applied sciences

MDPI

Article

Cable Temperature Prediction Based on RF-GPR for Digital Twin Applications

Weixing Han [1], Chunsheng Hao [1], Dejing Kong [2] and Guang Yang [2,*]

[1] State Grid Handan Electric Power Supply Company, Handan 056000, China; be_hwx@163.com (W.H.); hdgdhcs@126.com (C.H.)
[2] Hebei Provincial Key Laboratory of Power Transmission Equipment Security Defense, North China Electric Power University, Baoding 071000, China; ali_jing1021@163.com
* Correspondence: gascendcn@163.com; Tel.: +86-159-3121-5916

Abstract: With the wide application of power cables in the field of transmission and distribution and the increasing emphasis of power departments on the reliability, safety and stability of power cable operation, how to more accurately and quickly analyze the temperature distribution of power cables and how to evaluate the running state of power cables have become research hotspots. Through the combination of finite element calculation and the artificial intelligence method, an innovative method of digital twin cable temperature prediction based on RF-GPR is proposed in this paper. Firstly, the finite element method is used to calculate the coupling of the electromagnetic field and temperature field of a 10 kV AC cable laid in the cable trench, and a certain amount of basic data are provided through the finite element calculation results. Then, using the basic principle of the random forest (RF) variable importance score, the RF-GPR cable temperature prediction model is constructed using the series hybrid model and Gaussian process regression (GPR), the model prediction results are compared and analyzed, and the calculation time is improved by about 1500 times. Finally, a digital twinning platform for cable temperature calculation based on RF-GPR is designed, which provides technical support for the application of digital twinning.

Keywords: digital twin; power cable; machine learning; temperature prediction

check for updates

Citation: Han, W.; Hao, C.; Kong, D.; Yang, G. Cable Temperature Prediction Based on RF-GPR for Digital Twin Applications. *Appl. Sci.* **2023**, *13*, 7700. https://doi.org/10.3390/app13137700

Academic Editor: Dimitris Mourtzis

Received: 4 June 2023
Revised: 18 June 2023
Accepted: 21 June 2023
Published: 29 June 2023

1. Introduction

Under the goal of "reaching the peak of carbon and carbon neutrality", the construction of new power systems with new energy as the main body is the future development trend of power grids. Digital technology and intelligent application are the keys to support the construction of new power systems and the key technologies to achieve the safe and stable operation and intelligent maintenance of power equipment [1]. A digital twin is one of the key technologies to promote intelligent and digital development in the field of power equipment. Digital twin technology opens up the full-link process of the entity perception model application [2], realizes the comprehensive perception of the state of power equipment based on new sensor technology, and realizes sensor device evaluation and data in-depth governance according to the operation characteristics of power equipment. The digital twin power equipment model is built by big data analysis, data mining and other technologies to carry out the state differentiation evaluation, accurate fault diagnosis and accurate state prediction of power equipment and realize real-time interaction between the physical entity of power equipment and the digital twin, as well as information sharing between multiple digital twins of power equipment [3].

With the rise of the Internet of things and big data and the development of artificial intelligence and virtual reality technology, digital twin technology has received more and more technical support and engineering applications, mainly applied in intelligent agricultural production [4,5], design and manufacturing in the automobile and aerospace fields [6], system development and operation [7], preliminary analysis and research in the field of power equipment [8,9], etc.

Ref. [10] proposed a method of integrating the DS simulation model and Manufacturing Execution System (MES) to create a digital twin model and proposed two frameworks of integrating the digital twin based on MES. Ref. [11] summarized and reviewed the current trends and limitations of digital twin applications in additive manufacturing and provides technical support and method reference for further research on digital twin systems. Ref. [12] showed the reliability and effectiveness of the twin model through the real-time simulation of the digital twin system of wind turbines and the twinning simulation of unmeasurable field data. Ref. [13] deeply and comprehensively discussed the technical route of the real-time online analysis system of the power grid and prospects the application scenarios of the next-generation real-time control system with the second-level response. Ref. [14] analyzed the similarities and differences between digital twinning technology and big data from many aspects and proposed technologies and methods for how to promote the realization of intelligent manufacturing.

With the acceleration of the construction of the new power system network, the electrical connection between the transmission networks becomes closer and closer, and the security and stability of power transmission become more and more important. Currently, the electric energy transmission equipment in the power network framework of our country are mainly aerial lines and cable lines, while power cable is usually laid in the soil, tube galleries, tunnels and cable trenches, which have the advantages of being safe and stable, not occupying space and having high power transmission efficiency. In recent years, the proportion of power transmission efficiency has been increasing in the urban distribution network. Therefore, how to ensure the safe and reliable operation of power cables, as well as the accuracy of the monitoring and evaluation calculation results of key heating conditions such as the cable core, are hot issues in current research [15].

The calculation methods of cable temperature mainly include the numerical method and thermal method. Among them, the numerical method calculates the entire temperature field under the given cable electromagnetic parameters and environmental parameters, including the finite element method, the finite difference method, etc. The thermal path method is mainly based on IEC standards, and the cable structure is equivalent to a one-dimensional model using the equivalent model principle and then calculated according to the law of heat conduction. The thermal path method can only calculate the average temperature of each layer of the cable structure, and the error of the cable temperature value calculated under a poor laying environment is large. In comparison, the finite element coupling calculation method has the advantages of simulating complex models and high calculation accuracy. However, under the same cable laying conditions, the time used to calculate the cable temperature by the finite element method also increases. When constructing cable, the multi-physical field coupling model is applied to the digital twinning cable frame, and the model-driven theory represented by the finite element method cannot meet the real-time requirement of digital twinning application.

The digital twinning of cable can be realized not only by model drive but also by the artificial intelligence algorithm. The development of the artificial intelligence algorithm has provided a new research method for the early state recognition of cable. Machine learning and deep learning were applied in many power system fields [16,17]. In [18], a prediction model was built by the machine learning algorithm to realize the accurate identification of the partial discharge mode of electrical equipment. Ref. [19] constructed the state prediction model of a power transformer by using the short-short memory network algorithm and studied the corresponding relationship between the characteristic variables and transformer operating state. Ref. [20] realized the predictive control of an NSI system in dual-output mode and six-phase mode based on the model prediction method of the finite control set. Ref. [21] established a prediction model based on the recursive neural network principle to realize the load prediction of a power system. Under the framework of digital twin technology, this paper innovatively proposes a digital twin cable temperature prediction method based on machine learning for digital twin application by combining finite element calculation and artificial intelligence. A certain number of basic data are provided through finite element calculation, and then the artificial intelligence method is used to learn, establish a reliable model training process, generate a cable temperature prediction model, and realize the rapid and accurate calculation of the cable temperature distribution in the cable trench. The calculation results meet the real-time requirement of cable temperature calculation in the application of the digital twin and provide technical support for the application of the digital twin.

2. Finite Element Calculation Model of Cable Temperature

The cable laid in the cable trench is a cross-linked polyethylene three-core cable; the specific model is YJV22-8.7/10-3 $\times$ 240 mm^2. According to the internal structure of the cable, the cable can be divided into the following parts: cable core, semiconductor layer, insulation layer, metal shielding layer, filling layer, armor and sheath layer. Since the axial length of the cable is much larger than the radial length of the cable, the two-dimensional plane model is used in the analysis of the cable model. The finite element model is shown in Figure 1, and the cable model inside the cable trench is shown in Figure 2. The finite element software used in this paper was Comsol Multiphysics 5.6. Free triangle mesh was used for the cable part, and boundary layer mesh was used for the inner wall and surface of the cable channel. The total number of meshes in the finite element calculation model was 92,856.

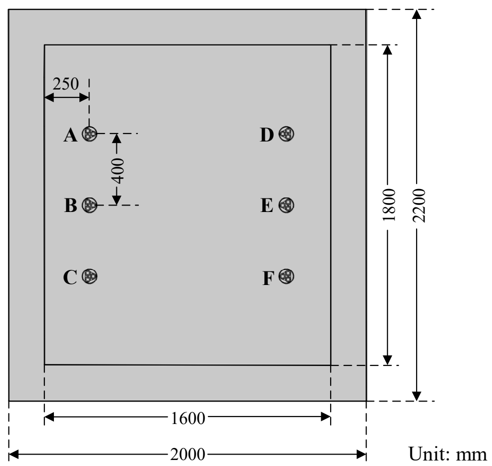

Figure 1. Finite element calculation model of cable trench laying.

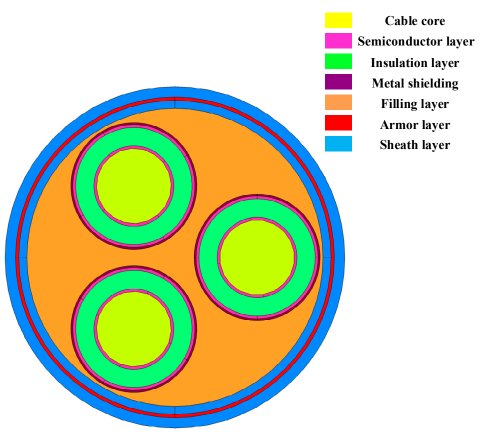

Figure 2. Finite element calculation model of cable temperature.

Two-dimensional steady-state electromagnetic and temperature field coupling calculation is carried out for 10 kV AC cables, in which the cable core loss, dielectric loss, metal shield loss and armor loss generated by the cables themselves are used as the heat source [22], and the partial differential equation of the temperature inside the cables can be obtained as shown below:

$$-\nabla \cdot (k \nabla T) = Q_h \tag{1}$$

where Q_h is the total cable loss, and k is the thermal conductivity.

In the cable trench laying environment set up in this paper, because there is contact between the cable and the air inside the cable trench, it is necessary to consider the influence of the air flow in the cable trench on the heat dissipation. The conductor loss, insulation loss and armor loss of the cable are the main sources of heat, which affect the flow of air in the cable trench and then transfer the heat through conduction and convection. The thermal conductivity differential equation is shown in Equation (2), and the convection differential equation is shown in Equation (3).

$$\lambda \left(\frac{\partial^2 T_s}{\partial x^2} + \frac{\partial^2 T_s}{\partial y^2} \right) + Q = 0 \tag{2}$$

where λ is the thermal conductivity of the medium, T_s is the medium temperature, and Q is the heating rate per unit volume of the medium.

$$Q = h \Delta t \tag{3}$$

where h is the convective heat transfer coefficient, and Δt is the temperature difference between a solid and a fluid.

Heat transfer between the cable trench and soil is a solid heat transfer problem, so the lower boundary of the finite element model of the cable trench belongs to the first type of boundary condition, and the temperature is set as 20 °C for the deep soil. The formula is shown as follows:

$$\begin{cases} T|_{\Gamma} = T_w \\ T|_{\Gamma} = f(x, y, z, t) \end{cases} \tag{4}$$

where Γ is the object boundary, T_w is the boundary temperature, and $f(x, y, z, t)$ is the boundary temperature function.

The left and right boundaries of the finite element calculation model of the cable trench belong to the second type of boundary conditions. The heat flux value of the boundary method phase is set as 0, so the boundary is regarded as an adiabatic surface. The formula is shown as follows:

$$\begin{cases} \lambda \frac{\partial T}{\partial n}\Big|_{\Gamma} + q_w = 0 \\ \lambda \frac{\partial T}{\partial n}\Big|_{\Gamma} + g(x, y, z, t) = 0 \end{cases} \tag{5}$$

where q_w is the known boundary heat flux value, and $g(x, y, z, t)$ is the boundary heat flux function.

The upper boundary of the finite element model of the cable trench belongs to the third type of boundary conditions. The external air temperature is set as the fluid temperature, and the heat transfer coefficient is the convective heat transfer coefficient in the air.

$$-\lambda \frac{\partial T}{\partial n}\Big|_{\Gamma} = \alpha \left(T - T_f \right)\Big|_{\Gamma} \tag{6}$$

where α is the convective heat transfer coefficient of the object surface, and T_f is the temperature of the surrounding fluid, K.

3. RF-GPR Prediction Model of Cable Temperature

3.1. Selection of Characteristic Variables

Considering the influence of the cable trench environmental factors, cable geometric structure and electromagnetic parameters on the temperature distribution inside the cable trench, different parameter combinations were selected in the parameters' setting range for finite element calculation, and, after obtaining several groups of calculation results, the data that can reflect the cable temperature change were extracted. Geometric parameters, electromagnetic parameters and environmental factors were taken as the input characteristic quantity, and the temperature distribution inside the cable trench was taken as the output characteristic quantity. The input quantity and output quantity were combined to obtain a data set that can reflect the change in temperature distribution in the cable trench, which was used as the data set of the machine learning model prediction. In this paper, the influence of various variables on the temperature distribution in the cable trench was analyzed through the finite element calculation results, and a 10-dimensional vector was constructed as a set of input characteristic quantities. The 10-dimensional data are shown in Table 1.

Table 1. Characteristic variables of 10-dimensional input.

Variable Numbers	Input Characteristic Variables	Unit
1	X coordinate of the point	m
2	Y coordinate of the point	m
3	x-axis coordinates of the cable core center	m
4	y-axis coordinates of the cable core center	m
5	Excitation current	A
6	Cable core conductivity	S/m
7	Relative permeability	1
8	Thermal conductivity of insulation	W/(m·K)
9	Convective heat transfer coefficient	W/(m^2·K)
10	Ambient temperature	K

3.2. Characteristic Variable Importance Score Calculation

3.2.1. Random Forest Algorithm Variable Importance Score

The random forest algorithm can be used to evaluate the importance of a set of feature data in the process of regression prediction, and its results can be used as the basis for the selection of feature variables. The conventional random forest variable importance score (VIM) calculation method can be obtained by calculating the *Gini* index [23]. Assuming that there are characteristic variables $X_1, X_2, \cdots, X_M$, then the importance score of variable X_j can be calculated, and the specific formula is shown as follows.

In a decision tree, the *Gini* index of node m is calculated, as shown in Equation (7).

$$G_m = \sum_{k=1}^{K} P_{mk}(1 - P_{mk}) \tag{7}$$

where G_m is the *Gini* index of node m, K is the number of sample categories in the total sample, and P_{mk} is the probability estimate of the sample belonging to class k at node m.

The importance of variable X_j in node m, that is, the calculation formula of the *Gini* index change before and after node m splitting, is shown in Equation (8).

$$V_{jm}^{Gini} = G_m - G_{ml} - G_{mr} \tag{8}$$

where V_{jm}^{Gini} is the importance of variable X_j in node m, and G_{ml} and G_{mr} represent the *Gini* index of the left and right nodes split by node m, respectively.

If M nodes in the i tree contain variable X_j, then the importance calculation formula of variable X_j in the i tree is shown in Equation (9).

$$V_{ij}^{Gini} = \sum_{m=1}^{M} V_{jm}^{Gini} \tag{9}$$

where V_{ij}^{Gini} is the importance of variable X_j in the i tree, and M is the number of nodes in the i tree containing the variable X_j.

If there are n trees in the random forest, the $Gini$ importance of variable X_j in the random forest is defined as the average importance of variable X_j in all trees in the random forest, and its calculation formula is shown in Equation (10).

$$V_j^{Gini} = \frac{1}{n} \sum_{i=1}^{n} V_{ij}^{Gini} \tag{10}$$

where V_j^{Gini} is the $Gini$ importance of variable X_j in the random forest, and n is the number of decision trees in the random forest.

3.2.2. Characteristic Variable Importance Calculation Results

In order to calculate the importance of each characteristic variable, the random forest model should be built first. The optimal parameters are set by the grid search method in the random forest model. The parameters include Max_features, Max_depth, Min_samples_split, Min_samples_leaf, Min_weight_fraction_leaf and Max_leaf_nodes [24]. The specific parameter settings of the random forest model are shown in Table 2.

Table 2. Parameter settings of random forest model.

Parameter Name	Numerical Value
Max_features	10
Max_depth	10
Min_samples_split	2
Min_samples_leaf	1
Min_weight_fraction_leaf	0
Max_leaf_nodes	5

After setting the specific parameters of the random forest, the random forest algorithm is used for model training, and the importance score of the 10-dimensional input feature variables is obtained through the generated prediction model. The importance degree of each feature variable in the model training is shown in Table 3.

Table 3. The importance of each feature variable in model training.

Variable Numbers	Input Characteristic Variables	Importance (%)
1	X coordinate of the point	14.68
2	Y coordinate of the point	37.10
3	x-axis coordinates of the cable core center	0.85
4	y-axis coordinates of the cable core center	1.23
5	Excitation current	28.52
6	Cable core conductivity	5.10
7	Relative permeability	0.65
8	Thermal conductivity of insulation	5.95
9	Convective heat transfer coefficient	4.45
10	Ambient temperature	1.48

3.3. Model Evaluation Index

The data set calculated by the finite element method for the cable temperature is preprocessed and randomly divided into a training set and a test set in proportion, of which the training set accounts for 80%, and the test set accounts for 20%. The training set is used to train the prediction model, and the test set is used to validate the model and evaluate the performance. The criterion to judge the accuracy of the model is expressed by accuracy and heel mean square deviation [25–27].

The calculation formula of accuracy r^2 is as follows:

$$r^2 = 1 - u/v \tag{11}$$

where u is the sum of squares of the difference between the true value and the predicted value, $u = \sum\left(T_{true} - T_{predict}\right)^2$; T_{true} is the real value; $T_{predict}$ is the predicted value; and v is the sum of squares of the difference between the true value and the true average.

The calculation formula of the root-mean-square error ($rmse$) is as follows:

$$rmse = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(T_{true}^{(i)} - T_{predict}^{(i)}\right)^2} \tag{12}$$

where n is the number of samples, T_{true} is the true value, and $T_{predict}$ is the predicted value.

In order to eliminate the difference between the magnitude and dimension of the feature quantities and improve the accuracy of the training results, the training data are normalized [28]. The normalized formula is as follows:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{13}$$

where X is the data set; X_{max} and X_{min} are the maximum and minimum values in the data, respectively; and X_{norm} is the normalized data value.

3.4. Construction of RF-GPR Cable Temperature Prediction Model

In the calculation of the importance score of the characteristic variables, it can be seen that among the 10-dimensional input characteristic variables, 6 input characteristic variables were more than 4% important in model training, and 4 input characteristic variables were less than 2% important in model training. Six characteristic variables with an important degree of more than 4% were selected as new input characteristic variables. The temperature prediction model of RF-GPR cable was constructed by using the Gaussian process regression algorithm in the form of the series mixing model. Gaussian process regression is a machine learning regression method developed in recent years. Its principle is based on statistical learning theory. It is a non-parametric model that takes the Gaussian process as a priori. The basic idea of Gaussian process regression is to map low-dimensional characteristic variables to high-dimensional characteristic variables and then carry out regression analysis on the processed high-dimensional characteristic variables. Gaussian process regression has good characteristics for dealing with nonlinear models and small sample data and is widely used in parameter control, model prediction and time series analysis [29].

During the construction of the cable temperature prediction model based on Gaussian process regression, the data set was first processed according to the normalization requirements, and the result after processing is shown in Formula (10).

$$D = \{(x_i, T_i)\}, i = 1, 2 \cdots n \tag{14}$$

where x_i is a group of six-dimensional input characteristic variables, and T_i is the output temperature value of the corresponding group.

Then, the basis function, kernel function, kernel size and Sigma parameters of the Gaussian process regression model were optimized. The basis function has decisive factors on the specific form of the prior mean function of the Gaussian process regression model. Generally, zero value function, constant function and linear function can be selected. Choosing different kernel functions affects the accuracy of the Gaussian process prediction model.

The mean function $m(x)$ and covariance function $k(x, x')$ together constitute the kernel functions in the Gaussian process regression, and the kernel functions affect all the statistical characteristics of Gaussian process regression. Among them, the covariance function can be divided into four types: the Mattern covariance function, square exponential covariance function, periodic covariance function and rational quadratic covariance function. Sigma is used to set the initial value of the standard deviation of observed noise [30]. For the characteristic variable type of the cable temperature output calculated by the finite element method in this paper, the optimizer type in Gaussian process regression model was set as Bayesian optimization, the basis function type as linear function, the mean function as zero mean function and the kernel function type as square exponential covariance function. The specific expression is shown as follows:

$$k(x_i, x_j) = \sigma_f^2 exp\left[-\sum_{m=1}^{n} \frac{(x_{im} - x_{jm})^2}{2\sigma_m^2}\right] \tag{15}$$

where x_{im} and x_{jm} are the variable in vectors x_i and x_j, respectively; σ_f is the variance; σ_m is the variance scale; and σ_f and σ_m constitute the hyperparameters in the covariance function.

After setting the specific parameters of the Gaussian process regression prediction model, the regression expression of the prediction model is obtained, as shown in Equation (16).

$$y \sim GP(0, k(x, x')) \tag{16}$$

Finally, the input data set of the Gaussian process regression prediction model is trained and tested. According to Bayes' principle, prior functions are established in data set D during model training, and n_* data sets are set to test the prediction model, as shown in Formula (17).

$$D_* = \{(x_i, T_i)\}, i = n+1, n+2 \cdots n+n_* \tag{17}$$

If the test results meet the set model error conditions, the cable temperature prediction model can be output; otherwise, if the test results do not meet the set model error conditions, it is necessary to readjust the hyperparameters and conduct model training again until the model error conditions are met. In the test data set, the input variable is x_*, and the output variable is y_*. The joint Gaussian distribution between the output value y of the training set and the output value y_* of the test set is obtained as follows:

$$\begin{bmatrix} y \\ y_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X,X) + \delta_n^2 I_n & K(X, x_*) \\ K(x_*, X) & k(x_*, x_*) \end{bmatrix}\right) \tag{18}$$

where $K(X, X)$ is the symmetric positive definite covariance matrix of order $n \times n$, δ_n^2 is the variance of white Gaussian noise, I_n is the unit matrix of order n, and $K(x_*, X) = K(X, x_*)^T$ is the covariance matrix of order $n \times 1$ between the prediction input variable x_* and the training input variable X. $k(x_*, x_*)$ is the covariance of the input variable x_*.

The regression equation of Gaussian process can be written as follows:

$$y_* | X, y, x_* \sim N(\overline{y}_*, \ cov(y_*)) \tag{19}$$

where $\overline{y}_*$ and $cov(y_*)$ are the mean and variance of test variable x_* to test result y_*, respectively.

4. Results and Discussion

4.1. Analysis of Prediction Results of Air Temperature in Cable Trench

After training and learning the data set of the cable trench finite element calculation model with RF-GPR, the corresponding machine learning prediction model of the air temperature distribution in the cable trench and the cable temperature distribution was obtained. Then, by setting new input characteristic quantity values different from the training set and test set, the RF-GPR cable temperature prediction model was used to predict the input characteristic quantity values. The finite element method was used to calculate the air temperature distribution in the cable trench and the cable temperature distribution under the same conditions. The resulting data were extracted and compared to verify the accuracy of the RF-GPR cable temperature prediction model. A comparison between the predicted value of the air temperature in the cable trench based on the RF-GPR cable temperature model and the calculated value of the finite element is shown in Figure 3. This paper also analyzed the prediction results of the cable temperature model based on the random forest and convolutional neural network (RF-CNN) and obtained a comparison between the predicted value of the air temperature in the cable trench based on the RF-CNN cable temperature model and the calculated value of the finite element, as shown in Figure 4. As can be seen from Figure 4, the prediction accuracy of the cable temperature model based on RF-GPR is higher. The prediction accuracy of the temperature model in the cable trench (R^2) is 0.9911, and the root-mean-square error ($RMSE$) is 0.7629. The overall temperature distribution of the air in the cable trench is between 20 °C and 60 °C. Compared with the temperature predicted by machine learning and calculated by the finite element, it can be concluded that the overall prediction effect of the air temperature in the cable trench is the best. The points of the predicted value and the calculated value fall on the diagonal line in Figure 3, that is, the predicted value is close to the calculated value. Under the same conditions, the prediction time of the cable temperature based on the Gaussian process regression algorithm is nearly 1500 times shorter than that of the finite element calculation, which effectively improves the real-time performance of cable temperature prediction and provides technical support for the application of the digital twinning evaluation of the cable running state.

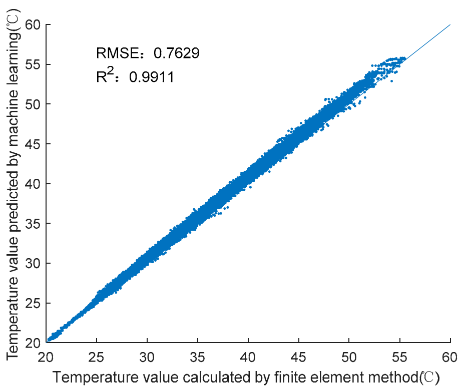

Figure 3. Accuracy analysis of RF-GPR model prediction results.

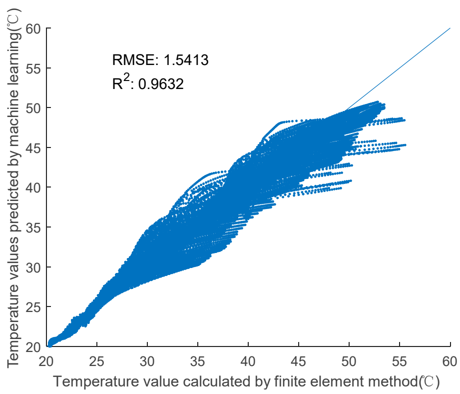

Figure 4. Accuracy analysis of RF-CNN model prediction results.

A comparison between the air temperature distribution in the cable trench based on the RF-GPR cable temperature prediction model and the air temperature distribution in the cable trench obtained by the finite element calculation method is shown in Figure 5. It can be seen from Figure 5 that the distribution trend of the model prediction results trained by the machine learning method is basically the same as that of the finite element calculation results. Through the error analysis of the temperature calculated by the finite element and predicted by machine learning, the relative error percentage of the machine learning model to predict the air temperature in the cable trench is obtained, as shown in Figure 6. It can be seen from the relative error distribution diagram that most of the regional error is below 4%, and the relative error of the lower part of the cable trench is smaller than that of the middle and upper parts. It can also be seen from the relative error distribution diagram that the relative error near the outer skin of the cable is larger than that far away from the cable. The interior of the cable belongs to solid heat transfer, and the air in the cable trench belongs to convective heat dissipation, which lead to a large temperature gradient near the cable skin, resulting in the poor prediction accuracy of the machine learning model; thus, the relative error near the cable skin is large.

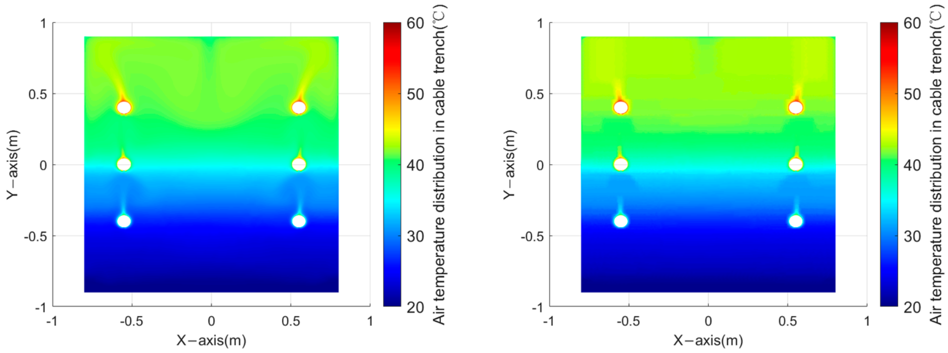

Figure 5. Comparison of finite element calculation results (**left**) and machine learning prediction results (**right**) of air temperature distribution in cable trench.

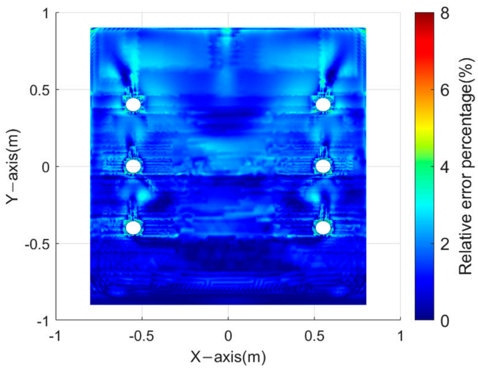

Figure 6. Relative error percentage of air temperature predicted by machine learning.

The relative errors of the machine learning predicted temperature values are further numerically analyzed below, and the relative error distribution of all data is shown in Table 4. According to the data analysis, the relative error of the temperature value predicted by machine learning is mostly between 1% and 2%, accounting for 41.22%, while the relative error of data above 5% accounts for only 0.08%. The relative error of most data is below 3%, which accounts for 91.93% of the whole data. The maximum relative error of the overall data is 6.18%, and the average relative error is 0.72%.

Table 4. Proportion of relative error distribution.

Error Distribution	0–1%	1–2%	2–3%	3–4%	4–5%	>5%
Proportion	22.38%	41.22%	28.33%	7.02%	0.97%	0.08%

4.2. Analysis of Prediction Results of Cable Temperature in Cable Trench

A comparison between the predicted cable temperature based on the RF-GPR cable temperature prediction model and the cable temperature calculated by the finite element is shown in Figure 7. It can be seen from Figure 7 that the distribution trend of the model prediction results trained by the machine learning method is basically the same as that of the finite element calculation results. Through the error analysis of the finite element calculation temperature value and the machine learning prediction temperature value, the relative error percentage distribution of the cable temperature distribution predicted by the RF-GPR cable temperature prediction model in the cable trench is obtained, as shown in Figure 8. The relative error of the cable temperature predicted by the machine learning model is minimum in the cable core part, and the relative error distribution is basically within 0.1%. The relative error distribution outside the cable shows an increasing trend, and the average relative error of the temperature prediction value is 0.17%. Near the cable sheath, the machine learning prediction results are not as accurate as the finite element calculation results. This is due to the fact that the cable jacket is located on the outer layer of the cable, and the temperature gradient drops considerably compared to the cable core temperature. The decrease in temperature gradient near the cable bottom skin is particularly obvious; thus, the machine learning model near the cable bottom skin has poor temperature prediction accuracy, and the relative error reaches the maximum of 1.54%. At the same time, the training feature data near the cable sheath are less than the data set near the cable core, which causes the prediction results to not be as accurate as those near the cable core. In the research on cable temperature calculation, the most noteworthy value is the temperature of the cable core. The influence of the temperature distribution at the cable sheath on cable operation is much smaller than that at the cable core temperature, so the prediction accuracy of a few points at the cable sheath is negligible.

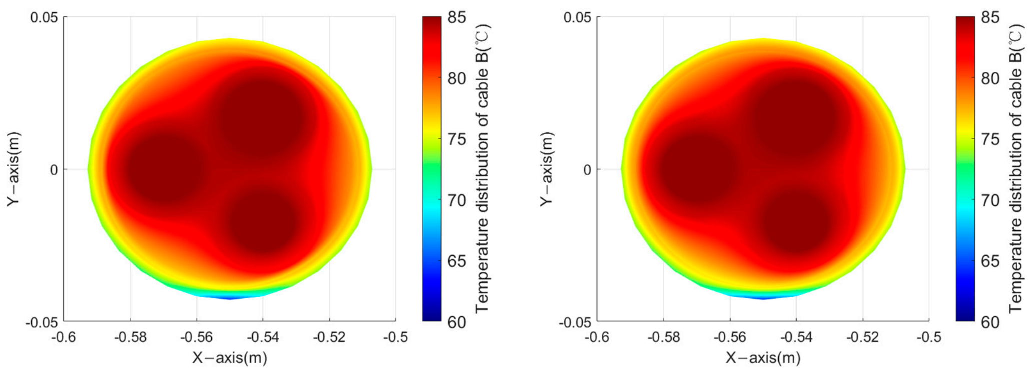

Figure 7. Comparison of finite element calculation results (**left**) and machine learning prediction results (**right**) of cable temperature distribution in cable trench.

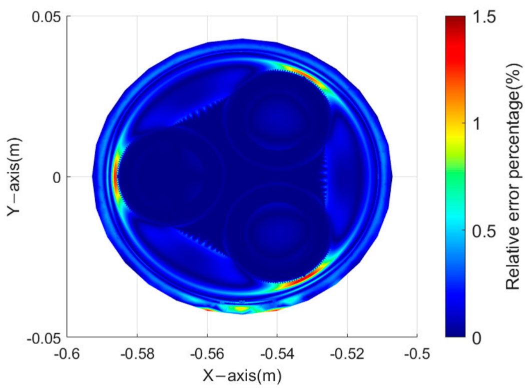

Figure 8. Relative error percentage of cable temperature predicted by machine learning.

5. Digital Twin Platform for Cable Temperature Calculation Based on RF-GPR

Based on the research of the cable temperature calculation method and artificial intelligence technology, the digital twin frame of cable temperature calculation is further designed, as shown in Figure 9. Referring to the five-dimensional model of digital twin technology in the industrial field, a digital twin cable model (DTCTM) is proposed, which includes five parts: physical entity, digital twin, twin data, virtual–real connection and intelligent application, as shown in Equation (20).

$$DTCM = \{P_{CT}, V_{CT}, DT_{Data}, VRC, DT\&A\} \quad (20)$$

where P_{CT} indicates the physical entity of the cable, V_{CT} represents a digital twin; DT_{Data} indicates the twin data, and VRC represents the connection between the virtual and real systems. $DT\&A$ represents the application provided by the twin system.

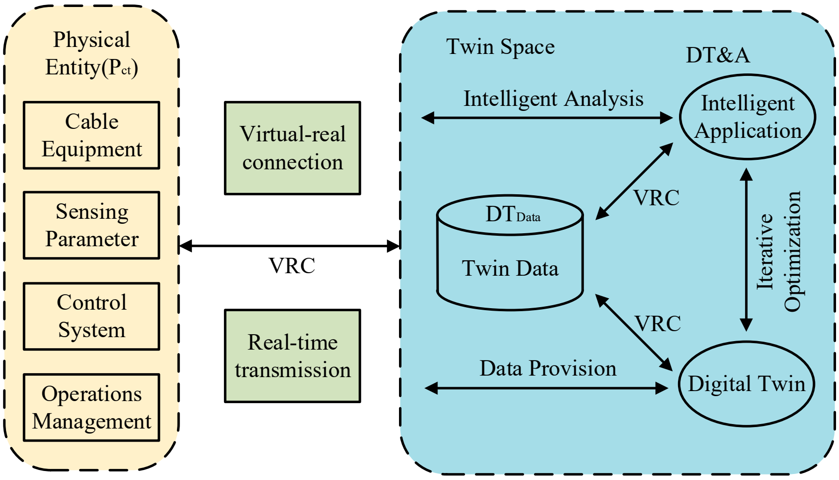

Figure 9. Design of digital twin frame for cable temperature calculation.

The digital twin of the cable temperature calculation is the development form of the coexistence of the physical space entity and the digital cable. By accepting the physical parameters and distributed sensing data from the physical cable entity, the digital space establishes a virtual space matching the physical cable entity, synchronously evolves with the physical cable entity, and then reflects the state of the cable model in the real environment in the form of dynamic monitoring, real-time diagnosis and accurate prediction. In this way, the real-time visualization of the full state of the cable model and the intelligent operation and maintenance are promoted to realize the comprehensive and accurate monitoring of the physical cable entity, and the analysis results such as the diagnostic data and predictive data are fed back into the cable's physical entity for the next step of scheduling control. Through the combination of the cable monitoring system and digital twin technology, the formation of the continuous learning and evolution of the intelligent cable situation management system, the reasonable control of the cable load transport and operation safety monitoring occurs, thus promoting the durable and safe operation of the cable system as a whole, opening up a new model of digital smart grid operation and maintenance [31,32].

Constructing the digital twin platform for cable temperature calculation is an important step in the design of the digital twin frame for cable temperature calculation, which should consider the connection between virtual space and the cable entity. This paper builds a data platform based on the RF-GPR cable temperature prediction model and uses digital technologies such as smart sensors and the Internet of things to digitally describe the characteristics, parameters and operating conditions of the cable trench's physical entity, forming a variable database including geometric parameters, electromagnetic parameters and material parameters. In addition, during the operation of the cable system, intelligent perception technology and model analysis technology should also be used to import sensing and simulation related models and operation data into the digital twin space to improve the data platform of the digital twin model for cable temperature calculation.

In the optimization of the operation process of the cable entity, the digital twin space carries out virtual–real interaction and information sharing with the cable entity through the connection technology of real-time transmission, uses machine learning prediction technology in the virtual space to conduct simulation analysis and accurate prediction, and then uses the optimization control technology to feedback the analysis results into the cable entity. It supports the safe and stable operation of cable entities, while providing twin system applications such as safety warning, temperature monitoring and fault analysis, and presents the analysis results to users of the digital twin platform. The interface of the digital twin platform for cable temperature calculation is shown in Figure 10.

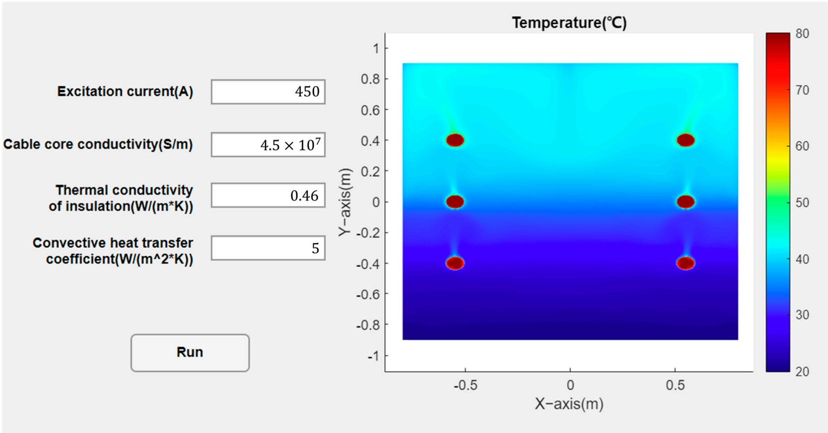

Figure 10. Cable temperature calculation digital twin platform interface.

6. Conclusions

To solve the problem that the finite element method takes a long time to calculate the multi-physical field coupling model, a cable temperature prediction model construction method based on machine learning is proposed. The random forest algorithm is used to calculate the importance score of the characteristic variables, and six-dimensional input characteristic variables with an importance that is greater than 4% in model training are selected. The cable temperature prediction model based on RF-GPR is constructed by the series mixing model. The model accuracy (R^2), root-mean-square error ($RMSE$) and average relative error were 0.9879 and 0.8564, respectively. The average relative error of the cable temperature prediction in the cable trench is 0.17%, which can replace the finite element method to predict the temperature distribution in the cable trench. Under the same conditions, the prediction time of the cable trench temperature based on the Gaussian process regression algorithm is nearly 1500 times shorter than that of the finite element calculation, which effectively improves the real-time performance of cable temperature prediction. A digital twinning platform for cable temperature calculation based on RF-GPR is designed to provide technical support for the application of digital twinning for cable running state evaluation.

Author Contributions: Conceptualization, W.H.; methodology, G.Y.; validation, C.H.; writing—original draft preparation, G.Y. and D.K.; writing—review and editing, D.K.; project administration, C.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of State Grid Hebei Electric Power Company (kj2021-060).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, Y.D.; Chen, S.; Cong, Z.H.; Yan, Y.J. Key technologies and application prospects of digital twin in power equipment industry. *High Volt Technol.* **2021**, *47*, 1539–1554.
2. Schleich, B.; Anwer, N.; Mathieu, L.; Wartzack, S. Shaping the Digital Twin for Design and Production Engineering. *CIRP Ann.* **2017**, *66*, 141–144. [CrossRef]
3. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* **2019**, *7*, 167653–167671. [CrossRef]
4. Kantaros, A.; Piromalis, D. Setting up a Digital Twin Assisted Greenhouse Architecture. *Am. J. Eng. Appl. Sci.* **2022**, *15*, 230–238. [CrossRef]
5. Chaux, J.D.; Sanchez-Londono, D.; Barbieri, G. A Digital Twin Architecture to Optimize Productivity within Controlled Environment Agriculture. *Appl. Sci.* **2021**, *11*, 8875. [CrossRef]
6. Piromalis, D.; Kantaros, A. Digital Twins in the Automotive Industry: The Road toward Physical-Digital Convergence. *Appl. Syst. Innov.* **2022**, *5*, 65. [CrossRef]
7. Stark, R.; Fresemann, C.; Lindow, K. Development and operation of Digital Twins for technical systems and services. *CIRP Ann.* **2019**, *68*, 129–132. [CrossRef]
8. Ma, C.; Liang, K.; Li, X.P.; Peng, Y.X.; Huang, Y.Y.; Chen, B. Application Exploration of Digital Twin Technology in High Voltage Cable Field. *High Volt. Electr. Appar.* **2023**, *59*, 70–76+85.
9. Qi, B.; Zhang, P.; Zhang, S.Q.; Zhao, L.J.; Wang, H.B. Application Status and Prospect of Digital Twin Technology in State Evaluation of Power Transmission and Transformation Equipment. *High Volt Technol.* **2021**, *47*, 1522–1538.
10. Negri, E.; Berardi, S.; Fumagalli, L.; Macchi, M. MES-integrated digital twin frameworks. *J. Manuf. Syst.* **2020**, *56*, 58–71. [CrossRef]
11. Kantaros, A.; Piromalis, D.; Tsaramirsis, G.; Papageorgas, P.; Tamimi, H. 3D Printing and Implementation of Digital Twins: Current Trends and Limitations. *Appl. Syst. Innov.* **2022**, *5*, 7. [CrossRef]
12. Shang, H.Y.; Liu, L.Q.; Qi, Y.S. Research on wind turbine modeling based on digital twin technology. *Acta Energ. Sol. Sin.* **2023**, *44*, 391–400.
13. Zhou, M.; Yan, J.; Zhou, X. Real-time online analysis of power grid. *CSEE J. Power Energy Syst.* **2020**, *6*, 236–238.
14. Qi, Q.; Tao, F. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access* **2018**, *6*, 3585–3593. [CrossRef]
15. Lux, J.; Czerniuk, T.; Olschewski, M.; Hill, W. Non-Concentric Ladder Soil Model for Dynamic Rating of Buried Power Cables. *IEEE Trans. Power Deliv.* **2021**, *36*, 235–243. [CrossRef]
16. Shen, C.; Cao, Q.N.; Huang, M.S.; Chen, Y.; Huang, S.W. Concept, Characteristics and Application Prospect of Digital twinning in Power System. *Proc. CSEE* **2022**, *42*, 487–499.
17. Chi, P.; Zhang, Z.; Liang, R.; Cheng, C.; Chen, S.K. A CNN recognition method for early stage of 10 kV single core cable based on sheath current. *Electric Power Syst. Res.* **2020**, *184*, 106292. [CrossRef]
18. Zeng, F.P.; Dong, Y.L.; Tang, J. Feature extraction and severity assessment of partial discharge under protrusion defect based on fuzzy comprehensive evaluation. *IET Gener. Transm. Distrib.* **2015**, *9*, 2493–2500. [CrossRef]
19. Dai, J.J.; Song, H.; Sheng, G.H. Prediction method for power transformer running state based on LSTM network. *High Volt. Eng.* **2018**, *44*, 1099–1106.
20. Gulbudak, O.; Gokdag, M. Finite control set model predictive control approach of nine switch inverter-based drive systems: Design, analysis, and validation. *ISA Trans.* **2021**, *110*, 283–304. [CrossRef]
21. Zheng, C.; Wang, S.R.; Liu, Y.L.; Liu, C. A novel RNN based load modelling method with measurement data in active distribution system. *Electr. Power Syst. Res.* **2019**, *166*, 112–124. [CrossRef]
22. Pu, L.; Duan, W.; Sun, J.; Zhao, X.F.; Xiong, L. Calculation of core temperature of Single-core cable Joint based on Finite Element Method. *Power Grid Clean Energy* **2021**, *37*, 57–63.
23. Wu, J.; Fang, L.C.; Liu, X.T. SOH estimation of lithium ion battery based on feature optimization and Random Forest algorithm. *J. Mech. Eng.* **2023**, 1–9.
24. Li, B.; Han, R.; He, Y.G.; He, X.Y.; Hou, J.B. Application of Improved Random Forest Algorithm in Motor Bearing Fault Diagnosis. *Proc. CSEE* **2006**, *40*, 1310–1319+1422.
25. Han, F.J.; Wang, X.H.; Qiao, J.; Shi, M.J.; Pu, T.J. A review of Load forecasting for new Power Systems based on Artificial Intelligence. *Proc. Chin. Soc. Electr. Eng.* **2023**, 1–24.
26. Giamarelos, N.; Zois, E.N.; Papadimitrakis, M.; Stogiannos, M.; Livanos, N.-A.I.; Alexandridis, A. Short-Term Electric Load Forecasting with Sparse Coding Methods. *IEEE Access* **2021**, *9*, 102847–102861. [CrossRef]
27. Jawad, M.; Sajjad Ahmed Nadeem, M.; Shim, S.-O.; Khan, I.R.; Shaheen, A.; Habib, N.; Hussain, L.; Aziz, W. Machine Learning Based Cost Effective Electricity Load Forecasting Model Using Correlated Meteorological Parameters. *IEEE Access* **2020**, *8*, 146847–146864. [CrossRef]
28. Zhao, Y.; Wang, H.M.; Kang, L. Short-term Power Load Prediction based on Time Convolutional Network. *Trans. China Electrotech. Soc.* **2022**, *37*, 1242–1251.
29. Zhang, Y.; Liu, B.J.; Shao, Z.G. General Harmonic Source Uncertainty Model based on Gaussian Process Regression. *Proc. CSEE* **2022**, *42*, 992–1002.

30. Yang, J.; Wang, L.; Song, D.R.; Dong, M.; Chen, S.F.; Huang, L.X. A Diagnosis method of zero drift of wind turbine yaw Angle Based on isolated forest and Sparse Gaussian Process Regression. *Proc. CSEE* **2021**, *41*, 6198–6212.
31. Wang, X.; Wang, L.; Yu, Y. Survey on Characteristics, Architecture and Applications of Digital Twin Power Grid. *J. Electron. Inf. Technol.* **2022**, *44*, 3721–3733.
32. Yang, S.C.; Li, Q.W.; Zhou, S.D. Construction method of digital twin battery for intelligent management. *J. Beijing Univ. Aeronaut. Astronsutics* **2022**, *48*, 1734–1744.