Machine Learning to Predict Joint Performance in Epoxy Composites Based on Process Parameters

Brennen M. Middleton[1], Tyler B. Hudson[2], Yi Lin[3], Jin Ho Kang[3], and Frank L. Palmieri[2,]*

[1]NASA Internships, Fellowships, and Scholarships Program, NASA Langley Research Center, Hampton, VA 23681
[2]NASA Langley Research Center, Hampton, VA 23681
[3]National Institute of Aerospace, 100 Exploration Way, Hampton VA 23666

[*]frank.l.palmieri@nasa.gov

Introduction

Polymer matrix composites are widely used in the aerospace industry. Due to their high specific strength, fatigue properties, and formability, they have begun to replace the aluminum structures typically used in the manufacturing of the fuselage and wings. Although specific strength is ideal for the aerospace industry, the creation of the complex shapes necessary for the wings and fuselage often require small, simple components be assembled with adhesives to overcome manufacturing and composites processing limitations. These adhesive joints can fail unexpectedly due to the unpredictable and undetectable occurrence of weak bonds, posing a reliability concern due to unpredictable, unmeasurable strength [1]. The Federal Aviation Administration (FAA) has placed certification requirements on commercial aircraft to satisfy certain strength criteria [1-2]. Due to this, manufacturers commonly install redundant fasteners to guarantee the strength of adhesively bonded composite parts. The number of fasteners installed varies among aircraft, but it is typically on the order of 10^5 for a single-aisle commercial transport airframe. These fasteners add a tremendous cost towards implementing composite structures in the aerospace industry and reduce the benefits obtained from the specific strength of composites.

Unlike co-cure processing, which is limited to parts with simple designs, the Adhesive Free Bonding of Composites (AERoBOND) method would allow for complex parts to be manufactured without the use of adhesives. This is done by joining and curing two partially cured composite surfaces, resulting in a monolithic composite structure with no discernable interface. This potentially reduces the number of fasteners dramatically, decreasing manufacturing time and cost. However, it is challenging to perform process optimization for the AERoBOND project, as the current set of process parameters are interdependent and complex. To assist with this process optimization, several machine learning (ML) algorithms were created to provide predictions for mechanical performance and characterization properties of the composite parts.

Experimental

Raw experimental data from fifteen previous AERoBOND experiments were utilized in the ML algorithms for analysis. Python™ 3.7 was utilized to perform all necessary preprocessing steps for their utilization in the ML algorithm scripts. The ML algorithms utilized Python™ 3.7 with XGBoost, scikit-learn, and scikit-image python modules.

The ML algorithms used gradient boosting decision trees with hyperparameter tuning to provide predictions of the mechanical and characterization properties of the composite materials. The ML algorithms were trained and tested utilizing an 80:20 training to testing split during model creation. Input parameters for the algorithms were data collected from the composite fabrication process [i.e., number of base laminate plies, fiber areal weight (FAW), ratio of molar equivalents of hardener to epoxy, primary and secondary cure cycle temperature-time piecewise functions (i.e., primary and secondary ramps and holds), and Fourier transform infrared (FTIR) spectroscopy data, etc.], leading to 29 input variables. Output parameters were non-precracked (NPC) and precracked (PC) mode II fracture toughness (G_{IIc}) values from end-notched flexure (ENF) testing, and time-of-flight (ToF) and amplitude ultrasonic C-scan images.

When hyperparameter tuning constraints were met by the ML algorithms, accuracy scores were presented, as well as a feature importance graph of the model. Feature importance graphs portray the input parameters utilized in decision tree nodes, or forks in the decision tree where data is separated to provide different predictions. The F score on this graph denotes the number of instances these parameters are used for a decision tree node. These accuracies and feature importance graphs allow the user to determine if the models are accurate and discard models that are overfitting or underfitting the data presented to the ML algorithms. Overfitting of the data can be assessed from the feature importance graph if a high number of input parameters are identified and the F scores are large relative to the number of total input parameters available. For example, overfitting is likely if the decision tree utilizes 19 of 29 input parameters and the F score for the input parameter of highest importance is 53. Conversely, underfitting can be identified if a low number of input parameters are identified relative to the total available input parameters and F scores are small. For example, underfitting is likely if the decision tree utilizes five input parameters out of 29 with a F score of three for the parameter of highest importance. If the model was determined to be accurate, the validity of the model was tested through a validation test.

Validation of the model was done using input parameters that have not been used in the training or testing set when creating the model. This process allows the user to determine if the model can predict output values outside the model's perspective. For the ongoing AERoBOND project, additional datasets were not yet available for the validation stage. Due to this, the values used in training and optimization were utilized by averaging similar experimental sets together to create this validation set. The model was then tested with the validation set to make predictions for this data. This process allowed the user to determine if the model could predict these values, as these average values were not inherently used during training and provided values that approximated future experimental data.

Results and Discussion

Approximately 200 random states, or model variations, were tested within the ML algorithms for each output parameter. Models that met the hyperparameter tuning constraints were then validated. The models from each ML algorithm were then analyzed based on their accuracies, feature importance graph, and the validation process to determine their overall accuracy in predicting model outputs. Models were selected if they met the hyperparameter tuning constraints, showed no underfitting or overfitting in the feature importance graph, and were determined to be accurate based on their accuracy scores and validation process

Figure 1, found in the appendix, shows the feature importance for the ENF NPC model. This figure shows no signs of underfitting or overfitting of the data. Additionally, this figure denotes the average FAW of panel 1 of the two epoxy rich panels as the parameter of most importance in the model.

Table 1 denotes the model predictions of the testing portion of the 80:20 split of the experimental data. Table 2 denotes the model predictions of the validation data.

Table 1. ENF NPC Model Testing Predictions

Experimental Value G_{IIc} (in-lb/in^2)	**Predicted Value** G_{IIc} (in-lb/in^2)	**Absolute Difference** G_{IIc} (in-lb/in^2)
2.83	2.07	0.76
5.46	5.47	0.01
0.71	0.30	0.41

Table 2. ENF NPC Model Validation Predictions

Experimental Value G_{IIc} (in-lb/in^2)	**Predicted Value** G_{IIc} (in-lb/in^2)	**Absolute Difference** G_{IIc} (in-lb/in^2)
5.28	4.28	1.00
0.54	2.00	1.46
5.41	4.79	0.62
2.08	3.56	1.48

Table 1 shows small differences in the testing and prediction outputs, while Table 2 shows slightly higher absolute differences between the validation and prediction outputs.

Figure 2, found in the appendix, shows the feature importance graph of the ENF PC model. This figure shows no signs of underfitting or overfitting the data. The height of the average CO stretch peak (823 cm^{-1}), identified in the normalized FTIR data, was highly important in the model.

Tables 3 and 4 denote the ENF PC model predictions of the testing portion of the experimental data and of the validation data, respectively. Both tables show small differences between their experimental and model prediction values, with all absolute differences below 1.

Table 3. ENF PC Model Testing Predictions

Experimental Value G_{IIc} (in-lb/in^2)	**Predicted Value** G_{IIc} (in-lb/in^2)	**Absolute Difference** G_{IIc} (in-lb/in^2)
0.00	0.01	0.01
3.22	3.20	0.02
1.60	0.93	0.67

Table 4. ENF PC Model Validation Predictions

Experimental Value G_{IIc} (in-lb/in^2)	**Predicted Value** G_{IIc} (in-lb/in^2)	**Absolute Difference** G_{IIc} (in-lb/in^2)
3.17	3.25	0.08
0.00	0.24	0.24
4.07	3.23	0.84
1.36	1.96	0.60

Figure 3, found in the appendix, shows the feature importance for the ToF C-scan model. This figure shows no signs of underfitting or overfitting of the data. Additionally, this figure denotes the "Unknown Peak" (peak with no assigned stretch at 1150 cm^{-1}), an FTIR peak identified in the experimental data, as having the highest importance in the model.

Tables 5 and 6 denote the ToF C-scan model predictions of the testing portions of the experimental data and for the validation data, respectively. Experimental data from these tables is represented as the fraction of pixels indicating good bonds (i.e., no voids or delamination) in the C-scan images. Pixels indicating good bonds were determined during the preprocessing step. Grayscale images were analyzed and the summed pixels indicating a good bond (having light coloration), are divided over the total pixels in the image, creating the fraction of good bonds in the laminate. Therefore, fractions closer to unity indicate a well bonded panel.

Table 5. ToF C-scan Model Testing Predictions

Experimental Value	Predicted Value	Absolute Difference
0.95	0.86	0.09
0.53	0.65	0.12
0.91	0.91	0.00

Table 6. ToF C-scan Model Validation Predictions

Experimental Value	Predicted Value	Absolute Difference
0.56	0.65	0.09
0.86	0.88	0.02
0.78	0.79	0.01
0.92	0.91	0.01

Both Table 5 and Table 6 show small differences between the experimental and model prediction values, with the second experimental value in Table 6 showing a slightly higher absolute difference relative to other experimental values.

Figure 4, found in the appendix, shows the feature importance for the Amplitude C-scan model. This figure shows no signs of underfitting or overfitting of the data. The CO stretch peak is denoted as the parameter of highest importance for this model.

Tables 7 and 8 denote the amplitude C-scan model predictions of the testing portions of the experimental data and for the validation data. Experimental data from these tables is represented as the fraction of pixels indicating good bonds (having light coloration) based on analysis of grayscale C-scan images.

Table 7: Amplitude C-scan Model Testing Predictions

Experimental Value	Predicted Value	Absolute Difference
0.45	0.54	0.09
0.03	0.16	0.13
0.54	0.54	0.00

Table 8: Amplitude C-scan Model Validation Predictions

Experimental Value	Predicted Value	Absolute Difference
0.23	0.25	0.02
0.36	0.47	0.11
0.57	0.64	0.07
0.32	0.34	0.02

Both Table 7 and Table 8 show small absolute differences between the experimental and model predictions, with the second experimental values in both tables showing slightly higher absolute differences than other experimental values in their relative tables.

Table 9 denotes the accuracy scores for each of the four models. These accuracy scores are based on the ability of the model to make accurate predictions of the testing portion of the 80:20 split of the experimental data. These accuracy scores were calculated utilizing the accuracy_score() function within the scikit-learn python module [3]. Data from tables 1, 3, 5, and 7 were utilized in this function to calculate the accuracy scores.

Table 9. Model Accuracy Scores

Model	Accuracy Score
ENF NPC	93%
ENF PC	91%
ToF C-scan	80%
Amplitude C-scan	84%

For all models, feature importance showed no signs of overfitting or underfitting the data. Additionally, the feature importance shows all models utilized data in their decision trees from the composite layup, cure cycle, and FTIR data categories. Additionally, all models had an accuracy score greater than or equal to 80%. All models showed small absolute differences between experimental and model prediction values, with three validation absolute difference values for the ENF NPC model above 1. However, all models showed high rates of predictability with accuracy during testing, and the ENF PC, ToF C-scan, and Amplitude C-scan models showed high rates of predictability with accuracy during validation. Additionally, it should be noted that increasing the sample size of the experimental data could potentially increase the accuracies of the models, and the performance of the ENF NPC model during the validation phase has the potential to improve in accuracy and reliability with the addition of experimental data.

Conclusions

After analysis, the ENF PC, ToF C-scan, and Amplitude C-scan models were shown to determine predictions from input parameters with accuracy. The ENF NPC model showed slight inconsistencies during the validation phase of testing, but performed with accuracy during the model testing phase. It was concluded that models created from the ML algorithms could accelerate a parametric study for the AERoBOND process by rapidly optimizing process parameters to achieve desired performance characteristics.

Acknowledgements

We thank Charles Liles and Tanner Griffen for support with accessing and instruction in machine learning software tools and methods.

References

1. F.L. Palmieri, T.B. Hudson, R.J. Cano, et al. Reliable

bonding of composite laminates using reflowable epoxy resins, *Proceedings of the 2019 International SAMPE Tech. Conf.* Seattle, WA, 2019 Society for the Advancement of Materials and Process Engineering.
2. Federal Aviation Administration, 14 CFR § 25.301-307.
3. Scikit-learn, 3.3. Metrics and scoring: quantifying the quality of predictions, from https://scikit-learn.org/stable/modules/model_evaluation.html.

Appendix

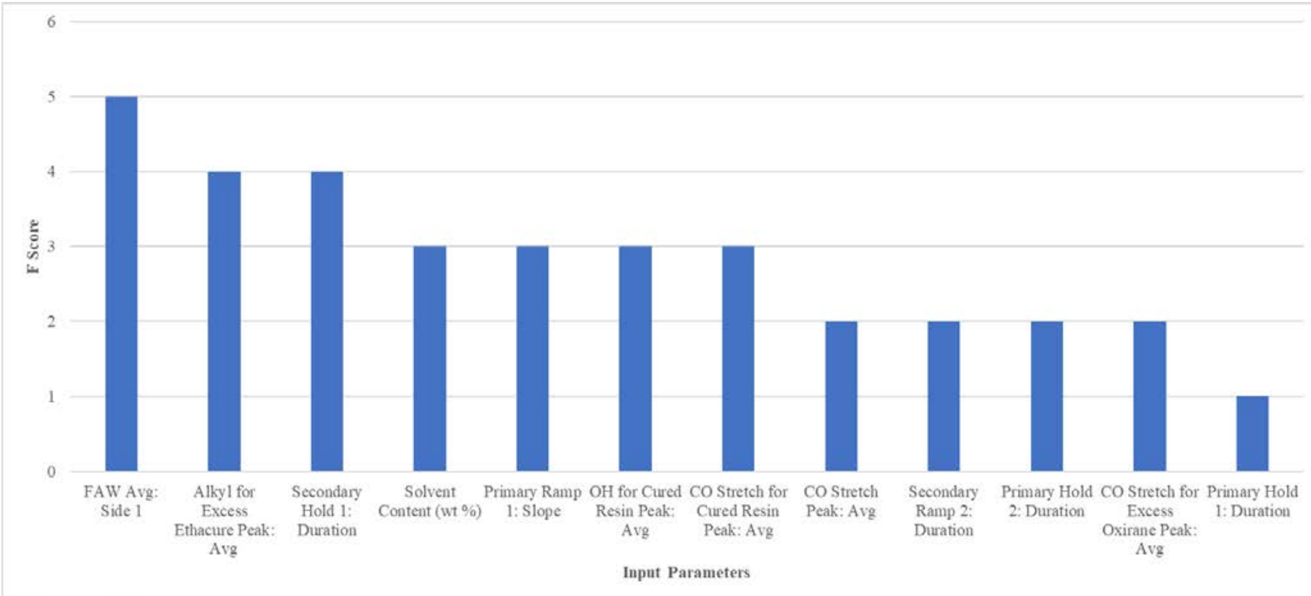

Figure 1: ENF NPC Feature Importance

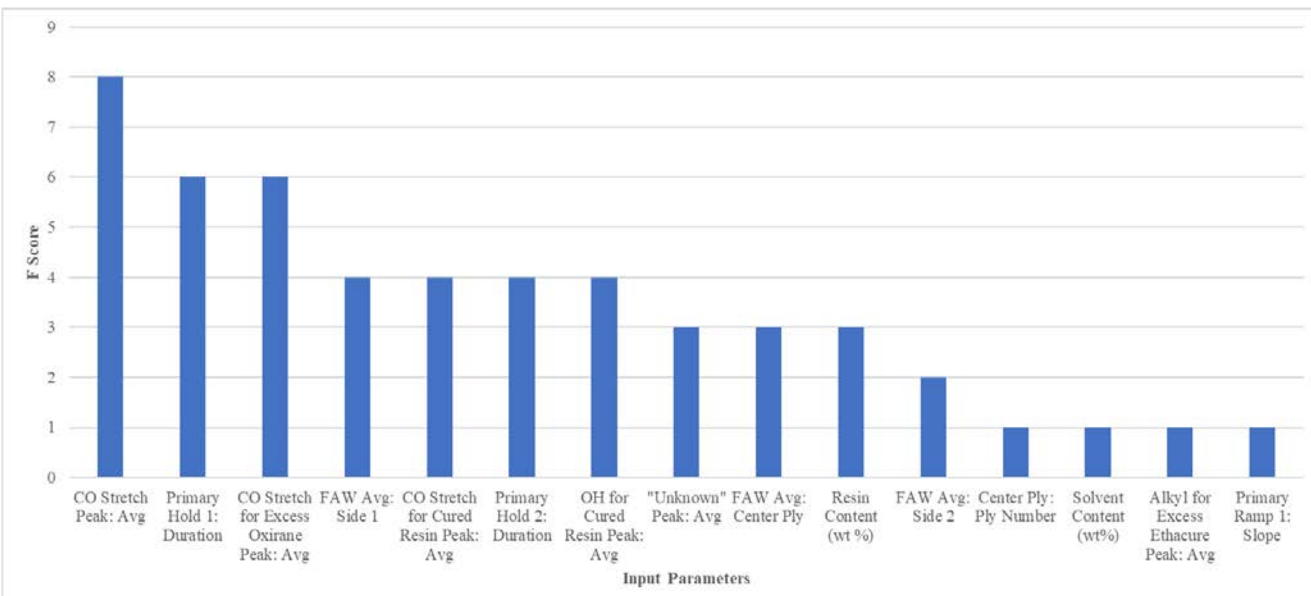

Figure 2: ENF PC Feature Importance

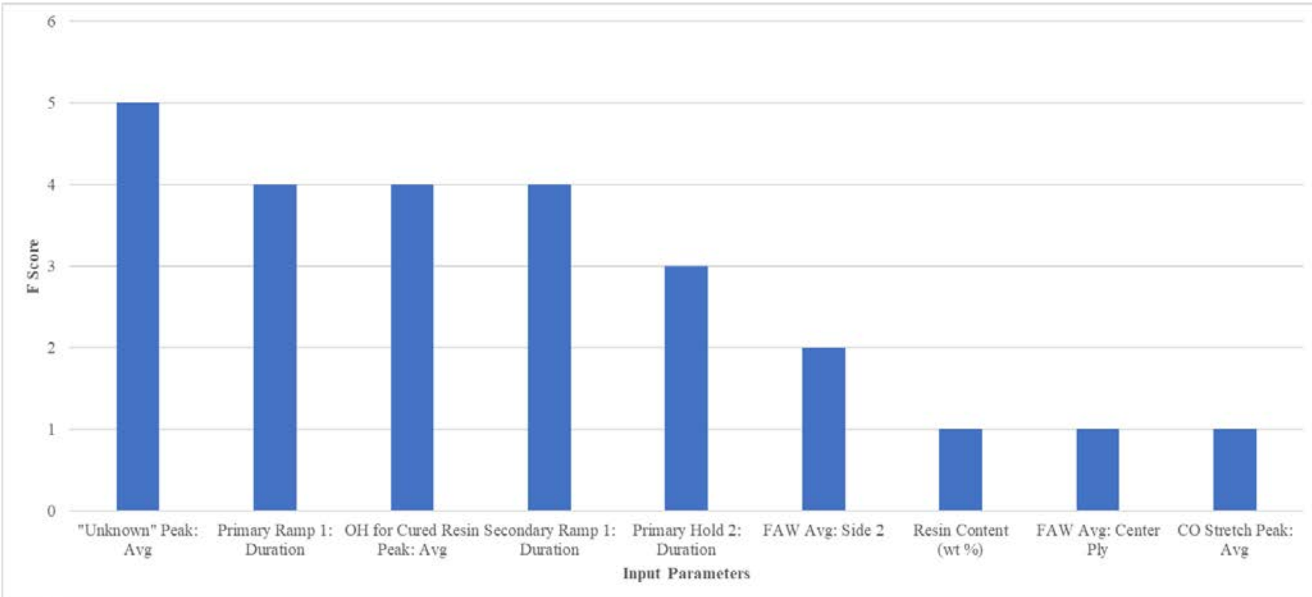

Figure 3: ToF C-scan Feature Importance

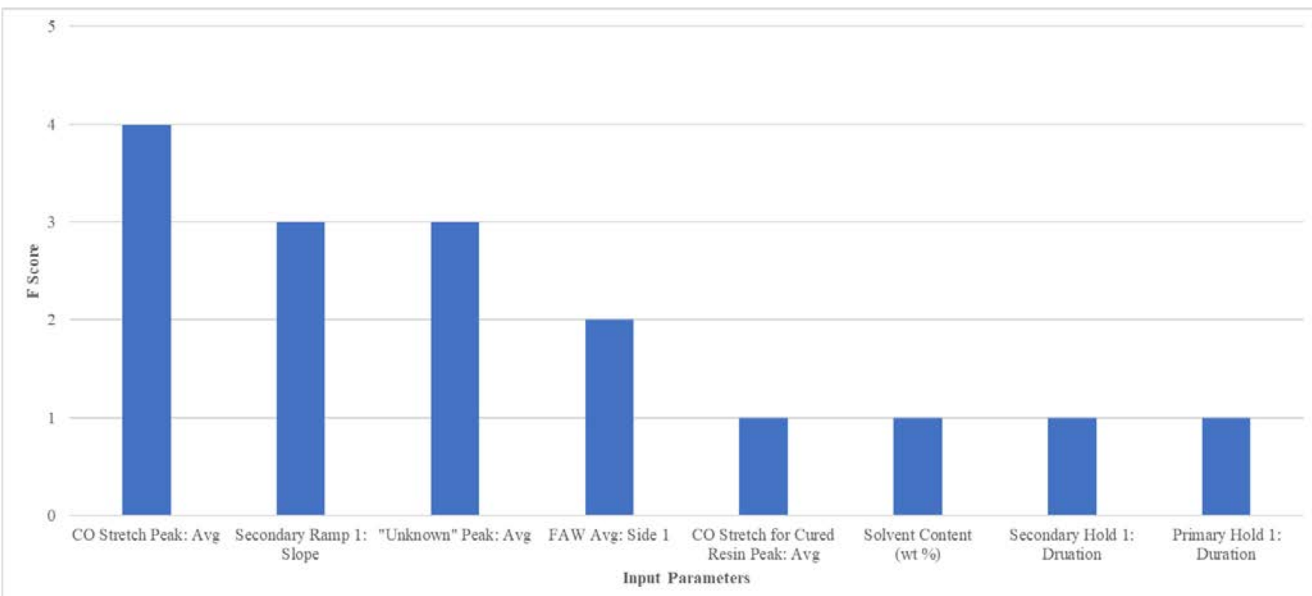

Figure 4: Amplitude C-scan Feature Importance