International Journal of Heat and Technology
Vol. 43, No. 1, February, 2025, pp. 103-111
Journal homepage: http://iieta.org/journals/ijht

Measurement Error Correction and Thermal Property Prediction Based on Raman Scattering and Neural Networks

Wanfang Gao*, Liqing Ren

School of Energy Engineering, Yulin University, Yulin 719000, China

Corresponding Author Email: gaowanfang@yulinu.edu

https://doi.org/10.18280/ijht.430112

Received: 9 June 2024
Revised: 7 December 2024
Accepted: 20 December 2024
Available online: 28 February 2025

Keywords:
Raman scattering, thermal property prediction, measurement error correction, neural networks, deep learning, TCN, SA mechanism

ABSTRACT

With the rapid development of materials science and engineering, the measurement and prediction of thermal properties have become a key research focus. Traditional methods for thermal property testing, while providing accurate experimental data, often require complex experimental conditions and lengthy testing periods, limiting their widespread application in real-time detection and efficient usage. Raman scattering, as a non-destructive and highly sensitive analytical technique, can effectively capture molecular structure and thermal property information of samples, making it increasingly popular in the field of thermal property measurement. However, Raman scattering signals are susceptible to interference from factors such as environmental temperature and laser power, leading to experimental errors that affect the accuracy of thermal property predictions. Currently, methods for predicting thermal properties based on Raman scattering signals primarily employ machine learning and deep learning techniques, but there remains significant room for improvement in accuracy, particularly when dealing with complex samples and highly nonlinear signals. Furthermore, existing error correction methods lack real-time adaptability, limiting their application in dynamic environments. To address these issues, this paper proposes a neural network-based model for Raman scattering measurement error correction and thermal property prediction. By constructing a thermal property prediction model and error correction model based on Random Forest (RF)-temporal convolutional network (TCN)-self-attention (SA), this study effectively improves prediction accuracy and the reliability of experimental data, providing theoretical support and technical assurance for the precise application of Raman scattering technology.

1. INTRODUCTION

With the continuous development of technology, the thermal property characteristics of materials play a crucial role in many fields such as materials science, energy engineering, environmental monitoring, etc. [1-5]. Traditional thermal property testing methods, such as steady-state and transient-state methods, while providing relatively accurate measurement data, often require complex experimental equipment and long experimental processes, which limits their application in certain high-efficiency and real-time monitoring scenarios [6-9]. Raman scattering technology, as a high-sensitivity and non-destructive testing method, has gradually become an important tool for studying thermal properties because it can provide information on molecular structure, composition, and thermal properties [10, 11]. With the advancement of sensor technology and the improvement of data processing capabilities, thermal property measurement based on Raman scattering technology has become one of the current research hotspots.

However, Raman scattering measurement technology faces some challenges during application, especially in terms of measurement errors and thermal property prediction accuracy. Since Raman scattering signals are easily affected by factors such as environmental temperature, laser power, and sample morphology, the actual measurement results often deviate [12, 13]. Additionally, due to the complex nonlinear relationship between Raman scattering signals and the thermal properties of the sample, extracting accurate thermal property information from Raman spectra has become a difficult research issue [14, 15]. Therefore, studying how to accurately predict thermal properties and correct measurement errors based on Raman scattering signals is of significant importance for enhancing the application value of Raman scattering technology.

Currently, many scholars have proposed various thermal property prediction models based on Raman scattering signals, and the introduction of machine learning and deep learning methods has brought new developments to this field. However, most existing methods still have considerable errors when facing complex samples, especially in modeling the nonlinear mapping relationship between Raman scattering signals and thermal properties, and still lack effective and accurate correction methods [16-18]. For example, traditional regression analysis methods often overlook signal noise and the interaction between multiple variables, resulting in poor generalization ability of the prediction models. Moreover, the real-time performance and adaptability of existing error

correction methods in dynamic monitoring are poor, making it difficult to meet the practical application needs in complex environments [19-21]. Therefore, improving the processing accuracy of Raman scattering signals and the prediction accuracy of thermal properties has become an urgent issue to address.

This paper aims to propose a comprehensive method for Raman scattering measurement error correction and thermal property prediction based on neural networks. Specifically, this paper first establishes a thermal property prediction model based on RF-TCN-SA, and by introducing TCN and the SA mechanism, effectively captures the features related to thermal properties in the Raman spectra, thereby enhancing prediction accuracy. Secondly, this paper constructs an error correction model for Raman scattering measurement, which uses neural network methods to perform real-time correction of measurement deviations in experiments, ensuring the accuracy of thermal property prediction results. Through the combination of these two models, not only can more accurate thermal property prediction be achieved, but errors that may occur during the experimental process can also be effectively reduced, improving the reliability of Raman scattering technology in practical applications. This research provides a new solution for the precise application of Raman scattering technology, with significant theoretical and practical implications.

2. ESTABLISHMENT OF THERMAL PROPERTY PREDICTION MODEL BASED ON RF-TCN-SA

The neural network-based thermal property prediction model constructed in this paper requires accurate input data sources, which must include information after the Raman scattering measurement error correction. First, the Raman scattering measurement data itself is an important input source. Since Raman scattering signals are affected by noise, instrument bias, and various other factors, data that has not been error-corrected may affect the subsequent thermal property prediction. Therefore, error correction in the data preprocessing step is crucial. Specifically, in the input of the neural network model, the Raman scattering spectral data, after correction, should include the correct peak positions, peak intensities, and their correlation with the material's thermal properties. Additionally, experimental control parameters such as temperature, pressure, and other conditions, as well as the chemical composition and structural features of the sample, are also indispensable input data. These factors influence the Raman scattering performance, indirectly affecting the thermal property prediction.

Specifically, this paper constructs a thermal property prediction model based on RF-TCN-SA. To improve the prediction efficiency and accuracy of the model, this paper uses feature selection methods to optimize the input data. Since Raman scattering spectral data contains a large number of potential features, some of these features may be redundant or irrelevant for thermal property prediction. Directly inputting all features into the model would greatly increase the computational load and may even affect the training efficiency and accuracy of the model. Therefore, this paper introduces traditional feature selection techniques such as the Pearson correlation coefficient method and gray relational analysis method. By analyzing the correlation between each feature and thermal property parameters, the features most closely related to thermal properties are selected. These methods can effectively remove weakly correlated or irrelevant features, retaining the key information that significantly impacts the prediction results, thus simplifying the input data, reducing computational load, and improving the training efficiency of the model. With the data after feature selection, this paper further uses the RF-TCN-SA model to provide a new solution for thermal property prediction. RF, as a powerful feature selection and evaluation tool, evaluates each feature's contribution to the prediction result, which helps further select the most important features, improving the model's robustness and prediction ability. The TCN, with its strong ability to model time-series data, can capture complex dynamic changes between different time or frequency points in the Raman spectrum. Finally, the addition of the SA mechanism helps the model focus on the most important features for thermal property prediction, automatically adjusting the weights of different input features, further enhancing the model's prediction accuracy.

2.1 RF

RF is an ensemble learning algorithm that enhances the model's generalization ability by combining the results of multiple decision trees. Its advantage lies in effectively addressing complex and nonlinear prediction problems, which makes it perform excellently in thermal property prediction. RF generates a large number of decision trees and integrates their prediction results, achieving better performance in prediction tasks compared to a single decision tree. In the thermal property prediction of this paper, the core role of RF is to help the model identify the features most closely related to thermal properties by performing feature selection, evaluation, and optimization on the Raman scattering spectral data.

During the training process of RF, the Bootstrap resampling technique is first used to randomly sample the training set from the corrected Raman scattering dataset with replacement. Suppose the original corrected dataset is L, and by randomly sampling $L(u)$ samples, multiple different subsample sets are formed. These subsample sets do not need to completely cover the original dataset, but through repeated sampling, the diversity of the training data is ensured. This method of sample generation ensures that each decision tree is trained on different subsets of data, avoiding the overfitting problem caused by a single dataset. During the training process of each decision tree, the RF algorithm does not use all the features, but randomly selects other appropriate subsets of features from the total feature set L as candidate features. This strategy reduces the correlation between features, avoids the model from overly relying on certain features, and enhances the diversity of each tree. In thermal property prediction, Raman scattering data may contain multiple redundant features, and by randomly selecting features, irrelevant or noisy features can be effectively removed, thereby improving the model's robustness and prediction ability.

During the training process of each decision tree, RF uses the Gini impurity minimization principle to select the best feature for node splitting. For each candidate feature subset, the algorithm calculates the Gini impurity that each feature may cause after splitting and selects the feature with the minimum Gini value for node division. In thermal property prediction, there is a complex nonlinear relationship between different features of the Raman scattering data and thermal

properties. By minimizing Gini impurity, the model can accurately capture the correlation between features and thermal properties, thereby improving prediction accuracy. During the training process of RF, each decision tree grows to its maximum extent until all samples are classified or a certain termination condition is met, without pruning. This strategy ensures that each tree can fully learn the information in the data. Especially in thermal property prediction, Raman scattering data may have complex patterns that are difficult to capture by simplified rules. Therefore, by avoiding pruning, the model can retain the predictive ability of each tree to the maximum extent. This is particularly helpful when dealing with complex and high-dimensional data, providing more information sources to aid the final prediction result.

After all decision trees are trained, RF aggregates the prediction results of the multiple decision trees according to the ensemble learning principle to obtain the final thermal property prediction value. In regression tasks, the aggregation method usually involves averaging, while in classification tasks, a voting mechanism is used. In the specific application of thermal property prediction, aggregating the prediction results of multiple trees can effectively reduce the errors or biases that may arise from a single tree. Especially when there is a lot of data noise, the aggregated result significantly improves the stability and accuracy of the prediction. Figure 1 shows the RF flowchart.

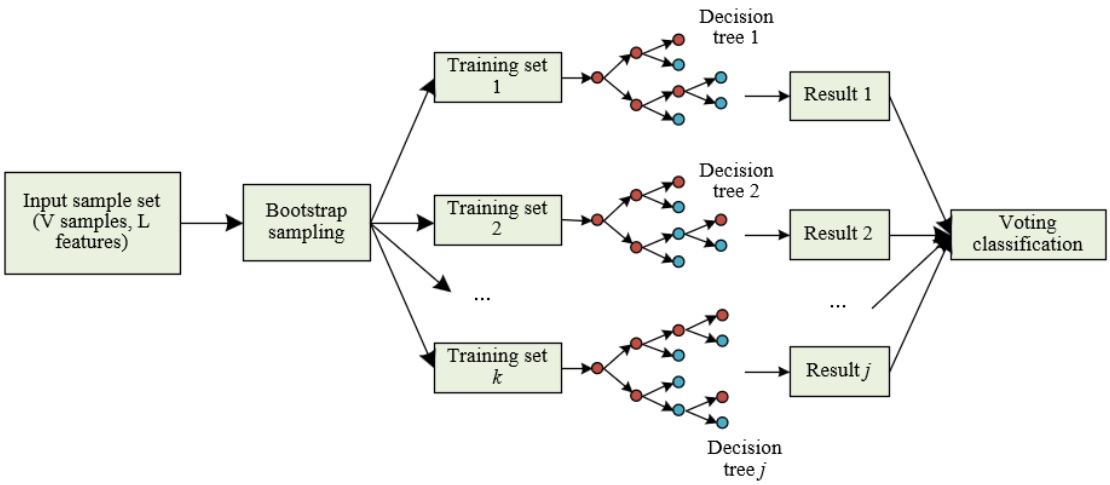


Figure 1. RF flowchart

2.2 Feature selection

In the model, feature selection is a crucial step, especially when processing the corrected Raman scattering spectral data. Feature selection not only improves the model's prediction accuracy but also reduces the data dimensionality, improves computational efficiency, and prevents overfitting. Below, we describe the three basic steps of feature selection for thermal property influencing factors data:

Step 1: Calculate Out-of-Bag (OOB) error rate

During the training process of RF, the Bootstrap resampling technique is used to generate multiple training sample sets, while keeping the samples that were not selected as OOB data. In the training process of each decision tree, the OOB data does not participate in training, so it can be used to evaluate the performance of the decision tree. Specifically, for a decision tree s_l in the RF, the misclassification rate is calculated using its OOB data. This misclassification rate reflects the model's prediction error on unseen data and is an important basis for feature selection. In the corrected Raman scattering spectral data, the OOB error rate can effectively assess the correlation between features and thermal properties and their impact on the model's prediction. Assuming there are v decision trees in total, the number of prediction errors is represented by $l_{ER}(a_u)$, and the total number of predictions is represented by $l_{AR}(a_u)$, then for a specific feature a_u randomly selected by a certain tree s_l in the RF, the following formula gives the misclassification rate of the decision tree $t_m s_l$:

$$P_l(a_u) = \frac{l_{ER}(a_u)}{l_{AL}(a_u)} \times 100\% \tag{1}$$

Step 2: Calculate feature accuracy decrease

When evaluating feature importance, keeping other features unchanged, a feature a_u is randomly selected and noise is added to it, i.e., the feature is randomly assigned a value, and the misclassification rate of decision tree s_l is calculated again. Suppose the misclassification rate before adding noise is OBB_{BF}, and the misclassification rate after adding noise is OBB_{AF}, then the accuracy decrease of the feature is $\Delta_{OOB}=OBB_{AF}-OBB_{BF}$. This accuracy decrease reflects the importance of feature a_u to the model prediction: if a feature is crucial to the model prediction, the misclassification rate will increase significantly after noise is added. In the corrected Raman scattering spectral data, this method can identify which features are most important for thermal property prediction. Suppose after randomly assigning values to feature a_u, the accuracy decrease of the decision tree is represented by $L_u(a_.)$, the OOB error rate of the original feature sample set is represented by $P_l(a_u)$, and the OOB error rate after random assignment of feature au is represented by $P_l'(a_u)$. The corresponding accuracy decrease of the feature can be calculated using the following formula:

$$L_u(a_u) = P_l'(a_u) - P_l(a_u) \tag{2}$$

Step 3: Calculate average accuracy decrease and feature selection

Repeat the above steps and calculate the accuracy decrease of feature a_u for each decision tree in the RF, then take the average as the importance score of the feature. These importance scores can be used to rank and select features, choosing the features with higher scores as the model's input

variables. In thermal property prediction, this process can effectively filter out the features most closely related to thermal properties, reducing data dimensionality and improving the model's computational efficiency and prediction accuracy. For Raman scattering spectral data, feature selection helps the model focus on the most informative features, such as the corrected peak positions and peak intensities, further enhancing the performance of thermal property prediction. Suppose the average accuracy decrease of feature a_u is represented by $L(a_u)$, and the corresponding average accuracy decrease of the feature can be calculated using the following formula:

$$L(a_u) = \frac{1}{v}\sum_{u=1}^{v} L_u(a_u) \tag{3}$$

2.3 SA mechanism

The actual Raman scattering data typically include multiple spectral features, which have complex correlations and potential temporal dependencies with the thermal property parameters of the substance. Traditional TCN models, although capable of effectively extracting temporal features through convolution operations, may not fully capture the correlations between long-distance inputs when dealing with long-term dependencies. The SA mechanism can dynamically adjust the weight of each time step by calculating the similarity between the *Query* of each input and the *Key* of other inputs, enabling effective focus on long historical sequence information. Figure 2 shows an illustration of the attention mechanism calculation process. Through this mechanism, the model can focus more attention on the time steps and features most crucial for thermal property prediction, significantly improving prediction accuracy. Let $A=[a_1,a_2,\ldots,a_v]$ represent the model input data, and the calculation process of the SA mechanism is as follows:

$$\begin{cases} W = Q^W A \\ J = Q^J A \\ N = Q^N A \end{cases} \tag{4}$$

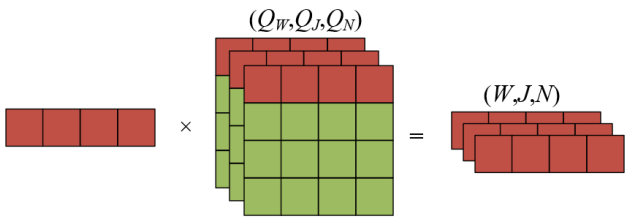


Figure 2. Illustration of the attention mechanism calculation process

Suppose the weight vector is represented by β_k, the adjustment parameter is represented by $1/(f_j)^{1/2}$, and the dimension of the Key matrix is represented by f_j. The calculation process of the SA mechanism's weight coefficient is as follows:

$$\beta_k = \text{softmax}\left(J^S w_k\right) \in \Re^l \tag{5}$$

Figure 3 shows the schematic diagram of the constructed TCN-SA model.

$$ATT(W,J,N) = \text{softmax}\left(\frac{WJ^S}{\sqrt{f_j}}\right)N \tag{6}$$

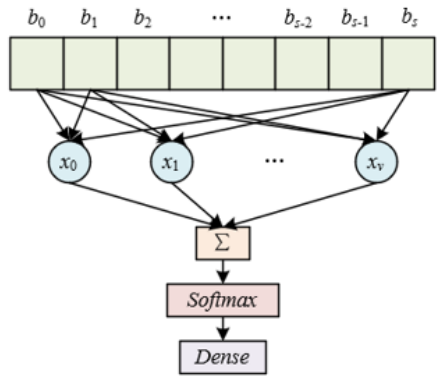


Figure 3. Schematic diagram of the TCN-SA model

3. RAMAN SCATTERING MEASUREMENT ERROR CORRECTION MODEL

The generation of Raman scattering measurement errors mainly results from the following three aspects, which not only affect the accuracy of Raman scattering data but also directly impact the precision of thermal property predictions.

(1) Instrumental Noise and Environmental Interference: Raman scattering signals are inherently weak and susceptible to interference from instrumental noise and external environmental factors. Instrumental noise, particularly from components such as the optical system, detectors, and laser sources, can cause amplitude deviations or spectral distortion. Additionally, fluctuations in environmental conditions, such as changes in temperature and humidity in the laboratory, can also affect the quality of the Raman scattering signal.

(2) Errors from Sample Preparation and Handling: The preparation and handling of the sample during the Raman scattering measurement process also have a significant impact on the results. The surface condition, thickness, uniformity of the sample, and the contact method between the sample and the detection system can all introduce errors. Furthermore, the physical and chemical properties of the sample may interfere with the Raman signal, causing fluorescence or other stray light to mix with the Raman signal.

(3) Complexity and High Dimensionality of Raman Scattering Spectral Data: Raman scattering spectral data itself has high complexity and dimensionality. Overlapping peaks, band width, and position shifts can lead to measurement errors. Raman scattering spectra from different substances often share similar peak characteristics, particularly in complex samples where peak interference is common. This complexity increases the difficulty of analyzing Raman scattering signals and can lead to incorrect signal interpretation. Additionally, due to the diversity of Raman scattering signals, the spectral data may contain redundant information and irrelevant features. If this redundant information is not effectively processed, it can introduce noise interference into the thermal property prediction model.

In order to improve the accuracy of thermal property prediction, this paper first calculates the original Raman scattering measurement error sequence using known Raman scattering spectral data and the corresponding thermal property data. This error sequence refers to the difference

between the preliminary predicted values and the true thermal properties, reflecting noise, bias, and sources of error in the Raman scattering spectral data. Through the analysis of this error sequence, the patterns and characteristics of the errors can be clarified, providing an important basis for subsequent error correction. Next, to carry out effective error correction, we use a neural network to construct a dynamic error correction model. This model can extract the inherent patterns of the time series from a large amount of historical Raman scattering data and use these patterns to correct future thermal property prediction values. The specific process is as follows:

First, based on the given original Raman scattering spectral data sequence $O=(a_1,a_2,...,a_v)$ and the model's preliminary thermal property prediction sequence $O'=(a'_1,a'_2,...,a'_v)$, the Raman scattering measurement error sequence is calculated. This error sequence reflects the difference between the preliminary predicted values and the true thermal properties, revealing the bias and noise in the Raman scattering spectral data. The formula for calculating the error sequence is:

$$r = O - O' \tag{7}$$

After obtaining the error sequence, this paper uses a neural network to construct the error correction model. Let the prediction time point be represented by s, and the prediction time scale by S. The error prediction value at time s is represented by $r'(e)$, and the error correction model is established as follows:

$$r'(s) = d\left[r(s-S), r(s-2S), ..., r(s-vS)\right] \tag{8}$$

The corrected thermal property prediction value is represented by O'', and the final output of the corrected model is:

$$O'' = O' + r' \tag{9}$$

To improve the prediction accuracy of the error correction model, this paper uses the Savitzky-Golay (SG) smoothing method to preprocess the error sequence. The SG smoothing method is a moving window least squares polynomial smoothing technique. The core idea is to perform polynomial fitting within a time window and calculate the smoothed value at the center point. Specifically, the SG smoothing method first selects a time window containing $v=2l+1$ data points, and performs polynomial fitting of order j as shown in the following formula, calculating the weighted coefficients of the center point with respect to the surrounding points, and then obtaining the smoothed estimate of the center point. The window moves continuously, providing the smoothed result for the entire error sequence. The larger the time window v, the better the data smoothing effect; the higher the polynomial fitting order j, the more detail in the original data can be preserved. The fitting formula is as follows:

$$b = x_0 + x_1 a + x_2 a^2 + ... + x_{j-1} a^{j-1} \tag{10}$$

Let the data before smoothing be represented by a_u, the data after smoothing by a_u^*, and the moving window smoothing coefficient by g_k, then:

$$a_u^* = \frac{\sum_{k=-l}^{l} a_{u+k} g_k}{G} \tag{11}$$

Figure 4 shows the prediction flowchart of RF-TCN-SA thermal property prediction based on error correction.

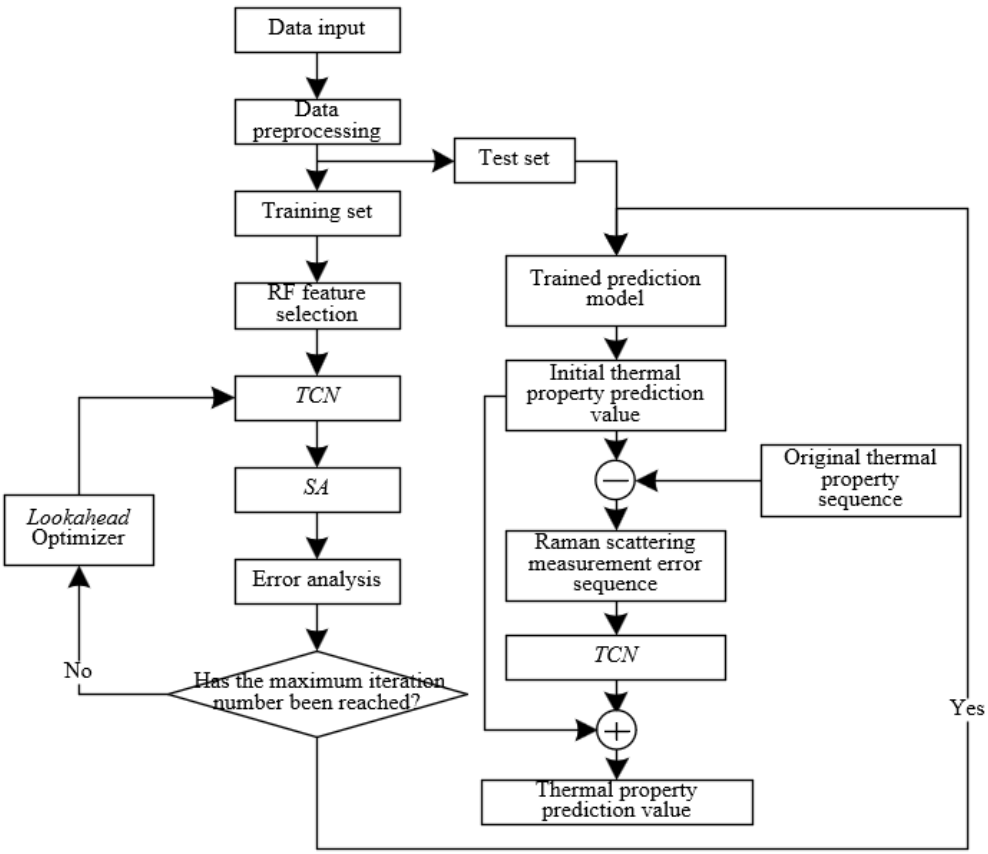


Figure 4. RF-TCN-SA thermal property prediction flowchart based on error correction

Table 1. Feature importance ranking of corrected Raman scattering spectral data (Sample set 1)

Pearson Correlation Coefficient Method		Grey Relational Degree		The Proposed Method	
Peak Position	0.82	Peak Area	0.81	Peak Position	0.6125
Peak Intensity	0.81	Peak Width	0.81	Peak Intensity	0.5415
Peak Width	0.75	Peak Intensity	0.81	Frequency Shift	0.4289
Peak Area	0.74	Peak Position	0.81	Peak Width	0.3895
Signal-to-Noise Ratio after Baseline Correction	0.076	Isotope Effect	0.76	Time-Resolved Feature	0.3215
Relative Peak Intensity	0.073	Time-Resolved Feature	0.76	Frequency Shift Polarization Effect	0.2862
Frequency Drift Dependency	0.072	Polarization Effect	0.73	Frequency Shift Dependency	0.1456
Frequency Drift	0.057	Signal-to-Noise Ratio after Baseline Correction	0.71	Isotope Effect	0.1428
Polarization Effect	0.019	Relative Peak Intensity	0.71	SNR after Baseline Correction	0.1326
Time-Resolved Feature	0.0093	Time-Resolved Feature	0.71	Time-Resolved Feature	0.1158
Isotope Effect	0.0054	Frequency Drift Dependency	0.72	Relative Peak Intensity	0.0985

Table 2. Ranking of feature information importance for corrected Raman scattering spectral data (Sample set 2)

Pearson Correlation Coefficient Method		Grey Relational Degree		The Proposed Method	
Peak Position	0.83	Peak Area	0.65	Peak Position	6.1256
Peak Intensity	0.81	Peak Width	0.65	Peak Intensity	4.2368
Peak Width	0.78	Peak Intensity	0.65	Peak Width	2.1524
Peak Area	0.73	Peak Position	0.65	Peak Area	1.3258
Frequency Shift Dependency	-0.47	Time-Resolved Feature	0.65	Polarization Effect	0.8254
Relative Peak Intensity	-0.46	Frequency Shift	0.65	Time-Resolved Feature	0.7895
SNR after Baseline Correction	-0.44	Polarization Effect	0.61	Frequency Shift	0.6123
Isotope Effect	-0.43	Isotope Effect	0.57	Frequency Shift Dependency	0.1758
Time-Resolved Feature	-0.062	SNR after Baseline Correction	0.57	Relative Peak Intensity	0.1526
Frequency Shift	-0.017	Relative Peak Intensity	0.57	SNR after Baseline Correction	0.1236
Polarization Effect	-0.015	Frequency Shift Dependency	0.57	Isotope Effect	0.1258

4. EXPERIMENTAL RESULTS AND ANALYSIS

According to the feature importance ranking obtained from the Pearson correlation coefficient method, grey relational degree, and the method proposed in this paper (as shown in Table 1), it can be observed that the contributions of different features to thermal property prediction differ. In the comparison of the three methods on Sample Set 1, peak position, peak area, and peak intensity consistently rank among the top, showing a significant impact on thermal property prediction. Specifically, the correlation of peak position is 0.82 in the Pearson correlation coefficient method, 0.6125 in the grey relational degree method, and 0.81 in the proposed method, all of which demonstrate high predictive ability. Peak area and peak intensity also show strong correlations, both with values of 0.81, ranking as high-priority features. Notably, peak width and frequency drift exhibit relatively consistent correlations across different methods, ranging from 0.75 to 0.81, indicating that these features have some influence in reflecting the thermal properties of the material. Features such as the signal-to-noise ratio after baseline correction, relative peak intensity, and time-resolved features generally show low correlations in all three methods, indicating that they contribute less to thermal property prediction and may have lower weight in practical applications.

Similarly, based on the ranking of feature information importance obtained through Pearson Correlation Coefficient, Grey Relational Degree, and the proposed method in Table 2, the importance of different features in thermal property prediction shows some differences. In Sample Set 2, the peak position consistently shows the highest importance across all three methods, with a Pearson correlation of 0.83, grey relational degree of 0.65, and a weight value of 6.1256 in the proposed method, highlighting its core role in thermal property prediction. The peak intensity and peak width also demonstrate high correlation in various methods, with Pearson correlation coefficients of 0.81 and 0.78, grey relational degree of 0.65, and weight values of 4.2368 and 2.1524 in the proposed method, respectively. These features' high correlation indicates their significant predictive ability for the thermal properties of the material. Conversely, features like frequency drift dependence, relative peak intensity, signal-to-noise ratio after baseline correction, and isotope effect show negative correlation in Pearson correlation coefficient, indicating their limited contribution to thermal property prediction. In the grey relational degree and proposed method, these features also exhibit significantly lower importance compared to peak position and peak intensity, confirming their secondary position in practical applications.

From the error sequence of the Raman scattering spectral data before and after smoothing in Sample Set 1 shown in Figure 5, it can be observed that smoothing significantly improves the stability and noise level of the signal. The data before smoothing exhibit large fluctuations and noise, with error fluctuations being wide at most sampling points. For instance, at sampling point 50, the error is 0.8, while at sampling point 200, the error is -4, indicating significant high-frequency noise in the original data. In contrast, the data after smoothing show reduced fluctuation amplitude, with error fluctuations stabilizing. For example, the error at sampling point 50 remains at 0.8, but the error at sampling point 200 is now stable at -2, showing a smoother trend.

Similarly, from the error sequence of the Raman scattering

spectral data before and after smoothing in Sample Set 2 shown in Figure 6, smoothing effectively reduces the fluctuation amplitude in the data and stabilizes the error change trend. The original data before smoothing exhibit sharp fluctuations, especially at sampling points 150 and 250, where the errors change drastically to -120 and -65, respectively. These extreme values indicate significant noise and instability in the signal. In comparison, after smoothing, the error changes become much smoother; for instance, the error at sampling point 150 decreases from -120 to -65, and at sampling point 250, the error decreases from -65 to -15. Overall, the error changes in the smoothed data are much more stable, with large fluctuations and outliers being effectively suppressed. From the experimental results, it can be seen that after smoothing, the error variations align more closely with the true signal trend and are less affected by noise, providing more reliable input data for thermal property prediction.

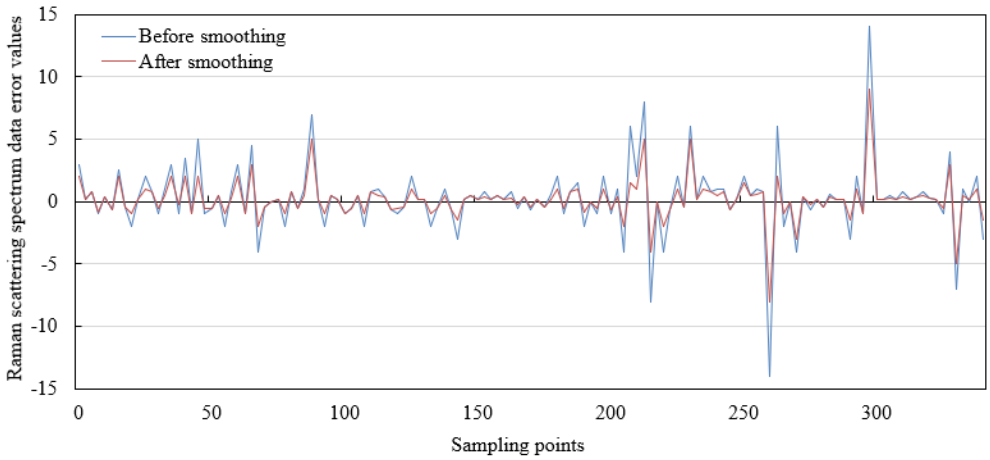


Figure 5. Error sequence of Raman scattering spectral data before and after smoothing (Sample set 1)

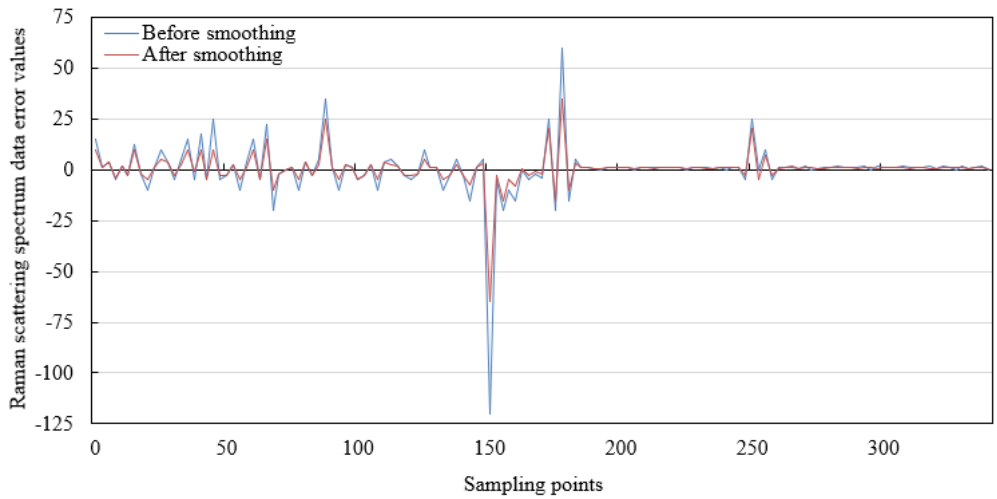
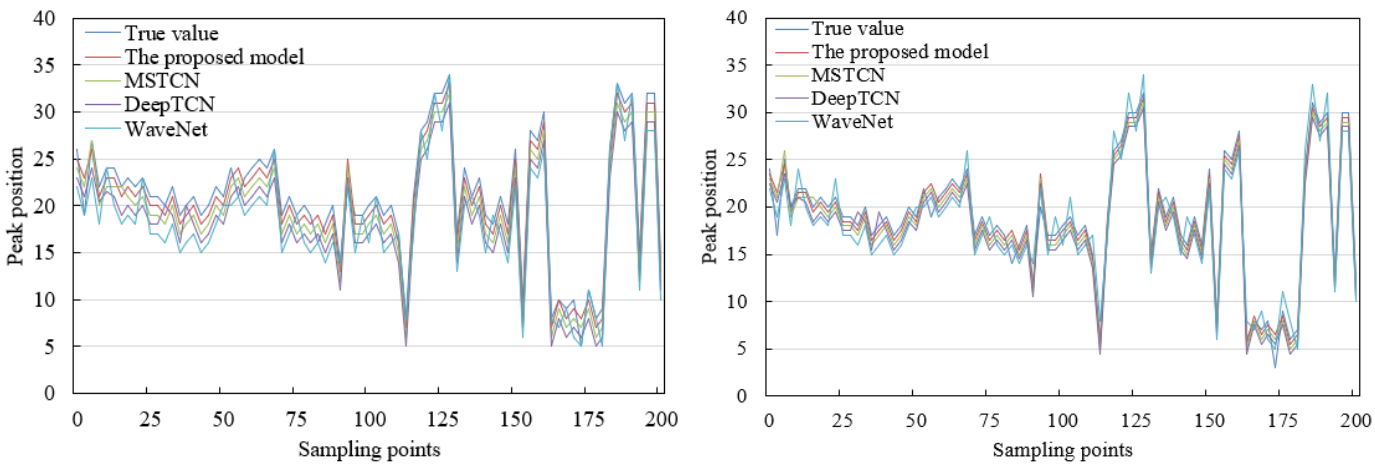


Figure 6. Error sequence of Raman scattering spectral data before and after smoothing (Sample set 2)



(a) Before error correction (b) After error correction

Figure 7. Comparison of thermal property prediction results of Raman scattering measurement model before and after error correction

From the data in Figure 7, it can be seen that the error between the predicted thermal property results and the real values significantly decreases after error correction, and the prediction accuracy is notably improved. Taking "this model" as an example, the prediction values before error correction show large deviations at multiple sampling points. For example, at sampling point 0, the prediction value is 25, with a difference of +1, while at sampling point 200, the prediction value is 23, with a difference of +2. These deviations indicate that the original prediction contains certain measurement errors and does not fully reflect the real situation. However, after error correction, the prediction results improved. For example, the prediction value at sampling point 0 decreased from 25 to 23.5, reducing the error by 0.5; the prediction value at sampling point 200 decreased from 23 to 20.5, reducing the error by 2.5, significantly improving the prediction accuracy. Overall, after error correction, the predicted results of all models are closer to the real values at each sampling point, and the error fluctuation tends to be smoother, showing higher prediction accuracy.

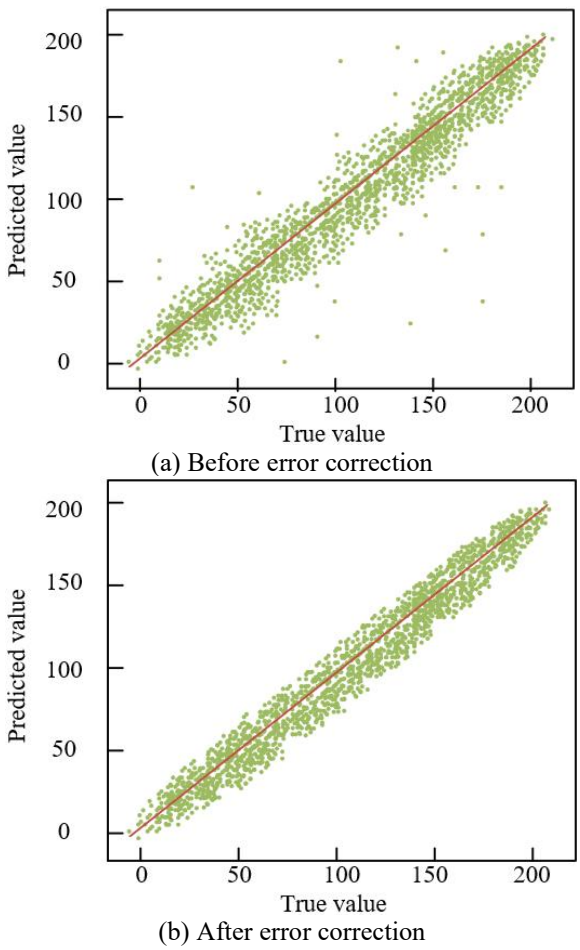


Figure 8. Prediction distribution map of thermal property prediction model based on RF-TCN-SA before and after Raman scattering measurement error correction

Figure 8 shows the prediction distribution map of the thermal property prediction model based on RF-TCN-SA before and after Raman scattering measurement error correction. From the figure, it is clear that before error correction, the predicted values and real values are more dispersed, with the green predicted points widely distributed on both sides of the diagonal, indicating a large deviation between the predicted and real values and a higher degree of dispersion in the predictions. After error correction, the predicted points are clearly more concentrated around the diagonal, and the degree of dispersion is significantly reduced. This visually indicates that after the error correction based on the neural network, the difference between the predicted values and real values of the RF-TCN-SA-based thermal property prediction model has significantly decreased.

The experimental results show that the error correction model constructed for Raman scattering measurement errors plays an effective role. By using a neural network to real-time correct experimental measurement deviations, the accuracy of the RF-TCN-SA-based thermal property prediction model is successfully improved. The change in the prediction distribution before and after error correction further confirms that combining the error correction model with the thermal property prediction model can effectively reduce errors occurring during the experiment, significantly improving the accuracy of thermal property prediction. This, in turn, enhances the reliability of Raman scattering technology in practical applications, providing more precise thermal property prediction results for research and applications in related fields.

5. CONCLUSION

This paper proposed a comprehensive method for Raman scattering measurement error correction and thermal property prediction based on neural networks, aiming to improve the accuracy and reliability of Raman scattering technology in thermal property prediction. First, the RF-TCN-SA-based thermal property prediction model was constructed, which effectively captures the complex features in Raman spectra by combining time convolution networks and SA mechanisms, thereby enhancing the prediction accuracy of thermal properties. Secondly, addressing common measurement errors in experiments, this paper proposed a neural network-based correction model that can real-time correct deviations in Raman scattering data, eliminating the interference of noise and outliers, thereby improving the accuracy of thermal property prediction. By combining these two models, this paper successfully achieved precise optimization of thermal property prediction while enhancing the reliability of Raman scattering technology, especially in dealing with experimental errors and data fluctuations, which has significant advantages.

However, despite the good experimental results achieved by the comprehensive method proposed in this paper, some limitations remain. First, the model training process requires a large amount of high-quality experimental data, and the quality and quantity of data directly affect the model's accuracy. Secondly, although neural network methods perform excellently in error correction and thermal property prediction, the training and adjustment process of the method requires considerable computational resources, and in specific cases, there may be issues with overfitting or excessive model complexity. Furthermore, the characteristics of Raman scattering data itself may be influenced by multiple factors, such as the uniformity of the sample and environmental conditions, which may limit the model's generalization ability under different experimental conditions. Therefore, future research can focus on further optimizing the structure of the neural network model, exploring more efficient training methods, and enhancing the robustness of the model under complex experimental conditions. At the same time, adding

more samples and multi-dimensional feature information, especially feedback data in practical applications, will further improve the practicality and accuracy of the method in various thermal property prediction tasks.

ACKNOWLEDGMENT

This paper was supported by Shaanxi Province Science and Technology Resource Sharing Platform (Grant No.: 2023-CX-PT-16).

REFERENCES

[1] Chen, Y., Zhang, G.Q., Song, A., You, P.B. (2024). Thermodynamic properties of composite material bridges under thermal cycling. International Journal of Heat and Technology, 42(1): 171-182. https://doi.org/10.18280/ijht.420118

[2] Jha, B.K., Musa, M.K., Ajibade, A.O. (2024). Natural convection flow in a thermally stratified fluid through an asymmetrically heated and cooled vertical channel with anisotropic porous material. Power Engineering and Engineering Thermophysics, 3(2): 116-133. https://doi.org/10.56578/peet030204

[3] Farrera-Vázquez, N., Aviles-Trujilo, L., Moreira-Acosta, J., García-Ramos, O., et al. (2022). Development of an insulating material based on Trametes elegans mycelium and the study of its hygrothermal properties. Green Materials, 11(1): 28-36. https://doi.org/10.1680/jgrma.21.00046

[4] Lazim, A.A., Daneh-Dezfuli, A., Habeeb, L.J. (2023). Enhancing heat transfer in heating pipes with Fe3O4 nanofluid under magnetic fields: A numerical study. Journal of Sustainability for Energy, 2(4): 207-216. https://doi.org/10.56578/jse020404

[5] Zhang, X.L., Liu, H.S., Ma, S.S., Yang, P.P. (2023). Thermal air aging and lifespan prediction of PVC-P geomembranes: An arrhenius equation-based approach. Power Engineering and Engineering Thermophysics, 2(4): 199-211. https://doi.org/10.56578/peet020402

[6] Gao, P., Zhang, Y., Yu, Z., Fang, J., Zhang, Q. (2015). Correlation study of shallow layer rock and soil thermal physical tests in laboratory and field. Geothermics, 53: 508-516. https://doi.org/10.1016/j.geothermics.2014.09.005

[7] Saq'an, S., Zihlif, A.M., Ragosta, G. (2008). Thermal, elastic, and electrical properties of talc-polypropylene composite. Journal of Thermoplastic Composite Materials, 21(5): 457-467. https://doi.org/10.1177/0892705708090826

[8] Andriani, G.F., Germinario, L. (2014). Thermal decay of carbonate dimension stones: Fabric, physical and mechanical changes. Environmental Earth Sciences, 72: 2523-2539. https://doi.org/10.1007/s12665-014-3160-6

[9] Mokrejs, P., Langmaier, F., Janacova, D., Mladek, M., Kolomaznik, K., Vasek, V. (2009). Thermal study and solubility tests of films based on amaranth flour starch–protein hydrolysate. Journal of Thermal Analysis and Calorimetry, 98(1): 299-307. https://doi.org/10.1007/s10973-009-0106-4

[10] Tian, S., Zhang, Z., Meng, F., Wang, Z., Luo, L. (2023). Recent advances in enhancement of Raman scattering intensity for biological applications. Chemical & Biomedical Imaging, 1(7): 575-589. https://doi.org/10.1021/cbmi.3c00017

[11] Liu, R., Liu, J.F., Zhou, X.X., Jiang, G.B. (2011). Applications of Raman-based techniques to on-site and in-vivo analysis. TrAC Trends in Analytical Chemistry, 30(9): 1462-1476. https://doi.org/10.1016/j.trac.2011.06.011

[12] Kant, K., Abalde-Cela, S. (2018). Surface-enhanced Raman scattering spectroscopy and microfluidics: Towards ultrasensitive label-free sensing. Biosensors, 8(3): 62. https://doi.org/10.3390/bios8030062

[13] Nie, Q., Liu, Z., Cheng, M., Pei, S., Yang, D., Guo, D., Yang, M. (2024). Review on hollow-core fiber based multi-gas sensing using Raman spectroscopy. Photonic Sensors, 14(4): 240412. https://doi.org/10.1007/s13320-024-0730-4

[14] Smith, Z.J., Huser, T.R., Wachsmann-Hogiu, S. (2012). Raman scattering in pathology. Analytical Cellular Pathology, 35(3): 145-163. https://doi.org/10.3233/ACP-2011-0048

[15] Hansen, M., Truong, J., Xie, T., Hahm, J.I. (2017). Spatially distinct Raman scattering characteristics of individual ZnO nanorods under controlled polarization: Intense end scattering from forbidden modes. Nanoscale, 9(24): 8470-8480. https://doi.org/10.1039/C7NR02672B

[16] Felinskyi, G., Dyriv, M. (2015). Noise suppression phenomenon in fiber Raman amplifier. Measurement Science Review, 15(3): 107-110. https://doi.org/10.1515/msr-2015-0016

[17] Wandinger, U. (1998). Multiple-scattering influence on extinction-and backscatter-coefficient measurements with Raman and high-spectral-resolution lidars. Applied Optics, 37(3): 417-427. https://doi.org/10.1364/ao.37.000417

[18] Kim, J.H., Lee, J., Kim, Y.G., Kim, S., Wang, J., Lee, J.H., Hwang, Y.S. (2024). Systematic error analysis of Thomson scattering system on VEST with rotational Raman calibration. IEEE Transactions on Plasma Science, 52(9): 4005-4010. https://doi.org/10.1109/TPS.2024.3385545

[19] Qiu, D., Minami, T., Kado, S., Inagaki, S., et al. (2023). Increased signal separation upgrade permits preliminary electron anisotropy measurements with Heliotron J multi-pass Thomson diagnostic. Review of Scientific Instruments, 94: 023510. https://doi.org/10.1063/5.0101815

[20] Richardson, D. R., Kearney, S. P., Guildenbecher, D. R. (2020). Three-beam rotational coherent anti-Stokes Raman spectroscopy thermometry in scattering environments. Applied Optics, 59(27): 8293-8301. https://doi.org/10.1364/ao.392110

[21] Hogg, R.E., Zlatkova, M.B., Chakravarthy, U., Anderson, R.S. (2007). Investigation of the effect of simulated lens yellowing, transparency loss and refractive error on in vivo resonance Raman spectroscopy. Ophthalmic and Physiological Optics, 27(3): 225-231. https://doi.org/10.1111/j.1475-1313.2007.00478.x