Journal of Science and Technology in Civil Engineering, HUCE (NUCE), 2022, 16 (3): 1–17

LONG SHORT-TERM MEMORY FOR NONLINEAR STATIC ANALYSIS OF FUNCTIONALLY GRADED PLATES

Dieu T. T. Do[a], Son Thai[b,c], Tinh Quoc Bui[d,*]

[a]*Duy Tan Research Institute for Computational Engineering, Duy Tan University, 254 Nguyen Van Linh street, Da Nang, Vietnam*
[b]*Faculty of Civil Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam*
[c]*Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc City, Ho Chi Minh City, Vietnam*
[d]*Department of Civil and Environmental Engineering, Tokyo Institute of Technology, Tokyo, Japan*

Article history:
Received 10/3/2022, Revised 21/3/2022, Accepted 01/4/2022

Abstract

This study presents an effective method based on long short-term memory to reduce the computational cost in nonlinear static analysis of functionally graded plates. Data points representing a load-deflection curve in a dataset are generated through isogeometric analysis. The order of these data points is always maintained as a sequential series of observations; therefore, it is referred to as a time series. Dataset is divided into three sets including training, testing, and prediction sets. Both training and testing sets are used for the training process by the long short-term memory to gain optimum weights. Based on these obtained weights, data points in the prediction set are directly predicted without using any analysis tools. The effectiveness and accuracy of the proposed method are demonstrated by comparing the obtained results to those of isogeometric analysis.

Keywords: long short-term memory; time series forecasting; nonlinear analysis; functionally graded plate.

https://doi.org/10.31814/stce.huce(nuce)2022-16(3)-01

1. Introduction

Functionally graded materials (FGMs) are known as advanced materials with properties that change in a specific direction, which are often made up of two or more components. Metal and ceramic materials are frequently used in structural applications because of their advantages. Specifically, the ceramic phase is excellent at withstanding high temperatures, while the metal phase has a high fracture toughness. Overcoming laminated composite drawbacks, the FGMs completely eliminate undesirable stress discontinuity between two layers in laminated composites. Therefore, FGMs have been used in a wide range of fields, including aircraft engineering, nuclear power plants, and electrical engineering as in Refs. [1–4].

Isogeometric analysis (IGA) [5] was proposed as a way to combine computer-aided design (CAD) and finite element analysis (FEA). The same non-uniform rational B-spline (NURBS) is used in IGA to represent exact CAD geometry and approximate FEA solution fields. Furthermore, even at the

*Corresponding author. *E-mail address:* tinh.buiquoc@gmail.com (Bui, T. Q.)

coarsest discretization level, the exact geometry is preserved, and this method is effective in reducing degrees of freedom (DOFs) for high-order elements. As a result, IGA has been widely applied in a variety of engineering fields [6–8]. For example, Kim et al. [6] proposed IGA for geometrically exact shell considering the first-order shear deformation. The accuracy and robustness of the approach have been demonstrated by the test results of numerical examples using the developed method, which show a higher convergence rate. In modeling of the plate structure, the formulation can be reduced to a linear problem. However, the real response of structures is usually deviated by linear solutions. Therefore, geometrically nonlinear analysis is necessary to investigate the plate behavior. To solve such nonlinear problems, many methods have been introduced such as Picard method [9], Newton–Raphson method [10], modified Newton–Raphson method [11], modified Riks method [12], arclength technique [13], and other methods. Most of the above methods require many iterations to obtain nonlinear solutions. As a result, it necessitates a significant amount of computational cost. In order to overcome this issue, this study proposes an approach that uses long short-term memory (LSTM) [14] to directly predict the nonlinear behavior of the FG plate without using any analysis tools. Therefore, it saves a significant amount of computational cost while maintaining the accuracy of the obtained solution. LSTM is a deep learning technique for time series prediction. Deep learning is currently one of the most popular fields, with numerous studies [15–19] etc. For example, Huynh et al. [15] proposed machine learning-assisted numerical methods for predicting compressive strength of fly ash-based geopolymer (FAGP) concrete. Different methods including artificial neural networks (ANN), deep neural networks (DNN), and deep residual networks (ResNet) were evaluated based on experimentally collected data in terms of R-squared (R2), root mean square error (RMSE), and mean absolute percentage error (MAPE). Lieu et al. [18] presented a simple and effective adaptive surrogate model based on DNN to structural reliability analysis. With only a few experiments, the surrogate model for MCS-based failure probability assessment becomes more precise.

LSTM is a type of recurrent neural network (RNN) that excels at learning and predicting sequential data. The ability of RNN to maintain long-term memory is limited. As a result, the LSTM was created to overcome this restriction by adding a memory structure that can maintain its state over time, with gates that determine what to remember, forget, and output. The LSTM performs well in a variety of inherently sequential applications, including speech recognition [20], language modeling and translation [21], speech synthesis [22], emotion recognition [23] and handwriting recognition [24], etc. For example, Chien et al. [20] proposed a combination of deep feedforward and recurrent neural networks for the acoustic model. Under various tasks and conditions, this combination improved noisy speech recognition performance. Do et al. [25] proposed LSTM and multi-layer neural network for predicting the crack propagation in risk assessment of engineering structures. This approach helps reduce the amount of computational cost when it can directly predict the propagation of a crack without using any analysis tools. According to the literature, the use of LSTM in the structural engineering field is still limited. Therefore, LSTM is proposed for the first time in this study to predict the nonlinear behavior of FG plates. This research will contribute to the filling of a gap in the literature.

In this study, a dataset is generated through analysis using IGA to ensure the accuracy of solution. Data points in the dataset represent the load-deflection curve of the FG plate. These data will be rearranged as a supervised learning problem; however, the order of data points is always maintained in the dataset. In each data pair in the supervised learning problem, the input and output variables are data from the previous time step and data from the next time step, respectively. The dataset is divided into three sets: training, testing and prediction sets. Training and testing sets are used to gain optimum weights by using LSTM. Based on these weights, data in the prediction set are directly predicted

without using any analysis tools. Specifically, data from the previous time step will be used to predict data for the subsequent time step. This approach helps to save a significant amount of computational cost while ensuring the accuracy of gained solutions. The effects of optimizers such as SGD, Adadelta [26], Adamax [27], RMSprop [28], and Adam optimizer [27], as well as the number of neurons in each hidden layer such as 10, 30, 50, 70, and 90 on method accuracy will be investigated in this study. The effectiveness and efficiency of the proposed method are demonstrated by comparing the obtained results to those of IGA.

2. Isogeometric nonlinear analysis of functionally graded plates

2.1. Functionally graded materials

Functionally graded materials (FGMs) are novel materials made by combining two distinct material phases, such as ceramic and metal, whose properties change gradually with their dimensions, as illustrated in Fig. 1. The volume fractions of the FGMs are assumed to be determined by a power-law function as follows:

$$V_c(z) = \left(\frac{1}{2} + \frac{z}{h}\right)^n, \quad V_c + V_m = 1 \tag{1}$$

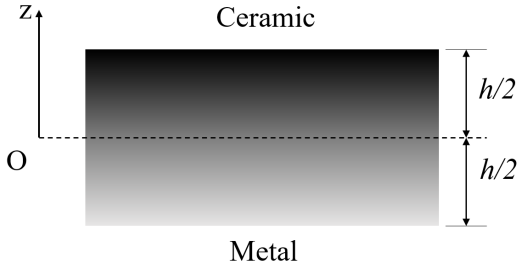


Figure 1. A functionally graded material layer

where n symbolizes power or gradient index; V_c and V_m represent ceramic and metal volume fractions; h is thickness.

The variation of material properties along the thickness of the plate is reflected by the rule of mixture as follows:

$$P_e = P_c V_c + P_m V_m \tag{2}$$

where P_e symbolizes the effective material properties such as Young's modulus (E), Poisson's ratio (ν), density (ρ); P_c and P_m symbolize the properties of ceramic and metal surfaces, respectively.

2.2. Plate formulation

The displacement field of an arbitrary point $\mathbf{u} = \{u, v, w\}^T$ according to the generalized shear deformation plate theory [7] can be described as follows:

$$\mathbf{u} = \mathbf{u}_1 + z\mathbf{u}_2 + f(z)\,\mathbf{u}_3 \tag{3}$$

in which $\mathbf{u}_1 = \{u_0, v_0, w_0\}^T$ represents displacement components in the x, y and z axes; $\mathbf{u}_2 = -\left\{w_{0,x}, w_{0,y}, 0\right\}^T$ and $\mathbf{u}_3 = \left\{\beta_x, \beta_y, 0\right\}^T$ represent the rotations in the xz, yz and xy planes, respectively; $f(z) = z - 4z^3/\left(3h^2\right)$.

In strain-displacement relations, the von Karman nonlinear theory is used as follows [29]:

$$\left\{ \begin{array}{c} \boldsymbol{\varepsilon} \\ \boldsymbol{\gamma} \end{array} \right\} = \left\{ \begin{array}{c} \boldsymbol{\varepsilon}_m \\ \mathbf{0} \end{array} \right\} + \left\{ \begin{array}{c} z\boldsymbol{\kappa}_1 \\ \mathbf{0} \end{array} \right\} + \left\{ \begin{array}{c} f(z)\,\boldsymbol{\kappa}_2 \\ f'(z)\,\boldsymbol{\beta} \end{array} \right\} \tag{4}$$

where

$$\begin{aligned}\boldsymbol{\varepsilon}_m &= \begin{bmatrix} u_{0,x} \\ v_{0,y} \\ u_{0,y}+v_{0,x} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} w_{0,x}^2 \\ w_{0,y}^2 \\ 2w_{0,x}w_{0,y} \end{bmatrix} = \boldsymbol{\varepsilon}_L + \boldsymbol{\varepsilon}_{NL}, \\ \boldsymbol{\kappa}_1 &= -\begin{bmatrix} w_{0,xx} \\ w_{0,yy} \\ 2w_{0,xy} \end{bmatrix}, \quad \boldsymbol{\kappa}_2 = \begin{bmatrix} \beta_{x,x} \\ \beta_{y,y} \\ \beta_{x,y}+\beta_{y,x} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_x \\ \beta_y \end{bmatrix}\end{aligned} \tag{5}$$

The nonlinear component of in-plane strain in Eq. (5) can be rewritten as follows:

$$\boldsymbol{\varepsilon}_{NL} = \frac{1}{2}\boldsymbol{A}_{\boldsymbol{\theta}}\boldsymbol{\theta} \tag{6}$$

where

$$\boldsymbol{A}_{\boldsymbol{\theta}} = \begin{bmatrix} w_{0,x} & 0 \\ 0 & w_{0,y} \\ w_{0,y} & w_{0,x} \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{Bmatrix} w_{0,x} \\ w_{0,y} \end{Bmatrix} \tag{7}$$

The constitutive relation for the FG plate is as:

$$\begin{Bmatrix} \boldsymbol{\sigma} \\ \boldsymbol{\tau} \end{Bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\gamma} \end{Bmatrix} \tag{8}$$

in which the matrices for the materials are as follows:

$$\mathbf{C} = \frac{E_e}{1-\nu_e^2}\begin{bmatrix} 1 & \nu_e & 0 \\ \nu_e & 1 & 0 \\ 0 & 0 & (1-\nu_e)/2 \end{bmatrix}, \quad \mathbf{G} = \frac{E_e}{2(1+\nu_e)}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{9}$$

Stress resultants are given by

$$\underbrace{\begin{Bmatrix} \mathbf{N} \\ \mathbf{M} \\ \mathbf{P} \\ \mathbf{Q} \end{Bmatrix}}_{\hat{\boldsymbol{\sigma}}} = \underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{E} & \mathbf{0} \\ \mathbf{B} & \mathbf{D} & \mathbf{F} & \mathbf{0} \\ \mathbf{E} & \mathbf{F} & \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}^s \end{bmatrix}}_{\hat{\mathbf{D}}} \underbrace{\begin{Bmatrix} \boldsymbol{\varepsilon}_m \\ \boldsymbol{\kappa}_1 \\ \boldsymbol{\kappa}_2 \\ \boldsymbol{\beta} \end{Bmatrix}}_{\hat{\boldsymbol{\varepsilon}}} \tag{10}$$

where

$$\begin{aligned}A_{ij}, B_{ij}, D_{ij}, E_{ij}, F_{ij}, H_{ij} &= \int_{-h/2}^{h/2} \left(1, z, z^2, f(z), zf(z), f^2(z)\right) C_{ij} dz, \\ D_{ij}^s &= \int_{-h/2}^{h/2} \left[f'(z)\right]^2 G_{ij} dz\end{aligned} \tag{11}$$

The variation of total energy of the plate can be calculated using the virtual displacement principle as follows:

$$\delta\Pi = \delta U_\varepsilon - \delta V = \int_\Omega \delta\hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\sigma}} d\Omega - \int_\Omega \delta\mathbf{u}^T f_z d\Omega = 0 \tag{12}$$

in which f_z symbolizes the transverse load.

2.3. Isogeometric nonlinear analysis

In this study, isogeometric approach (IGA) [5] is utilized for the nonlinear analysis of the FG plates because it is more computationally efficient than the traditional finite element method [30]. In IGA, the knot vector, which is given as a set of parameters $\Xi = \left\{\xi_1, \xi_2, ..., \xi_i, ..., \xi_{n+p+1}\right\}, \xi_i \leq \xi_{i+1}$ with n symbolizing the number of the basis functions, is utilized to construct the B-splines basis function of degree p. The univariate B-spline basis functions $N_i^p(\xi)$ are determined recursively on the corresponding knot vector using the Cox-de Boor algorithm as follows:

$$\begin{aligned} &N_i^p(\xi) = \frac{\xi - \xi_i}{\xi_{i+p} - \xi_i} N_i^{p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1}^{p-1}(\xi), \\ &\text{For } p = 0, \quad N_i^0(\xi) = \begin{cases} 1, & \text{if } \xi_i \leq \xi < \xi_{i+1} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{13}$$

In general, the multivariate B-spline basis functions are created as:

$$N_i^p(\xi) = \prod_{\alpha=1}^{d} N_{i_\alpha}^{p_\alpha}(\xi^\alpha) \tag{14}$$

where $d = 1, 2, 3$ corresponds to 1D, 2D and 3D spaces, respectively.

Non-uniform rational B-spline (NURBS) provides a more generalized approach in the form of rational functions for some conic shapes (e.g., circles, spheres, ellipses, etc.):

$$R_i^p(\xi) = N_i^p(\xi)\zeta_i \Big/ \sum_j N_j^p(\xi)\zeta_j \tag{15}$$

in which ζ_i is the weight corresponding to B-spline basis function $N_i^p(\xi)$.

The displacement variables are defined based on NURBS basis functions as follows:

$$u^h(\xi) = \sum_A R_A(\xi)\mathbf{q}_A \tag{16}$$

where $\mathbf{q}_A = \left[u_{0A}, v_{0A}, \beta_{xA}, \beta_{yA}, w_{0A}\right]^T$ stands for the vector of nodal degrees of freedom related to the control point $\mathbf{P}_A$.

The generalized strains can be expressed by substituting Eq. (16) into Eq. (5) as

$$\hat{\boldsymbol{\varepsilon}} = \left(\mathbf{B}^L + \frac{1}{2}\mathbf{B}^{NL}\right)\mathbf{q} \tag{17}$$

where $\mathbf{B}^L$ stands for the linear infinitesimal strain

$$\mathbf{B}_A^L = \begin{bmatrix} \left(\mathbf{B}_A^m\right)^T & \left(\mathbf{B}_A^{b1}\right)^T & \left(\mathbf{B}_A^{b2}\right)^T & \left(\mathbf{B}_A^s\right)^T \end{bmatrix}^T \tag{18}$$

in which

$$\begin{aligned} &\mathbf{B}_A^m = \begin{bmatrix} R_{A,x} & 0 & 0 & 0 & 0 \\ 0 & R_{A,y} & 0 & 0 & 0 \\ R_{A,y} & R_{A,x} & 0 & 0 & 0 \end{bmatrix}; \quad \mathbf{B}_A^{b1} = -\begin{bmatrix} 0 & 0 & R_{A,xx} & 0 & 0 \\ 0 & 0 & R_{A,yy} & 0 & 0 \\ 0 & 0 & 2R_{A,xy} & 0 & 0 \end{bmatrix} \\ &\mathbf{B}_A^{b2} = \begin{bmatrix} 0 & 0 & 0 & R_{A,x} & 0 \\ 0 & 0 & 0 & 0 & R_{A,y} \\ 0 & 0 & 0 & R_{A,y} & R_{A,x} \end{bmatrix}; \quad \mathbf{B}_A^s = \begin{bmatrix} 0 & 0 & 0 & R_A & 0 \\ 0 & 0 & 0 & 0 & R_A \end{bmatrix} \end{aligned} \tag{19}$$

And $\mathbf{B}^{NL}$ denotes the nonlinear strain

$$\mathbf{B}_A^{NL}(\mathbf{q}) = \begin{bmatrix} \mathbf{A}_0 \\ \mathbf{0} \end{bmatrix} \mathbf{B}_A^g \tag{20}$$

where

$$\mathbf{B}_A^g = \begin{bmatrix} 0 & 0 & R_{A,x} & 0 & 0 \\ 0 & 0 & R_{A,y} & 0 & 0 \end{bmatrix} \tag{21}$$

From Eqs. (12) and (17), the governing equation of the problem can be described as follows:

$$(\mathbf{K}_L + \mathbf{K}_{NL})\,\mathbf{q} = \mathbf{F} \tag{22}$$

in which $\mathbf{K}_L$ and $\mathbf{K}_{NL}$ symbolize the linear and nonlinear stiffness matrices, respectively; $\mathbf{F}$ is the load vector. They can be expressed as:

$$\begin{aligned} \mathbf{K}_L &= \int_\Omega \left(\mathbf{B}^L\right)^T \hat{\mathbf{D}} \mathbf{B}^L d\Omega, \\ \mathbf{K}_{NL} &= \frac{1}{2}\int_\Omega \left(\mathbf{B}^L\right)^T \hat{\mathbf{D}} \mathbf{B}^{NL} d\Omega + \int_\Omega \left(\mathbf{B}^{NL}\right)^T \hat{\mathbf{D}} \mathbf{B}^L d\Omega + \frac{1}{2}\int_\Omega \left(\mathbf{B}^{NL}\right)^T \hat{\mathbf{D}} \mathbf{B}^{NL} d\Omega, \\ \mathbf{F} &= \int_\Omega \mathbf{R}^T f_z d\Omega \end{aligned} \tag{23}$$

2.4. Solution procedure

An iterative Newton–Raphson technique is utilized for solving the nonlinear equilibrium equation in Eq. (22). The residual force depicting the error in the approximation and tending to zero can be given as follows:

$$\boldsymbol{\varphi}(\mathbf{q}) = (\mathbf{K}_L + \mathbf{K}_{NL}(\mathbf{q}))\,\mathbf{q} - \mathbf{F}^{ext} \to 0 \tag{24}$$

If the approximate trial solution at the ith iteration ${}^{i}\mathbf{q}$ produces an unbalance residual force, a solution ${}^{i+1}\mathbf{q}$ is introduced as

$${}^{i+1}\mathbf{q} = {}^{i}\mathbf{q} + \Delta\mathbf{q} \tag{25}$$

where $\Delta\mathbf{q}$ is the increment displacement, which is defined as follows:

$$\Delta\mathbf{q} = \left[\mathbf{F} - \left(\mathbf{K}_L + {}^{i}\mathbf{K}_{NL}\left({}^{i}\mathbf{q}\right)\right){}^{i}\mathbf{q}\right] \Big/ \mathbf{K}_T \tag{26}$$

where tangent stiffness matrix $\mathbf{K}_T$ can be defined as

$$\mathbf{K}_T = \frac{\partial \boldsymbol{\varphi}\left({}^{i}\mathbf{q}\right)}{\partial {}^{i}\mathbf{q}} = \tilde{\mathbf{K}}_{NL} + \mathbf{K}_g \tag{27}$$

in which

$$\begin{aligned} \tilde{\mathbf{K}}_{NL} &= \int_\Omega \left(\mathbf{B}^L + \mathbf{B}^{NL}\right)^T \hat{\mathbf{D}} \left(\mathbf{B}^L + \mathbf{B}^{NL}\right)^T d\Omega, \\ \mathbf{K}_g &= \int_\Omega \left(\mathbf{B}_g\right)^T \begin{bmatrix} N_x & N_{xy} \\ N_{xy} & N_y \end{bmatrix} \left(\mathbf{B}_g\right) d\Omega. \end{aligned} \tag{28}$$

3. Time series forecasting

Time series forecasting involves fitting a model to historical data in a time series to predict future values. A time series is a set of time-sequenced observations. To solve such problem, time series forecasting will be re-framed as a supervised learning problem. In this problem, the input and output variables are the values from previous time steps and the values from the next time steps, respectively. As a result, values from previous time steps are used to forecast values for subsequent time steps. The regular linear and nonlinear machine learning algorithms can be accessed more easily with this re-framing.

In the supervised learning, an algorithm is used for learning the mapping function between available input (x) and output (y) variables as $y = f(x)$. The algorithm performs predictions on the training data, which includes input and output variables, iteratively and is corrected through updates. Once the algorithm has reached a satisfactory level of performance, it will be terminated. It is noticed that the order of observations is always maintained in the dataset.

After the dataset has been reorganized as a supervised learning problem, classical or machine learning methods will be used to resolve the problem. Classical methods such as Autoregression (AR), Moving Average (MA), Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA), and others have several drawbacks, including no support for missing or corrupted data, and only being effective for linear relationships. While machine learning methods such as long short-term memory, recurrent neural networks, and others can overcome these limitations of classical methods. As a result, long short-term memory for solving nonlinear static analysis of functionally graded plates has been examined in this study. LSTM will be presented in the subsequent section. For the sake of brevity, readers can refer to the detailed solving process of the supervised learning problem in some studies such as [31–33].

4. Long short-term memory

Long short-term memory (LSTM) [14] is a temporal sequence model that is based on an recurrent neural network (RNN) extension that basically extends its memory. Specifically, RNNs are a type of artificial neural network that can be used to process sequential data. This algorithm is the first to use internal memory to remember its input. As a result, RNN is well-suited to problems involving sequential data in machine learning.

The units of an LSTM are used to construct layers, which is often referred to as an LSTM network. LSTM stores their data in memory and can read, write, and delete from that memory, which is similar

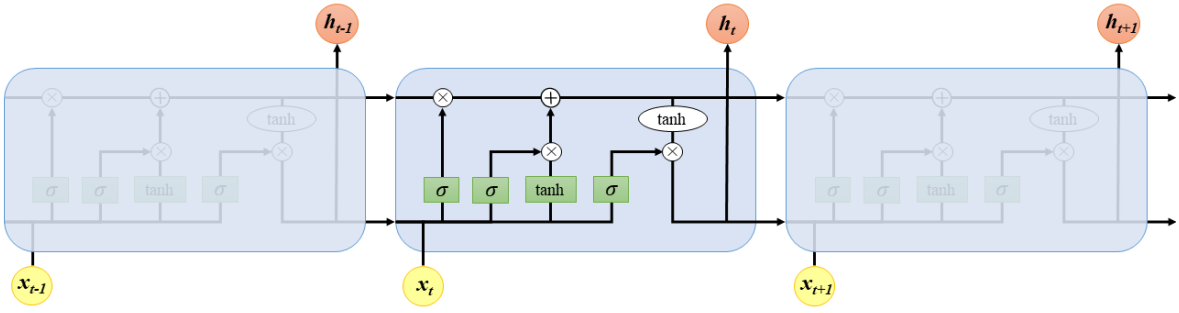


Figure 2. A LSTM's repeating module including four interacting layers

to that of a computer. As a result, the LSTM can remember information for longer periods of time. In contrast to RNN, this method distinguishes between relevant and irrelevant data. The input, forget, and output gates make up a typical LSTM unit. The input gate determines how much new data flows into the cell, the forget gate determines how much data remains in the cell, and the output gate determines how much data from the cell is used to compute the LSTM unit's output activation. LSTM architecture is often unfolded into the form of a chain of repeating modules of neural networks to make analysis easier, as shown in Fig. 2.

The LSTM diagram will be explained step by step. The first step in the LSTM shown in Fig. 3 is to determine that what information from the cell state will be discarded. A sigmoid layer known as the "forget gate layer" will make this decision. It examines hidden state h_{t-1} and input in the$t$th state of the model named x_t, and returns a number between 0 and 1 for each number in the cell state c_{t-1}. A 1 means "keep it completely" while a 0 means "get rid of it completely". The activation of the forget gate at time t is represented by f_t, which is computed as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{29}$$

where σ is the sigmoid function; W_f and b_f denote the weight and bias parameters, respectively; h_{t-1} is the hidden state vector, also known as the LSTM unit's output vector, and it is a set of values that summarizes all of the information about the system's past behavior that is required to provide a description of its future behavior; and x_t denotes the LSTM unit's input vector.

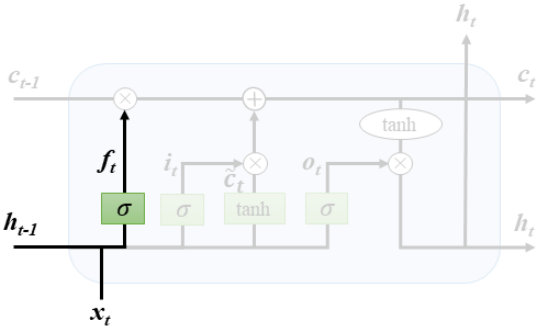


Figure 3. The first step in LSTM

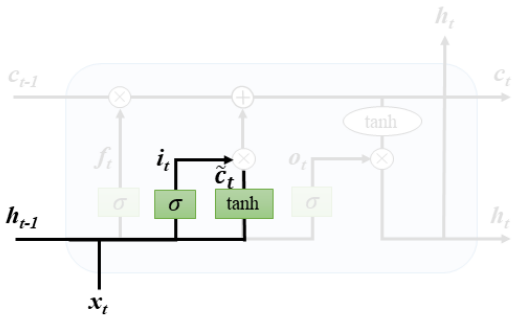


Figure 4. The second step in LSTM

Second, it will be decided whether or not to store new information in the state cell. This procedure is divided into two parts: The first step is to determine which values will be updated by a sigmoid layer known as the "input gate layer" as shown in Eq. (30); A tanh layer then creates a vector of new candidate values $\tilde{c}_t$ to add to the state, as in Eq. (31). The procedure is depicted in Fig. 4.

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{30}$$

$$\tilde{c}_t = \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \tag{31}$$

The next step is depicted in Fig. 5. The old cell state c_{t-1} has been updated into the new cell state c_t, as shown in this figure. To create an update to the state, the values obtained in the previous step have been combined, $i_t * \tilde{c}_t$. After multiplying c_{t-1} by f_t, the new state is constructed by adding $i_t * \tilde{c}_t$ as Eq. (32).

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \tag{32}$$

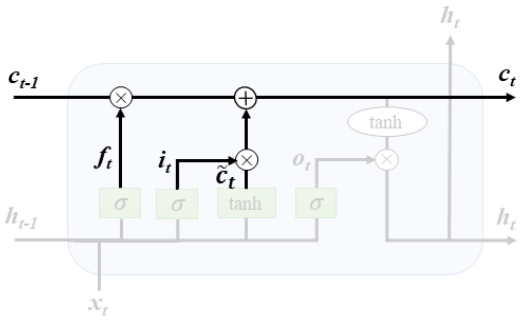


Figure 5. The third step in LSTM

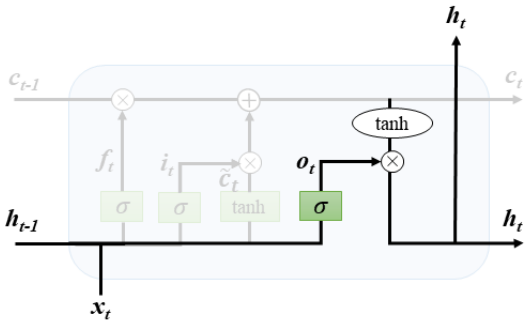


Figure 6. The last step in LSTM

The last step is to choose an output, as shown in Fig. 6. As in Eq. (33), parts of the cell state will be determined to give output by sigmoid layer. As shown in Eq. (34), the cell state is passed through the tanh layer and multiplied by the generated output.

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{33}$$

$$h_t = o_t * \tanh\left(c_t\right) \tag{34}$$

The sigmoid and tanh functions are expressed as follows:

- The sigmoid function:

$$f\left(x\right) = \frac{1}{1 + e^{-x}} \tag{35}$$

- The tanh function:

$$f\left(x\right) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{36}$$

5. Numerical examples

In this study, two boundary conditions for nonlinear analyzing of the plates have been considered as follows:

- Simply supported condition with movable edge (SSSS1):

$$\begin{cases} v_0 = w_0 = \beta_y = 0 \text{ on } x = 0, L \\ u_0 = w_0 = \beta_x = 0 \text{ on } y = 0, W \end{cases} \tag{37}$$

- Simply supported condition with immovable edge (SSSS3):

$$u_0 = v_0 = w_0 = 0 \text{ on all edges} \tag{38}$$

The non-dimensional deflection $\bar{w}$ and load parameter $\bar{P}$ are given as follows:

$$\bar{w} = \frac{w}{h}; \quad \bar{P} = \frac{f_z a^4}{E_m h^4} \tag{39}$$

Firstly, a moderate isotropic square plate with $L/h = 10$, $a = b = 10$ in, Young's modulus $E = 7.8 \times 10^6$ psi, and Poisson's ratio $\nu = 0.3$ as in the study by Reddy [29] is considered, which is subjected to a uniformly distributed load. Fig. 7 shows how the central deflection $\bar{w}$ of this plate

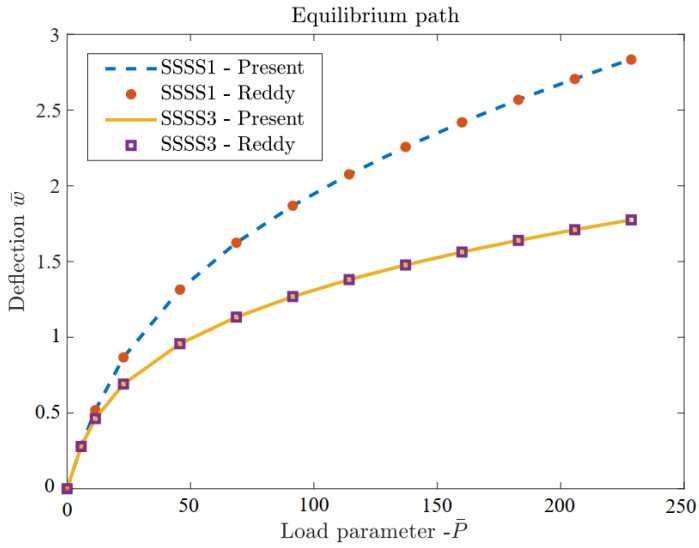


Figure 7. The load-deflection curves of the isotropic plate

varies with the load parameter $\bar{P}$ under two different boundary conditions: SSSS1 and SSSS3. From the figure, it can be seen that the current solutions are very similar to those reported by Reddy [29] using FEM.

This load-deflection curve of the isotropic square plate under the SSSS1 boundary condition has been used to validate the effectiveness of LSTM in predicting the nonlinear behavior of the plate. This curve is divided into 125 data points which are considered as data in a dataset. The dataset is divided into three sets: training, testing, and prediction. Training and testing sets are used for training process of LSTM to obtain optimum weights. Based on these weights, data in prediction set will be predicted. LSTM's architecture with one hidden layer, batch-size of 2 and 1500 epochs has been used for training and predicting the nonlinear behavior of the plate. In addition, the first example considers the impact of optimizers such as SGD, Adadelta, Adamax, RMSprop, and Adam optimizer, as well as the number of neurons in LSTM such as 10, 30, 50, 70, and 90 on method accuracy. In all subsequent examples, the most appropriate optimizer will be chosen based on the obtained results.

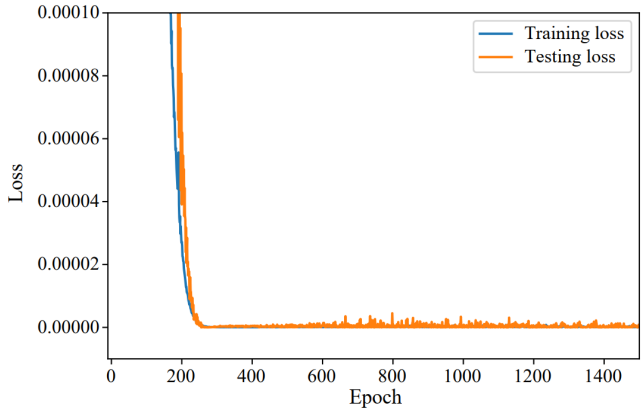


Figure 8. The convergence history of the loss function of the chosen LSTM model

The first example uses 30% of the 125 data for training and testing, and 70% for prediction. The training set accounts for 80% of the data in the training process, while the testing set accounts for 20%. Mean absolute percentage error (MAPE) for training, testing and prediction process, and computational time for all three processes are tabulated in the Table 1. Based on the investigation, the LSTM architecture with Adam optimizer and 30 neurons in one hidden layer is chosen for subsequent examples. Because this option helps to save computational cost (26.6313 seconds) while still ensuring the accuracy of the method (the accuracy achieves 99.6% for training, 99.8% for testing, and 99.5% for prediction). Fig. 8 depicts the convergence history of the chosen model's loss function with 30% of the training data.

Table 1. The effect of optimizers and the number of neurons on the effectiveness of LSTM in terms of error and computational time

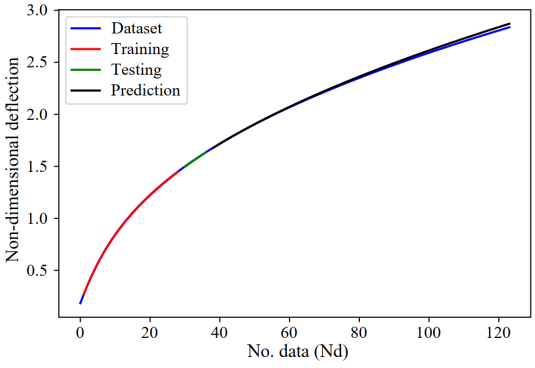
Optimizer	No. neurons	Mean absolute percentage error			Time (second)
		Training	Testing	Prediction	
SGD	10	33.1541	28.6092	43.7454	25.1310
	30	35.6226	30.8990	46.9065	25.7333
	50	35.0143	30.0525	45.8447	26.4011
	70	37.7320	32.5965	49.5027	27.0310
	90	35.3496	30.3923	46.2894	28.1742
Adadelta	10	1.1961	0.3358	0.7126	26.4610
	30	1.1584	1.0962	2.1038	26.8710
	50	1.0372	1.0508	2.0512	28.5641
	70	1.7448	1.4741	2.4434	30.9525
	90	0.3950	0.6831	1.7406	32.6538
Adamax	10	0.2383	0.2600	0.8842	25.9476
	30	0.4673	0.2298	0.5096	26.3728
	50	0.1995	0.2199	0.9091	28.0496
	70	0.2659	0.0913	0.5559	28.3083
	90	0.1606	0.1187	0.7008	30.0339
RMSprop	10	1.6880	1.0726	0.4040	25.4014
	30	1.4149	0.8884	1.1799	26.1134
	50	1.0193	0.7772	0.5514	27.0045
	70	2.7480	1.4023	1.3136	28.7788
	90	3.0675	1.1976	0.8437	30.0582
Adam	10	0.3090	0.3521	1.1557	26.1953
	30	**0.3851**	**0.1833**	**0.5057**	**26.6313**
	50	0.5561	0.3588	0.9157	27.6547
	70	0.0762	0.0425	0.4958	28.7211
	90	0.4495	0.1938	0.4714	30.9207

The chosen LSTM architecture has been used to train and predict nonlinear static behavior of the isotropic plate using 30%, 50%, and 80% of training data under SSSS1, and SSSS3 boundary conditions. The effectiveness of the method is demonstrated in Table 2 by the obtained results. From

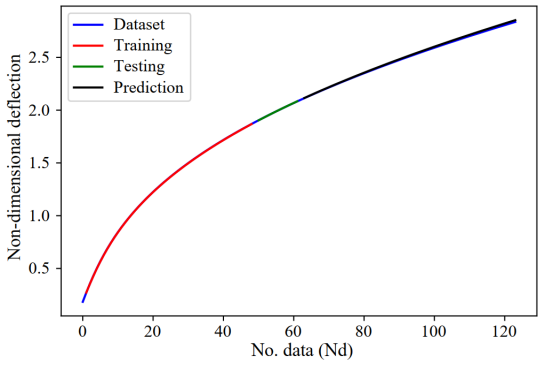
Table 2. Error and computational time obtained using the chosen LSTM architecture for the isotropic plate

Boundary condition	Training data	Mean absolute percentage error			Training and prediction time (second)	Analysis time (second)
		Training	Testing	Prediction		
SSSS1	30%	0.3851	0.1833	0.5557	26.6313	963.8210
	50%	0.1432	0.0212	0.3148	38.0769	
	80%	0.0553	0.1103	0.2307	57.9372	
SSSS3	30%	0.2667	0.1794	0.8900	26.6801	915.6035
	50%	0.2273	0.3524	0.9132	39.7864	
	80%	0.2550	0.3831	0.5819	57.1417	

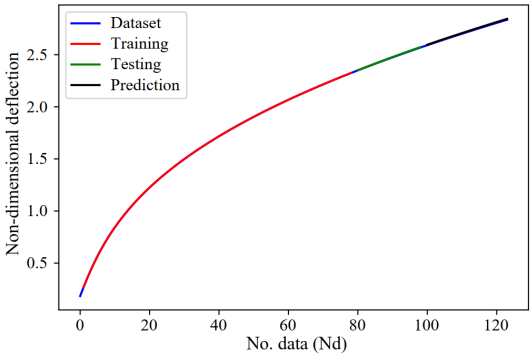
the table, it can be seen that LSTM correctly predicted more than 99% of the nonlinear behavior of the isotropic plate without using any analytical tools. Moreover, the current method saves a significant amount of computational time when compared to the traditional analytical method. The nonlinear static behaviors of the isotropic plate in these cases predicted by LSTM have been shown in Fig. 9. As shown in the figure, the LSTM successfully predicted the load-deflection curve in comparison to the IGA solution. Load parameter corresponding to the number of data (Nd) in x-axis can be calculated as $\bar{P} = \dfrac{-250 \times Nd}{125}$. In which, 250 is the maximum magnitude of the nondimensional load, 125 is the number of data points in the dataset.



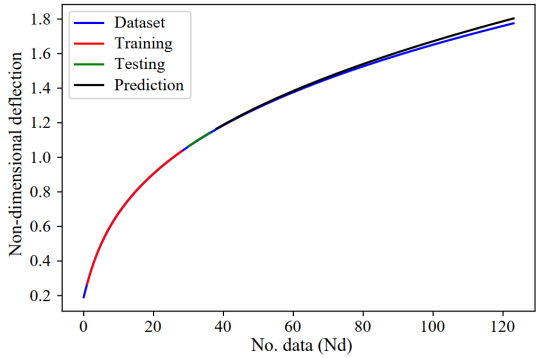
(a) SSSS1, 30% of training data



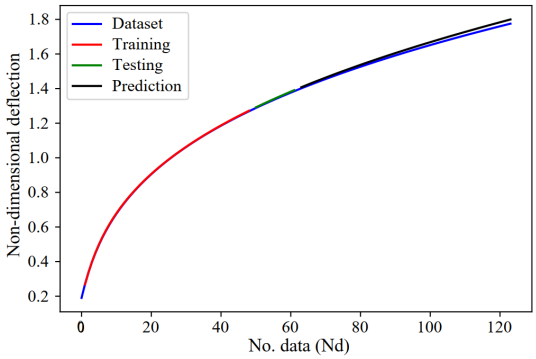
(b) SSSS1, 50% of training data



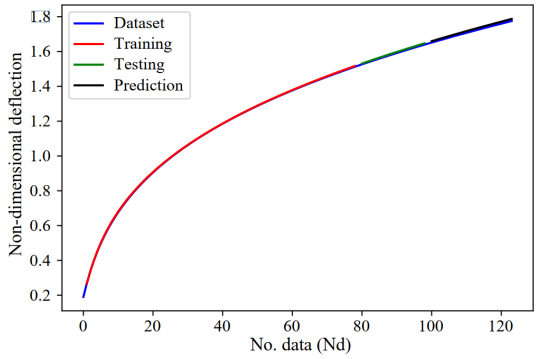
(c) SSSS1, 80% of training data



(d) SSSS3, 30% of training data



(e) SSSS3, 50% of training data



(f) SSSS3, 80% of training data

Figure 9. The load-deflection curves of the isotropic plate predicted by LSTM

The nonlinear behavior of an Al/ZrO_2 plate (E_{Al} = 70 GPa, ν_{Al} = 0.3, ρ_{Al} = 2707 kg/m^3, E_{ZrO_2} = 151 GPa, ν_{ZrO_2} = 0.3, ρ_{ZrO_2} = 3000 kg/m^3) with a length of 0.2 m and a thickness of 0.01 m under SSSS1 boundary condition is then investigated. This plate is subjected to a uniformly distributed load that is increased in 200 steps until it reaches $f_z = -10^7$ N/m^2. In this example, 200 data points have been used for investigating. The effectiveness of LSTM in predicting the nonlinear behavior of the FG plate with the variation of power index n is demonstrated in Table 3. In the table, two cases of percentage of training data are examined. The load-deflection curve of the FG plate was accurately predicted by LSTM with over 99 percent accuracy, as shown in the table. Furthermore, this method has helped to save a significant amount of computational time.

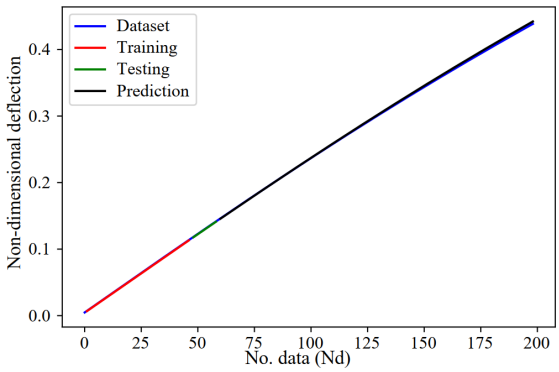
Table 3. Obtained error and computational time using the chosen LSTM architecture for the FG plate

Power index (n)	Training data	Mean absolute percentage error			Training and prediction time (second)	Analysis time (second)
		Training	Testing	Prediction		
0	50%	0.2419	0.3793	0.9130	57.0963	1739.8815
	80%	0.0491	0.0519	0.1770	89.1806	
0.5	50%	0.4648	0.1887	0.6180	59.1823	1799.2372
	80%	0.0477	0.0651	0.1924	89.8288	
1	50%	0.3370	0.2976	0.5721	59.8002	1757.3990
	80%	0.0634	0.0744	0.2021	90.3943	
2	50%	0.6563	0.3238	0.5150	57.7218	1860.9100
	80%	0.1425	0.0753	0.1979	90.4622	
5	50%	0.6354	0.2982	0.4602	58.0472	1878.2629
	80%	0.0498	0.0725	0.2008	90.3810	
10	50%	0.4957	0.2413	0.3735	58.1756	1919.5960
	80%	0.0868	0.1180	0.2451	90.1580	
inf	50%	0.2140	0.0098	0.0281	56.7398	1971.3872
	80%	0.1664	0.0766	0.1942	86.9989	

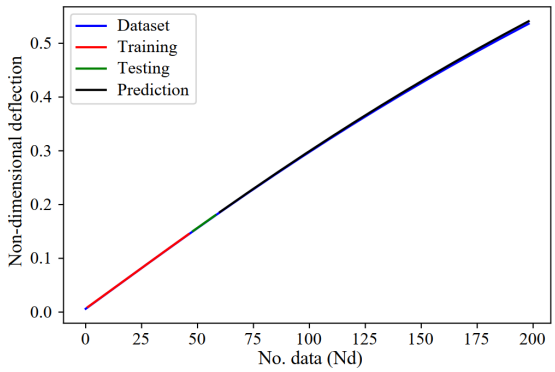
Although the chosen LSTM method is highly accurate in the above cases (over 99 percent), it is less accurate in the case of 30% of the training dataset. For this problem, the LSTM structure is enhanced with two hidden layers, each with 30 neurons, and 5000 epochs. The obtained results have been tabulated in Table 4, and one can see that the current architecture of LSTM can overcome the difficulty and the load-deflection curve is well captured with over 99.3% accuracy. Although the architecture of LSTM has been enhanced, it still saves significant computational time when compared

Table 4. Obtained error and computational time using LSTM for the FG plate with 30% of the training data

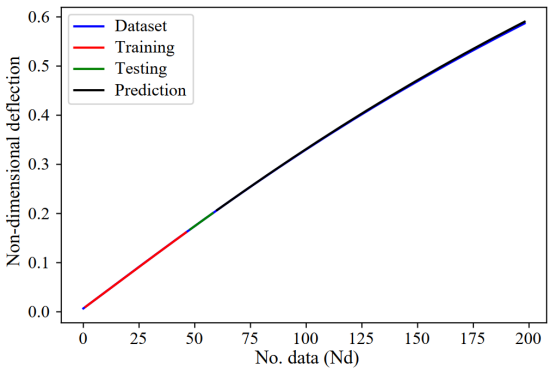
Power index (n)	Mean absolute percentage error			Training and prediction time (second)	Analysis time (second)
	Training	Testing	Prediction		
0	1.2794	0.4004	0.4000	119.0740	1739.8815
0.5	0.7587	0.2793	0.6243	118.3229	1799.2372
1	0.1616	0.0197	0.3758	117.3758	1757.3990
2	1.7952	0.3215	0.3244	120.4145	1860.9100
5	0.5020	0.2037	0.4536	119.5008	1878.2629
10	1.0916	0.2409	0.2367	119.7014	1919.5960
inf	1.0920	0.4104	0.1363	118.5780	1971.3872



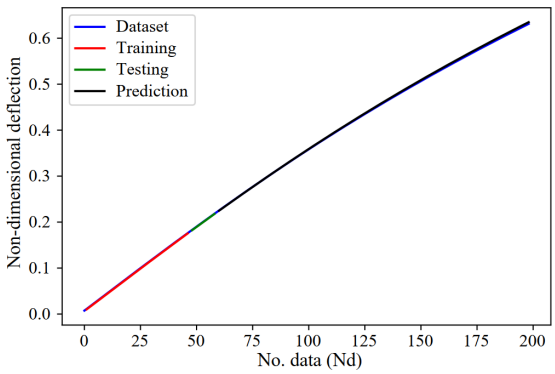
(a) $n = 0$



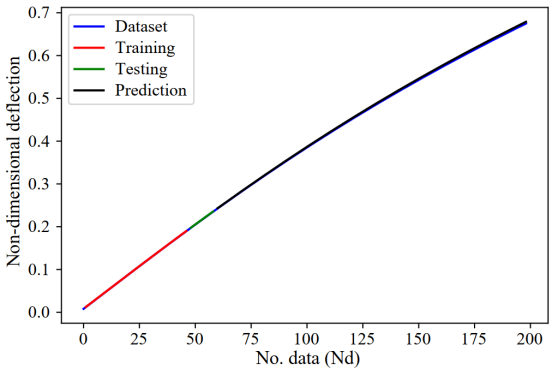
(b) $n = 0.5$



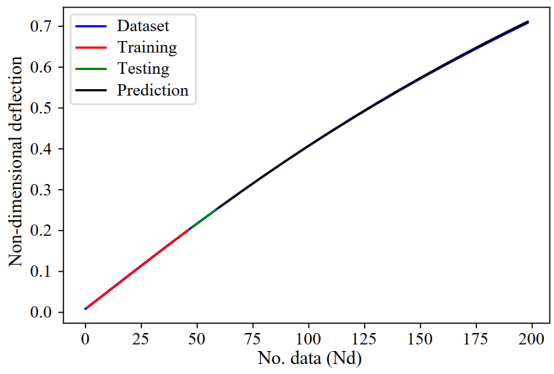
(c) $n = 1$



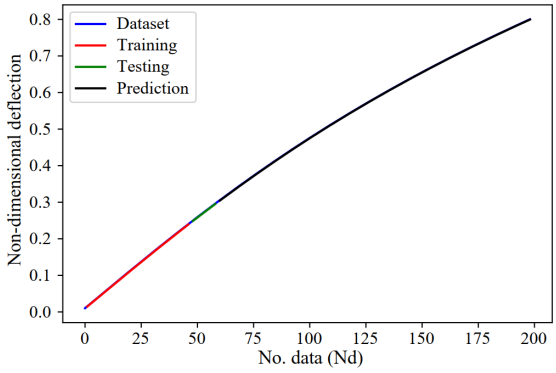
(d) $n = 2$



(e) $n = 5$



(f) $n = 10$



(g) $n = \text{inf}$

Figure 10. The load-deflection curves of the FG plate predicted by LSTM

to the IGA. Load-deflection curves of the FG plate predicted by LSTM have been shown in Fig. 10. The computed results of the current method agree well with those obtained by the IGA. Load parameter corresponding to the number of data (Nd) in x-axis can be calculated as $\bar{P} = \dfrac{-22.8571 \times Nd}{200}$. In which, 22.8571 is the maximum magnitude of the nondimensional load, 200 is the number of data points in the dataset. From the above discussions, it is clear that LSTM is a powerful tool that can be used to replace analytical tools. The LSTM also is potential when it comes to solving complex problems by enhancing the LSTM architecture.

6. Conclusions

We have presented an effective approach based on the LSTM for predicting the nonlinear static behavior of isotropic, and functionally graded plates with high accuracy and less effort. Therefore, the LSTM can be used in place of traditional analytical tools to save computational cost. With 50 percent and 80 percent of the data for training process, an LSTM architecture with a hidden layer of 30 neurons and 1500 epochs can accurately predict the nonlinear behavior of the plate. For more complex problems with only 30% of the training data, the LSTM architecture enhanced with two hidden layers, each layer including 30 neurons, and 5000 epochs successfully predicted the nonlinear behavior of the FG plate. Expanding the proposed method to more complex problems is promising.

Acknowledgements

We would like to thank Duy Tan University for their support.

References

[1] Wei, Z., Liu, S., Liu, D., Wu, J., Xia, H., Wang, B., Shi, Z. (2022). Fabrication and properties of symmetrical W/Si3N4/W functionally graded materials by spark plasma sintering. *Journal of Alloys and Compounds*, 896:163077.

[2] Son, T., Huu-Tai, T. (2019). Free-vibration analysis of multi-directional functionally graded plates based on 3D isogeometric analysis. *Journal of Science and Technology in Civil Engineering (STCE) - HUCE*, 13(2):1–11.

[3] Van, V. T., Hieu, N. V. (2022). Buckling analysis of functionally graded sandwich plates resting on Pasternak foundation using a novel refined quasi-3D third-order shear deformation theory. *Journal of Science and Technology in Civil Engineering (STCE) - HUCE*, 16(1):68–79.

[4] Guo, W., Jiang, Z., Zhang, C., Zhao, L., Jiang, Z., Li, X., Chen, G. (2021). Fabrication process of smooth functionally graded materials through a real-time inline control of the component ratio. *Journal of the European Ceramic Society*, 41(16):256–265.

[5] Hughes, T. J. R., Cottrell, J. A., Bazilevs, Y. (2005). Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering*, 194(39-41):4135–4195.

[6] Kim, M.-G., Lee, G.-H., Lee, H., Koo, B. (2022). Isogeometric analysis for geometrically exact shell elements using Bézier extraction of NURBS with assumed natural strain method. *Thin-Walled Structures*, 172:108846.

[7] Thai, C. H., Kulasegaram, S., Tran, L. V., Nguyen-Xuan, H. (2014). Generalized shear deformation theory for functionally graded isotropic and sandwich plates based on isogeometric approach. *Computers & Structures*, 141:94–112.

[8] Lieu, Q. X., Lee, J. (2019). An isogeometric multimesh design approach for size and shape optimization of multidirectional functionally graded plates. *Computer Methods in Applied Mechanics and Engineering*, 343:407–437.

[9] Nguyen, T. N., Ngo, T. D., Nguyen-Xuan, H. (2017). A novel three-variable shear deformation plate formulation: Theory and Isogeometric implementation. *Computer Methods in Applied Mechanics and Engineering*, 326:376–401.

[10] Thai, S., Thai, H.-T., Vo, T. P., Nguyen-Xuan, H. (2017). Nonlinear static and transient isogeometric analysis of functionally graded microplates based on the modified strain gradient theory. *Engineering Structures*, 153:598–612.

[11] Crisfield, M. A. (1979). A faster modified newton-raphson iteration. *Computer Methods in Applied Mechanics and Engineering*, 20(3):267–278.

[12] Crisfield, M. A. (1981). A fast incremental/iterative solution procedure that handles "snap-through". *Computers & Structures*, 13(1-3):55–62.

[13] Riks, E. (1972). The application of Newton's method to the problem of elastic stability. *Journal of Applied Mechanics*, 39(4):1060–1065.

[14] Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.

[15] Huynh, A. T., Nguyen, Q. D., Xuan, Q. L., Magee, B., Chung, T., Tran, K. T., Nguyen, K. T. (2020). A machine learning-assisted numerical predictor for compressive strength of geopolymer concrete based on experimental data and sensitivity analysis. *Applied Sciences*, 10(21):7726.

[16] Mai, H. T., Lieu, Q. X., Kang, J., Lee, J. (2022). A robust unsupervised neural network framework for geometrically nonlinear analysis of inelastic truss structures. *Applied Mathematical Modelling*, 107: 332–352.

[17] Lee, S., Park, S., Kim, T., Lieu, Q. X., Lee, J. (2021). Damage quantification in truss structures by limited sensor-based surrogate model. *Applied Acoustics*, 172:107547.

[18] Lieu, Q. X., Nguyen, K. T., Dang, K. D., Lee, S., Kang, J., Lee, J. (2022). An adaptive surrogate model to structural reliability analysis using deep neural network. *Expert Systems with Applications*, 189:116104.
[19] Tsuneki, M. (2022). Deep learning models in medical image analysis. *Journal of Oral Biosciences*.
[20] Chien, J.-T., Misbullah, A. (2016). Deep long short-term memory networks for speech recognition. In *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, IEEE.
[21] Nugaliyadde, A., Wong, K. W., Sohel, F., Xie, H. (2019). Language modeling through long term memory network. *arXiv preprint arXiv:1904.08936*.
[22] Bollepalli, B., Juvela, L., Airaksinen, M., Valentini-Botinhao, C., Alku, P. (2019). Normal-to-Lombard adaptation of speech synthesis using long short-term memory recurrent neural networks. *Speech Communication*, 110:64–75.
[23] Chen, T., Yin, H., Yuan, X., Gu, Y., Ren, F., Sun, X. (2021). Emotion recognition based on fusion of long short-term memory networks and SVMs. *Digital Signal Processing*, 117:103153.
[24] Chherawala, Y., Roy, P. P., Cheriet, M. (2017). Combination of context-dependent bidirectional long short-term memory classifiers for robust offline handwriting recognition. *Pattern Recognition Letters*, 90:58–64.
[25] Do, D. T. T., Lee, J., Nguyen-Xuan, H. (2019). Fast evaluation of crack growth path using time series forecasting. *Engineering Fracture Mechanics*, 218:106567.
[26] Zeiler, M. D. (2012). Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.
[27] Kingma, D. P., Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
[28] Tieleman, T., Hinton, G. (2012). Lecture 6.5-rmsprop, coursera: Neural networks for machine learning. Technical report, University of Toronto.
[29] Reddy, J. N. (2003). *Mechanics of Laminated Composite Plates and Shells*. CRC Press.
[30] Nguyen, V. P., Anitescu, C., Bordas, S. P. A., Rabczuk, T. (2015). Isogeometric analysis: An overview and computer implementation aspects. *Mathematics and Computers in Simulation*, 117:89–116.
[31] Do, D. T. T., Nguyen, T.-T., Nguyen, Q.-H., Bui, T. Q. (2021). Analysis of non-uniform hexagonal cross-sections for thin-walled functionally graded beams using artificial neural networks. *Journal of Science and Technology in Civil Engineering (STCE) - HUCE*, 15(3):1–12.
[32] Do, D. T. T., Lee, D., Lee, J. (2019). Material optimization of functionally graded plates using deep neural network and modified symbiotic organisms search for eigenvalue problems. *Composites Part B: Engineering*, 159:300–326.
[33] Do, D. T. T., Nguyen-Xuan, H., Lee, J. (2020). Material optimization of tri-directional functionally graded plates by using deep neural network and isogeometric multimesh design approach. *Applied Mathematical Modelling*, 87:501–533.