incorporated into an improved RTPM in future work. During the measurements, the rail was installed on concrete sleepers with a thermal conductivity of 0.13 $W \cdot m^{-1} \cdot k^{-1}$ [49]. Although the installation of concrete sleepers is increasing with the development of high-speed trains, wooden sleepers remain common in many parts of the world. The thermal conductivity of wood is 2.0 $W \cdot m^{-1} \cdot k^{-1}$ [49,50]. The different thermal conductivities of these two materials will degrade the universal applicability of CNU RTPM. Concrete sleepers with low thermal conductivity will conduct a lower heat flux than wooden sleepers. For this reason, when the solar irradiance is high, a rail installed on wood sleepers will reach a lower temperature than a rail installed on concrete sleepers.

Furthermore, the health status of rails can be integrated into an improved RTPM. The health status of the rail is directly related to its thermal properties, such as solar absorptivity, and mechanical properties, which include hardness and density [7,51]. For example, according to a previous study, solar absorptivity is important for predicting rail temperatures [7]. The solar absorptivity expresses how much solar irradiation affects the temperature change of the rail. The rail surface condition affects the solar absorptivity. Usually, the solar absorptivity of unused rail is 98.7%, whereas that of used rail is 81.1%. This difference comes from the rail surface and paint erosion. Thus, the performance of RTPM can be improved in future works by considering the health status of rail.

7. Conclusions

Herein, we developed novel rail-temperature prediction models (CNU RTPMs) and a TSLAM using weather forecast data alone, which can predict the rail temperature over the entire network.

- The CNU RTPMs were developed with different machine learning methods using the long-term (over 10 months) measured data from all seasons. Such long-term data collection ensures a reliable model.
- To improve the prediction performance, the CNU RTPMs combine standard weather features with newly suggested solar effect features. These features originate from the analysis of the thermal environment around the rail. Additionally, they are easily obtained from global weather forecasts and additional calculations on the weather forecast data. Precisely, the solar effect features significantly improved prediction performance at the high-rail-temperature range.
- In a performance comparison, the CNU RTPM–XGBoost emerged as the best predictor of the rail temperature among the machine learning methods. The proposed CNU RTPM–XGBoost, which delivered higher performance, reliability, and versatility than previous RTPMs, was suggested as a new model for predicting rail temperature over the entire network. The CNU RTPM–XGBoost is applicable worldwide because its features are globally available in weather forecast data.
- The visualization application, TSLAM, maps the predicted rail temperatures, which assist railway safety officers (if necessary) in planning safety measures.

We expect that CNU RTPM–XGBoost and TSLAM will significantly improve both train safety and train timeliness.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Abbreviations

The following abbreviations are used in this article:

CWR	Continuous welded rail
BoM	Bureau of meteorology
RTPM	Rail-temperature-prediction model
RMSE	Root mean square error
MAE	Mean absolute error
R^2	Coefficient of determination
ANN	Artificial neural network
SVM	Support vector machine
RF	Random forest
XGBoost	Extreme gradient boosting
TSLAM	Train-speed-limit alarm-map
DAQ	Data acquisition system
CNU	Chungnam national university
KMA	Korea meteorological administration
TSI	Total solar irradiance
GHI	Global solar irradiance
DHI	Diffuse horizontal irradiance
DNI	Direct normal irradiance
API	Application programming interface
OWM	Open weather map
GUI	Graphical user interface
PM	Particulate matter

Appendix A

The summary of the dataset is shown in Table A1. This dataset is consisted of measured data (air temperature, relative humidity, wind speed, rainfall amount, and rail temperature) and borrowed data (cloud cover). Table A1 shows the maximum value, minimum value, average value, and standard deviation.

Table A1. Summary of the dataset (from August of 2016 to May of 2017).

Features	Maximum Value	Minimum Value	Average Value	Standard Deviation
Air temperature	38.00 °C	−10.10 °C	11.03 °C	10.30 °C
Relative humidity	98.00%	14.00%	67.59%	18.64%
Wind speed	19.30 m/s	0.00 m/s	1.64 m/s	2.12 m/s
Rainfall amount	3.81 mm	0.00 mm	0.03 mm	0.98 mm
Rail temperature	53.04 °C	−11.67 °C	13.63 °C	13.08 °C
Cloud cover	10	1	6.90	2.84
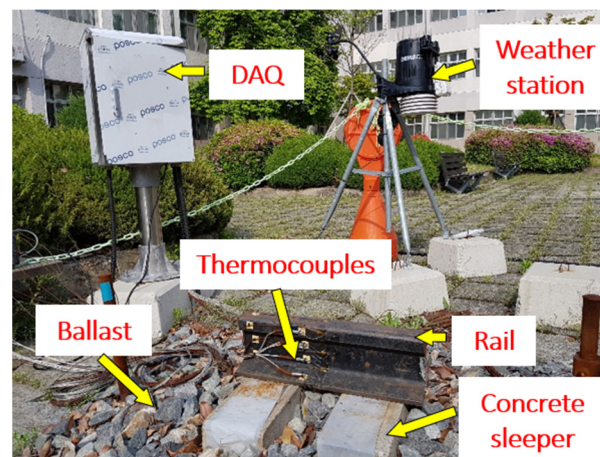
Figure A1. The installed measurement system.

Appendix B

The global horizontal irradiance (GHI) is closely related to rail temperature. The comparison of the measured rail temperature and GHI at the measurement station during the 2016 summer (16.08.05–16.08.30) are shown in Figure A2. Figure A2a shows the positive correlation between the measured rail temperature and GHI. This result is more clearly shown in Figure A2b. According to Figure A2b, the coefficient of determination (R^2) was 0.6870.
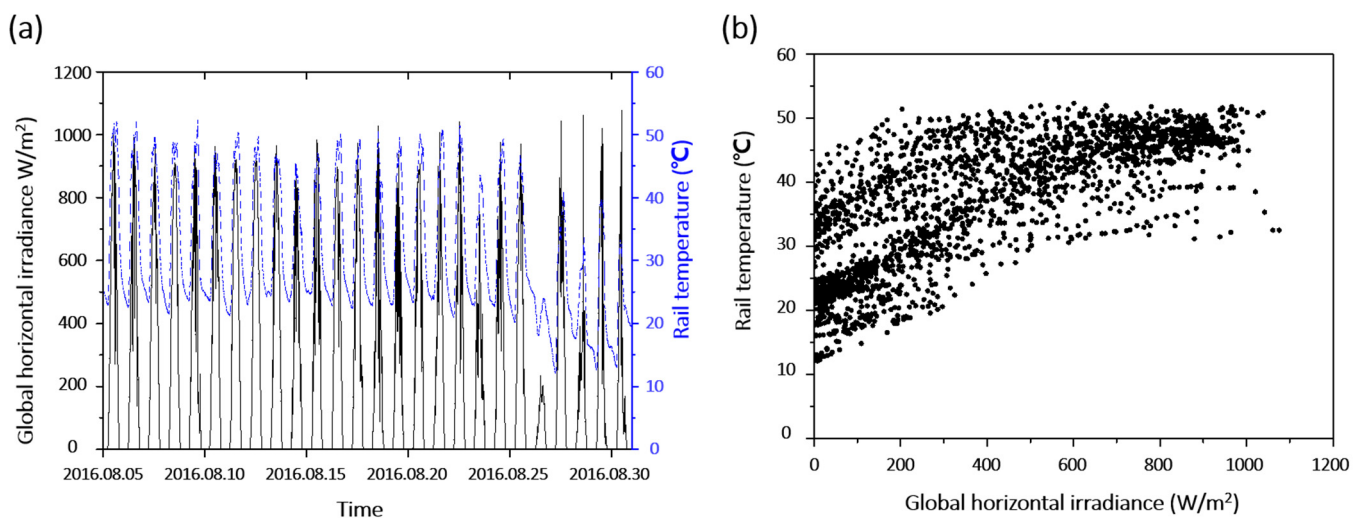
Figure A2. (**a**) Measured rail temperatures and global horizontal irradiance data. (**b**) Correlation between rail temperature and global horizontal irradiance at the study site in summer (August of 2016).

Appendix C

The total solar irradiance (TSI) is governed by the elliptical orbit of the Earth. The TSI is maximized and minimized at the perihelion (January) and apex (July) of the Earth, respectively, and is calculated as follows:

$$\theta = (d_n - d_p) \times \frac{360}{365.256363004} \times \frac{\pi}{180}, \tag{A1}$$

$$D = \frac{\alpha \times (1 - e^2)}{1 + e \times \cos(\theta)}, \tag{A2}$$

$$I_t = I_{sc} \times \left(\frac{D_{avg}}{D}\right)^2. \tag{A3}$$

In the above expressions, d_n and d_p denote the elapsed day from January 1 and perihelion day (January 2–5), respectively, and θ is the circumferential angle from perihelion. D and D_{avg} are the actual and mean sun–Earth distances, respectively, α is the semi-major axis of the ellipse (149,598,261 km), e is the eccentricity of the Earth's orbit (0.01671123), I_{sc} is the solar constant (1367 W/m^2), and It is the TSI.

Note that d_n is the only variable in the TSI calculation.

Appendix D

The optimized hyperparameters of the extreme gradient boosting (XGBoost) were *n_estimators*, *colsample_bytree*, *learning_rate*, *max_depth*, and *subsample*. The value range of hyperparameters were described in the list of candidate hyperparameters in Table A2. As a result, tuning was performed to find the optimal value for the combination of hyperparameters in the 3750 preset value for the combination of hyperparameters. The results of the hyperparameters tuning are described in Table A2. Since the results were difficult to describe completely, the top six results were shown based on the best score. The score was calculated based on the coefficient of determination (R^2). As a result, the optimal hyperparameters set was *n_estimators* = 3000, *colsample_bytree* = 1.0, *learning_rate* = 0.03, *max_depth* = 9, and *subsample* = 0.4.

Table A2. The list of candidate parameters and parameters tuning results for finding the optimal parameter set in XGBoost.

List of Candidate Hyperparameters					
n_estimators	*Colsample_bytree*	*Learning_rate*	*Max_depth*	*Subsample*	
{500, 1000, 1500, 2000, 2500, 3000}	{0.2, 0.4, 0.6, 0.8, 1.0}	{0.01, 0.03, 0.05, 0.07, 0.09}	{5, 6, 7, 8, 9}	{0.2, 0.4, 0.6, 0.8, 1.0}	
Hyperparameters Tuning Results (Top 6)					
n_estimators	*Colsample_bytree*	*Learning_rate*	*Max_depth*	*Subsample*	**Best Score**
3000	1.0	0.03	9	0.4	0.998009
2500	1.0	0.03	9	0.4	0.997989
3000	1.0	0.03	8	0.4	0.997984
2000	1.0	0.03	9	0.4	0.997951
2500	1.0	0.03	8	0.4	0.997945
3000	1.0	0.05	8	0.4	0.997924

The hyperparameter of the support vector machine (SVM) was optimized using GridSearchCV same as XGBoost. The optimized hyperparameters are the kernel coefficient γ and penalty coefficient C. The value range of hyperparameters and result of hyperparameters tuning are described in Table A3. Table A3 shows the optimal hyperparameters set is $\gamma = 0.1$ and $C = 10$.

Table A3. The list of candidate parameters and parameters tuning results for finding the optimal parameter set in SVM.

List of Candidate Hyperparameters		
γ	C	
{ 0.001, 0.01, 0.1, 1, 10}	{ 0.001, 0.01, 0.1, 1, 10}	
Hyperparameters Tuning Results (top 6)		
γ	C	**Best Score**
0.1	10	0.960011
1	10	0.958223
1	1	0.957759
1	0.1	0.956218
0.1	1	0.953065
0.01	10	0.944773

The hyperparameter of the random forest (RF) was optimized using GridSearchCV same as SVM and RF. The optimized hyperparameters are the number of variables to pick to split on at each node (M_{try}) and the number of trees in the run (N_{tree}). The value range of hyperparameters and result of hyperparameters tuning are described in Table A4. Table A4 shows that the optimal hyperparameters set was $M_{try} = 11$ and $N_{tree} = 3000$.

Table A4. The list of candidate parameters and parameters tuning results for finding the optimal parameter set in RF.

List of Candidate Hyperparameters		
M_{try}	N_{tree}	
{7, 8, 9 10, 11}	{500, 1000, 1500, 2000, 2500, 3000}	
Hyperparameters Tuning Results (Top 6)		
M_{try}	N_{tree}	**Best Score**
11	3000	0.993880
11	1000	0.993865
11	1500	0.993865
11	2000	0.993861
11	2500	0.993859
11	500	0.993845

To optimize the number of nodes in one hidden layer artificial neural network (ANN), we compared the losses generated by changing the number of nodes. This ANN model used MSE as a loss function. The optimization of ANN seeks to minimize a loss function. Figure A3 shows the tuning result in ANN. In this graph, node 22 shows the minimum loss value. This means node 22 is an optimal hyperparameter.
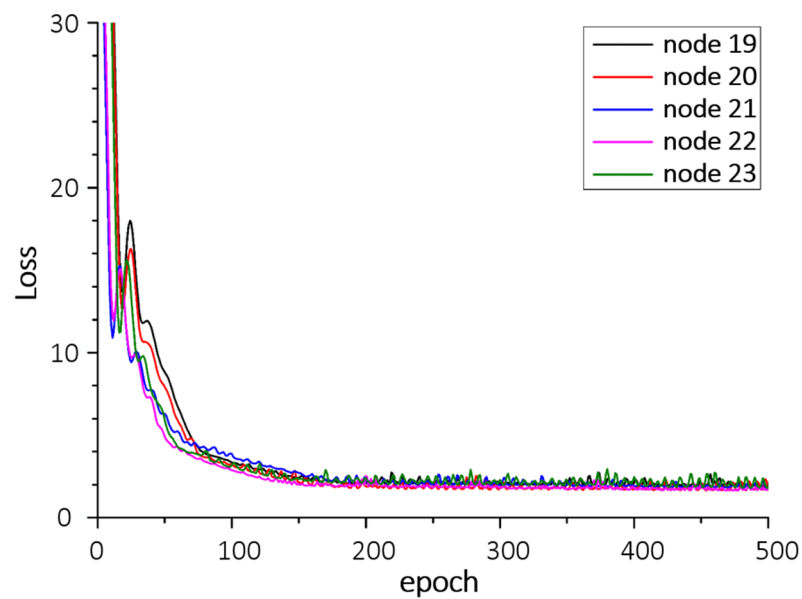
Figure A3. Tuning result of the ANN.

Appendix E

To prevent buckling-induced train derailment due to the rise of the rail temperature, the Korea Railroad Operation Safety Managers Association (KORSMA), made a rule on the train speed limit according to the rail temperature [15]. The rule is divided into part of regular trains and Korea Train eXpress (KTX). The rule is summarized in Table A5.

Table A5. Train driving rules according to rail temperature in KORSMA.

Rail Temperature / Train Type	55–60 °C	60–64 °C	Exceed 64 °C
Regular train	-	<70 km/h	<20 km/h
KTX (Korea Train eXpress)	<230 km/h	<70 km/h	Stop

References

1. Wu, Y.; Munro, P.; Rasul, M.; Khan, M. A review on recent developments in rail temperature prediction for use in buckling studies. In Proceedings of the CORE 2010: Rail, Rejuvenation and Renaissance, Wellington, Australia, 12–15 September 2010.
2. Kish, A.; Samavedam, G. *Risk Analysis Based CWR Track Buckling Safety Evaluations*; John A. Volpe National Transportation Systems Center: Cambridge, MA, USA, 2001.
3. Kish, A.; McWilliams, R.; Harrison, H. Track buckling hazard detection and rail stress management. In Proceedings of the 9th World Congress on Railway Research, Lille, France, 22–26 May 2011.
4. Ferranti, E.; Chapman, L.; Lowe, C.; Mcculloch, S.; Jaroszweski, D.; Quinn, A. Heat-Related Failures on Southeast England's Railway Network: Insights and Implications for Heat Risk Management. *Weather Clim. Soc.* **2016**, *8*, 177–191. [CrossRef]
5. Jaroszweski, D.; Fu, Q.; Easton, J. A data model for heat-related rail buckling: Implications for operations, maintenance and long-term adaptation. In Proceedings of the 12th World Congress on Railway Research, Tokyo, Japan, 28 October–1 November 2019.
6. Ferranti, E.; Chapman, L.; Lee, S.; Jaroszweski, D.; Lowe, C.; Mcculloch, S.; Quinn, A. The hottest July day on the railway network: Insights and thoughts for the future. *Meteol. Appl.* **2017**, *25*, 195–208. [CrossRef]
7. Hong, S.U.; Kim, H.U.; Lim, N.H.; Kim, K.H.; Kim, H.; Cho, S.J. A Rail-Temperature-Prediction Model Considering Meteorological Conditions and the Position of the Sun. *Int. J. Precis. Eng. Manuf.* **2019**, *20*, 337–346. [CrossRef]
8. Hunt, G. *An Analysis of Track Buckling Risk*; British Railways: London, UK, 1994.
9. Esveld, C. *Modern Railway Track*; Delft University of Technology: Zaltbommel, Netherlands, 2001.
10. Wu, Y.; Rasul, M.; Powell, J.; Micenko, P.; Khan, M. Rail temperature prediction model. In Proceedings of the CORE 2012 Global Perspectives Conference on railway engineering, Brisbane, Australia, 10–12 September 2012.

11. Zhang, Y.; Lee, S. Modeling rail temperature with real-time weather data. Surf. Trans. *Weather Snow. Remove Ice Cont. Tech.* **2008**, *37*, 37–48.

12. Kesler, K.; Zhang, Y. System and Method for Predicting Future Rail Temperature. U.S. Patent Application US2007/0265780 A1, 15 November 2007.

13. Chapman, L.; Thornes, J.E.; Huang, Y.; Cai, X.; Sanderson, V.L.; White, S.P. Modelling of rail surface temperatures: A preliminary study. *Theor. Appl. Clim.* **2007**, *92*, 121–131. [CrossRef]

14. Han, S.H.; Hwang, S.H.; Kim, H.J.; Cho, S.J.; Lim, N.H. Numerical study on the characteristics of temperature distribution in continuous welded rail by solar radiation and rail orientation. *J. Mech. Sci. Technol.* **2020**, *34*, 4819–4829. [CrossRef]

15. The Rules of High Speed Rail for Operation Rule. Available online: http://www.krosma.co.kr/36/?q=YToxOntzOjEyOiJrZXl3b3 JkX3R5cGUiO3M6MzoiYWxsIjt9&bmode=view&idx=1696661&t=board (accessed on 5 June 2021).

16. Dobney, K.; Baker, C.J.; Chapman, L.; Quinn, A.D. The future cost to the United Kingdoms railway network of heat-related delays and buckles caused by the predicted increase in high summer temperatures owing to climate change. *Proc. Inst. Mech. Eng. J. Rail Rapid Trans.* **2009**, *224*, 25–34. [CrossRef]

17. Mateo, F.; Carrasco, J.J.; Sellami, A.; Millán-Giraldo, M.; Domínguez, M.; Soria-Olivas, E. Machine learning methods to forecast temperature in buildings. *Exp. Syst. Appl.* **2013**, *40*, 1061–1068. [CrossRef]

18. Abdel-Aal, R.E.; Elhadidy, M.A. Modeling and Forecasting the Daily Maximum Temperature Using Abductive Machine Learning. *Weather Forecast* **1995**, *10*, 310–325. [CrossRef]

19. Ryan, M. *Rail Temperature Study*; Rail Safety and Standards Board: Derby, UK, 2005.

20. Hong, S.; Jung, H.; Park, C.; Lee, H.; Kim, H.; Lim, N.; Bae, H.; Kim, K.; Kim, H.; Cho, S.J. Prediction of a representative point for rail temperature measurement by considering longitudinal deformation. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Trans.* **2019**, *233*, 1003–1011. [CrossRef]

21. Chapman, L.; Thornes, J.E.; White, S.P. Thermal imaging of railways to identify track sections prone to buckling. *Proc. Inst. Mech. Eng. J. Rail Rapid Trans.* **2006**, *220*, 317–327. [CrossRef]

22. Fröhling, R.; Koker, J.D.; Amade, M. Rail lubrication and its impact on the wheel/rail system. *Proc. Instit. Mecha. Eng. J. Rail Rapid Transit* **2009**, *223*, 173–180. [CrossRef]

23. Stein, J.S.; Hansen, C.W.; Reno, M.J. *Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis*; Sandia National Laboratories: Albuquerque, NM, USA, 2012.

24. Rohsenow, W.M.; Hartnett, J.P.; Cho, Y.I. *Handbook of Heat Transfer*; McGraw-Hill: New York, NY, USA, 1998.

25. Nematollahi, O.; Kim, K.C. A feasibility study of solar energy in South Korea. *Renew. Sustain. Energy Rev.* **2017**, *77*, 566–579. [CrossRef]

26. Sengupta, M.; Habte, A.; Kurtz, S.; Dobos, A.; Wilbert, S.; Lorenz, E.; Stoffel, T.; Renné, D.; Gueymard, C.; Myers, D.; et al. *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications*; National Renewable Energy Laboratory: Alexandria, NM, USA, 2017.

27. Gueymard, C.A. The sun's total and spectral irradiance for solar energy applications and solar radiation models. *Sol. Energy* **2004**, *76*, 423–453. [CrossRef]

28. Davies, J.A.; Mckay, D.C. Estimating solar irradiance and components. *Sol. Energy* **1982**, *29*, 55–64. [CrossRef]

29. Robledo, L.; Soler, A. Luminous efficacy of global solar radiation for clear skies. *Energy Conver. Manag.* **2000**, *41*, 1769–1779. [CrossRef]

30. Rizvi, A.; Addoweesh, K.; El-Leathy, A.; Al-Ansary, H. Sun position algorithm for sun tracking applications. In Proceedings of the IECON 2014—40th Annual Conf. IEEE Industrial Electronics Society, Dallas, TX, USA, 29 October–1 November 2014.

31. Spencer, J. Comments on the Astronomical Almanacs Algorithm for Approximate Solar Position (1950–2050). *Sol. Energy* **1988**, *42*, 353. [CrossRef]

32. Wang, W.; Shi, Y.; Lyu, G.; Deng, W. Electricity Consumption Prediction Using XGBoost Based on Discrete Wavelet Transform. In Proceedings of the 2nd International Conference on Artificial Intelligence and Engineering Applications, Guilin, China, 23–24 September 2017.

33. Pan, B. Application of XGBoost algorithm in hourly PM2.5 concentration prediction. *OP Conf. Ser. Earth Environ. Sci.* **2018**, *113*, 012127. [CrossRef]

34. Fan, J.; Wang, X.; Wu, L.; Zhou, H.; Zhang, F.; Yu, X.; Lu, X.; Xiang, Y. Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. *Energy Conver. Manag.* **2018**, *164*, 102–111. [CrossRef]

35. Zhang, L.; Ai, H.; Chen, W.; Yin, Z.; Hu, H.; Zhu, J.; Zhao, J.; Zhao, Q.; Liu, H. CarcinoPred-EL: Novel models for predicting the carcinogenicity of chemicals using molecular fingerprints and ensemble learning methods. *Sci. Rep.* **2017**, *7*, 1–14. [CrossRef]

36. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.

37. Friedman, J.H. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **2002**, *38*, 367–378. [CrossRef]

38. Hastie, T.; Friedman, J.; Tisbshirani, R. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2017.

39. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

41. Chen, Y.-Y.; Lin, Y.-H.; Kung, C.-C.; Chung, M.-H.; Yen, I.-H. Design and Implementation of Cloud Analytics-Assisted Smart Power Meters Considering Advanced Artificial Intelligence as Edge Analytics in Demand-Side Management for Smart Homes. *Sensors* **2019**, *19*, 2047. [CrossRef]

42. Huang, G.; Huang, G.; Song, S.; You, K. Trends in extreme learning machines: A review. *Neural Net.* **2015**, *61*, 32–48. [CrossRef]

43. Song, C.; Zeng, P.; Wang, Z.; Zhao, H.; Yu, H. Wearable Continuous Body Temperature Measurement Using Multiple Artificial Neural Networks. *IEEE Trans. Ind. Inform.* **2018**, *14*, 4395–4406. [CrossRef]

44. Claesen, M.; De Moor, B. Hyperparameter Search in Machine Learning. Available online: https://arxiv.org/abs/1502.02127 (accessed on 5 June 2021).

45. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.

46. Bergstra, J.; Pinto, N.; Cox, D. Machine learning for predictive auto-tuning with boosted regression trees. In Proceedings of the 2012 Innovative Parallel Computing (InPar), San Jose, CA, USA, 13–14 May 2012.

47. Olston, C.; Najork, M. *Web Crawling*; Now Publish: Boston, MA, USA, 2010.

48. Luo, H.; Han, Y.; Lu, C.; Yang, J.; Wu, Y. Characteristics of Surface Solar Radiation under Different Air Pollution Conditions over Nanjing, China: Observation and Simulation. *Adv. Atmos. Sci.* **2019**, *36*, 1047–1059. [CrossRef]

49. Ilomets, S.; Kuusk, K.; Paap, L.; Arumägi, E.; Kalamees, T. Impact Of Linear Thermal Bridges On Thermal Transmittance Of Renovated Apartment Buildings. *J. Civ. Eng. Manage* **2016**, *23*, 96–104. [CrossRef]

50. Silva, É.; Pokropski, D.; You, R.; Kaewunruen, S. Comparison of structural design methods for railway composites and plastic sleepers and bearers. *Aust. J. Struct. Eng.* **2017**, *18*, 160–177. [CrossRef]

51. Ding, S.; Guiyun, T.; Ruslee, S. Non-destructive hardness prediction for 18CrNiMo7-6 steel based on feature selection and fusion of Magnetic Barkhausen Noise. *NDT E Int.* **2019**, *107*, 102138. [CrossRef]