Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

Theory-guided machine learning for optimal autoclave co-curing of sandwich composite structures

Tania Lavaggi[1], Mina Samizadeh[2], Navid Niknafs Kermani[1], Mohammad Mahdi Khalili[2], and Suresh G. Advani[1]

[1]Center for Composite Materials and Department of Mechanical Engineering, University of Delaware, Newark, Delaware, USA
[2]Department of Computer and Information Sciences, University of Delaware, Newark, Delaware, USA

Correspondence
Suresh G. Advani, Center for Composite Materials and Department of Mechanical Engineering, University of Delaware, Newark, DE, USA.
Email: advani@udel.edu

Abstract

The bonding of a honeycomb core to the thermoset prepreg facesheets by co-curing them allows one to manufacture composite sandwich structures in a single operation. However, the process is strongly dependent on the prescribed autoclave cure cycle. A previously developed physics-based simulation can predict the bond quality as a function of the process parameters. The disadvantage of physics-based simulations is the high computational effort needed to identify the optimal cure cycle to fabricate sandwich structures with desired bond-line properties. Theory guided machine learning (TGML) methods have demonstrated their capabilities to reduce the computational effort for different applications. In this work, three TGML models are trained on a data set produced from physics-based simulations to predict the co-cure process of honeycomb sandwich structures. The accuracy of the TGML models were compared to select the best performing predictive tool. In addition to reduction of computational time by orders of magnitude, we demonstrate how the TGML tools can also quantify the contribution of each process parameter on the properties of the fabricated part. The most accurate model was implemented in an optimization routine to tune the input process parameters to obtain the desired properties such as the bond-line porosity and facesheet consolidation level. This methodology could be extended to any process simulation of composites manufacturing processes.

Keywords

composites, computer modeling, curing of polymers, modeling, voids

1 | INTRODUCTION

Honeycomb sandwich structures are light and stiff materials that provide high mechanical performance for aerospace and automotive applications.[1,2,3] They are fabricated by two skins (facesheets), consisting of thermoset pre-impregnated fabrics, or prepreg, bonded to a honeycomb core (Figure 1). Different methods have been used to bond the skins to the core. In the secondary bonding manufacturing process, first, the prepreg facesheets are consolidated. This is followed by using an adhesive film to bond the consolidated facesheets to the honeycomb core in a separate operation. Thus, the process occurs in several steps, which prolong the manufacturing cycle and moreover, introduce difficulties in fitting the consolidated facesheet to the more complex, curved, geometries. The co-cure process involves simultaneous (i) consolidation of facesheets; (ii) bonding process by a film adhesive and (iii) concurrent

Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

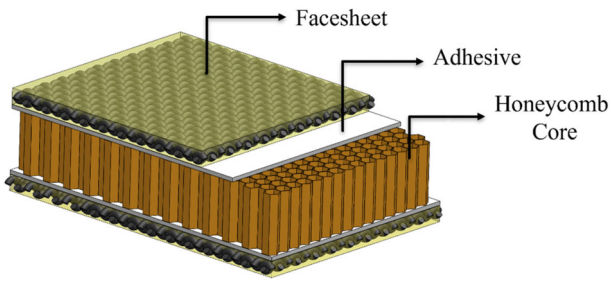


FIGURE 1 Honeycomb composite sandwich structures

curing of all materials. In this case, the fabrication is achieved in a single operation. Also, the geometrical mismatch between the skins and the core is no longer a hurdle due to the higher flexibility of the uncured facesheets.[4,5]

Autoclave processing is one of the most common manufacturing methods adopted to fabricate honeycomb sandwich composites, either by the secondary bonding or the co-cure process. In the latter case, the uncured prepreg sheets and the adhesive are placed on one or both sides of the honeycomb core and sealed under a vacuum bag subjected to a prescribed temperature and pressure cycle to co-cure the prepregs and the adhesive bonding the prepregs to the core. The co-cure process cycle in the autoclave is described by setting the pressure and temperature in the autoclave and the pressure in the vacuum bag over time. Two configurations are used for the co-cure process (i) equilibrated or (ii) closed core. In the first case, the prepreg facesheets are co-cured on one side of the core, while, on the other side, the core is connected to a precured perforated tool. Therefore, during the process, the core pressure equals the pressure inside the vacuum bag. However, in the closed core configuration, the honeycomb core is simultaneously bonded to the skins on both sides.

The quality of the sandwich structures is highly influenced by the quality of the bonding between the skins and the honeycomb core. During the co-cure, the interactions between the polymer resin and the adhesive resin may lead to an unacceptable amount of porosity in the bond-line, if the autoclave process parameters are not set properly.[6,7] The typical strategy for the design of the autoclave cure cycle is to adopt process parameters suggested by the materials supplier, although this approach is not optimal and requires multiple trials. This is because of the dominant coupled physical phenomena occurring during the co-cure process such as the bond-line fillet formation, prepreg facesheet consolidation and core pressure dynamics may result in an excessive presence of porosity within the bond-line weakening the bond between the facesheet and the core.[8,9] In a co-cured honeycomb sandwich structure, the bond-line is composed by the adhesive polymer and the resin that bleeds out of the prepreg skins due to consolidation and the applied pressure gradient. Thus, the cure cycle must be properly tuned to (i) avoid large number of voids within the bond-line, and (ii) to guarantee the desired level of consolidation.

Numerical simulations with physics-based tools have been developed in the composite manufacturing field to predict the effect of the process parameters on the final part in order to optimize the fabrication method.[10,11,12] Niknafs Kermani et al. developed the physics-based model that accounts for the bond-line fillet formation, prepreg facesheet consolidation and void dynamics during the co-cure process and have implemented in a software tool *SANDWICH* to simulate the co-cure of honeycomb sandwich structures with a specified autoclave cure cycle. The simulation inputs in addition to material rheology are the pressure and temperature in the autoclave and the vacuum pressure as a function of time. The simulation predicts the fiber volume fraction and resin pressure profile across the thickness of the facesheet, the adhesive fillet shape, the core pressure, the bond-line porosity, and volume of resin bled out of the prepreg during the equilibrated co-cure.[9] Zebrine et al.[13] performed validation experiments to assess the consistency of the predictions by *SANDWICH*. They fabricated three specimens by equilibrated co-cure with different processing conditions in terms of autoclave and vacuum pressures, and number of layers of prepreg. The manufactured samples were characterized to evaluate the porosity content and the fillet shape. Simulations by *SANDWICH* with the same conditions were performed to compare the results. The experiments confirmed that the tool *SANDWICH* correctly captures the trends of the response for the adhesive fillet shape and the bond-line porosity in response to the changes to the cure cycle parameters. Niknafs et al.[14] integrated an optimization routine to the software to find the optimal autoclave process parameters to fabricate a honeycomb sandwich composite with minimum content of porosity in the bond-line and maximum consolidation level of the skins. The multiobjective optimization routine consisted of one cost function in which the simulation outcomes were embedded. The optimal parameters were identified by performing several physics-based simulations by *SANDWICH*, to minimize the cost function. The typical time necessary to complete one physics-based simulation by *SANDWICH* is an average of 380 s.[15] Thus, identifying the optimal cure cycle is limited, as the number of runs to ensure the optimum result is of the order of thousands and with the physics-based computational cost, the necessary time would take months on a traditional PC.

In the last decade, machine learning (ML) methods became a viable tool to extract useful information from large amounts of data. It is more common to substitute

science-based approaches, relying on numerical methods such as finite element (FE) or finite differences (FD) simulations to solve the governing physical laws, with data-driven modeling techniques. Several studies explored the use of ML methods for processing of composite materials.[16] Ozkan et al.[17] predicted the mechanical properties of nanocellulose films by ML. They demonstrated how ML tools help in identifying combinations of the components to tailor the nanocellulose composites to achieve higher performances. ML methods are also used to predict the temperature differences and the degree of cure during compression molding of thermosetting prepregs. Hou et al.[18] trained several ML algorithms on a data set produced by FE simulations of the process in order to provide a tool with low computational cost for the prediction of the effect of several process parameters on the outcome of the compression molding process. The accuracy of the predictions by their ML tools was affected by the availability of limited data set.

Two strategies can be adopted to create a ML tool to predict the desired outputs of a manufacturing process (i) conducting a large enough set of experiments so the ML algorithms can capture the possible correlations between inputs and outputs; or (ii) using physics-based models implemented in simulations to generate a data set to train the ML models. In such cases, the ML algorithm is informed of the process physics through the experienced phenomena recorded in the data set. The first strategy is clearly not viable, since composites material and their manufacturing process are expensive and time consuming. Therefore, new studies involving the second strategy investigate the application of theory-guided ML (TGML) tools for the composite materials processing.[19] TGML tools have been developed to calibrate a continuum damage FE model to predict the intralaminar progressive damage for quasi-isotropic composite laminates.[20]

In this paper, we present TGML models to predict the co-cure process of honeycomb sandwich structures. Several TGML algorithms are trained on a data set produced from physics-based simulations of the co-cure process to perform regressions, and their performances are compared. The best regressor is selected to predict the simulation results of the co-cure process. The data set consists of 5056 points. The simulations are performed with the model developed by Niknafs Kermani et al.[9] and are considered as the actual data for the training, testing, and validation of the TGML models. Twelve process parameters are defined to uniquely characterize the selected cure cycles and are used as features for the TGML regressors. This is followed by integrating the models with an optimization routine to obtain the optimal process parameters. In this study, the predicted output for a specific cure cycle is represented by the maximum fiber volume fraction and minimum bond-line porosity developed during the cure. The multi-objective optimization process identifies the process parameters that result in co-cured honeycomb sandwich structures with minimum bond-line porosity and highest consolidation level with highest possible pressure that can be applied to the part without crushing the core. Unlike previous studies that evaluated limited number of cure cycles in the optimization process due to the computational limitations, 1000 cure cycles are evaluated in order of minutes using the TGML method.

2 | DEVELOPMENT OF THE DATA SET

2.1 | Materials and process

The prepreg facesheets selected for this study are Hexcel 8225, and the adhesive is Henkel Loctite EA 9658 AERO. The software *SANDWICH* is used to simulate the co-cure process for the honeycomb sandwich structures in the equilibrated configuration.[9] The physical parameters used to generate the data points for adhesive thickness, number of prepreg layers, and honeycomb core cell size are in agreement with the validation experiments performed by Zebrine et al.[13]

2.2 | Data set structure

The data set for this study is generated by simulating different cure cycles as the input for the physic-based simulation, *SANDWICH*. The cure cycles are designed with two temperature ramps and two temperature dwells and characterized by six parameters to describe the temperature profile and six parameters to describe the pressure profiles in the autoclave and in the vacuum bag.[9] As shown in Figure 2, the temperature cycle is defined by the two heat rates (x_1, x_4), the two ramp durations (x_2, x_5) and the two dwell durations (x_3, x_6). The vacuum cycle is described by the vacuum pressure (x_7), the time of application of vacuum (x_8) and the duration of the vacuum dwell (x_9), and similar parameters are defined for the autoclave pressure profile. For this study, the pressures are assumed to change only once during the cure.

To generate the data set, 5056 different combinations of the 12 inputs parameters are designed to reasonably cover several possible physical scenarios in the co-cure process and offer to the TGML regressors a comprehensive description of the phenomena involved in the

Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

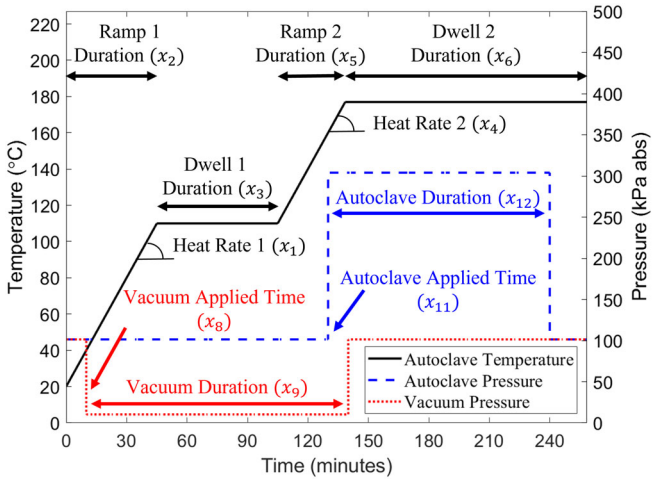


FIGURE 2 Selection of the parameters to describe the temperature and pressure cycles in the autoclave and in the vacuum bag during the co-cure process

TABLE 1 Parameters of the temperature cycle selected to generate the input for the simulation *SANDWICH* and the data set for the training of the machine learning algorithms

	x_1 [°C/min]	x_2 [min]	x_3 [min]	x_4 [°C/min]	x_5 [min]	x_6 [min]
Temperature cycle	2	45	60	2	33.5	120
			100			
	3	30	60	3	22.5	120

TABLE 2 Pressure values selected to generate the input for the simulation, *SANDWICH* and the data set for the training of the machine learning algorithms

Vacuum pressure x_7 [Pa]	$0.01P_{atm}$	$0.1P_{atm}$	$0.5P_{atm}$	P_{atm}
Autoclave pressure x_{10} [Pa]	P_{atm}	$2P_{atm}$	$3P_{atm}$	$4P_{atm}$

co-curing process. A large number of data points guarantees the validity of the predictions and the robustness of the regressors, but the generation of the data is limited by the computational effort required by the machine. Indeed, the 5056 simulations were performed in about 500 h.

Three temperature cycles with different heating ramps and different hold durations are combined with five different autoclave and vacuum pressures. Five different start times and duration of the pressure cycles are also considered. To generate the various input cure cycles, permutations of the 12 parameters are used by fixing one temperature cycle and one of the vacuum or autoclave cycle parameters. The considered temperature cycles and vacuum and autoclave pressures are listed in Tables 1 and 2.

As shown in Table 1, the two cure cycles with heating ramps of 2°C/min are designed to test the influence of the temperature dwell on the outputs of the co-cure. The third cycle is designed with higher heating rates (x_1, x_4) and different values for heating ramp duration (x_2, x_5).

Four pressure values are used, for both the vacuum bag and the autoclave, to generate the input cycles. The remaining input parameters, that is, vacuum and autoclave applied times (x_8, x_{11}) and durations (x_9, x_{12}), are designed separately for the different temperature cycles by considering the prepreg and adhesive gel times.

The outputs of the simulations for the data set are the porosity of adhesive (φ_{AD}) and bled resin (φ_{PR}), effective porosity of the bond-line (φ_{EFF}), and maximum fiber volume fraction (FVF). Thus, the data set consists of data points characterized by 12 inputs and four outputs. The arrangement of the input and output parameters in the data set is presented in Figure 3.

The data points are then fed to the machine learning models for training and validation to generate the predictions. A schematic of the data generation and processing is presented in Figure 4.

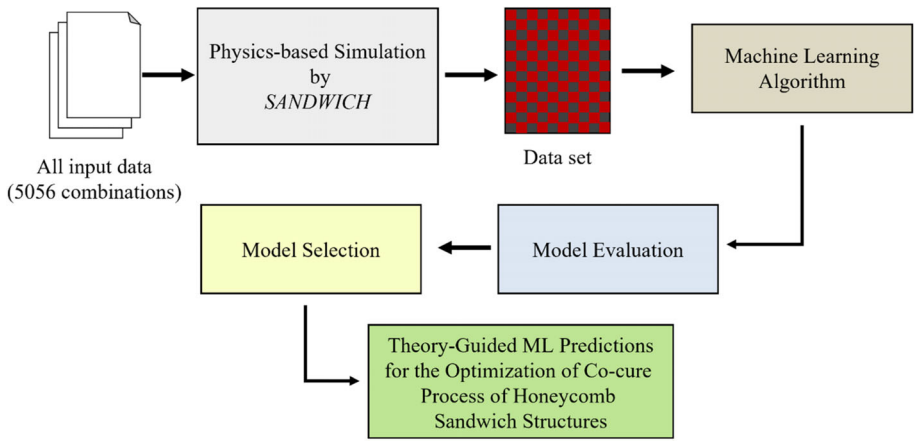
Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

FIGURE 3 Definition of the input and output variables for the data set

Inputs											
Temperature Profile						Vacuum Pressure Profile			Autoclave Pressure Profile		
Heat Rate 1 [°C/min]	Ramp 1 Duration [min]	Dwell 1 Duration [min]	Heat Rate 2 [°C/min]	Ramp 2 Duration [min]	Dwell 2 Duration [min]	Vacuum Pressure (VP) [Pa]	Vacuum Applied Time (VAT) [min]	Vacuum Duration (VD) [min]	Autoclave Pressure (AP) [Pa]	Autoclave Applied Time (AAT) [min]	Autoclave Duration (AD) [min]
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}

Outputs			
φ_{AD}	φ_{PR}	φ_{EFF}	FVF
y_1	y_2	y_3	y_4

FIGURE 4 Schematic of the implemented methodology for the theory-guided machine learning framework for the optimization of the co-cure process



3 | MACHINE LEARNING MODELS FOR THE REGRESSION

To predict the four outputs, three different TGML algorithms are trained on the generated data set (Figure 4). These models include decision tree (DT), random forest (RF), and extremely randomized trees (ERT). Four different regressors are trained by each of the three supervised ML models to predict each output. A total of 12 TGML regressors, four for each ML method, are generated and 10-fold cross validation is performed to evaluate them. The performances of the three ML models for each output are compared by evaluating the average root mean squared error (RMSE) of the 10-fold cross validation. Then, the four regressors for the four outputs of each ML model that allow the predictions of each of the outputs with the highest accuracy are selected to be integrated in the optimization routine. Figure 5 shows a schematic of the generation of the TGML regressors for the four outputs.

3.1 | Decision tree

Decision tree (DT) is a supervised learning method widely used for data regression. A DT consists of decision nodes and leaves. The DT algorithms grows a connected branch between the decision nodes and it is arrested at the leaves. To grow the branch, the feature space is partitioned in cells and the branches of the DT are grown by decisions made by the features on how to split the data based on the value that the nodes can assume. The split criteria to grow a DT are learned by the model from statistical measurements of features in the training data.[21]

In this study, four DT regressors are generated for the four outputs. We set the squared error as a function to measure the quality of splits, and all the other parameters are the default parameters of the "DecisionTreeRegressor" function in the SciKit Learn Python 3.0 library. The algorithm is also set to select the best split at each node and all of the models are trained with the same parameters.

3.2 | Decision tree based ensemble methods: Random Forest

Ensemble methods in machine learning use several base learners to construct a regressor that generates the final decision based on a voting mechanism. The results of the base models are aggregated

Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

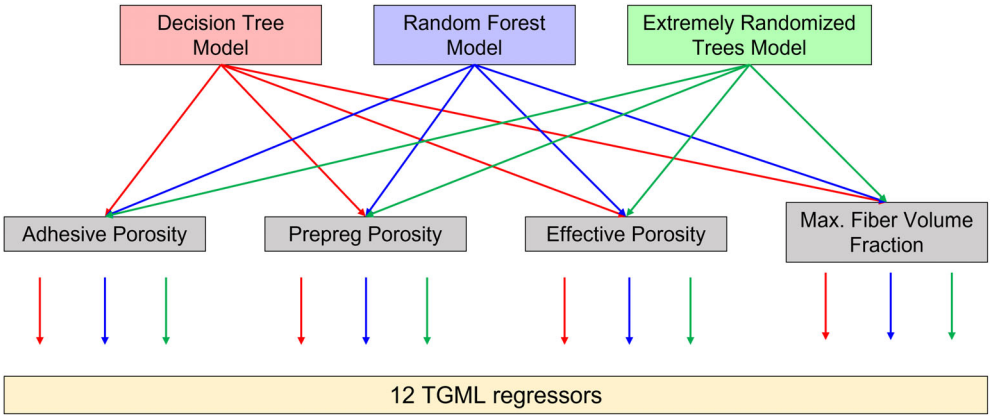


FIGURE 5 Schematic of the training of the 12 theory guided machine learning regressors. Three ML models are selected and trained on the four outputs separately. Therefore, for each output three TGML regressors are generated and compared by 10-fold cross-validation

to make more accurate predictions. DT are typically used as weak learner. For this study, two ensemble methods are implemented to develop the regressors.

Random forest (RF) is a tree-based ensemble method, which consists of a large number of DTs creating a forest in an ensemble. Each DT uses a random subset of the available features and a random sample set of the data set to build the model. The predictions of each of the trees in the forest are then averaged by arithmetic mean to perform the regression.[22]

In this work, four RF regressors are developed, each of them built with the same parameters. The built-in function "*RandomForestRegressor*" of the library Scikit Learn in Python 3.0 is used to generate the four models. The algorithm is set to include 500 DT in the forest. The mean squared error is used as a function to evaluate the quality of a split.

3.3 | Decision tree based ensemble methods: Extremely randomized trees

Extremely randomized trees (ERT) is an ensemble method similar to RF. Several DTs as part of an ensemble are trained on the whole data set and the splitting point of features at each node is randomly selected.[23] For our study, we generated four ERT regressors, one for each of the outputs, each with 100 DT as part of the forest. We implemented the models with the built-in function "ExtraTreeRegressor" of the Scikit Learn library in Python 3.0. The mean squared error was used as a function to evaluate the quality of the splits.

3.4 | Ten-fold cross validation and RMSE

Cross-validation is a technique used to evaluate the performances of a ML algorithm. In this study, 10-fold cross-validation is adopted. The data set is split in 10 folds, where 10% of the data is randomly selected to be assigned in the test fold and the remaining data is split in the nine training folds. After the training, the models are validated on the test fold. The process is iterated 10 times to capture a robust performance error and at each step the accuracy of the model is evaluated by means of percent RMSE:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (x_i - x_i^*)^2}{N}} * 100 \tag{1}$$

where x_i and x_i* represent, respectively, the actual output value in the data set and the predicted value by the ML model, over the N data points. At the end of the procedure, the RMSEs of the 10 iterations are averaged and the mean value and standard deviation are computed. The process is performed for the 12 regressors that are generated for the outputs, in four groups of three regressors each relative to one output, in order to compare the performances of the TGML models.

4 | OPTIMIZATION

To construct the optimization routine, we assumed that the facesheet consolidation level is represented by the highest fiber volume fraction (FVF) across the facesheets. Therefore, we built a cost function (Equation 2) for the optimization as a linear combination of φ_{EFF} and FVF. The optimization algorithm will proceed to minimize the cost function with the iterations. Therefore, to guarantee that FVF is maximized, the term related to it in the cost function is built with the form 1-FVF. The range of possible values for the porosity φ_{EFF} are between 0 and 1, while the span for the FVF is usually between 0.55 and 0.59. To equalize the sensitivity of the cost function to the two parameters, the term 1-FVF is normalized with

VF_{max} and VF_{min}, that are the expected highest or initial fiber volume fraction, respectively, in the simulation depending on the selected process parameters (x).

$$f_{(x)} = w_1 \varphi_{EFF(x)} + w_2 \left(\frac{VF_{max} - FVF_{(x)}}{VF_{max} - VF_{min}} \right) \tag{2}$$

The weight functions w_i, with $i = 1, 2$ and $\sum w_i = 1$, define the sensitivity of the cost function to the two output parameters. The objective function is highly non-linear and non-convex, as both $\varphi_{EFF(x)}$ and $FVF_{(x)}$ are computable by means of non-linear functions (finite differences or finite element methods). The optimization parameters are described by vector (x) as follows:

$$x = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}] \tag{3}$$

where the parameters x_i describe the cure cycle as defined in section 2.

Several inequality constraints are set to guarantee the appropriate search for the optimal process parameters. In this study, the total time of the cure cycle is defined by the temperature cycle. Therefore, the vacuum and autoclave applied times (x_8 and x_{11}, respectively) are bounded by the temperature cycle. The vacuum and autoclave release times ($x_8 + x_9$ and $x_{11} + x_{12}$, respectively), are bounded by the total time for the cure cycle as well. Moreover, the vacuum and autoclave release times cannot happen before the vacuum and autoclave applied times. Such specifications result in the following constraints implemented in the optimization code:

$$f_{1(x)} = x_8 - x_2 - x_3 - x_5 - x_6 \leq 0 \tag{4}$$

$$f_{2(x)} = x_{11} - x_2 - x_3 - x_5 - x_6 \leq 0$$

$$f_{3(x)} = -x_9 \leq 0$$

$$f_{4(x)} = x_8 + x_9 - x_2 - x_3 - x_5 - x_6 \leq 0$$

$$f_{5(x)} = -x_{12} \leq 0$$

$$f_{6(x)} = x_{11} + x_{12} - x_2 - x_3 - x_5 - x_6 \leq 0$$

The non-linear constraints are set to guarantee that both polymers of adhesive and prepreg are properly cured when a trial optimal cycle is generated by the optimization routine. A check on the degree of cure for both materials is implemented and the values are compared with the degree of cure when the resin or adhesive gels.

$$f_{7(x)} = \alpha_{AD,Gel} - \alpha_{AD} \leq 0 \tag{5}$$

$$f_{8(x)} = \alpha_{PR,Gel} - \alpha_{PR} \leq 0$$

where α_{AD} and α_{PR} represent the final degrees of cure for the adhesive and prepreg resin respectively, and $\alpha_{AD,Gel}$ and $\alpha_{PR,Gel}$ are the degrees of cure when the gelation for adhesive and prepreg will initiate respectively.[14]

For the optimization, upper and lower bounds for the process parameters are defined considering that autoclaves typically operate pressures up to 400 kPa and temperatures up to 180°C for the co-cure of honeycomb sandwich structures.[8,24] The algorithm for the optimization is implemented in MATLAB as the Surrogate Optimization function "surrogateopt." The function provides a global search for the optimum value for functions with high computation costs. The algorithm is arrested after 1000 iterations.

5 | RESULTS AND DISCUSSION

5.1 | Evaluation of the performances

The performances of the models evaluated with 10-fold cross-validation by RMSE are reported in Table 3.

As shown in Table 3, the average percent errors of the 12 regressors are very small, confirming the robustness of the regression models across all the outputs. The DT regressors have the lowest error for the predictions of φ_{AD} and φ_{PR}. The ERT regressors perform the best to predict the φ_{EFF} and the FVF. In Figure 6, the scatter plots of the predicted values with respect to the actual values for the outputs are reported for the different ML models. The scatter plots confirm the results presented in Table 3. The 12 TGML regressors are comparable in terms of accuracy. In particular for the φ_{EFF} and the FVF, it is clear how the ERT algorithms have superior performances than the other models. Thus, considering that the cost function for the optimization is defined in Equation 2, the software *SANDWICH* is substituted with the ERT regressors for the prediction of φ_{EFF} and FVF.

5.2 | Importance of the inputs

To design a cure cycle for the co-cure process of a sandwich structure with the desired properties, it is necessary to identify the correlation between the process parameters and the physics of the phenomena involved in the co-curing of prepreg and adhesive. Although the software

Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

TABLE 3 Percent error in mean RMSE of 10-folds cross-validation using different ML algorithms

Model	Adhesive porosity φ_{AD}	Prepreg porosity φ_{PR}	Effective porosity φ_{EFF}	Max. fiber volume fraction FVF
DT average RMSE	1.03E-09 ± 8.60E-12	0.22 ± 3.42E-03	1.95 ± 9.62E-03	0.10 ± 1.60E-04
RF average RMSE	0.04 ± 2.54E-04	0.78 ± 4.14E-03	1.32 ± 1.53E-03	0.08 ± 5.72E-05
ERT average RMSE	0.31 ± 6.48E-03	0.75 ± 3.19E-03	1.02 ± 8.64E-04	0.07 ± 7.37E-05

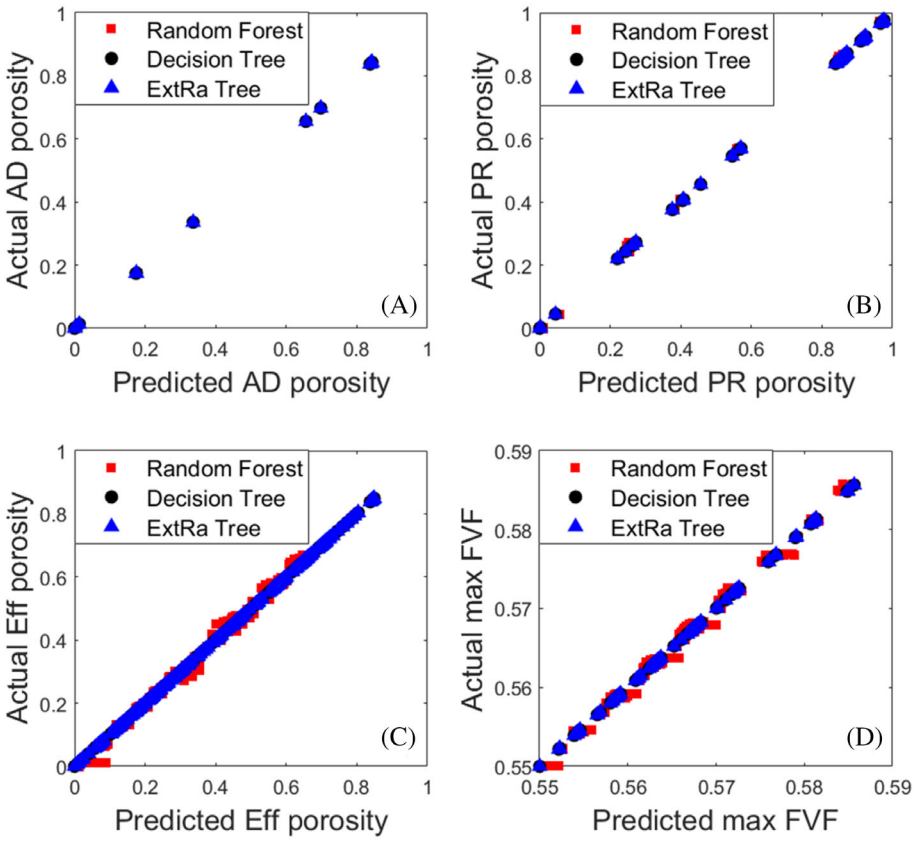


FIGURE 6 Scatter plot of outputs predicted by the theory guided machine learning regressors versus the actual outputs in the data set: (A) adhesive porosity; (B) prepreg porosity; (C) effective porosity; (D) max fiber volume fraction

SANDWICH is able to characterize the underlying multiphysics of the co-curing process and predict the outcome of a prescribed cure cycle, it is not designed to capture the influence of each of the inputs on the output parameters. One of the capabilities of ML tools is the recognition of the contribution of each feature on the outputs. For this study, in parallel to the development of the TGML regressors, we implemented the attribute function "feature_importances" of the built-in regressor "RandomForestRegressor" of the Python 3.0 module sklearn.ensemble from the library Scikit Learn. The function is able to rank the input features according to the level of influence on each output, and thus in each specific physical phenomenon of the co-curing process.

Figure 7 presents the results of ranking of the input parameters. For the φ_{AD}, the input parameters that have the highest rank, as shown in the pie chart in Figure 7A, are related to the vacuum pressure and autoclave temperature cycles. This is in accordance with the knowledge from the literature: the autoclave pressure has no influence on the growth or collapse of the voids within the adhesive, as they are subjected to the pressure within the core cells. The autoclave temperature is the parameter affecting the adhesive viscosity and as the core is connected to the vacuum bag in the equilibrated configuration, the growth or shrinkage of voids is dependent on the vacuum pressure.[25,26]

The pie chart in Figure 7B presents the rank of the inputs on the output φ_{PR}. The first parameter in rank is the vacuum pressure. Indeed, similarly for the case of φ_{AD}, the growth of voids within the bled prepreg resin is governed mostly by the vacuum pressure. The RF model is also able to capture the influence of the autoclave cycle on φ_{PR} as second in rank. In fact, the volume of resin that

Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

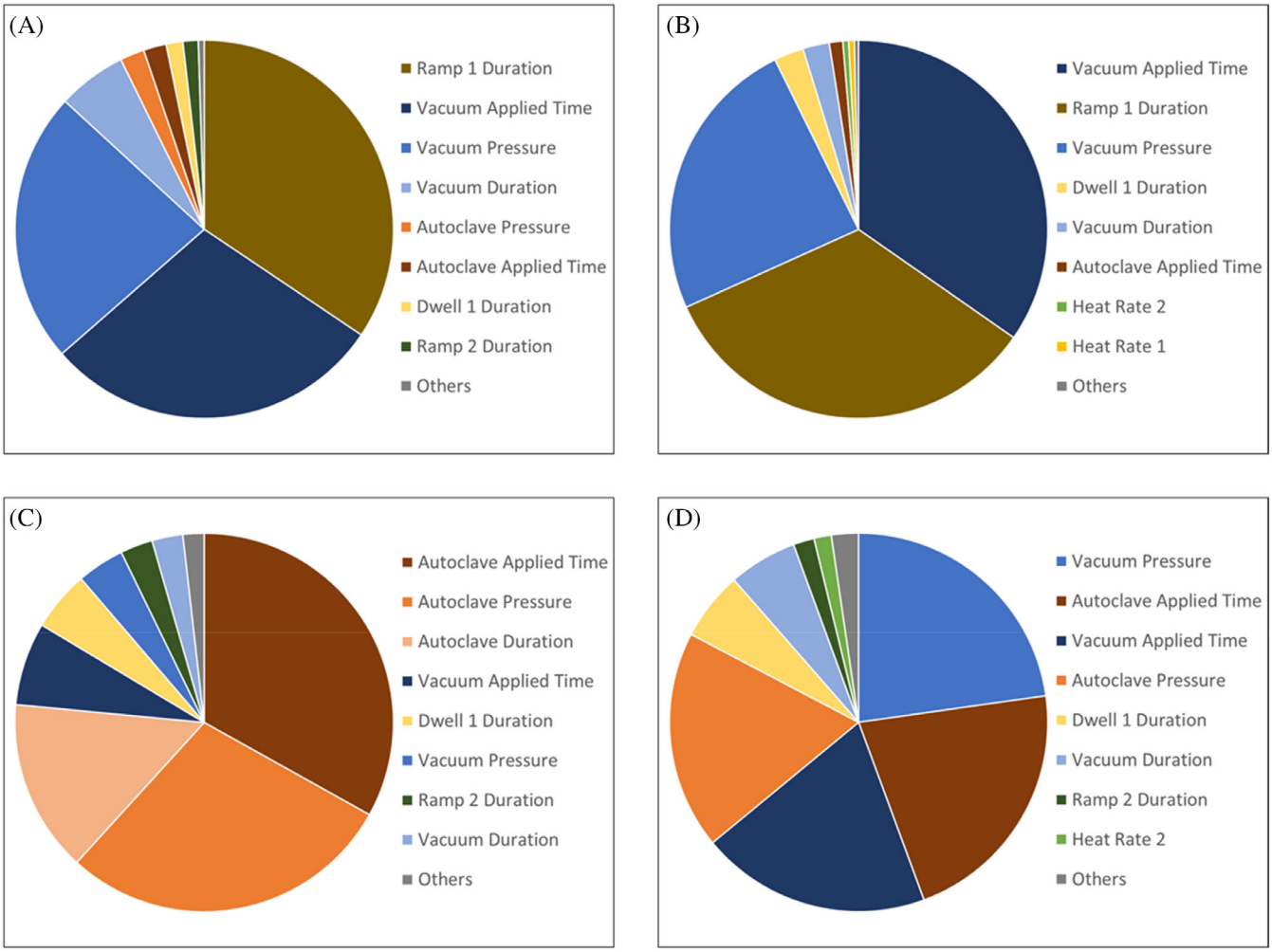


FIGURE 7 Feature ranks: (A) adhesive porosity; (B) prepreg porosity; (C) effective porosity; (D) max. fiber volume fraction

bleeds is dependent on the consolidation process, and thus on the autoclave cycle. The fifth ranked parameter is one of the parameters of the temperature cycle, to mark that the bled prepreg resin is also influenced by the temperature in the autoclave which determines its viscosity.

For φ_{EFF} in the bond-line, as shown in Figure 7C, the results obtained for φ_{AD} and φ_{PR} are combined. Indeed, the bond-line is formed by the adhesive and the bled resin from the prepreg facesheets. The first parameter in rank is associated with the temperature cycle, identifying the influence of the viscosity of both the polymers on the void growth and collapse. Then, from the second position in rank are the parameters of the vacuum cycle, and finally fifth in rank is the autoclave pressure. These results are justified by the fact that the bled resin in the bond-line is less in volume with respect to the adhesive.[27]

The rank of the inputs for the FVF are summarized in Figure 7D. The TGML tool was able to understand that the consolidation process is dependent most of all on the autoclave pressure cycle. Indeed, the three parameters associated with it are first in rank. Then, the fourth feature in rank is related to the vacuum pressure cycle, and it is followed by one of the parameters that characterizes the temperature cycle. The rank for the FVF is in accordance with the physics of the process. The boundary conditions at the top and bottom of the facesheets during the consolidation of the sandwich panel are the autoclave pressure and the vacuum pressure, respectively, in the equilibrated configuration. Then, the consolidation is also dependent on the temperature of the process, since the volume of resin that bleeds from the prepreg depends on the viscosity of the polymer.[26]

5.3 | Optimization

To evaluate the performances of the TGML optimization for the co-cure process, similar optimization routines as the test cases presented by Niknafs et al.[14] are designed.

Three tentative optimizations are constructed: the first to optimize the autoclave pressure process parameters only, the second to optimize the vacuum pressure process parameters only, and the third to optimize both. Two temperature cycles are specified as shown in Figure 8 and kept fixed for the optimization runs. The temperature cycles are designed with two different heating ramps (2 and 3°C/min) with isothermal dwells at the same temperatures (110 and 170°C).

First, the capabilities of the TGML regressors to formulate an optimization strategy able to find a cure cycle to minimize the bond-line porosity are explored and tested. The weights of the cost function described in section 4 are set as $w_1 = 1$ and $w_2 = 0$. The resulting cost function is presented in Equation (6).

$$f_{(x)} = \varphi_{\mathrm{EFF}}(x) \tag{6}$$

In this case, the optimization is conducted on the vacuum parameters (VP, VAP and VD), while the autoclave parameters are relaxed and listed in Table 4.

Figure 9A shows the results of the optimization, where the evolution of φ_{EFF} is plotted with regard to the iterations. The optimal value of the porosity for the case of heat rate of 2°C/min is 1.21% and for the case of heat rate of 3°C/min is 1.89%. The trend of the results is similar to what Niknafs et al. previously presented.[14] The TGML predictive models responded to the change of parameters in a similar fashion to the physics-based optimization, identifying the optimal process parameters to minimize φ_{EFF}.

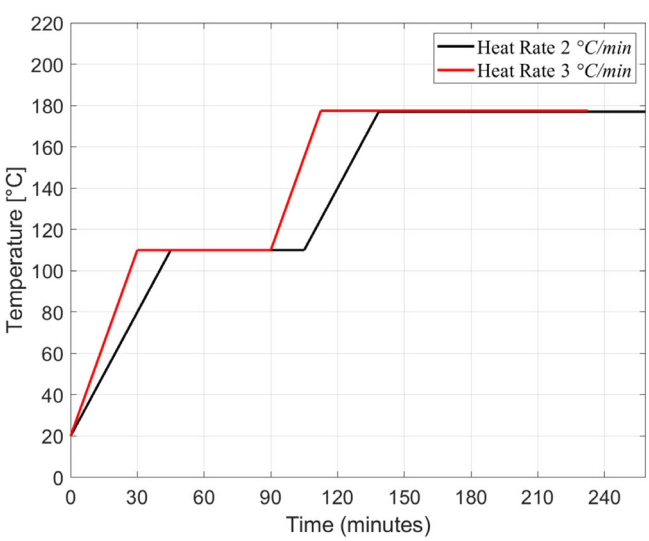


FIGURE 8 Temperature cycles used to optimize with the theory guided machine learning models

TABLE 4 Process parameters for optimizing with objective function stated in Equation (6)

$x_{10} = 370$ kPa	$x_{11} = 1$ min	$x_{12} = 170$ min

A second test is performed to optimize the consolidation level. In this case, the cost function is built with weights $w_1 = 0$ and $w_2 = 1$, as shown in Equation (7).

$$f_{(x)} = \frac{\mathrm{VF}_{\max} - \mathrm{FVF}_{(x)}}{\mathrm{VF}_{\max} - \mathrm{VF}_{\min}} \tag{7}$$

For this trial, the autoclave pressure process parameters (AP, AAP, AD) are optimized, and the vacuum process parameters are kept fixed and are listed in Table 5.

The results of the optimization for FVF are shown in Figure 9B. The optimal value of the fiber volume fraction is 58.27% for both temperature cycles. Also in this case, the FVF evolves in a similar fashion to what Niknafs et al.[15] presented for the physics-based optimization. The TGML tool performance was comparable to *SANDWICH* to predict the outputs for the optimization routine and it was able to change the three inputs to maximize the FVF in a correct way in seconds as compared to hours and days needed by *SANDWICH*.

For the final test, both autoclave pressure and vacuum pressure process parameters are used as input vector for the search of the optimal values to simultaneously minimize the bond-line porosity and maximize the consolidation level. Therefore, the optimization is carried out by varying six parameters. The cost function is modified as a linear combination of φ_{EFF} and FVF, with weights $w_1 = 0.5$ and $w_2 = 0.5$ (Equation 8).

$$f_{(x)} = \frac{1}{2}\varphi_{\mathrm{EFF}(x)} + \frac{1}{2}\left(\frac{\mathrm{VF}_{\max} - \mathrm{FVF}_{(x)}}{\mathrm{VF}_{\max} - \mathrm{VF}_{\min}}\right) \tag{8}$$

Figure 9C presents the evolution of φ_{EFF} and FVF w.r.t. the iterations. The results of the optimization with the TGML tool are again similar to the results of the physics-based optimization.[14] The minimum φ_{EFF}, outputs shows no porosity with the optimal process parameters after 1000 iterations, for both cure cycles and the maximum FVF approach 58.56% in both cases as shown in Figure 9.

Table 6 presents the optimal parameters for obtaining the minimum effective bond-line porosity (Figure 9A) in which the optimization parameters define the vacuum pressure cycle (x_7,x_8,x_9). The result is in agreement with the previous study in which, as expected, the minimum porosity corresponds to the cure cycles with very low vacuum pressure to make the majority of voids grow to the

Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

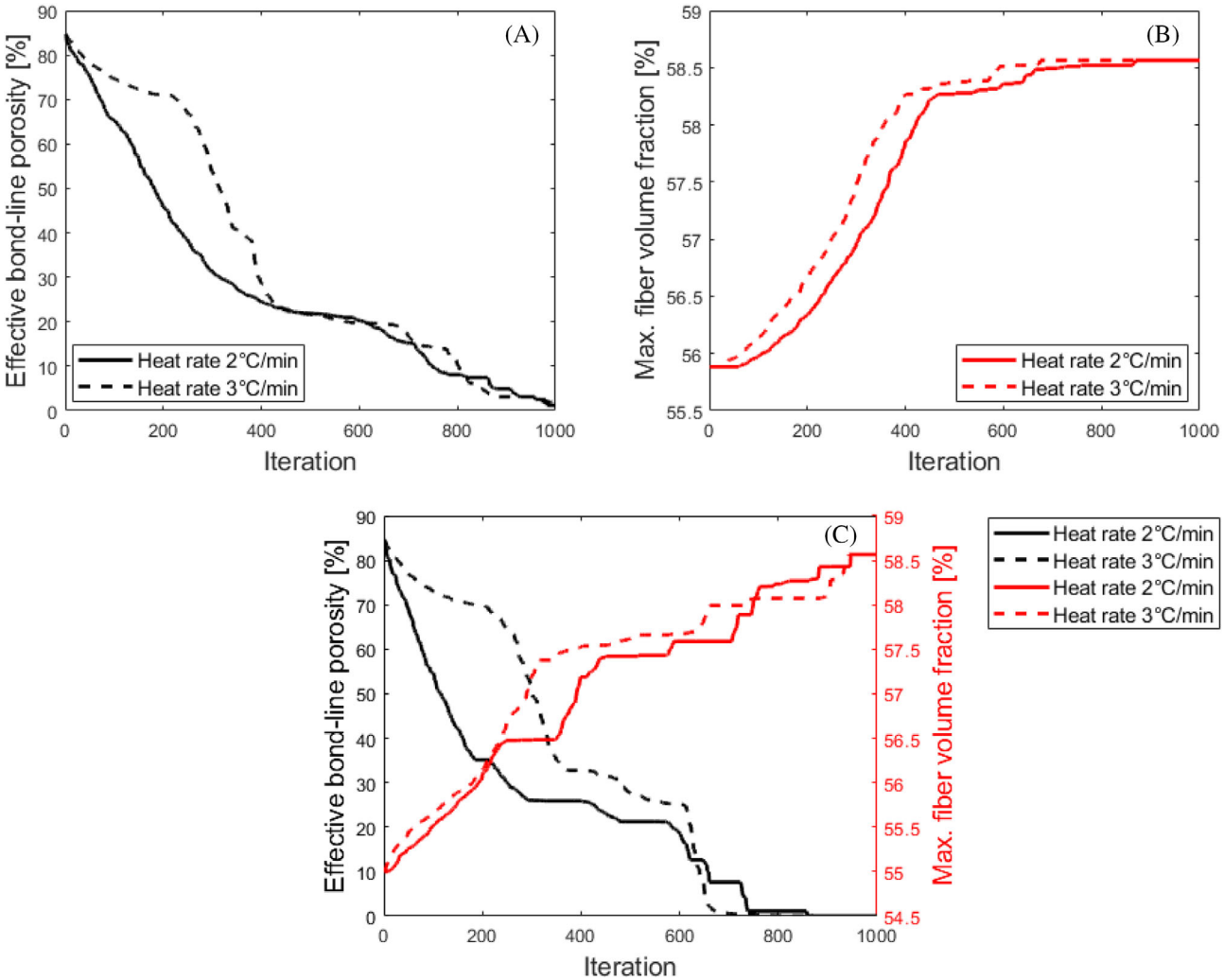


FIGURE 9 Results of optimization using theory guided machine learning models: (A) optimization on cost function in Equation (6), (B) optimization on cost function in Equation (7), (C) optimization on cost function in Equation 8

TABLE 5 Process parameters for optimizing with the objective function stated in Equation (7)

$x_7 = 1.18$ kPa	$x_8 = 1$ min	$x_9 = 170$ min

extent that they leave the bond-line caused by the escape phenomena.[14]

In the second case, optimization of autoclave pressure cycle calls for applying the highest consolidation pressure permissible until the prepreg gels to achieve the highest consolidation level across the facesheet, which results in the highest fiber volume fraction as shown in Table 7 which is consistent with previous investigations and observations.

The results in the previous cases show the expected trends in the optimization process. Finally, the study to obtain the optimum parameters over the vacuum and autoclave pressure cycles are listed in Table 8.

The optimum bond-line porosity and fiber volume fraction in this study shows the benefit of the defined cost function in this study compared to the previous one. Unlike the previous study,[14] the cost function here is sensitive to both bond-line porosity and facesheet consolidation level in the simultaneous optimization of these two outputs over vacuum and autoclave pressure cycles.

The TGML regressors developed with the ERT algorithm for φ_{EFF} and FVF show results in agreement with the physics-based simulations in all the tests performed, both for the optimization on three and six parameters. Therefore, the TGML tool is a suitable substitute for the physics-based software as a prediction tool in the optimization routine. The benefits associated with the use of TGML tools are related to the computational time. One single iteration of the surrogateopt algorithm with the predictions by *SANDWICH* occurs on an average in 386.54 s. This translates into 231,924 s to perform

Accepted Manuscript
Version of record at: https://doi.org/10.1002/pc.26829

TABLE 6 Optimal input parameters to obtain minimum bond-line porosity

	Heat rate [°C/min]	x_7 [kPa]	x_8 [min]	x_9 [min]	y_3 [%]
$w_1 = 1$	2	1.5	80	146	1.2
$w_2 = 0$	3	1.5	59	198	1.4

TABLE 7 Optimal input parameters to obtain maximum fiber volume fraction

	Heat rate [°C/min]	x_7 [kPa]	x_8 [min]	x_9 [min]	y_4 [%]
$w_1 = 0$	2	370	73	184	58.6
$w_2 = 1$	3	370	6	258	58.6

TABLE 8 Optimal input parameters to reduce the bond-line porosity and maximize fiber volume fraction

	Heat rate [°C/min]	x_7 [kPa]	x_8 [min]	x_9 [min]	x_{10} [kPa]	x_{11} [min]	x_{12} [min]	y_3 [%]	y_4 [%]
$w_1 = 0.5$	2	1.5	80	120	370	1	170	0	58.6
$w_2 = 0.5$	3	1.2	61	196	370	25	199	1.8	58.3

600 iterations for the search of the optimal parameters for each of the three tests.[15] One iteration of the optimization with TGML regressors lasts on an average for 0.77 s. Therefore, 1000 iterations of the optimization code were completed within 770 s. The possibility to perform thousands of iterations in few minutes justifies the use of TGML regressors for the design of new cure cycles to fabricate honeycomb sandwich structures by the co-cure process. In fact, this methodology could be used to optimize other process simulations of manufacturing processes and reduce the prediction time by orders of magnitude.

6 | CONCLUSIONS

TGML regressors to predict the characteristics of the bond-line of honeycomb composite sandwich structures were developed. Each regressor was trained to establish the correlation between the cure cycle prescribed in the autoclave to perform the co-cure with the parameters related to the final porosity of the bond-line and the prepreg facesheets consolidation.

A total of 5056 data points were generated through physics-based simulations by the software *SANDWICH*. The input cure cycles were designed with 12 parameters, six describing the autoclave temperature cycles in terms of heating rates, heating ramp durations and isothermal dwell durations, and six describing the vacuum pressure and autoclave pressure cycles in terms of pressure values, applied times and dwell durations. Three outputs were collected to characterize the bond-line porosity (i) adhesive porosity; (ii) bled prepreg resin porosity and (iii) effective porosity, and one output related to the final consolidation level, which was assumed to be related to the maximum fiber volume fraction across the facesheets.

The data set was used to train 12 TGML regressors with three different ML algorithms. Respectively, three regressors were generated separately for each of the four outputs, and the best four regressors were then selected by evaluating the RMSE of the 10-fold cross-validation.

The input features were ranked for each of the outputs and from the results it is possible to notice how the TGML tools can identify information related to the physical phenomena involved in the co-cure process of sandwich composites through the regression. The most influencing parameters for the porosity were identified as the vacuum pressure and temperature which is consistent what has been reported. The TGML tools also identified the influence of the autoclave pressure cycle on the void growth and collapse, due to the fact that the amount of bled resin is determined during the consolidation process. The most influencing parameters in the consolidation process were identified as the autoclave pressure and temperature, results again in agreement with the literature.

The selected best regressors were used in an optimization routine to find the best cure cycle to minimize the bond-line porosity content and maximize the consolidation of the composite skins. The results of the optimization are in complete agreement with the results of optimizations performed with the physics-based tools. Therefore, the TGML is a valuable tool for the prediction of the output of the co-cure of a honeycomb sandwich composite fabricated under a prescribed curing cycle in an autoclave. Future applications may extend the comprehension of the simulations by constructing the optimization on a higher number of input parameters, such as the parameters related to the temperature cycle.

ORCID

Suresh G. Advani https://orcid.org/0000-0002-2670-903X

REFERENCES

[1] J. Leijten, H. E. N. Bersee, O. K. Bergsma, A. Beukers, *Compos. Part A: Appl. Sci. Manuf.* **2009**, *40*(2), 164.
[2] M. Shifa, F. Tariq, A. D. Chandio, *J. Sandw. Struct. Mater.* **2019**, *23*(1), 222.
[3] A. Pavlovic, D. Sintoni, G. Minak, C. Fragrassa, *Compos. Struct.* **2020**, *248*, 112523.
[4] M. Anders, D. Zebrine, T. Centea, S. Nutt, *Compos. A: Appl. Sci. Manuf.* **2019**, *116*, 24.
[5] N. Niknafs Kermani, P. Simacek, S. G. Advani, Proceedings of the SAMPE 2019 Conference, Society for the Advancement of Material and Process Engineering, **2019**.
[6] N. Niknafs Kermani, P. Simacek, S. G. Advani, Proceedings of the American Society of Composites, **2019**.
[7] J.-D. Nam, J. C. Seferis, S. W. Kim, K. J. Lee, *Polym. Compos.* **1995**, *16*(5), 370.
[8] M. Anders, D. Zebrine, T. Centea, S. Nutt, *J. Manuf. Sci. Eng.* **2017**, *139*(11), 111012.
[9] N. Niknafs Kermani, P. Simacek, S. G. Advani, *Compos. Part A: Appl. Sci. Manuf.* **2020**, *139*, 106071.
[10] M. Yun, E. Lopez, F. Chinesta, S. G. Advani, *Compos. Sci. Technol.* **2018**, *168*, 238.
[11] M. Yun, C. A. Martin, P. Giormini, F. Chinesta, S. G. Advani, *Entropy* **2019**, *22*(1), 30.
[12] M. Li, Y. Li, Z. Zhang, Y. Gu, *Polym. Compos.* **2008**, *29*(5), 560.
[13] D. Zebrine, N. Niknafs Kermani, P. Simacek, T. A. Cender, S. G. Advani, S. Nutt, *Adv. Manuf.: Polym. Compos. Sci.* **2022**, 1. https://doi.org/10.1080/20550340.2022.2077890
[14] N. Niknafs Kermani, P. Simacek, S. G. Advani, Proceedings of the American Society for Composites, **2021**.
[15] N. Niknafs Kermani, Unpublished dissertation for the degree of Doctor of Philosophy in Mechanical Engineering, University of Delaware, 2021.
[16] C.-T. Chen, G. X. Gu, *MRS Commun* **2019**, *9*, 556.
[17] M. Ozkan, A. Karakoc, M. Borghei, J. Wiklund, O. J. Rojas, J. Paltakari, *Polym. Compos.* **2019**, *40*, 4013.
[18] J. Hou, B. You, J. Xu, T. Wang, *Polym. Compos.* **2022**, *43*, 1749.
[19] J. Reiner, R. Vaziri, N. Zobeiry, *Compos. Struct.* **2021**, *273*, 114290.
[20] N. Zobeiry, J. Reiner, R. Vaziri, *Compos. Struct.* **2020**, *246*, 112407.
[21] J. R. Quinlan, *Mach. Learn.* **1986**, *1*(1), 81.
[22] L. Breiman, *Mach. Learn.* **2005**, *45*(1), 5.
[23] P. Geurts, D. Ernst, L. Wehenkel, *Mach. Learn.* **2006**, *63*(1), 3.
[24] S. G. Advani, E. M. Sozer, *Process modeling in composites manufacturing*, CRC Press, New York. **2002**.
[25] F. Boey, S. Lye, *J. Mater. Process. Technol.* **1990**, *23*, 121.
[26] P. Simacek, S. G. Advani, *Compos. Sci. Technol.* **2020**, *186*, 107892.
[27] N. Niknafs Kermani, P. Simacek, S. G. Advani, *Compos. A: Appl. Sci. Manuf.* **2020**, *132*, 105824.

How to cite this article: T. Lavaggi, M. Samizadeh, N. Niknafs Kermani, M. M. Khalili, S. G. Advani, *Polym. Compos.* **2022**, 1. https://doi.org/10.1002/pc.26829