Machine Learning Algorithms for Identification of Cohesive Zone Parameters for Mixed-Mode Fracture in Composite Sandwich Structure

Arash Ramian, Rani F Elhajjar*

Department of Civil & Environmental Engineering, 3200 North Cramer St., The College of Engineering & Applied Science, University of Wisconsin, Milwaukee, WI 53211, USA

Abstract

We investigate three machine-learning algorithms to identify the cohesive zone parameters used in the fracture analysis of a honeycomb/carbon-epoxy sandwich structure. Determining the cohesive parameters from experimental results can be complicated because multiple parameters can fit the experimental results equally well, resulting in non-unique solutions that do not capture the correct values. Numerically, determining the cohesive zone parameters can be time-consuming due to modeling challenges, such as fine meshes needed near the crack tip. In this study, we apply three machine learning algorithms, namely, Support Vector Regression (SVR), Random Forest (RF), and Artificial Neural Network (ANN), to identify the interfacial cohesive parameters between the honeycomb core and carbon-epoxy facesheets. Our study represents the interfacial fracture using a cohesive zone model in a finite element (FE) simulation of the Asymmetric Double Cantilever Beam (ADCB) specimen configuration. The input variables include four cohesive zone parameters: the maximum normal contact stress, critical fracture energy for normal separation, maximum equivalent tangential contact stress, and critical fracture energy for tangential slip. FE simulations with ranges of cohesive model parameters produce a database of load-displacement responses that serve as training data for the machine learning algorithms. The results show that the suggested methods demonstrate remarkable accuracy for situations where the interfacial characteristics lie within the training dataset or beyond and reveal the importance of specific

parameters. Additionally, we perform hyperparameter tuning and optimization techniques to enhance the model's performance.

Keywords: Cohesive Zone Model, Composite Sandwich Structure, Machine Learning Algorithms, Asymmetric Double Cantilever Beam

Corresponding author: Rani F Elhajjar, Email: elhajjar@uwm.edu

Nomenclature

a_0	Initial crack
a_i^{in}, a_i^{out}	Activation Units
b	Bias Parameter
C	Penalty factor
d	Damage Parameter
f	Activation Function
k	Interfacial Stiffness
G_{II}	Critical Fracture Energy for Tangential Slip
n	Number of Data Points
L	Length of sandwich structure
h_c	Height of core
G_I	Critical Fracture Energy for Normal Separation
N	Number of Trees
h_f	Height of facesheet
ξ_i, ξ_i^*	Slack variables
G_{Tc}	Critical Total Fracture Energy
G_{Td}	Actual Dissipated Energy
φ_x	Mapping Function
w	Weight vector
x_i	Input Value
y	Predicted Value
$\bar{y}$	Mean of The Actual Values
δ^*	Displacement at the Maximum Load
δ_m	Equivalent Relative Displacement
σ_m	Resulting Mixed-Mode Traction
σ_I	Maximum Normal Contact Stress
σ_{II}	Maximum Equivalent Tangential Contact Stress
$\sigma_{u,I}$	Local Strength for Mode I
$\sigma_{u,II}$	Local Strength for Mode II

Abbreviation

ADCB	Asymmetric Double Cantilever Beam
ANN	Artificial Neural Network
CFRP	Carbon Fiber Reinforced Plastic
CZM	Cohesive Zone Model
FE	Finite Element
ML	Machine Learning
MAPE	Mean Absolute Percentage Error
RF	Random Forest
RMSE	Root Mean Square Error
SVM	Support Vector Machine
SVR	Support Vector Regression

1. Introduction

Composite sandwich structures have become essential in structural engineering for applications where stiffness, strength, and weight optimization are crucial considerations. These adaptable structures, which offer compelling advantages in terms of efficiency and reduced environmental impact, find application in various industries, including wind energy, automotive, aerospace, and marine industries [1, 2]. However, despite their outstanding performance, they face a significant obstacle: the complex interaction between the facesheet and core materials and debonding failure modes. Debonding at this interface can significantly affect structural integrity, potentially jeopardizing performance and safety [3]. Researchers have proposed advanced modeling techniques to comprehend and reduce facesheet to core debonding in composite sandwich structures. Of these methods, the Cohesive Zone Model (CZM) has found wide application in the field. When simulating fracture processes numerically, CZMs are essential because they provide a framework for understanding the complex interactions between the different materials [4]. Fundamentally, the CZM identifies surfaces of discontinuity where abrupt changes in displacement occur and relates these changes to appropriate tractions via traction-separation laws. In addition to simplifying the examination of the bond-slip relationship comprehensively, this method also naturally considers debonding factors like crack initiation, propagation, and arrest. CZM is adaptable in tackling a broad range of fracture-related problems and has practical applications in different fields, for instance, delamination in composite structures [5-7], adhesively bonded joints [8], and peeling [9].

CZMs using different constitutive laws [10-14] have been proposed for various material systems. These CZMs offer a valuable toolbox for examining the fracture process in multiple materials because of their distinctive shapes and properties. In the context of composite structures, Akhmet

et al. [15] utilized a mixed-mode bilinear CZM to study the stress distribution behavior of corrugated sandwich structures subjected to a three-point bending load. Airoldi et al. [16] applied a trilinear CZM to analyze composite structures' delamination and crack advancement using a double cantilever beam test. Han et al. [17] employed a mixed-mode exponential CZM to investigate debonding at the facesheet and core interfaces. Additionally, the authors examined the effects of various loading conditions. In CZM, specific parameters are essential in determining the cohesive law, and these parameters can be found in multiple ways. Researchers have recently proposed different techniques to assess traction-separation characteristics and enable numerical simulations to predict cohesive failure. Several studies in the literature have analyzed CZM parameters analytically.

By considering the elastic behavior in the adhesive, the double cantilever beam test was proposed by Jumel et al. [18]. They calculated the double cantilever beam specimen's stiffness and critical energy release rate using this modified model. In addition, based on the beam on elastic foundation theory and assuming the hyperelastic behavior of the adhesive layer, Cabello et al. [19] proposed a model to investigate the mechanical responses of a double cantilever beam specimen with a thick, flexible adhesive. Cabello et al. [20] proposed an analytical approach based on the elastic foundation approach to predict the fracture behavior of the double cantilever beam specimen. Simplifying assumptions for some analytical techniques may lead to inaccuracies when dealing with complex geometry, such as sandwich structures with nonlinear materials or intricate failure modes [21]. Inverse identification is another approach for determining CZM parameters. This approach has attracted the attention of numerous researchers. Valoroso et al. [22] researched an inverse method for determining CZM parameters using the DCB test by analyzing the outcomes from various measurement approaches. Jaillon et al. [23] used an indirect method with uncertainty

estimation of the CZM constants. Xu et al. [24] identified these mixed-mode fracture parameters using the inverse process. In this investigation, the authors used simulations and experiments to examine both the load-displacement curves and the crack path. Shen and Paulino [25] investigated using the full-field displacement and an optimization strategy within an FE framework to identify the elastic characteristics and CZM parameters.

Machine learning (ML) algorithms have become an increasingly effective tool for obtaining valuable data from large datasets, rapidly increasing their successful use in composites [26-30]. More specifically, in the area of composite sandwich structures, many studies have been conducted on different aspects, such as delamination [31, 32], damage detection [33], and vibrations [34, 35]. ML can anticipate and identify new insights in the characterization of composites using the complex relationship between input and output data. For example, Sikdar et al. [36] investigated the identification and categorization of the damage for carbon-fiber reinforced composite panels using neural networks and deep learning. The authors proposed a deep learning architecture based on neural networks to automatically extract damage characteristics from images and utilize them to categorize damage-source locations in the composite panel. Viotti and Gomes [31] studied delamination identification for composite sandwich structures using artificial neural networks, random forests, decision trees, and support vector machine approaches. The researchers used algorithms to detect, locate, and size damages in the structure's core and facesheets. By applying artificial neural networks, Su et al. [37] researched CZM parameters for composites bonded to concrete using ANN to identify two parameters on the CZM model.

In this study, we investigate the identification of CZM parameters for mixed-mode fractures in composite sandwich structures by exploring three ML algorithms. This research enables the extraction of CZM parameters for mixed-mode fractures and compares the ML algorithms to determine the one best suited for composite fracture analysis. To achieve this, we model the Asymmetric Double Cantilever Beam (ADCB) specimen, where we explore four parameters in the study. The parameters are maximum normal contact stress, critical fracture energy for normal separation, maximum equivalent tangential contact stress, and critical fracture energy are required. After calibration and validation of the current results, the ML algorithms' load-displacement curve is the input dataset. We discuss the benefits and drawbacks of the three machine learning algorithms, SVM, RF, and ANN, that we trained and optimized to identify the abovementioned CZM constants for fracture and determine the critical input parameters that impact the cohesive zone model parameters most.

2. Cohesive zone modeling

In this study, we identify the CZM parameters based on the analysis of the load-displacement curve from the specimen tests of the ADCB specimen. The ADCB is a simple but effective specimen for measuring fracture properties where the position of the crack in the specimen results in a different percentage of model-I and mode-II load. Unlike linear relationships, this method uses FE analysis findings to construct a series of ML models. Our suggested process consists of four major stages: (1) Data Collection: We collect data from the ADCB test using the FE model-generated load-displacement curves. This data serves as a basis for our further modeling attempts. (2) Data Preprocessing: Before model construction, we preprocessed the data to ensure the quality and effectiveness of the data for ML algorithms. This process involves data cleaning, feature selection, and data scaling. Our models use the maximum load, displacement at the maximum load, area under the curve, and slope of the load-displacement curve as the four input parameters for the ML algorithms. (3) Model Training and Optimization: In this stage, we investigate three different ML algorithms, namely, Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN). These algorithms are trained and optimized to capture the most essential data correlations. (4) Prediction and Identification: These ML methods are processed to identify the CZM parameters.

2.1 Cohesive Zone Model

In this study, we use the sandwich ADCB specimen to assess the fracture behavior occurring at the interface between the facesheet and core materials of the structure. Due to the geometrical characteristics of the sandwich structure and the differing materials in the facesheets and core, mixed-mode loading comprising both Mode I and Mode II, rather than purely Mode I loading, is induced. Numerical simulations of the ADCB specimen with a linear softening law (Figure 1) for

CZM are implemented in the ANSYS Workbench 2022R2 [38]. The CZM that was relies on a quadratic stress criterion to simulate the initiation of damage [39, 40]:

$$\left(\frac{\sigma_I}{\sigma_{u,I}}\right)^2 + \left(\frac{\sigma_{II}}{\sigma_{u,II}}\right)^2 = 1 \quad (1)$$

(σ_I, σ_{II}) represent the loading components for modes I and II, while ($\sigma_{u,I}$, $\sigma_{u,II}$) are the local strengths. Also, the softening relation during the propagation is as follows:

$$\sigma_m = (1 - d)k\delta_m \quad (2)$$

where

$$\delta_m = \sqrt{\delta_I^2 + \delta_{II}^2} \quad (3)$$

In the above equations, *k* is the interfacial stiffness, δ_m is the equivalent mixed-mode I+II relative displacement, and σ_m is the resulting mixed-mode I+II traction. The calculation of the damage parameter d involves taking the ratio between the calculated actual accumulated dissipated energy (G_{Td}) and the critical total fracture energy (G_{Tc}) specific to the current mode ratio (Figure 1). The linear energetic loading criteria, including Mode I and II, simulated damage propagation is:

$$\frac{G_I}{G_{Ic}} + \frac{G_{II}}{G_{IIc}} = 1 \quad (4)$$

which enables the defining of the G_{Tc} for each mode ratio $\beta = G_{II}/G_T$ as being:

$$G_{Tc} = \frac{G_{Ic}G_{IIc}}{(1-\beta)G_{IIc} + \beta G_{Ic}} \quad (5)$$

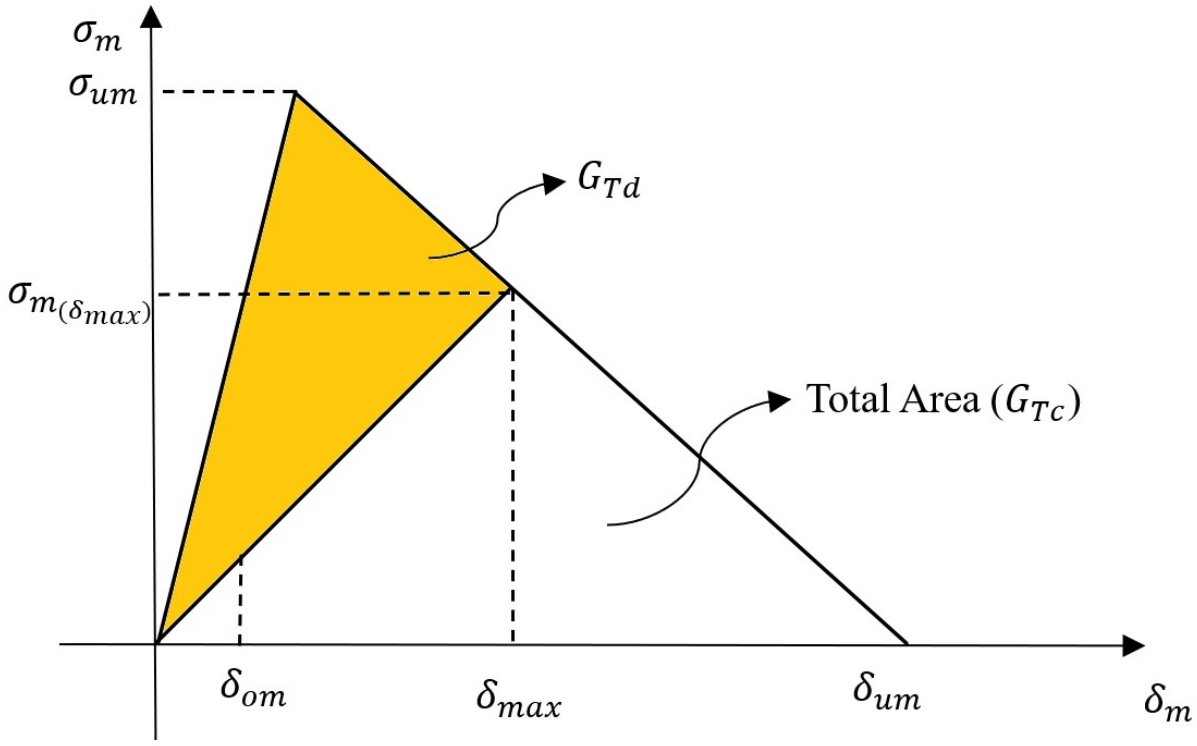

Figure 1: Schematic of the linear softening cohesive law for mixed-mode I+II fracture process. δ_{om}= damage onset relative displacement. δ_{um}= ultimate relative displacement. σ_{um}= local strength.

2.2 Numerical simulation of the ADCB specimen

The finite element model with the cohesive zone model can effectively represent the fracture characteristics of asymmetric structures with different materials, including composite sandwich structures with honeycomb cores and carbon fiber facesheets. In this context, we have developed a nonlinear ADCB finite element model with CZM to capture the interfacial failure process of the ADCB specimen. The model is then compared to experimental results in the literature. The material properties for both core and facesheets are presented in Table 1. Our approach involves a numerical simulation of the ADCB sandwich structure specimen, as shown in Figure 2, using the CZM with a linear softening law based on fracture energy. To effectively capture the mixed-mode fracture behavior of the sandwich structure, this tool requires the definition of four essential

parameters: maximum normal contact stress, critical fracture energy for normal separation, maximum equivalent tangential contact stress, and critical fracture energy for tangential slip. In our analysis, the first phase involves using geometry tools to produce an accurate schematic illustration of the structure to accurately depict the sandwich composite's design, dimensions, and layer orientations. A uniform element size of 1.5 mm was selected, discretizing the model into 8405 elements, making it possible for the simulation to capture stress concentrations accurately and reliably. The ACP (ANSYS Composite PrepPost) module appropriately represented the composite facesheets and the core. With the help of this module, composite materials can be made with the desired orientations, stacking sequences, and properties. The NOMEX® Honeycomb core and facesheets made of carbon-fiber-reinforced polymer (CFRP) prepreg were precisely modeled using ACP, considering their particular material characteristics.

2.3 Geometry and materials properties

The sandwich structure consists of unidirectional facesheets comprising ten layers of prepreg carbon-fiber-reinforced polymer (CFRP) prepreg (SEAL® Texipreg HS 160 RM). The excellent mechanical properties of CFRP drive this choice. This type of material offers high strength-to-weight ratios, exceptional stiffness, and superior fatigue resistance. These properties are vital for applications where structural integrity, weight, and durability are paramount. Each skin has a thickness of 1.5 mm. The core material is NOMEX® Honeycomb with a thickness of 10 mm. The honeycomb structure provides both lightweight and robust support to the facesheets. Table 1 shows the material parameters of the facesheets and core. Figure 2 illustrates a schematic of the ADCB test for composite sandwich structure. The height of the facesheets and core are h_f and h_c, respectively, and the initial crack is depicted by a_0.

Table 1: Elastic properties of the facesheets and core of the sandwich structure [39].

Material	Properties								
	E_1	E_2	E_3	v_{12}	v_{13}	v_{23}	G_{12}	G_{13}	G_{23}
CFRP	130 GPa	8200 MPa	8200 MPa	0.27	0.27	0.41	4100 MPa	4100 MPa	3200 MPa
NOMEX	0.45 MPa	0.45 MPa	258 MPa	0.9956	0.0005	0.0005	0.11 MPa	38.62 MPa	63.12 MPa

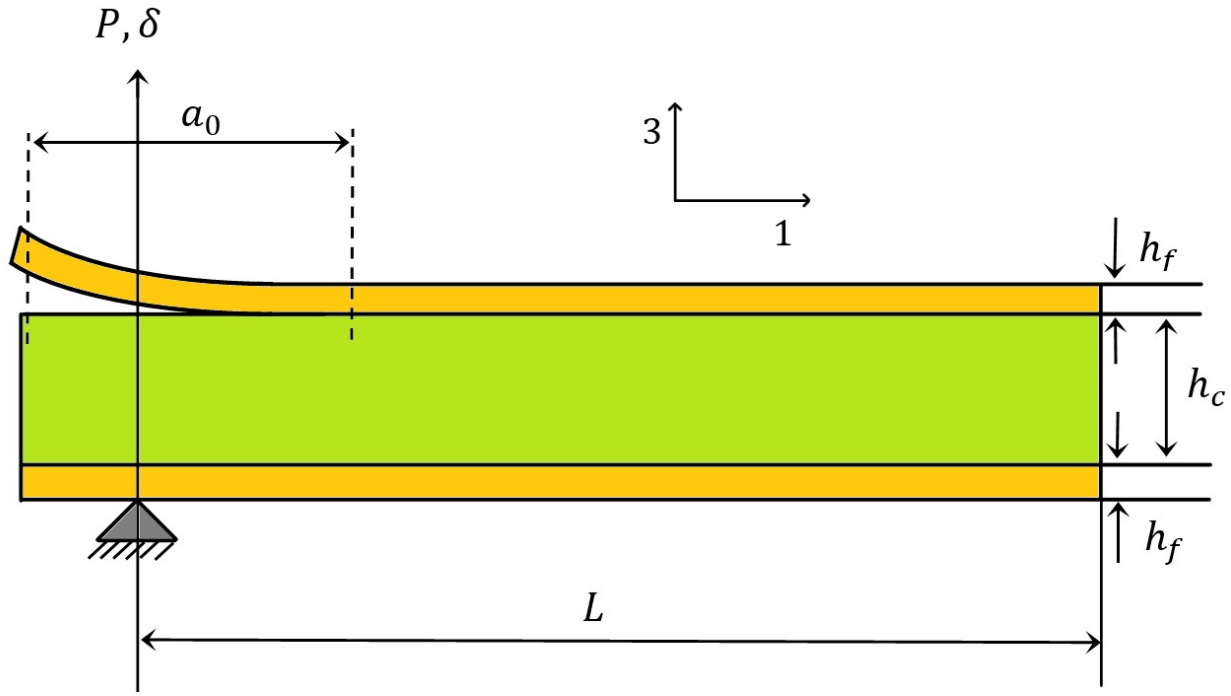

Figure 2: Schematic of the ADCB specimen for composite sandwich structure used to calibrate the cohesive zone models and determine the fracture properties; : L = 230 mm, a_0 = 50 mm, h_c = 10 mm, h_s = 50 mm.

3. Machine Learning Algorithms

This study uses three ML algorithms (SVR, RF, ANN) to train and predict the cohesive zone parameters. These algorithms consist of three main stages: data input, a learning algorithm, and output. The algorithm learns from historical data to generalize and predict new and unseen data, allowing continuous improvement. SVM algorithms are generally practical for binary classifications because they are based on mathematical optimization. On the other hand, RF is an ensemble method that uses multiple decision trees to improve accuracy and reduce overfitting. Finally, ANN is a deep learning method that uses multiple interconnected layers resembling the neuron connections in the human brain but typically requires large data sets. A description of the methods used is defined below.

3.1 Support Vector Regression

The SVR approach extends the Support Vector Machine (SVM), to build a machine learning model. SVM is known for its efficiency, and the most widely used SVR model is the ε-SVR introduced by Vapnik [41]. Notably, prior researchers have demonstrated the successful application of the SVR algorithm within the area of composite structures [42-44]. For our data set, we have defined them as $\{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_N, y_N)\}$, where each x_i is an input vector and its associated y_i represents the corresponding output value. Finding a mapping function $f(x)$ to describe the correlation of input and output values is the regression problem that needs to be solved. In our specific application, the input vector x_i is a collection of four parameters: δ^*, load, area, and slope. Correspondingly, the output y_i is a set of four values: $\sigma_I, G_I, \sigma_{II}$ and G_{II}. The formulation for the SVR approach in the linear regression is defined as follows [44]:

$$y = f(x) = w \cdot \varphi_x + b \tag{6}$$

where φ_x, $w = \{w_1, \dots, w_i, \dots, w_N\}$, and b are the mapping function, the adjustable weight vector, and the adjustable bias parameter, respectively. The SVR optimization problem can be written as follows:

$$\min \|w\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \tag{7}$$

$$subject\ to \begin{cases} y_i - w \cdot \varphi(x_i) - b \leq \varepsilon + \xi_i \\ -y_i + w \cdot \varphi(x_i) + b \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* > 0 \end{cases} \qquad i = 1, 2, \dots, n \tag{8}$$

where ε corresponds to the width of the ε-tube, which defines a zone around the predicted values where errors are allowed without penalty. C is a penalty factor, and ξ_i and ξ_i^* are slack variables that measure the distances between the predicted value and the nearest boundary of the ε-tube. Figure 3 illustrates the schematic of the linear SVR model used in this study.

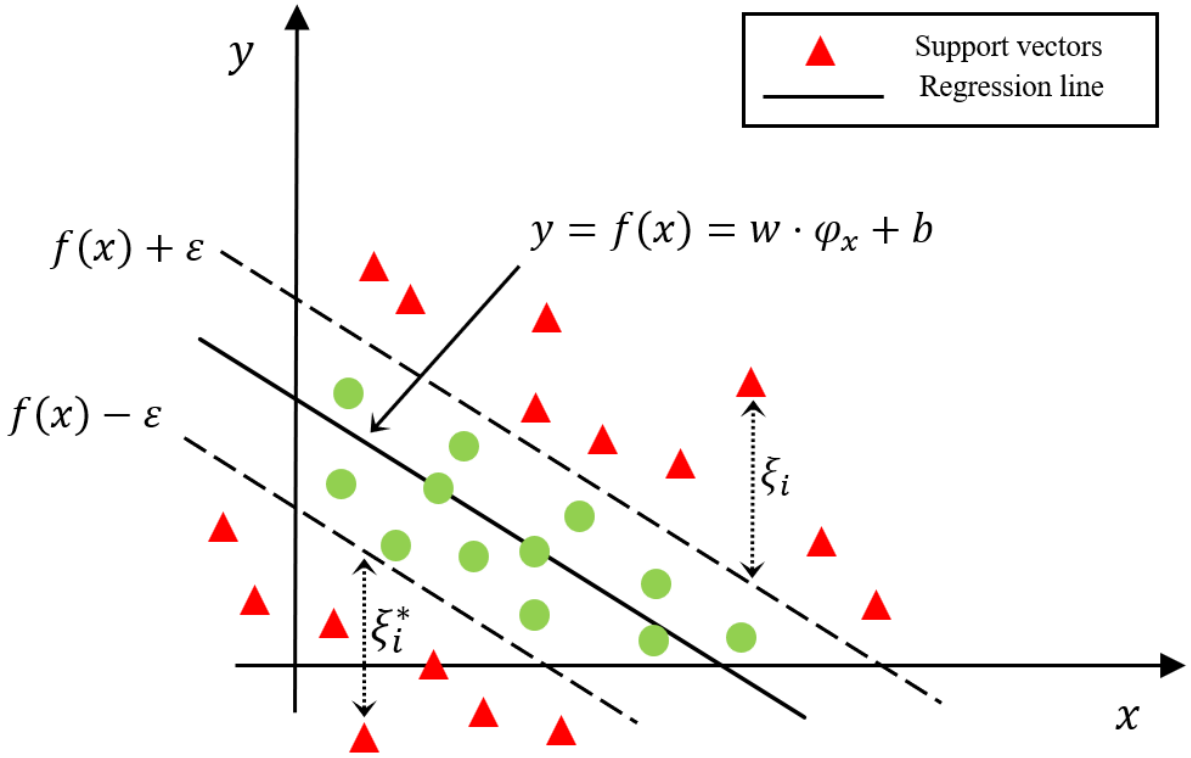

Figure 3: Representation of the linear SVR model showing support vectors and regression line

3.2 Random Forest

Classification and regression problems can be addressed using the ensemble machine learning technique called the random forest (RF) [45, 46]. Bootstrapping is one of the fundamental concepts in RF. This procedure describes randomly selecting samples using replacement to create ***n*** subsamples from the original dataset. This process implies that some samples might be chosen more than once in a subset while others might not. All these bootstrapped datasets train a single decision tree for RF. The randomness that RF adds to the feature selection process is one of its primary characteristics. It merely considers a random subset of characteristics rather than all features at each node to identify the optimal split. Every bootstrapped subsample provides the foundation for an individual decision tree, like Tree 1, Tree 2, Tree 3, and so on. Though part of the same forest, these trees are distinct in their structure and decision pathways. When all the decision trees are built, the regression prediction for a new data point can be obtained by aggregating the predictions from individual trees.

In our problem, the algorithm draws a bootstrap sample from our dataset, which includes specific values for δ^*, load, area under the curve, and slope. Next, for each node in a tree, the model randomly selects a subset of our features (δ^*, load, area, and slope). For instance, one node may prioritize splitting data depending on load values, while another may prioritize the area under the curve. Each tree is trained to understand the relationships between these features and the $\sigma_I, G_I, \sigma_{II}$ and G_{II} values. The schematic representation of the RF model is indicated in Figure 4. The mathematical formulation for the regression method is as follows:

$$f^N(x) = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{9}$$

where $f^N(x)$ is the predicted output for a new data point, *N* is the number of trees in the RF, and y_i is the prediction from the *i*th tree.

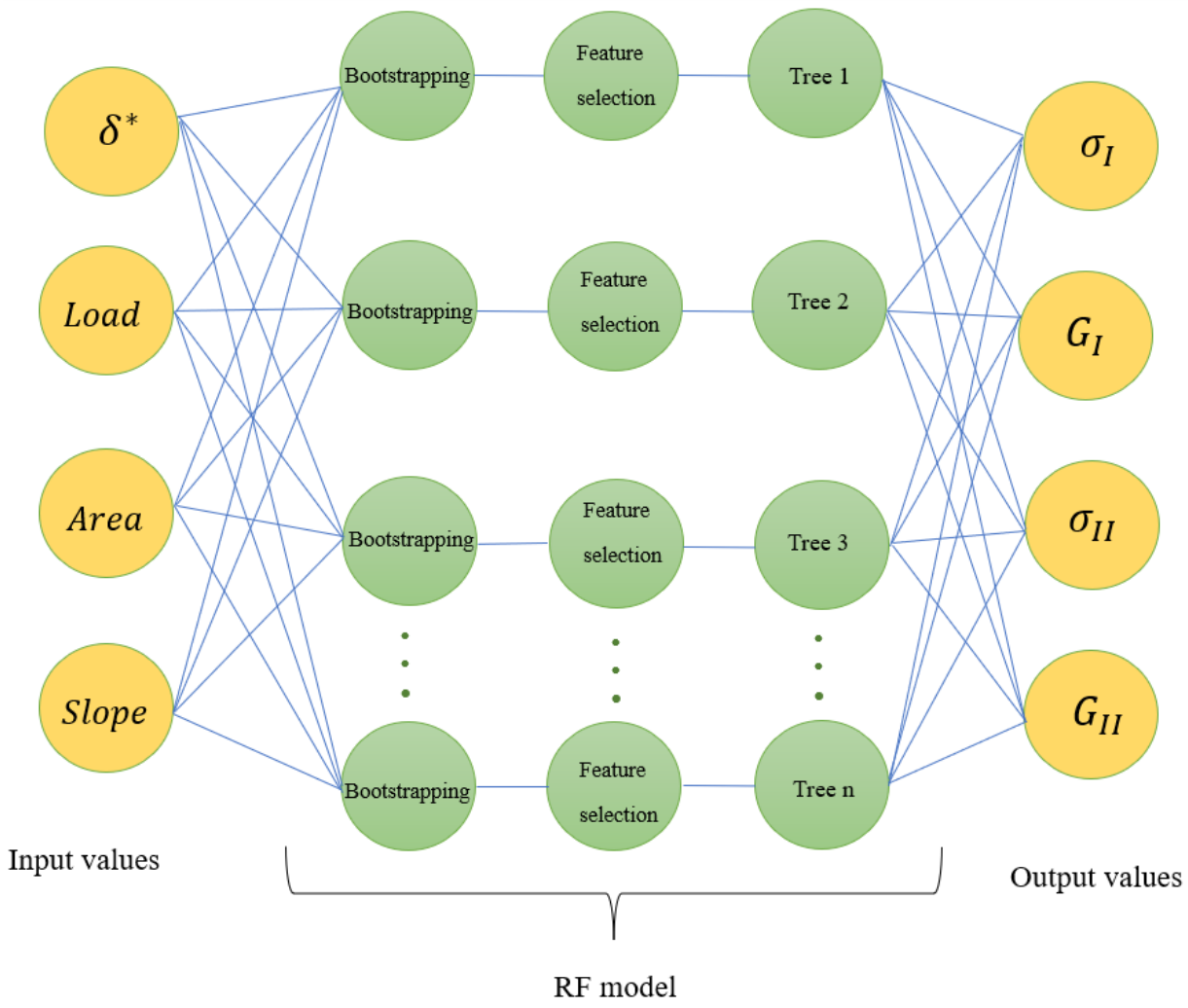

Figure 4: Representation of RF model for determining cohesive zone model parameters

3.3 Artificial Neural Network

Artificial Neural Networks (ANN) are popular machine learning algorithms widely used in engineering applications [47, 48]. ANNs are networks of artificial or natural neurons that are connected by a computational or mathematical model for processing information. This algorithm consists of Three essential parts: the input, hidden, and output layers. The input layer, the first part

of an artificial neural network, collects and encodes any raw data or characteristics from external sources. Each neuron in this layer represents a different input characteristic and acts as the network's entrance point for information. The hidden layers of an artificial neural network are the intermediary levels placed between the input and output layers. Hidden layers use interconnected neurons to carry out extensive mathematical operations on the input data, enabling the network to identify complex patterns and correlations. The output layer is the last layer of an artificial neural network. It is in the position of making predictions and classifications based on the data processed in the hidden layers. Each of our inputs (δ^*, load, area, slope) is represented as activation units from a_1^{in} to a_4^{in} (Figure 5). Each unit will process the input through their weights $W_{m,d}^{in}$. Each neuron in the hidden layer, represented as a_d^1 to a_f^3, processes information from the previous layer through connection weights like $W_{d,e}^1$ and transmits it forward. These weights determine the significance or influence of one neuron's output on another. By training this network on our dataset, these weights adjust themselves to optimize the prediction of outputs ($\sigma_I, G_I, \sigma_{II}$ and G_{II}) based on the inputs (δ^*, load, area, and slope).

This study uses a three-layer ANN machine-learning technique, as depicted in Figure 5. Within this framework, x_i represents the δ^*, load, area, and slope values, while y denotes $\sigma_I, G_I, \sigma_{II}$ and G_{II} values. The start of the ith neuron, in the lth level, is denoted as a_i^l, and the connections between the kth sector in level l and the jth sector in the subsequent level $(l+1)$ are represented by $w_{k,j}^l$. To illustrate, a_i^{in} and a_i^{out} correspond to the activation units. The output in level, $(l+1)$ is a function of the sum of their weighted inputs.

$$a^{l+1} = f(a^l W_k^l + b^l) \tag{10}$$

In Eq. (6), the activation function denoted as f determines whether to activate the neuron, while b^l is a vector representing bias for the $(l+1)$th level. a^l is a vector with the activation units for level l. Furthermore, Eq. (8) defines W^l as a weight matrix with dimensions $r \times s$, where s is the number of units in the $(l+1)$th layer.

$$a^l = [a_1^l, a_2^l, \dots, a_r^l] \tag{11}$$

$$W^l = \begin{bmatrix} w_{1,1}^l & \cdots & w_{1,s}^l \\ \vdots & \ddots & \vdots \\ w_{r,1}^l & \cdots & w_{r,s}^l \end{bmatrix} \tag{12}$$

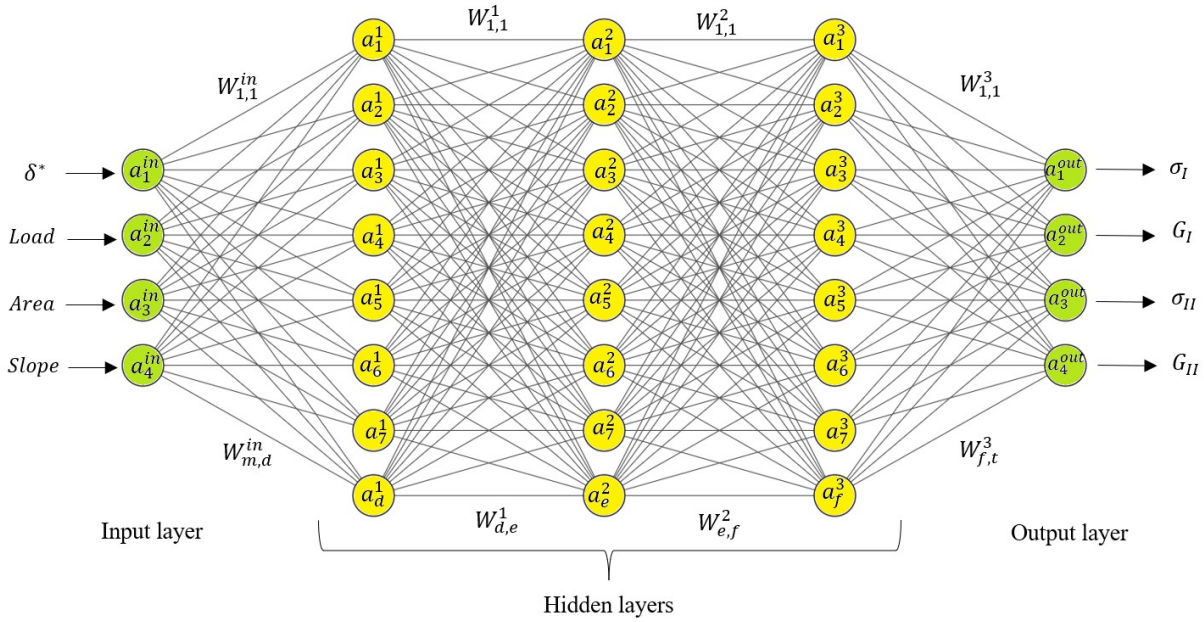

Figure 5: Representation of ANN architecture for identifying cohesive zone parameters

4. Results and Discussion

4.1 Validation of modeling approach

The finite element (FE) results of the ADCB specimen with the cohesive zone model are compared with experimental results to show the ability of the model to capture the fracture propagation behavior. Since the load-displacement curve is a required input for the ML methods utilized in the present study, we conduct a validation to compare our FE outcomes for this curve with those in the existing literature. The load-displacement curve is a graphical representation that depicts the relationship between the applied load and the resulting displacement or deformation in the structure. It can be seen that in Figure 6, the present results are in reasonable agreement with the experimental test which is presented in [39].

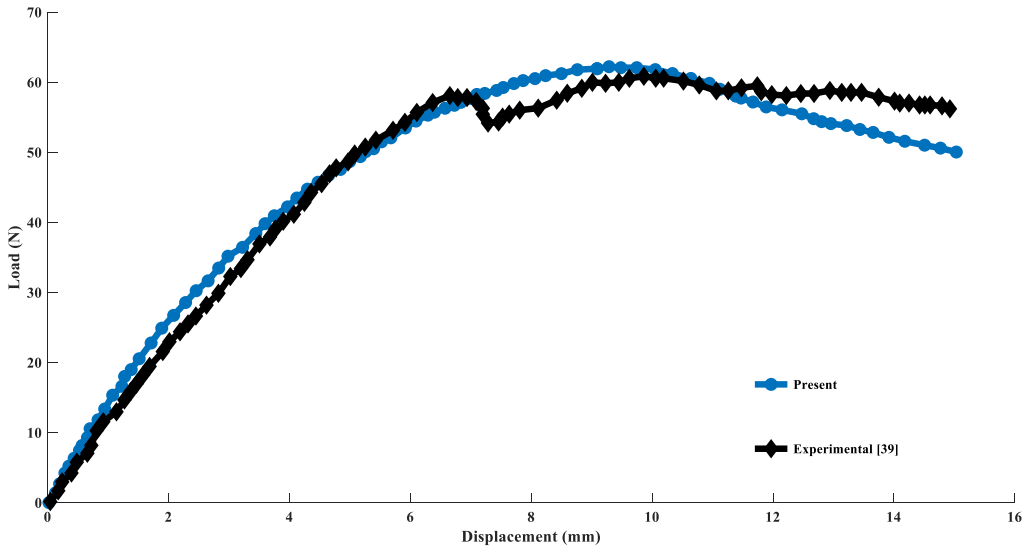

Figure 6: Validation study showing the load-displacement curve obtained from the FE solution for the sandwich structure using the ADCB specimen.

4.2 Dataset collection for cohesive zone model parameters

The CZM requires the specification of four parameters for mixed-mode fracture properties. The following parameter ranges were used based on previous research and practical experience: σ_I runs from 0.15 MPa to 0.35 MPa, G_I is between 280 J/m^2 and 600 J/m^2, σ_{II} is between 1.6 MPa and 3.5 MPa, and G_{II} is between 700 J/m^2 and 980 J/m^2. In Figure 7, it can be observed that this results in data points that represent 51 distinct cohesive characteristics generated by these parameter combinations. Figure 7 visually represents four dimensions because CZM contains four parameters. The x, y, and z axes correspond to σ_I, G_I, and σ_{II}, respectively, while G_{II} is depicted on the fourth color dimension. Moreover, the circle, triangle, and square symbols represent either a point in the dataset, a point within the gap of the dataset, or a point outside the dataset, respectively. A series of finite element nonlinear analyses were performed to capture the debonding for this dataset. Consequently, our database contains 51 load-displacement curves from this modeling, some depicted in Figure 8.

4.3 Hyperparameters tuning

The next step in our approach is hyperparameter tuning, which involves systematically adjusting the settings and configurations of a machine learning algorithm to optimize its performance on the given dataset. Hyperparameters are parameters not learned from the data but set before training. We discuss the techniques we used in this process for each ML algorithm. In SVR hyperparameter tuning, the $\boldsymbol{\varepsilon}$ value and the box constraint are essential for optimal performance. The error tolerance is controlled by $\boldsymbol{\varepsilon}$, which establishes the margin size so that errors do not penalize the model. This affects the model's sensitivity to variations in the training set.

In contrast, the box constraint is pivotal in balancing the tradeoff between the model's complexity and its ability to fit training data errors. Its correct calibration ensures that the model neither

overfits by capturing unnecessary noise nor underfits by being too generalized. As shown in Table 2, which thoroughly examines how these configurations affect the behavior of the SVR model, different combinations of epsilon and box constraint values are presented in detail. The MAPE and R^2 are shown and compared in this table.

Table 2: Hyperparameter tunning for SVR

ε	Mean R^2	Mean MAPE	Box Constraint	Mean R^2	Mean MAPE
0.01	0.5308	**6.0558**	0.1	0.4658	7.2371
0.02	**0.5376**	6.1988	0.7	0.5359	6.2742
0.03	0.5233	7.1526	1	0.5376	**6.1987**
0.04	0.5168	7.4750	2	**0.5424**	6.3128
0.05	0.5024	7.8319	3	0.5380	6.4652

Results of a K-fold cross-validation hyperparameter tuning procedure for the random forest method are presented in Table 3. Finding the model configuration with the best Mean Absolute Percentage Error (MAPE) requires optimizing the hyperparameters. MAPE measures predictive accuracy; the lower the MAPE, the more effective the model. With MAPE as the primary assessment metric, we investigated how changing the number of trees in the random forest affects the model's performance. The investigation used a grid search technique to examine the ensemble's performance for tree counts ranging from 5 to 650. It is important to note that the quantity of trees directly impacts the stability of the predictions in the RF [49, 50]. Although a single decision tree can be susceptible to noise in its training data, an ensemble of trees can mitigate individual errors, leading to a more robust and stable model. However, there is a tradeoff that a model with too few trees may have a high variation, and a model with too many trees may have computational inefficiencies without improving accuracy. One important finding from our analysis is that, based on the lowest MAPE value, 82 trees seem to be the ideal number for this model. This suggests that the average of the model's prediction accuracy is higher than for other configurations. The performance of each design is analyzed over ten distinct cross-validation folds.

Furthermore, as the number of trees in the model increases, we observe that the MAPE tends to stabilize when evaluated without employing the k-fold technique. However, the introduction of k-fold cross-validation introduces variability in the MAPE values. The reason for this variation is that the method used in k-fold cross-validation chooses different subsets of data for testing and training. This naturally introduces diversity into the results, emphasizing the impact of different data subsets on the model's accuracy.

Table 3: Cross-validated MAPE scores of the 8 RFs.

NO.	NO. of Trees	Number of Folds (%)										Mean (%)
		1	2	3	4	5	6	7	8	9	10	
1	5	8.4952	9.6051	11.9662	8.6994	7.4333	9.2395	6.1857	9.9078	11.5513	11.249	9.4332
2	10	8.7319	10.6674	9.7386	8.9025	9.3058	11.1510	9.1339	7.0797	9.5477	10.2851	9.4544
3	36	8.4919	10.4474	8.5306	10.7308	11.5945	8.2028	8.0728	10.4431	10.0279	10.8345	9.7376
4	82	8.4929	11.4597	9.1358	9.6219	9.7889	8.4278	11.4191	5.7201	10.1255	7.8462	9.2038
5	140	8.4924	9.8091	7.1622	10.7415	7.4258	10.4358	8.9860	10.9316	9.8295	10.3503	9.4167
6	260	8.4918	11.4592	11.1199	7.7427	9.6577	10.3049	7.2246	9.4013	9.8909	7.2202	9.2517
7	480	8.4918	10.2073	8.7420	8.9150	8.7455	8.3842	11.089	10.9269	13.2572	13.8415	10.2604
8	650	8.4918	10.7436	9.9257	8.5966	12.4392	6.6686	10.0478	6.5755	11.3680	8.3010	9.3161

In the ANN algorithm, the configuration of network layers and the number of neurons in each layer are considered hyperparameters. Optimizing the model features is also called hyperparameter tuning. The dataset is initially divided into training and testing subsets to aid this process. In the current study, we use 80% of the dataset for the training. In our attempt at hyperparameter tuning, we use a ten fold cross-validation approach. This approach allows for avoiding the overfitting the machine learning algorithm. With this validation method, the dataset for the training is randomly divided into ten unique, non-overlapping subsets, known as folds. During each iteration, nine folds are utilized to train the model, while one fold is an independent evaluation set. This process is iterated ten times, yielding ten performance assessments for the training dataset. Finding the model configuration with the best MAPE requires optimizing the hyperparameters for our three-

hidden-layer artificial neural network (ANN). We used k-fold cross-validation throughout the hyperparameter optimization process to ensure our findings were reliable and robust. We employed a random search strategy to manage the computational complexity arising from the extensive search space, randomly selecting hyperparameter combinations for evaluation. Using this method, we effectively explored the vast hyperparameter space and optimized the ANN's architecture without considering every possible variety. It is a sensible option for hyperparameter tuning in complex machine-learning models because it balances computational resources and model performance. The best-performing hyperparameter configurations are listed in Table 4, along with the corresponding MAPE values. The leading design (24, 136, 68), is made of three layers, achieving an average MAPE of 1.0321%. Based on these results and in conjunction with the purelin transfer activation function, we selected the (24, 136, 68) design as the machine learning model for our study.

Table 4: MAPE levels of the eight ANNs studied with 3 hidden layers.

Number.	Design	Number of Folds (%)										Mean (%)
		1	2	3	4	5	6	7	8	9	10	
1	(24, 136, 68)	0.9660	1.0791	1.0517	0.9845	1.0195	1.1307	1.0552	1.0303	1.0629	1.0409	1.0321
2	(128, 128, 64)	1.0464	1.0807	0.9970	1.0213	1.0448	1.0183	1.0091	1.0912	1.0286	1.0598	1.0417
3	(64, 128, 32)	0.9523	1.1049	1.0135	1.0064	0.9675	1.1326	1.1764	1.0116	1.0382	1.0859	1.0489
4	(64, 32, 16)	0.9965	0.9698	1.0894	1.1165	1.0789	0.9931	1.0096	1.0806	1.2289	1.0737	1.0518
5	(25, 14, 30)	1.0566	1.2809	1.0140	1.0373	0.9576	1.1982	1.0863	0.9965	1.0097	1.0804	1.0718
6	(6, 12, 18)	1.1260	1.1050	0.9924	1.4297	0.9799	1.0727	1.0918	1.0369	1.0854	1.0713	1.0991
7	(12, 12, 12)	1.1769	1.2491	3.1212	1.2077	1.1493	1.1373	1.1298	1.0434	0.9577	1.0871	1.3260
8	(4, 4, 16)	1.0675	1.0979	2.2337	1.4208	1.0566	2.6262	1.6814	1.1751	1.0834	1.0336	1.4476

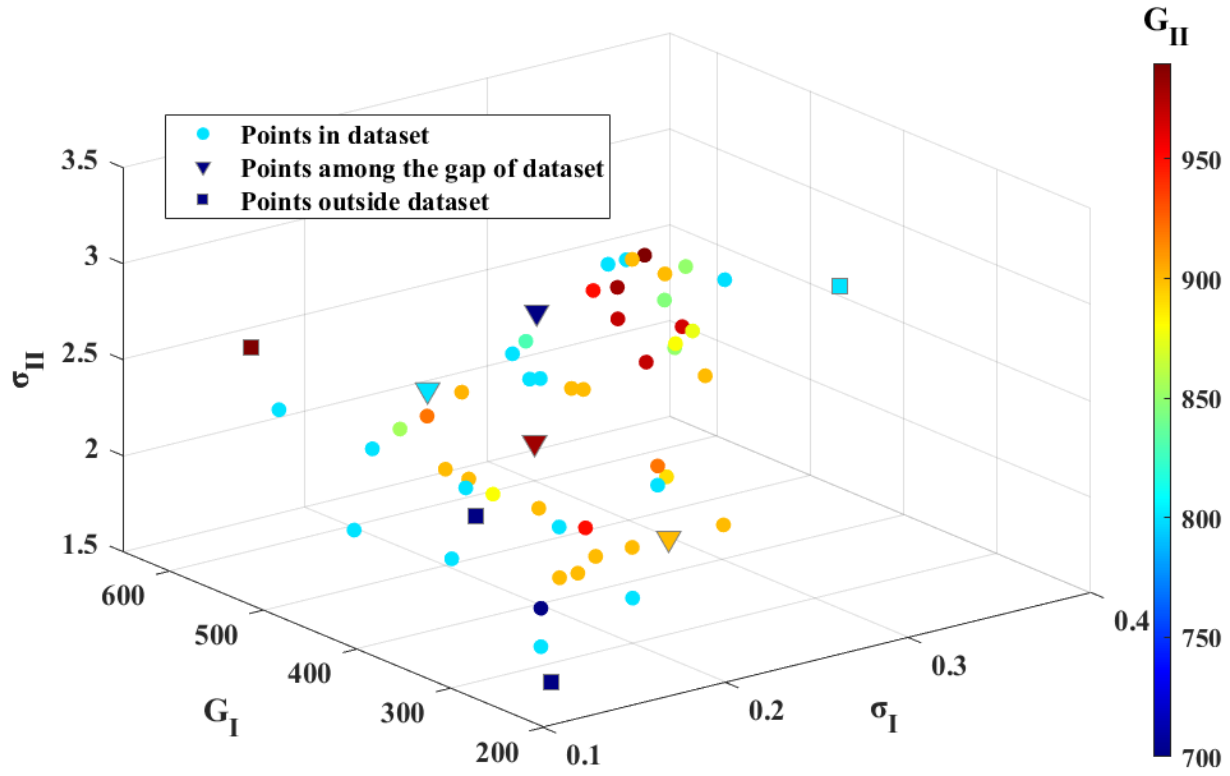

Figure 7: Cohesive properties cases in the whole database.

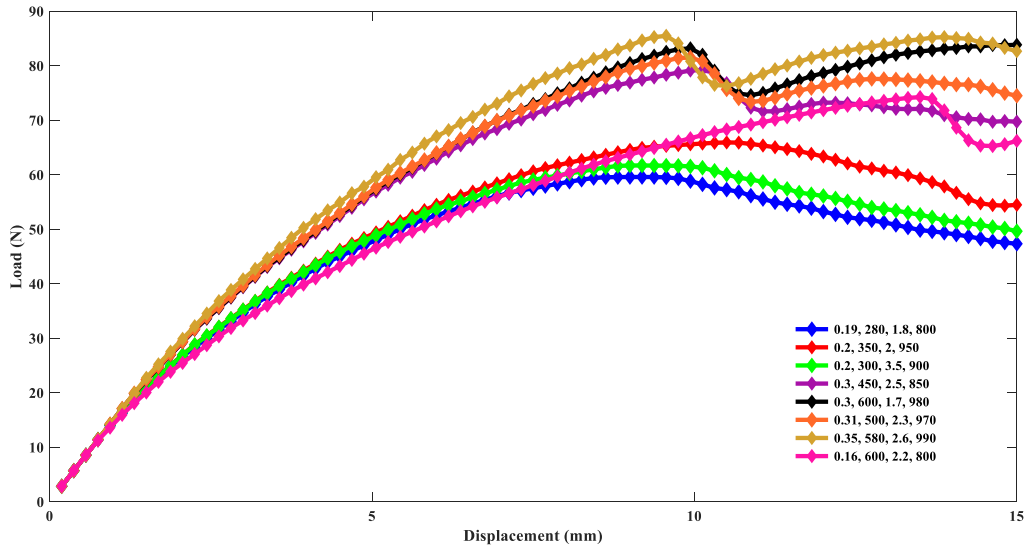

Figure 8: Typical load versus displacement curves in the database from ADCB fracture simulations.

4.4 Evaluation of Machine Learning Algorithms

In the evaluation of the performance of the ML algorithms, we compare the actual and predicted output for the training and test sets in the SVR, RF, and ANN algorithms, respectively. This comparison provides a direct perception of each algorithm's accuracy in approximating the CZM parameters. Figures 9 to 14 illustrate a comparative analysis between the CZM parameters predicted by the machine learning models and the true values during the training and testing phases. The results display how closely the model matches the actual dataset and shed light on the model's variability. Compared to the ground truth values for the training and testing sets, the results show that all the machine learning models produce remarkably accurate predictions. This substantial agreement highlights the effectiveness of the machine learning method for CZM parameter identification, which primarily relies on the load-displacement responses of the ADCB. Notably, it's important to note that the training sets are randomly chosen. It's also worth mentioning that the training set constitutes 80% of the data in this study.

Figures 9 and 10 display comparative regression analysis of the SVR model for training and testing datasets, respectively. In these figures, there's a noticeable variance in the σ_{II} parameter during both the training and testing. Certain areas in the training set where the data points diverge from the reference line suggest possible weaknesses in the model. Comparably, the testing set exhibits noticeable variations in particular regions, which may indicate that the model could not fully capture the intricacies of the training data. Although many of the training data points for G_{II} closely match the reference line, there are some deviations, particularly as values rise. Similar trends may also be seen in the testing data, with predictions being less reliable in particular value ranges. These variations suggest the SVR model could face challenges in capturing specific ranges. Thus, the

SVR model demonstrates good performance for σ_I and G_I there are evident challenges with the σ_{II} and G_{II} parameters. These variations could be attributed to inherent complexities in the dataset. It is worth noting that this phenomenon can also be seen in Figure 15, which indicates the R^2 and MAPE values, and it reveals that SVR underperforms compared to ANN and RF for σ_{II} and G_{II}.

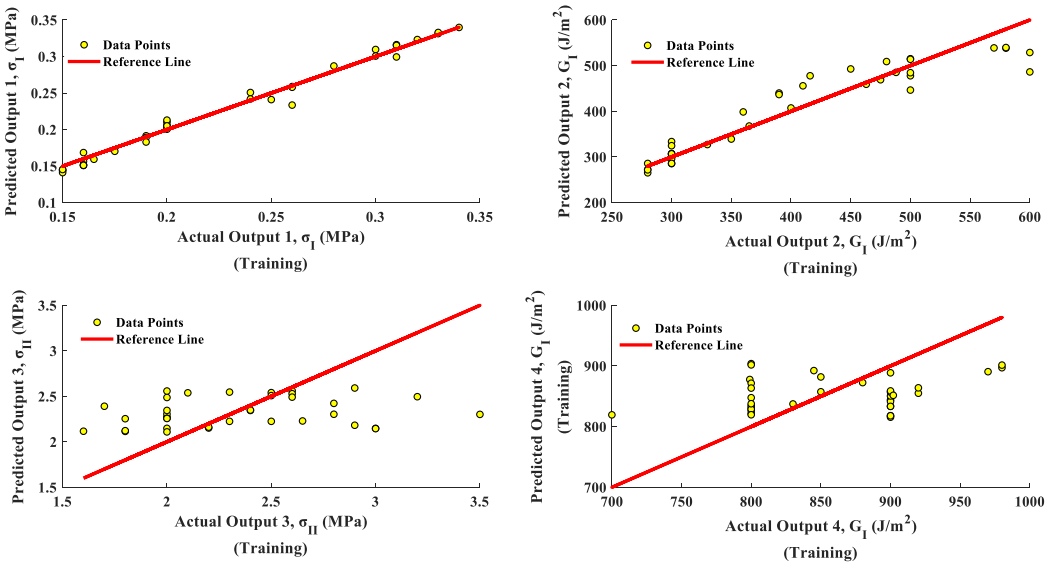

Figure 9: Results showing the training set of different outputs for SVR

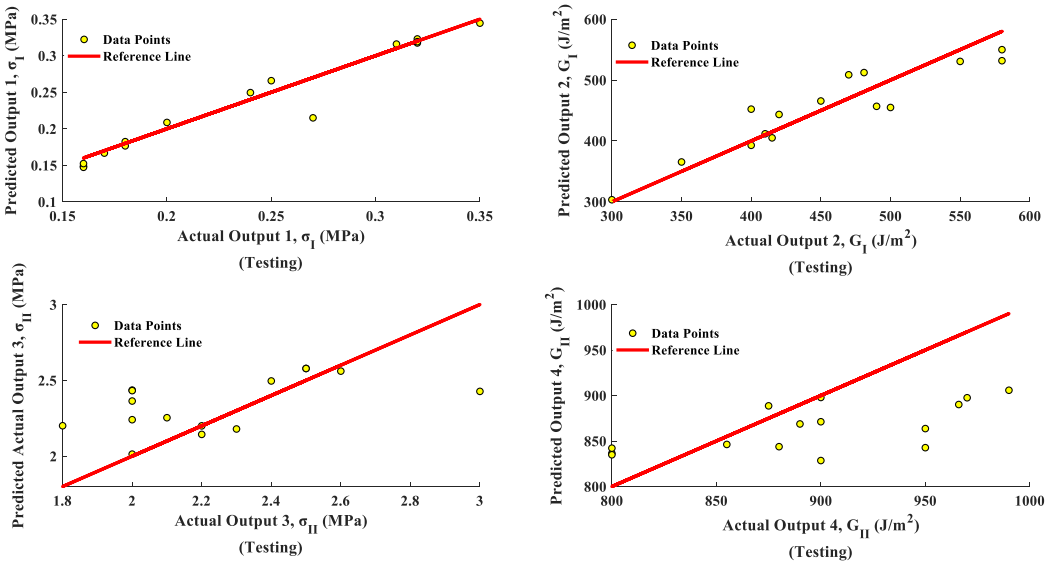

Figure 10: Results showing a test set of different outputs for SVR

The training set for the RF model is displayed in Figure 11, where the predicted and actual outputs for $\sigma_I, G_I, \sigma_{II}$ and G_{II} are well aligned. The test set is shown in Figure 12, with data points largely following the reference line but showing minor variations in G_{II}. Both figures indicate the RF model's predictive capability with slight variations in testing. Figures 13 and 14 show comparative regression analysis of the ANN model for training and testing datasets, respectively. Both figures indicate scatter plots of actual versus predicted outputs, with the red reference line showing an ideal one-to-one correspondence. Figure 13 demonstrates that throughout the training phase, the data points for outputs $\sigma_I, G_I, \sigma_{II}$ and G_{II} are closely clustered around the reference line, indicating a well-fitting model. The closeness of data points to the reference line for each output in Figure 14 also shows the model's resilience and capacity for generalization during testing.

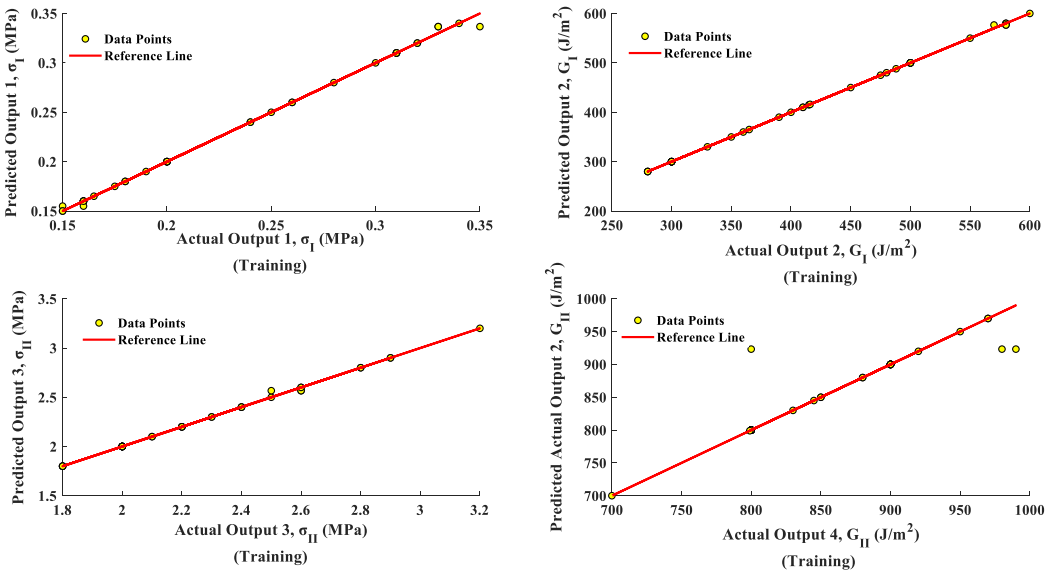

Figure 11: Results showing the training set of different outputs for RF

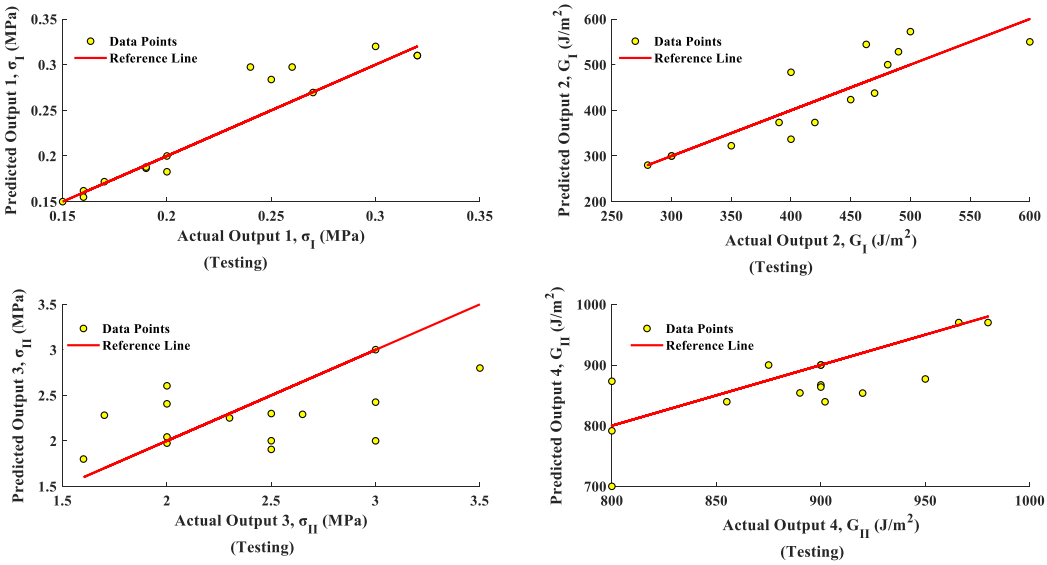

Figure 12: Results showing the test set of different outputs for RF

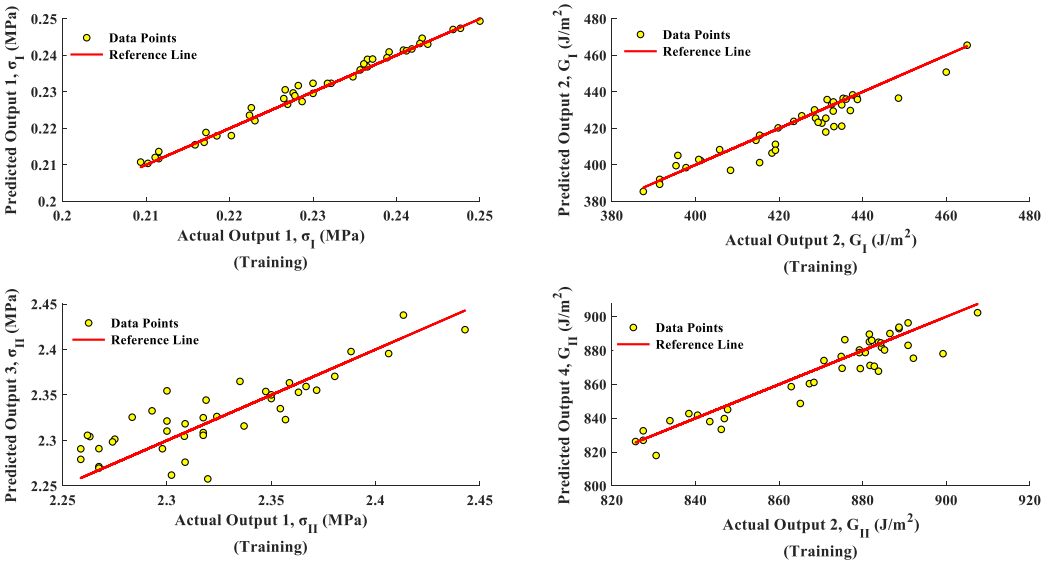

Figure 13: Results showing the training set of different outputs for ANN

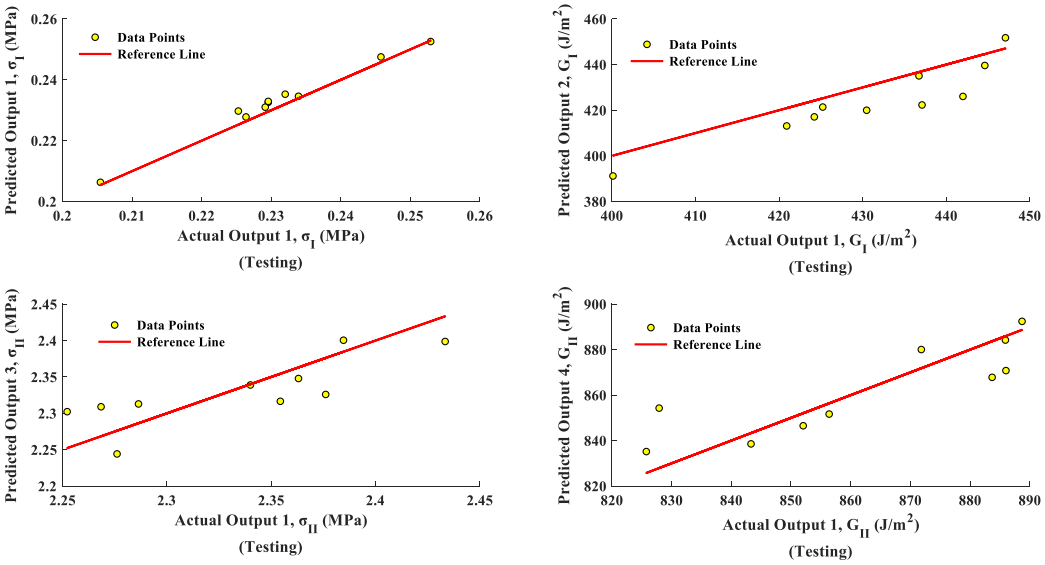

Figure 14: Results showing the test set of different outputs for ANN

A comparative analysis of the three ML algorithms uses the R-squared (R^2) and the Mean Absolute Percentage Error (MAPE) metrics. The normalized index R^2, shown in Eq. (13), is frequently used to assess algorithms for various problems.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{13}$$

where n is the number of data points, y_i represents the actual CZM parameter values $\hat{y}_i$ represents the predicted CZM parameter values, and $\bar{y}$ is the mean of the actual CZM parameter values. The method's accuracy typically increases as the index R^2 gets closer to 1. Another important performance metric in evaluating the accuracy of our machine learning algorithms is the Mean Absolute Percentage Error (MAPE). The percentage deviation between predicted and actual CZM parameters is measured comprehensively by MAPE.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{14}$$

It should be mentioned that the method's accuracy tends to improve as the MAPE value approaches 0%. Notably, because the unit and magnitude of the CZM parameters differ, the comparison of Root Mean Square Error (RMSE) is not included in this study because RMSE is sensitive to the scale of the data [44]. Figure 15 compares the R^2 and MAPE values produced from three distinct ML methods. All three methods consistently have R^2 values greater than 0.83 for σ_I. Notably, ANN has an R^2 value of almost 0.98, which is the highest. R^2 values for all methods continue to exceed 0.7 in G_I. SVR has the greatest R^2 value of 0.85 in this output. However, for σ_{II} and G_{II}, a noticeable trend appears. Only the ANN algorithm maintains an R^2 value near to one, whereas the R^2 values of the other two methods decline significantly. When the performance of these three methods is compared, ANN constantly surpasses the other two, especially for σ_{II} and G_{II}, where it achieves near-perfect fit with a R^2 value close to 1. Overall, the average R^2 for all outputs is 0.7932 for ANN, 0.5224 for SVR, and 0.4908 for RF. Similar trends emerge when comparing the MAPE values generated by three different machine learning approaches. Evidently, the MAPE values for the ANN algorithm consistently outperform those of the other two methods across all outputs. Notably, the MAPE value for the RF algorithm increases by about 16% in σ_{II}, making it the highest among all MAPE values. When the average MAPE values for all outputs are considered, ANN has a score of 0.9578, whereas SVR and RF lag with values of 6.1093 and 8.4951, respectively.

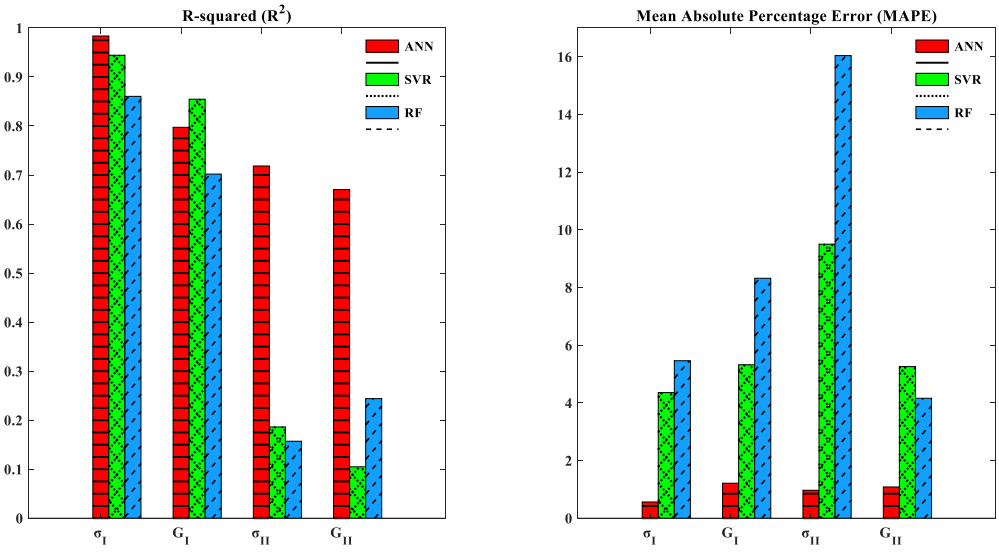

Figure 15: Comparison of R^2 and MAPE for various ML algorithms and different outputs.

Figure 16 visually represents the feature importance for all ML algorithms. Given the four inputs in this study, the figure shows how these inputs influence the overall responses. The impact of these features varies between algorithms. In the case of SVR and RF, for example, the area under the curve significantly influences the responses. In the ANN model, however, maximum load, area under the curve, and slope of the curve all contribute roughly equally to shaping the responses.

Furthermore, in SVR, the maximum load and the slope of the curve are roughly equal in importance in determining the responses. Figure 17 compares the CZM parameters obtained via various ML algorithms and the ground truth values obtained through ANSYS for load-displacement curves of mixed-mode fractures in composite sandwich structures. We used four distinct points (represented by triangular shapes in Figure 7) located among the gap of the dataset. The figure depicts an excellent agreement between the CZM parameters predicted by the ML algorithms and those obtained through rigorous finite element simulations, confirming the

reliability and precision of the ML-based approach for characterizing mixed-mode fractures in such complex structures.

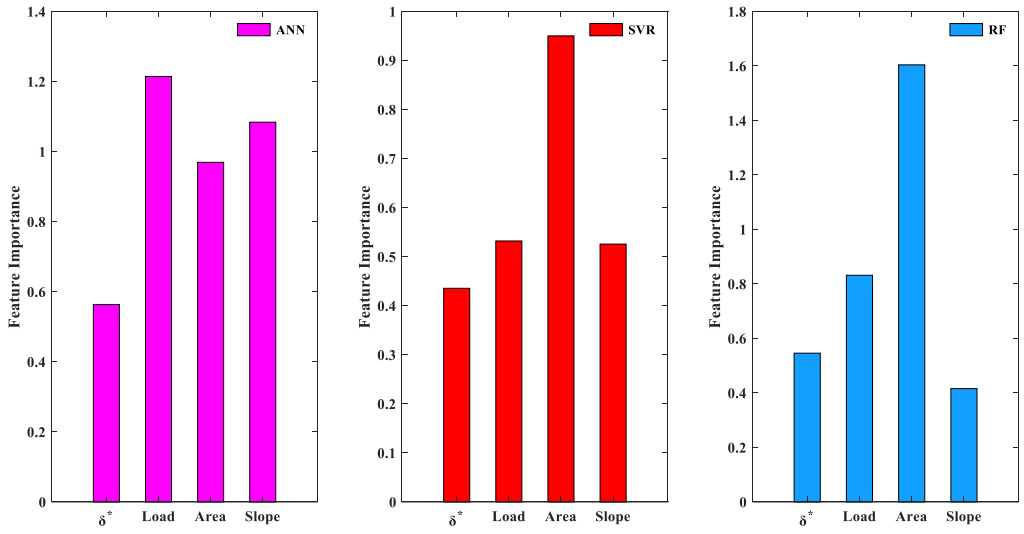

Figure 16: Feature importance for the machine learning algorithms

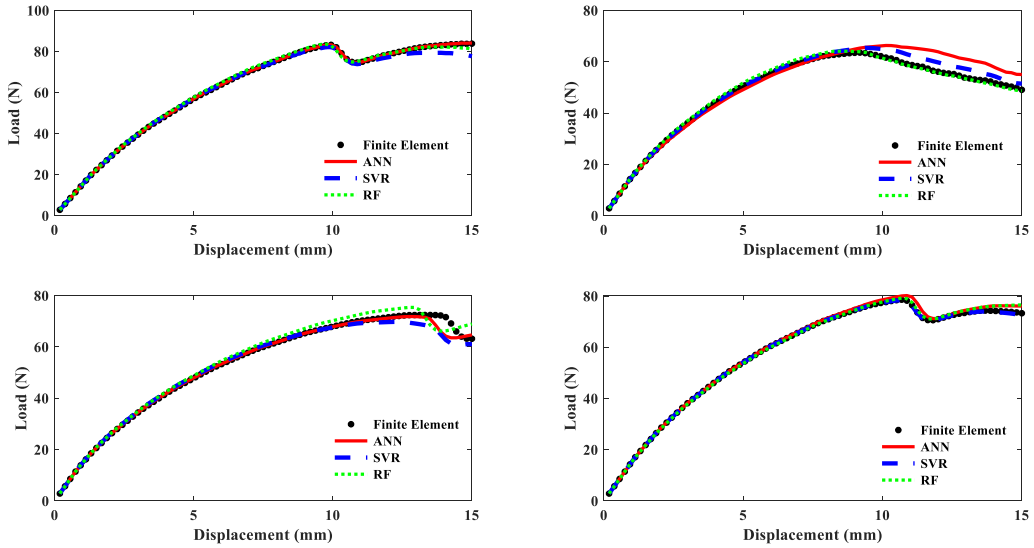

Figure 17: Comparison of load-displacement curves from different ML algorithms

We also chose four points outside the dataset to demonstrate the adaptability of our ML algorithms in predicting CZM parameters. Figure 18 shows the ML algorithms' ability to effectively predict CZM parameters even when dealing with data points outside the dataset boundaries. It should be noted that all ML algorithms can produce accurate outcomes in predicting CZM parameters for this out-of-dataset point, denoted by a square marker in Figure 7. This demonstrates the robustness and generalizability of our ML-based approaches to characterizing mixed-mode fractures beyond the boundaries of the training data.

A correlation heatmap can also investigate the relationships between input features and outputs. Figure 19 includes correlations between inputs, correlations between outputs, and correlations between inputs and outputs, providing a view of the relationships in the dataset. The heatmap utilizes a color scale to represent different levels of correlation, with warmer red colors representing positive correlations and cooler blue ones representing negative correlations. A closer assessment reveals notable patterns within the heatmap. Specifically, the area under the curve feature has a strong positive correlation with G_I, indicating that as this feature increases, G_I also increases.

In contrast, there is almost no visible correlation between the same feature and σ_{II}. Furthermore, the area under the curve has a nearly strong positive correlation with both σ_I and G_I. The displacement at maximum load feature, on the other hand, has a negative correlation with σ_I, σ_{II}, and G_{II}, indicating that as this feature increases, σ_I, σ_{II}, and G_{II} tend to decrease.

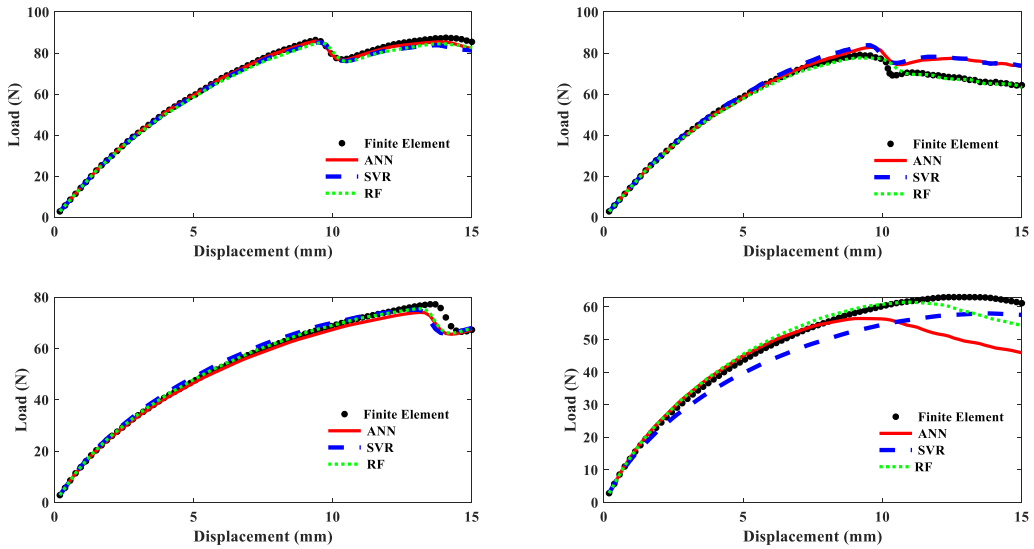

Figure 18: Comparison results for the points outside the dataset.

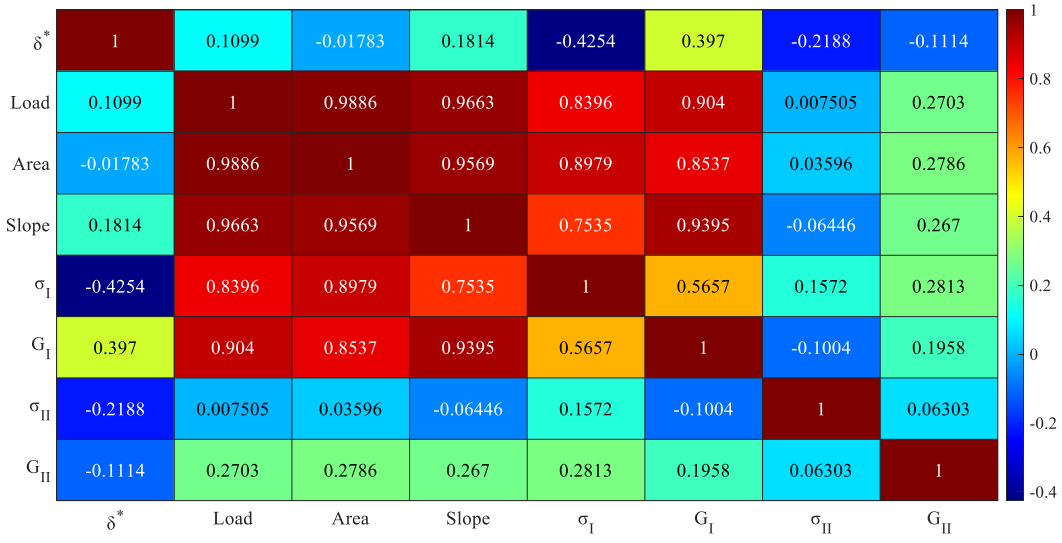

Figure 19: Correlation heatmap for input features and outputs.

5. Conclusions

In this study, three different machine learning algorithms, Support Vector Regression (SVR), Random Forest (RF), and Artificial Neural Networks (ANN), are developed and evaluated to identify the Cohesive Zone Method (CZM) parameters for finite element simulation of mixed mode fracture of composite sandwich structures using the ADCB specimen. Our training dataset consisted of 51 load versus displacement responses generated from the numerical models of the ADCB specimen. For input features, we considered the displacement at the maximum load (δ^*), the area under the curve, the maximum load, and the slope. Hyperparameter tuning was performed for the ML algorithms to reach their best possible performance. Analysis of the relationships between these inputs and the corresponding outputs ($\sigma_I, G_I, \sigma_{II}, G_{II}$) were indicated using a correlation heatmap showing the importance of the various input parameters. Notably, we saw that the machine learning (ML) algorithms demonstrated strong predictive capabilities for CZM parameters at the interface of facesheet and core in sandwich structures. Among the algorithms, ANN consistently showed superior performance when compared to RF and SVR. The overall Mean Absolute Percentage Error (MAPE) and R-squared (R^2) values for ANN were 0.9578 and 0.7924, respectively, outperforming both RF and SVR. It's also important to note that our proposed machine learning models demonstrated the ability to accurately predict points outside the dataset, indicating their ability to calculate within a high-dimensional space outside the training set. Another significant contribution of our approach is the ability to simultaneously predict and identify interface characteristics with varying scales and magnitudes.

Contribution Statement

Arash Ramian: Conceptualization, Methodology, Validation, Simulation analysis, Software, Visualization, Writing– original draft, Writing– review & editing. Rani Elhajjar: Conceptualization, Methodology, Resources, Supervision, Project administration, Writing– review & editing.

Declaration of competing interest

The authors declare no potential competing interests.

Data availability

Data will be made available on request.

References

[1] Chen Y, Zeng X, Deng Y, Wei G. Investigation on manufacturing and low-velocity impact performance of all-composite sandwich structure with S-type foldcore. Composite Structures. 2022;290:115539.

[2] Lv H, Shi S, Chen B, Ma J, Sun Z. Low-velocity impact response of composite sandwich structure with grid–honeycomb hybrid core. International Journal of Mechanical Sciences. 2023;246:108149.

[3] Xue P, Wei X, Li Z, Xiong J. Face-core interfacial debonding characterization model of an all-composite sandwich beam with a hexagonal honeycomb core. Engineering Fracture Mechanics. 2022;269:108554.

[4] Poblete FR, Mondal K, Ma Y, Dickey MD, Genzer J, Zhu Y. Direct measurement of rate-dependent mode I and mode II traction-separation laws for cohesive zone modeling of laminated glass. Composite Structures. 2022;279:114759.

[5] Zangoei AR, Rahimi GH, Gazor MS, Vahidibidhendi M, Azarniya O. Experimental and numerical study of delamination phenomenon in sandwich structures with hybrid corrugated core in the mode II of crack growth. Journal of Composite Materials. 2023;57(6):1199-210.

[6] Fu C, Wang X. Simulating delamination in composite laminates involving large scale fiber bridging based on the mixed-mode three-linear cohesive zone model. Theoretical and Applied Fracture Mechanics. 2022;117:103164.

[7] Rafiee R, Sotoudeh S. A cohesive zone model for predicting the initiation of Mode II delamination in composites under cyclic loading. Journal of Reinforced Plastics and Composites. 2021;40(5-6):179-92.

[8] Freed Y, Salviato M, Zobeiry N. Implementation of a probabilistic machine learning strategy for failure predictions of adhesively bonded joints using cohesive zone modeling. International Journal of Adhesion and Adhesives. 2022;118:103226.

[9] Diehl T. On using a penalty-based cohesive-zone finite element approach, Part II: Inelastic peeling of an epoxy-bonded aluminum strip. International Journal of Adhesion and Adhesives. 2008;28(4-5):256-65.

[10] Allix O, Ladeveze P, Corigliano A. Damage analysis of interlaminar fracture specimens. Composite structures. 1995;31(1):61-74.

[11] Needleman A. An analysis of tensile decohesion along an interface. Journal of the Mechanics and Physics of Solids. 1990;38(3):289-324.

[12] Mi Y, Crisfield MA, Davies G, Hellweg H. Progressive delamination using interface elements. Journal of composite materials. 1998;32(14):1246-72.

[13] Donadon MV, De Almeida SFM, Arbelo MA, de Faria AR. A three-dimensional ply failure model for composite structures. International Journal of Aerospace Engineering. 2009;2009.

[14] Tvergaard V, Hutchinson JW. The relation between crack growth resistance and fracture process parameters in elastic-plastic solids. Journal of the Mechanics and Physics of Solids. 1992;40(6):1377-97.

[15] Akhmet G, Yu Y, Hu P, Hou W-b, Han X. Analysis of the performance of adhesively bonded corrugated core sandwich structures using cohesive zone method. Journal of Sandwich Structures & Materials. 2020;22(1):104-24.

[16] Airoldi A, Novembre E, Mirani C, Gianotti G, Passoni R, Cantoni C. A model for damage and failure of carbon-carbon composites: development and identification through Gaussian process regression. Materials Today Communications. 2023;35:106059.
[17] Han X, Cai H, Sun J, Wei Z, Huang Y, Meng L. Behaviors of composite laminates under low-energy impact using a novel analytical framework. International Journal of Applied Mechanics. 2022;14(02):2250004.
[18] Jumel J, Budzik MK, Shanahan ME. Beam on elastic foundation with anticlastic curvature: Application to analysis of mode I fracture tests. Engineering Fracture Mechanics. 2011;78(18):3253-69.
[19] Cabello M, Zurbitu J, Renart J, Turon A, Martínez F. A non-linear hyperelastic foundation beam theory model for double cantilever beam tests with thick flexible adhesive. International Journal of Solids and Structures. 2016;80:19-27.
[20] Cabello M, Zurbitu J, Renart J, Turon A, Martínez F. A general analytical model based on elastic foundation beam theory for adhesively bonded DCB joints either with flexible or rigid adhesives. International Journal of Solids and Structures. 2016;94:21-34.
[21] Ereiz S, Duvnjak I, Jiménez-Alonso JF. Review of finite element model updating methods for structural applications. Structures: Elsevier; 2022. p. 684-723.
[22] Valoroso N, Sessa S, Lepore M, Cricri G. Identification of mode-I cohesive parameters for bonded interfaces based on DCB test. Engineering Fracture Mechanics. 2013;104:56-79.
[23] Jaillon A, Jumel J, Lachaud F, Paroissien E. Mode I cohesive zone model parameters identification and comparison of measurement techniques based on uncertainty estimation. International Journal of Solids and Structures. 2020;191:577-87.
[24] Xu Y, Li X, Wang X, Liang L. Inverse parameter identification of cohesive zone model for simulating mixed-mode crack propagation. International Journal of Solids and Structures. 2014;51(13):2400-10.
[25] Shen B, Paulino GH. Identification of cohesive zone model and elastic parameters of fiber-reinforced cementitious composites using digital image correlation and a hybrid inverse technique. Cement and Concrete Composites. 2011;33(5):572-85.
[26] Ribeiro Junior RF, Gomes GF. On the Use of Machine Learning for Damage Assessment in Composite Structures: A Review. Applied Composite Materials. 2023:1-37.
[27] Azad MM, Kim S, Cheon YB, Kim HS. Intelligent structural health monitoring of composite structures using machine learning, deep learning, and transfer learning: a review. Advanced Composite Materials. 2023:1-27.
[28] Jakkamputi L, Devaraj S, Marikkannan S, Gnanasekaran S, Ramasamy S, Rakkiyannan J, et al. Experimental and computational vibration analysis for diagnosing the defects in high performance composite structures using machine learning approach. Applied Sciences. 2022;12(23):12100.
[29] Ferdousi S, Advincula R, Sokolov AP, Choi W, Jiang Y. Investigation of 3D printed lightweight hybrid composites via theoretical modeling and machine learning. Composites Part B: Engineering. 2023;265:110958.
[30] Liu X, Tian S, Tao F, Yu W. A review of artificial neural networks in the constitutive modeling of composite materials. Composites Part B: Engineering. 2021;224:109152.
[31] Viotti ID, Gomes GF. Delamination identification in sandwich composite structures using machine learning techniques. Computers & Structures. 2023;280:106990.

[32] Li Y, Liu H, Zhou K, Qin H, Yu W, Liu Y. Machine learning approach for delamination detection with feature missing and noise polluted vibration characteristics. Composite Structures. 2022;287:115335.
[33] Khan A, Kim N, Shin JK, Kim HS, Youn BD. Damage assessment of smart composite structures via machine learning: a review. JMST Advances. 2019;1:107-24.
[34] Khan A, Ko D-K, Lim SC, Kim HS. Structural vibration-based classification and prediction of delamination in smart composite laminates using deep learning neural network. Composites Part B: Engineering. 2019;161:586-94.
[35] Avci O, Abdeljaber O, Kiranyaz S, Hussein M, Gabbouj M, Inman DJ. A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. Mechanical systems and signal processing. 2021;147:107077.
[36] Sikdar S, Ostachowicz W, Kundu A. Deep learning for automatic assessment of breathing-debonds in stiffened composite panels using non-linear guided wave signals. Composite Structures. 2023;312:116876.
[37] Su M, Peng H, Yuan M, Li S. Identification of the interfacial cohesive law parameters of FRP strips externally bonded to concrete using machine learning techniques. Engineering Fracture Mechanics. 2021;247:107643.
[38] Kohnke P. ANSYS theory reference, Canonsburg, PA, USA. Ansys Inc, release. 2020.
[39] de Moura MF, Moreira RD, Rocha RJ, Oliveira CF. Determination of the fracture energy under mode I loading of a honeycomb/carbon-epoxy sandwich panel using the asymmetric double cantilever beam test. Journal of Sandwich Structures & Materials. 2022;24(6):1977-92.
[40] De Moura M, Gonçalves J, Silva F. A new energy based mixed-mode cohesive zone model. International Journal of Solids and Structures. 2016;102:112-9.
[41] Vapnik V. The nature of statistical learning theory: Springer science & business media; 1999.
[42] Liu Q, Wang F, Li J, Xiao W. A hybrid support vector regression with multi-domain features for low-velocity impact localization on composite plate structure. Mechanical Systems and Signal Processing. 2021;154:107547.
[43] Lu S, Jiang M, Wang X, Yu H, Su C. Damage degree prediction method of CFRP structure based on fiber Bragg grating and epsilon-support vector regression. Optik. 2019;180:244-53.
[44] Lyu F, Fan X, Ding F, Chen Z. Prediction of the axial compressive strength of circular concrete-filled steel tube columns using sine cosine algorithm-support vector regression. Composite Structures. 2021;273:114282.
[45] Bergmayr T, Höll S, Kralovec C, Schagerl M. Local residual random forest classifier for strain-based damage detection and localization in aerospace sandwich structures. Composite Structures. 2023;304:116331.
[46] Revathy G, Rajendran V, Rashmika B, Kumar PS, Parkavi P, Shynisha J. Random forest regressor based superconductivity materials investigation for critical temperature prediction. Materials Today: Proceedings. 2022;66:648-52.
[47] Zara A, Belaidi I, Khatir S, Brahim AO, Boutchicha D, Wahab MA. Damage detection in GFRP composite structures by improved artificial neural network using new optimization techniques. Composite Structures. 2023;305:116475.
[48] Peng X, Qiu C, Li J, Wu H, Liu Z, Jiang S. Multiple-scale uncertainty optimization design of hybrid composite structures based on neural network and genetic algorithm. Composite Structures. 2021;262:113371.
[49] Probst P, Wright MN, Boulesteix AL. Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: data mining and knowledge discovery. 2019;9(3):e1301.

[50] Pal K, Patel BV. Data classification with k-fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques. 2020 fourth international conference on computing methodologies and communication (ICCMC): IEEE; 2020. p. 83-7.