Identification of the stick and slip motion between contact surfaces using artificial neural networks

Jakub Górski · Andrzej Klepka · Kajetan Dziedziech · Jakub Mrówka · Rafał Radecki · Ziemowit Dworakowski

Received: 2 February 2019 / Accepted: 3 February 2020 / Published online: 15 February 2020
© The Author(s) 2020

Abstract The paper presents work related to nonlinear system parameters identification. The research is focused on systems with hysteretic stiffness characteristics. The identification procedure is developed with use of artificial neural networks. The presented method assumes two separate clusters of neural networks, which are supported by additional signal processing block. Such approach gives an advantage over the conventional identification methods due to its small restrictions. The validation process considers structural responses in time and frequency domains as well as the restoring force plane of the dynamic structure. First, verification of the identification method is performed on the numerical simulation of the system with hysteretic stiffness. Next, the identification of the real dynamic system with contact-related nonlinearity is carried out. The steel samples with contacting surfaces were used in the experiment. Electromagnetic shaker was used to excite the structure and enforce a relative shear motion between surfaces in contact. The system response was recorded using the Polytec laser vibrometer.

Keywords Identification · Neural networks · Multilayer perceptron · Ensemble · Restoring force · Stiffness hysteresis

1 Introduction

The contact phenomenon is common to many problems encountered in practical engineering applications. In case of mechanical systems, it occurs everywhere, where two surfaces move with respect to each other in either tangent, normal or both directions. The contact mechanics problems can be applied to all phenomena associated with interaction of solid bodies. It includes plasticity, elasticity and contact stress field problems, effects of surface roughness and friction phenomenon. The fundamental law related to contact stress is the Hertz law [1]. It assumes that if two elastic bodies, whose surface profile is described by the second-order mathematical surface, are pressed into contact, they can slide over each other with the presence of friction. The normal pressure of contact is assumed to be constant. Based on that, the theory linking size, shape of contact and distribution of contact pressure were developed. This theory is applicable only in the case of normal load conditions. However, it does not include the case, when the shear force is too weak to cause a slip effect. In the early 1920s, the problem of contact path phenomenon with different motion regions was investigated [2]. It was explained how to recognize the stick and slip regions, where the displacement

J. Górski · A. Klepka (✉) · K. Dziedziech · J. Mrówka · R. Radecki · Z. Dworakowski
Department of Robotics and Mechatronics, AGH
University of Science and Technology, Al. Mickiewicza
30, 30-059 Kraków, Poland
e-mail: klepka@agh.edu.pl

Fig. 1 Breathing action for contact-type damage (a), Bilinear stiffness for tension and compression phase (b)

between contact interfaces was very small and associated with local elasticity. In 1989, Mindlin [3] published his work, where problem of partial slip of surfaces in contact was solved. It was observed that for small movements between bodies in contact, a dissipative hysteresis behavior can be obtained at the local scale.

The above theories were developed over the years, bringing many interesting and scientifically important conclusions. They were also used to study phenomena occurring in metallic structures and composite materials [4–7]. These types of failures are typical for many engineering structures and exhibit behavior subjected to contact mechanics. Three main forms of the movement between the contact surfaces can be distinguished. The first type of the movement occurs when the tension/compression load, normal to the contact surfaces, is applied to the structure. As a result of the harmonic excitation, a periodic change in stiffness can be observed. Due to the tensional load applied to the structure with damage, an opening of the gap between the damage surfaces will be exhibited. Such a phenomenon can be described as a local stiffness reduction. In the next step, due to the harmonic excitation, a compressional load will be applied to the structure. As a result, the gap between the damage surfaces will close. For the ideal case, the closing phenomenon can be described by setting the local stiffness value as for the undamaged structure. This event is known as “breathing” or “clapping” phenomenon, and it can be represented by the

bilinear relation between the force and displacement [8]. Figure 1 presents bilinear stiffness for breathing action of damage interfaces.

The second type of the movement of the damage interfaces considered in this section is the relative sliding motion. Such a phenomenon can be characterized with the following two force–displacement relations. If the amplitude of the harmonic excitation is too small to break the static friction force, the micro-slip motion will be exhibited between the contact surfaces [8]. To characterize such a motion, a linear, piecewise function can be assumed as a relation between the force and displacement. However, in the case, where the amplitude of the harmonic excitation is large enough to overcome the static friction force, the relative motion between the contact surfaces will vary between two modes: stick and slip [9]. This type of the motion is represented by the hysteretic relation between the force and displacement. Figure 2 presents sliding action of contact interfaces and characteristics for micro-slip and stick and slip modes.

The third type of the movement of the damage interfaces is the relative sliding motion parallel to the leading edge of the crack. Although the movement is in a different plane, the stiffness characteristics can be described similarly as for sliding action shown in Fig. 2.

The presented characteristics generate different effects in the signals’ spectrum. In the first case, a bilinear stiffness characteristic, resulting from a different strain field for tension and another for compression

Fig. 2 Sliding action of damage interface (a), stiffness characteristic for micro-slip mode (b), stiffness characteristic for stick and slip mode (c)

sion, is provided. This effect was called “clapping non-linearity.” For this reason, the stiffness of the system is pulse modulated. Based on the modulation theory,

it is possible to determine the amplitudes of subsequent harmonics (even and odd in this case). Due to the pulse-type stiffness variation, the harmonic amplitudes are always modulated by the *sinc* envelope function. For the second type of motion (tangential), the stiffness changes between some value (for stick phase) and zero (for slip phase) twice over the input strain. It causes symmetrical variation in tangential interface stiffness. This results in only odd harmonics in the response spectrum [8]. The presented above models are the ideal cases. For real-life structures, the output signal can be modified due to additional phenomena like cyclic hardening and/or creeping, boundary conditions, environmental changes and many others. Additionally, the different nonlinearity types can occur simultaneously or can change for different amplitudes of excitation. All these factors make the stiffness characteristics or nonlinearity-type identification process

Fig. 3 Frequency domain response for systems with different stiffness type: Bounc-Wen model (a) and corresponding response spectrum (c), model described in Sect. 4 (b) and corresponding response spectrum (d)

extremely difficult because in many cases, different stiffness characteristics produce very similar effects in frequency domain. An example can be a comparison of the piecewise linear and hysteretic characteristics. For both cases, only odd harmonics are predicted in the response signal. Of course, for “ideal system” the distribution of subsequent harmonics amplitudes should be different for both characteristics. In the case of real systems where many types of nonlinearity can occur or disturbances caused by other contact-related phenomena (e.g., change of temperature, excitation of structural frequencies due to contact) the approach based on frequency domain signal representation is impossible. The same problem concerns hysteresis in general—different types of hysteresis (e.g., Bounc–Wen model, Duhem model) are practically indistinguishable when harmonics-based approach is used, although their shape is completely different. It is shown in Fig. 3, where two systems with different types of hysteresis stiffness are presented: a hysteresis based on the Bounc–Wen model and the model described in detail in Sect. 4 of this article. In both cases, the odd harmonics are dominant when the spectra of the response are observed.

Only the amplitude of subsequent harmonics is slightly different in the case of these two models. Unfortunately, in the case of real structures and measured data from real structure, this cannot be considered as reliable information for determining the type of hysteresis. For these reasons, identification of hysteretic systems is generally a non-trivial task. Additional information about nonlinear effects related to contact-type damage can be found in [10–12].

This article presents a novel method of stiffness characteristic identification. The procedure was used to determine the hysteresis stiffness characteristics of the system with stick and slip motion. The case with the relative motion in the tangent direction of the contact surface is analyzed. The method is based on Artificial Neural Networks (ANNs) and model of mechanical system with hysteretic stiffness change. The structure of ANN is tailored empirically for both numerical simulation and experimental data. The reliability of the identification procedure is increased by using an ensemble approach. The proposed method has not been used for identification of hysteretic parameters before.

In the following section, the developed model based on Artificial Neural Networks is presented. Section 4 shows validation of the identification method on data

collected from numerical model. In Sect. 5 model identification is presented for two steel specimens, which model contact phenomenon related to surfaces in contact. The work is concluded in the final section.

2 Artificial neural network-based model for identification

ANN is a parallel computing system inspired by the biological neural networks [13]. The most popular ANN structure, namely the Multilayered Perceptron (MLP) contains at least three layers: input, output and at least one hidden layer. Each of the layers is composed of neurons, which are basic computation units in the network. One hidden layer is enough to perform any nonlinear mapping, i.e., it allows to solve any classification or regression problem provided that the number of neurons in the layer is high enough [14]. However, additional hidden layers allow to arrange network input signals into more advanced features (with higher level of abstraction), rendering the training process to be easier and capable of solving more complex tasks. The number of hidden layers and number of neurons in each layer can be defined depending on the user’s requirements, while observing overfitting problem: Quantity of adjustable parameters (weights) of the network should be much smaller than the number of data points used to train the network. A popular “rule of thumb” is to use ten data samples for each weight in the MLP [15]. In this article, MLPs are applied for function approximation. The approximation capacity depends on the number of hidden layers and number of neurons. The schematic of a typical MLP is illustrated in Fig. 4.

The equation for k th neuron can be written as:

$$y_k = f \left(\sum_{i=1}^n w_{i,j} x_i + b_j \right) \quad (1)$$

where x_i is received input by the neuron, b_j is bias factor, $w_{i,j}$ is corresponding to input x_i weight factor, n is number of inputs to a neuron and $f(*)$ is an activation function. Typical examples of activation functions can be found in [16].

Before the ANN can approximate any nonlinear function, it needs to be trained. The goal of this process is to minimize the value of the loss function, which depends on weight factors. In neural networks, this problem does not have analytical solution. Therefore,

Fig. 4 Feedforward artificial neural network' schematic

training algorithms based on numerical methods, like gradient descent or Newton's method, need to be used. One of the fastest convergence training procedure for small networks is the Levenberg–Marquardt algorithm [17], which is used in the presented identification system.

The training procedure of the ANN is usually a non-deterministic process, i.e., it starts from a random point in ANN parameter space and with an aim to minimize the error, it finishes in a local minimum of an objective function. For that reason, two networks of the same structure trained on the same data may differ in efficiency. On the other hand, it is often difficult to assess the efficiency of the trained ANN before it is employed in practice. One of the many methods to solve this issue require training of several ANNs, then choice of few that are best in validation data assessment and finally, in operational phase, average these responses to obtain an ensemble output. This ensemble approach usually boosts reliability and accuracy of the ANN-based diagnose [18] and was successfully used in various mechanical systems research scenarios [19–21].

Artificial neural networks are often used in the field of nonlinear systems modeling. The most common approach to identify the dynamical properties of systems is restoring force plane (RFP) analysis. Masari et al. [22] developed an ANN to identify restoring force of single degree-of-freedom (SDOF) Duffing oscillator. For training purpose, authors used deterministic signals and for validation, stochastic signals. Based on their work, Liang et al. [23] extended the application to MDOF systems. Furthermore, Liang et al. [24] improved the convergence of model presented in [23] was improved by connecting backpropagation algorithm with fuzzy rules. Pei et al. [25,26] presented novel approach presented for designing ANNs to approximate nonlinear hysteretic restoring force without memory. They explained physical meaning of trained networks' parameters for system identification and damage detection purpose. The approach presented in [27,28] is a continuation of Pei's studies, where they presented a developed initialization procedure and neurons' number selection to approximate nonlinear functions in context of restoring force planes identification.

All these papers cover restoring force identification, which is described by nonlinear function. However, there is a wide group of nonlinearities in structural dynamical system, which are modeled by hysteretic relation with memory [29]. In the review [29], authors presented five classical models describing hysteretic relation in piezoelectric materials and shape memory alloys, focusing on possible applications, while underlining the limitations. As a result, many researchers tried to design even better models. Throughout many years, researchers managed to develop [30,31,33,34] methods based on artificial neural networks, to identify systems with hysteresis. In Kosmatopoulos et al. [30], a

Fig. 5 Structure of the identification model

Fig. 6 Structure of the validation model

Fig. 7 Structure of model for ensemble approach

robust algorithm for restoring force identification, with the use of the Volterra/Wiener neural network, was presented. Kosmatopoulos et al. applied the algorithm to MDOF system with memory-based hysteresis. Pei et al. [31] continued to study on this algorithm to improve its computational efficiency and to select more appropriate ANN structure. A different approach was developed in Park et al. [32], i.e., a combined ANN with the Preisach hysteresis model. Such a solution allowed to identify the system with memory-based hysteresis. Another approach was presented in Nakshatharan et al. [34], where authors proposed hysteresis system identification based on MLP. In work, authors presented networks with two inputs and one output. One of their

inputs was a tag, informing about the trend of the second input signal. Farrokh and Dizaji [33] continued work with MLPs. In this work, authors took into account Madelung's description of hysteresis relation, to develop identification algorithm. The results reveal flexibility of the proposed method for different fields. Based on their work, the identification structure presented in this article was developed, which contained artificial neural networks and signal processing block. The proposed model was applied to identify the hysteretic relationship occurring at the contact of two sliding surfaces. This problem is not well studied, which is simultaneously the motivation to apply the following model for such a purpose.

3 Identification model

In this paper, the ANNs were applied to identify SDOF nonlinear dynamical system with hysteresis. The equation of motion for such system is written as:

$$m\ddot{x} + R(\dot{x}, x, t) = F(t) \quad (2)$$

where $\ddot{x}$, $\dot{x}$ and x are, respectively, the acceleration, velocity and displacement responses of the nonlinear system. The F is the excitation force, and R is the nonlinear hysteretic restoring force to be identified. Rewriting Eq. (2) obtains:

$$R(\dot{x}, x, t) = F(t) - m\ddot{x} \quad (3)$$

The identification procedure requires four signals: displacement, velocity, acceleration and force. Direct identification of restoring force from Eq. (3) by MLP does not give good results, because it is unable to describe hysteresis with memory, i.e., for particular set of inputs, two different outputs might be required. Based on work Farrokh and Dizaji [33] and Pei et al. [35], a hysteretic relation as a three argument function was developed:

$$R_h = g(x, x_0, R_{h0}) \quad (4)$$

where x is displacement, R_h is hysteresis restoring force value and $[x_0, R_{h0}]$ are coordinates of the point on the displacement–restoring force surface, where the first derivative of the input signal changes sign. In this approach, however, a non-hysteretic part of identified system is not taken into account. For further improvement in identification performance, the model described following equation was developed:

$$R(\dot{x}, x, t) = f(\dot{x}, x, R_h) \quad (5)$$

Considering Eq. (5), the identification model has the structure as presented in Fig. 5.

The presented model consists of two neural networks and a signal preprocessing block. The inputs to the model are displacement $x[k]$ and velocity $\dot{x}[k]$, and the output of the model is restoring force $R[k]$. The purpose of hysteresis network is to approximate the hysteresis part of the model [see Eq. (4)].

The restoring network is trained to combine hysteretic and non-hysteretic parts of modeled system using as inputs, signals of displacement $x[k]$, velocity $\dot{x}[k]$ and hysteresis network's output $R_h[k]$ [see Eq. (5)].

In other words, the task is divided into two components: A general approximation of the restoring force by the hysteresis network that does not take into account the measured velocity and a fine-tuning of that approximation by the restoring network based on added velocity measurement. We therefore ask the first network a question of “where in the (x, R) space are we in general given the current $x[t]$ and starting point of the hysteresis curve?” and ask the second network a question of “what exactly is the value of R given the general estimate of R (namely R_h), displacement $x[t]$ and velocity $\dot{x}[t]$?”.

The purpose of the signal preprocessing block is to provide to the hysteresis network input displacement signal $x[k]$, the extracted displacement $x_0[k]$ and network response $R_{h0}[k]$ values when the value of the velocity $\dot{x}[k]$ changes sign. The discrete time equations for this block can be written as:

$$x_0[k] = \begin{cases} x[k] & \text{when } \dot{x}[k-1] \dot{x}[k] \leq 0 \\ x_0[k-1] & \text{otherwise} \end{cases} \quad (6)$$

$$R_{h0}[k] = \begin{cases} R_h[k-1] & \text{when } \dot{x}[k-1] \dot{x}[k] \leq 0 \\ R_{h0}[k-1] & \text{otherwise} \end{cases} \quad (7)$$

The identification procedure is composed of three major steps. First, the mass value of the system is estimated by weight measurement of the moving element. Next, the restoring force is calculated from Eq. (3). From the obtained sampled restoring force— R , displacement— X and velocity— $\dot{X}$, signal preprocessing block calculates vectors of coordinates for hysteresis network— X_0 and R_0 . From obtained discrete signals described above, the vectors of training data for hysteresis network U_1 and T_1 can be written as:

$$U_1 = [X, X_0, R_0]^T, T_1 = R \quad (8)$$

There is no possibility of direct measurement of the R_h value (the component that contributes to restoring force and are caused purely by the presence of hysteresis); thus, in training of the hysteresis network the R values are used instead. It is for that reason expected that the values returned from hysteresis network will contain an error caused by the lack of knowledge of

Table 1 Numerical model parameter values

m (kg)	B (Ns/m)	K_1 (N)	K_s (N)
1	1	1000	20,000

Fig. 8 Blok diagram of SimulinkTM model

Fig. 9 Part of training data divided into: train data (black circle), validation data (blue diamond) and test data (magenta square) for **a** hysteresis network, **b** restoring network. (Color figure online)

Table 2 Training and test data division for simulated data

Experiment	Data purpose	Data division	Data samples
Experiment 1	Training data	70/15/15	50,000
Experiment 2	Independent test data	No division	20,000

measured velocity. The hysteresis network has MLP structure and is trained using a Levenberg–Marquardt algorithm [17]. As result of this process, the restoring force vector R_h estimated by hysteresis network is obtained. Combining it with displacement vector X and velocity vector $\dot{X}$ yields an input vector U_2 for the restoring network. The target vector T_2 for this network is again restoring force signal R . Both vectors can be written as:

$$U_2 = [X, \dot{X}, R_h]^T, T_2 = R \quad (9)$$

The restoring network is trained by the same algorithm as hysteresis network.

After the training process, prepared networks are inserted into the model presented in Fig. 5. Next the identification model inserted into block diagram presented in Fig. 6 for final validation on test dataset. During this process, the model is stimulated with the same excitation as the identified structure. Then, the model response is compared to registered restoring force signals by calculating the mean square error (MSE).

Fig. 10 Comparison of hysteretic restoring force projection on force–displacement plane for simulation model (solid black line) and neural network model (dashed magenta line)

Since ANN training is a non-deterministic process, ANNs trained on the same data may exhibit different performances in test. On the other hand, averaging answers of many classifiers tend to increase accuracy provided that the classifiers are diverse and accurate [36,37]. Such approach, known under *ensemble* name, is used in this paper as well. A simple *overproduce and choose* approach is used: A fraction of ANNs that scored best during training are stored as ensemble members [21,38]. In Fig. 7, the block diagram of ensemble is presented.

4 Simulation tests

In order to check the proposed identification structure, the dynamic model with hysteresis is prepared. The motion equation of simulated system can be written as:

$$m\ddot{x}(t) + B\dot{x}(t) + K_1x(t) + K_h(t) = F(t) \quad (10)$$

where K is linear stiffness coefficient, B is the damping factor, m is the mass, $F(t)$ is force excitation and

Fig. 11 Comparison between time responses of simulation model (solid black line) and neural network model (dashed magenta line) for frequency of: **a** 20 Hz, **b** 30 Hz, **c** 40 Hz, **d** 50 Hz

Fig. 12 Comparison of power spectrum for simulation model (solid black line) and neural network model (dashed magenta line) for frequency of: **a** 20 Hz, **b** 30 Hz, **c** 40 Hz, **d** 50 Hz

$K_h(t)$ is hysteresis relation, which is based on mathematical model described in Pecorari and Solodov [8]. Through a change of the stress–strain relation to a force–displacement one, the following system of equations is obtained:

$$K(t) = 0.5 \cdot K_s \{1 - \text{sign}(\dot{x}(t)) \cdot \text{sign}[x(t) + x_1 \text{sign}(\dot{x}(t))]\} \quad (11)$$

$$K_h(t) = K(t)x(t) + K(t)x_1 \text{sign}(\dot{x}) + 0.5K_s(x_0 - x_1)\text{sign}(\dot{x}) \quad (12)$$

where K_s is hysteretic stiffness. The $K(t)$ is two state variables, equal to zero when hysteresis reaches output limit and equal to K_s when the limit is not reached. The x_0 , x_1 can be described as displacement values when the hysteresis restoring force reaches its constant limit.

The each factor in Eq. (12) has the following meaning. The $K(t)x(t)$ factor is responsible for steepness of hysteresis restoring force edge. The $K(t)x_1 \text{sign}(\dot{x})$ prevents from discontinuities, when velocity value changes sign. The last factor $0.5K_s(x_0 - x_1)\text{sign}(\dot{x})$ determines the hysteresis restoring force constant

value. The parameters values of the introduced model are presented in Table 1.

The mathematical model using Eqs. (10–12) was implemented in MATLAB Simulink™ environment and solved using the fourth-order Runge–Kutta method with a fixed step length of $\Delta t = 0.1$ ms. The block diagram of the test model is presented in Fig. 8.

To train MLPs, one dataset was collected. For that purpose, force excitation [in Eq. (10)] was introduced as a chirp signal with linearly increasing frequency in the range of 0–100 Hz with added white noise of signal-to-noise ratio (SNR) equal to 34 dB. Six tests with different amplitude levels (15, 30, 45, 60, 75 and 105 N) were carried out. Before training process, the value of mass was arbitrarily assumed to be equal to 1 kg. The number of neurons and hidden layers was selected through a trial and error method, by gradually increasing the number of neurons based on validation dataset performance. To simplify the network structure optimization step, it was arbitrary decided that both networks have the same structure. When it was observed that further increase in the number of neurons did not

Fig. 13 Experimental test rig: **a** schematic view **b** close-up photo **(b)** and the whole experimental setup **(c)**

improve the accuracy, the procedure was stopped. As a result, the proposed networks contain two hidden layers with five neurons in each and hyperbolic tangent activation function. For the training purpose, 50,000 collected data samples were randomly divided into training, validation and testing sets in proportion 70/15/15. A part of training data, and its division on three subset, is presented in Fig. 9 and Table 2.

For such subdivided dataset, each neural network in model was trained by Levenberg–Marquardt algorithm [17]. The initial weights values were chosen randomly. Next, the validation of the trained networks was per-

formed. The second dataset was generated from simulation model for four force harmonic signals. The frequencies were 20, 30, 40 and 50 Hz with a constant amplitude of 80 and 90 N. Although the frequencies were included in the chirp signal, the signal amplitudes were different and no noise was added to the test data. For that reason, there was no risk of a hidden overfit phenomenon, in which training and testing data are too similar to properly assess the generalization capabilities of the ANN. The results of validation are illustrated in Figs. 10, 11 and 12. The identified model exhibits good fit for hysteresis loops and for time responses.

Table 3 Training and test data division for measured data

Experiment	Data purpose	Data division	Data samples
Experiment 1	Training data	70/15/15	51,200
Experiment 2	Independent test data	No division	102,396

Fig. 14 Block diagram of signal processing procedure for measured signal

The frequency components are reproduced almost perfectly. The test presented in this section justifies the decision of using the ANN-based model proposed by the authors for identification of a SDOF system with hysteresis.

5 Identification of the motion between contact surfaces

5.1 Description of the experimental setup

For further validation of identification model, the test is conducted on the experimental data. Figure 13 presents scheme of the experimental system considered in this study. It included two C45 steel specimens. One sample was firmly attached to the steel base plate. The second was mounted at the end of threaded rod. The upper sample was pressed to other by loading mass of 0.2 kg weight, in order to induce a firm contact between two surfaces in normal direction. Furthermore, hysteretic measurements were taken by impedance head PCB 288D01 attached to the specimen, which measured force and acceleration responses. The two heads of vibrometer system PSV Polytec - 400 SLDV measured the velocity and displacement responses. Due to the fact that measurement points were very close, the structure could be modeled as SDOF vibrational system described by Eq. (2).

5.2 Experiments and signal processing

For the identification purpose, two series of signals were collected. In the first one, the structure was excited

with chirp signal with the frequency range of 10–500 Hz and five levels of amplitudes (0.10, 0.12, 0.15, 0.18, 0.20 V_{pp}). The measurement sampling frequency was equal to 51,200 Hz. A total of 51,199 data samples were collected in this series. In order to test trained model on a separate set of data, four signals were collected. The force excitations were harmonic signals of frequencies equal to 20, 30, 40 and 50 Hz. The amplitude of the force was linearly increasing from 0.18 to 0.2 V_{pp} with same sampling frequency as chirp signals. Totally, 102,396 data samples were collected in this series. The summary of collected dataset is presented in Table 3.

The signal processing block diagram for measured data is presented in Fig. 14. The procedure starts from band-pass filtering by zero-phase digital filter, to eliminate high- and low-frequency disturbances. Finally, the sampling of the response signal was reduced by 4 to shorten the calculation time.

5.3 Restoring force identification

Before model identification, the training dataset was divided into three subsets: training, validation and test in proportion 70/15/15. The value of mass for identified model was assumed as 0.43 kg. Network structure was optimized using a grid search approach. The authors have investigated three different MLP depths (1, 2 and 3 hidden layers) and five different hidden layer sizes (2, 5, 8, 10 and 13 neurons), making in total 15 MLP structures under investigation. Each of these structures was tested 30 times. In each test, the net was reinitialized to random weights, trained using data from experiment 1 and finally tested using independent dataset (acquired in experiment 2). In each test, the structure of hysteresis network and restoring network was the same. The nets were using a hyperbolic tangent activation function. The results provided by the trained networks are presented in Fig. 15 and Table 4.

The box plots presented in Fig. 15 show networks' performance on independent test dataset. The best

Fig. 15 Performance on test data for network: **a** with one hidden layer, **b** with two hidden layers, **c** with three hidden layers

Table 4 Network performance value features

Number of neurons in layer	Lower whisker (9%)	Lower quantile (25%)	Median (50%)	Upper quantile (75%)	Upper whisker (91%)
2	1.25	1.49	1.63	1.73	1.85
5	1.32	1.57	1.62	1.77	1.89
8	1.48	1.57	1.63	1.72	1.85
10	1.47	1.58	1.75	1.83	2.8
13	1.27	1.55	1.67	1.84	1.98
2/2	0.81	0.89	0.99	1.11	1.21
5/5	0.82	0.87	0.96	1.05	1.16
8/8	0.93	1.02	1.1	1.13	1.50
10/10	0.94	1.02	1.1	1.18	1.29
13/13	0.95	1.12	1.26	1.47	3.42
2/2/2	1.03	1.93	7.71	8.64	8.7
5/5/5	0.85	1.16	5.67	8.7	8.79
8/8/8	1.48	2.92	7.84	8.65	8.69
10/10/10	1.68	3.33	8.62	8.64	8.69
13/13/13	1.37	3.82	8.62	8.65	8.77

results are obtained by the two-hidden-layer networks. In this category, the number of neurons have a minor influence on the network performance, with an exception for the largest network evaluated—for which a performance degradation is observed. The shallow MLPs with only one hidden layer perform notably worse for all the cases investigated—with no clear difference between particular networks. Similarly, a dras-

tic change in results is observed for a relatively deep three-hidden-layer network configurations. It is therefore shown that proper choice of hidden layers' number is significantly more important than choice of neuron numbers for particular depth of the network. On the basis of this observation, two-hidden-layer structure is deemed optimal for the problem at hand.

Fig. 16 Mass influence on performance for two-layer networks with: **a** 2 neurons, **b** 5 neurons, **c** 8 neurons, **d** 10 neurons, **e** 13 neurons in hidden layer

After selecting the number of layers for neural networks, calculations were conducted to select the optimal mass value. A series of tests were carried out for various values of mass related to value adopted at the beginning of subsection. The results are presented in Fig. 16.

From the presented results, the following conclusions are derived: The method is definitely mass-sensitive. While a small reduction in the adopted mass

does not deteriorate the performance, a slight increase causes a deterioration of model response. From the above observations, one can deduce that the proper value of mass can be selected from a relative range of -10 to 0% . Therefore, the mass value was correctly assumed. The possible explanation for the fact that mass sensitivity is not symmetrical around 0 is that the mass was not assumed correctly in the first place.

Fig. 17 Comparison of hysteretic restoring force projection on force and displacement for measured data (solid black line), ANN model (dashed blue line) and ANN ensemble model (dotted magenta line). (Color figure online)

A correct value should probably be estimated at 96% of current estimation.”

For all investigated cases, the spread of the results is very high with mean value equal to 0.98 for the best-case scenario. It means that for each structure

under investigation, there were networks that scored reasonably well and others that performed poorly—only because of a random nature of the training process. The observed spread of the results is a good motivation for using an ensemble approach—to average answers of a group of networks. For further evaluation, the seventh ANN (5/5) is chosen as a classic (non-ensemble) solution, while 10% of nets that performed best in training (when assessed on the “test” subset of the initial set of data) are used to form an ensemble. In Fig. 15b, the results obtained by the selected network are marked with gray and magenta arrows for best network and sets of ensemble members, respectively. The plotted predictions for these model are presented in Figs. 17, 18 and 19.

In Fig. 17, the measured hysteretic relation between the restoring force and displacement is compared with the hysteresis of the same variables obtained from the MLP model. Although the experimental hysteresis is not recreated perfectly with the use of the ANNs, there are still some major qualitative similarities between the two presented curves. The main reason behind such discrepancies is that it was not possible to obtain pure stick

Fig. 18 Comparison of time responses for experimental structure (solid black line), neural network model (dashed blue line) and ensemble model (dotted magenta line) for frequency of: **a** 20 Hz, **b** 30 Hz, **c** 40 Hz, **d** 50 Hz

Fig. 19 Comparison of time responses for experimental structure (solid black line), neural network model (dashed blue line) and ensemble model (dotted magenta line) for frequency of: **a** 20 Hz, **b** 30 Hz, **c** 40 Hz, **d** 50 Hz

and slip motion in the experimental investigation. It can be observed from the numerical model (see Sect. 4) that for such a type of motion, only odd harmonics should be present in the systems responses. This is not the case in the experimental approach, where a strong presence of the even harmonics can be noticed as shown in Fig. 18. Thus, considering the design of the experiment, the observation can be stated that apart from the stick-slide motion, another type of motion could be obtained, namely the disconnection between the surfaces. Unfortunately, such a complex surfaces contact behavior cannot be simulated with a SDOF model. On the other hand, a good agreement is obtained between the experimental response and recreated one in both time (Fig. 18) and frequency (Fig. 19) domains, which lay a basis for a good identification approach based on MLP.

6 Conclusions

In the presented article, a novel method for the identification of the hysteretic stiffness (stick and slip motion)

is introduced. The proposed model is composed of two MLP networks. The first network combined with signal processing block is trained to approximate hysteretic behavior of the dynamic system. The second network approximates non-hysteretic dependencies in the identified system.

The method is first tested on a numerical SDOF system with hysteretic stiffness and then verified on the basis of experimental data. In both scenarios, a good fit for time and frequency domain signals is obtained. The hysteresis loop shape is recreated almost perfectly in the first scenario. In the second case, however, observable discrepancies between estimated and measured hysteresis loops are acquired. A possible cause of this phenomenon is the fact that obtained stick and slip motion is not pure. The model used did not assume less dominant phenomena, which may be revealed as a result of contact such as impact inducing vibrations, or the effect related to forces normal to the contact surfaces. As a result, it may cause errors in the identification process.

Tailoring of the ANN structure resulted in choice of two-hidden-layer ANNs. It is found that this depth

of the network is significantly better than one-hidden-layer or three-hidden-layer solutions. It is also found that choice of ANN depth is far more important than the choice of number of neurons in hidden layers.

Estimation of a mass value has an impact on the results and, when performed poorly, may deteriorate the performance of the system. In the investigated scenario, it was found that the similar results are obtained for mass estimated in range of 90% to 100% of measured mass. Out of this range, the method starts to score significantly lower results. It is expected that errors exceeding the range covered in this work (i.e., more than 10% of mass change) would continue that trend. The limitation of the method is that it can only be used if the excitation is known or measured.

Acknowledgements The work presented in this paper was performed within the scope of the research Project 2015/17/B/ST8/03399 financed by the Polish National Science Centre. The third author would like to acknowledge support from Foundation for Polish Science via START 21.2018 stipend for young researchers.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Johnson, K.L.: One hundred years of Hertz contact. *Proc. Inst. Mech. Eng.* (1982). <https://doi.org/10.1243/PIME>
- Carter, F.W.: On the action of a locomotive driving wheel. *Proc. R. Soc. Ser. A*. **112**, 151–157 (1926)
- Mindlin, R.D.: Compliance of elastic bodies in contact. In: *The Collected Papers of Raymond D. Mindlin Volume I* (1989)
- Iwaniec, J., Uhl, T., Staszewski, W.J., Klepka, A.: Detection of changes in cracked aluminium plate determinism by

- recurrence analysis. *Nonlinear Dyn.* **70**, 125–140 (2012). <https://doi.org/10.1007/s11071-012-0436-9>
- Pieczonka, L., Ukowski, P., Klepka, A., Staszewski, W.J., Uhl, T., Aymerich, F.: Impact damage detection in light composite sandwich panels using piezo-based nonlinear vibro-acoustic modulations. *Smart Mater. Struct.* **23**, 105021 (2014). <https://doi.org/10.1088/0964-1726/23/10/105021>
- Pieczonka, L., Klepka, A., Martowicz, A., Staszewski, W.J.: Nonlinear vibroacoustic wave modulations for structural damage detection: an overview. *Opt. Eng.* **55**, 011005 (2015). <https://doi.org/10.1117/1.OE.55.1.011005>
- Brewick, P.T., Masri, S.F., Carboni, B., Lacarbonara, W.: Data-based nonlinear identification and constitutive modeling of hysteresis in NiTiNOL and steel strands. *J. Eng. Mech.* (2016). [https://doi.org/10.1061/\(asce\)em.1943-7889.0001170](https://doi.org/10.1061/(asce)em.1943-7889.0001170)
- Pecorari, C., Solodov, I.: Nonclassical nonlinear dynamics of solid surfaces in partial contact for NDE applications. *Univ. Nonclassical Nonlinearity* **309–326**, 19 (2006). https://doi.org/10.1007/978-0-387-35851-2_
- Klepka, A., Dziedzic, K., Spytek, J., Mrówka, J., Górski, J.: Experimental investigation of hysteretic stiffness related effects in contact-type nonlinearity. *Nonlinear Dyn.* **95**, 1513–1528 (2018)
- Zaitsev, V.Y., Gusev, V., Castagnède, B.: Observation of the “Luxemburg–Gorky effect” for elastic waves. *Ultrasonics* **40**, 627–631 (2002)
- Joglekar, D.M., Mitra, M.: Time domain analysis of nonlinear frequency mixing in a slender beam for localizing a breathing crack. *Smart Mater. Struct.* (2017). <https://doi.org/10.1088/1361-665X/26/2/025009>
- Dziedzic, K., Pieczonka, L., Adamczyk, M., Klepka, A., Staszewski, W.J.: Efficient swept sine chirp excitation in the non-linear vibro-acoustic wave modulation technique used for damage detection. *Struct. Heal. Monit.* (2017). <https://doi.org/10.1177/1475921717704638>
- Gerven, M., Van, Bohte, S.: Artificial neural networks as models of neural information processing. *Front. Comput. Neurosci.* **11**, 114 (2018)
- Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366 (1989). [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- Worden, K., Staszewski, W.J., Hensman, J.J.: Natural computing for mechanical systems research: a tutorial overview. *Mech. Syst. Signal Process.* **25**, 4–111 (2011). <https://doi.org/10.1016/j.ymssp.2010.07.013>
- Haykin, S.: *Neural Networks: A Comprehensive Foundation*. Pearson Education, Delhi (2005)
- Hagan, M.T., Menhaj, M.B.: Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **5**, 989–993 (1994). <https://doi.org/10.1109/72.329697>
- Chen, L., Kamel, M.S.: Multiple Classifier Systems, 6th International Workshop, MCS 2005, Seaside, CA, USA, June 13–15, 2005, Proceedings. In: *Multiple Classifier Systems* (2005)
- Sbarufatti, C., Manson, G., Worden, K.: A numerically-enhanced machine learning approach to damage diagnosis using a Lamb wave sensing network. *J. Sound Vib.* (2014). <https://doi.org/10.1016/j.jsv.2014.04.059>

20. Dworakowski, Z., Stepinski, T., Dragan, K., Jablonski, A., Barszcz, T.: Ensemble ANN classifier for structural health monitoring. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2016)
21. Dworakowski, Z., Dragan, K., Stepinski, T.: Artificial neural network ensembles for fatigue damage detection in aircraft. *J. Intell. Mater. Syst. Struct.* (2017). <https://doi.org/10.1177/1045389X16657428>
22. Masri, S.F., Chassiakos, A.G., Caughey, T.K.: Identification of nonlinear dynamic systems using neural networks. *ASME J. Appl. Mech.* **60**, 123–133 (1993). <https://doi.org/10.1115/1.2900734>
23. Liang, Y.C., Zhou, C.G., Wang, Z.S.: Identification of restoring forces in non-linear vibration systems based on neural networks. *Adm. Sci. Q.* **15**, 1106–1107 (1997). [https://doi.org/10.1016/0002-9610\(92\)90118-B](https://doi.org/10.1016/0002-9610(92)90118-B)
24. Liang, Y.C., Feng, D.P., Cooper, J.E.: Identification of restoring forces in non-linear vibration systems using fuzzy adaptive neural networks. *J. Sound Vib.* **242**, 47–58 (2001). <https://doi.org/10.1006/jsvi.2000.3348>
25. Pei, J., Smyth, A.W.: New approach to designing multilayer feedforward neural network architecture for modeling nonlinear restoring forces. I: formulation. *J. Eng. Mech.* **132**, 1290–1300 (2006). [https://doi.org/10.1061/\(ASCE\)0733-9399\(2006\)132:12\(1290\)](https://doi.org/10.1061/(ASCE)0733-9399(2006)132:12(1290))
26. Pei, J., Smyth, A.W.: New approach to designing multilayer feedforward neural network architecture for modeling nonlinear restoring forces. II: applications. *J. Eng. Mech.* **132**, 1290–1300 (2006). [https://doi.org/10.1061/\(ASCE\)0733-9399\(2006\)132:12\(1290\)](https://doi.org/10.1061/(ASCE)0733-9399(2006)132:12(1290))
27. Pei, J.S., Mai, E.C., Wright, J.P., Masri, S.F.: Mapping some basic functions and operations to multilayer feedforward neural networks for modeling nonlinear dynamical systems and beyond. *Nonlinear Dyn.* **71**, 371–399 (2013). <https://doi.org/10.1007/s11071-012-0667-9>
28. Pei, J.S., Masri, S.F.: Demonstration and validation of constructive initialization method for neural networks to approximate nonlinear functions in engineering mechanics applications. *Nonlinear Dyn.* **79**, 2099–2119 (2014). <https://doi.org/10.1007/s11071-014-1797-z>
29. Bertotti, G., Mayergoyz, I.D., Damjanovic, D.: The science of hysteresis. *Sci. Hysteresis* **1**, 337–465 (2006). <https://doi.org/10.1016/B978-012480874-4/50022-1>
30. Kosmatopoulos, E.B., Smyth, A.W., Masri, S.F., Chassiakos, A.G.: Robust adaptive neural estimation of restoring forces in nonlinear structures. *J. Appl. Mech.* **68**, 880 (2001). <https://doi.org/10.1115/1.1408614>
31. Pei, J.S., Smyth, A.W., Kosmatopoulos, E.B.: Analysis and modification of Volterra/Wiener neural networks for the adaptive identification of non-linear hysteretic dynamic systems. *J. Sound Vib.* **275**, 693–718 (2004). <https://doi.org/10.1016/j.jsv.2003.06.005>
32. Park, J.K., Washington, G., Yoon, H.S.: A hybrid approach to model hysteretic behavior of PZT stack actuators. *J. Intell. Mater. Syst. Struct.* **20**, 467–480 (2009). <https://doi.org/10.1177/1045389X08093547>
33. Farrokh, M., Dizaji, M.S.: Adaptive simulation of hysteresis using neuro-Madelung model. *J. Intell. Mater. Syst. Struct.* **27**, 1713–1724 (2016). <https://doi.org/10.1177/1045389X15604283>
34. Nakshatharan, S.S., Punning, A., Aabloo, A.: A neural network modeling and sliding mode control of self-sensing ionic polymer–metal composite actuator. *J. Intell. Mater. Syst. Struct.* **28**, 3163–3174 (2017). <https://doi.org/10.1177/1045389X17733234>
35. Pei, J.S., Gay-Balmaz, F., Wright, J.P., Todd, M.D., Masri, S.F.: Dual input-output pairs for modeling hysteresis inspired by mem-models. *Nonlinear Dyn.* **88**, 2435–2455 (2017). <https://doi.org/10.1007/s11071-017-3388-2>
36. Sharkey, A.J.C.: *Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems*, vol. 298. Springer, Berlin (2012)
37. Dietterich, T.G.: *Ensemble Methods in Machine Learning*. Presented at the (2000)
38. Sharkey, A.J.C., Sharkey, N.E., Gerecke, U., Chandroth, G.O.: The “Test and Select” Approach to Ensemble Combination. Presented at the (2000)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.