

This is the author-created version of the following work:

Chalak Qazani, Mohamad Reza, Sajed, Moosa, Pedrammehr, Siamak, and Seyedkashi, Seyed Mohammad Hossein (2025) *Mechanical properties estimation of multi-layer friction stir plug welded aluminium plates using time-series neural network models*. Soft Computing, 29 pp. 1147-1168.

Access to this file is available from:

<https://researchonline.jcu.edu.au/86698/>

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2025

Please refer to the original source for the final version of this work:

<https://doi.org/10.1007/s00500%2D025%2D10429%2Dx>

Mechanical Properties Estimation of Multilayer Friction Stir Plug Welded Aluminium Plates Using Time-series Neural Network Models

Mohammad Reza Chalak Qazani¹, Moosa Sajed², Siamak Pedrammehr³, Seyed Mohamad Hossein Seyedkashi⁴

Abstract—A multilayer friction stir plug welding can be used to fix the thick aluminium plates. Optical microscopy and tensile tests are utilised to study the microstructural and mechanical characteristics of the welded aluminium plates. However, finding the relation between the indexes of the process and the mechanical properties would be challenging. The present work aims to devise a time-series machine learning model including a recurrent neural network (RNN) and nonlinear autoregressive network with the external state (NARX) to estimate the mechanical properties of the repaired aluminium plate using the force-extension plot. The ultimate tensile strength, yield strength, impact energy and elongation of the repaired aluminium plate can be calculated based on a force-extension plot trained and extracted using the developed networks. In addition, the Bayesian technique is employed to recalculate the optimal hyperparameters of RNN and NARX, targeting the lowest root mean square error (RMSE) between the target and the estimated force during the testing. The investigated methods (RNN and NARX), in addition to classical estimation methods, including decision tree and support vector regression, are modelled in MATLAB. The outcomes prove the proposed NARX model efficiency in terms of lower RMSE than support vector regression, decision tree and RNN.

Index Terms—mechanical properties estimation, recurrent neural network, nonlinear autoregressive, Bayesian method.

(Corresponding author: Mohammad Reza Chalak Qazani)

M. R. C. Qazani with the Institute for Intelligent Systems Research and Innovation, Deakin University, Geelong, VIC 3216, Australia (e-mail: m.r.chalakqazani@gmail.com)

M. Sajed is with the Mechanical Engineering Department, Azarbaijan Shahid Madani University, Tabriz, Iran (e-mail: sajed@azaruniv.ac.ir)

S. Pedrammehr is with the Faculty of Design, Tabriz Islamic Art University, Tabriz, Iran, and he is with the Institute for Intelligent Systems Research and Innovation, Deakin University, Geelong, VIC 3216, Australia (e-mail: s.pedrammehr@gmail.com)

H. Seyedkashi is with the Mechanical Engineering Department, University of Birjand, Birjand 97175-376, Iran (e-mail: seyedkashi@birjand.ac.ir)

I. INTRODUCTION

Nowadays, aluminium and its alloys are getting more and more attention because of their distinctive properties. Thus, it is necessary to find novel methods for repairing and maintaining parts made from aluminium alloys, considering the high oxidation rate and low weldability in high temperatures. Solid-state welding solves the low weldability of aluminium alloys of specific grades [1]. As a recent solid-state welding method, friction stir welding (FSW) is an effective process for joining aluminium alloys [2]. FSW can be employed in various operations such as lap, butt, and spot welding [3]. In addition to the necessity of using aluminium alloys in industry, there are different types of time-consuming tests on the materials, such as compression tests, tensile tests, creep tests, etc. As a result, machine learning methods are employed to estimate the material property information faster than the previous time-consuming traditional methods. Machine learning methods are applied in material science to predict physical properties, new compound discovery, crack location prediction, fatigue life, twin nucleation prediction, and automatic microstructure analysis.

Friction hydro-pillar processing (FHPP) and Friction taper plug welding are two solid-state operations for thick steel plate maintenance. A cylindrical hole is filled via a rotating stud in FHPP under a load. The load is implemented to produce the continuous plasticised layer. The plasticised layer develops faster than the stud feed rate. Then, it is left beneath a dynamically recrystallised deposit material. A further forge load is required at the end of the stage to reach a finer microstructure. Friction taper plug welding is quite similar to the FHPP. The mentioned operations have been studied from various aspects of literature. Buzzatti *et al.* [4] investigated the material's toughness that FHPP had welded. They discovered that the higher axial forces increase the friction hydro pillar processed offshore mooring chain steel toughness properties. Landell *et al.* [5] studied FHPP by reporting a coupled theoretical model with the addition of a three-dimensional X-ray tomography investigation. They [5] discovered the deformation of the plugin's central portion. Wukai *et al.* [6] proposed a numerical study for the temperature field and frictional interface in FHPP. Xu *et al.* [7] studied overlap welding using the FHPP method. Sajed and Seyedkashi [8] developed a variant of FHPP to process the micro-alloying in the depth of thick steel sheets. FHPP has been considered for steel, and it cannot directly be applied to aluminium because of its high oxidation rate behaviour and reduced high-temperature strength [9]. A specific heat amount is needed to fill a void for aluminium softening. The extreme frictional heat of FHPP lowers the mechanical properties and structure of aluminium alloys. Gadakh and Adepu [10] developed the

heat generation model in FSW for a taper cylindrical pin via regenerated heat from the pin tip surface, shoulder surface and taper probe surface. Using the same condition, they discovered that taper cylindrical pin generates less heat than a straight cylindrical pin. It proves the essential need to develop the process to repair voids in aluminium alloys. Mahoney et al. [9] proposed an aluminium alloy voids as a new repair method. They targeted removing the aluminium plate's keyhole, which remains after FSW. They proposed void machining to prepare a tapered hole with some sidewall ridges and grooves. However, this approach cannot be considered for deep and big holes, according to the conducted study by Sajed and Seyedkashi [11]. Sajed [12] proposed a two-step friction stir spot welding operation to avoid keyholes in spot joint production. Multilayer friction stir plug welding (FSPW) was introduced by Sajed and Seyedkashi [11] to repair voids and cracks in thick aluminium plates. Their proposed method is based on FSW to address issues regarding aluminium alloys FHPP.

Machine learning and deep learning methods have recently been designed and have gained remarkable interest across different studies, such as manufacturing, economics, energy, and climate [13-16]. The learning-based model, including machine and deep learning methods, can be used to calculate the appropriate imitation model of the system using the experimental or finite element dataset without any need to understand the system's complicated mathematical model or the procedure's physics. Then, they can yield models with a higher degree of precision than classical mechanistic models. That is the main motivation for researchers who used learning-based methods to understand the concept of FSW and the mechanical behaviour of final products. Hwang *et al.* [17] investigated a four-layer multilayer perceptron (MLP) for extraction of the mechanical properties of the rolled steel bar. Alkayem *et al.* [18] used soft computing methods (including MLP and evolutionary optimisation methods) to calculate the optimal process parameters of friction stir welding. The process parameters are yield stress, tensile strength, percentage elongation, hardness, and bending angle. They used the full factorial Taguchi design experiment to extract the dataset for training and developing their proposed model. The results are validated using the practical experiment, and a lower margin error proves the efficiency of their proposed research method. Pilia et al. [19] were able to fasten the materials property predictions via kernel ridge regression. Rupp *et al.* [20] used the traditional mathematical model, including Kernel ridge regression and nonlinear regularised regression, to calculate the quantum mechanical properties of atoms in molecules. Ma *et al.* [21] used two-step propagation to evaluate the design properties of mechanical

products. Karina *et al.* [22] estimated the tensile strength of corroded steel plates using a simple feedforward neural network to simplify, accurise and reduce the cost of the process. Feedforward neural networks can be trained using mathematical or evolutionary algorithms. Evans *et al.* [23] employed the gradient boosting regressor to estimate the zeolite frameworks' mechanical properties. Altarazi *et al.* [24] predicted and classified the tensile-strength behaviour of the different compositions' polymer films as a process function using decision tree (DT), k-nearest neighbours (KNN), support vector machine (SVM), AdaBoost, random forest (RF), feedforward neural network, stochastic gradient descent and linear regression. Rakshe *et al.* [25] reviewed the adhesively bonded spar-wingskin joints for a loaded composite. Gong *et al.* [26] used MLP, SVM, multivariate adaptive regression spline, and multivariate linear regression to estimate mechanical rock properties. Golkarnarenji *et al.* [27] used MLP and SVM to forecast carbon fibre mechanical properties. Qi *et al.* [28] employed DT to calculate the mechanical properties of carbon fibre. Zhang *et al.* [29] KNN, SVM DT and MLP for mechanical properties of graphene. Wang *et al.* [30] RF to forecast the mechanical properties of tungsten disulphide monolayer. Xu *et al.* [31] forecasted the tensile properties of AZ31 magnesium alloys based on alloy compositions and processing parameters using a feedforward neural network and SVM. Behnood and Golafshani [32] used a Gradient-boosted tree regression model to estimate the mechanical properties of concretes containing waste foundry sand. Chaabene *et al.* [33] used SVM, MLP, and DT with an evolutionary algorithm to forecast the concrete's mechanical properties. Xiong *et al.* [34] employed RF, KNN, MLP, and linear least squares to estimate the mechanical properties of steels using production process parameters. Fathi *et al.* [35] used a group method of data handling to estimate the mechanical properties prediction of wood. Deng *et al.* [36] used standard linear regression, MLP, SVM, sequential minimal optimisation-based SVM and puk kernel to predict the mechanical properties of Cu-Al alloy based on the production process parameters. Mamun *et al.* [37] modelled the rupture strength of austenitic alloys and iron-based martensitic. They used gaussian process regression, DT and feedforward neural network methods. Xie *et al.* [38] used machine learning techniques to design the mechanical properties online estimation for the hot rolled steel plate. Güçlüer *et al.* [39] used MLP, SVM, DT and linear regression for concrete compressive strength estimation. Ferreño *et al.* [40] used KNN, DT, RF, MLP, SVM, multilinear regression and gradient boosting to predict the rail pads' mechanical properties under in-service conditions. Xiong *et al.* [41] used RF to calculate mechanical properties in concentrated/complex alloys. Nafees *et al.* [42] used DT and SVM to

estimate the mechanical properties of silica fume-based green concrete. Stoll *et al.* [43] used a linear regression model and RF to estimate the mechanical properties of different materials. Shah *et al.* [44] employed Optimised multi-expression programming based on particle swarm optimisation for predicting sustainable concrete mechanical properties. Demirci *et al.* [45] used an image processing technique to detect the damage of carbon face sheet nomex sandwich composites. Huang *et al.* [46] used SVM and MLP to predict the mechanical properties of carbon nanotube-reinforced cement composite. Dabiri *et al.* [47] used nonlinear regression, ridge regression and feedforward neural network to estimate the tensile strength of steel bars which is used in reinforced concrete structures. Hayashi and Ohsaki [48] used reinforcement learning and graph embedding for the optimal cross-sections of steel frames. Najjar *et al.* [49] employed random vector functional link network and red fox algorithm for estimating the mechanical properties of Cu–Al₂O₃ nanocomposites. Liu *et al.* [50] used SVM and MLP for prediction of mechanical properties of graphene/aluminium nanocomposite. Diao *et al.* [51] used SVM with radial basis function kernel, MLP, KNN, RF, and kernel ridge regression for mechanical properties prediction of carbon steels. Table I is the reviewed literature in this study, which used machine learning models to estimate the mechanical properties of materials in different applications.

In contrast to the literature [17-51], our approach leverages the tensile-extension behaviour of the final product to estimate mechanical properties, thereby mitigating the reliance on process parameters. By capturing the force-extension behaviour of repaired aluminium plates through tensile testing, our method provides a non-destructive means of estimating mechanical properties. Specifically, we utilise time-series neural network models trained on the complete tensile-extension plot of the final product. This allows us to predict the material's tensile-extension behaviour during the plastic deformation stage using elastic deformation data acquired from tensile tests.

Our method offers several advantages over traditional approaches:

1. Independence from process parameters: By focusing on the material's tensile-extension behaviour, our model reduces the need for extensive experimentation related to process parameters, streamlining the model development process.
2. Non-destructive testing: Eliminating the need to destroy materials for mechanical property calculation, our method relies solely on the elastic behaviour observed during tensile testing, enabling the extraction of the final product's complete tensile extension plot.

3. Applicability across materials: As our model depends on the material's intrinsic behaviour, it can be applied to materials with similar metrological structures.

In this study, we employ two highly efficient time-series neural networks, recurrent neural network (RNN) and nonlinear autoregressive with the existence of external states (NARX), alongside traditional methods, including decision tree (DT) and support vector regression (SVR). We introduce a novel global optimisation method to determine optimal hyperparameters for RNN and NARX models, enhancing prediction accuracy. Additionally, we apply the optimisation algorithm to DT and SVR methods to ensure a fair comparison. Leveraging Bayesian optimisation, our approach efficiently achieves optimal model performance.

It's imperative to highlight that our dataset, obtained from tensile tests, is divided into 70% training and 30% testing sets for developing and evaluating our globally optimised RNN (Opt-RNN) and globally optimised NARX (Opt-NARX) models.

The overview of the multilayer FSPW for aluminium plates is described below. The methodology of this research, which includes RNN, NARX, and Bayesian optimisation, is explained in Section III. The proposed techniques are investigated in MATLAB, and the results are shown and discussed in Section IV. Section V presents the most remarkable results and conclusions.

II. OVERVIEW OF MULTILAYER FRICTION STIR PLUG WELDING

There is no appropriate operation for voids and crack repair at the depth of thick aluminium plates. Multilayer FSPW is a new method introduced by Sajed and Seyedkashi [11] using stir welding to present problems relevant to the FHPP of aluminium alloys. A multilayer friction stir plug welding can be used to fix the thick aluminium plates. This process overcomes the problems related to the aluminium alloys' friction hydro-pillar process, which consists of a high oxidation rate and excessive softening at raised temperatures. A cross-section hole is drilled to eliminate the cracks and voids and filled via multiple plugs with the same base material composition. The plugs are welded, and the base material is via a non-consumable tool.

The heat input parameter is essential and must be controlled in operation. Thus, dwell time and rotation speed are the basic parameters that control the heat during the operation [52]. Also, the contact pressure between the tool and welded workpieces significantly influences heat generation [53]. The aluminium plate has drilled a hole to eliminate the void or crack. The mentioned hole has been filled using non-consumable tools and some plugs. The smallest plug has been utilised at the

hole bottom. The diameter of the plugs is increased to fill the hole from bottom to top. The plug should be stirred and welded to the sidewall at every process step. Also, the diameter of the lower plug should be a bit larger than the hole's local diameter in FSW. In each step, the keyhole is filled with the material of the upper plug. A simple without-pin shoulder is employed to fill the last keyhole on the top. Any crack and void could be effectively repaired with any dimension and at any depth using this process.

A. Materials and Methods

The investigated material in this study is an ASTM E1251 standard. The ASTM E1251 was investigated via optical emission spectroscopy, and the outcomes are presented in Table II. It should be noted that the studied aluminium alloy has no heat treatment. The test parts centre drilled a specific hole with 10 (mm) overall depth. The dimensions for the machined hole are shown in Fig. 1a. To drill the hole, end ball mills with 12 and 16 (mm) diameters and 10 and 6 (mm) penetration depths were used. The hole's spherical form leads to a smooth material flow. The hole dimensions and form have been considered according to error and trial tests.

According to the operation steps, four shoulders have been produced from the CK45 steel bar and heat-treated. The shoulder without a pin is just used to cover the step. However, the other three have been fabricated with different conical pins for the welding operation. According to Fig. 1b, the tool diameters are $11 \times 15 \times 18$ (mm³), and all pins' noses diameter is 4 (mm). The length of the pins is 4 and 5 (mm) for 2 and 1 tools, respectively. A pin with a length of 5 (mm) has been used to combine more material from the hole to obtain better bonding and mixing. The shoulder without a pin is concave with an 18 (mm) diameter and 5° angle for better material flow. The shank edges have also been machined for better heat transfer.

Multilayer friction stir plug welding (MFSPW) is a newly developed solid-state repair method that can be applied on thick aluminium plates and is presented step by step in Fig. 2. The common repair technique for thick steel plates is friction hydro-pillar processing. It cannot be applied to aluminium alloys due to the significant drop in the strength of aluminium alloys at elevated temperatures. This phenomenon is explained comprehensively in our previous paper [10]. In this method, a hole with a predefined geometry is drilled to eliminate the crack or void in the depth of the aluminium plate, see Fig. 2a. Then, considering the depth of the hole, a number of plugs are inserted in the hole and welded to the hole wall and base metal one after another using a set of friction stir welding tools. Each keyhole is refilled when the next plug is welded (Fig. 2b-g). The

keyhole of the upper layer is refilled using a non-pin tool according to a procedure called two-stage refilled friction stir spot welding developed by [12] see Fig. 2h.

Three 5 (mm) height plugs with 8, 12, and 14 (mm) diameters have been employed. The plugs' material is the same as the base. The plug with a diameter of 8 (mm) is placed at the hole bottom. Afterwards, the 11 (mm) diameter shoulder is utilised to weld the plug inside the hole. It should be noted that the total tool penetration depth should be more than the hole depth, which is 12 (mm), to combine the base material with the plug. Then, the hole is filled via the 12 (mm) diameter plug and welded via the 15 (mm) diameter tool. This plug is welded to both the lower plug and the sidewall of the hole. Therefore, the previous step's keyhole is filled in this stage.

It should be noted that the total tool penetration depth is 9 (mm). Afterwards, the plug with 14 (mm) diameter was placed and welded by the 18 (mm) diameter tool. In this step, the total tool penetration is 6 (mm). The keyhole is finally filled using the non-pin shoulder with 3 (mm) penetration depth. All the dimensions and welding depths are considered according to primitive experiments. The whole process is shown in Fig. 3.

B. Data Acquisition

The dimensions of specimens are 100×30×30 (mm) to be used in tensile tests. The tools' rotational speed and dwelling time are two essential parameters in FSW that affect frictional heat generation. For all the specimens, the rotational speed of the tool is fixed in all steps. The critical point for calculating the dwell time is satisfying the tool penetration depth in every step. In the full factorial design of experimental tests, 1000 and 2000 (rpm) rotational speeds and 5 and 10 (seconds) dwell times have been considered. The FP4M milling machine has been used for the welding operation. The summary of the experimental tests is listed in Table III.

Three experiments have been carried out for the tensile test in each setting. A typical prepared tensile test specimen and its dimensions and the cutting line are presented in Fig. 4. The tensile test is accomplished using Samantha tensile test machine (Zwick controller) under the standard of ASTM E8/E8M-08 using 10 (mm/min) for tensile rate. As the main contribution of this study is to estimate the mechanical properties of the products without destroying them, then the extracted tensile extension via tensile test is employed as the dataset in a time-series neural network. The time-series neural networks can predict the material's tensile-extension behaviour based on its behaviour during the elastic deformation without destroying it.

In this process, four different zones are detectable in the welding zone, including the bonding

zone (BZ), plug thermomechanical affected zone (PTMAZ), base metal thermomechanical affected zone (BTMAZ), and heat-affected zone (HAZ), as presented in Fig. 5a-d [10]. PTMAZ consists of plug material stirred by the tool and experiences a thermomechanical process and dynamic recrystallisation. Fine microstructure and homogenous distribution of precipitates are the main characteristics of this zone. The bonding zone also experiences a heavy thermomechanical process. The rotary friction between the plug and base metal generates heat, facilitating the material flow and bonding when the pressure is applied through the tool shoulder. Thus, the microstructure in this zone will be similar to the PTMAZ. However, the grain is larger due to the tool pin not stirring it. BTMAZ is a zone from the base metal which experiences thermomechanical work. The heat is generated due to friction between the plug and the hole wall, rotation of the tool shoulder, and plastic deformation. The resulting temperature rise leads to a material with lower flow stress than the base metal and plastic deformation in this zone. HAZ is a zone in the base metal that experiences almost no plastic deformation and, due to the heat conduction, experiences a significant temperature rise. The result is a coarse microstructure that is evident in Fig. 5d. The different zones of the welding nugget are presented in Fig. 6 for all four specimens. Based on the outcomes of the tensile strength, the specimens welded by 3000 (rpm) rotational tool speed and 5 (seconds) dwelling time have the highest strength. The comparison of welding nuggets in different specimens using Fig. 6 indicates that the width of the bond zone is bigger in this specimen, which could be the reason for the higher strength. The HAZ is considered the zone where the crack starts to grow. Thus, the grain size in the HAZ zone is a critical index in determining the strength of the joint.

Fig. 7 shows the force variation based on the extension of the sample during the tensile test for four samples (repaired aluminium). The input of the developed time-series networks is the elastic deformation of the final product. At the same time, the proposed methods try to predict the plastic deformation part of the tensile-extension plot. The study can be divided into two parts: training and testing. During the network training process, the stress-strain plot is used to develop the time-series machine learning models (RNN and NARX). It is possible to use the stress-strain plot as a time-series signal if we consider the strain as the time interval of the signal. For using the developed time-series machine learning methods in real-life applications, the stress-strain behaviour of the final products during the elastic deformation is captured without destroying the final products. Then, this dataset is used as input for the developed model to calculate the final product's stress during plastic deformation. The combination of experimentally captured elastic deformation of the material with

estimated plastic deformation results in the whole stress-strain plot, which can be used to calculate the final products' mechanical behaviours.

III. METHODOLOGY

This paper aims to estimate the force based on extension data to calculate the mechanical properties of the repaired aluminium using the multilayer FSPW method. The employed data inside the neural network model is obtained via tensile test for the four mentioned and non-repaired samples. The time-series neural network model has not yet been employed in force-extension data. It is the first time time-series data has been employed for the force-extension dataset. The reason for using the time-series neural network model for the force-extension dataset is the behaviour of the data. The current tension or force of the sample ($T(t+1)$) has a relation with the previous tension or force sampling data ($T(t)$) in addition to the previous normal extension or extension of data ($x(t)$). The problem is that the extension did not show the behaviour the same way as time because the difference between them is not constant.

However, we have used the Simulink environment of MATLAB to recalculate tension or force regarding the equal extension difference. Then, after data-pre-processing, the extension can be treated as time in our dataset. Fig. 8 represents the schematic proposed technique in this paper as the main contribution, combining a time-series neural network (including RNN and NARX) and the Bayesian optimisation method. According to Fig. 8a, the process starts with repairing the crack of the aluminium plate using the multilayer FSPW.

The repaired aluminium plate is then machined to be prepared for the tensile test (Fig. 4). The tensile test has proceeded, and the results have been recorded. Fig. 8b shows the implementation of our newly developed optimised RNN and NARX on force-extension data. The data pre-processing is used inside the developed algorithm to eliminate the out-of-range data and enhance the algorithm's accuracy. The normalisation/standardisation is then used to reduce the system's complexity for a network. It should be noted that the extension is a time to consider the stress-extension plot as a time-series signal. In other words, the captured stress-extension plot should be extracted as a constant extension interval to substitute the strain with a time interval. In addition, the purpose of the developed algorithm is to calculate the final product's stress based on the product's strain interval during the plastic deformation. In addition, the Bayesian optimisation method selects the RNN/NARX optimal hyperparameters to reach the lowest root mean square error (RMSE) between

the actual and predicted data. There is no need for sequencing data as the employed networks in this paper are time-series networks.

A. *Pre-Tuning of Algorithm*

Three stages should be applied to the data before applying the dataset and developing the model. Initially, there were some unrealistic data (negative force), according to Fig. 8. It should be noted that there were 298, 569, 456, and 357 datasets for samples 1-4, respectively. These data should be eliminated because they are produced as a result of the malfunction of the system. Because of the sensor's fault, the out-of-range data should also be removed before employing it inside the model.

Normalisation/standardisation is used in the second step to reduce the complexity of the dataset. In this research, both techniques are used to enhance the network's input data's complexity and the system's accuracy. The standardised dataset (input/output) can be calculated as follows:

$$\sigma_{x_i} = \frac{x_i - \bar{x}}{\sigma_x} \quad (1)$$

where σ_{x_i} and x_i are the i th standardised and raw input data, respectively. σ_x and $\bar{x}$ are also the standard deviation of the data and functions for extracting the average. The following formula is used to calculate the normalised dataset:

$$n_{x_i} = \frac{x_i - \underline{x}}{\bar{x} - \underline{x}} \quad (2)$$

where $\underline{x}$, $\bar{x}$ and n_{x_i} are the minimum and maximum of the dataset and the normalised input dataset. The dataset is divided to the training (70%) and testing (30%) process of the network as the final stage of the pre-tuning process. Four models, including DT, SVR, RNN and NARX, have been employed in this research to estimate the next step force based on the history of the sample extension. As RNN and NARX have the delay block inside the network, there is no need for a sequencing block in the algorithm.

B. *Recurrent Neural Network*

A directed circular connection exists between the neurons in an RNN. Inside the network, the inner states are made with the ability to reveal model nonlinear dynamic behaviour. Compared to the non-time series neural network models, RNN can better tackle free input sequences. RNN utilises the inputs' sequential information to obtain better outcomes, as contrasted to DT and SVM, in which the outputs and inputs are not dependent. In an RNN, in the previous states, the outputs are dependent. RNN can obtain past sequences because it has a memory block that records the previous

states of the system.

Fig. 9a shows an RNN schematic structure including the output, hidden, and input states at time t , respectively, as o_t , s_t and x_t [54]. It should be noted that the current input u_{xt} is the current tensile stress of the material at the current time interval (extension interval). In addition, the output is the tensile stress of the next time-step (extension interval). The network's current memory is calculated using the current input and previous memory of the network as the following:

$$s(t) = f(u_{xt} + w_{st-1}) \quad (3)$$

The Levenberg Marquardt (LM) method is used for RNN training. LM can properly balance Newton's method and the steepest descent convergence [55].

An LM-trained RNN can be used in electricity usage prediction and electrical information-optimised control [55]. Although RNN has an acceptable capability for handling time-series data, gradient vanishing and exploding issues are its disadvantages.

It should be noted that in Fig. 9a, x_t is the extension data, and T_t is the force of the repaired aluminium plate at time t . Also, n is the delay number for output. The explained RNN should be designed and developed parametrically based on the number of layers, neurons, feedback step, and delay. These parameters are selected via the optimiser as the hyperparameters of the RNN to reach the highest efficient network for estimation of the mechanical properties of the repaired aluminium plate.

C. *Nonlinear Autoregressive with Existence of External States*

Chow and Leung [56] developed the nonlinear autoregressive network to tackle time-series datasets. Fig. 9b illustrates a NARX using n time-step delays. The structure consists of three layers: input, hidden and output. The output can be obtained as the following:

$$T(t) = f(T_{t-1}, \dots, T_{t-n_y}, x_{t-1}, \dots, x_{t-n_u}) + \epsilon_t \quad (4)$$

where x , T and ϵ are the external variable, output and model noise. n_x and n_T are respectively the delay of the external variable and output. Nonlinear function f is utilised to estimate the outputs via the external variable and past signals. The LM approach is utilised for network training as LM is the least computationally expensive method to achieve adequate outcomes. The Hessian matrix $\mathbf{H}$ can be obtained using the Jacobian matrix $\mathbf{J}$ as the following:

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} \quad (5)$$

$\mathbf{J}$ can be calculated as:

$$\mathbf{g} = \mathbf{J}^T \mathbf{e} \quad (6)$$

in which $\mathbf{e}$ is the model's error in each sampling time. The Jacobian matrix is calculated using the backpropagation method to reduce the training process's computation load as follows:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{J}^T \mathbf{J} + \mu \mathbf{I})^{-1} \mathbf{J}^T \mathbf{e} \quad (7)$$

in which μ is related to the network learning rate. The LM method's adjustment parameters have been selected according to the research by Moré [57] to reach a higher model performance. In Fig. 9b, x_t is the extension data, and T_t is the force of the repaired aluminium plate at time t . The explained NARX should be designed and developed parametrically based on the number of layers, neurons, feedback step, and delay. These parameters are selected via the optimiser as the hyperparameters of the RNN to reach the highest efficient network for estimation of the mechanical properties of the repaired aluminium plate.

D. Bayesian Optimisation method

Thomas Bayes is the inventor of Bayes' theorem, which is employed to calculate the probability of an incident using statistical information from the system [58]. A Bayesian optimisation technique is a solver with the ability to consider the limitation to extract the optimal state of the system to reach the minimum cost function [59]. The cost function is typically deterministic or stochastic. Also, the state can be continuous reals, integers, or categorical. The deep learning technique employs the Bayesian optimisation technique [60, 61] for extracting the automatic tuning application. The tuning of the machine learning method is often a black art that needs expert knowledge to reach high performance. The Bayesian method works through an iterative development of the global statistical model of the unknown objective function. The optimisation starts with a likelihood and prior over functions. A posterior distribution is calculated according to each iteration's previous objective function evaluation. The elements in the Bayesian optimisation method are a Gaussian model of the cost function, a Bayesian updating procedure, and an acquisition function. An acquisition function assesses the location of the state using the posterior distribution function.

The Bayesian optimisation technique is employed to recalculate the hyperparameters of the RNN and NARX, including the number of layers, number of neurons, feedback step and feedback delay. The search range of the Bayesian method is defined as follows:

1. The number of layers is between 1 and 20.
2. The number of neurons is between 1 and 400.
3. The feedback step is between 1-5.

4. The feedback delay is between 1-20.

The proper definition of cost function is essential in calculating the relevant results based on the system's needs. The modular definition of the sequencing block and MLP model is used inside the Bayesian optimisation method. Then, the model is trained and assessed using the predefined values via the optimisation algorithm. The cost function is defined to decrease prediction error. Then, the mean square error (MSE) between the actual T and predicted value Y during the testing process is defined as the cost function for the Bayesian optimisation algorithm.

$$J = MSE(T - Y) \quad (8)$$

IV. RESULTS AND DISCUSSIONS

Four mentioned models in Section III, including DT, SVR, RNN and NARX, have been modelled in MATLAB to show the higher efficiency of our newly proposed Opt-NARX compared to the Opt-RNN and two other traditional prediction models (DT and SVR). It should be noted that the Bayesian optimisation method is employed automatically using the MATLAB function (OptimizeHyperparameters) to extract the optimal DT and SVR model.

A. Verification

DT and SVR are prepared using the previous studies [42]. The Opt-RNN has been prepared according to the presented mathematical model in Sections III.D and III.F. The Opt-NARX has also been prepared based on the presented mathematical model in Sections III.E and III.F. It is worth mentioning that the required data from the tensile test is divided into training (70%) and testing (30%) of the proposed methods.

MATLAB software's `fitr tree`, `fitrsvm`, `layrecnet` and `narxnet` functions are used to design the DT, SVR, RNN and NARX models. The algorithms mentioned in Section III are designed using MATLAB software, and all the results are represented via the `plot` function. Furthermore, the Bayesian optimisation method is designed using the `bayesopt` function of MATLAB with the definition of 14 hours and 200 iteration checks as the boundaries of the method.

The Bayesian optimisation method extracts the optimal sequencing and indexing parameters of the proposed RNN and NARX. Table V represents the optimised hyperparameters of the proposed RNN and NARX for five testing samples.

In addition, the different sets of hyperparameters are chosen for the sample 1 model to be

compared with the selected optimised hyperparameters using the Bayesian optimisation method and show the influence of different parameters. Table VI shows the RMSE between the actual and predicted tensile of sample 1 during the network testing process using the tensile test dataset using the number of layers 1 and 3, a number of neurons 2 and 4, and sequencing 1,4 and 1,2,3,4. The optimised solution based on Table V is the number of layers=2, number of neurons=2, and sequencing 1,4. Initially, the represented results prove the efficiency of the Bayesian optimisation method in extracting the optimal solution in a short period. Also, the higher number of layers has the most significant influence on the algorithm's accuracy. The mean of RMSEs between the actual and predicted tensile of sample 1 using the NARX with 1 layer and the different number of neurons and sequencings is 7.9207×10^{-4} MPa. At the same time, it is 9.62×10^{-2} MPa using the NARX with 3 layers. Also, the higher number of neurons results in a higher RMSE between the actual and predicted tensile of repaired aluminium via FSW. The results also show less dependency on the sequencing strategy of the network.

Fig. 10a presents the fitness function of the RNN for sample 1 data using the Bayesian optimisation method concerning the four hyperparameters mentioned above (number of layers, number of neurons, feedback step and feedback delay). Fig. 10b also shows the fitness function for recalculating the optimal indexing parameters of the NARX. After developing these three models in MATLAB software, they have been compared together to show the efficiency of our newly proposed model in predicting 1-day volatility with better accuracy and less error.

B. Validation

Five parameters have been used to validate our proposed method compared with the traditional methods, including correlation coefficient (CC), MSE, RMSE, normalised root mean square error (NRMSE), and R-square.

These validation parameters are calculated as follows:

$$CC = \frac{\sum_{i=1}^n (x_i - \bar{x}) (T_i - \bar{T})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (T_i - \bar{T})^2}} \quad (9)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (T_i - \hat{T}_i)^2 \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (T_i - \hat{T}_i)^2}{n}} \quad (11)$$

$$\text{NRMSE} = \frac{\text{RMSE}}{\bar{T}} \quad (12)$$

$$R^2 = 1 - \frac{\sum (T_i - \hat{T}_i)^2}{\sum (T_i - \bar{T})^2} \quad (13)$$

where n , x_i , $\bar{x}$, T_i , $\bar{T}$, and $\hat{T}_i$ are the number of samples, the i^{th} input, the mean of inputs, the i^{th} output, the mean of outputs, and the i^{th} predicted value.

Table VII shows the outcomes of the four techniques investigated during the training/testing processes. Based on the shown results, SVR is the weakest algorithm during the training and testing stage of the data, with the lowest CC and R-square and highest MSE, RMSE, and NRMSE. NARX can reach the highest accuracy during the network training and testing process. For samples 1, 2, 3 and 4, the highest CC during the testing process was achieved via Opt-NARX with 0.99924, 0.99925, 0.98698, and 0.9978, the perfect shape similarity between the two signals.

The CC of another 2 traditional prediction methods is not satisfactory for all samples during the testing process of the models. The lowest MSE for samples 1-4 are required via Opt-NARX with 1.0927×10^{-4} , 1.319×10^{-6} , 5.4335×10^{-7} and 8.338×10^{-6} Newton, based on the represented results in Table VII. In the same way, Opt-NARX can beat the DT, SVR and RNN with the lowest RMSE and NRMSE values for all samples. Opt-NARX is also able to increase the R-square from -4.6443, 0.15326, 0.028548, and 0.28824 to 0.99866, 0.99845, 0.97853 and 0.99986 compared with RNN for sample 1-4, respectively.

Fig. 11 shows the results of the proposed Opt-NARX during the training stage in a visualised format. The actual and estimated force of Opt-NARX during the training stage of sample 1 has been illustrated in Fig. 11a. It shows that the CC is 0.99989 between these two signals. Fig. 11b shows the actual and estimated force of Opt-NARX during the training stage of sample 2, which reached 0.99971 for the CC between these two signals. Fig. 11c represents the actual and estimated force of Opt-NARX during the training stage of sample 3, which reached 0.99992 for the CC between these two signals. Also, the CC between the actual and estimated force of Opt-NARX during the training stage of sample 4 is 0.99998.

Fig. 12 shows the force's actual and predicted one-step estimation using Opt-NARX for 4 samples during the testing process. According to Table VII, the prediction power of Opt-NARX is much higher than that of the other three methods (RNN, SVM, and DT). Then, the estimated force of repaired aluminium using RNN, SVM and DT is not mentioned in Fig. 12. Opt-RNN cannot reach satisfactory results for sample 3 regarding the represented results in Table VII. In contrast, the results

of Opt-NARX for sample 3 are pretty convincing, according to Fig. 12c. The RMSE between the actual and estimated force for samples 1 to 4 using Opt-NARX decreased 97.91%, 98.23%, 92.09% and 98.84% compared with Opt-RNN, respectively. It is possible because of the higher capability of the Opt-NARX model in handling complex signals.

Fig. 13 shows the regression of the proposed Opt-NARX during the training and testing stages. The vertical and horizontal axes in Fig. 13 represent the actual and predicted one-step-ahead force based on the previous extension of the specimens. The R-square of Opt-NARX during all datasets (training and testing) stage of 4 samples are 0.99929, 0.99998, 0.99998 and 0.99996.

To fairly evaluate the effectiveness of the proposed Opt-NARX, the completely unseen dataset should be implemented in the network. Then, the new sample, known as sample 5, is produced and tested based on the explained method in Section II. Then, the specimens from sample 5 are extracted, and a destructive tensile test is implemented to extract the force-extension behaviour of the new sample, as shown in Fig. 14 a. The network has not been developed using this dataset, and it is an entirely new dataset for the proposed Opt-NARX. In the next step, the elastic part of the dataset, shown as the training dataset in Fig. 14 b, is provided to the Opt-NARX to start calculating the behaviour of the specimens during the plastic deformation. Fig. 14 b shows that during the testing process (plastic deformation part of the specimens), the MSE, RMSE, CC, and R-square between the experimentally generated and predicted by Opt-NARX are 4.8327×10^{-3} , 6.9517×10^{-2} , 0.95845, and 0.75953, respectively. The extracted accuracy should be below the training and development process of the network in the previous part of this Section. However, the accuracy of the network in calculating the plastic deformation behaviour of the sample is still high when calculating the different mechanical properties of the final product.

To fairly evaluate the effectiveness of the proposed Opt-NARX, the completely unseen dataset should be implemented in the network. Then, the new sample, known as sample 5, is produced and tested based on the explained method in Section II. Then, the specimens from sample 5 are extracted, and a destructive tensile test is implemented to extract the force-extension behaviour of the new sample, as shown in Fig. 14 a. The network has not been developed using this dataset, and it is an entirely new dataset for the proposed Opt-NARX. In the next step, the elastic part of the dataset, shown as the training dataset in Fig. 14 b, is provided to the Opt-NARX to start calculating the behaviour of the specimens during the plastic deformation. Fig. 14 b shows that during the testing process (plastic deformation part of the specimens), the MSE, RMSE, CC, and R-square between

the experimentally generated and predicted by Opt-NARX are 4.8327×10^{-3} , 6.9517×10^{-2} , 0.95845, and 0.75953, respectively. The extracted accuracy should be below the training and development process of the network in the previous part of this Section. However, the accuracy of the network in calculating the plastic deformation behaviour of the sample is still high when calculating the different mechanical properties of the final product.

The represented results in this Section prove the efficiency of our newly proposed Opt-NARX compared with traditional prediction methods (DT and SVR) and Opt-RNN in terms of higher CC and R-square and lower MSE, RMSE, and NRMSE.

V. CONCLUSION

This paper studied a new time-series neural network model to estimate the force of repaired aluminium plates via a multilayer FSPW process. The global optimisation technique (Bayesian method) extracted the optimal hyperparameters of the time-series neural network model. The RNN and NARX were the best methods because they could handle complex time-series nonlinear data. Two traditional prediction models, SVR and DT, were studied to validate our proposed method. The investigated methods are trained using 70% of the tensile test gathered data and have been tested using the rest 30% of the dataset. The proposed models were validated using the force-extension plot of the samples calculated via tensile test. The forecasting power of the proposed models (Opt-RNN and Opt-NARX) was specified using different validation indexes, including MSE, RMSE, NRMSE, R-square, and CC. The four investigated techniques were modelled in MATLAB, and the outcomes were then calculated using the division of data for training (70%) and testing (30%). The extracted results prove that the Opt-NARX can achieve better results during training and testing. The proposed method can benefit the operator in setting the exact control indexes on the aluminium plate's multilayer friction stir plug welding process. Therefore, the quality of repaired aluminium plates can be improved, and plate production will be economical. For future works, an ensemble method and evolutionary optimisation can be considered to increase the accuracy of the prediction model [62, 63].

DECLARATIONS

Conflict of interest: Authors are not in any conflict of interest.

REFERENCES

- [1] S. Dmytro, "The study of welding requirements during construction and installation of seismic-resistant steel structures," *Journal of Research in Science, Engineering and Technology*, vol. 8, no. 2, pp. 17-20, 2020.
- [2] S. Sattari, H. Bisadi, and M. Sajed, "Mechanical properties and temperature distributions of thin friction stir welded sheets of AA5083," *International Journal of Mechanics and Applications*, vol. 2, no. 1, pp. 1-6, 2012.
- [3] M. Sajed and H. Bisadi, "Experimental failure study of friction stir spot welded similar and dissimilar aluminum alloys," *Welding in the World*, vol. 60, no. 1, pp. 33-40, 2016.
- [4] D. T. Buzzatti *et al.*, "Toughness properties of a friction hydro pillar processed offshore mooring chain steel," *Journal of Materials Research and Technology*, vol. 8, no. 3, pp. 2625-2637, 2019.
- [5] R. Landell, L. F. Kanan, D. Buzzatti, B. Vicharapu, A. De, and T. Clarke, "Material flow during friction hydro-pillar processing," *Science and Technology of Welding and Joining*, vol. 25, no. 3, pp. 228-234, 2020.
- [6] W. Li, D. Jiang, L. Yang, J. Pan, R. Qu, and T. Sun, "Numerical simulation of temperature field and prediction of microstructure in friction hydro pillar processing," *Journal of Materials Processing Technology*, vol. 252, pp. 370-380, 2018.
- [7] Y. Xu, H. Jing, Q. Jia, Y. Han, and L. Xu, "Microstructures and mechanical properties of friction hydro-pillar processing overlap welding in API 5L X65 pipeline steel," *Welding in the World*, vol. 62, no. 2, pp. 325-338, 2018.
- [8] M. Sajed and S. H. Seyedkashi, "Solid-state local micro-alloying of thick st37 steel plates with SiC powder using a modified friction hydro-pillar process," *Journal of Materials Research and Technology*, vol. 9, no. 4, pp. 7158-7171, 2020.
- [9] N. G. T. M.W. Mahoney, W.H. Bingel, F.E. Long, R.A. Spurling, G.S. Steele, "Method to repair voids in aluminum alloys," Unitate State of America, 1999.
- [10] V. S. Gadakh and K. Adepu, "Heat generation model for taper cylindrical pin profile in FSW," *Journal of Materials Research and Technology*, vol. 2, no. 4, pp. 370-375, 2013.
- [11] M. Sajed and S. H. Seyedkashi, "Multilayer friction stir plug welding: A novel solid-state method to repair cracks and voids in thick aluminum plates," *CIRP Journal of Manufacturing Science and Technology*, vol. 31, pp. 467-477, 2020.
- [12] M. Sajed, "Parametric study of two-stage refilled friction stir spot welding," *Journal of Manufacturing Processes*, vol. 24, pp. 307-317, 2016.
- [13] B. Aslam, A. Zafar, and U. Khalil, "Development of integrated deep learning and machine learning algorithm for the assessment of landslide hazard potential," *Soft Computing*, vol. 25, no. 21, pp. 13493-13512, 2021.
- [14] R. Tansuchat, U. Pham, and C. Van Le, "On soft computing with random fuzzy sets in econometrics and machine learning," *Soft Computing*, vol. 25, no. 12, pp. 7745-7751, 2021.
- [15] S. Gino Sophia, V. Ceronmani Sharmila, S. Suchitra, T. Sudalai Muthu, and B. Pavithra, "Water management using genetic algorithm-based machine learning," *Soft computing*, vol. 24, no. 22, pp. 17153-17165, 2020.
- [16] A. Rodan, A. F. Sheta, and H. Faris, "Bidirectional reservoir networks trained using SVM \$\$+ \$\$+ privileged information for manufacturing process modeling," *Soft computing*, vol. 21, no. 22, pp. 6811-6824, 2017.
- [17] R.-C. Hwang, Y.-J. Chen, and H.-C. Huang, "Artificial intelligent analyser for mechanical properties of rolled steel bar by using neural networks," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3136-3139, 2010.
- [18] N. F. Alkayem, B. Parida, and S. Pal, "Optimisation of friction stir welding process parameters using soft computing techniques," *Soft Computing*, vol. 21, no. 23, pp. 7083-7098, 2017.
- [19] G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, "Accelerating materials property predictions using machine learning," *Scientific reports*, vol. 3, no. 1, pp. 1-6, 2013.
- [20] M. Rupp, R. Ramakrishnan, and O. A. Von Lilienfeld, "Machine learning for quantum mechanical properties of atoms in molecules," *The Journal of Physical Chemistry Letters*, vol. 6, no. 16, pp. 3309-3313, 2015.
- [21] S. Ma, Z. Jiang, and W. Liu, "Evaluation of a design property network-based change propagation routing approach for mechanical product development," *Advanced Engineering Informatics*, vol. 30, no. 4, pp. 633-642, 2016.
- [22] C. Karina, P.-j. Chun, and K. Okubo, "Tensile strength prediction of corroded steel plates by using machine learning approach," *Steel Compos. Struct.*, vol. 24, no. 5, pp. 635-641, 2017.
- [23] J. D. Evans and F.-X. Coudert, "Predicting the mechanical properties of zeolite frameworks by machine learning," *Chemistry of Materials*, vol. 29, no. 18, pp. 7833-7839, 2017.
- [24] S. Altarazi, R. Allaf, and F. Alhindawi, "Machine learning models for predicting and classifying the tensile strength of polymeric films fabricated via different production processes," *Materials*, vol. 12, no. 9, p. 1475, 2019.
- [25] S. Rakshie, S. Nimje, and S. Panigrahi, "Optimisation of Adhesively Bonded Spar-Wingskin Joints of Laminated FRP Composites Subjected to Pull-Off Load: A Critical Review," *Progress in Adhesion and Adhesives*, vol. 6, pp. 29-46, 2021.
- [26] Y. Gong*, M. Mehana, I. El-Monier, F. Xu, and F. Xiong, "Machine learning for estimating rock mechanical properties beyond traditional considerations," in *Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July 2019*, 2019: Unconventional Resources Technology Conference (URTeC); Society of ..., pp. 466-480.
- [27] G. Golkarnarenji, M. Naebe, K. Badii, A. S. Milani, R. N. Jazar, and H. Khayyam, "A machine learning case study with limited data for prediction of carbon fiber mechanical properties," *Computers in Industry*, vol. 105, pp. 123-132, 2019.
- [28] Z. Qi, N. Zhang, Y. Liu, and W. Chen, "Prediction of mechanical properties of carbon fiber based on cross-scale FEM and machine learning," *Composite Structures*, vol. 212, pp. 199-206, 2019.
- [29] Z. Zhang, Y. Hong, B. Hou, Z. Zhang, M. Negahban, and J. Zhang, "Accelerated discoveries of mechanical properties of graphene using machine learning and high-throughput computation," *Carbon*, vol. 148, pp. 115-123, 2019.
- [30] X. Wang, D. Han, Y. Hong, H. Sun, J. Zhang, and J. Zhang, "Machine learning enabled prediction of mechanical properties of tungsten disulfide monolayer," *ACS omega*, vol. 4, no. 6, pp. 10121-10128, 2019.
- [31] X. Xu, L. Wang, G. Zhu, and X. Zeng, "Predicting tensile properties of AZ31 magnesium alloys by machine learning," *Jom*, vol. 72, no. 11, pp. 3935-3942, 2020.
- [32] A. Behnood and E. M. Golafshani, "Machine learning study of the mechanical properties of concretes containing waste foundry sand," *Construction and Building Materials*, vol. 243, p. 118152, 2020.

- [33] W. B. Chaabene, M. Flah, and M. L. Nehdi, "Machine learning prediction of mechanical properties of concrete: Critical review," *Construction and Building Materials*, vol. 260, p. 119889, 2020.
- [34] J. Xiong, T. Zhang, and S. Shi, "Machine learning of mechanical properties of steels," *Science China Technological Sciences*, vol. 63, pp. 1247-1255, 2020.
- [35] H. Fathi, V. Nasir, and S. Kazemirad, "Prediction of the mechanical properties of wood using guided wave propagation and machine learning," *Construction and Building Materials*, vol. 262, p. 120848, 2020.
- [36] Z.-h. Deng *et al.*, "Machine-learning-assisted prediction of the mechanical properties of Cu-Al alloy," *International Journal of Minerals, Metallurgy and Materials*, vol. 27, no. 3, pp. 362-373, 2020.
- [37] O. Mamun, M. Wenzlick, J. Hawk, and R. Devanathan, "A machine learning aided interpretable model for rupture strength prediction in Fe-based martensitic and austenitic alloys," *Scientific Reports*, vol. 11, no. 1, pp. 1-9, 2021.
- [38] Q. Xie, M. Suvarna, J. Li, X. Zhu, J. Cai, and X. Wang, "Online prediction of mechanical properties of hot rolled steel plate using machine learning," *Materials & Design*, vol. 197, p. 109201, 2021.
- [39] K. Güçlüer, A. Özbeyaz, S. Göymen, and O. Günaydin, "A comparative investigation using machine learning methods for concrete compressive strength estimation," *Materials Today Communications*, vol. 27, p. 102278, 2021.
- [40] D. Ferreño *et al.*, "Prediction of mechanical properties of rail pads under in-service conditions through machine learning algorithms," *Advances in Engineering Software*, vol. 151, p. 102927, 2021.
- [41] J. Xiong, S.-Q. Shi, and T.-Y. Zhang, "Machine learning of phases and mechanical properties in complex concentrated alloys," *Journal of Materials Science & Technology*, vol. 87, pp. 133-142, 2021.
- [42] A. Nafees *et al.*, "Modeling of Mechanical Properties of Silica Fume-Based Green Concrete Using Machine Learning Techniques," *Polymers*, vol. 14, no. 1, p. 30, 2021.
- [43] A. Stoll and P. Benner, "Machine learning for material characterisation with an application for predicting mechanical properties," *GAMM-Mitteilungen*, vol. 44, no. 1, p. e202100003, 2021.
- [44] M. I. Shah, S. A. Memon, M. S. Khan Niazi, M. N. Amin, F. Aslam, and M. F. Javed, "Machine learning-based modeling with optimisation algorithm for predicting mechanical properties of sustainable concrete," *Advances in Civil Engineering*, vol. 2021, 2021.
- [45] I. Demirci and I. Saritas, "Damage Detection of Carbon Face Sheet Nomex Sandwich Composites with Image Processing Technique," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 9, no. 4, pp. 282-287, 2021.
- [46] J. Huang, J. Liew, and K. Liew, "Data-driven machine learning approach for exploring and assessing mechanical properties of carbon nanotube-reinforced cement composites," *Composite Structures*, vol. 267, p. 113917, 2021.
- [47] H. Dabiri, A. Kheyroddin, and A. Faramarzi, "Predicting tensile strength of spliced and non-spliced steel bars using machine learning-and regression-based methods," *Construction and Building Materials*, vol. 325, p. 126835, 2022.
- [48] K. Hayashi and M. Ohsaki, "Graph-based reinforcement learning for discrete cross-section optimisation of planar steel frames," *Advanced Engineering Informatics*, vol. 51, p. 101512, 2022.
- [49] I. Najjar, A. Sadoun, G. S. Alsorjji, M. Abd Elaziz, and A. Wagih, "Predicting the mechanical properties of Cu-Al₂O₃ nanocomposites using machine learning and finite element simulation of indentation experiments," *Ceramics International*, vol. 48, no. 6, pp. 7748-7758, 2022.
- [50] J. Liu, Y. Zhang, Y. Zhang, S. Kitipornchai, and J. Yang, "Machine learning assisted prediction of mechanical properties of graphene/aluminium nanocomposite based on molecular dynamics simulation," *Materials & Design*, vol. 213, p. 110334, 2022.
- [51] Y. Diao, L. Yan, and K. Gao, "A strategy assisted machine learning to process multi-objective optimisation for improving mechanical properties of carbon steels," *Journal of Materials Science & Technology*, vol. 109, pp. 86-93, 2022.
- [52] S. Verma and J. Misra, "Study on temperature distribution during Friction Stir Welding of 6082 aluminum alloy," *Materials Today: Proceedings*, vol. 4, no. 2, pp. 1350-1356, 2017.
- [53] M. Djurdjanovic, M. Mijajlovic, D. Milcic, and D. Stamenkovic, "Heat generation during friction stir welding process," *Tribology in industry*, vol. 31, no. 1&2, p. 8, 2009.
- [54] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [55] X. Fu, S. Li, M. Fairbank, D. C. Wunsch, and E. Alonso, "Training recurrent neural networks with the Levenberg-Marquardt algorithm for optimal control of a grid-connected converter," *IEEE transactions on neural networks and learning systems*, vol. 26, no. 9, pp. 1900-1912, 2014.
- [56] T. Chow and C.-T. Leung, "Nonlinear autoregressive integrated neural network model for short-term load forecasting," *IEE Proceedings-Generation, Transmission and Distribution*, vol. 143, no. 5, pp. 500-506, 1996.
- [57] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical analysis*: Springer, 1978, pp. 105-116.
- [58] R. N. Longenecker, *Paul, Apostle of Liberty*. Wm. B. Eerdmans Publishing, 2015.
- [59] M. Pelikan, D. E. Goldberg, and E. Cantú-Paz, "BOA: The Bayesian optimisation algorithm," in *Proceedings of the genetic and evolutionary computation conference GECCO-99*, 1999, vol. 1: Citeseer, pp. 525-532.
- [60] Z. Chiba, M. S. E. K. Alaoui, N. Abghour, and K. Moussaid, "Automatic Building of a Powerful IDS for The Cloud Based on Deep Neural Network by Using a Novel Combination of Simulated Annealing Algorithm and Improved Self-Adaptive Genetic Algorithm."
- [61] K. Adane and B. Beyene, "Machine Learning and Deep Learning Based Phishing Websites Detection: The Current Gaps and Next Directions," *Review of Computer Engineering Research*, vol. 9, no. 1, pp. 13-29, 2022.
- [62] M. Ghalandari, A. Ziamolki, A. Mosavi, S. Shamshirband, K.-W. Chau, and S. Bornassi, "Aeromechanical optimisation of first row compressor test stand blades using a hybrid machine learning model of genetic algorithm, artificial neural networks and design of experiments," *Engineering Applications of Computational Fluid Mechanics*, vol. 13, no. 1, pp. 892-904, 2019.
- [63] S. Srinivasareddy, Y. Narayana, and D. Krishna, "Sector beam synthesis in linear antenna arrays using social group optimisation algorithm," *National Journal Of Antennas And Propagation*, vol. 3, no. 2, pp. 6-9, 2021.

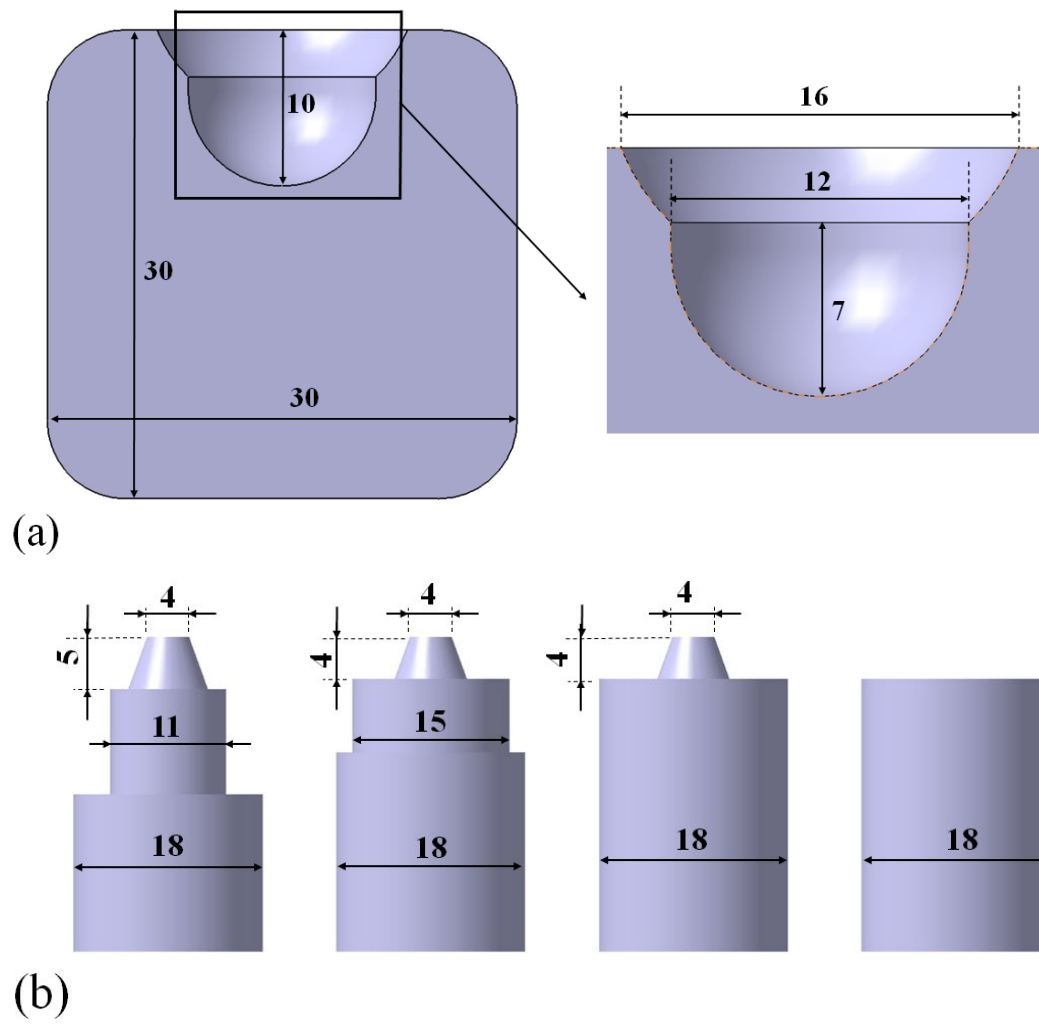


Fig. 1. Dimensions of the (a): hole; (b): welding tools.

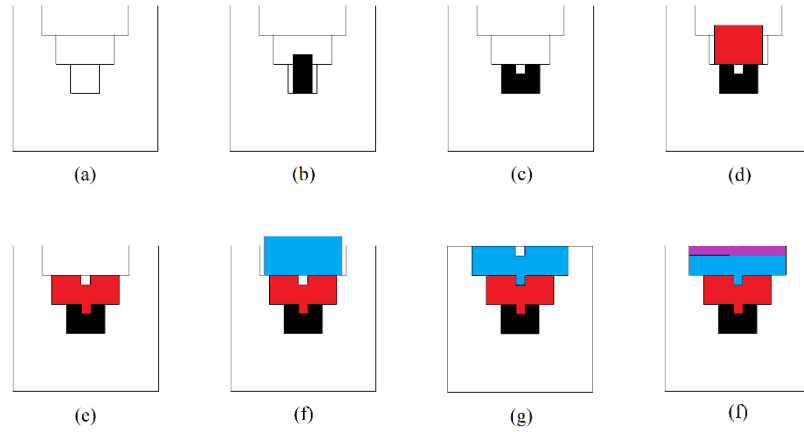


Fig. 2. The schematic explanation of steps of the multilayer friction stir plug welding (MFSPW)

(a): predrilling the hole; (b): injecting the bottom plug; (c): welding the bottom plug; (d): inserting the middle plug; (e): welding the middle plug; (f): injecting the top plug; (g): welding the top plug; (h): refilling the keyhole of the top layer.

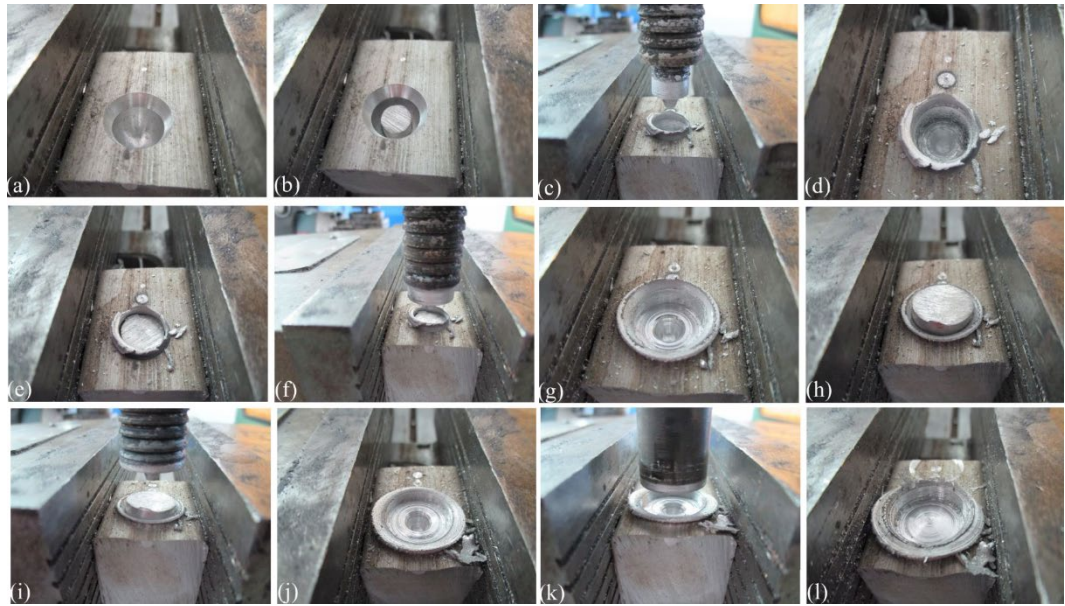


Fig. 3. Multilayer friction stir plug welding process (a) the hole; (b) the 8 (mm) diameter plug inserted into the hole; (c) welding the hole with the 11 (mm) diameter shoulder; (d) the first step welded sample; (e) the 12 (mm) diameter plug inserted to the hole; (f) welding the sample with the 15 (mm) diameter shoulder; (g) the second step welded sample; (h) the 14 (mm) diameter plug inserted to the hole; (i) welding the sample with the 18 (mm) diameter shoulder; (j) the final welded sample; (k) covering the keyhole with the non-pin tool; (l) the final product.

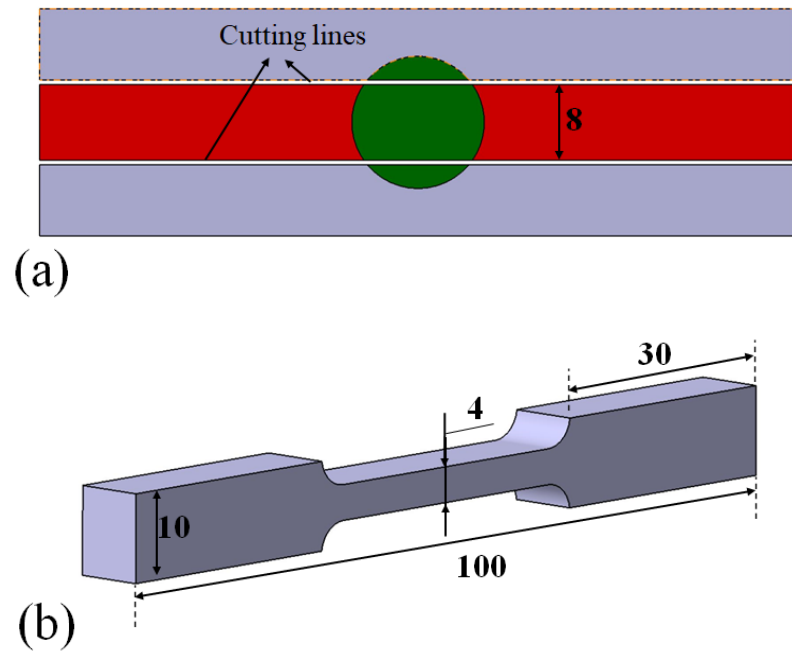


Fig. 4. A schematic sample of the tensile test: (a) the top view of the cutting line; (b) its position and dimensions inside view.

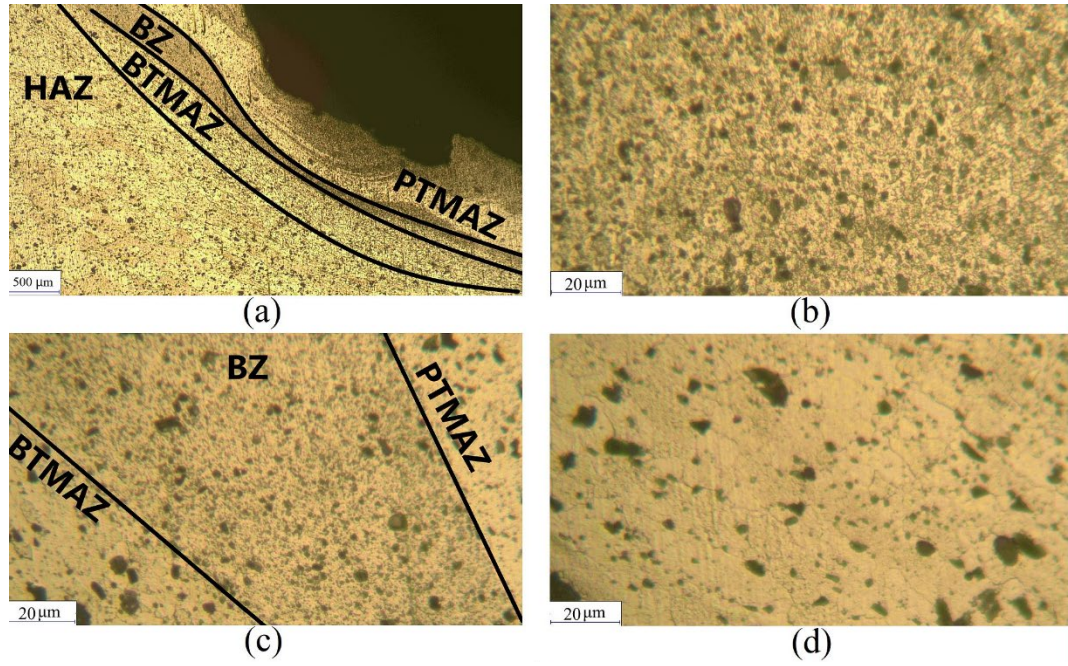


Fig. 5. (a): Different zones of the samples that are welded with the 1000 (rpm) rotational speed and 10 (seconds) dwelling time, including (b): the affected zone of a plug, thermomechanically; (c): the thermomechanical affected zone between the bonding zone and base metal; (d): the zone which is affected by heat.

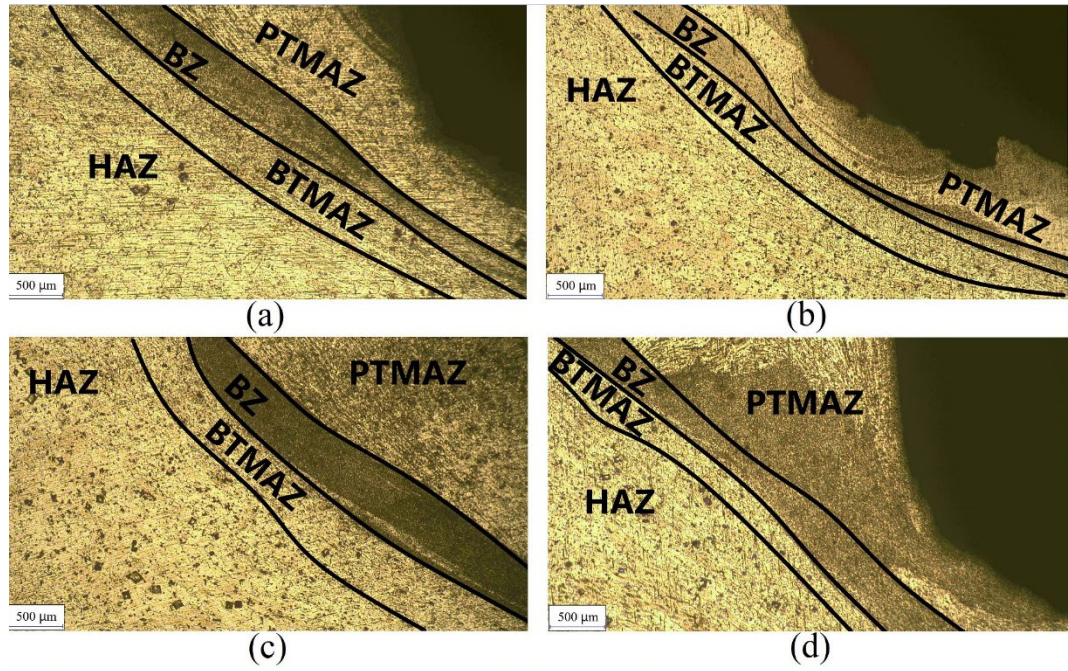


Fig. 6. Different Zones in the welding nugget of the upper plug, including bonding zone (BZ), plug thermomechanical affected zone (PTMAZ), heat-affected zone (HAZ) and base metal thermomechanical affected zone (BTMAZ) (a): the specimen welded with a 1000 (rpm) tool rotational speed, and 5 (seconds) dwelling time; (b): the specimen welded with a 1000 (rpm) tool rotational speed, and 10 (seconds) dwelling time; (c): the specimen welded with a 3000 (rpm) tool rotational speed, and 5 (seconds) dwelling time; (d): the specimen welded with a 3000 (rpm) tool rotational speed, and 10 (seconds) dwelling time.

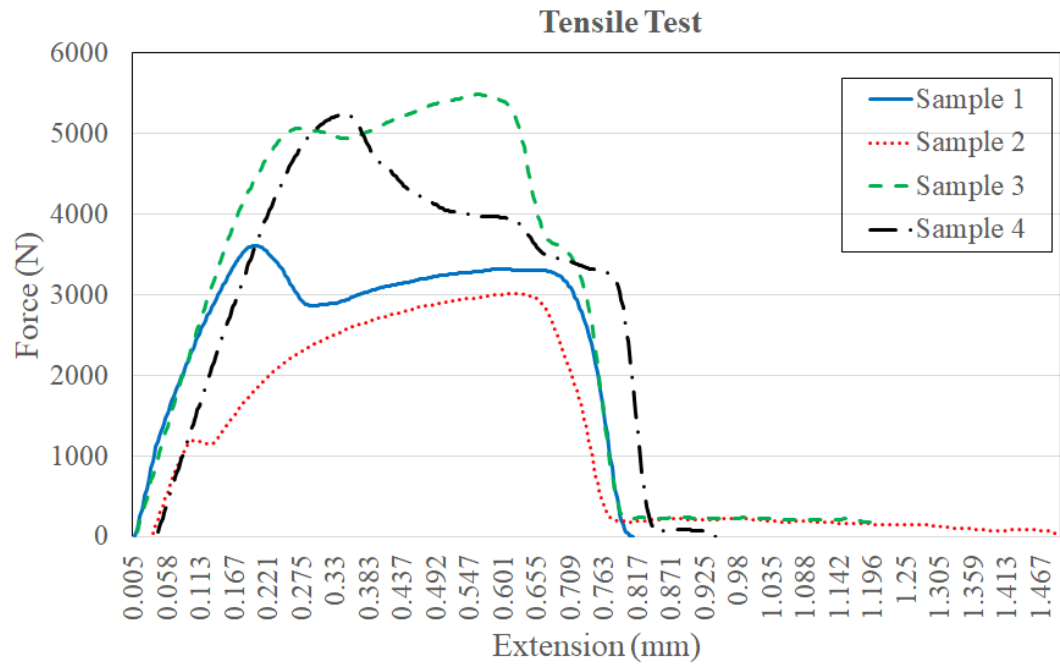


Fig. 7. The obtained dataset from the tensile test is to be used in the investigated machine learning method.

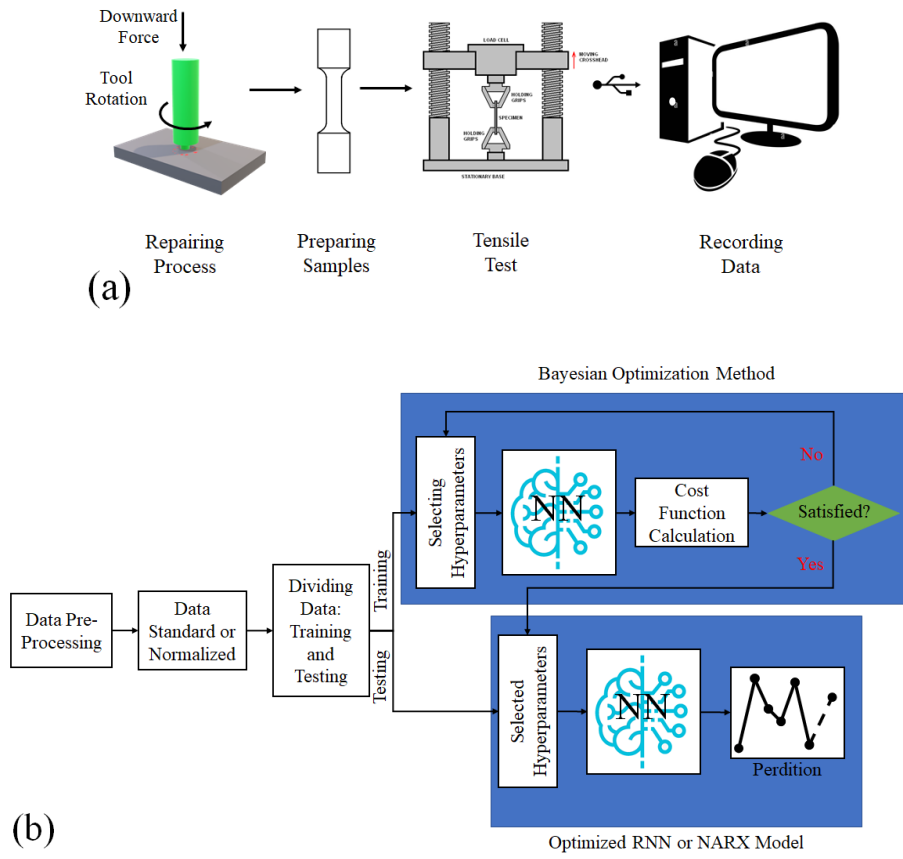
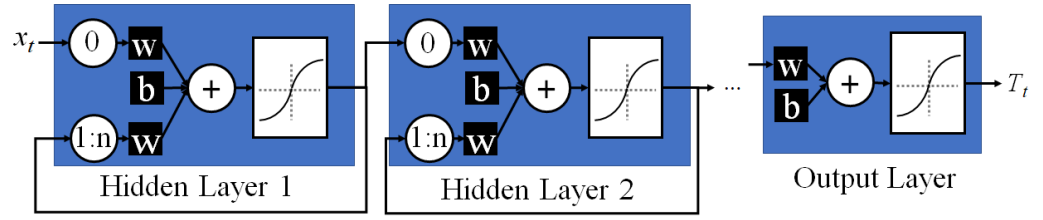
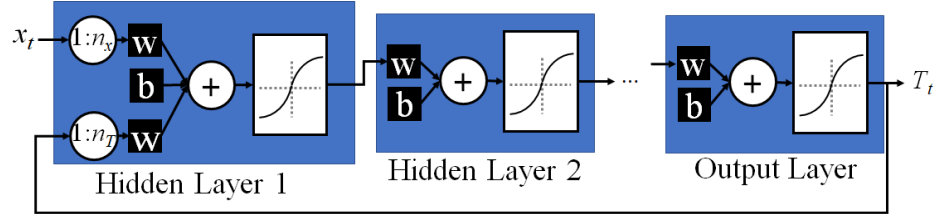


Fig. 8. The obtained dataset from the tensile test is to be used in the investigated machine learning method.



(a)



(b)

Fig. 9. The structure of (a) RNN; (b) NARX.

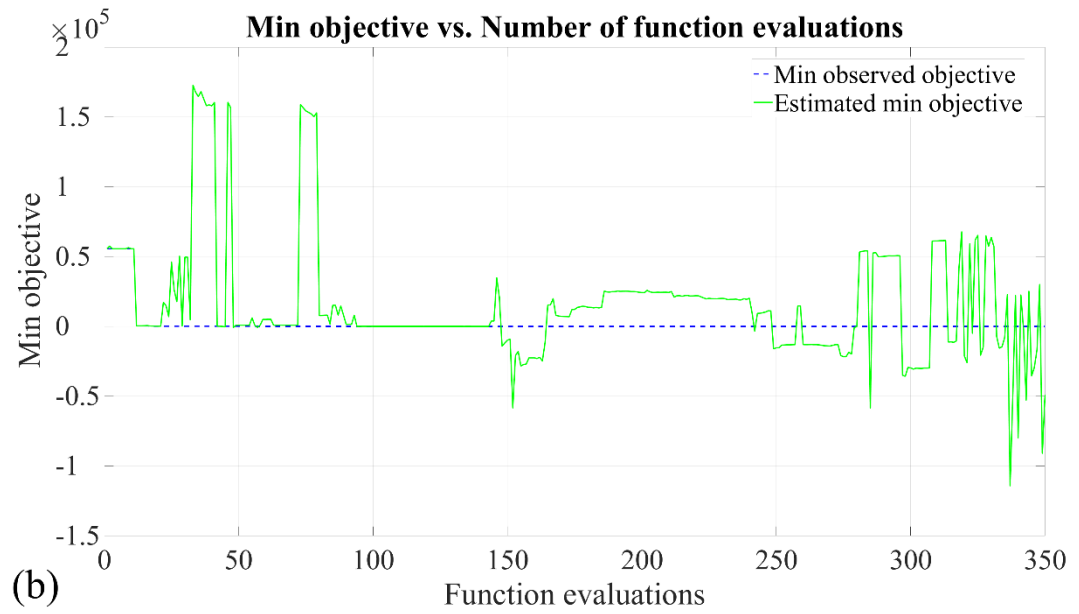
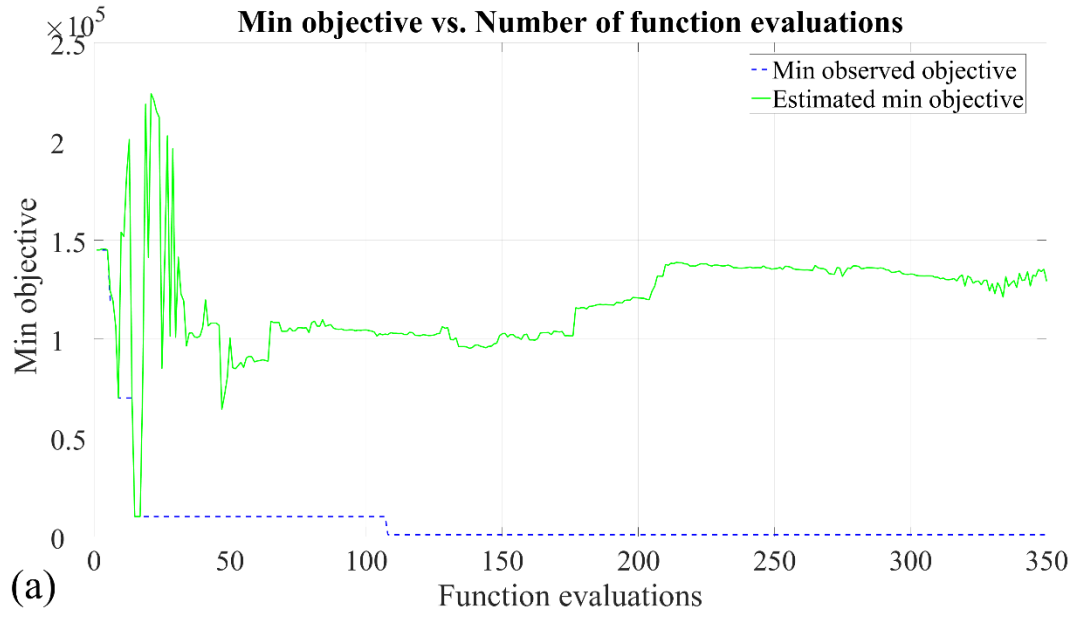


Fig. 10. The obtained dataset from the tensile test is to be used in the investigated machine learning method.

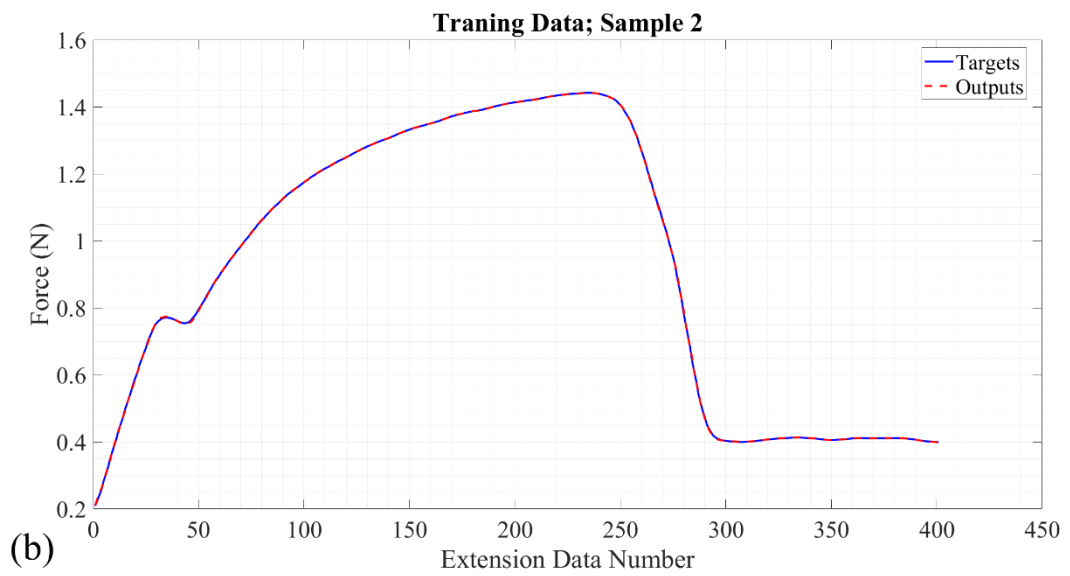
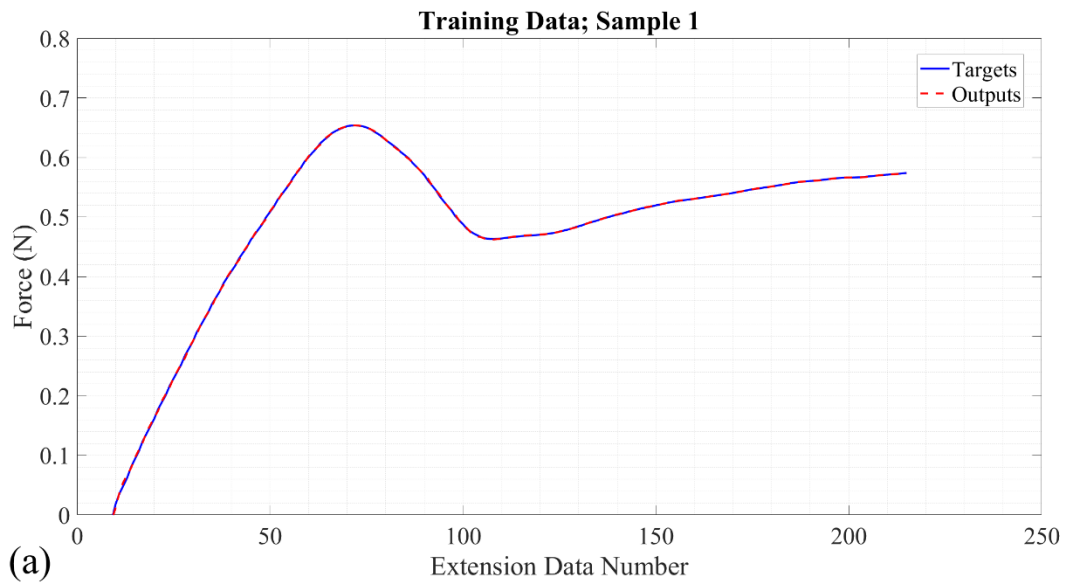
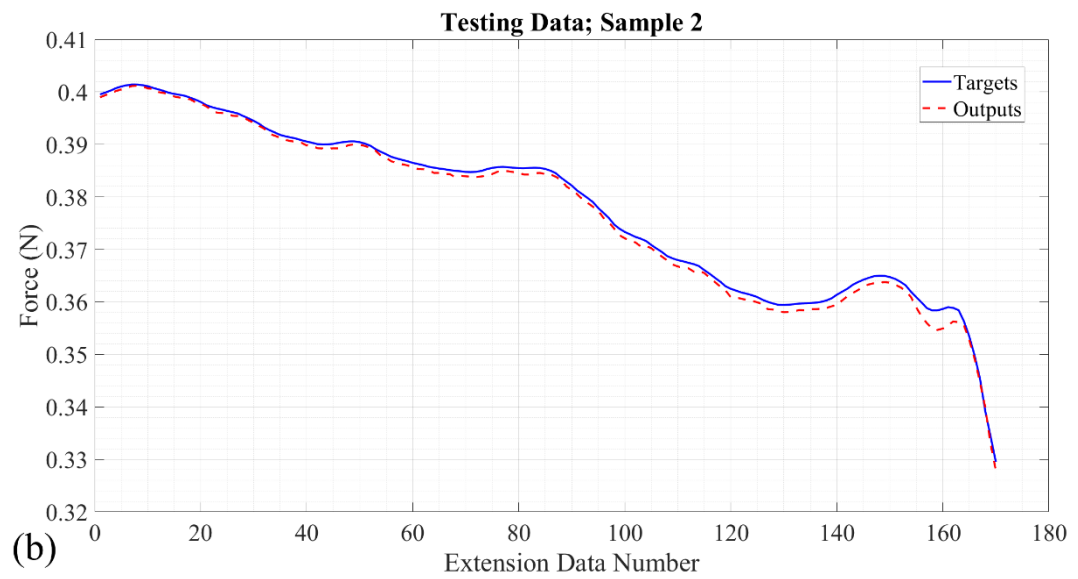
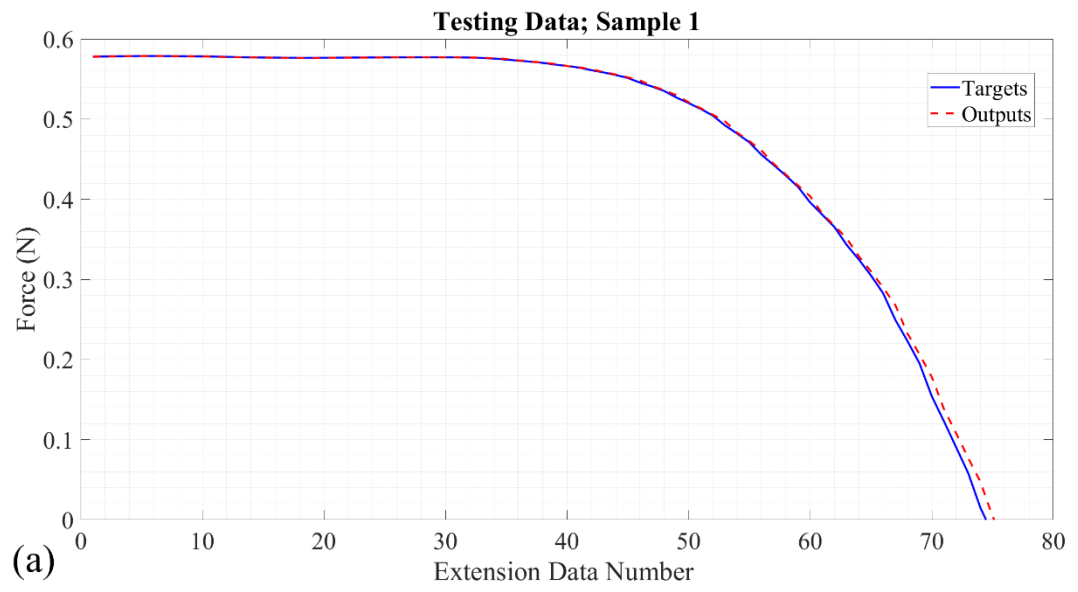




Fig. 11. The target and estimated data during the training process of Opt-NARX for (a) sample 1; (b) sample 2; (c) sample 3; (d) sample 4.



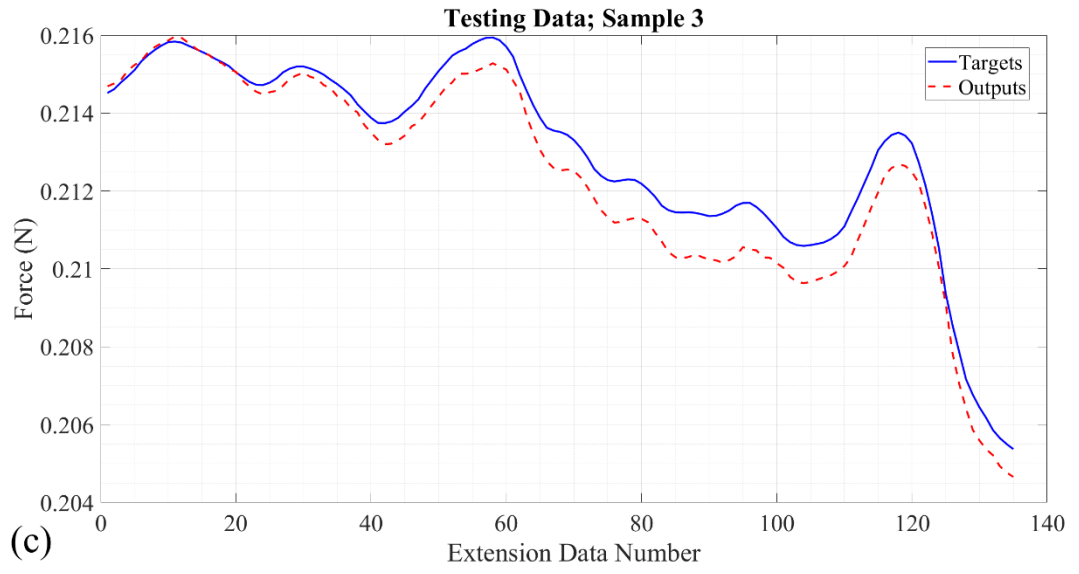


Fig. 12. The target and estimated data during the testing process of Opt-RNN and Opt-NARX for (a): sample 1; (b): sample 2; (c): sample 3; (d): sample 4.

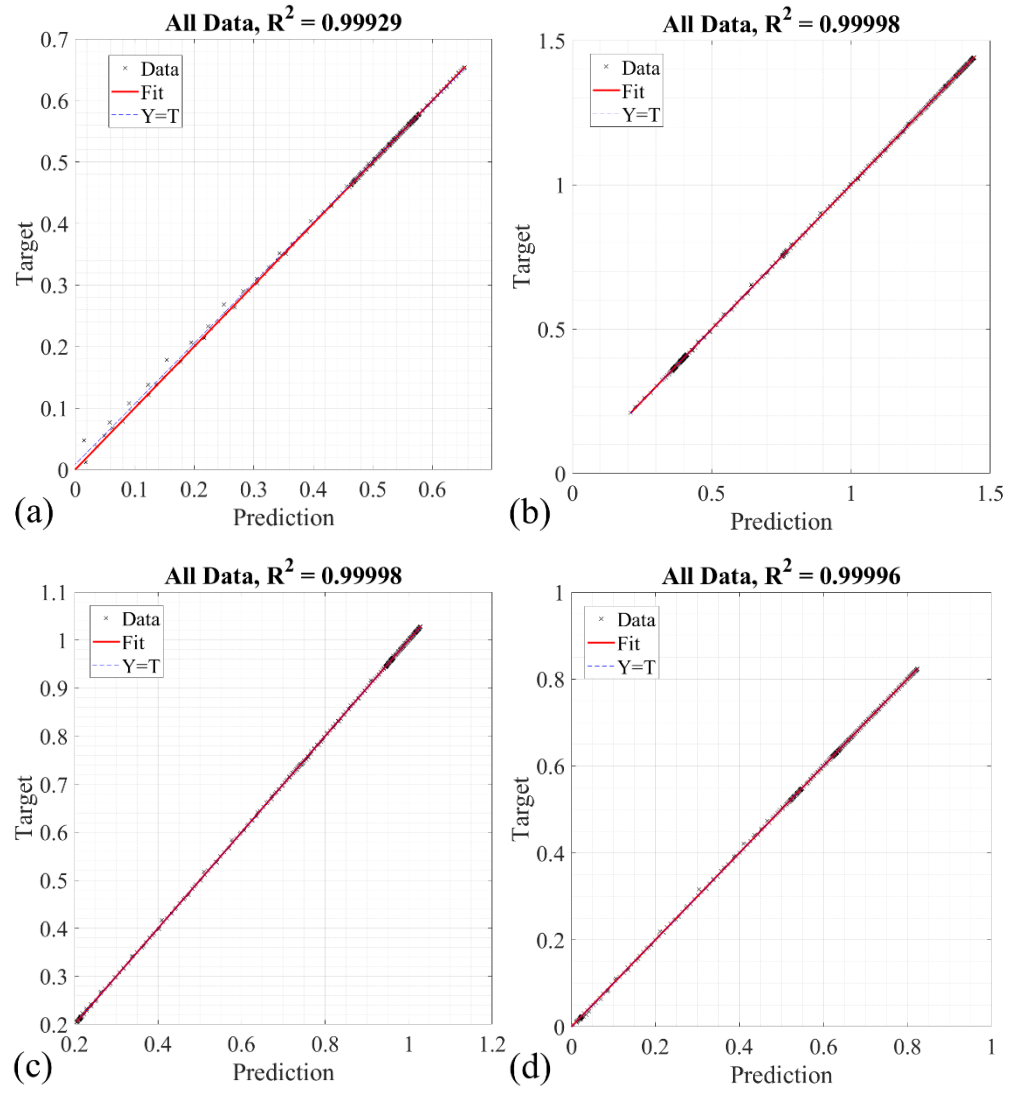
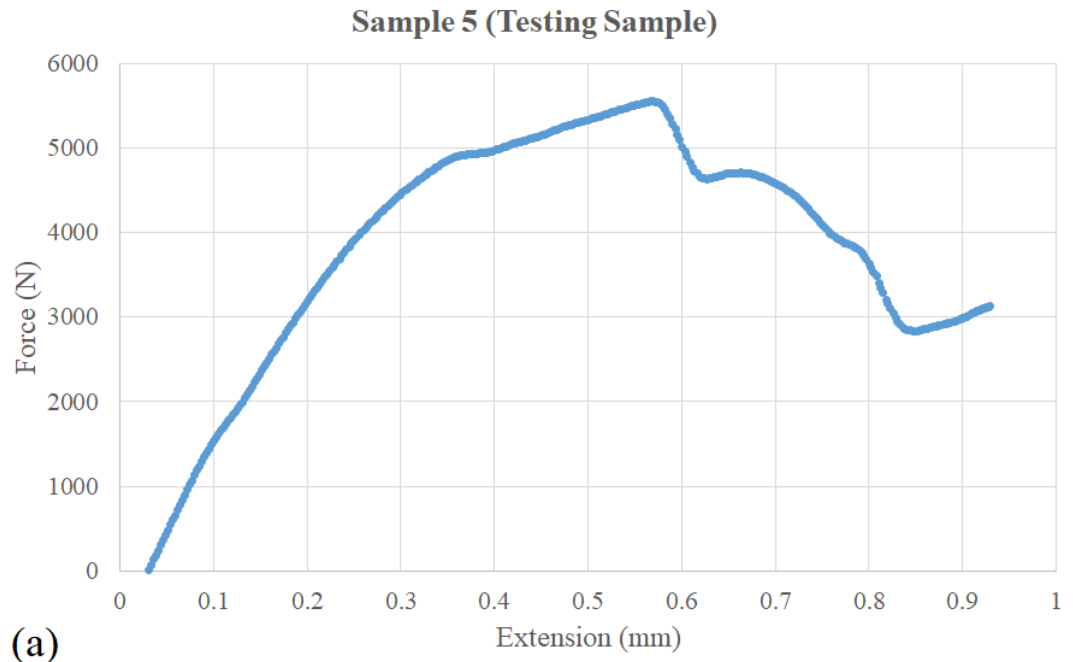
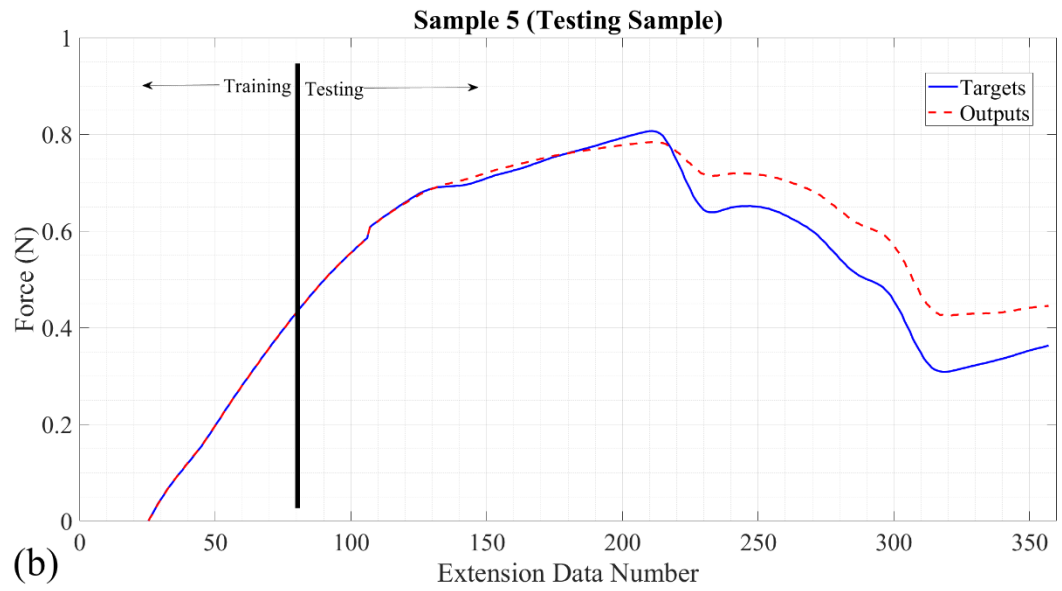


Fig. 13. The regression of the data during the Opt-NARX training and testing process (a) sample 1; (b) sample 2; (c) sample 3; (d) sample 4.



(a)



(b)

Fig. 14. The extracted force extension data for testing sample 5 using (a) actual destructive experiment;

(b) The proposed Opt-NARX.

TABLE I

The literature reviews in the field of mechanical properties prediction using different machine learning methods.

Authors	Year	Machine Learning Methods	Applications
Rupp et al. [20]	2015	Kernel ridge regression and nonlinear regularised regression	Quantum mechanical properties prediction of atoms in molecules
Evans et al. [23]	2017	Gradient boosting regressor	mechanical properties prediction of zeolite frameworks
Gong et al. [26]	2019	Multilayer perceptron, support vector machine, multivariate linear regression and multivariate adaptive regression spline	Estimating rock mechanical properties
Golkarnarenji et al. [27]	2019	Support vector machine and multilayer perceptron.	Prediction of carbon fibre mechanical properties
Qi et al. [28]	2019	Decision Tree	Prediction of mechanical properties of carbon fiber
Zhang et al. [29]	2019	k-nearest neighbours, support vector machine, decision tree and multilayer perceptron.	Mechanical properties prediction of graphene
Wang et al. [30]	2019	Random forest.	Prediction of mechanical properties of tungsten disulphide monolayer
Behnood and Golafshani [32]	2020	Gradient-boosted tree regression model	mechanical properties prediction of concretes containing waste foundry sand
Chaabene et al. [33]	2020	Multilayer perceptron, support vector machine, decision tree for prediction. Evolutionary algorithms for optimised value selection of operation	Prediction of mechanical properties of concrete
Xiong et al. [34]	2020	Random forest, k-nearest neighbours, multilayer perceptron and linear least-square.	Mechanical properties prediction of steel
Fathi et al. [35]	2020	Group method of data handling	Mechanical properties prediction of wood
Deng et al. [36]	2020	standard linear regression, multilayer perceptron, sequential minimal optimisation algorithm for support vector regression, puk kernel and support vector machine	Prediction of the mechanical properties of Cu-Al alloy

Güçlüer et al. [39]	2021	Multilayer perceptron, support vector machine, decision tree and linear regression	concrete compressive strength estimation
Ferreño et al. [40]	2021	Multilinear regression, K nearest neighbours, regression tree, random forest, gradient boosting, Multilayer perceptron and support vector machine	Prediction of mechanical properties of rail pads under in-service conditions
Xie et al. [38]	2021	Multilayer perceptron	Prediction of mechanical properties of hot rolled steel plate
Xionget al. [41]	2021	Random forest regression	Mechanical properties prediction in complex concentrated alloys
Nafees et al. [42]	2021	Decision tree and support vector machine	Mechanical Properties of Silica Fume-Based Green Concrete
Stoll et al. [43]	2021	Linear regression model and random forest	Predicting mechanical properties of different materials
Shah et al. [44]	2021	Optimised multi-expression programming based on particle swarm optimisation	Predicting mechanical properties of sustainable concrete
Huang et al. [46]	2021	Support vector machine and multilayer perceptron	mechanical properties prediction of carbon nanotube-reinforced cement composite
Najjar et al. [49]	2022	random vector functional link network and red fox algorithm	Predicting the mechanical properties of Cu–Al ₂ O ₃ nanocomposites
Liu et al. [50]	2022	Support vector machine and multilayer perceptron	Prediction of mechanical properties of graphene/aluminium nanocomposite
Diao et al. [51]	2022	support vector regression with radial basis function kernel, multilayer perceptron, k-nearest neighbour, kernel ridge regression, and random forest.	Mechanical properties prediction of carbon steels

TABLE II

Chemical composition of the welded Aluminium alloy (WT. %)

Al	Ni	Cr	Sb	Pb	Ti	Zn	Mg	Mn	Cu	Si	Fe
Balance	0.0247	0.0403	0.0209	0.291	0.0235	2.9	1.55	0.593	1.17	0.683	2.01

TABLE III

The experiment setup.

No.	Tool rotational speed (rpm)	Dwell time (seconds)	Average Joint strength (MPa)
First	1000	5	94
Second	1000	10	67
Third	3000	5	124
Fourth	3000	10	114

TABLE IV

The tensile test setup for 4 samples.

No.	Thickness (mm)	Width (mm)	Gage Length (mm)	Test Speed (mm/min)
1	6.1	6.3	50	10
2	6.1	7.4	50	10
3	6.2	7.15	50	10
4	5.8	7.93	50	10

TABLE V

The extracted hyperparameters of RNN and NARX using the Bayesian optimisation method for 4 samples.

Method and No.		Number of Layers	Number of neurons	Feedback Step	Feedback Delay
RNN	1	2	1	4	20
	2	2	3	2	10
	3	1	1	5	11
	4	1	3	5	14
NARX	1	1	2	3	4
	2	1	4	1	20
	3	3	8	2	2
	4	1	1	3	6

TABLE VI

Using different sets of hyperparameters, the RMSE between the predicted and actual tensile of the repaired products via FSW for sample 1.

Number of Layers	Number of Neurons	Sequencing	RMSE
1	2	1-4	2.0229×10^{-5}
1	4	1-4	4.9901×10^{-5}
1	2	1-2-3-4	1.3261×10^{-5}
1	4	1-2-3-4	3.0849×10^{-3}
3	2	1-4	2.5087×10^{-1}
3	4	1-4	5.202×10^{-2}
3	2	1-2-3-4	5.9179×10^{-2}
3	4	1-2-3-4	2.2637×10^{-2}

TABLE VII

The extracted results of Opt-RNN and Opt-NARX using the Bayesian optimisation method for 4 samples.

Method	Sample No	Process	MSE	RMSE	NRMSE	CC	R ²
DT	1	Training	8.3956×10^{-4}	2.8975×10^{-2}	6.1982×10^{-2}	0.99759	0.97169
	1	Testing	0.1237	0.3517	0.93668	0.5	-2.5005×10^{31}
	2	Training	3.0907×10^{-4}	1.758×10^{-2}	1.9103×10^{-2}	0.98478	0.99823
	2	Testing	1.1147×10^{-3}	3.3387×10^{-2}	8.8224×10^{-2}	0.5	-9.4752×10^{28}
	3	Training	2.2855×10^{-4}	1.5118×10^{-2}	1.9619×10^{-2}	0.99506	0.99677
	3	Testing	1.8907×10^{-5}	4.3482×10^{-3}	2.0412×10^{-2}	0.5	-5.1524×10^{26}
	4	Training	3.358×10^{-4}	1.8325×10^{-2}	3.2431×10^{-2}	0.99424	0.99351
	4	Testing	0.15289	0.39102	1.3802	0.5	-5.9561×10^{30}
SVR	1	Training	370752.6585	608.8946	0.21116	0.50593	-6.9638
	1	Testing	2145698.115	1464.8202	0.58317	-0.97292	-40.11331
	2	Training	1545645.6304	1243.24	0.77687	0.30014	-0.094995
	2	Testing	1677868.4299	1295.3256	9.9856	0.97848	0.65177
	3	Training	3790037.4917	1946.8019	0.64558	0.51343	0.13954
	3	Testing	10133085.4295	3183.2508	15.6874	0.41782	0.09323
	4	Training	3553396.4984	1885.0455	0.52753	-0.38889	-11.6135
	4	Testing	5677840.9437	2382.8221	1.3596	0.96869	-13.6676
RNN	1	Training	2.0568×10^{-4}	1.4341×10^{-2}	2.8657×10^{-2}	0.95829	0.98453
	1	Testing	0.24897	0.49897	1.4965	-0.79747	-4.6443
	2	Training	5.9574×10^{-5}	7.7184×10^{-3}	8.2863×10^{-3}	0.99181	0.99965
	2	Testing	4.1927×10^{-3}	6.4751×10^{-2}	0.17152	0.97664	0.15326
	3	Training	5.8541×10^{-3}	7.6512×10^{-2}	9.76×10^{-2}	0.77761	0.90261
	3	Testing	8.6848×10^{-3}	9.3192×10^{-2}	0.43776	0.85822	0.028548
	4	Training	1.378×10^{-5}	3.7121×10^{-3}	6.375×10^{-3}	0.99936	0.99968
	4	Testing	6.149×10^{-2}	0.24797	0.95661	0.96031	0.28824
NARX	1	Training	1.8925×10^{-6}	1.3757×10^{-3}	2.9459×10^{-3}	0.99989	0.99994
	1	Testing	1.0927×10^{-4}	1.0453×10^{-2}	2.8345×10^{-2}	0.99924	0.99866
	2	Training	3.6956×10^{-6}	1.9224×10^{-3}	2.083×10^{-3}	0.99971	0.99998
	2	Testing	1.319×10^{-6}	1.1485×10^{-3}	3.0367×10^{-3}	0.99925	0.99845
	3	Training	2.4414×10^{-6}	1.5625×10^{-3}	2.0231×10^{-3}	0.99992	0.99997
	3	Testing	5.4335×10^{-7}	7.3712×10^{-3}	3.4608×10^{-3}	0.98698	0.97853
	4	Training	9.4275×10^{-7}	9.7095×10^{-4}	1.7181×10^{-3}	0.99998	0.99998
	4	Testing	8.338×10^{-6}	2.8876×10^{-3}	1.0378×10^{-2}	0.9978	0.99986

