IJAME
Int. J. of Applied Mechanics and Engineering, 2024, vol.29, No.4, pp.78-95
DOI: 10.59441/ijame/192278

OPTIMISATION OF MATERIAL COMPOSITION IN FUNCTIONALLY GRADED PLATES USING A STRUCTURE-TUNED DEEP NEURAL NETWORK

Ryoichi Chiba* and Takuya Kishida
Mechanical Engineering, Sanyo-Onoda City University, JAPAN
E-mail: chiba@rs.socu.ac.jp

Ryoto Seki and Seiya Sato
Mechanical Systems Engineering, National Institute of Technology, Asahikawa College, JAPAN

This study presents a neural network (NN)-based approach for optimising material composition in multi-layered functionally graded (FG) plates to minimise steady-state thermal stress. The focus is on the metal-ceramic composition across the thickness of heat-resistant FG plates, with the volume fractions of the ceramic constituent in each layer treated as design variables. A fully-connected NN, configured with an open-source Bayesian optimisation framework, is employed to predict the maximum in-plane thermal stress for various combinations of design variables. The optimal distribution of material composition is determined by applying a backpropagation algorithm to the NN. Numerical computations on ten- and twelve-layered FG plates demonstrate the usefulness of this approach in designing FG materials. NNs trained with *8000* samples enable the successful acquisition of valid optimal solutions within a practical timeframe.

Key words: neural network; optimal design; functionally graded material; thermal stress; material design; multi-layered material.

1. Introduction

Functionally graded materials (FGMs) are composite materials in which the composition and/or structure vary along specific direction(s) in a continuous or stepwise manner [1]. Among FGMs, those composed of metal and ceramic, known as *FGMs for thermal stress relaxation*, are designed to withstand large temperature drops. Generally, the performance of FGMs heavily depends on their material composition profiles, thus making the profile optimisation crucial for maximising desired functions, such as thermal stress reduction.

Optimal design methods for FG structures have been extensively reviewed, focusing on structural shape [2] and chronological order [3]. Gradient-based approaches, such as sequential quadratic programming [4-6], and metaheuristic approaches including differential evolution [7], genetic algorithm [8-12], particle swarm optimisation [13-18], and others [18-21] have been commonly employed for FGM composition optimisation. Additionally, there is a growing interest in utilising neural networks (NNs) for surrogate modelling to enhance the efficiency of the optimisation [22-27].

On the other hand, Ootao *et al.* [28-31] proposed a simple NN-based optimisation technique for FGM composition in the late 90s, which integrates both the analysis and optimisation processes into a single framework, eliminating the need for separate preparation of an analyser and optimiser. However, this approach assumed power-law type functions for spatially varying material composition, which was found to be overly restrictive in representing the material gradation in certain cases [4]. This assumption may have been influenced by the limited computing power at the time, which constrained the number of design variables that could be accommodated by specifying the function form.

* To whom correspondence should be addressed

Advancements in computing power and deep NN technology have prompted a re-evaluation of the FGM composition optimisation problem. This study aims to overcome the previous limitations and leverage recent technological developments for optimal material design in multi-layered FG plates. The choice of multi-layered FGMs here is driven by practical and interpretive considerations. Technically, achieving a continuously varying material composition is challenging due to manufacturing constraints, such as limitations in material handling and processing accuracy [32]. Furthermore, advanced thermal barrier coatings often take the form of multi-layered FGM [33]. The discrete nature of multi-layered FGMs also enhances interpretability, making the optimisation results clearer and more applicable. Importantly, the optimisation results for multi-layered FGMs can provide a basis for designing compositionally continuous FGMs; the stepwise distribution can be smoothed to approximate continuous gradation, thus bridging practical constraints with the goal of seamless material variation.

This study applies a NN-based optimisation method, building on the work of Ootao *et al.* [28], and utilises a deep NN with a structure tuned via Bayesian optimisation. Numerical examples include ten- and twelve-layered FG plates made of nickel and alumina [13, 22, 34] and carbon and silicon carbide [14]. Unlike previous approaches, this method does not impose restrictions on material gradation profiles, allowing for more flexible representation and accurate optimisation of FGM composition. The objective is to minimise the absolute value of in-plane thermal stress under a steady temperature field.

2. Thermoelastic analysis

2.1. Temperature field

As an analytical model, let us consider a multi-layered FG plate with a thickness of h and infinite lengths along the x- and y-directions, as illustrated in Fig.1. In the FG plate, the thermal conductivity λ, Young's modulus E, thermal expansion coefficient α and Poisson's ratio ν vary in a stepwise manner along the thickness (or z-axis) direction, depending on the composition profile in the same direction. The plate surfaces at $z=0$ and $z=h$ are subjected to known temperatures T_0 and T_n, respectively. These thermal boundary conditions represent typical operational scenarios for FG plate structures (e.g. thermal barrier coatings) exposed to high and relatively stable temperatures, as found in aerospace applications, gas turbines and industrial furnaces.

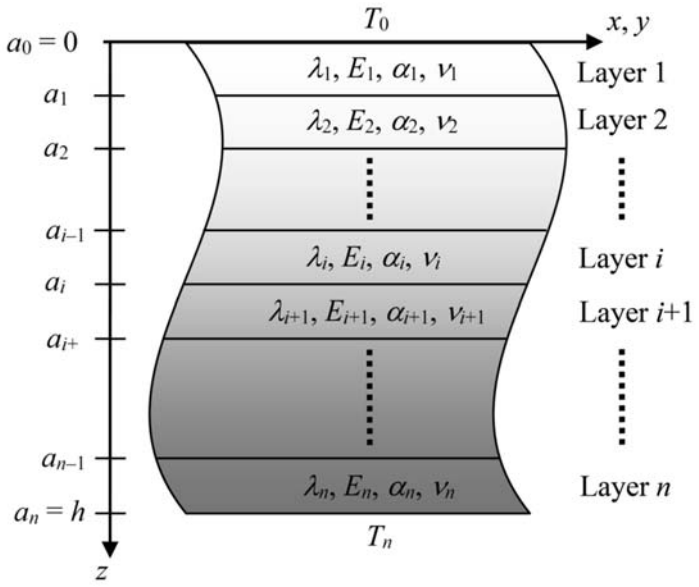

Fig.1. Analytical model of functionally graded infinite plate and coordinate system.

Assuming one-dimensional steady heat conduction across the plate thickness, the interface temperature T_i at $z = a_i \left(i = 1, 2, \ldots, n-1\right)$ can be obtained as follows:

$$T_i = T_{i-1} - \frac{T_0 - T_n}{\sum_{k=1}^{n} \frac{a_k - a_{k-1}}{\lambda_k}} \frac{a_i - a_{i-1}}{\lambda_i} \quad \text{for } i = 1, 2, \ldots, n-1 . \tag{2.1}$$

No thermal contact resistance is considered at the mating surfaces, as the bonding between layers is assumed to be perfect. The temperature distribution in each layer is given as a linear function that connects the temperatures at both interfaces, resulting in a continuous multi-linear temperature profile throughout the plate.

2.2. Thermal stress field

Consider the thermal stress field in a traction-free plate where the temperature varies only along the thickness direction (Fig.1). The plate is assumed to have an arbitrary planform with a constant thickness *h*. Both the length and width of the plate are considered sufficiently large compared to its thickness, classifying it as a thin plate. Stress components $\sigma_{zz} = \sigma_{xz} = \sigma_{yx} = \sigma_{zy}$ are assumed to be zero, which is a reasonable approximation for this analysis. In the case of an arbitrarily non-homogeneous plate across the thickness, the in-plane thermal stresses can be evaluated using an analytical solution derived by Sugano [35]. Therefore, the non-zero thermal stress components in the multi-layered FG plate under consideration, with surfaces free of tractions, are expressed as follows:

$$\sigma_{xx} = \sigma_{yy} = \frac{E(z)}{1-\nu(z)} \left[-\alpha(z)\Delta T(z) + \frac{\left(I_2 z - I_3\right)\Phi_1 + \left(I_2 - I_1 z\right)\Phi_2}{I_2^2 - I_1 I_3} \right], \tag{2.2a}$$

where

$$I_j = \int_0^h \frac{z^{j-1} E(z)}{1-\nu(z)} \mathrm{d}z , \quad j = 1, 2, 3, \tag{2.2b}$$

$$\Phi_j = \int_0^h \frac{z^{j-1} E(z)\alpha(z)\Delta T(z)}{1-\nu(z)} \mathrm{d}z , \quad j = 1, 2, \quad \text{with} \quad \Delta T = T - T_0 . \tag{2.2c}$$

Equation (2.2a) was derived through a two-step process: (i) first, by determining the form of stress components necessary to satisfy the compatibility equations expressed in terms of stress components; and (ii) second, by identifying two unknown constants to ensure that the resultant force and moment (per unit length) produced by σ_{xx} and σ_{yy} are zero over the edges of the plate. According to Saint-Venant's principle, this ensures that Eq.(2.2a) provides an accurate approximation for traction-free edges at distances from these edges greater than approximately one plate thickness.

Since all the shear stress components and σ_{zz} are zero, the absolute value of each of σ_{xx} and σ_{yy} is consistent with the von Mises stress, an important scalar measure of the stress tensor. For the generality of analysis, dimensionless stresses are introduced as $\left\{\bar{\sigma}_{xx}, \bar{\sigma}_{yy}\right\} = \left\{\sigma_{xx}, \sigma_{yy}\right\} / \left(E_m \alpha_m \Delta t\right)$, where E_m and α_m denote the Young's modulus and thermal expansion coefficient of the non-ceramic phase, respectively, and Δt is a reference temperature difference.

3. Volume fraction optimisation using a deep neural network

3.1. Formulation of the optimisation problem

The goal here is to determine the optimal volume fraction of ceramic in each layer of multi-layered FGMs, including metal-ceramic and carbon-ceramic compositions. As such, the optimisation problem, along with the necessary constraint conditions, can be expressed in the following form:

$$\begin{aligned} &\text{Design variables} && \boldsymbol{V}, \\ &\text{Minimise} && f(\boldsymbol{V}), \\ &\text{Subject to} && 0 \le V_i \le 1, \quad i = 1, 2, \ldots, n, \quad V_i \le V_{i+1}, \quad i = 1, 2, \ldots, n-1, \end{aligned} \tag{3.1}$$

where $\boldsymbol{V} = [V_1, V_2, \ldots, V_n]$ represents the vector of the ceramic volume fractions in the respective layers, and $f(\boldsymbol{V})$ denotes the objective function to be minimised. In this study, the objective function is defined as the maximum absolute value of $\bar{\sigma}_{xx}$ evaluated throughout the thickness [13, 22, 34], i.e.

$$f(\boldsymbol{V}) = \max_{z \in [0,h]} \left| \bar{\sigma}_{xx}(z) \right|. \tag{3.2}$$

3.2. Building a deep neural network

A NN predictor (or analyser) for the maximum thermal stress in the *n*-layered FG plate, represented by Eq.(3.2), is constructed using a fully-connected NN with *M* hidden layers, as illustrated in Fig.2. The input layer contains *n+1* neurons for the input of features $\boldsymbol{V}$ and an additional bias. Each hidden layer can have a variable number of neurons, denoted as $N_i + 1$ for the *i*th hidden layer, where "+*1*" represents the inclusion of one bias in the layer. The output layer comprises a single neuron that generates the predicted value of the maximum thermal stress.

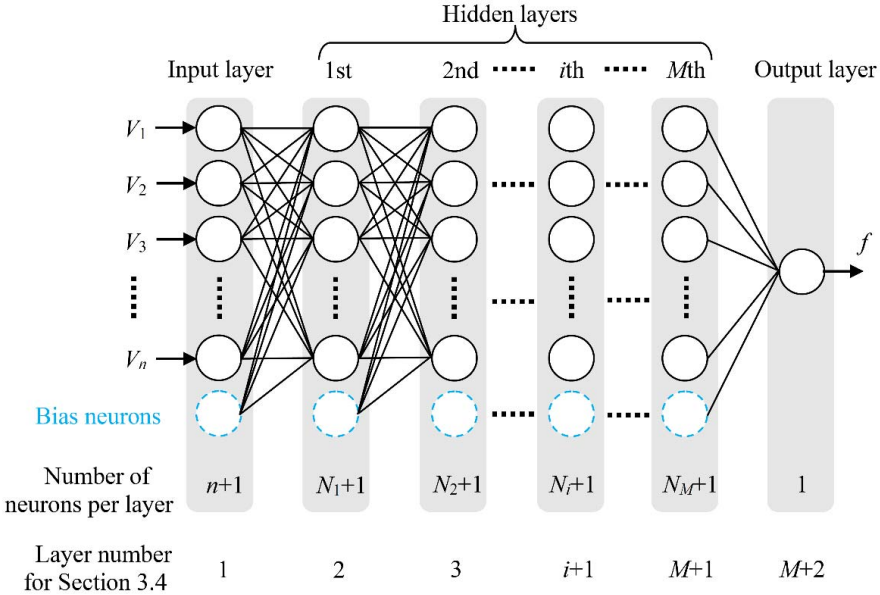

Fig.2. Structure of fully-connected deep neural network with bias neurons.

A dataset consisting of *12000* pairs of randomly generated $\boldsymbol{V}$ (excluding fixed values of $V_1 = 0$ and $V_n = 1$) and their corresponding maximum absolute value of $\bar{\sigma}_{xx}$, evaluated using Eqs (2.1) and (2.2), is prepared. Details of the dataset are provided in Tab.1. Subsequently, the dataset is split into *8000* data pairs for training, *2000* for validation and *2000* for testing.

Table 1. Details of dataset used for training, validating and testing the neural network.

No.	V_1	V_2	V_3	⋯	V_{n-1}	V_n	$\lvert\bar{\sigma}_{xx}\rvert_{\max}$
1	0	0.068	0.121	⋯	0.988	1	0.2285
2	0	0.150	0.237	⋯	0.888	1	0.1386
3	0	0.023	0.027	⋯	0.987	1	0.2211
⋮	⋮	⋮	⋮		⋮	⋮	⋮
12000	0	0.084	0.089	⋯	0.690	1	0.1938

In the training phase of the NN, the mean squared error (MSE) of mini-batch version is employed as a metric to assess the prediction error. Specifically, the MSE between the target values and predicted values of $f(\boldsymbol{V})$ serves as the loss function. To update the network weights and bias values, Adaptive Moment Estimation, commonly referred to as Adam, is utilised. The specific Adam configuration parameters are detailed in Tab.2. Additionally, the rectified linear unit (ReLU) is employed as the activation function in all the hidden layers of the network. The network undergoes mini-batch training for up to *3000* epochs. The normalisation of input data is not necessary because all components of $\boldsymbol{V}$ fall within the range from *0* to *1*.

Table 2. Adam configuration parameters for neural network training.

Parameter	Moment decay rate		Small constant for numerical stability	Learning rate
	First	Second		
Value	0.9	0.999	10^{-8}	0.001

3.3. Structure tuning of the neural network

In this study, the tuning of the NN structure is conducted using Bayesian optimisation. This process aims to optimise the hyperparameters that define the architecture of the NN to achieve optimal performance. To facilitate this, we utilise the Tree-structured Parzen Estimator (TPE) algorithm [36], which is provided by Optuna [37] – an open-source hyperparameter optimisation framework for machine learning. TPE is a Bayesian optimisation algorithm that efficiently explores the solution space of possible NN structures and identifies the configuration that yields optimal results. Optuna provides a user-friendly interface for defining hyperparameter search spaces and also offers extensive documentation and community support, thus facilitating integration into the workflow.

Specific hyperparameters targeted for tuning are M and $N_i\,(i = 1, 2, \ldots, M)$, which denote the number of hidden layers and the number of neurons for each hidden layer (excluding bias neuron), respectively, as shown in Fig.2. Constraints are imposed as $1 \le M \le 7$ and $10 \le N_i \le 50$, in increments of *10*. These constraints help manage the trade-offs between thorough exploration of the solution space and computational burden required for the tuning. During the tuning process, Optuna conducts *200* trials, each comprising *100* epochs,

to thoroughly explore the space of possible NN structures. Thus, the objective value of Optuna is the MSE calculated for the validation dataset after the completion of *100*-epoch training.

Additionally, we also employ Optuna to determine the optimal mini-batch size from a predefined set of options, including *32*, *64*, *96* and *128*. This ensures identification of the mini-batch size, another hyperparameter, that best suits the dataset characteristics and maximises training efficiency.

3.4. Optimisation calculation

Once the training of the NN is completed, the optimal volume fractions of ceramic in the respective layers are estimated through a backpropagation algorithm while keeping all network weights and bias values fixed [28, 38]. This approach was selected for its efficiency and effectiveness in gradient-based optimisation tasks, allowing precise adjustment of the volume fractions with minimal computational cost. Thus, this streamlines the workflow and simplifies the optimisation process. The following provides a concise description of the optimisation procedure.

(1) With a vector $\boldsymbol{V}^{ini} = \left[0,\ V_2^{ini},\ V_3^{ini},\ ...,\ V_{n-1}^{ini},\ 1 \right]$ as input, the forward propagation is initiated from the input layer of the trained NN. During the propagation to the output layer, the internal states of neurons u_j^k are recorded, where the superscript k denotes the layer number ranging from *1* to *M+2* (as shown in the bottom of Fig.2), and the subscript j represents the neuron number in each layer.

(2) Recursive calculations are employed based on the following equations from k equals M to *1*, which allow us to obtain the values of $\delta_1^1,\ \delta_2^1,\ ...,\ \delta_n^1$:

$$\begin{bmatrix} \delta_1^{M+1} \\ \delta_2^{M+1} \\ \vdots \\ \delta_q^{M+1} \end{bmatrix} = -\varepsilon \cdot \boldsymbol{W}^{M+1}, \tag{3.3a}$$

$$\begin{bmatrix} \delta_1^k \\ \delta_2^k \\ \vdots \\ \delta_q^k \end{bmatrix} = \boldsymbol{W}^k \begin{bmatrix} \delta_1^{k+1} F(u_1^{k+1}) \\ \delta_2^{k+1} F(u_2^{k+1}) \\ \vdots \\ \delta_p^{k+1} F(u_p^{k+1}) \end{bmatrix}, \tag{3.3b}$$

$$\boldsymbol{W}^k = \begin{bmatrix} W_{1,1}^{k+1,k} & W_{2,1}^{k+1,k} & W_{3,1}^{k+1,k} & \cdots & W_{p,1}^{k+1,k} \\ W_{1,2}^{k+1,k} & W_{2,2}^{k+1,k} & W_{3,2}^{k+1,k} & \cdots & W_{p,2}^{k+1,k} \\ W_{1,3}^{k+1,k} & W_{2,3}^{k+1,k} & W_{3,3}^{k+1,k} & \cdots & W_{p,3}^{k+1,k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ W_{1,q}^{k+1,k} & W_{2,q}^{k+1,k} & W_{3,q}^{k+1,k} & \cdots & W_{p,q}^{k+1,k} \end{bmatrix}, \tag{3.3c}$$

where ε is a small positive constant, and function $F(\cdot)$ represents the derivative of the activation function. In this study, $F(\cdot)$ is equivalent to the unit step function because ReLU is used as the activation function in all hidden layers. Moreover, $W_{j,i}^{k+1,k}$ denotes the weight of the connection between the ith neuron of

layer number k and the jth neuron of layer number $(k+1)$. The symbol p in $\boldsymbol{W}^k$ represents the number of neurons (excluding the bias neuron) for layer number $(k+1)$, and the symbol q represents that for layer number k; thus, $\boldsymbol{W}^k$ is a variable-sized matrix depending on the value of k.

(3) The amounts of correction δ_i^1 are added to V_i^{ini} for $i = 2,\ 3,...,\ n-1$ to update the input, i.e.

$$V_i^{new} \leftarrow V_i + \delta_i^1, \quad i = 2,\ 3,...,\ n-1, \tag{3.4}$$

while V_1 and V_n remain fixed at *0* and *1*, respectively.

(4) The updated input $\boldsymbol{V}^{new} = \left[0,\ V_2^{new},\ V_3^{new},\ ...,\ V_{n-1}^{new},\ 1\right]$ is provided to the input layer, and forward propagation to the output layer is executed, resulting in the update of the internal states of neurons.

(5) Steps (2)-(4) are repeated until a termination condition is met.

This algorithm serves as the optimiser of the present approach. It should be noted that during step (3), if the addition of correction amounts results in the values of V_i^{new} exceeding *1*, they are constrained to a maximum value of *1*. Similarly, if the addition of correction amounts causes the values of V_i^{new} to fall below *0*, they are adjusted to a minimum value of *0*.

4. Numerical results and discussion

4.1. Uniformly cooled multi-layered FG plate

4.1.1. Ni-Al_2O_3 heat-resisting FG plate

In the first numerical example, the optimisation of ceramic volume fraction distribution aims to minimise the residual thermal stress of a uniformly cooled Ni-Al_2O_3 FGM [13, 22, 34]. The FGM is a twelve-layered FG plate with a thickness of *10 mm*. The first layer is pure Ni, while the twelfth layer consists of pure Al_2O_3, each with a thickness of *1 mm*. Therefore, $n = 12$ and $a_1 / h = (a_{12} - a_{11}) / h = 0.1$ in Fig.1. The remaining ten layers have a uniform thickness of *0.08* in dimensionless units. This FG plate is subjected to a uniform temperature drop from *300 K* to *100 K* $(\Delta t = -200\ K)$.

Table 3 presents thermoelastic material data for Ni and Al_2O_3 [34]. The modified rule of mixture method is used to estimate the effective material properties, facilitating a direct comparison in optimisation results with previous work [34]. Further information regarding these estimates can be found in [39].

Table 3. Material properties of nickel and alumina [34].

Material properties	Ni	Al_2O_3
Young's modulus [GPa]	199.5	393
Poisson's ratio [-]	0.3	0.25
Thermal expansion coefficient [$\times 10^{-6}$/K]	15.4	7.4

First, the NN structure is optimised using Optuna with selectable mini-batch sizes. The structure tuning process yields a network structure such that $M = 4$ and $\{N_1, N_2, N_3, N_4\} = \{50,\ 50,\ 40,\ 30\}$ for a mini-batch size of *32*. Next, the MSE loss from the NN is evaluated during training over *3000* epochs on both training and

validation datasets. Figure 3 illustrates the dual learning curves obtained through *5*-fold cross-validation. After *3000* epochs of training, the validation loss decreases to approximately 5×10^{-5}. Moreover, despite the absence of dropout or regularisation techniques, the simple NN demonstrates no signs of overfitting.

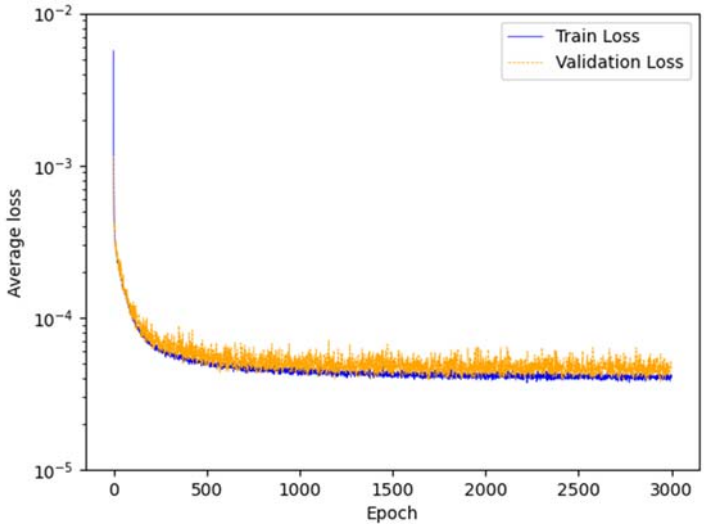

Fig.3. Training and validation learning curves averaged over *5*-fold cross-validation for neural network with $(50-50-40-30)$ hidden layer structure, for a mini-batch size of *32*.

However, as the numbers of neurons in half of the hidden layers have already reached the predetermined maximum value of 50, it is reasonable to anticipate that expanding the exploration space by increasing the upper limit could potentially result in improved network structures. Consequently, an extended structure tuning process is conducted by raising the upper limit to *100*, yielding a larger-scale structure with $M=4$ and $\{N_1,N_2,N_3,N_4\}=\{60,\ 100,\ 90,\ 70\}$ for a mini-batch size of *64*.

Figure 4 shows the dual learning curves produced from this updated NN. As compared to the learning curves of the previous NN with a hidden layer structure of *50-50-40-30*, the validation loss is reduced by half.

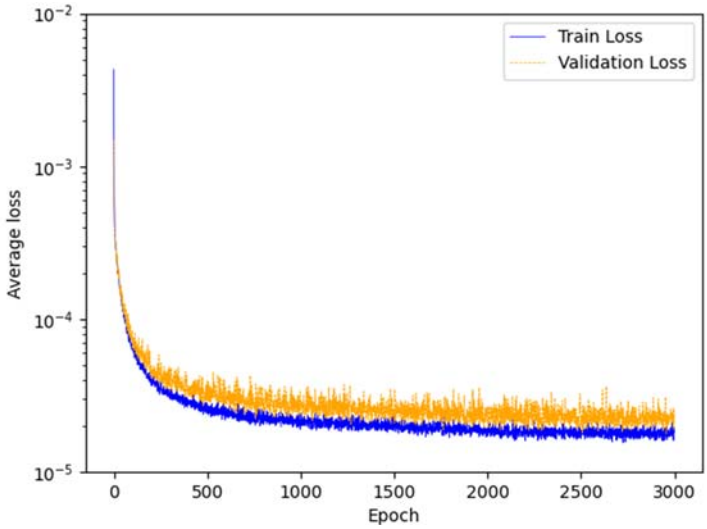

Fig.4. Training and validation learning curves averaged over *5*-fold cross-validation for neural network with $(60-100-90-70)$ hidden layer structure, for a mini-batch size of *64*.

Figure 5 illustrates the actual values of maximum thermal stress for the test dataset alongside those predicted by the NN model. The close clustering of data points around the ideal line reflects the network's capacity to generalise well beyond the training set, as supported by the near-perfect R^2 value.

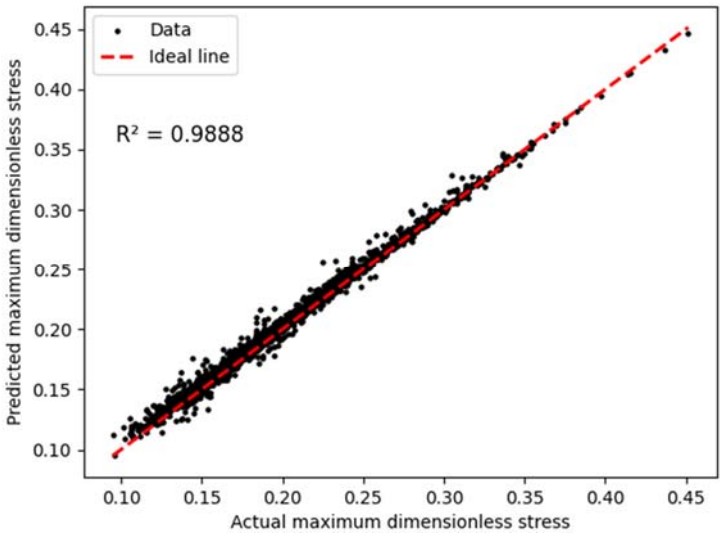

Fig.5. Comparison of actual and predicted maximum thermal stress values for the test dataset.

Subsequently, the optimisation calculation described in Section 3.4 is conducted using the trained NN. In this computation, two different values of ε, namely $\varepsilon = 0.01$ and 0.1, are adopted, and the maximum number of iterations is limited to 1000. The initial volume fraction vector $\boldsymbol{V}^{ini}$ is set to $\boldsymbol{V}^{ini} = [0,\ 0.5,\ 0.5,\ \ldots,\ 0.5,\ 1]$ based on Cho *et al.* [34].

The initial and optimised distributions of the volume fraction and the corresponding thermal stress profiles are compared in Fig.6 and Fig.7, respectively. Both figures include the optimisation results previously

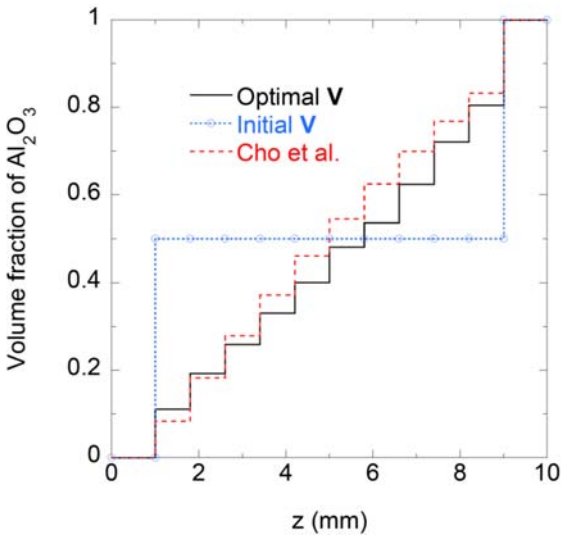

Fig.6. Comparison of initial and optimal distributions of the alumina volume fraction along the z-axis.

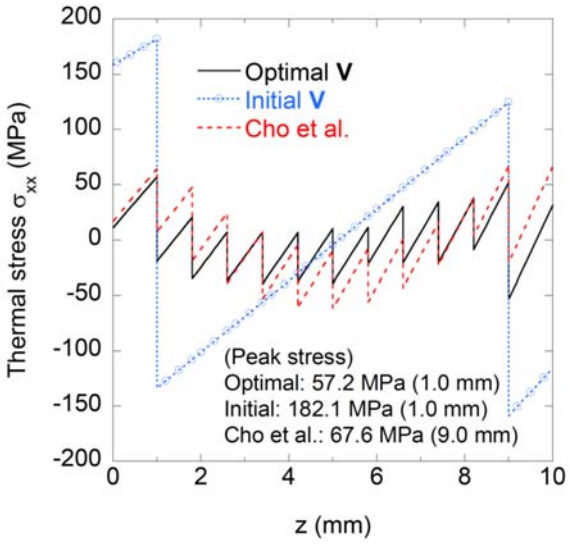

Fig.7. Comparison of in-plane thermal stress distributions along the z-axis.

presented by Cho *et al.* [34]. Because of the stepwise discontinuous volume fractions, the stress distributions exhibit sharp jumps at all interfaces between layers. Although optimisation results do not perfectly match previous results in terms of both volume fraction and thermal stress distributions, a remarkable similarity is observed. This discrepancy arises because the analytical model used by Cho *et al.* is based on a semi-infinite plate, whereas the model here considers an infinite plate. Their model assumes a finite plate length along the *x*-direction, specifically equal to *100 mm*, which is ten times larger than the thickness. However, the distributions are remarkably similar, validating the optimisation results.

4.1.2. C/SiC FGM coating of C/C composite

In contrast to the previous example, which involved a linear variation in ceramic volume fraction, this section addresses a different FGM composition optimisation problem. The optimal ceramic volume fraction distribution is expected to exhibit a sharp increase at a specific location, followed by a change characterised by a power-law type function of the position coordinate. The focus is on minimising the residual thermal stress induced in a C/SiC FGM coating applied to carbon/carbon (C/C) composites, following Xu *et al.* [14].

Consider a C/SiC FGM coating consisting of *11* layers, deposited on a *20-mm*-thick C/C composite substrate. The total thickness of the coating is *2.881 mm*, with the outermost layer being a *1-mm*-thick pure SiC layer. Therefore, $n = 12$ layers including the substrate, where $a_1 / h = 0.874$ and $\left(a_{12} - a_{11}\right) / h = 0.044$ in Fig.1. The remaining ten layers have a uniform thickness of *0.0082* in dimensionless units. This C/C composite coated with a C/SiC FGM undergoes cooling from a sintering temperature of *1573 K* to room temperature $\left(298\ K\right)$ under a uniform temperature field $\left(\Delta t = -1275\ K\right)$.

Table 4 presents material properties of the constituents [14]. In Xu *et al.* study [14], the FGM coating system is assumed to be a strip in a plane stress state in depth (i.e. the *y*-direction in Fig.1 of the present paper); thus, the Poisson's ratios of the constituents are not provided. However, for quantitative comparisons of numerical results, we assume an arbitrary constant Poisson's ratio and convert thermal stress values for a planar coating in a plane strain state, as illustrated in Fig.1, into those for the strip-like coating [40]. The effective material properties are estimated using the linear rule of mixture to ensure consistency with the estimation method employed in [14].

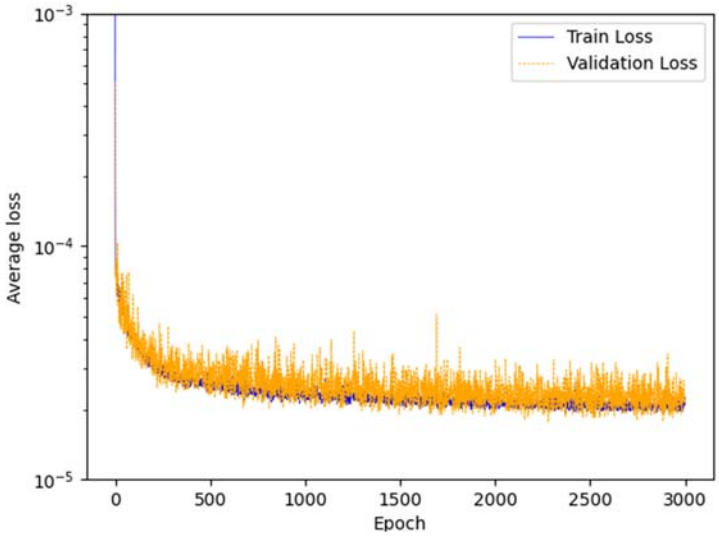

Fig.8. Training and validation learning curves averaged over *5*-fold cross-validation for neural network with $\left(80-80-70\right)$ hidden layer structure, for a mini-batch size of *32*.

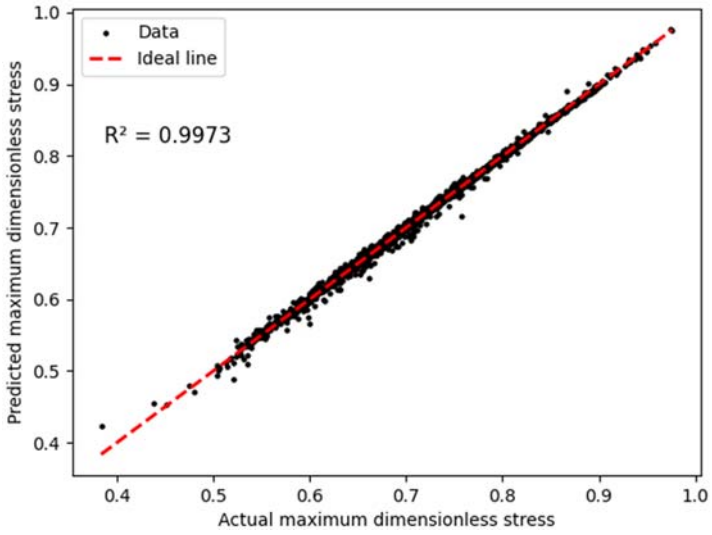

Fig.9. Comparison of actual and predicted maximum thermal stress values for the test dataset.

Table 4. Material properties of carbon and SiC [14].

Material properties	Carbon	SiC
Young's modulus [GPa]	23	466
Thermal expansion coefficient [$\times 10^{-6}$/K]	1.0	4.5

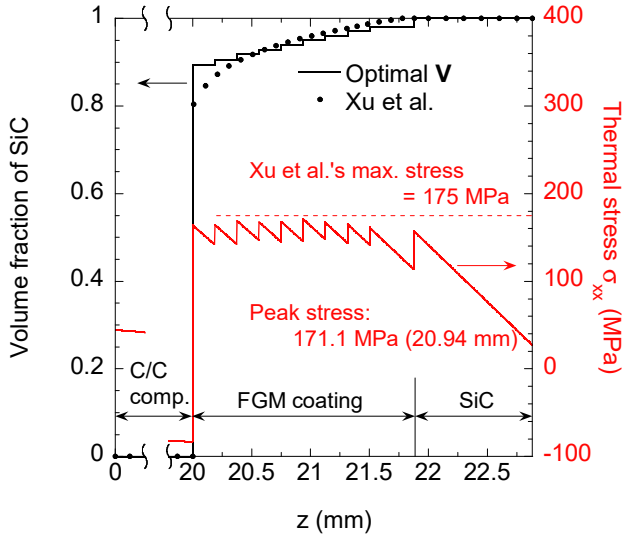

Fig.10. Comparison of optimised distributions of the SiC volume fraction obtained by two different optimisation approaches and resulting thermal stresses.

Similar to Section 4.1.1, first, the optimal NN structure and mini-batch size are determined using Optuna. Because all target values in the datasets exceed *1*, these values are normalised by dividing each by their maximum value $(=33.1)$ to ensure numerical stability. The structure tuning process results in a network architecture of $M=6$ with $\{N_1,N_2,N_3,N_4,N_5,N_6\}=\{50,\ 30,\ 40,\ 50,\ 50,\ 40\}$ for a mini-batch size of *32*. Additionally, increasing the upper limit of the number of neurons per hidden layer up to *100* yields a more compact structure

of $M = 3$ with $\{N_1, N_2, N_3\} = \{80, 80, 70\}$ for the same mini-batch size. Learning curves from the NN and prediction accuracy results after training for *3000* epochs are illustrated in Figs 8 and 9, respectively.

Subsequently, the optimisation calculation described in Section 3.4 is performed using the trained NN. Figure 10 compares our optimised distribution of SiC volume fraction with that of Xu *et al.* [14], obtained using a particle swarm optimisation algorithm. Thermal stresses resulting from both distributions are also depicted on the secondary Y-axis. Xu *et al.* assumed a power-law type function for the SiC volume fraction in the FGM coating and determined the optimal power index to be *0.0739*. Although they did not present the stress distribution in the FGM coating system, they reported a maximum stress of *175 MPa*. In contrast, our optimised volume fraction distribution results in a peak stress below *175 MPa*, indicating a potentially superior distribution. This serves as a clear example that power-law type functions may overly restrict the representation of material gradation.

4.2. Multi-layered FG plate subjected to a temperature gradient

In this numerical example, the optimisation of ceramic volume fraction distribution aims to minimise the thermal stress of a Ni-Al_2O_3 FG plate heated on one surface. The FG plate consists of ten layers of equal thickness, with the first layer being pure Ni and the *10*th layer being pure Al_2O_3, i.e. $n = 10$, $a_1 / h = (a_2 - a_1) / h = \ldots = (a_{10} - a_9) / h = 0.1$ in Fig.1.

Table 5 presents the thermoelastic material data for Ni and Al_2O_3 [41] used in this numerical example. Effective material properties are estimated using the linear rule of mixture, primarily for simplicity.

Table 5. Material properties of nickel and alumina [41].

Material properties	Ni	Al_2O_3
Thermal conductivity [W/(m·K)]	90.1	28.2
Young's modulus [GPa]	218	360
Poisson's ratio [-]	0.31	0.23
Thermal expansion coefficient [$\times 10^{-6}$/K]	14.0	5.4

Moreover, Fig.11 provides a histogram illustrating the maximum of the absolute $\bar{\sigma}_{xx}$ values corresponding to the *10000* randomly generated $\boldsymbol{V}$ for training and validating the NN.

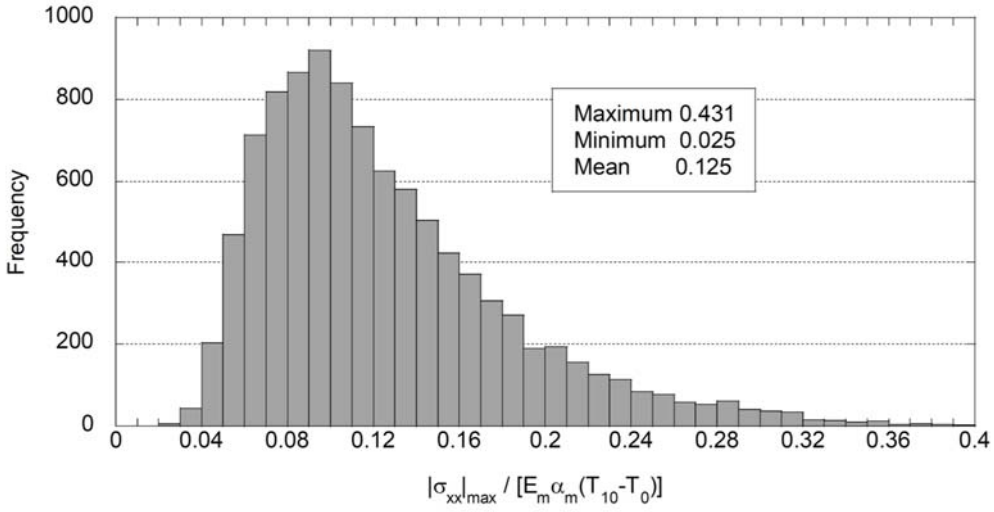

Fig.11. Histogram of $|\bar{\sigma}_{xx}|_{\max}$ for all samples included in the training and validation sets.

Similar to the previous numerical examples, the NN structure is optimised using Optuna for four different mini-batch size options. The resulting network structure is $M = 4$ with $\{N_1, N_2, N_3, N_4\} = \{50,\ 50,\ 50,\ 40\}$ for a mini-batch size of *96*. Figure 12 depicts the dual learning curves obtained during the training process. The training progresses successfully without overfitting; however, a notable *loss spike* is observed in the validation loss, with loss values fluctuating by half an order of magnitude.

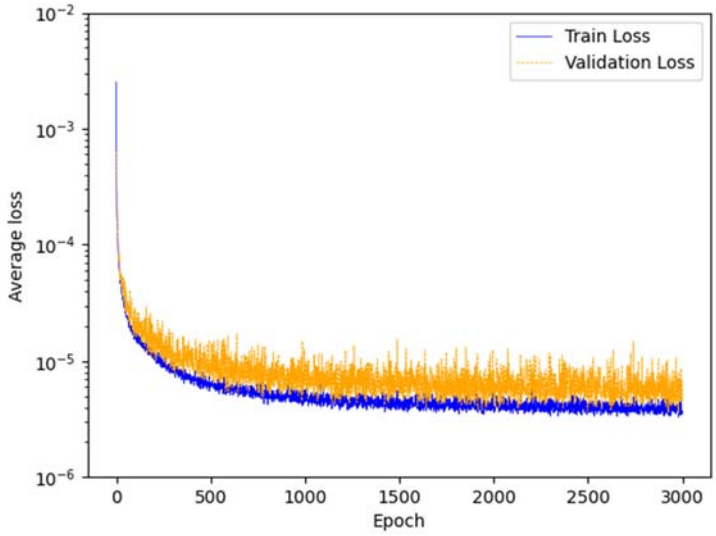

Fig.12. Training and validation learning curves averaged over *5*-fold cross-validation for neural network with $(50-50-50-40)$ hidden layer structure, for a mini-batch size of *96*.

Considering that the numbers of neurons in most of the hidden layers in the tuned NN have reached the pre-set upper limit of *50*, it is plausible that the significant fluctuation in the validation loss may be attributed to the NN's limited expressive power. To address this, a subsequent structure tuning is carried out using Optuna, with a raised upper limit of N_i set to *100*. Through the structure tuning process, a more shallow NN is obtained such that $M = 3$ and $\{N_1, N_2, N_3\} = \{80,\ 90,\ 90\}$ for a mini-batch size of *64*.

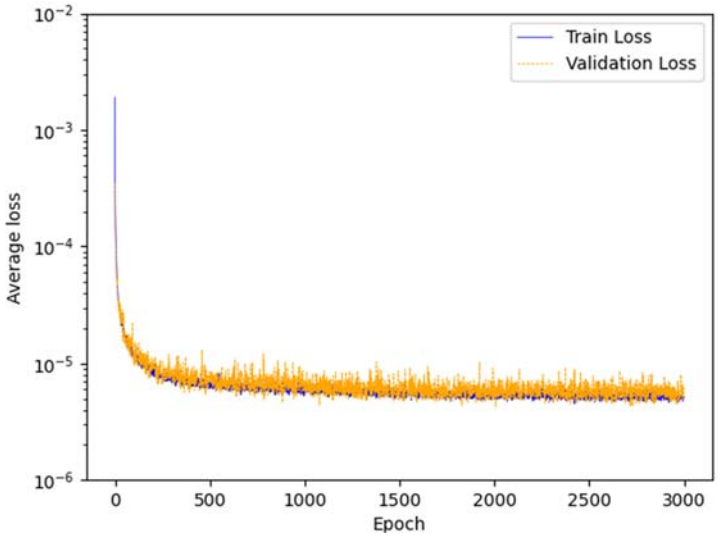

Fig.13. Training and validation learning curves averaged over *5*-fold cross-validation for neural network with $(80-90-90)$ hidden layer structure, for a mini-batch size of *64*.

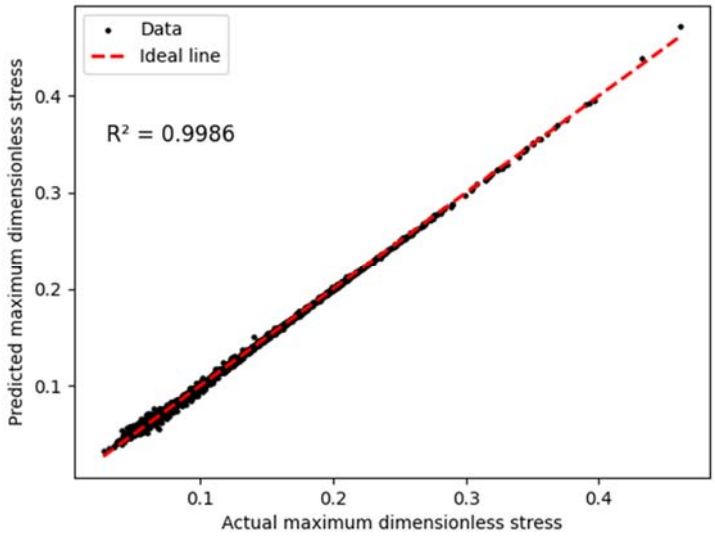

Fig.14. Comparison of actual and predicted maximum thermal stress values for the test dataset.

Figure 13 displays the dual learning curves generated by this NN. The validation loss closely approaches the train loss, and its spike is significantly reduced compared to that shown in Fig.12. Additionally, Fig.14 underscores the trained NN's precision, with an R^2 of *0.9986* highlighting its accuracy in predicting the maximum thermal stress for FG plates with unseen data of $\boldsymbol{V}$. The close correlation between predicted and actual stresses confirms the model's reliability for the subsequent optimisation process based on it.

The optimisation calculation using this NN produces the optimal distribution of Al_2O_3 volume fraction and the corresponding thermal stress distribution, as shown in Figs 15 and 16, respectively. For consistency with the first example, $\boldsymbol{V}^{ini}$ is set as $[0, 0.5, 0.5, \ldots, 0.5, 1]$. The variation of the objective function throughout the iteration during the optimisation process is illustrated in Fig.17. The optimal $\boldsymbol{V}$ shown in Fig.15 represents a plausible solution leading to the most significant reduction in thermal stress.

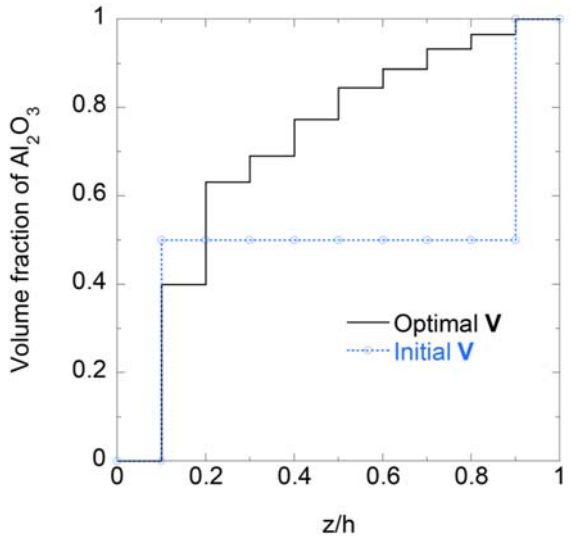

Fig.15. Comparison of initial and optimal distributions of the alumina volume fraction along the thickness direction (*z*-axis)

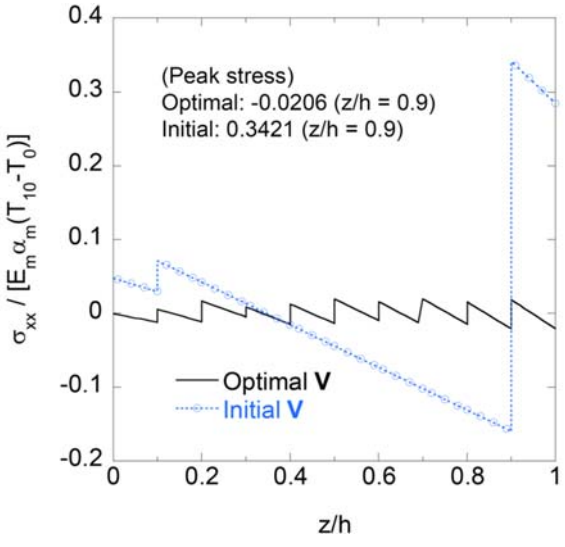

Fig.16. Comparison of in-plane thermal stress distributions along the thickness direction (*z*-axis) in dimensionless units.

This is supported by the fact that the dimensionless $|\sigma_{xx}|_{max}$ value for the optimal $\boldsymbol{V}$ is *0.0206*, as shown in Fig.16, whereas the minimum among those corresponding to the *10000* preliminarily generated $\boldsymbol{V}$ is *0.0250*, as shown in Fig.11. This indicates an approximately *18%* decrease in dimensionless $|\sigma_{xx}|_{max}$ value.

In the numerical computation described above, training the NN over *3000* epochs takes *14* minutes in the Google Colaboratory computation environment, utilising a GPU (NVIDIA Tesla T4) with *16 GB* RAM. Furthermore, the optimisation calculation using the trained NN requires around *22* minutes for *1000* iterations. However, incorporating a convergence criterion in the termination condition can reduce the computational cost to approximately *30%*, by the look of the convergence observed at around *300* iterations, as shown in Fig.17.

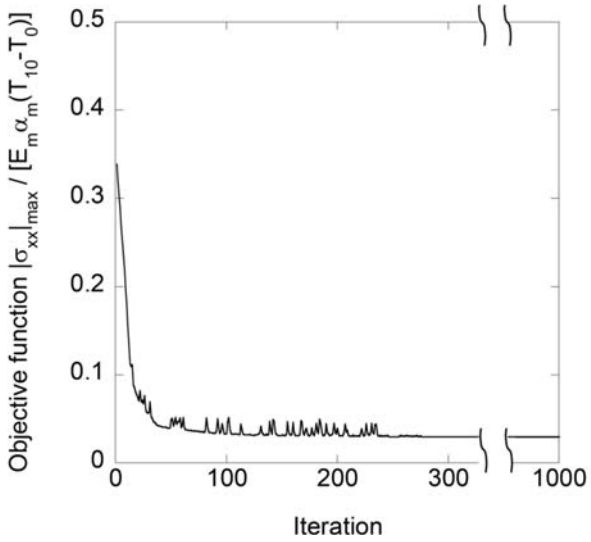

Fig.17. Convergence of objective function over iterations for ε=*0.1*.

5. Conclusions

This study applied a NN-based optimisation method, which builds upon the pioneering work of Ootao *et al.* [28], to determine optimal material composition distributions for minimising steady-state thermal stresses in multi-layered heat-resistant FG plates. The approach treats the volume fractions of the ceramic constituent in each layer as design variables, enabling a precise optimisation process. By applying a backpropagation algorithm to a NN predictor, the optimal set of volume fractions is efficiently identified to reduce the maximum in-plane thermal stress.

A key novelty of this study lies in the use of Optuna, an open-source hyperparameter optimisation tool powered by the TPE, a Bayesian optimisation algorithm. This advanced framework facilitates the determination of hyperparameters associated with the NN structure, thereby enhancing the efficiency and effectiveness of the optimisation process.

Numerical computations conducted on ten- and twelve-layered FG plates subjected to temperature changes demonstrate the practical applicability of the presented optimisation approach. The NN model, trained on a dataset of *8000* samples, consistently yields valid optimal solutions for optimisation problems involving around ten design variables within a reasonable timeframe on a GPU. Network structure tuning with Optuna indicates that NNs with up to *100* neurons per hidden layer are suitable for solving material composition optimisation problems of this scale, provided that *8000* samples are used for NN training.

Implementing the NN-based algorithms is relatively straightforward in a Python environment and incurs minimal costs, as many libraries and tools are available for free. This ease of implementation and cost-effectiveness makes the NN-based approach a practical choice for optimisation problems.

Nomenclature

a – interface position
E – Young's modulus
f – objective function
F – derivative of activation function
h – thickness
i – index w.r.t. layer number
j – index w.r.t. neuron number
k – index w.r.t. network layer number
M – number of hidden layers
n – number of layers
N – number of neurons
R^2 – coefficient of determination
T – temperature
u – internal state of neuron
V – ceramic volume fraction
$\boldsymbol{V}$ – vector of ceramic volume fractions
W – connection weight between different neurons
x – length direction coordinate
y – width direction coordinate
z – thickness direction coordinate
α – thermal expansion coefficient
δ – correction amount
Δt – reference temperature difference
ε – small positive constant
λ – thermal conductivity
ν – Poisson's ratio
σ – stress
$\bar{\sigma}$ – dimensionless stress

References

[1] Neubrand A. and Rödel J. (1997): *Gradient materials: an overview of a novel concept.*– Int. J. Materials Research, vol.88, No.5, pp.358-371, 10.3139/ijmr-1997-0066.

[2] Nikbakht S., Kamarian S. and Shakeri M. (2019): *A review on optimization of composite structures Part II: Functionally graded materials.*– Composite Structures, vol.214, pp.83-102, 10.1016/j.compstruct.2019.01.105.

[3] Nayak P. and Armani A. (2022): *Optimal design of functionally graded parts.*– Metals, vol.12, No.8, pp.1335, 10.3390/met12081335.

[4] Bobaru F. (2007): *Designing optimal volume fractions for functionally graded materials with temperature-dependent material properties.*– Trans. ASME J. Applied Mech., vol.74, No.5, pp.861-874, 10.1115/1.2712231.

[5] Taheri A.H., Hassani B. and Moghaddam N.Z. (2014): *Thermo-elastic optimization of material distribution of functionally graded structures by an isogeometrical approach.*– Int. J. Solids Structures, vol.51, No.2, pp.416-429, 10.1016/j.ijsolstr.2013.10.014.

[6] Abdalla H.M.A., Casagrande D. and Moro L. (2020): *Thermo-mechanical analysis and optimization of functionally graded rotating disks.*– J. Strain Analysis Eng. Design, vol.55, No.5-6, pp.159-171, 10.1177/0309324720904793.

[7] Roque C.M.C. and Martins P.A.L.S. (2015): *Differential evolution for optimization of functionally graded beams.*– Composite Structures, vol.133, pp.1191-1197, 10.1016/j.compstruct.2015.08.041.
[8] Ootao Y., Kawamura R., Tanigawa Y. and Ishimaru O. (1998): *Optimization of material composition of hollow circular cylinder of functionally graded material for thermal stress relaxation making use of genetic algorithm.*– Trans. Japan Soc. Mech. Eng. Series A, vol.64, No.626, pp.2645-2652, 10.1299/kikaia.64.2645.
[9] Shimojima K., Yamada Y., Mabuchi M., Saito N., Nakanishi M., Shigematsu I., Nakamura M., Asahina T. and Igarashi T. (1999): *Optimization method of FGM compositional distribution profile design by genetic algorithm.*– Materials Science Forum, vol.308-311, pp.1006-1011.
[10] Ootao Y., Tanigawa Y. and Ishimaru O. (2000): *Optimization of material composition of functionally graded plate for thermal stress relaxation using a genetic algorithm.*– J. Thermal Stresses, vol.23, No.3, pp.257-271, 10.1080/014957300280434.
[11] Goupee A.J. and Vel S.S. (2006): *Two-dimensional optimization of material composition of functionally graded materials using meshless analyses and a genetic algorithm.*– Computer Methods Applied Mech. Eng., vol.195, No.44, pp.5926-5948, 10.1016/j.cma.2005.09.017.
[12] Chiba R. and Sugano Y. (2012): *Optimisation of material composition of functionally graded materials based on multiscale thermoelastic analysis.*– Acta Mechanica, vol.223, No.5, pp.891-909, 10.1007/s00707-011-0610-z.
[13] Fereidoon A., Sadri F. and Hemmatian H. (2012): *Functionally graded materials optimization using particle swarm-based algorithms.*– J. Thermal Stresses, vol.35, No.4, pp.377-392, 10.1080/01495739.2012.663688.
[14] Xu Y., Zhang W., Chamoret D. and Domaszewski M. (2012): *Minimizing thermal residual stresses in C/SiC functionally graded material coating of C/C composites by using particle swarm optimization algorithm.*– Computational Materials Science, vol.61, pp.99-105, 10.1016/j.commatsci.2012.03.030.
[15] Kou X.Y., Parks G.T. and Tan S.T. (2012): *Optimal design of functionally graded materials using a procedural model and particle swarm optimization.*– Computer-Aided Design, vol.44, No.4, pp.300-310, 10.1016/j.cad.2011.10.007.
[16] Eldeeb A., Shabana Y. and Elsawaf A. (2023): *Thermoelastic stress mitigation and weight reduction of functionally graded multilayer nonuniform thickness disc.*– J. Strain Analysis Eng. Design, vol.58, No.8, pp.661-671, 10.1177/03093247231165091.
[17] Eldeeb A.M., Shabana Y.M. and Elsawaf A. (2023): *Particle swarm optimization for the thermoelastic behaviors of functionally graded rotating nonuniform thickness sandwich discs.*– Arabian J. Science Eng., vol.48, No.3, pp.4067-4079, 10.1007/s13369-022-07351-x.
[18] Kamgar R., Rahmani F. and Rahgozar R. (2024): *Geometrical and material optimization of the functionally graded doubly-curved shell by metaheuristic optimization algorithms.* Structures, vol.62, pp.106254, doi.org/10.1016/j.istruc.2024.106254.
[19] Mozafari H., Ayob A. and Kamali F. (2012): *Optimization of functional graded plates for buckling load by using imperialist competitive algorithm.*– Procedia Technology, vol.1, pp.144-152, 10.1016/j.protcy.2012.02.028.
[20] Cross S.R., Woollam R., Shademan S. and Schuh C.A. (2013): *Computational design and optimization of multilayered and functionally graded corrosion coatings.*– Corrosion Science, vol.77, pp.297-307, 10.1016/j.corsci.2013.08.018.
[21] Lieu Q.X. and Lee J. (2017): *Modeling and optimization of functionally graded plates under thermo-mechanical load using isogeometric analysis and adaptive hybrid evolutionary firefly algorithm.*– Composite Structures, vol.179, pp.89-106, 10.1016/j.compstruct.2017.07.016.
[22] Cho J.R. and Shin S.W. (2004): *Material composition optimization for heat-resisting FGMs by artificial neural network.*– Composites Part A, vol.35, No.5, pp.585-594, 10.1016/j.compositesa.2003.12.003.
[23] Yas M.H., Kamarian S. and Pourasghar A. (2014): *Application of imperialist competitive algorithm and neural networks to optimise the volume fraction of three-parameter functionally graded beams.*– J. Experimental Theoretical Artificial Intelligence, vol.26, No.1, pp.1-12, 10.1080/0952813X.2013.782346.
[24] Do D.T.T., Lee D. and Lee J. (2019): *Material optimization of functionally graded plates using deep neural network and modified symbiotic organisms search for eigenvalue problems.*– Composites Part B, vol.159, pp.300-326, 10.1016/j.compositesb.2018.09.087.
[25] Truong T.T., Lee S. and Lee J. (2020): *An artificial neural network-differential evolution approach for optimization of bidirectional functionally graded beams.*– Composite Structures, vol.233, pp.111517, 10.1016/j.compstruct.2019.111517.
[26] Do D.T.T., Nguyen-Xuan H. and Lee J. (2020): *Material optimization of tri-directional functionally graded plates by using deep neural network and isogeometric multimesh design approach.*– Applied Mathematical Modelling, vol.87, pp.501-533, 10.1016/j.apm.2020.06.002.

[27] Truong T.T., Lee J. and Nguyen-Thoi T. (2021): *Multi-objective optimization of multi-directional functionally graded beams using an effective deep feedforward neural network-SMPSO algorithm.*– Structural Multidisciplinary Optimization, vol.63, No.6, pp.2889-2918, 10.1007/s00158-021-02852-z.

[28] Ootao Y., Kawamura R., Tanigawa Y. and Nakamura, T. (1998): *Neural network optimization of material composition of a functionally graded material plate at arbitrary temperature range and temperature rise.*– Archive of Applied Mechanics, vol.68, No.10, pp.662-676, 10.1007/s004190050195.

[29] Ootao Y., Tanigawa Y. and Nakamura T. (1999): *Optimization of material composition of FGM hollow circular cylinder under thermal loading: a neural network approach.*– Composites Part B, vol.30, No.4, pp.415-422, 10.1016/S1359-8368(99)00003-7.

[30] Ootao Y., Kawamura R., Tanigawa Y. and Imamura R. (1999): *Optimization of material composition of nonhomogeneous hollow sphere for thermal stress relaxation making use of neural network.*– Computer Methods Applied Mech. Eng., vol.180, No.1, pp.185-201, 10.1016/S0045-7825(99)00055-9.

[31] Ootao Y., Kawamura R., Tanigawa Y. and Imamura R. (1999): *Optimization of material composition of nonhomogeneous hollow circular cylinder for thermal stress relaxation making use of neural network.*– J. Thermal Stresses, vol.22, No.1, pp.1-22, 10.1080/014957399281020.

[32] Eldeeb A., Shabana Y. and Elsawaf A. (2021): *Thermo-elastoplastic behavior of a rotating sandwich disc made of temperature-dependent functionally graded materials.*– J. Sandwich Structures Materials, vol.23, No.5, pp.1761-1783, 10.1177/1099636220904970.

[33] Fathi R., Wei H., Saleh B., Radhika N., Jiang J., Ma A., Ahmed M.H., Li Q. and Ostrikov K.K. (2022): *Past and present of functionally graded coatings: Advancements and future challenges.*– Applied Materials Today, vol.26, pp.101373, 10.1016/j.apmt.2022.101373.

[34] Cho J.R. and Ha D.Y. (2002): *Volume fraction optimization for minimizing thermal stress in Ni–Al2O3 functionally graded materials.*– Materials Science Eng. A, vol.334, No.1, pp.147-155, 10.1016/S0921-5093(01)01791-9.

[35] Sugano Y. (1987): *An expression for transient thermal stress in a nonhomogeneous plate with temperature variation through thickness.*– Ing. Arch., vol.57, No.2, pp.147-156, 10.1007/BF00541388.

[36] Bergstra J., Bardenet R., Bengio Y. and Kégl B. (2011): *Algorithms for hyper-parameter optimization.*– In the 24th Int. Conf. Neural Information Processing Systems. (NY, USA: Curran Assosiates Inc.), pp.2546-2554.

[37] Akiba T., Sano S., Yanase T., Ohta T. and Koyama, M. (2019): *Optuna: a next-generation hyperparameter optimization framework.*– In the 25th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining. (Anchorage, USA), pp.2623-2631.

[38] Tanaka M., Hanahara K. and Seguchi, Y. (1992): *Configuration control of the truss-type parallel manipulator by the modular neural network model.*– JSME Int. J. Series 3, vol.35, No.1, pp.89-95, 10.1299/jsmec1988.35.89.

[39] Cho J.R. and Ha D.Y. (2001): *Averaging and finite-element discretization approaches in the numerical analysis of functionally graded materials.*– Materials Science Eng. A, vol.302, No.2, pp.187-196, 10.1016/S0921-5093(00)01835-9.

[40] Zhang X.C., Xu B.S., Wang H.D., Jiang Y. and Wu Y.X. (2006): *Modeling of thermal residual stresses in multilayer coatings with graded properties and compositions.*– Thin Solid Films, vol.497, No.1-2, pp.223-231, 10.1016/j.tsf.2005.09.184.

[41] Awaji H., Takenaka H., Honda S. and Nishikawa T. (2001): *Temperature/stress distributions in a stress-relief-type plate of functionally graded materials under thermal shock.*– JSME Int. J. Series A, vol.44, No.1, pp.37-44, 10.1299/jsmea.44.37.

Received: April 25, 2024
Revised: August 13, 2024