



Article

Fatigue Damage Monitoring of Composite Structures Based on Lamb Wave Propagation and Multi-Feature Fusion

Feiting Zhang^{1,2}, Kaifu Zhang¹, Hui Cheng¹, Dongyue Gao^{3,4,*} and Keyi Cai⁴

¹ School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China

² Beijing Institute of Astronautical Systems Engineering, Beijing 100076, China

³ School of Fiber Engineering and Equipment Technology, Jiangnan University, Wuxi 214122, China

⁴ School of Aeronautics and Astronautics, Sun Yat-sen University, Guangzhou 510006, China;

caiky3@mail2.sysu.edu.cn

* Correspondence: gaody@jiangnan.edu.cn

Abstract: To address the challenges associated with fatigue damage monitoring in load-bearing composite structures, we developed a method that utilizes Lamb wave propagation and partial least squares regression (PLSR) for effective monitoring. Initially, we extracted diverse characteristics from both the time and frequency domains of the Lamb wave signal to capture the essence of the damage. Subsequently, we constructed a PLSR model, leveraging Lamb wave multi-feature fusion, specifically tailored for in-service fatigue damage monitoring. The efficacy of our proposed approach in quantitatively monitoring fatigue damage was thoroughly validated through rigorous standard fatigue tests. In practical applications, our model effectively mitigated the impact of multicollinearity among feature variables on model accuracy. Furthermore, the PLSR model demonstrated superior accuracy compared to the PCR model, given an equal number of principal components. To strike a harmonious balance between efficiency and precision, we optimized the size of the feature variable. The results show that the optimized PLSR model achieved an R-squared value exceeding 97% in predicting the in-service damage area. This underscores the robustness and reliability of our method in accurately monitoring fatigue damage in load-bearing composite structures.

Keywords: partial least squares regression; Lamb wave; multi-feature fusion; multicollinearity; fatigue damage monitoring



Citation: Zhang, F.; Zhang, K.; Cheng, H.; Gao, D.; Cai, K. Fatigue Damage Monitoring of Composite Structures Based on Lamb Wave Propagation and Multi-Feature Fusion. *J. Compos. Sci.* **2024**, *8*, 423. <https://doi.org/10.3390/jcs8100423>

Academic Editor: Raj Das

Received: 16 August 2024

Revised: 23 September 2024

Accepted: 29 September 2024

Published: 14 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Fiber-reinforced resin matrix composites have gained widespread popularity in aerospace applications due to their exceptional lightness, efficiency, and design flexibility. Initially, these composite components were primarily utilized in non-load-bearing aircraft structures, rendering fatigue damage a less critical concern. However, as the need for weight reduction in aerospace structures has intensified, composites have increasingly been employed in load-bearing roles [1,2]. During the operational lifespan of an aircraft, load-bearing structures are inevitably subjected to prolonged cyclic loading. Consequently, aircraft designers and operators must contend with the potential failures of composite structures in fatigue environments. In contrast to traditional materials, the fatigue damage behavior exhibited by composites is often more diverse and subtle, encompassing various forms of damage such as fiber fracture, matrix cracking, fiber buckling, fiber-matrix debonding, and delamination. To fully harness the benefits of composites in aircraft structures, it is imperative to employ structural health monitoring techniques that can predict damage conditions in real time. Such techniques enable a proactive approach to maintaining aircraft safety and operational efficiency, ultimately contributing to the reliable and sustainable use of composite materials in aerospace applications [3,4].

Damage identification techniques based on Lamb waves have many advantages, such as the ability to propagate over considerable distances; high sensitivity to anomalies and

inhomogeneities near the wave propagation path; and the ability to identify changes in boundary conditions, material properties, or structural geometry by analyzing the scattered wave signals [5]. Therefore, Lamb waves have gained great interest in the identification of fracture and debonding damage in plate, shell, and rod structures [6–9]. For composite structures under fatigue loading, typical damage types include delamination, cracking, and fiber breakage, which will lead to phenomena such as attenuation of wave energy, changes in dispersion characteristics, and conversion of propagation modes [10,11]. Therefore, based on appropriate signal feature extraction and modeling analysis methods, Lamb wave signals can be transformed into structural damage information, including damage location, damage type, and impact range. In recent years, researchers have developed some representative damage probability imaging algorithms in order to identify and image damage by analyzing the changes of Lamb wave signals under healthy and damage conditions. Janarthan B. and Mitra M. [12] described a Lamb wave-based technique where root mean square deviation based damage index (RMSD-DI) was generated by analyzing A0 mode Lamb waves. The ability of DI to predict damage levels was experimentally validated using various specimens with known damage locations. Zhanjun Wu [13] summarized the sensitivity of various damage indices to different typical composite materials damages. From theoretical analysis to sensor design, as well as the physical analysis model for damage monitoring and signal processing, the damage diagnosis of aerospace composite structures based on Lamb waves was explored. Ductho Le [14] described the dispersion behavior of Lamb waves in composites reinforced with unidirectional fibers and obtained dispersion curves. The derived computational results contribute to the development of ultrasound-based techniques for nondestructive evaluation of composite structures. Yan [15] provided a method to detect various damages in composite plates using the Lamb wave tilt effect with several Lamb wave modes selected on the basis of a comb-type transducer. The feasibility of the proposed method was verified by defect detection tests in carbon-fiber-reinforced epoxy resin composite plates. Distinguishing it from the single-feature model used in other studies, Zhao, GQ [16] established a diagnostic method for the delamination of composite double cantilever beams using Lamb wave multi-feature fusion technique. To describe the stratification length, valid linear Lamb wave parameters and nonlinear Lamb wave parameters were retrieved. The results show that the multi-feature damage identification model can effectively distinguish different types of damage.

In recent years, composite damage quantification techniques based on Lamb waves and intelligent models have also attracted the interest of researchers; the research includes signal processing techniques, nonlinear analysis methods, and big data-driven machine learning methods. Tang, JF [17] provided an experimental study using Lamb wave signals to detect and monitor fatigue damage of composite materials under cyclic loading. These Lamb wave signals were characterized by wavelet packet transform (WPT) to extract features and damage indices indicating accumulated internal fatigue damage were calculated. Zhao, JL [18] proposed a linear-nonlinear feature fusion technique for Lamb wave damage detection in composite materials. The phase velocities of Lamb waves in composites were measured using a laser-based generation imaging (LGBI) system. The recorded phase velocities were then inverted using a genetic algorithm (GA) to determine the elastic modulus of the specimen. Yan and Gang [19] provided a statistical multivariate outlier analysis method for determining the presence of fatigue damage and describing its progression using Mahalanobis squared distances. The usability and usefulness of Lamb waves for continuous monitoring of fatigue damage in composite structures was demonstrated.

Indeed, the practical application of composite-load-bearing structures poses significant challenges due to the combined mechanical–thermal and vibration environment they operate in. A key issue is the low signal-to-noise ratio of the monitoring signal, which can make accurate damage detection difficult. When building multi-feature fusion models for fatigue damage monitoring, internal multicollinearity can be a problem. This multicollinearity can lead to noise amplification, thereby compromising the accuracy of the model. To address these challenges, researchers are exploring advanced signal processing techniques and

algorithms that can effectively filter out noise and enhance the signal quality. Additionally, the development of more robust and reliable models that can handle multicollinearity and noise is ongoing. Through these efforts, we aim to improve the accuracy and reliability of fatigue damage monitoring in composite structures, ensuring their safe and efficient operation in complex environments.

This work introduces a method for monitoring fatigue damage in composite structures, leveraging Lamb wave propagation and partial least squares regression (PLSR). The technique addresses the challenges posed by the loading environment, effectively predicting composite damage using a multi-feature model. Notably, the model mitigates the influence of multicollinearity among feature variables on model accuracy, ensuring more reliable predictions. Compared to the principal component regression (PCR) model with the same number of principal components, the PLSR model demonstrates superior accuracy. To strike a balance between efficiency and accuracy, the feature variable size is optimized, resulting in a refined and efficient prediction framework. To validate the overall proposed technique, standardized run-to-failure experiments were conducted on CFRP panels. The results confirm the effectiveness of the method in accurately monitoring fatigue damage in composite materials, providing a robust tool for structural health monitoring and maintenance. This work offers a promising approach for enhancing the safety and reliability of composite structures in various engineering applications.

2. Method

2.1. Lamb Wave Signal Feature Extraction and Fusion

Regression models based solely on single signal features can compromise their stability during the dimensionality reduction phase. This instability can lead to significant parameter fluctuations, even with minor perturbations in the sample data. Therefore, it is crucial to establish multi-signal feature modeling techniques that preserve as much signal matrix information as possible, especially in low signal-to-noise ratio environments. In this study, a comprehensive approach was adopted, incorporating both time domain and frequency domain features to model damage monitoring. Five key features were extracted from the time domain signal: peak value, time of flight (TOF), root mean square, signal standard deviation, and margin index. These features capture diverse aspects of the signal, such as energy propagation efficiency, mode transition effects, signal energy, dispersion degree, and peak extremity.

Simultaneously, three features were selected from the frequency domain: center of gravity frequency, standard deviation, and main lobe energy ratio. These features provide insights into the power spectrum, frequency dispersion, and signal energy distribution within a specific frequency range. By leveraging these combined features, a more robust and comprehensive regression model can be constructed, enhancing the accuracy and stability of damage monitoring in composite structures. Table 1 presents the preliminary selected feature expressions of damage signals, providing a clear overview of the utilized features and their respective expressions. This approach offers a promising framework for improving structural health monitoring in various engineering applications.

2.2. PLSR-Based Damage Monitoring Modeling

Multicollinearity among Lamb wave signal features poses a challenge for damage monitoring modeling. Both principal component regression (PCR) and PLSR can overcome the problem of multicollinearity [20].

Principal component regression (PCR) is a regression method that employs principal components as predictor variables, rather than relying on original characteristics. PCR solely considers the variations exhibited by the predictor variables and reflected through the principal components during their computation, excluding any consideration of response variables. Consequently, a significant drawback of PCR lies in the fact that the principal component exhibiting the greatest variation may fail to accurately predict the response variable. PLSR shares certain similarities with the principal component regres-

sion algorithm; however, the key distinction lies in its approach. Instead of seeking a hyperplane that maximizes variance between the response variable and the independent variable, PLSR constructs a linear regression model by projecting both predictor and observable variables into a new space. This approach potentially yields higher accuracy. In this paper, we compared the predictive capabilities of PLSR and PCR models to assess their relative performance.

Table 1. The time and frequency domain features of the signal representing the damage.

Signal Features	Feature Extraction	Remarks
Time domain signal features:		
Peak of signal	Peak($X_i(t)$)	Maximum of Lamb wave signal (removing crosstalk signal)
Time of flight	TOF($X_i(t)$)	Time of peak
Root mean square	$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N X_i^2(t)}$	
Signal standard deviation	$STD = \sqrt{\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N [X_i(t) - \mu_X(t)]^2}$	
Margin index	$K_f = \frac{x_p}{x_r}$	x_p : peak of signal $x_r = \left(\frac{1}{n} \sum_{i=1}^n \sqrt{ x }\right)^2$: mean square amplitude
Frequency domain signal features:		
Barycenter frequency	$FC = \frac{\int_0^{+\infty} fP(f)df}{\int_0^{+\infty} P(f)df}$	$P(f)$: frequency domain signal
Standard deviation of frequency domain signal	$RVF = \sqrt{\frac{\int_0^{+\infty} (f-FC)^2 P(f)df}{\int_0^{+\infty} P(f)df}}$	
Main lobe signal energy ratio	$SMLE = \frac{\int_{f_2}^{f_1} P(f)df}{\int_0^{+\infty} P(f)df}$	Main lobe: 225 kHz–275 kHz

The damage area was set as the dependent variable Y ; signal features are independent variables x_1, x_2, \dots, x_k ; the sample number is n ; independent variable matrix $X = [x_1, x_2, \dots, x_k]_{n \times k}$; and the dependent variable matrix $Y = [y]_{n \times 1}$.

The X with Y standardized processing obtains independent variable matrix E_0 and the dependent variable matrix F_0 .

The first principal component $t_1 = E_0 w_1$ is extracted from E_0 , where w_1 is the first principal axis of E_0 , i.e., $\|w_1\| = 1$. Both E_0 and F_0 are normalized matrices, then

$$w_1 = \frac{E_0^T F_0}{\|E_0^T F_0\|} = \frac{1}{\sqrt{\sum_{j=1}^k r^2(x_j, y)}} \begin{bmatrix} r(x_1, y) \\ M \\ r(x_k, y) \end{bmatrix} \tag{1}$$

$$t_1 = E_0 w_1 = \frac{1}{\sqrt{\sum_{j=1}^k r^2(x_j, y)}} [r(x_1, y)E_{01} + r(x_2, y)E_{02} + \dots + r(x_k, y)E_{0k}] \tag{2}$$

where $E_{0i} (i = 1, 2, \dots, k)$ represents the i th columns of the independent variable matrix E_0 , and $r(x_j, y) (j = 1, 2, \dots, k)$ represents the correlation coefficient between x_j and y . The regression coefficient p_1 between E_0 and t_1 is

$$E_0 = t_1 p_1^T + E_1 \tag{3}$$

$$p_1 = \frac{E_0^T t_1}{\|t_1\|^2} \tag{4}$$

where p_1 is the regression coefficient and E_1 is the residual matrix.

$$E_1 = E_0 - t_1 p_1^T \tag{5}$$

Repeat the modeling steps above, replacing E_1 with E_0 , replacing F_1 with F_0 , and similarly $w_2 = \frac{E_1^T F_1}{\|E_1^T F_1\|}$; E_1 is no longer a normalized matrix:

$$w_2 = \frac{E_1^T F_1}{\|E_1^T F_1\|} = \frac{1}{\sqrt{\sum_{j=1}^p cov^2(E_{1j}, y)}} \begin{bmatrix} cov(E_{1j}, y) \\ M \\ cov(E_{1p}, y) \end{bmatrix} \tag{6}$$

$$t_2 = E_1 w_2 \tag{7}$$

$$p_2 = \frac{E_1^T t_2}{\|t_2\|^2} \tag{8}$$

$$E_2 = E_1 - t_2 p_2^T \tag{9}$$

where $cov(E_{1j}, y)$ is covariance between E_{1j} and y .

The number of components extracted in PLSR can be determined by cross validity.

In implementing the standard of the dependent variable matrix F_0 to principal component t_1, t_2, \dots, t_m , the regression of

$$F_0 = r_1 t_1 + r_2 t_2 + \dots + r_m t_m \tag{10}$$

Because the principal component t_1, t_2, \dots, t_m is a linear combination of E_0 ,

$$t_h = E_{h-1} w_h = E_0 \prod_{j=1}^{h-1} (I - w_j p_j^T) w_h = E_0 w_h^* \tag{11}$$

where $w_h^* = \prod_{j=1}^{h-1} (I - w_j p_j^T) w_h$.

$$F_0 = r_1 E_0 w_1^* + \dots + r_m E_0 w_m^* = E_0 (r_1 w_1^* + \dots + r_m w_m^*) \tag{12}$$

$x_j^* = E_0 w_j^*$, $y^* = F_0$, $\alpha_j = \sum_{h=1}^m r_h w_{hj}^*$ ($j = 1, 2, \dots, k$). The standardized variable y^* about x_j^* for the regression equation is

$$\hat{y}^* = \alpha_1 x_1^* + \alpha_2 x_2^* + \dots + \alpha_k x_k^* \tag{13}$$

Finally, the regression equation of y with respect to x_j is obtained by the inverse:

$$\hat{y} = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \tag{14}$$

where β_i ($i = 1, 2, \dots, k$) is the regression coefficient of y with respect to x_j .

2.3. Damage Prediction Framework Based on PLSR

In this study, PLSR was used to predict structural damage from Lamb wave signals. As shown in Figure 1, The model was built and validated through a series of structural fatigue test procedures to fill the required Lamb wave signal space. The Lamb wave signal and X-ray images of the structure were measured at each loading step. The signal characteristics (autocovariance matrix $X = [x_1, x_2, \dots, x_k]_{n \times k}$) and damage condition (dependent variable matrix $Y = [y]_{n \times 1}$) were extracted. The obtained dataset was separated into calibration data and a smaller subset of validation data and used to build and validate the PLSR model and accompanying R scripts described in detail in the text. Once the datasets were separated, the optimal number of components was determined, and the calibration model

and validation were built. Twenty percent of the remaining dataset was used as the cross-validation dataset for model checking. VIP (variable impact on prediction) was used to measure the importance of a given predictor on the predicted response variable. In this paper, $VIP \geq 1$ was considered as the critical value for highly significant predictor variables.

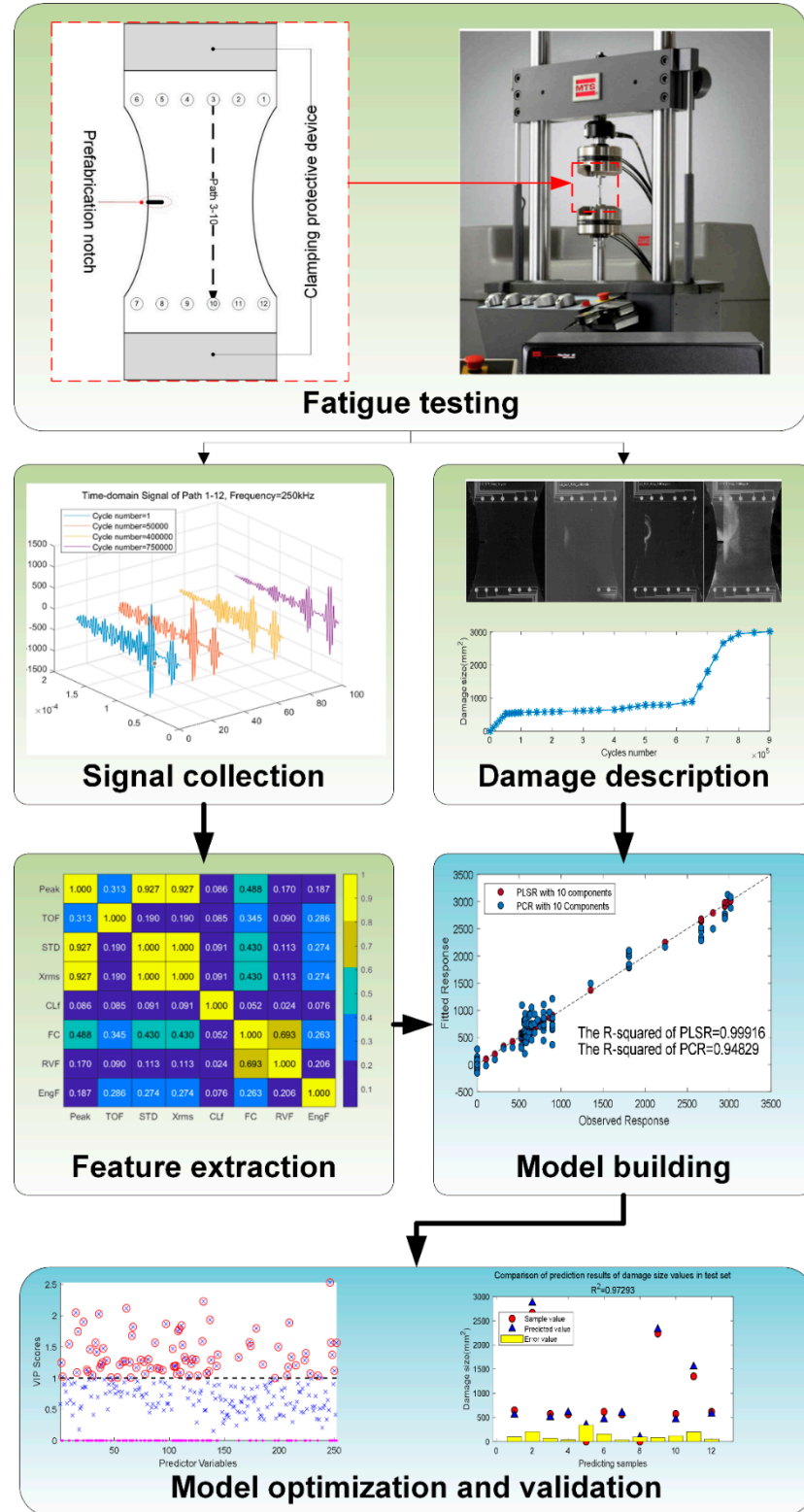


Figure 1. Flowchart of structural fatigue damage prediction based on Lamb wave signal and PLSR.

3. Experiment

The data used in this study came from a public dataset of run-to-failure experiments on CFRP panels conducted at the Stanford Structures and Composites Laboratory (SACL) in collaboration with the NASA Ames Research Center Prognostic Center of Excellence (PCoE). In the experiments, internal damage growth under tension–tension fatigue were captured by periodic measurements. The monitoring data consist of Lamb wave signals from a network of piezoelectric (lead zirconate titanate—PZT) sensors and multiple triaxial strain gages. Additionally, periodic X-rays were taken to characterize internal damage as ground truth information. Three different layups were tested. In this experiment, a set of specimens were performed to a tension–tension fatigue test at a frequency of 50 [Hz] and a stress ratio of $R \approx 0.14$. Torayca T700G unidirectional carbon-prepreg material was used to create stress concentration in 15.24 [cm] \times 25.4 [cm] specimens with dog-bone geometry and a prefabrication notch (5.08 [mm] \times 19.3 [mm]), as shown in Figure 2. In order to protect the clamping position of both ends, the surface of the specimen was pasted with a glass fiber composite plate, as shown in Figure 2. For this investigation, the No. 17 specimen of Layup2 (layup configurations: $[0/90_2/45/45/90]_s$) was used in this paper [21].

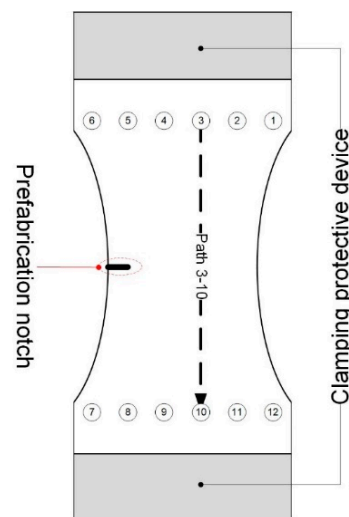


Figure 2. Schematic diagram of the fatigue test specimen and sensor network.

To quantify fatigue damage, two groups of six SMART Layer[®] sensors were arranged on the specimen surface, as shown in Figure 2. One group of SMART Layer[®] was used to excite the Lamb waves, and the other group was responsible for receiving and monitoring the Lamb wave propagation behavior in the specimen through the signal. The sensor numbers on the upper part of the specimen were 6#, 5#, 4#, 3#, 2#, and 1# from left to right, and the sensor numbers on the lower part were 7#, 8#, 9#, 10#, 11#, and 12#. The average input voltage and gain for each of the 36 actuator-sensor paths were 50 volts and 20 dB, respectively.

All testing procedures were conducted utilizing the MTS machine in accordance with ASTM standards D3039 and D3479. During fatigue cycling experiments, the PZT sensor data were collected for all paths and excitation frequencies at intervals of every 50,000 cycles. Additionally, stain X-rays were captured from the specimens to enhance the X-ray absorption quality.

The fatigue data of the specimens were gathered under three distinct boundary conditions: type 1, where the specimens were loaded with an average load; type 2, where the specimens were unloaded but clamped in place; and type 3, where the specimens were removed entirely from the test machine, resulting in an absolute zero load. At a load of 4 Kips, the documentation noted a decline in the signal in the notched region, yet no audible sound was reported. Concurrently, the documentation indicated “No audible sounds of

matrix cracking were heard; X-ray inspection will follow". The subsequent X-ray report stated "Minimal damage observed in X-rays; will subject to fatigue with 7 Kips". This suggests that the ultrasonic wave had already emitted a precursor warning signal before the damage became detectable through other means. After enduring 10,000 fatigue cycles at a load of 4 Kips, the signal exhibited a universal decline, with the X-ray inspection revealing significant delamination damage around the notch. As the cycle count reached 50,000, the documentation described a shift and reduction in the signals near the notched area, indicating that the time-domain amplitude, time of flight (TOF), and energy of the Lamb wave signal serve as tangible indicators of damage progression. In various states such as "Loaded, Clamped, and Traction Free", the Lamb wave signal characteristics may exhibit slight variations due to factors like the vibration of the testing apparatus and the crack breathing effect induced by loading. During the model training phase, these subtle differences can be magnified by the presence of multicollinearity, making feature optimization and the mitigation of multicollinearity pivotal in the practical deployment of Lamb wave-based damage prediction algorithms. Following 750,000 cycles of loading, the sensor experienced extensive failure, with the record stating "No Signal Anywhere". The accompanying figure illustrates the X-ray transmission images depicting the specimen's damage progression through different cycles [21].

It can be seen from the Figure 3 that the damage extended along the tectonic fracture direction from Path 6#-7# to Path 3#-10# to Path 1#-12# as the number of cycles increased. The damage forms were diverse, including delamination, cracking, and fiber fracture. When the number of cycles approached 750 K, the specimens showed longitudinal penetration damage and failed completely. According to the experimental description, the damage extension trend with increasing cycle times is shown in Figure 4.

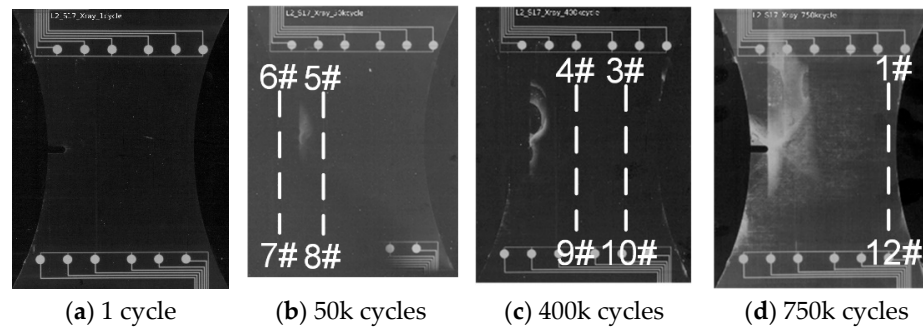


Figure 3. Test results of the test X-ray under different cycles.

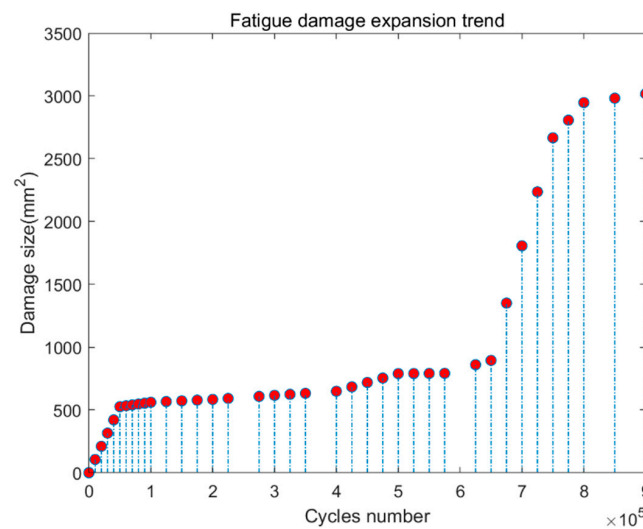


Figure 4. Fatigue damage expansion trend of a typical specimen.

As shown in Figure 4, the fatigue damage showed a stepwise trend with increasing number of cycles, with two damage step changes after 0.5×10^5 and 6×10^5 cycles. Figure 5 shows the signals in the time and frequency domains for the representative paths at different cycles.

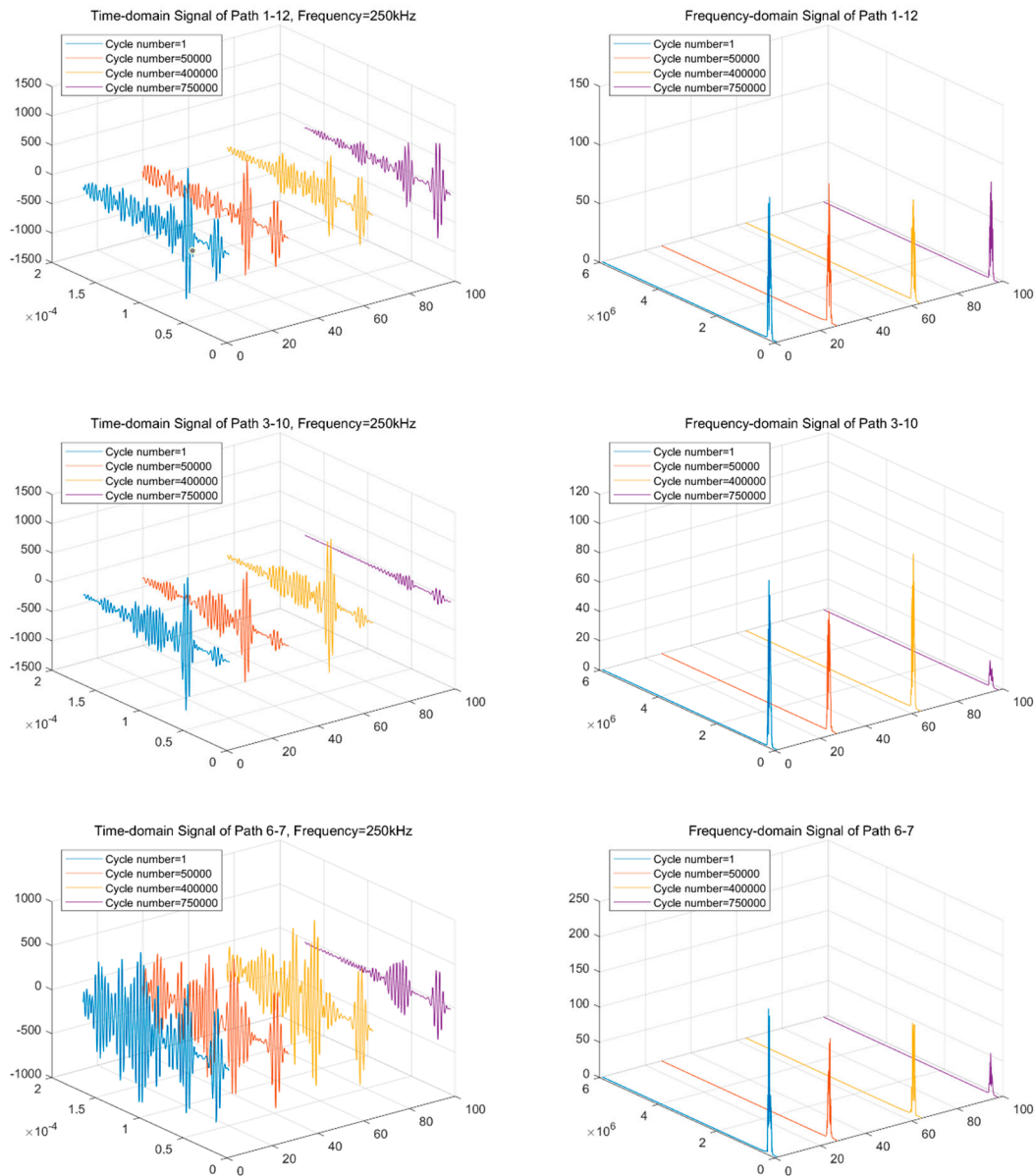


Figure 5. Time domain signal and frequency domain signal of representative paths under different cycle loading numbers.

As the fatigue damage progressed, the Lamb wave propagation environment underwent gradual alterations. According to the measurements from Figure 5, the amplitude and energy of the path signals proximate to the damage extension boundary exhibited a decline over time. Specifically, upon the occurrence of penetrating damage, paths 6#-7# and 3#-10# within the damaged region exhibited modal shifts. Conversely, paths 1#-12#, which were unaffected by the damage, remained relatively unaffected.

In the frequency domain, the signal energy of the primary flap underwent changes as the damage level varied. At the periphery of the damage extension, the signal’s main peak value increased. However, as the damage area expanded along the line, the peak value of the signal decreased. It is evident that a singular signal feature alone is insufficient to

explain the signal variations across diverse damage scenarios. Instead, the damage can only be accurately characterized through the development of a model that incorporates numerous signal features. Nevertheless, not every segment of every signal path was damage-related, necessitating the optimization of features to establish an effective model for damage prediction.

When considering the correlation coefficient, a value below 0.3 suggests no linear correlation, while a value above 0.3 indicates a linear relationship. Specifically, a coefficient between 0.3 and 0.5 represents a low correlation, 0.5 to 0.8 signifies a significant correlation, and above 0.8 denotes a high correlation. As is evident in Figure 6, the correlation between features across numerous paths often exceeded 0.5 and, in some cases, even reached 0.8, indicating strong multicollinearity among these signal features. Furthermore, the correlation between STD and XRMS was found to be perfect at 1, indicating that they are redundant features. Therefore, in the modeling process, only XRMS was selected to represent the energy characteristics of the signals, eliminating the need for both STD and XRMS to avoid redundancy.

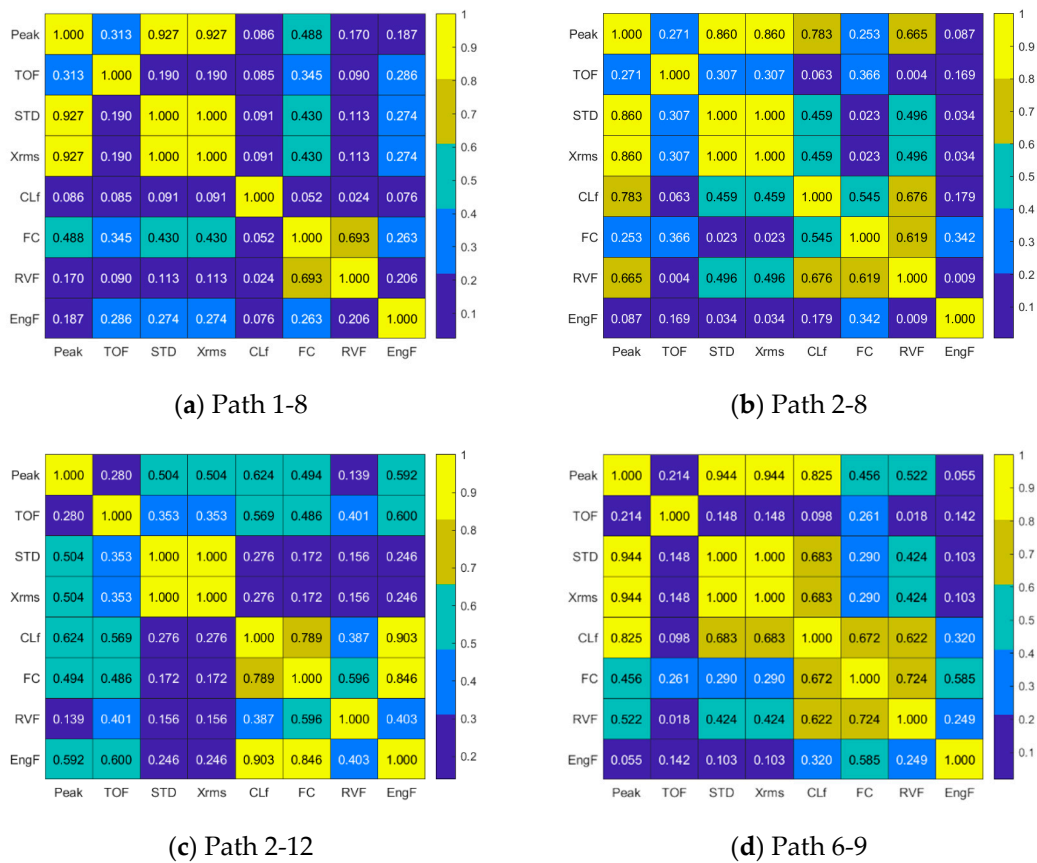


Figure 6. Multi-collinearity matrix of typical path signal features.

4. Result and Discussion

Using PLSR and principal component regression (PCR) models, we examined the effectiveness and precision of monitoring fatigue damage dimensions in specimens under fatigue loading conditions. For this investigation, we specifically utilized the dataset from the seventeenth specimen. Based on previous research findings, we extracted various features including peak values, time of flight (TOF), XRMS, edge index, gravity frequency, standard deviation ratio (SDR) of the frequency domain signal, and the main flap energy ratio within the range of 225 kHz to 275 kHz. These features were then compiled into a matrix of independent variables. To ensure consistency and comparability across the modeling process, each eigenvalue was normalized.

The data file of this specimen had 112 sets of test data, and 4 sets of faulty data were excluded according to the test records. Among the 108 sets of valid data, 96 sets were randomly selected as the training group, and the rest of the data were used as the test group. The contribution of principal components is shown in Figure 7.

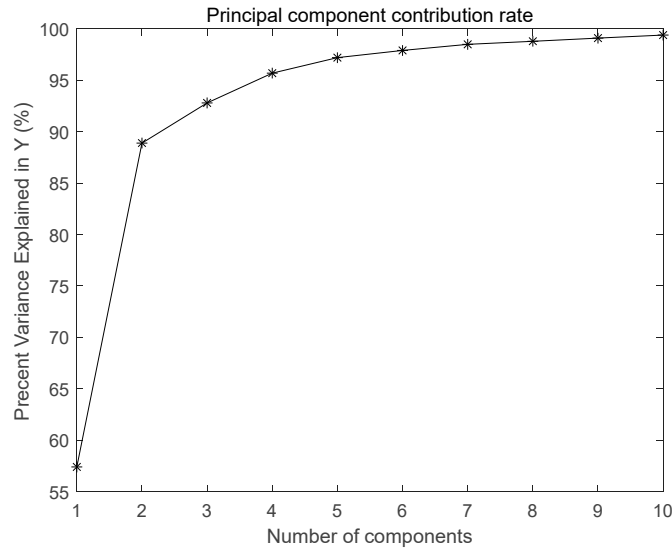


Figure 7. Principal component contribution rate.

PLSR analysis with five components appeared to capture a significant proportion of the variance within the observed damage area data. This conclusion was drawn from the determination of the response values associated with the chosen five-component model. To further validate the effectiveness of this model, Figure 8 provides visual confirmation through observed–fit response plots, which compare the predicted value variables with a VIP score close against the actual observed damage areas. Additionally, a similar comparison was made with a principal component regression (PCR) model utilizing five components. These plots allow for a visual assessment of the accuracy and reliability of both models in replicating the observed data patterns, thereby providing insights into their predictive capabilities.

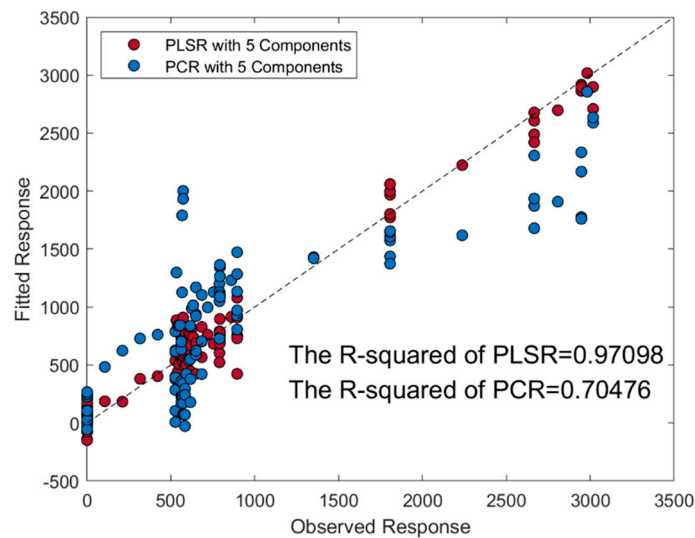


Figure 8. Observed–fitted response maps of the PLSR and PCR models (five components).

As depicted in Figure 8, the observed–fit response plots for both the PLSR and the five-component PCR model are presented. The blue dots represent the PLSR model, while the

red dots represent the five-component PCR model. The R-squared value for the PLSR model was 0.97, indicating a strong correlation between the observed and predicted values. In contrast, the R-squared value for the five-component PCR model was 0.70, which suggests a less robust fit. With the same number of components, it is evident that the PLSR model provided a superior regression compared to the five-component PCR model. In fact, the five-component PCR model performed barely better than a constant model, indicating that it may not be capturing the underlying structure of the data effectively. The R-squared values of the two regressions confirm this observation. The higher R-squared value for PLSR suggests that it is able to explain a larger proportion of the variation in the observed damage area matrix, making it a more suitable model for monitoring fatigue damage dimensions under fatigue loading conditions.

As shown in Figure 9, the observation that the PCR curves were consistently higher indicates that, in terms of fit, the multicomponent PCR performed at an inferior level to PLSR. As the number of components increased in both PCR and PLSR models, a tighter fit to the original data was inevitable since the majority of predictive information contained within the independent variable matrix was captured by the principal components. This trend is clearly exemplified in the figure below, which demonstrates that the residuals of both methods decrease significantly when using 10 components compared to when only 5 components are employed.

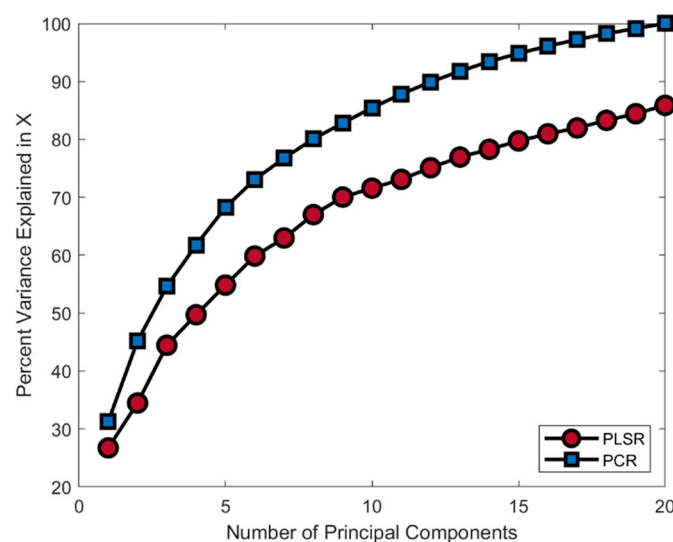


Figure 9. Percentage variance explained of PLSR and PCR models under different numbers of components.

However, it is important to note that while adding more components may lead to a better fit, it can also increase the complexity of the model and potentially lead to overfitting. Therefore, it is crucial to strike a balance between model complexity and predictive accuracy. In this context, PLSR often offers a more robust and interpretable solution by effectively combining features from the independent variable matrix while minimizing the number of components required to achieve a good fit.

According to Figure 10, the R-squared value for PLSR stood at an impressive 0.99. This metric indicates that PLSR provides a much closer fit to the observed data, suggesting its superiority in capturing the underlying relationships. To further validate the model's damage prediction capabilities, we tested it using an independent test set of data, as shown in the accompanying figure. This step is crucial in assessing the model's generalization ability—its capacity to predict outcomes on unseen data. The test set data offer a realistic evaluation of the model's performance in real-world scenarios.

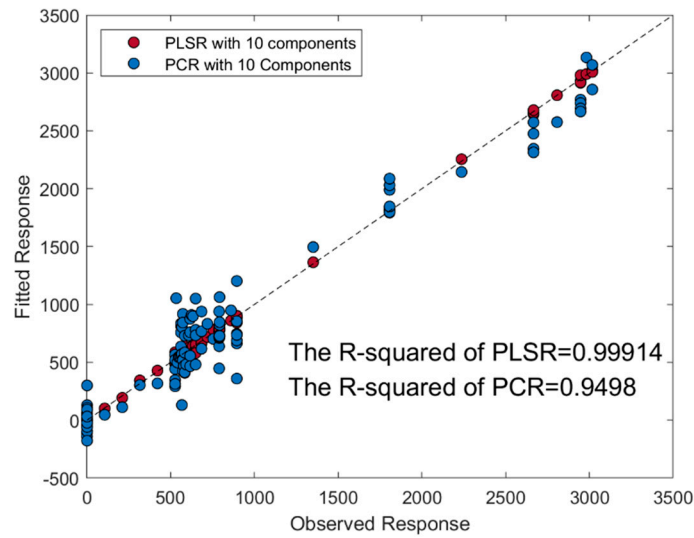


Figure 10. Observed–fitted response maps of the PLSR and PCR models (ten components).

If the model’s predictions on the test set align closely with the actual damage outcomes, it would further confirm the predictive power of PLSR in damage prediction tasks. Such validation is essential in ensuring the reliability and accuracy of the model in practical applications.

As shown in Figure 11, the red dots signify the actual damage magnitude recorded in the random test dataset. In contrast, the blue triangles represent the damage magnitude predicted by the signal features extracted from the same dataset. Furthermore, the yellow bars illustrate the discrepancy between the observed and predicted values, providing a visual representation of the prediction errors.

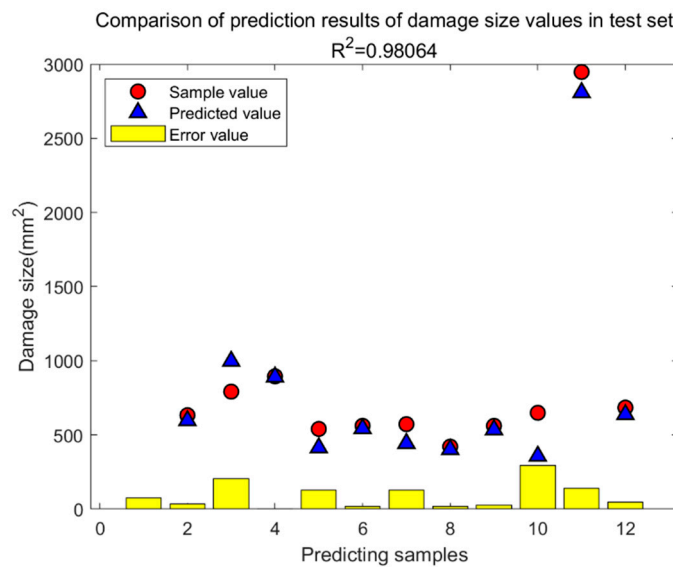


Figure 11. Comparison of prediction results of damage size values in the test set.

For this analysis, the final 10 components were chosen for making predictions. The results revealed that the prediction model achieved an R-squared value of 98% for damage magnitude, indicating a strong correlation between the predicted and actual values. This high R-squared score demonstrates the model’s excellent predictive capabilities and its ability to accurately estimate damage magnitude based on the signal features present in the dataset.

To improve the efficiency of the model run, the signal feature matrix was optimized by VIP scoring.

In order to improve the efficiency of the model operation, the signal feature matrix was optimized by VIP scores.

$$VIP_j = \sqrt{\frac{\sum_{a=1}^h R^2(y, t_a) (w_{aj} / \|w_a\|)^2}{(1/p) \sum_{a=1}^h R^2(y, t_a)}} \tag{15}$$

where w_{aj} is weight of the j th feature variable in component a and $R^2(y, t_a)$ is a fraction of variance in y explained by the component a . VIP scores estimate the importance of each variable in the projection used in a PLS model and are often used for variable selection. A variable with a higher VIP score indicates that it is more correlated with the predicted response. A VIP score close to or greater than 1 is considered important in the given model; otherwise, it will be considered less important and may be excluded from the model of good candidates. Figure 12 shows the VIP scores of the signal features in the PLSR regression model.

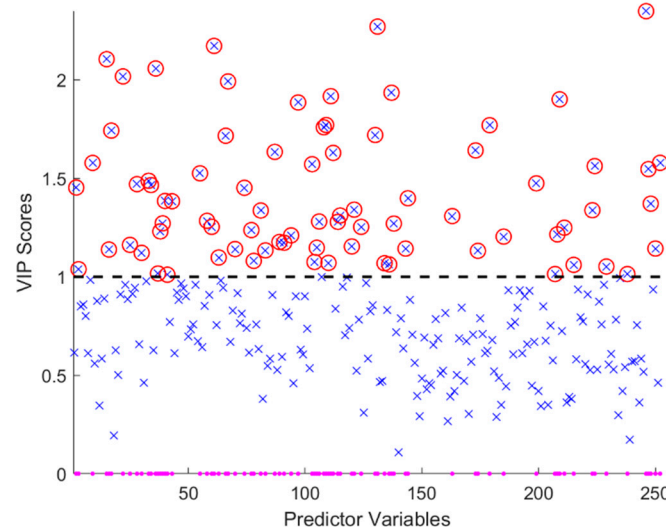


Figure 12. VIP scores of signal features for the PLSR regression model (the × symbol represents the VIP score of the signal feature, and the circle represents the feature whose VIP score is close to or greater than 1, that is, the selected feature value).

The selection of features for the prediction model was carefully optimized to guarantee the validity and accuracy of the predictions. As part of this process, VIP (importance of variables in prediction) scores were computed specifically for the PLSR regression model. These scores provide a quantitative measure of the significance of each feature in contributing to the model’s predictive power.

Out of the 86 features evaluated, those with VIP scores exceeding 1 were deemed crucial and are highlighted with red circles. This threshold ensures that only the most influential variables are retained, enhancing the model’s predictive performance. In the context of multicollinearity, where variables may exhibit strong correlations, the selection of these high-VIP variables is particularly important. It helps to identify and prioritize those predictor variables that contribute most significantly to the model, while minimizing the impact of redundant or less informative features. By focusing on these key features with VIP scores greater than 1, the PLSR regression model is able to make more accurate predictions. This optimized feature selection not only improves the model’s predictive accuracy but also enhances its interpretability, making it easier to understand the relationships between the selected features and the predicted outcome.

As shown in Figure 13, the test results revealed an impressive accuracy level of 97% in quantifying fatigue damage through the application of the optimized model. This significant figure underscores the reliability and precision of the proposed method in accurately assessing the condition of composite materials, offering a robust tool for structural reliability analysis.

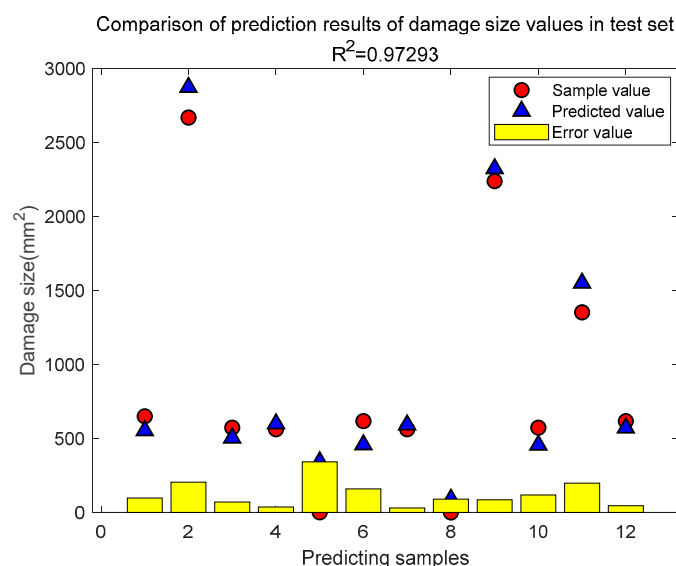


Figure 13. Comparison of the optimization model prediction results of damage size values.

5. Conclusions

This paper introduces an innovative fatigue damage monitoring method specifically tailored for composite materials. The core of this approach lies in the utilization of guided wave propagation, in conjunction with PLSR. The Lamb wave propagation technique emerges as a highly effective tool for internally characterizing fatigue damage within composite materials. Leveraging NASA's publicly available metamaterial fatigue test database, a quantitative fatigue damage monitoring model was crafted, incorporating multi-feature fusion and PLSR.

Upon analyzing the experimental results, the following key insights were derived:

- A significant multicollinearity exists among various signal features, with the correlation between multiple path features exceeding 0.5 and even reaching 0.8.
- In terms of fitting accuracy, the multi-component PLSR (99%) outperformed PCR (94%), demonstrating its superiority in capturing the intricate relationships within the data.
- Through cross-validation, the 10-component model achieved an optimal balance between efficiency and accuracy, ensuring both speedy computations and reliable predictions.
- After optimizing the variable importance projection, 86 signal features remained relevant, while the model's R-squared value remained consistently above 97%, a testament to its predictive power.

The proposed method stands out for its baseline-free nature, its disregard for the potential influence of diagnostic techniques or ambient temperature, and its suitability for practical working conditions. As such, it serves as a solid foundation for structural reliability analysis, offering new avenues for enhancing the safety and durability of composite materials in various applications.

Future research needs to address such problems as damage identification in noisy environments, complex damage scenarios, increasing computational requirements with the increase of model complexity, and the problem that the effectiveness of injury diagnosis depends on the quality and location of the sensor.

Author Contributions: Conceptualization, F.Z.; Methodology, F.Z. and D.G.; Validation, K.Z.; Formal analysis, H.C.; Investigation, K.Z.; Data curation, F.Z. and K.C.; Writing—original draft, H.C.; Writing—review & editing, D.G.; Funding acquisition, D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Key Program of National Natural Science Foundation of China (U2341235), the China National Building Material Group special funding project to tackle key problems (2023SJYL01), the National Key R&D Program of China (2018YFA0702800), and the Jiangnan University Basic Research Program youth fund (JUSRP123003).

Data Availability Statement: The original data presented in the study are openly available in <https://phm-datasets.s3.amazonaws.com/NASA/2.+Composites.zip>.

Acknowledgments: The authors would like to thank the NASA Ames' Center for Predictive Excellence and Stanford University's Structures and Composites Laboratory for providing the experimental data that partially supported this work.

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

References

- Govindasamy, M.; Kamalakannan, G.; Kesavan, C.; Meenashisundaram, G.K. Damage detection in glass/epoxy laminated composite plates using modal curvature for structural health monitoring applications. *J. Compos. Sci.* **2020**, *4*, 185. [[CrossRef](#)]
- Demo, L.B.; Tronci, E.M.; Feng, M.Q. Multifunctional fiber-reinforced polymer composites for damage detection and memory. *J. Compos. Sci.* **2023**, *7*, 383. [[CrossRef](#)]
- Caldwell, S.P.; Radford, D.W. Composite single lap shear joint integrity monitoring via embedded electromechanical impedance sensors. *J. Compos. Sci.* **2023**, *7*, 53. [[CrossRef](#)]
- Cilento, F.; Curcio, C.; Martone, A.; Liseno, A.; Capozzoli, A.; Giordano, M. Effect of graphite nanoplatelets content and distribution on the electromagnetic shielding attenuation mechanisms in 2D nanocomposites. *J. Compos. Sci.* **2022**, *6*, 257. [[CrossRef](#)]
- Yang, Z.L.; Elgamal, H.M.; Zhang, J.W. Lamb-Wave-Based Damage Identification in Laminated Composite Plates. *Adv. Mater. Res.* **2014**, *1014*, 3–8. [[CrossRef](#)]
- Moll, J.; Kexel, C.; Kathol, J.; Fritzen, C.P.; Moix-Bonet, M.; Willberg, C.; Herrmann, A. Guided Waves for Damage Detection in Complex Composite Structures: The Influence of Omega Stringer and Different Reference Damage Size. *Appl. Sci.* **2020**, *10*, 3068. [[CrossRef](#)]
- Barouni, A.K.; Rekatsinas, C.S. Study on the propagation of stress waves in natural fiber composite strips. *J. Compos. Sci.* **2021**, *5*, 34. [[CrossRef](#)]
- Liu, H.; Huang, Z.; Yin, Z.; Sun, M.; Bo, L.; Li, T.; Tian, Z. Investigation of Viscoelastic Guided Wave Properties in Anisotropic Laminated Composites Using a Legendre Orthogonal Polynomials Expansion—Assisted Viscoelastodynamic Model. *Polymers* **2024**, *16*, 1638. [[CrossRef](#)] [[PubMed](#)]
- Mei, H.; Haider, M.F.; James, R.; Giurgiutiu, V. Pure S0 and SH0 detections of various damage types in aerospace composites. *Compos. Part B Eng.* **2020**, *189*, 107906. [[CrossRef](#)]
- Wang, Y.; He, M.; Sun, L.; Wu, D.; Wang, Y.; Qing, X. Weighted adaptive Kalman filtering-based diverse information fusion for hole edge crack monitoring. *Mech. Syst. Signal Process.* **2022**, *167 Pt A*, 108534. [[CrossRef](#)]
- Wang, Y.; Gao, T.; Liu, D.; Sun, H.; Miao, B.; Qing, X. Propagation characteristics of ultrasonic weld-guided waves in Friction stir welding joint of same material. *Ultrasonics* **2020**, *102*, 106058. [[CrossRef](#)] [[PubMed](#)]
- Janarthan, B.; Mitra, M.; Mujumdar, P.M. Lamb Wave Based Damage Detection in Composite Panel. *J. Indian Inst. Sci.* **2013**, *93*, 715–734.
- Wu, Z.; Liu, K.; Wang, Y.; Zheng, Y. Validation and evaluation of damage identification using probability-based diagnostic imaging on a stiffened composite panel. *J. Intell. Mater. Syst. Struct.* **2014**, *26*, 2181–2195. [[CrossRef](#)]
- Le, D.; Lee, J.; Cho, Y.; Dao, D.K.; Nguyen, T.G.; Phan, H. Ultrasonic guided waves in unidirectional fiber-reinforced composite plates. In *Advances in Condition Monitoring and Structural Health Monitoring*; Springer: Singapore, 2021.
- Yan, F.; Rose, J.L. Composite Plate Inspection Using a Novel Guided Wave Skew Effect Method. *AIP Conf. Proc.* **2009**, *1096*, 1049–1056.
- Zhao, G.; Wang, B.; Wang, T.; Hao, W.; Luo, Y. Detection and monitoring of delamination in composite laminates using ultrasonic guided wave. *Compos. Struct.* **2019**, *225*, 111161. [[CrossRef](#)]
- Tang, J.F.; Lu, X.; Yan, G. Monitoring the Evolution of Fatigue Damage for Composite Using Guided Waves. In *Proceedings of the 25th Australasian Conference on Mechanics of Structures and Materials, Brisbane, Australia, December 2018*; Springer: Singapore, 2020; pp. 385–393.
- Zhao, J.; Chillara, V.; Cho, H.; Qiu, J.; Lissenden, C. Evaluation of fatigue damage accumulation in composites via linear and nonlinear guided wave methods. *AIP Conf. Proc.* **2016**, *1706*, 120007.

19. Yan, G.; Lu, X.; Tang, J. Guided Wave-Based Monitoring of Evolution of Fatigue Damage in Glass Fiber/Epoxy Composites. *Appl. Sci.* **2019**, *9*, 1394. [[CrossRef](#)]
20. Dodge, Y.; Daniel, C. *The Oxford Dictionary of Statistical Terms*; Oxford University Press on Demand: Oxford, UK, 2006.
21. Saxena, A.; Goebel, K.; Larrosa, C.C.; Chang, F.K. *CFRP Composites Data Set*; NASA Ames Research Center: Moffett Field, CA, USA.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.