applied sciences

MDPI

Article

# Nondestructive Evaluation of Thermal Barrier Coatings' Porosity Based on Terahertz Multi-Feature Fusion and a Machine Learning Approach

**Rui Li [1,2,3,4,5], Dongdong Ye [1,2,3,4,5,*], Qiukun Zhang [2], Jianfei Xu [6] and Jiabao Pan [1,*]**

1 School of Mechanical Engineering, Anhui Polytechnic University, Wuhu 241000, China; lirui1011@stu.ahpu.edu.cn
2 Fujian Provincial Key Laboratory of Terahertz Functional Devices and Intelligent Sensing, School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, China; qk_zhang@fzu.edu.cn
3 School of Artificial Intelligence, Anhui Polytechnic University, Wuhu 241000, China
4 Huzhou Key Laboratory of Terahertz Integrated Circuits and Systems, Yangtze Delta Region Institute (Huzhou), University of Electronic Science and Technology of China, Huzhou 313001, China
5 Anhui Key Laboratory of Detection Technology and Energy Saving Devices, Anhui Polytechnic University, Wuhu 241000, China
6 Department of Automotive Engineering and Intelligent Manufacturing, Wanjiang College of Anhui Normal University, Wuhu 241008, China; wuhuuniversityxjf@163.com
* Correspondence: ddyecust@ahpu.edu.cn (D.Y.); panjiabao@ahpu.edu.cn (J.P.)

**Abstract:** Thermal barrier coatings (TBCs) play a crucial role in safeguarding aero-engine blades from high-temperature environments and enhancing their performance and durability. Accurate evaluation of TBCs' porosity is of paramount importance for aerospace material research. However, existing evaluation methods often involve destructive testing or lack precision. In this study, we proposed a novel nondestructive evaluation method for TBCs' porosity, utilizing terahertz time-domain spectroscopy (THz-TDS) and a machine learning approach. The primary objective was to achieve reliable and precise porosity evaluation without causing damage to the coatings. Multiple feature parameters were extracted from THz-TDS data to characterize porosity variations. Additionally, correlation analysis and $p$-value testing were employed to assess the significance and correlations among the feature parameters. Subsequently, the dung-beetle-optimizer-algorithm-optimized random forest (DBO-RF) regression model was applied to accurately predict the porosity. Model performance was evaluated using K-fold cross-validation. Experimental results demonstrated the effectiveness of our proposed method, with the DBO-RF model achieving high precision and robustness in porosity prediction. The model evaluation revealed a root-mean-square error of 1.802, mean absolute error of 1.549, mean absolute percentage error of 8.362, and average regression coefficient of 0.912. This study introduces a novel technique that presents a dependable nondestructive testing solution for the evaluation and prediction of TBCs' porosity, effectively monitoring the service life of TBCs and determining their effectiveness. With its practical applicability in the aerospace industry, this method plays a vital role in the assessment and analysis of TBCs' performance, driving progress in aerospace material research.

**Keywords:** thermal barrier coatings; porosity characterization; terahertz time-domain spectroscopy; nondestructive evaluation; multi-feature fusion; machine-learning-based prediction; aerospace materials

check for updates

**Citation:** Li, R.; Ye, D.; Zhang, Q.; Xu, J.; Pan, J. Nondestructive Evaluation of Thermal Barrier Coatings' Porosity Based on Terahertz Multi-Feature Fusion and a Machine Learning Approach. *Appl. Sci.* **2023**, *13*, 8988. https://doi.org/10.3390/app13158988

Academic Editors: Zhigang Zhang, Hui Wang, Qiuwei Yang, Xi Peng and Fengjiang Qin

Received: 13 July 2023
Revised: 28 July 2023
Accepted: 3 August 2023
Published: 5 August 2023



## 1. Introduction

Thermal barrier coatings (TBCs), due to their excellent high temperature resistance, low thermal conductivity, corrosion protection, and anti-wear properties, are considered crucial protective layers. They have been extensively applied in the aerospace industry for hot-section components such as aero-engine blades [1,2]. Blades are exposed to extreme service environments, such as the impact of high-temperature combustion gases, oxidizer erosion, and thermal stresses. These factors lead to increased blade surface temperature and material thermal expansion, resulting in issues such as thermal fatigue, oxidation, and mechanical failure [3]. The application of TBCs has effectively mitigated these detrimental effects on the blades, enhancing their heat resistance, corrosion resistance, and mechanical strength [4]. They have played a crucial role in reducing surface temperatures, improving component durability, and extending service life. With the rapid development of the aerospace industry and increasing demands for engine performance, the quality and performance of TBCs have become a focal point of attention [5]. TBCs have typically been applied with a thickness ranging from 100 to 600 μm and consisting of three layers: a topcoat (TC) comprising an yttria-stabilized zirconia (YSZ) ceramic material, a band coat (BC) made of MCrAlY (where M represents Ni, Co, or other materials) as the metallic bond layer, and a high-temperature nickel-based alloy substrate. The topcoat possesses a low thermal conductivity and high thermal resistance, effectively reducing heat conduction and mitigating the surface temperature of the blades. The bond coat acts as an intermediary layer, providing adhesion and overall stability between the TBCs and the blade substrate [6]. However, as thermal barrier coatings are exposed to harsh operating conditions for prolonged periods, they might experience various forms of failure due to factors such as thermal fatigue, thermal stress mismatch, coating oxidation and corrosion, and particle impact [7–9]. As the TBCs are subjected to service conditions, the formation of thermal growth oxides (TGOs) occurs between the top ceramic layer and the bond coat [10]. Nevertheless, the presence of pore structures in TBCs is closely associated with various factors. The formation of these pores could be attributed to processes such as high-temperature spraying and sintering, which could result in particle accumulation and melting. Under thermal cycling conditions, TBCs experience complex effects such as temperature gradients, thermal stresses, and mechanical loads, leading to the formation of microcracks and delamination within the coating and subsequently generating pores. The porosity of TBCs significantly influences their performance. TBCs have typical porosity values which range from 3 to 20 vol%. Increasing the porosity of the coating is advantageous as it reduces its thermal conductivity, thereby enhancing its thermal insulation capabilities. The existence of air within the pores leads to a reduced effective thermal conductivity, acting as a thermal barrier and impeding the heat transfer through the coating. Nevertheless, excessive porosity could negatively impact the mechanical integrity and durability of the coating, rendering it susceptible to cracking and delamination [11,12]. Therefore, the accurate evaluation of porosity in TBCs is of crucial importance to ensure coating quality, performance, and reliability.

Accurate determination of the porosity assessment in TBCs has emerged as a key issue in research [13–17]. Various methods have been proposed to assess the porosity of these coatings, leading to significant research progress. Surface or cross-sectional observations and analyses of the coatings can be performed using a scanning electron microscope (SEM) [18] or transmission electron microscope (TEM) [19]. Quantitative measurements of pore size, shape, and distribution can be obtained through image processing and analysis techniques, enabling the calculation of porosity. However, these microscopic observation methods are considered destructive, requiring sample embedding and polishing procedures. Therefore, nondestructive testing (NDT) methods present a superior choice. NDT is a vital technology that could be applied to the evaluation of various materials and components. Utilizing various testing techniques detects surface and internal defects, damages, or structural changes without causing sample destruction. Due to its ability to provide timely and accurate information, NDT has found extensive applications in industries such as aerospace, automotive, marine, nuclear, and construction, facilitating quality control and fault diagnosis, as well as ensuring preventive maintenance, structural safety, and reliability [20,21]. Furthermore, NDT methods hold significant potential for the evaluation of porosity in TBCs. For instance, NDT of the coatings can be achieved using ultrasonic waves [22], infrared thermography [23], and X-rays [24]. These methods offer advantages such as rapid assessment, non-contact measurement, and comprehensive evaluation of the coatings. However, they still encounter complex challenges in signal processing and data analysis during practical application. Ultrasonic testing has limited resolution and cannot accurately detect micro-sized pores. Infrared thermography requires thermal cycling of the samples, which can affect the detection results. As the sample undergoes heating and subsequent cooling cycles, it experiences expansion and contraction, which can affect its thermal conductivity, heat capacity, and thermal diffusivity. X-ray technology involves radiation and poses certain safety risks to operators. Despite these challenges, the methods utilized for NDT of thermal barrier coatings are still quite accurate. In recent years, many scholars have employed machine learning techniques to further enhance the accuracy of NDT for thermal barrier coatings. Utilizing machine learning in NDT has further improved accuracy, enabling more efficient and precise evaluation of critical components such as TBCs in aerospace applications [25–27]. Therefore, combining a superior NDT method with machine learning techniques enables higher accuracy in detecting the TBCs' porosity.

The terahertz time-domain spectroscopy (THz-TDS) technique offers unique advantages in the evaluation of TBCs' porosity [28,29]. When the terahertz pulse interacts with the sample, it undergoes changes in amplitude and phase due to the sample's material properties, such as its refractive index and absorption characteristics. The altered terahertz pulse is then combined with the reference pulse at the detector, and the resulting electric field is measured as a function of time. Operating within the frequency range of 0.1 THz to 10 THz, corresponding to wavelengths ranging from 30 μm to 3 mm, terahertz waves occupy an intermediate position between infrared light and microwaves. This nondestructive technique provides high resolution, sensitivity, and real-time capabilities and the ability to conduct comprehensive analyses of multiple parameters. Moreover, it can penetrate non-metallic ceramic layers without causing material damage or altering their properties, making it suitable for continuous and repeatable testing of TBCs. For instance, Davit et al. [30] proposed a nondestructive evaluation method using THz-TDS to assess the dielectric constant of $Al_2O_3$ ceramic samples, allowing for the evaluation of critical parameters such as porosity, grain size, and impurity content. This demonstrates the potential application of terahertz radiation in the field of ceramics with high dielectric constants. Some researchers have already started exploring the application of machine learning as an advanced signal-processing method in the field of nondestructive testing [31,32]. Ye et al. [33,34] utilized THz-TDS to detect TBCs' porosity, employing effective medium theory and a time-domain broadening ratio for porosity evaluation. Additionally, methods based on machine learning and image processing have been applied to assess the quality of thermal barrier coatings. Li et al. [35,36] conducted analysis and processing of terahertz time-domain data using data-driven models and machine learning algorithms, facilitating feature extraction and classification for the evaluation of TBCs. This represented a major leap forward in combining machine learning with signal processing. The potential of THz-TDS combined with machine learning algorithms was emphasized to improve the accuracy of detecting TBCs, thus promoting the development of the field of NDT.

In summary, the current methodologies employed for assessing TBCs' porosity using THz-TDS present opportunities for improvement. The porosity of such coatings is influenced by multiple factors, demanding the comprehensive consideration of various key feature parameters. To achieve a thorough evaluation and prediction of porosity, this research conducted a detailed analysis of terahertz time-domain data, extracting vital information regarding the porosity structure. Section 2 elaborates on the extraction procedures for multiple feature parameters from terahertz time-domain data, encompassing correlation and significance analyses. Additionally, dung-beetle-optimization-algorithm-optimized random forest (DBO-RF) prediction models and cross-validation techniques are presented. In Section 3, an exploration of the visual representation of the correlation between multiple feature parameters and porosity is undertaken, integrating these fused feature parameters as inputs for the machine learning model to predict porosity. The model's superior performance is evidenced through evaluation indicators. The prediction accuracy is over 90%.

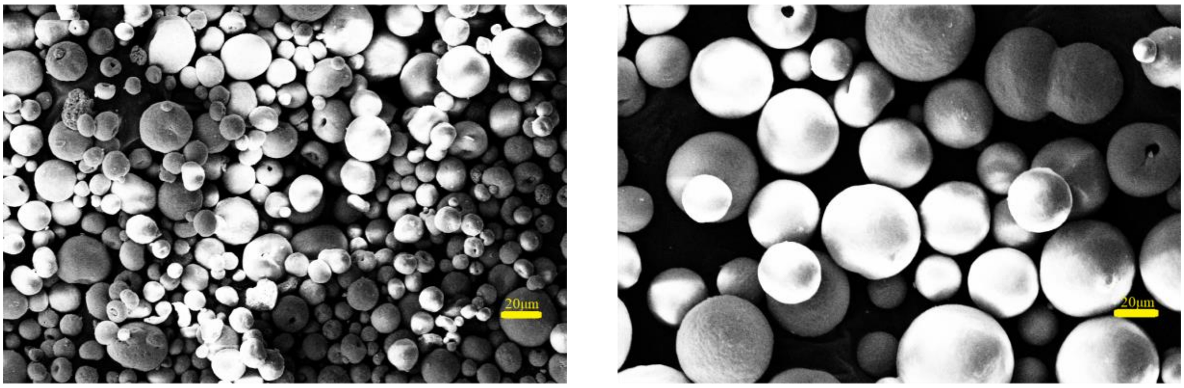
The primary focus of this study is to develop an effective approach for the nondestructive assessment and prediction of TBCs' porosity. By combining multiple feature parameters obtained from terahertz data and employing advanced machine learning algorithms, this research aims to establish a practical method for the quality control and evaluation of TBCs, particularly in the application of coating materials on aero-engine blades. The proposed approach holds significant potential for enhancing the understanding of TBCs' porosity and ensuring the performance and reliability of these coatings in critical engineering applications. Through this interdisciplinary investigation encompassing materials science and information sciences, we aim to contribute to the advancement of nondestructive evaluation methods for TBCs, enhancing their performance and reliability in critical aerospace applications.

## 2. Materials and Methods

### *2.1. Preparation of Thermal Barrier Coatings*

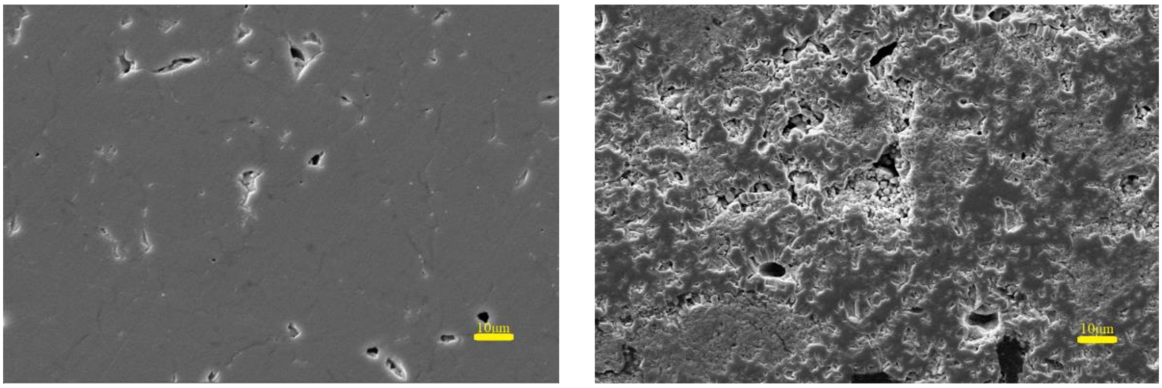
To obtain thermal barrier coating samples with varying porosity, atmospheric plasma spraying (APS) equipment was used to deposit TBCs onto carbon steel disc substrates [37,38]. The diameter of the substrate was 25.4 mm, and the thickness was 3 mm. Yttria-stabilized zirconia (8% $Y_2O_3$-$ZrO_2$, 8YSZ) was selected as the material for TBCs, known for its excellent thermal stability and suitability for high-temperature applications. Additionally, 8YSZ exhibited a reduced susceptibility to phase transformation, particularly the tetragonal-to-monoclinic phase transformation, which could lead to microcracking and degradation of the coating under thermal cycling conditions. The incorporation of yttria ($Y_2O_3$) as a stabilizing agent effectively inhibited this phase transformation, thereby enhancing the overall mechanical durability and longevity of the TBCs. By utilizing 8YSZ as the ceramic powder, we aimed to ensure the long-term reliability and performance of the TBCs, especially in high-temperature aerospace applications. Moreover, the excellent thermal and mechanical properties of 8YSZ were conducive to achieving the desired porosity levels and microstructural characteristics essential for our research on nondestructive evaluation and prediction of TBCs' porosity. The ceramic powder had two different particle sizes, as shown in Figure 1, namely 25–55 μm and 45–106 μm. These specific particle sizes were selected based on their relevance to the porosity evaluation, as they could influence the formation of pores during the spraying process. TBC porosity samples were controlled by varying the spraying parameters. Specifically, the spraying distance and power were adjusted. The spraying distance refers to the distance between the nozzle of the plasma spraying system and the substrate, and different spraying distances of 50 mm, 70 mm, 90 mm, 110 mm, and 130 mm were chosen. The spraying power levels were at the rates of 33 kW, 38 kW, and 40 kW. These parameter variations effectively regulated the TBCs' porosity. The justification for these choices lay in the need to investigate the impact of different spraying parameters on the resulting porosity levels of the TBCs. By systematically varying the spraying distance and power, we could better understand the relationship between the porosity and the process parameters, providing valuable insights for quality control and evaluation.

For quantitative characterization of TBC samples with varying porosities, scanning electron microscopy (SEM) was utilized to capture cross-sectional images of the samples. The SEM images facilitated the examination of the microstructure and pore distribution of TBCs. Figure 2 shows an SEM image of the TBC cross-section sprayed by ceramic powder with different particle sizes at the same spraying distance. It is clear that with the increase in powder particle size, the number of pores in the thermal barrier coating increases significantly, and the pore size also increases. This proves that it was valid to obtain samples containing different porosities by adjusting spraying parameters. To accurately assess the porosity, ImageJ module was employed for image analysis and processing, providing a set of image processing tools that enable the measurement of pore area fraction and calculation of the corresponding porosity. By employing the experimental preparation and quantitative characterization methods described above, TBC samples with different porosities were obtained and then utilized for subsequent THz-TDS detection, allowing for precise evaluation and predictive analysis of the porosity of TBCs. In this study, a total of 20 kinds of samples containing different porosities, ranging from 6.94 to 22.09%, were prepared.




(a) (b)

**Figure 1.** Particle size of powder: (**a**) 25–55 μm; (**b**) 45–106 μm.




(a) (b)

**Figure 2.** SEM cross-section prepared at 70 mm spraying distance of TBCs: (**a**) 25–55 μm; (**b**) 45–106 μm.

*2.2. Terahertz Time-Domain Spectroscopy*

Terahertz time-domain spectroscopy (THz-TDS) is a novel NDT technique [39,40] known for its strong penetration capability, transparency to various materials, and non-ionizing nature when applied to non-metallic materials. In this study, a reflection-based THz-TDS technique was utilized to examine the thermal barrier coating samples, aiming to obtain THz time-domain data. The propagation of reflected terahertz waves within the thermal barrier coating is shown in Figure 3.

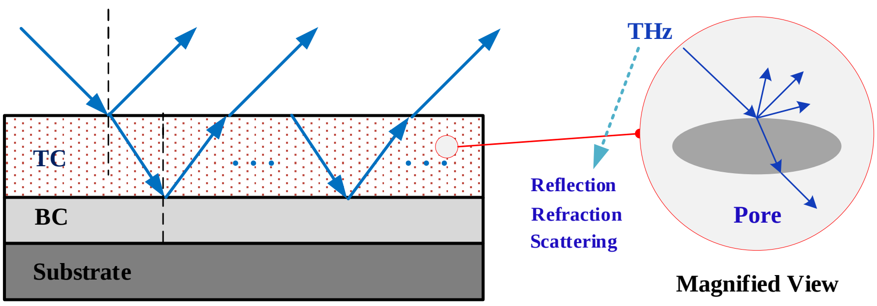



**Figure 3.** Schematic diagram of terahertz propagation in thermal barrier coatings.

The fundamental principles of THz-TDS involve observing the propagation of short-pulsed electromagnetic waves within the tested object, as shown in Figure 4. To generate broad-spectrum terahertz pulses, femtosecond lasers and optical pulse sources are commonly used as light sources, employing nonlinear optical mechanisms. These pulses propagate through the sample under the guidance of terahertz optical components, interacting with the object being tested. The test sample exhibited distinct absorption, transmission, and reflection characteristics toward terahertz waves, which were closely linked to their physical, chemical, and structural properties. Consequently, analyzing the interaction between terahertz waves and the test sample allowed for the retrieval of valuable information about the sample. Subsequently, the terahertz pulses passed through the sample and their changes in amplitude and phase were measured using a detector.

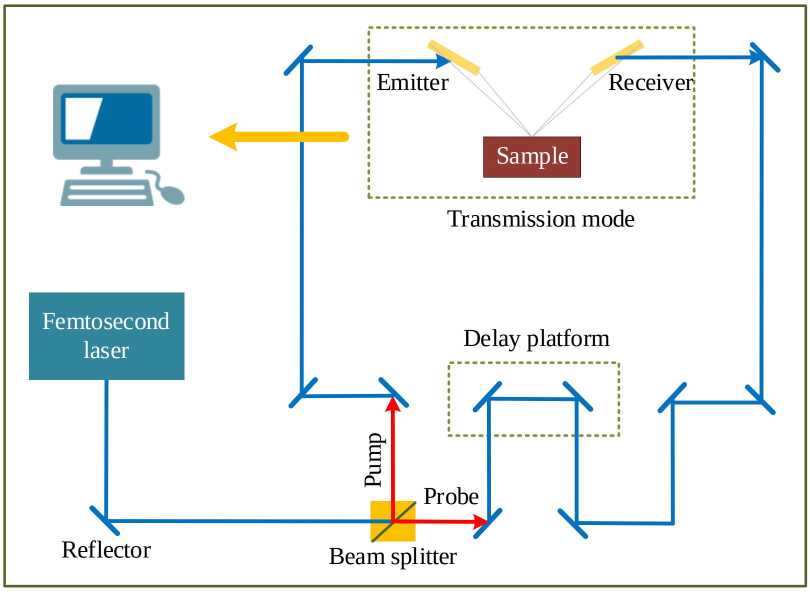



**Figure 4.** Schematic diagram of reflection terahertz time-domain spectroscopy technique.

The propagation and interaction of terahertz waves within the TBCs were influenced by factors such as the coating's porosity, distribution of material composition, and microstructure. Hence, THz-TDS offered a highly sensitive approach to detecting the TBCs' porosity and its reflective properties at different frequencies [41]. In this study, THz-TDS was employed to investigate the TBCs' porosity on aero-engine blades, with a focus on obtaining time-domain data. Specifically, the obtained terahertz time-domain data were subjected to fast Fourier transform (FFT) to obtain frequency-domain spectra, phase spectra, and reflectance spectra. Subsequently, data analysis and extraction of characteristic parameters were conducted on these spectra.

### *2.3. Terahertz Time-Domain Data Processing and Spectral Analysis*

THz-TDS allowed for the characterization of the material's electromagnetic wave propagation in the terahertz frequency range. However, to further investigate the influence of porosity on terahertz wave propagation, additional analysis of the time-domain data was required. In this study, as shown in Figure 5a–d, FFT was employed to convert the terahertz time-domain spectrum into frequency-domain spectrum, phase spectrum, and reflectance spectrum [42,43]. Different porosity levels correspond to varying energy losses and absorption characteristics in the frequency spectrum, resulting in distinct energy distribution patterns at different frequencies. Changes in porosity modify the propagation velocity and phase delay of terahertz waves within the TBCs, manifested as variations in the slope of the phase spectra. Pores present in the coating increase scattering and multiple reflections at interfaces, resulting in modifications to the intensity and shape of the reflectance spectra. Frequency-domain spectra, phase spectra, and reflectance spectra all provide insights into the variation of porosity in the thermal barrier coating samples. Analyzing the characteristics of these spectra allows for obtaining quantitative or qualitative information about the porosity of the thermal barrier coating, which assists in further evaluating the quality and performance of the coating.

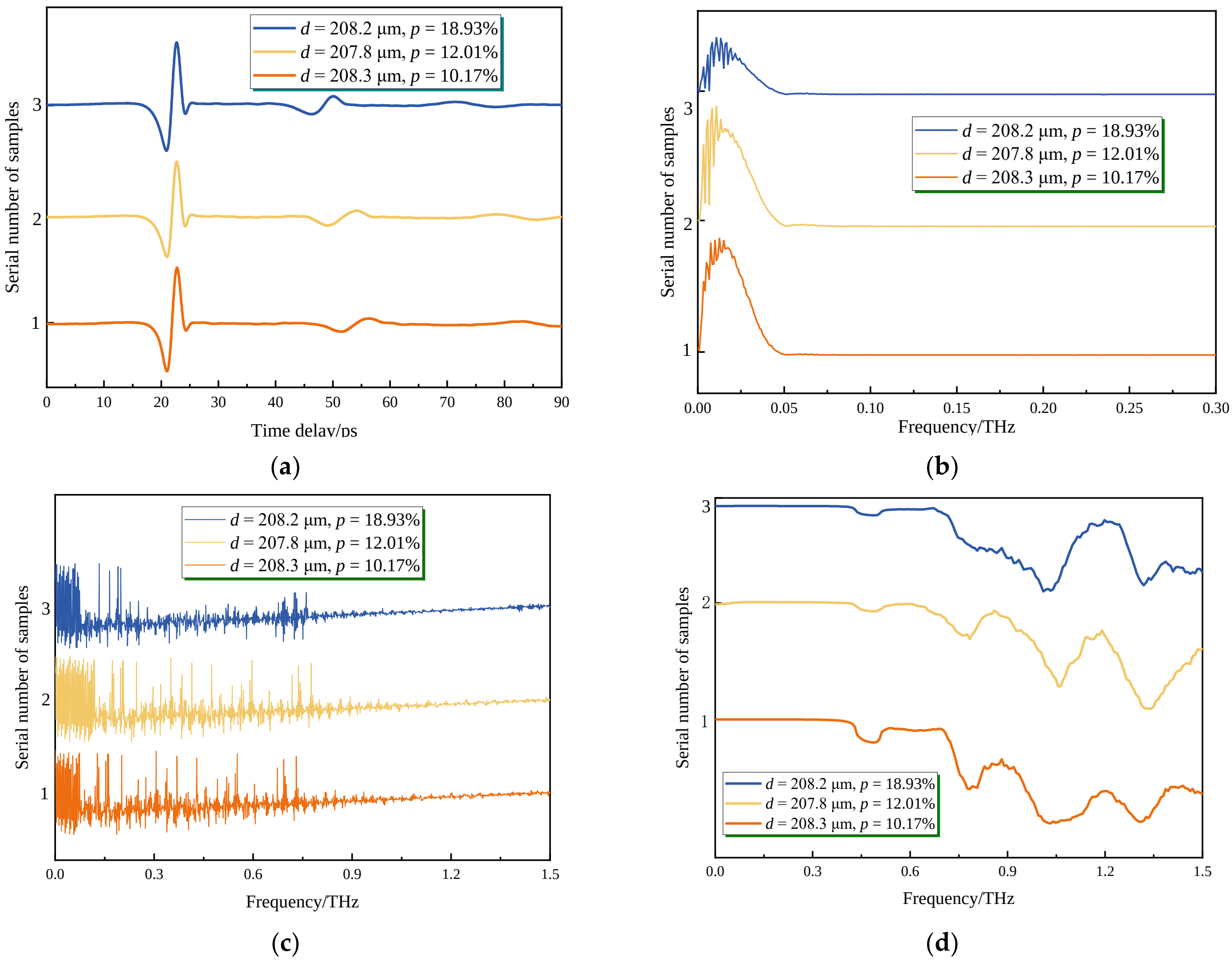



**Figure 5.** Results of terahertz time-domain data analysis: (**a**) time-domain spectrum; (**b**) frequency-domain spectrum; (**c**) phase spectrum; (**d**) reflectance spectrum.

The detailed steps are as follows.

The frequency-domain signal $T(w)$ is obtained by fast Fourier transform of terahertz time-domain signal $E(t)$.

$$T(w) = FFT(E(t)) \tag{1}$$

The amplitude spectrum and phase spectrum of the frequency-domain signal are calculated as follows:

$$A(w) = |T(w)| \tag{2}$$

$$\phi(w) = tan^{-1}\left(\frac{\Im\{T(w)\}}{\Re\{T(w)\}}\right) \tag{3}$$

where $A(w)$ represents the amplitude spectrum of the signal in the frequency domain, $\Im$ and $\Re$ represent the real and imaginary parts, respectively, and $\phi(w)$ represents the phase spectrum of the signal in the frequency domain.

The reflectance spectrum $R(w)$ is calculated, and $r_s(\theta)$ is the reflection coefficient:

$$r_s(\theta) = \frac{n_1 cos(\theta) - n_2\sqrt{1-(\frac{n_1}{n_2}sin(\theta))^2}}{n_1 cos(\theta) + n_2\sqrt{1-(\frac{n_1}{n_2}sin(\theta))^2}} \tag{4}$$

where $n_1$ and $n_2$ are the refractive indices of air and sample, respectively.

$$E_i(w) = \frac{1}{2}E_0(1+cos(\theta))e^{-ik(d+z_0 cos(\theta))} \tag{5}$$

$$E_r(w) = -\frac{1}{2}E_0(1-cos(\theta)r_s(\theta))e^{-ik(d-z_0 cos(\theta))} \tag{6}$$

Here, $E_0$ is the incident field intensity, $\theta$ is the incident angle, $d$ is the sample thickness, and $z_0$ is the free impedance.

$$\Gamma(w) = \frac{E_r(w)}{E_i(w)} \tag{7}$$

$$R(w) = |\Gamma(w)|^2 \tag{8}$$

Here, $E_i$ and $E_r$ are the electric fields of terahertz entering and leaving the sample surface, respectively, and $\Gamma(w)$ is the reflectance coefficient in the complex frequency domain.

## *2.4. Feature Extraction and Multi-Feature Fusion Analysis*

### 2.4.1. Extraction of Time-Domain Spectrum Features

THz-TDS, as an essential data source for characterizing the TBCs' porosity, encompasses complex information. Processing and feature extraction were required in this study, and the fast independent component analysis (Fast ICA) algorithm was employed to process the terahertz time-domain data. The objective was to extract kurtosis as a parameter to represent the variations in TBCs' porosity [44]. Fast ICA was a statistical signal-processing method used to extract independent components from mixed signals. THz-TDS was considered as mixed signals, with different components corresponding to distinct physical processes and features. By applying the Fast ICA algorithm, these components can be separated, and the kurtosis of each independent component vector can be computed.

Kurtosis, a statistical measure that quantifies the skewness of a probability distribution, was a significant statistical feature used to describe the shape of signal data distribution. A higher kurtosis value indicates a greater degree of skewness in the distribution. In this study, kurtosis was utilized as a feature parameter to represent porosity. TBCs' porosity influenced the propagation of terahertz waves within the material. As porosity increased, there was an enhanced occurrence of emission and scattering of the waves, resulting in changes in peak positions and kurtosis in the THz-TDS. Specifically, by calculating the kurtosis of each independent component vector, the degree and trend of porosity variations in TBC samples can be indirectly reflected. This approach enabled the quantitative analysis and comparison of variations among TBC samples with different levels of porosity.

The specific processing process is as follows.

It is assumed that the original terahertz signal data are $x$, $M$ represents the number of signal sampling points, and $N$ represents the number of signals.

$$x = \begin{bmatrix} x_{1,1} & \cdots & x_{M,1} \\ \vdots & \ddots & \vdots \\ x_{1,N} & \cdots & x_{M,N} \end{bmatrix} \tag{9}$$

Then, processing of time-domain data is centralized:

$$x_c(t) = x(t) - \frac{1}{N}\sum_{i=1}^{N} x_i \tag{10}$$

where $x_c(t)$ is the time-domain data after centralization, $t$ is the time, and $x_i$ is the column $i$ data.

The centralized time-domain data are processed by Fast ICA, and the independent component vector $s(t)$ is obtained:

$$s(t) = Ax_c(t) \tag{11}$$

where $A$ is the mixing matrix, representing the linear mixing process of the signal.

The kurtosis $k$ of each independent component vector is calculated:

$$k = \frac{1}{T}\int_{-\infty}^{\infty}\left(\frac{x_c(t)}{\sigma}\right)^4 dt - 3 \tag{12}$$

where $T$ is the total time length of the signal, and $\sigma$ is the standard deviation of the signal.

Kurtosis is used as the characteristic parameter of porosity. The feature parameter vector is obtained:

$$T\_parameter = [k_1, k_2, k_3 \ldots, k_n] \tag{13}$$

where $n$ is the number of independent components, and $k_n$ is the kurtosis of the $n$-th independent component vector. $T\_parameter$ is the feature parameter vector of time-domain data.

#### 2.4.2. Extraction of Frequency-Domain Spectrum Features

The porosity of TBCs influenced the terahertz frequency-domain signals, requiring the extraction of spectral features for porosity characterization. Specifically, the center frequency, bandwidth, and average power spectral density (APSD) were extracted from the spectra for this purpose [45]. The center frequency corresponded to the peak frequency in the spectrum data, indicating the primary energy distribution of the terahertz signal in the frequency domain. The bandwidth represented the frequency range where the spectrum decreased to a certain extent on both sides of the peak. By computing the bandwidth, the impact of porosity on the extent of spectral broadening in the terahertz signal within the TBC sample can be quantified. The average power spectral density (PSD) is obtained by averaging the squared amplitude of the signal at each frequency point, reflecting the changes in terahertz signal intensity. By extracting the center frequency, bandwidth, and APSD as feature parameters, the effects of porosity variations in the thermal barrier coating sample on the frequency characteristics of terahertz signals can be explored. The trends exhibited by these parameters will offer essential insights for subsequent prediction and analysis of porosity in TBCs.

The specific processing process is as follows.

Smoothing of frequency spectrum data is performed:

$$smoothed\_data_{i,j} = \frac{1}{N}\sum_{k=j-w}^{j+w} f\_data_{i,k} \tag{14}$$

where $i$ is the $i$-th signal, $j$ is the $j$-th frequency point, $w$ is the size of the smoothing window, $N$ is the number of data points in the window, $f\_data$ is the frequency spectrum data, and $smoothed\_data$ is the smoothed frequency spectrum data.

A Gaussian curve fit is performed for the $i$-th signal. Gaussian distribution is a continuous distribution function that accurately describes the main characteristics of the spectral data. Gaussian fitting is performed on the smoothed spectral data to calculate their center frequency and bandwidth:

$$y_i = A_i exp^{-\frac{(x-\mu_i)^2}{2\sigma_i{}^2}} + b_i \tag{15}$$

where $x$ is the frequency, $A_i$ is the peak intensity, $\mu_i$ is the peak frequency, $\sigma_i$ is the frequency bandwidth, and $b_i$ is the constant term of the fitted curve.

Amplitude spectrum $A(f)$ is obtained by taking the absolute value of the original frequency spectrum data $f\_data$ and squaring it.

$$A(f) = |f\_data|^2 \tag{16}$$

The average power spectrum is obtained by averaging the amplitude spectrum along the frequency axis:

$$APSD_i = \frac{1}{N}\sum_{j=1}^{N} A_{ij} \tag{17}$$

$$F\_parameters = \begin{bmatrix} freq\_center_1 & freq\_center_2 & \cdots & freq\_center_i \\ freq\_bandwidth_1 & freq\_bandwidth_2 & \cdots & freq\_bandwidth_i \\ APSD_1 & APSD_2 & \cdots & APSD_i \end{bmatrix} \tag{18}$$

where $APSD_i$ is average power spectral density. The fitted $\mu_i$ is the center frequency $freq\_center$, and $\sigma_i$ is the frequency bandwidth $freq\_bandwidth$. $F\_parameters$ is the feature parameter vector of frequency-domain data.

### 2.4.3. Extraction of Phase Spectrum Features

The frequency spectrum of a signal is comprised of the phase spectrum and amplitude spectrum, which offer valuable insights into the signal's frequency components, phase information, and facilitating signal reconstruction. However, during the analysis of the phase spectrum, the occurrence of phase jumps was a common issue. This phenomenon arose from the periodic nature of the phase spectrum and the utilization of the Fourier transform, which operated on periodic signals. Consequently, the phase spectrum may demonstrate discontinuous jumps, referred to as phase wrapping or phase discontinuity, which can present challenges in signal processing and analysis. To address this issue, a frequently employed approach was phase unwrapping [46], also referred to as phase jump removal, which aimed to transform the discontinuous points in the phase spectrum into corresponding continuous points, thereby achieving a smoother phase spectrum curve that was more amenable to analysis and processing.

The specific processing process is as follows.

Phase unwrapping was conducted on each dataset to extract the extremal points of first-order and second-order derivatives. The first-order derivative was calculated as the rate of change between adjacent sampling points, whereas the second-order derivative was calculated as the acceleration of change between three adjacent sampling points. The purpose was to capture the characteristics of phase transformations, specifically the prominence and depression features in the phase spectrum, and to quantify the peak and valley points in the phase spectrum. These measurements were utilized to assess the spatial distribution characteristics of pore spaces, specifically examining their uniformity.

$$P\_parameters_{i,j} = \sum_{\text{pks}>\text{thre}} \text{pks} + \sum_{\text{valleys}<-\text{thre}} \text{valleys} \tag{19}$$

Here, the characteristic parameter matrix $P\_parameters_{i,j}$ contains the $j$-th derivative extremum points of the $i$-th sample, $pks$ and $valleys$ are the points of the first-order and second-order derivative extremum points, and $thre$ is the set threshold.

### 2.4.4. Extraction of Reflectance Spectrum Features

The reflectance spectrum was a set of data utilized to investigate the changes in reflectance at different frequencies, providing insights into the optical, electrical, and physical properties of materials. In this study, the reflectance spectrum served as an informative source concerning the reflection characteristics exhibited by TBCs within the terahertz frequency domain. Extracting feature parameters through statistical analysis and examination of the reflectance spectrum aimed to uncover the underlying patterns of reflectance variation across frequencies. Moreover, it enabled the evaluation of changes in TBCs' porosity, thereby facilitating inferences related to the heterogeneity and distribution of porosity [47]. These derived feature parameters not only offer a feasibility basis for evaluating TBCs but also lay the groundwork for exploring the relationship between variations in porosity and other performance metrics associated with TBCs.

The extraction of the mean reflectance aimed to capture the average reflection performance of the TBC sample across the entire frequency range. The reflectance standard deviation quantified the dispersion of reflectance values, indicating the fluctuation of reflectance in the frequency domain. The maximum and minimum reflectance values represented the extreme reflection conditions observed in TBCs across different frequencies. The reflectance peak highlighted the intense reflection characteristics of the TBC sample at specific frequencies, thereby reflecting the frequency response of porosity to terahertz signals.

The specific calculation formula is as follows:

$$\overline{R} = \frac{\int_{w_{\mathrm{L}}}^{w_{\mathrm{H}}} R(w)dw}{\int_{w_{\mathrm{L}}}^{w_{\mathrm{H}}} dw} \tag{20}$$

$$\sigma_R = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(R_i - \overline{R}\right)^2} \tag{21}$$

$$R_{\max} = \max_{w\in[w_{\mathrm{H}},w_{\mathrm{L}}]} R(w) \tag{22}$$

$$R_{\min} = \min_{w\in[w_{\mathrm{H}},w_{\mathrm{L}}]} R(w) \tag{23}$$

$$R_{\mathrm{p}} = \max_{w\in[w_{\mathrm{H}},w_{\mathrm{L}}]}\left\{\frac{R(w-\Delta w)+R(w)+R(w+\Delta w)}{3}\right\} \tag{24}$$

$$R\_parameters = \left[\overline{R}, \sigma_R, R_{\max}, R_{\min}, R_{\mathrm{p}}\right] \tag{25}$$

where $w_{\mathrm{L}}$ and $w_{\mathrm{H}}$ are the lowest and highest frequencies in the reflectivity spectrum, $N$ is the number of data points in the reflectivity spectrum, and $\Delta w$ is a small frequency- domain interval for smoothing processing. $\overline{R}$, $\sigma_R$, $R_{\max}$, $R_{\min}$, and $R_{\mathrm{p}}$ are the mean reflectance, the standard deviation of reflectance, maximum reflectance, minimum reflectance, and peak reflectance, respectively. *R_parameters* is the feature parameter vector of reflectance spectrum data.

### 2.4.5. Analysis of Multi-Feature Fusion

In this section, a correlation analysis was conducted on the 11 feature parameters extracted from the time-domain spectrum, frequency-domain spectrum, phase spectrum, and reflectance spectrum [48]. The purpose of correlation analysis was to evaluate both the strength and direction of the correlation among multiple feature parameters. This analysis utilized a statistical method to quantify the linear relationship between variables. A correlation heatmap was utilized to visualize the correlation levels between multiple feature parameters. This approach facilitated a precise assessment of the strength and direction of relationships between different variables, resulting in a comprehensive understanding of the interrelationships within the dataset. The application of a heatmap not only identified linear correlations between variables but also revealed complex patterns of relationships among multiple variables. The colors in the correlation heatmap indicated the strength of the correlations, providing insights into the associations between different feature parameters. Positive correlation indicated similar changing trends between variables, negative correlation indicated opposite changing trends, and no correlation indicated the absence of a linear relationship between two variables.

The correlation heatmap was solely capable of illustrating the correlation between variables and does not establish a causal relationship between them. Therefore, it was necessary to consider other factors when making decisions or inferences. The *p*-value is a crucial statistical indicator that measures the consistency between the observed data and the null hypothesis. Additionally, *p*-value testing could be utilized to examine the significant differences among the feature parameters, thereby indicating that the relationships between these parameters are not coincidental and possess statistical significance. Hence, by conducting correlation analysis and *p*-value testing, the revealed correlations among multiple feature parameters hold essential significance in characterizing the heterogeneity of porosity. Exploring the feature parameters within the dataset bears significant application value for a comprehensive investigation of the porosity structural characteristics of TBCs.

### *2.5. Machine Learning Prediction and Performance Evaluation*

#### 2.5.1. Dung Beetle Optimization Algorithm

Dung beetle optimizer (DBO) was an innovative swarm intelligence optimization algorithm proposed in late 2022, primarily inspired by the dung beetle's behaviors of ball rolling, dancing, foraging, stealing, and breeding [49]. In nature, dung beetles have been observed compacting feces into a spherical ball and utilizing celestial cues for navigation to ensure the dung ball is rolled in a straight line. However, in the absence of any light source, the trajectory of dung beetles deviates from a straight line.

To mimic the rolling ball behavior and employ solar navigation, the DBO algorithm enables dung beetles to move in a predetermined direction across the entire search space. Additionally, during the rolling process, the dung beetle's position is updated using the following formula:

$$x_i(t+1) = x_i(t) + \alpha \times k \times x_i(t-1) + b \times \Delta x \tag{26}$$

$$\Delta x = |x_i(t) - X^w| \tag{27}$$

where $t$ is the number of the current iteration, $x_i(t)$ is the position information of the $i$-th dung beetle at the $t$ iteration, $k\epsilon(0, 0.2]$ is a constant representing the deflection coefficient, $b\epsilon(0, 1)$, $\alpha$ is the natural coefficient of 1 or $-1$, $X$ is the global worst position, and $\Delta x$ is the change in simulated light intensity.

When confronted with an obstacle, the beetle reorients itself. Therefore, the location is updated:

$$x_i(t+1) = x_i(t) + \tan\theta |x_i(t) - x_i(t-1)| \tag{28}$$

where $|x_i(t) - x_i(t-1)|$ is the difference between the $t$ and $t-1$ iterations of the $i$-th beetle.

Dung beetles will roll the dung ball to a safe location. As shown in Figure 6, a boundary selection strategy is proposed to simulate the area where female dung beetles lay their eggs:

$$R = 1 - t/T_{\max} \tag{29}$$

$$Lb^* = \max(X^* \times (1-R), Lb) \tag{30}$$

$$Ub^* = \min(X^* \times (1+R), Ub) \tag{31}$$

where $R$ is the scale coefficient, and $T_{\max}$ is the maximum number of iterations. $X$ is the current local optimal location, and $Lb^*$ and $Ub^*$ are the lower and upper bounds of the spawning area, respectively. $Lb$ and $Ub$ are the lower and upper bounds of the optimization problem, respectively.

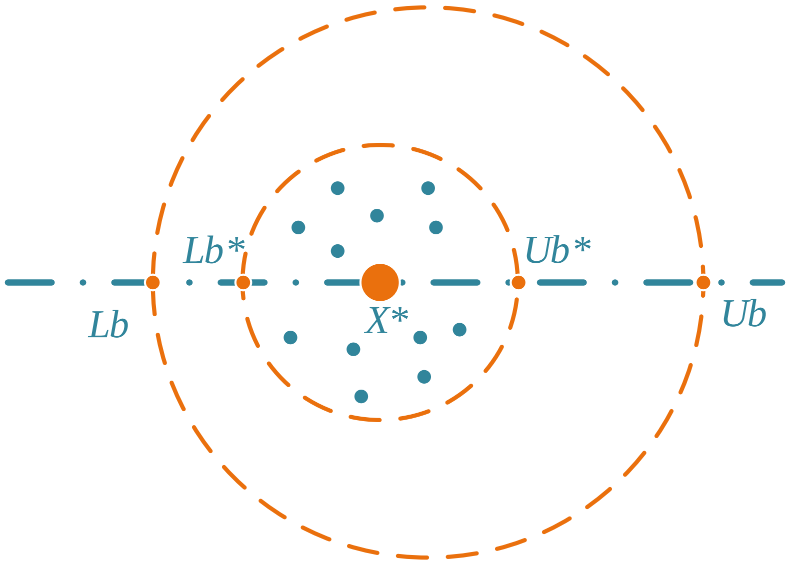



**Figure 6.** Selected rules for boundary limits.

During the iteration process, the position of the oocytes changes dynamically:

$$B_i(t+1) = X^* + b_1 \times (B_i(t) - Lb^*) + b_2 \times (B_i(t) - Ub^*) \tag{32}$$

where $B_i$ is the position of the oosphere, and $b_1$ and $b_2$ are random vectors of $1 \times$ D dimension.

The optimal feeding area is established, and the young dung beetles are guided to forage. The optimal feeding area is defined as follows:

$$Lb^b = \max\left(X^b \times (1-R), Lb\right) \tag{33}$$

$$Ub^b = \min\left(X^b \times (1+R), Ub\right) \tag{34}$$

$$x_i(t+1) = x_i + C_1 \times \left(x_i(t) - Lb^b\right) + C_2 \times \left(x_i(t) - Ub^b\right) \tag{35}$$

where $X^b$ is the global optimal position, $Lb^b$ and $Ub^b$ are the lower and upper bounds of the optimal feeding area, respectively, and the position of little dung beetles is updated. $x_i$ is the beetle's position, $C_1$ is a random number with a normal distribution, and $C_2$ is a random vector.

Assuming $X^b$ is the best place to compete for food, the thieving beetles are updated.

$$x_i(t+1) = X^b + S \times g \times \left(|x_i(t) - X^*| + \left|x_i(t) - X^b\right|\right) \tag{36}$$

All agents are randomly initialized, and the settings are distributed as shown in Figure 7. According to the different types of agents, the positions of the rolling dung beetle, the egg ball, the small dung beetle, and the thief are updated by choosing the appropriate way. Finally, the optimal position $X^b$ and its fitness value are output.

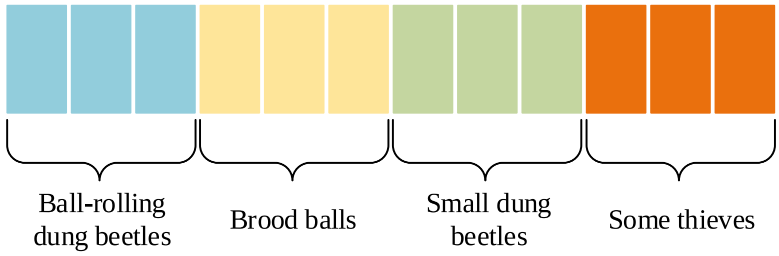



**Figure 7.** Rules for the division of populations.

#### 2.5.2. Random Forest Algorithm

Random forest (RF) regression is an ensemble learning algorithm that constructs multiple decision trees and integrates their predictions to perform regression tasks [50]. In the random forest approach, each decision tree is constructed independently and trained on randomly selected sub-samples, which effectively reduces the risk of overfitting. The predictions of multiple decision trees in the random forest are then averaged or weighted averaged to obtain the final regression result.

The fundamental principles of RF regression are outlined below (as shown in Figure 8).

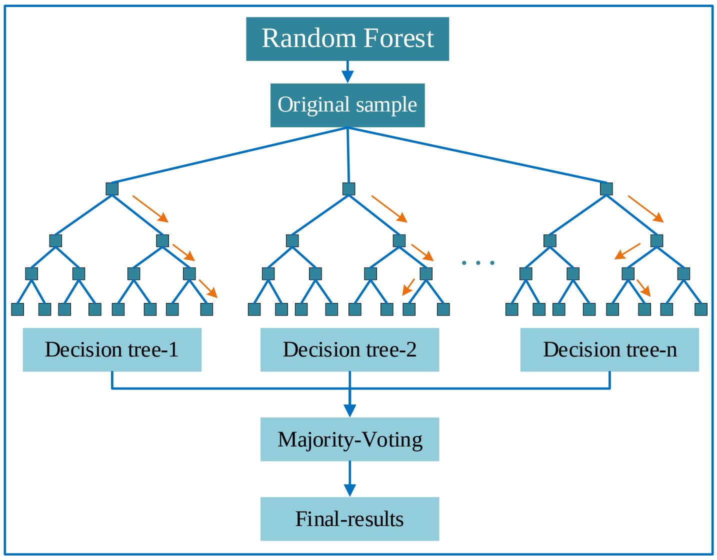



**Figure 8.** Flow chart of random forest algorithm.

Random sample selection: A subset of samples from the original training set is randomly selected to create a sub-sample. This ensures that each decision tree is trained on a different subset of samples, thereby increasing the model's diversity.

Random feature selection: For each node in every decision tree, when determining the best feature for splitting, only a randomly selected subset of features is considered. This prevents certain features from exerting excessive influence on the overall model, thereby enhancing its robustness.

Decision tree construction: A decision tree algorithm (such as the CART algorithm) is utilized to construct a decision tree on each sub-sample. During the growth process of the decision tree, the best splitting feature is typically chosen recursively to partition the dataset into subsets with minimal impurity.

Ensemble prediction: by averaging or weighted averaging the predictions of multiple decision trees, the final regression result is obtained for new input samples.

#### 2.5.3. Random Forest Model Optimized by Dung Beetle Optimization Algorithm

Dung beetle optimizer (DBO) algorithm is utilized to perform iterative searches for finding the optimal solution. This algorithm simulates the interactions among dung beetles in the solution space, with the objective of striking a balance between global exploration and local exploitation. The optimization process of the DBO algorithm focuses on optimizing the parameters of the RF algorithm model. More precisely, the DBO algorithm improves the performance and generalization capability of the random forest model by adaptively modifying feature subset selection and the count of decision trees. By utilizing iterative searching and interplays among dung beetle individuals, this algorithm can identify the optimal combination of feature subsets and the number of decision trees, thereby enhancing the accuracy and robustness of the model.

The specific optimization steps are as follows.

1. Initialization: The number of dung beetle individuals, maximum iteration count, and relevant parameters (such as perception radius and step size) are set. The positions and directions of each dung beetle individual are randomly initialized. Additionally, the probability of information exchange needs to be determined, which dictates the likelihood of information exchange among dung beetle individuals;
2. Fitness calculation: Using the positions of each dung beetle individual as parameter configurations, the random forest model is trained to predict the porosity of thermal barrier coatings. The fitness value of each individual is calculated by assessing the difference between the predicted results and the actual porosity data. The fitness value is a function of the prediction error, typically measured as the root-mean-square error (RMSE);
3. Iterative update: In each iteration, each dung beetle individual can perceive the positions and directions of neighboring dung beetle individuals. The dung beetle individual's position is adjusted based on the specified movement step size and direction. The hyperparameters of the random forest, such as the number of decision trees and the size of feature subsets, are adapted based on the final positions and directions of the dung beetle individuals;
4. Updating the optimal solution: The fitness value of the new position is compared with the fitness value of the current optimal solution. If the fitness value of the new position is better than the current optimal solution, the current optimal solution is updated to the new position, and the corresponding parameter configuration is recorded. During the optimization process, the dung beetle optimization algorithm progressively updates and iterates the optimal solution to search for an improved RF model;
5. Iterative optimization process: Steps 2 to 4 are repeated until the predefined maximum iteration count is reached or a stopping criterion is met. The stopping criterion can be reaching the maximum iteration count or achieving a small change in fitness value after consecutive iterations, indicating convergence to a stable optimal solution;
6. Outputting the optimal solution: Upon completion of the iterative optimization process in the DBO algorithm, the obtained optimal solution corresponds to the optimized parameter configuration of the RF algorithm. The RF model associated with this optimal solution demonstrates enhanced performance and the ability to accurately predict the TBCs' porosity.

In this study, the DBO was employed to adjust and optimize the number of decision trees and the size of feature subsets in a random forest model. More specifically, the internal parameters of the random forest model were treated as adjustable variables, and the DBO algorithm was iteratively applied to determine the optimal combination of these parameters. Afterwards, the RF model was trained using the optimal parameter combination, leading to a significant improvement in performance.

To summarize, the DBO algorithm primarily aimed at optimizing the internal parameters of the random forest model to enhance its predictive accuracy and stability. This algorithm facilitated a global search for the optimal solution and improved the quality of solutions by continuously updating the candidate solutions.

2.5.4. Cross-Validation

Cross-validation was a statistical technique employed to assess the performance of machine learning models and select appropriate parameters [51,52]. Among these techniques, K-fold cross-validation was the most commonly utilized approach. The dataset was randomly divided into K subsets, with K − 1 subsets used for training the model and the remaining subset used for evaluating the model's performance. This process was repeated K times, during which each iteration employed a different subset as the validation set. Finally, the results of the K evaluations were averaged to obtain the final performance indicator.

The specific process is as follows (as shown in Figure 9).

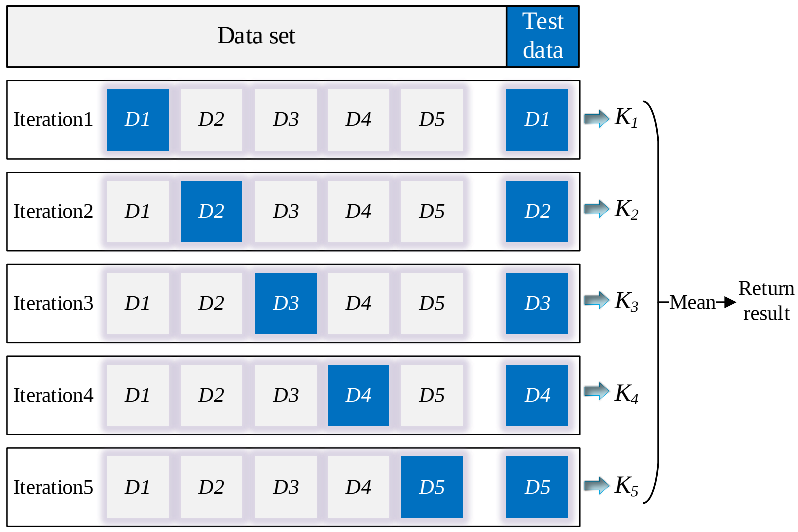



**Figure 9.** Cross-validation schematics.

Data partitioning: The procedure of K-fold cross-validation includes segregating the training dataset into K subsets, wherein K = 5 is chosen in this study. Each subset is used as a part of the validation set, while the remaining K − 1 subsets constitute the training set.

Model training and verification: For each iteration of cross-validation, the random forest model optimized using the DBO optimization algorithm is trained. K − 1 subsets are employed as the training set during each iteration, and one of the remaining subsets is used as the validation set to assess the model's performance.

Performance evaluation: For each cross-validation iteration, the performance metrics of the model on the validation set, i.e., accuracy and precision, are recorded. Finally, the results of the K evaluations are averaged to obtain the final performance metrics as the evaluation results of the model.

Parameters tuning: Cross-validation can be utilized to select appropriate parameters. By trying different parameter combinations and evaluating the model's performance using cross-validation results, the best-performing parameter combination can be chosen.

Using cross-validation performance metrics provided a more objective reflection of the model's generalization ability. It effectively avoided overfitting or underfitting issues on local data and improved the model's generalization ability and stability. In this study, a 5-fold cross-validation method was employed, and the results from each iteration were recorded and statistically analyzed to further validate the reliability and accuracy of the constructed prediction model.

## 3. Results and Discussion

### *3.1. Nondestructive Evaluation of Porosity*

In this study, the terahertz time-domain spectra were processed using the fast independent component analysis (Fast ICA) algorithm, yielding a set of kurtosis values for

20 samples. Kurtosis, a statistical metric quantifying the sharpness and peakedness of signal waveforms, was employed to describe the signal characteristics. The extracted kurtosis values were then analyzed in association with the corresponding porosity and visually presented. As shown in Figure 10a, after ICA treatment of individual components, it was observed that the waveform, amplitude, and frequency of each component were significantly different. It was evident that these components exhibited complex and redundant characteristics, making it challenging to discern clear patterns that could effectively represent porosity. To address this issue, we proceeded to calculate the kurtosis of each component, which served as a representative feature parameter for porosity assessment. As shown in Figure 10b, a decreasing trend in kurtosis values was observed with increasing porosity. This indicated that as the TBCs' porosity increased, the waveform of its terahertz time-domain spectra became smoother. This phenomenon can be attributed to the structural changes occurring within the thermal barrier coatings due to the increasing porosity, which consequently led to variations in the internal permittivity and refractive indicators of the samples.

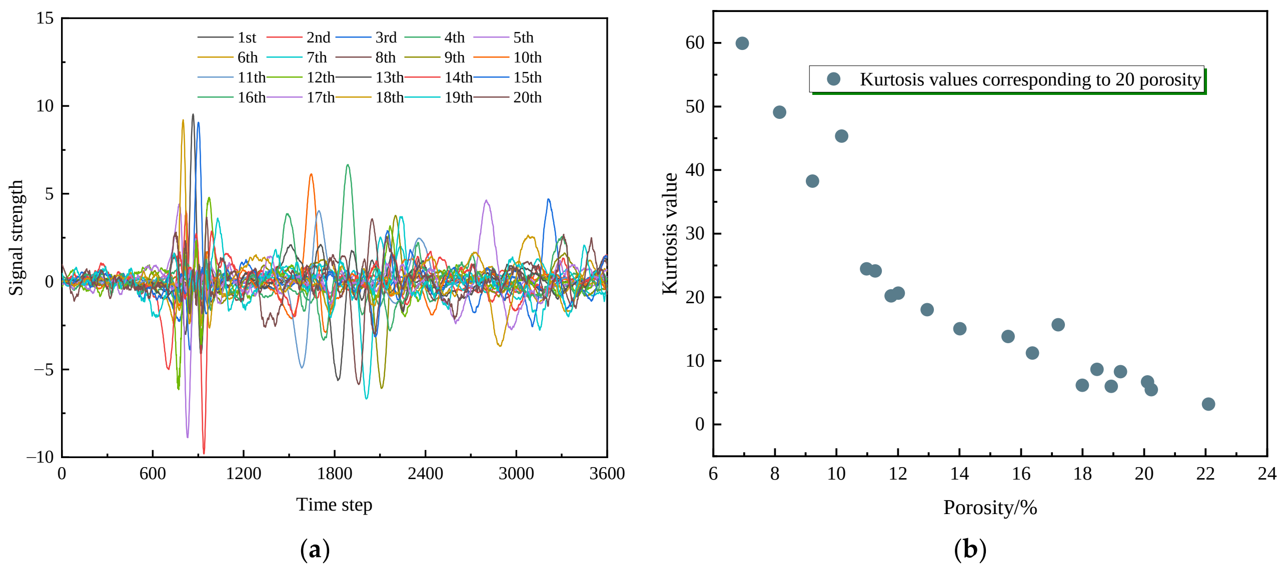



**Figure 10.** Results of Fast ICA: (**a**) ICA separation visualization; (**b**) kurtosis evaluation of porosity.

More precisely, the augmentation of voids and pore quantity caused the propagation path of the terahertz signal within the coating to become more intricate and convoluted, thereby diminishing the sharpness and steepness of the waveform. As the porosity varied, the complexity and heterogeneity within the TBCs became more prominent, consequently causing the interaction between the terahertz wave and the medium to exhibit a greater degree of complexity and diversity. Correspondingly, the kurtosis feature parameter values of the spectra also underwent changes. Hence, the fluctuations in kurtosis can serve as a reflection and evaluation of the porosity.

Furthermore, as shown in Figure 11, the spectrum was subjected to a smoothing process, followed by Gaussian fitting of the main peak to extract additional feature parameters, including center frequency, bandwidth, and APSD. These feature parameters provided crucial information about the frequency components and energy distribution of the signal, which were then correlated with the porosity. As shown in Figure 12a, where the small figure was a curved representation of the smooth spectrum, it was another form of Figure 11. The center frequency was the extreme point of the smoothed curve, and the center frequency decreased with the increase in porosity. As shown in Figure 12b, the frequency bandwidth was the frequency broadening when the peak dropped to half, and the frequency bandwidth increased with the increase in porosity. The changes in center frequency and frequency bandwidth with porosity showed strong regularity, which was the result of Gaussian fitting. Figure 13a shows a single power spectrum used to calculate

APSD. Figure 13b shows that the value of APSD decreased as a whole with the increase in porosity.

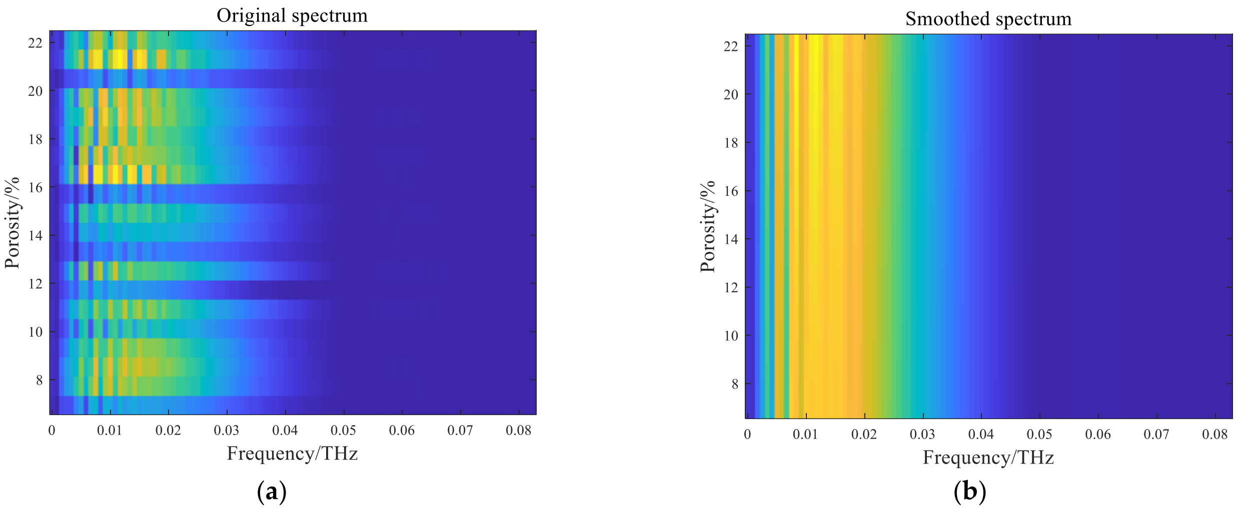



**Figure 11.** Frequency-domain smoothing: (**a**) original frequency spectrum; (**b**) smoothed frequency spectrum.

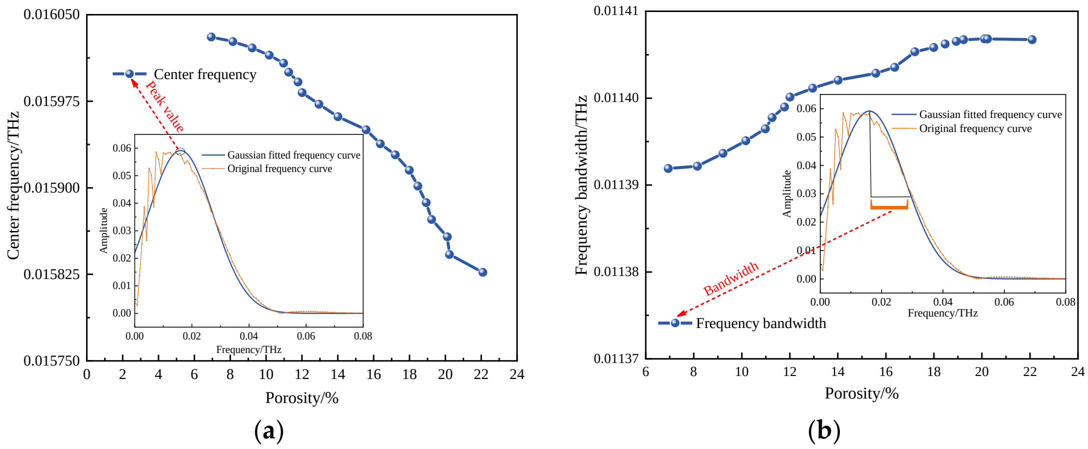



**Figure 12.** Gaussian fitting extraction: (**a**) center frequency; (**b**) frequency bandwidth.

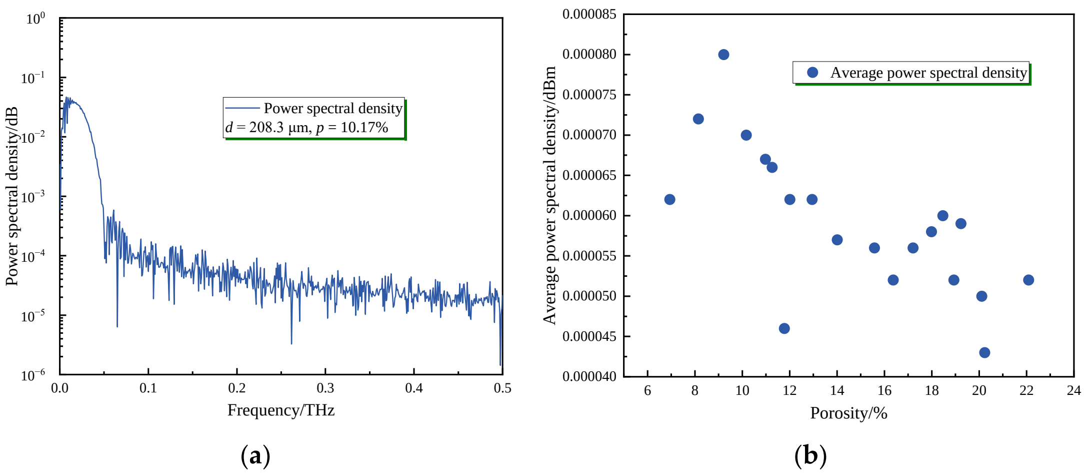



**Figure 13.** Average power spectral density extraction: (**a**) power spectral density; (**b**) APSD corresponding to different porosity.

More precisely, the variation in porosity led to changes in the dielectric properties within the TBCs. As porosity increased, scattering effects were enhanced, resulting in a decrease in the dielectric constant and a reduction in wave velocity. Consequently, the overall frequency decreased, causing the center frequency to shift towards lower frequencies. Moreover, an increased porosity introduced more scattering and multipath effects, making the signal transmission more complex. This broadened the signal spectrum, indicating a wider distribution of energy across a range of frequencies, thus increasing the bandwidth. Additionally, the enhanced heterogeneity of the TBCs due to increased porosity led to THz signal scattering and attenuation, thereby reducing the overall energy and resulting in a decrease in APSD. By extracting features from the frequency-domain data, the frequency-response characteristics and energy distribution of the TBCs could be quantitatively analyzed. The parameter variations associated with porosity provided insights into the internal structure and physical properties, thereby offering crucial information for performance assessment and optimization of the thermal barrier coatings.

Subsequently, the occurrence of phase jumps was attributed to the periodic and discontinuous nature of the signal. To mitigate these jumps, a phase unwrapping method was applied to the phase spectrum, as shown in Figure 14a, resulting in a reconstructed phase curve that was continuous and smooth. Additionally, feature parameters were derived from the number of extremal points in the first-order and second-order derivatives. As shown in Figure 14b, a slight decrease in the number of extremal points was observed with increasing porosity.

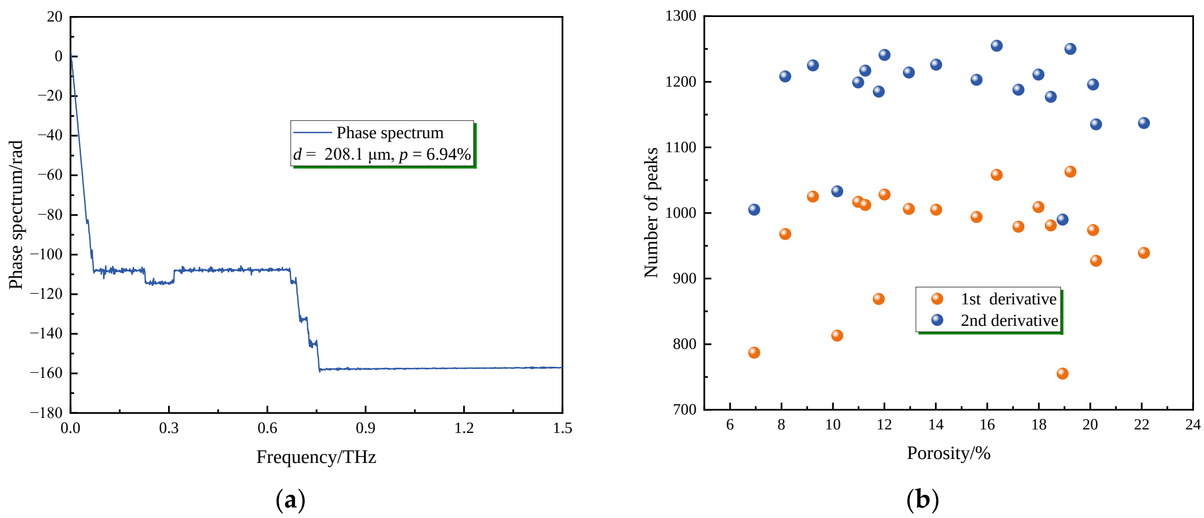



**Figure 14.** Feature extraction of phase spectrum: (**a**) removal of the phase jump; (**b**) extreme points of the first and second derivatives.

More precisely, despite the limited significance of the number of extremal points in the first-order and second-order derivatives in reflecting the variation pattern of the phase curve with increasing porosity, the feature parameters still provided information regarding the rate of change in the phase curve. As porosity increased, the phase curve tended to exhibit a greater number of local extremal points. Although these extremal points might be affected by the complexity of the phase curve, they still captured the overall trend and characteristics of the phase curve. Therefore, the investigation of porosity evaluation using terahertz analysis remained necessary as it served as a strong foundation for terahertz signal analysis. By combining these feature parameters with others, more comprehensive and integrated information was obtained to collectively characterize the changes in porosity within the TBCs. Such a comprehensive analysis contributed to a more accurate assessment of the performance and characteristics of TBCs.

Reflectance spectrum contained essential information regarding the material's electromagnetic response and exhibited a certain level of sensitivity towards porosity. The extracted feature parameters, including average reflectance, reflectance standard deviation, maximum reflectance, minimum reflectance, and reflectance peak, were employed to quantitatively characterize the variations in porosity. Furthermore, these parameters were visually represented to showcase their correlation with porosity, as shown in Figure 15a–e.

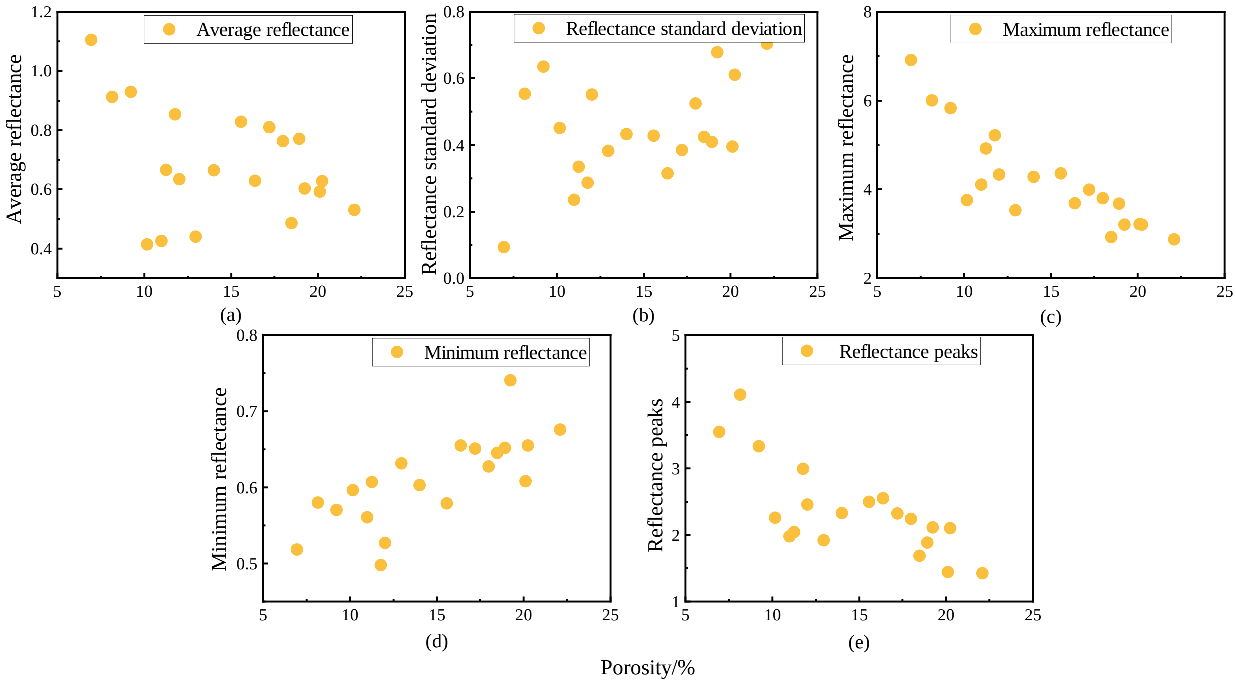



**Figure 15.** Feature extraction of reflectance spectrum: (**a**) average reflectance; (**b**) reflectance standard deviation; (**c**) maximum reflectance; (**d**) minimum reflectance; (**e**) reflectance peaks.

More precisely, increasing porosity resulted in greater attenuation of the terahertz signal within the coating, leading to a decrease in overall reflectance. The reflectance standard deviation measured the degree of variation between individual reflectance values and the average reflectance in the reflectance spectrum. As porosity increased, the internal structure and composition of the coating became more complex and heterogeneous, causing greater fluctuations and differences in the terahertz signal reflectance, thus increasing the reflectance standard deviation. The maximum reflectance exhibited a decreasing trend, while the minimum reflectance showed an increasing trend. With increasing porosity, the dielectric interface of the coating increased, and the synergistic effect of multiple interfaces affected the terahertz reflection properties. This increased the reflection and scattering, causing a decrease in the maximum reflectance. Simultaneously, in the porous regions, the penetration capability of the terahertz signal was enhanced, resulting in an increase in the minimum. With increasing porosity, the reflectance peak became smoother and displayed a lower peak intensity at specific frequencies, indicating a weakened characteristic of strong reflection.

The correlation heatmap was employed to visually depict the relationships among multiple terahertz feature parameters by representing the correlation coefficients between corresponding variables. The color intensity in each cell indicated the magnitude of the correlation coefficient, with darker colors indicating stronger correlations and lighter colors indicating weaker correlations. The elongation of the elliptical shape denoted higher correlation strength, with leftward tilting indicating negative correlation and rightward tilting indicating positive correlation. Additionally, the significance of the relationships between feature parameters was assessed using $p$-value tests: $p < 0.05$ indicated a probability of occurrence of the feature parameter samples of less than 5%, signifying a significant statistical difference; $p < 0.01$ indicated a highly significant difference; $p < 0.001$ indicated an extremely significant difference, as shown in Figure 16.

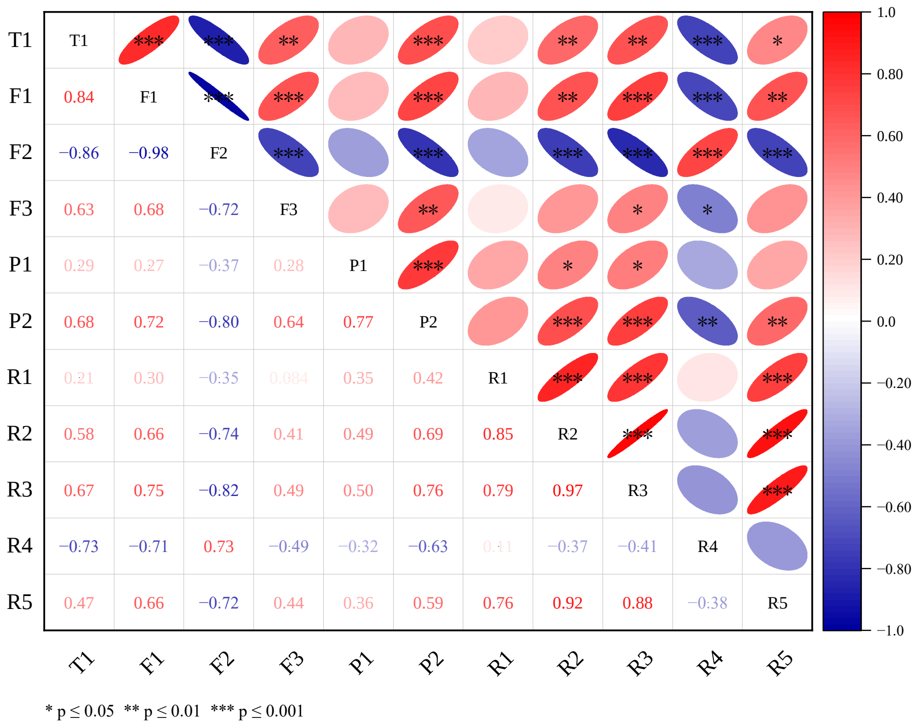



**Figure 16.** Correlation and significance analysis of multi-feature fusion. T1 is kurtosis; F1 is center frequency; F2 is frequency bandwidth; F3 is APSD; P1 and P2 are the first and second derivatives; R1, R2, R3, R4, and R5 are average reflectance, reflectance standard deviation, maximum reflectance, minimum reflectance, and reflectance peaks, respectively.

According to the outcomes of the correlation analysis, noticeable correlations were found among multiple feature parameters, and these correlations demonstrated significant statistical differences. The results of the correlation analysis demonstrated the feasibility of utilizing multiple feature parameters to describe and characterize variations in porosity, and differences were not attributable to random factors. Furthermore, these correlated feature parameters provided effective input features for subsequent machine learning prediction models, resulting in more accurate porosity predictions. Additionally, the correlation analysis served as a foundation for exploring the complex relationships between the feature parameters, allowing for further in-depth investigations into the mechanisms of their interactions.

### *3.2. Machine Learning Prediction of Porosity*

In order to predict the TBCs' porosity using machine learning, the dung beetle optimizer algorithm was employed to optimize the random forest regression model. More precisely, the previous sections involved conducting correlation analysis among the THz multi-feature parameters through the construction of a correlation heatmap and conducting *p*-value tests. The purpose of the correlation analysis was to determine the relationships between the feature parameters, identify the influential features for porosity prediction, and eliminate any potential redundant information among them. This approach effectively reduced the feature dimensionality, leading to improved model efficiency and predictive capability.

During the porosity prediction process, the random forest regression model was optimized using the dung beetle optimizer algorithm. The DBO, which simulated the foraging behavior of fireflies, was employed to search for the optimal solution through an adaptive search and iterative update process. By optimizing the model parameters, the RF model was able to fit the porosity data more accurately. This optimization aimed to improve the RF model's ability to predict porosity and enhance its predictive performance. Subsequently, the 11 extracted features were merged as input, and the optimized

RF regression model was utilized to predict porosity. Figure 17 shows the three curves depicting the test values of porosity and the predicted values of the RF model and DBO-RF model, respectively. Comparing these curves, it was evident that the predicted values of the DBO-RF model closely matched the test values, while the RF model exhibited significantly lower predictive accuracy compared to the DBO optimized model. This observation clearly demonstrated the remarkable optimization effect of the DBO algorithm, which led to a substantial improvement in predictive accuracy.

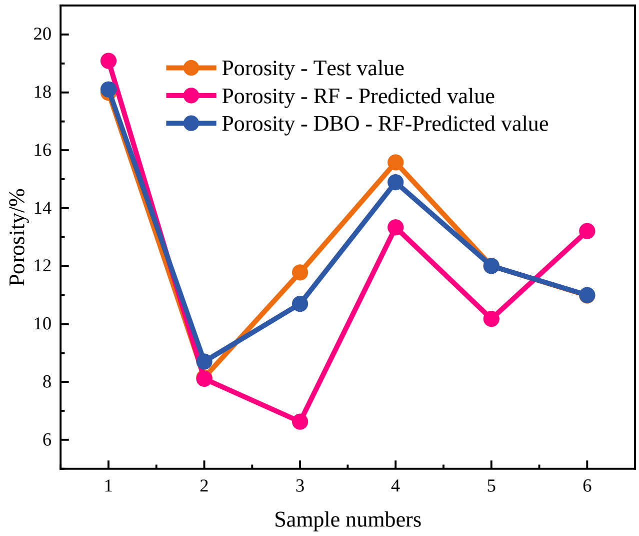



**Figure 17.** Prediction of porosity by DBO-RF model.

Following the feature extraction, correlation and significance analysis, and model optimization, an optimized model was obtained for porosity prediction. To objectively evaluate and validate the performance of the model, K-fold cross-validation was employed. Cross-validation was a prevalent approach for assessing model performance that mitigated dependence on a single training and testing set to split, thereby enhancing the model's stability and reliability. In this study, five-fold cross-validation was implemented to evaluate the model's consistency and generalization ability.

Figure 18 exhibited the iteration diagram for each fold, allowing for the assessment of the convergence behavior and performance metrics of the DBO-RF model. The figure highlighted that the five folds attained convergence after 159, 274, 154, 262, and 237 iterations, respectively, underscoring a relatively swift convergence rate. After cross-validation, the accuracy and robustness of the DBO-RF prediction model need to be evaluated. Through evaluation indicators, root-mean-square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and regression coefficient ($R^2$) were used to measure the performance of the model. Additionally, Table 1 provides a comprehensive evaluation indicator, including performance metrics for each fold and the overall average evaluation index of DBO-RF-Crossvalind. Obviously, the prediction error of each fold was very small, and the regression coefficient was about 0.9. The mean error values of RMSE, MAE, and MAPE were 1.802, 1.549, and 8.362, respectively. The mean regression coefficient was 0.912. The model has excellent prediction accuracy.

By performing statistical analysis and examination of the iteration counts and prediction results for each fold, a more comprehensive, accurate, and reliable predictive model was obtained. Within the five-fold cross-validation, this approach demonstrated high accuracy and robustness, providing strong support for practical applications. Additionally, the analysis of evaluation indicators indicated the superiority of the DBO-RF model for porosity prediction in TBCs. These analytical findings will serve as important references for further model optimization and predictive applications.

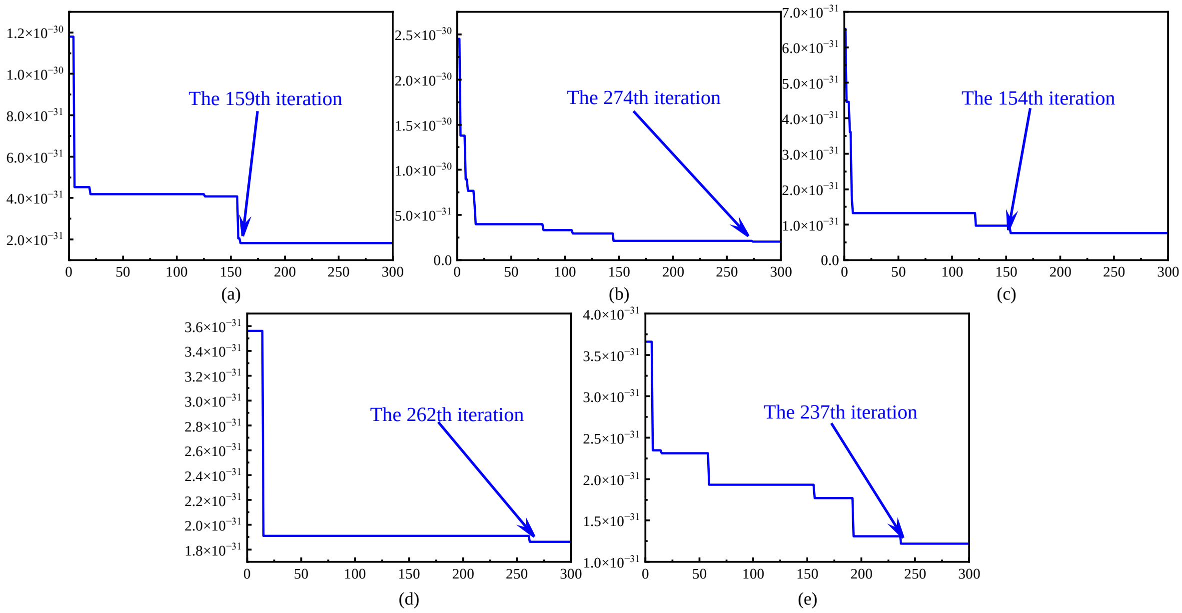



**Figure 18.** Five-fold iteration results: (**a**) first fold; (**b**) second fold; (**c**) third fold; (**d**) fourth fold; (**e**) fifth fold.

**Table 1.** DBO-RF-Crossvalind model performance evaluation.

| K-Fold | RMSE | MAE | MAPE | $R^2$ |
|---|---|---|---|---|
| K1 | 1.077 | 0.962 | 6.108 | 0.942 |
| K2 | 3.329 | 3.219 | 10.261 | 0.872 |
| K3 | 1.211 | 0.982 | 7.052 | 0.928 |
| K4 | 1.764 | 1.375 | 9.725 | 0.906 |
| K5 | 1.629 | 1.205 | 8.665 | 0.911 |
| Average | 1.802 | 1.549 | 8.362 | 0.912 |

Based on the analysis of terahertz signals, it was observed that the TBCs' porosity was influenced by multiple feature parameters to varying extents. In the model construction phase, multi-feature fusion parameters were adopted as the input, and the random forest model was trained and optimized using the DBO algorithm. Compared to the predictive models relying on single feature parameters, the predictive model incorporating the multi-feature feature parameters allowed for a more comprehensive exploration of the relationship between porosity and several other parameters, leading to enhanced prediction accuracy and stability.

Through the application of terahertz spectroscopy, terahertz multi-feature parameters were obtained and utilized as inputs to construct a DBO-RF prediction model for accurately predicting the TBCs' porosity. Experimental validation was conducted, yielding reliable and precise prediction results. The five-fold cross-validation methodology was employed to evaluate the model, and performance evaluation metrics such as RMSE, MAE, MAPE, and $R^2$ were utilized. The results demonstrated that the constructed DBO-RF prediction model exhibited high accuracy and robustness during the five-fold cross-validation, as indicated by the attainment of favorable levels of average RMSE, MAE, and MAPE. Furthermore, the regression coefficients provided evidence of the model's excellent fitting performance.

Our study incorporates multiple terahertz feature parameters to enable precise evaluation of porosity in TBCs. Furthermore, when combined with machine learning models, this approach facilitates accurate porosity prediction. Unlike some researchers who rely solely on single feature parameters to characterize porosity, our approach provides a more comprehensive representation. Additionally, machine learning models offer broader applicability compared to the limitations of linear fitting and regression for porosity calculations. They have consistently demonstrated outstanding performance and robustness, providing valuable insights for enhancing the performance and quality control of TBCs. Furthermore, our research contributes to the interdisciplinary field by integrating THz-TDS technology with machine learning algorithms from materials science and information science. This comprehensive approach paves the way for future studies in nondestructive testing and porosity evaluation across various materials, offering a promising avenue for research.

## 4. Conclusions

In this study, an efficient solution was proposed for the nondestructive evaluation and prediction of TBCs' porosity, based on the terahertz multi-feature fusion and machine learning approach. The extraction of multiple terahertz feature parameters was carried out, and subsequent correlation and significance analyses facilitated the fusion of these parameters for the nondestructive evaluation of porosity variations. Moreover, a robust random forest regression model, optimized using the dung beetle algorithm, was successfully developed to ensure accurate prediction of TBCs' porosity.

Based on the analysis of the experimental results, it was observed that the approach of terahertz multi-feature fusion comprehensively characterized the TBCs' porosity and effectively revealed the internal pore structure's characteristics. Subsequently, the DBO-RF model was employed, utilizing the fused feature parameters as inputs and porosity as the output. The experimental findings demonstrated the precise prediction of TBCs' porosity. Furthermore, through the implementation of five-fold cross-validation and analysis of evaluation indicators, the superior performance of the DBO-RF model in porosity prediction was confirmed. The mean error values of RMSE, MAE, and MAPE were 1.802, 1.549, and 8.362, respectively. The mean regression coefficient was 0.912, indicating the model's high accuracy and robustness in accurately predicting porosity in TBCs.

In conclusion, our study proposes an efficient solution for the nondestructive evaluation and prediction of TBCs' porosity. This solution is based on a quantitative characterization method that integrates terahertz multi-feature fusion and machine learning. The results demonstrate a porosity prediction accuracy exceeding 90%, highlighting its significant potential for practical applications in evaluating and analyzing the microstructure of TBCs. This method enables effective monitoring of the thermal barrier coating's service life and determination of its validity. Additionally, our research paves the way for expanding the application of this method to assess other aerospace materials, contributing to the advancement of materials science and related fields. Looking ahead, the continuous development and refinement of nondestructive techniques for evaluating various materials hold great promise. Further research in this direction has the potential to drive advancements in material characterization methodologies and contribute to a more scientific understanding of material performance and reliability.

**Author Contributions:** Conceptualization, D.Y.; data curation, R.L. and D.Y.; funding acquisition, D.Y. and J.P.; investigation, R.L., Q.Z. and J.X.; methodology, R.L., D.Y., Q.Z. and J.X.; project administration, D.Y.; resources, D.Y.; supervision, D.Y. and J.P.; visualization, Q.Z. and J.X.; writing—original draft, R.L.; writing—review and editing, D.Y. and J.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (52205547), Key Research and Development Projects in Anhui Province (2022a05020004, 202104a05020002), the Open Research Fund of Fujian Provincial Key Laboratory of Terahertz Functional Devices and Intelligent Sensing, Fuzhou University (FPKLTFDIS202301), the Open Research Fund of Huzhou Key Laboratory of Terahertz Integrated Circuits and Systems (HKLTICY23KFO5), the Open Research Fund of Anhui Key Laboratory of Detection Technology and Energy Saving Devices (JCKJ2022A08), the Anhui Institute of Future Technology Enterprise Cooperation Project (2023qyhz04), and the Scientific Research Starting Foundation of Anhui Polytechnic University (2021YQQ029).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

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

## References


1. Padture, N.P.; Gell, M.; Jordan, E.H. Thermal barrier coatings for gas-turbine engine applications. *Science* **2002**, *296*, 280–284. [CrossRef] [PubMed]
2. Wang, Y. Review and prospects for 20-year development of thermal sprayed nanocoatings. *Surf. Technol.* **2016**, *45*, 1–9.
3. Wang, R.; Dong, T.; Di, Y.; Wang, H.; Li, G.; Liu, L. High temperature oxidation resistance and thermal growth oxides formation and growth mechanism of double-layer thermal barrier coatings. *J. Alloys Compd.* **2019**, *798*, 773–783. [CrossRef]
4. Huang, J.; Wang, W.; Li, Y.; Fang, H.; Ye, D.; Zhang, X.; Tu, S. Novel-structured plasma-sprayed thermal barrier coatings with low thermal conductivity, high sintering resistance and high durability. *Ceram. Int.* **2020**, *47*, 5156–5167. [CrossRef]
5. Liu, Q.; Huang, S.; He, A. Composite ceramics thermal barrier coatings of yttria stabilized zirconia for aero-engines. *J. Mater. Sci. Technol.* **2019**, *35*, 2814–2823. [CrossRef]
6. Mehboob, G.; Liu, M.; Xu, T.; Hussain, S.; Mehboob, G.; Tahir, A. A review on failure mechanism of thermal barrier coatings and strategies to extend their lifetime. *Ceram. Int.* **2020**, *46*, 8497–8521. [CrossRef]
7. Amirhossein, P.; Fariborz, S.; Amirhossein, E.; Leila, B.; Mehdi, R.N.; Milad, B.; Kamalan, K. Failure mechanisms and structure tailoring of YSZ and new candidates for thermal barrier coatings: A systematic review. *Mater. Des.* **2022**, *222*, 111044.
8. Xu, G.; Yang, L.; Zhou, Y. Investigating the interface cracking mechanism of CMAS-corroded thermal barrier coatings based on the cohesive zone model. *Corros. Sci.* **2022**, *204*, 110337. [CrossRef]
9. Yang, S. Multiscale modeling of chemo-thermo-mechanical damage of EB-PVD thermal barrier coatings. *J. Mech. Phys. Solids.* **2022**, *158*, 104667. [CrossRef]
10. Feng, Y.; Dong, T.; Li, G.; Wang, R.; Zhao, X.; Liu, Q. High temperature oxidation resistance and TGO growth mechanism of laser remelted thermal barrier coatings. *J. Alloys Compd.* **2020**, *828*, 154266. [CrossRef]
11. Burger, N.; Laachachi, A.; Ferriol, M.; Lutz, M.; Toniazzo, V.; Ruch, D. Review of thermal conductivity in composites: Mechanisms, parameters and theory. *Prog. Polym. Sci.* **2016**, *61*, 1–28. [CrossRef]
12. Qiu, Q. Effect of internal defects on the thermal conductivity of fiber-reinforced polymer (FRP): A numerical study based on micro-CT based computational modeling. *Mater. Today Commun.* **2023**, *36*, 106446. [CrossRef]
13. Ye, D.; Wang, W. Research progress in terahertz non-destructive testing of thermal barrier coatings. *Surf. Technol.* **2020**, *49*, 126–137.
14. Emine, B.; Daniel, E.; Georg, M.; Robert, M.; Robert, V. Porosity–property relationships of plasma-sprayed $Gd_2Zr_2O_7$/YSZ thermal barrier coatings. *J. Am. Ceram. Soc.* **2015**, *98*, 2647–2654.
15. Lv, B.; Xie, H.; Xu, R.; Fan, X.; Zhang, W.; Wang, T. Effects of sintering and mixed oxide growth on the interface cracking of air-plasma-sprayed thermal barrier coating system at high temperature. *Appl. Surf. Sci.* **2016**, *360*, 461–469. [CrossRef]
16. Zhu, Y.; Yan, B.; Cai, D.; Wu, K.; Zhang, X. Structural parameter study on stress intensity factors of interfacial crack in thermal barrier coatings. *Ceram. Int.* **2021**, *47*, 14354–14365. [CrossRef]
17. Wellington, U.; Mattias, O.; Nicolaie, M.; Mohit, G.; Petra, H.; Zdeněk, V. Influence of spray angle on microstructure and lifetime of suspension plasma-sprayed thermal barrier coatings. *J. Therm. Spray Technol.* **2022**, *31*, 2068–2090.
18. Satyapal, M.; Céline, R.; Nicholas, C.; Jonas, H.; Stefan, B.; Nicolaie, M.; Per, N. Understanding the effect of material composition and microstructural design on the erosion behavior of plasma sprayed thermal barrier coatings. *Appl. Surf. Sci.* **2019**, *488*, 170–184.
19. Liu, Y.; Hu, X.; Zhu, Y.; Wei, H.; Dravid, V.; Zhang, W. Effects of isothermal oxidation on microstructure and mechanical properties of thermal barrier coatings. *Ceram. Int.* **2019**, *45*, 8815–8823. [CrossRef]
20. Khedmatgozar Dolati, S.S.; Caluk, N.; Mehrabi, A.; Khedmatgozar Dolati, S.S. Non-destructive testing applications for steel bridges. *Appl. Sci.* **2021**, *11*, 9757. [CrossRef]
21. Khedmatgozar Dolati, S.S.; Malla, P.; Ortiz, J.D.; Mehrabi, A.; Nanni, A. Identifying NDT methods for damage detection in concrete elements reinforced or strengthened with FRP. *Eng. Struct.* **2023**, *287*, 116155. [CrossRef]
22. Ma, Z.; Zhang, W.; Luo, Z.; Sun, X.; Li, Z.; Li, L. Ultrasonic characterization of thermal barrier coatings porosity through BP neural network optimizing Gaussian process regression algorithm. *Ultrasonics* **2020**, *100*, 105981. [CrossRef] [PubMed]
23. Cernuschi, F. Can TBC porosity be estimated by non-destructive infrared techniques? A theoretical and experimental analysis. *Surf. Coat. Technol.* **2015**, *272*, 387–394. [CrossRef]
24. Tian, H.; Wang, C.; Guo, M.; Gao, J.; Cui, Y.; Jin, G.; Liu, E.; Wang, F. Corrigendum to "Preparation and performance of thermal barrier coatings made of BNw-containing modified $Nd_2O_3$-doped yttria-stabilized zirconia". *Ceram. Int.* **2020**, *46*, 6999.
25. Taheri, H.; Gonzalez Bocanegra, M.; Taheri, M. Artificial intelligence, machine learning and smart technologies for nondestructive evaluation. *Sensors* **2022**, *22*, 4055. [CrossRef]
26. Mishra, M.; Bhatia, A.S.; Maity, D. Predicting the compressive strength of unreinforced brick masonry using machine learning techniques validated on a case study of a museum through nondestructive testing. *J. Civ. Struct. Health* **2020**, *10*, 389–403. [CrossRef]

27. Yee, T.S.; Shrifan, N.H.M.M.; Al-Gburi, A.J.A.; Isa, N.A.M.; Akbar, M.F. Prospect of using machine learning-based microwave nondestructive testing technique for corrosion under insulation: A Review. *IEEE Access.* **2022**, *10*, 88191–88210. [CrossRef]
28. Ye, D.; Wang, W.; Huang, J.; Lu, X.; Zhou, H. Nondestructive interface morphology characterization of thermal barrier coatings using terahertz time-domain spectroscopy. *Coatings* **2019**, *9*, 89. [CrossRef]
29. Ye, D.; Wang, W.; Zhou, H.; Fang, H.; Huang, J.; Li, Y.; Gong, H.; Li, Z. Characterization of thermal barrier coatings microstructural features using terahertz spectroscopy. *Surf. Coat. Technol.* **2020**, *394*, 125836. [CrossRef]
30. Davit, H.; Maher, H.; Olivier, R.; Anthony, B.; Alexis, M.; Laurence, B.; Olivier, D.; Emmanuel, A. Non-destructive evaluation of ceramic porosity using terahertz time-domain spectroscopy. *J. Eur. Ceram. Soc.* **2022**, *42*, 525–533.
31. Sun, F.; Cao, B.; Fan, M.; Liu, L. Physics-based deep learning framework for terahertz thickness measurement of thermal barrier coatings. *SSRN Electron. J.* **2023**, 4373227. [CrossRef]
32. Luo, M.; Zhong, S.; Yao, L.; Tu, W.; Nsengiyumva, W.; Chen, W. Thin thermally grown oxide thickness detection in thermal barrier coatings based on SWT-BP neural network algorithm and terahertz technology. *Appl. Opt.* **2020**, *59*, 4097–4104. [CrossRef] [PubMed]
33. Ye, D.; Wang, W.; Zhou, H.; Li, Y.; Fang, H.; Huang, J.; Gong, H.; Li, Z. Quantitative determination of porosity in thermal barrier coatings using terahertz reflectance spectrum: Case study of atmospheric-plasma-sprayed YSZ coatings. *IEEE Trans. Terahertz Sci. Technol.* **2020**, *10*, 383–390. [CrossRef]
34. Ye, D.; Wang, W.; Zhou, H.; Huang, J.; Wu, W.; Gong, H.; Li, Z. In-situ evaluation of porosity in thermal barrier coatings based on the broadening of terahertz time-domain pulses: Simulation and experimental investigations. *Opt. Express.* **2019**, *27*, 28150–28165. [CrossRef] [PubMed]
35. Li, R.; Ye, D.; Xu, Z.; Yin, C.; Xu, H.; Zhou, H.; Yi, J.; Chen, Y.; Pan, J. Nondestructive evaluation of thermal barrier coatings thickness using terahertz time-domain spectroscopy combined with hybrid machine learning approaches. *Coatings* **2022**, *12*, 1875. [CrossRef]
36. Li, R.; Ye, D.; Xu, J.; Pan, J. Multi-scale analysis of terahertz time-domain spectroscopy for inversion of thermal growth oxide thickness in thermal barrier coatings. *Coatings* **2023**, *13*, 1294. [CrossRef]
37. Zhao, Y.; Ge, Y.; Jin, X.; Koch, D.; Vaßen, R.; Chen, Y.; Fan, X. Oxidation behavior of double-ceramic-layer thermal barrier coatings deposited by atmospheric plasma spraying and suspension plasma spraying. *Ceram. Int.* **2022**, *48*, 23938–23945. [CrossRef]
38. Chen, L.; Meng, G.; Li, C.; Yang, G. Critical scale grain size for optimal lifetime of TBCs. *J. Mater. Sci. Technol.* **2022**, *115*, 241–250. [CrossRef]
39. Park, H.; Son, J. Machine learning techniques for THz imaging and time-domain spectroscopy. *Sensors* **2021**, *21*, 1186. [CrossRef]
40. Stoik, C.; Bohn, M.; Blackshire, J. Nondestructive evaluation of aircraft composites using transmissive terahertz time domain spectroscopy. *Opt. Express.* **2008**, *16*, 17039–17051. [CrossRef]
41. Yuan, B.; Wang, W.; Ye, D.; Zhang, Z.; Fang, H.; Yang, T.; Wang, Y.; Zhong, S. Nondestructive evaluation of thermal barrier coatings thickness using terahertz technique combined with PCA–GA–ELM algorithm. *Coatings* **2022**, *12*, 390. [CrossRef]
42. Qin, Y.; Zhang, C.; Zhu, D.; Zhu, Y.; Guo, H.; You, G.; Tang, S. Engineered nonlinear photonic quasicrystals for multi-frequency terahertz manipulation. *Opt. Express.* **2009**, *17*, 11558–11564. [CrossRef] [PubMed]
43. Kehrt, M.; Monte, C.; Beyer, J.; Hollandt, J. A highly linear superconducting bolometer for quantitative THz Fourier transform spectroscopy. *Opt. Express.* **2015**, *23*, 11170–11182. [CrossRef] [PubMed]
44. Yin, X.; Zhang, Y.; Cao, J.; Wu, J.; Hadjiloucas, S. Exploring the complementarity of THz pulse imaging and DCE-MRIs: Toward a unified multi-channel classification and a deep learning framework. *Comput. Meth. Prog. Biomed.* **2016**, *137*, 87–114. [CrossRef]
45. Liu, P.; Shi, W.; Xu, D.; Zhang, X.; Yao, J.; Norwood, R.; Peyghambarian, N. High-power high-brightness terahertz source based on nonlinear optical crystal fiber. *IEEE J. Sel. Top. Quantum Electron.* **2016**, *22*, 360–364. [CrossRef]
46. Lei, T.; Yang, S.; Tobin, B.; O'Reilly, C.; Sun, D. A measurement framework using THz Time-Domain sensing for wood quality assessment across tree ring samples. *Comput. Electron. Agric.* **2022**, *202*, 107437. [CrossRef]
47. Liu, H.; Chen, Y.; Bastiaans, G.; Zhang, X. Detection and identification of explosive RDX by THz diffuse reflection spectroscopy. *Opt. Express* **2006**, *14*, 415–423. [CrossRef]
48. Pradipta, M.; Ankita, M. Multimodal omics data integration using max relevance-max significance criterion. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 1841–1851.
49. Xue, J.; Shen, B. Dung beetle optimizer: A new meta-heuristic algorithm for global optimization. *J. Supercomput.* **2022**, *79*, 7305–7336. [CrossRef]
50. Schonlau, M.; Zou, R. The random forest algorithm for statistical learning. *Stata. J.* **2020**, *20*, 3–29. [CrossRef]
51. Little, M.; Varoquaux, G.; Saeb, S.; Lonini, L.; Jayaraman, A.; Mohr, D.; Kording, K. Using and understanding cross-validation strategies. Perspectives on Saeb et al. *GigaScience* **2017**, *6*, gix020. [CrossRef] [PubMed]
52. Tušar, T.; Gantar, K.; Koblar, V.; Ženko, B.; Filipič, B. A study of overfitting in optimization of a manufacturing quality control procedure. *Appl. Soft Comput.* **2017**, *59*, 77–87. [CrossRef]