www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

ARCHIVE OF MECHANICAL ENGINEERING

DOI: 10.24425/ame.2026.158899 Early Access; pp. 1–23

Arnaud BLAISE [1], Guilhem BLES [2], Ali TOURABI [1]

Recognition of mechanical behaviour laws through the supervised learning of an artificial neural network

Received 9 October 2025, **Revised** 20 January 2026, **Accepted** 6 February 2026, **Published online** 20 April 2026

Keywords: nonlinear elasticity, polyamide, yarns, artificial intelligence, neural networks

In this paper, we put forward a classifier to automatically identify the most suitable mechanical behaviour law among two candidates, enabling optimal modelling of experimental data from a uniaxial tensile test, represented as elongation–specific stress curves. An Artificial Neural Network (ANN) is employed to perform this classification task and assist the modeler, even when the resulting curves from different models are very close. This paper compares two different methods that enable supervised learning of the neural network from a training dataset labelled with specific stress values as a function of elongation, without requiring any external data. The learning process is then validated by testing the network's ability to correctly identify the most suitable model from experimental data it has never encountered before. This approach could potentially pave the way for adaptations to more complex tasks in a multidimensional context or to feature recognition in imaging, in the frame of various fields of materials mechanics.

1. Introduction

Artificial neural networks (ANN) are versatile and powerful tools that can be used in a wide range of applications to solve complex classification and prediction problems using learning (supervised, unsupervised or by reinforcement) from a certain amount of data. Since the early 21st century, significant progress has been made in artificial intelligence, notably with the introduction of Recurrent Neural Networks (RNN), convolutional neural networks (CNN) and physics-informed neural networks (PINN). Concerning engineering sciences applied to the mechanics

✉ Arnaud BLAISE, e-mail: arnaud.blaise@univ-grenoble-alpes.fr

[1]Grenoble-INP, Université Grenoble Alpes, France

[2]ENSTA, Institut Polytechnique de Paris, UMR CNRS 6027, IRDL, F-29200 Brest, France

of materials and industrial mechanical systems, the use of ANN, CNN, RNN and other more sophisticated neural networks such as TANN (Thermodynamics-based Artificial Neural Networks) and RTAN (Recurrent Thermodynamical Artificial Network) have proven promising for tackling certain challenges focused mainly on the maintenance and/or performance of mechanical or industrial systems as well as the prediction of the mechanical behaviour and structural characterization of complex materials. Table 1 provides a bibliography bringing together the different aspects described above.

Table 1. Bibliography in the field of engineering sciences applied to mechanics of material and industrial mechanical systems, using neural networks such as ANN, CNN, RNN, TANN and RTAN

Neural network type	Objective		
	Maintenance (fault diagnosis, damage detection, . . .)	Prediction (structural characterization, mechanical response, mechanical behaviour, thermodynamical approaches, . . .)	Performance by optimization of parameters (efficiency, lifetime, reduction of undesirable effects, control and sharing knowledges between devices, . . .)
ANN	[1–3]	[4–7]	[8–10]
CNN	[11–14]	[15, 16]	
RNN		[17, 18]	
PINN		[19–21]	
TANN		[22]	
RTAN		[23]	

In this paper, we propose two simple and effective models recognition methods intended for people of the mechanics community, who have little knowledge of artificial intelligence. The simplicity of the method is a key consideration in our approach, aimed to facilitate its implementation and make it accessible to users. In this framework, we propose methods using classification algorithms whose core is an Artificial Neural Network (ANN), with the aim of recognizing one mechanical model among two, best accounting for the mechanical behaviour of a material during a uniaxial tensile test. This classification task is done independently of the number of parameters of the optimal model that best describes the experimental data. Thus, modelling accuracy is the only parameter considered for the choice of the optimal model, without optimizing the number of parameters. We explore, within the frame of supervised learning, its effectiveness in being trained on labelled data and then, we evaluate its recognition capacity on external data that it has never seen before. The relevance of the obtained results is validated with a simple least squares regression method. In this paper, we set up a method to recognize the good mechanical model among two proposed in the case of modelling the behaviour of a PA66 polyamide yarn during a uniaxial tensile test.

2. Materials and methods

2.1. Material mechanical behaviour modelling

Polyamide, commonly known as Nylon, is a thermoplastic polymer with a semi-crystalline structure and high thermal stability. This material is very useful due to its high flexibility, high resistance to wear, low coefficient of friction and very high temperature and impact resistance properties [24]. In addition, this material is quite inexpensive. We define the Cauchy stress in a PA66 yarn, such as:

$$\sigma = \frac{f}{a}, \tag{1}$$

where, f is the current applied force and a is the current cross-sectional area of the material. However, it is complex to measure a section a for this type of material. This problem arises whether the material is 1D (yarn, rope, strap) or 2D (woven material, canvas) [25]. In this case, instead of the geometric section a, we use the mass to quantify stress. We can then use the specific stress tensor $\underline{\Sigma}$ used in other works of the scientific community [26–33], defined such as:

$$\underline{\Sigma} = \frac{1}{\rho^t} \cdot \underline{\sigma} = \frac{1}{\rho} \cdot \underline{\tau}, \tag{2}$$

where ρ and ρ^t define respectively the initial and the current density of the material and $\underline{\sigma}$ and $\underline{\tau}$ define, respectively, the Cauchy stress tensor and the Kirchhoff stress tensor. The yarn is considered as a one-dimensional (1D) material, which does not exhibit any contraction effect during longitudinal traction. It has neither bending nor torsional rigidity. It also does not exhibit any longitudinal or transverse shear effect. The only non-zero stress and strain components are those along the yarn axis. Thus, the initial and current densities can be written as:

$$\rho = \frac{\overline{\overline{\rho}}}{a} \quad \text{and} \quad \rho^t = \frac{\overline{\overline{\rho^t}}}{a}, \tag{3}$$

where $\overline{\overline{\rho}}$ and $\overline{\overline{\rho^t}}$ are respectively the initial and the current mass per unit length. Therefore, the 1D tensile specific stress Σ of the yarn is written, considering the Eq. (1), (2) and (3):

$$\Sigma = \frac{a}{\overline{\overline{\rho^t}}} \cdot \sigma = \frac{f}{\overline{\overline{\rho^t}}}. \tag{4}$$

This relationship allows us to calculate the specific stress Σ, without resorting to the measurement of the cross-section a of the yarn. Nevertheless, the measure of $\overline{\overline{\rho^t}}$ is difficult to get and it is necessary to obtain a new formulation to define

the specific stress in the yarn. For this, by considering the gradient tensor of the transformation $\underline{F}$ and its Jacobian J, we can write:

$$J = \frac{\rho}{\rho^t} = \frac{\overline{\overline{\rho}}}{\overline{\overline{\rho^t}}} = \det\left(\underline{F}\right) = \lambda, \tag{5}$$

where λ corresponds to the elongation in the longitudinal direction defined as:

$$\lambda = 1 + \frac{\Delta L}{L_0}, \tag{6}$$

where ΔL stands for the crosshead displacement.

By replacing the expression of $\overline{\overline{\rho^t}}$ in Eq. (4), in using Eq. (5), we arrive at a formulation of the specific stress Σ in the longitudinal direction, which depends on the elongation λ, the applied force f and the initial mass per unit length $\overline{\overline{\rho}}$ such as:

$$\Sigma = \frac{\lambda \times f}{\overline{\overline{\rho}}}. \tag{7}$$

Given Eq. (7), the specific stress is expressed in J/g.

We propose two behaviour models to account for the experimental results with the hypothesis of a 1D behaviour of the yarn. These models are inspired, respectively, by the Ogden model and the Neo Hookean model, established in the three-dimensional case with an incompressibility hypothesis where the Jacobian J is equal to 1 [34, 35]. We adopt the form of a 3D model even though our material is considered as a 1D material and, to avoid confusion, we will call these laws respectively the "Ogden-like model" and the "Neo Hookean-like model" and we will use the subscript letters "O" and "NH" to differentiate certain quantities concerning them.

The Ogden-like model is defined such as:

$$\Sigma_{\mathrm{O}}(\lambda) = \mu \cdot \left[\lambda^{\alpha} - \lambda^{\frac{-\alpha}{2}}\right], \tag{8}$$

where μ (in J/g) and α (unitless) are the two parameters of the model.

The Neo Hookean-like model is defined such as:

$$\Sigma_{\mathrm{NH}}(\lambda) = C_1 \cdot \left[\lambda^2 - \lambda^{-1}\right], \tag{9}$$

where C_1 (in J/g) is the only parameter of the model.

2.2. Methods for recognizing the most suitable mechanical model

2.2.1. Presentation of the artificial neural network

Fig. 1 presents the artificial neural network (ANN) used in this work. This perceptron network has only two layers: an input layer and an output layer. The

input layer uses a variable number n of neurons, each of which can take a numerical value. If we consider a specific stress signal Σ as a function of the elongation λ, then this signal is discretized according to the value of n and each input neuron is associated, in order, from top to bottom, with a specific stress value Σ_k with $k \in [\![1, n]\!]$ which is normalized by a value V which is of the order of magnitude of the specific stresses involved. This normalization makes it possible to (i) make the inputs dimensionless and (ii) avoid harmful saturation of the neural network in the case where the specific stress values are high, which could compromise its learning. Thus, the first neuron i_1 is associated with the normalized specific stress value for zero elongation and the last neuron i_n is associated with the normalized specific stress value for the maximum elongation value. We therefore build an input vector I composed of n elements: i_1 to i_n such as $i_k = \Sigma_k / V$ with $k \in [\![1, n]\!]$. The values of the parameters n and V adopted for this study are given in Appendix A.

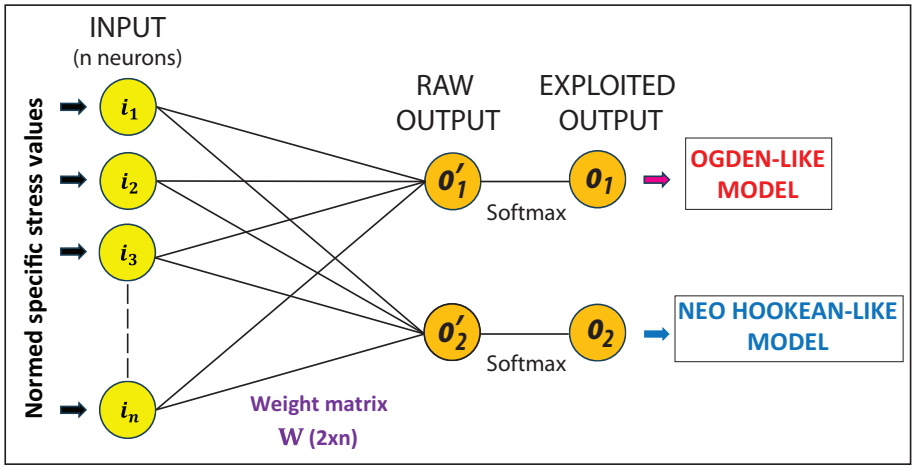


Fig. 1. Artificial Neural Network (ANN) used in the method of recognition for two behaviour laws

The output layer is built in two phases. First, a raw output vector O' with two elements: o'_1 and o'_2 is calculated by a matrix product between the weights matrix W (size $2 \times n$) and the input vector I such as:

$$O' = W \times I \quad \Longleftrightarrow \quad \begin{pmatrix} o'_1 \\ o'_2 \end{pmatrix} = \begin{pmatrix} w_{11} & \dots & w_{1n} \\ w_{21} & \dots & w_{2n} \end{pmatrix} \times \begin{pmatrix} i_1 \\ \vdots \\ i_n \end{pmatrix}. \tag{10}$$

We then recalculate the exploited output vector O whose elements are o_1 and o_2, using a softmax activation function that is commonly used in the last layer of a neural network to perform a multi-class classification task. We thus obtain Eq. (11):

$$\begin{pmatrix} o_1 \\ o_2 \end{pmatrix} = \frac{1}{\exp\left(o'_1\right) + \exp\left(o'_2\right)} \cdot \begin{pmatrix} \exp\left(o'_1\right) \\ \exp\left(o'_2\right) \end{pmatrix}. \tag{11}$$

The values of o_1 and o_2 are between 0 and 1 and only two outcomes are possible for the exploited output vector O. We subsequently consider that:

- if $o_1 > o_2$, then the model recognized by the neural network is the Ogden-like model;
- if $o_2 > o_1$, then the model recognized by the neural network is the Neo Hookean-like model.

If we consider a vector of fixed specific stress values at the input layer, it becomes clear that only the elements of the weights matrix W can change and be selected in such a manner that the network outputs correspond to the recognition of an Ogden-like or a Neo-Hookean-like model. The neural network learning process involves determining a set of elements from the matrix W that achieve the following goal: if a vector of experimental results based on the Ogden-like model is the input, the output o_1 should be greater than o_2. If the vector corresponds to the Neo-Hookean-like model, then the output o_2 should be greater than o_1.

2.2.2. Learning method of the artificial neural network

• Definition of a bounded validity zone

To reduce computation times, during the learning process, it is necessary to define a bounded validity zone around the experimental data to be modelled. This zone is characterized by envelope curves which respectively define minimum and maximum specific stress limits, as a function of the elongation, which should not be exceeded. It is in the bounded validity zone that we guarantee the recognition of the best model for the available data. The definition of the bounded validity zone requires preliminary work, considering the available computing power, the experimental data to be modelled and the equations of the two proposed models. This bounded validity zone is associated with a domain of parameter variations for the two adopted models, in which each parameter evolves within specific minimum and maximum bounds. The Ogden-like model (Eq. (8)) involves 4 values: $\mu_{\min}$, $\mu_{\max}$, $\alpha_{\min}$ and $\alpha_{\max}$ and the Neo Hookean-like model (Eq. (9)) involves 2 values: $C_{1_{\min}}$ and $C_{1_{\max}}$. For each of the two models, a specific validity zone is associated with its parameter variation domain. The envelope curves that delimit these specific validity zones are associated with the parameter sets $(\mu_{\min}, \alpha_{\min})$ and $(\mu_{\max}, \alpha_{\max})$ for the Ogden-like model and the parameters $C_{1_{\min}}$ and $C_{1_{\max}}$ for the Neo Hookean-like model. In each of these two specific validity zones, recognition of the corresponding model should be ensured. The final bounded validity zone, where we wish to guarantee the recognition of the most suitable model among the two, is represented by the intersection of the recognition zones specific to each model. It is therefore important that these specific validity zones have a strong connectivity. Section 4.1 and Appendix A give the parameters values chosen for this study.

• Overview of the learning process with labelled data

Training the neural network requires having a mass of reference data that is introduced into the input layer I, to allow the identification of a matrix W allowing

the network to perform the requested classification task. This mass of reference data is usually called the training set. This training set can be obtained by a mass of experimental results where the identification of the two models is carried out and is a part of the results themselves. The choice adopted in this study is to perform the learning without resorting to experimental data. The strategy adopted consists in using a training set of data generated from the theoretical equations of Ogden-like and Neo Hookean-like models. To get closer to the real signals, we add to these input data a random noise with a uniform distribution (this corresponds to the simplest choice for simulating experimental results), parameterized by a value b in % around the mean value. The data from either theoretical equation are characterized by parameter values chosen for learning. In this context, the corresponding model and the values of the parameters adopted are known in advance. This is called the labelled data for learning. In this study, we chose a sober training set to reduce computation times as much as possible without compromising learning efficiency. This training set consists of six input vectors or reference signals. These come from the set of four parameters of the Ogden-like model: $(\mu_{\min}, \alpha_{\min})$, $(\mu_{\min}, \alpha_{\max})$, $(\mu_{\max}, \alpha_{\min})$ and $(\mu_{\max}, \alpha_{\max})$ and of the two parameters of the Neo Hookean-like model like $C_{1_{\min}}$ and $C_{1_{\max}}$.

The learning process of the neural network is composed of two essential phases. The first one consists in performing the network learning, using the training set defined above, which allows a first selection of a matrix W. The second phase consists in testing the relevance of this first selection over the entire bounded validity zone in the specific stress. To limit computation time, the learning process is repeated a maximum number of times S (see Appendix A). This supervised learning process is illustrated in the form of an algorithm presented in Fig. 2 and a description is given in the following.

Description of phase 1: Training the neural network on the training set and selection of a matrix W

Two methods are proposed here. The method A consists in a pure random search of the elements of the matrix W linked to an iterative process repeated P times (see Appendix A). The method B involves research for a matrix W with a backpropagation algorithm using a stochastic hill climbing type optimization (see Appendix B for details concerning this process). The method A allows us using the neural network presented in Fig. 1 like a Feedforward Neural Network (FNN).

Initially, each element of the matrix W is randomly chosen in the interval $[-W_A, W_A]$ where W_A is the parameter of the neural network whose value is given in Appendix A. For a given reference signal from the training set, representative of the Ogden-like model (respectively Neo Hookean-like), in the ideal case, we wish to obtain at the output O of the neural network: $o_1 = 1$ and $o_2 = 0$ (respectively $o_1 = 0$ and $o_2 = 1$). In a non-ideal case, to quantify the differences between what is obtained and what was expected at the output, once the entire training set is introduced into the input layer, we define a loss indicator parameter, noted L in

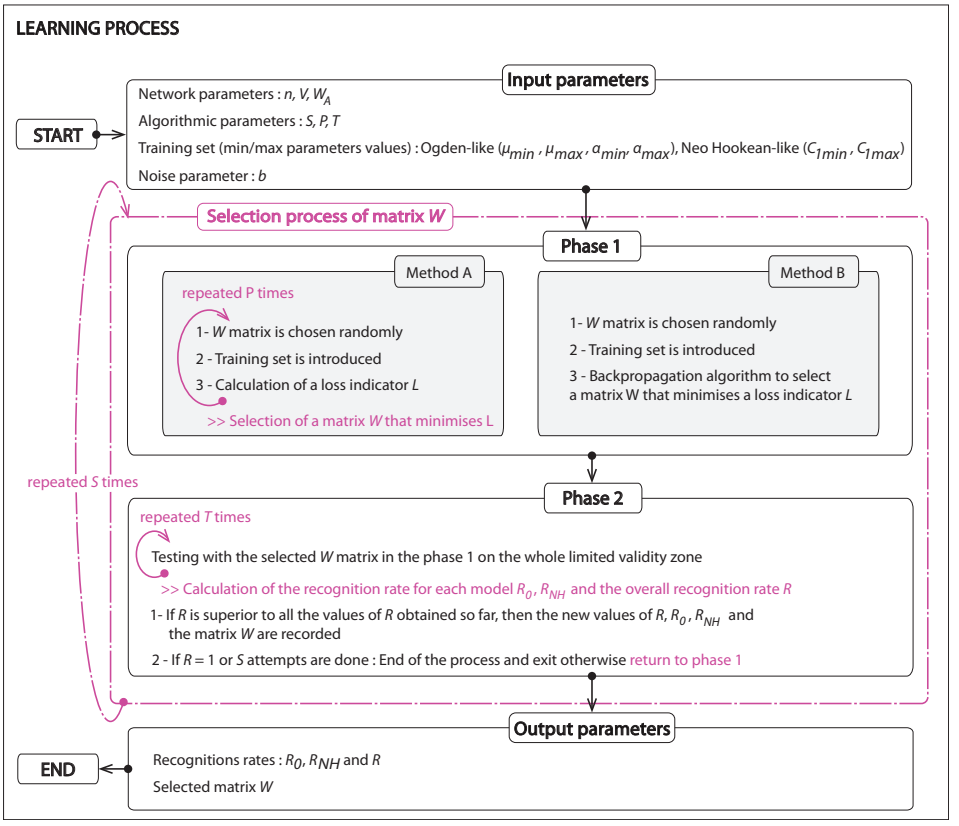


Fig. 2. Algorithm standing for the learning process of the neural network in the case of the recognition of two behaviour laws (Ogden-like and Neo Hookean-like models)

Appendix C. The objective of both methods A and B is to find a weights matrix W that minimizes the value of L, which is then retained for the rest of the process in phase 2.

Description of phase 2: Validation of the matrix W selected in phase 1, by tests carried out on the entire validity zone

Phase 2 allows us to verify whether the selected matrix W in phase 1 can cover the bounded validity zone in its entirety. To do this, we organize a certain number T of tests also with labelled data. A test consists of first randomly choosing a model among Ogden-like and Neo Hookean-like. Then, we perform a random draw of the parameters such as: $\mu \in [\mu_{\min}, \mu_{\max}]$ and $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ if the chosen model is the Ogden-like model and $C_1 \in \left[C_{1_{\min}}, C_{1_{\max}}\right]$ if the chosen model is the Neo Hookean-like model. A specific stress signal, noisy in the same way as in phase 1, is thus constituted and we introduce it into the input layer I of the neural network. We recover, at the output O of the network, the values o_1 and o_2. Two cases then arise:

- if the chosen model is the Ogden-like model and $o_1 > o_2$ then the test is successful;
- if the chosen model is the Neo Hookean-like model and $o_2 > o_1$ then the test is successful.

We count the successes relative to the number of times a model has been selected and thus, we obtain the recognition rates for each of the models: R_O for the Ogden-like model, R_{NH} for the Neo Hookean-like model. We also define R, the recognition rate for both models, defined by the total number of successes divided by the total number of tests. If the value of R obtained is strictly higher than all those obtained so far, then the values of R, R_O, R_{NH} and the associated matrix W are recorded. In the case where a value of R equal to 1 is obtained, the selection process stops: the learning on the labelled training data is successful and the matrix W is definitively chosen. Otherwise, the selection process is restarted by returning to the beginning of phase 1. If after S attempts, it is not possible to obtain a value of R equal to 1, the matrix W that made it possible to obtain the highest possible value of R is adopted. Let us note that this strategy is only possible because the underlying constitutive equations are known, which allows the generation of controlled and consistent synthetic data. This setup encourages the network to learn discriminative patterns associated with each constitutive model rather than dataset statistical artifacts. As a result, if the value of R is close to 1, it means that the network demonstrates strong generalization despite an exceptionally small size of the training set.

3. Experimental background

3.1. Experimental setup

A PA66 yarn of mass per unit length $\overline{\overline{\rho}} = 33dtex = 3.3{\cdot}10^{-3}$ g/m is used in our study (Fig. 3c). To carry out experimental tensile tests on this yarn, we used a single-column tensile press, brand INSTRON 2 kN Microtester 5944, commonly used for tensile tests on medical devices, biomaterials, textiles, elastomers, yarns, cables and plastic films (Fig. 3a). This tensile machine consists of a frame supporting a movable crosshead. The yarn, clamped between the jaws, is anchored at its lower part to the base of the machine and at its upper part to the movable crosshead. The jaws are specific to the traction of yarns (Fig. 3b). The load cell for measuring the force applied to the yarn is designed to carry out tests up to 2 kN. The test bench provides a platform for producing high-quality force and displacement data using the Bluehill® quasi-static and dynamic testing software. At the start of the tensile test, the jaw position is set from the stop. From the point of contact between the two jaws, the crosshead is raised by a distance of 100 mm, corresponding to the initial length L_0 of the yarn. The yarn is placed at this point, then it is tensioned by pre-load, with a force of the order of 0.01 N. The uniaxial tensile tests cited in this

paper are carried out at a constant crosshead speed such as $\Delta\dot{L} = 0.01$ mm/s which corresponds to a strain rate $\dot{\varepsilon} = \Delta\dot{L}/L_0 = 10^{-4}$ s^{-1}.

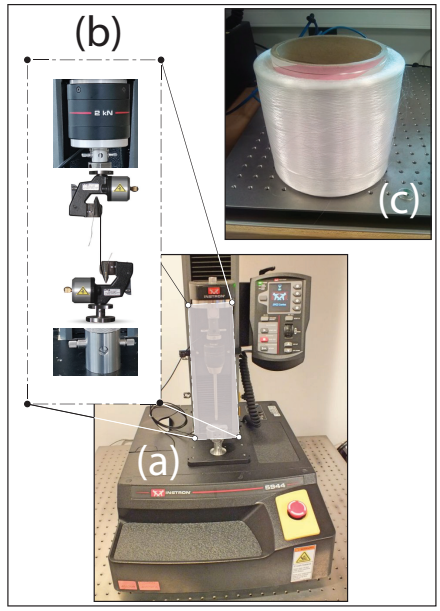


Fig. 3. (a) INSTRON 2 kN Microtester 5944 tensile machine; (b) zoom on the jaws where the yarn is fixed; (c) PA66 yarn spool used in this study

3.2. Mechanical behaviour modelling of experimental results

Five uniaxial tensile tests (the tests numbered 1, 2, 3, 4 and 5 in this paper) are performed under identical experimental conditions. Fig. 4 shows both the experimental curves (in the order of their numbering) as well as the fit and identification of the parameters of the two models with the associated RMSE (Root Mean Square Error) values (adjustment by the least squares method). The curves present disparities in terms of specific stress (Fig. 4a) related to the existence of measurement noise, because the way of attaching the yarn may be slightly different from one test to another and the fact that the material is probably not homogeneous from one sample to another. The model of Ogden-like models test 1 (Fig. 4b) and test 2 (Fig. 4c) performs very well: the RMSE values are low. It also models test 3 well (Fig. 4d) but with a higher RMSE value. For test 4 (Fig. 4e), the modelling becomes more complicated and the RMSE value for the Ogden-like model is lower, but still very close to that obtained for the Neo-Hookean-like model. For test 5 (Fig. 4f), the Neo-Hookean-like model clearly allows better modelling. In summary, identifying the most suitable model is easy for curves 1, 2 and 5. It is moderately easy for curve 3 and difficult for curve 4.

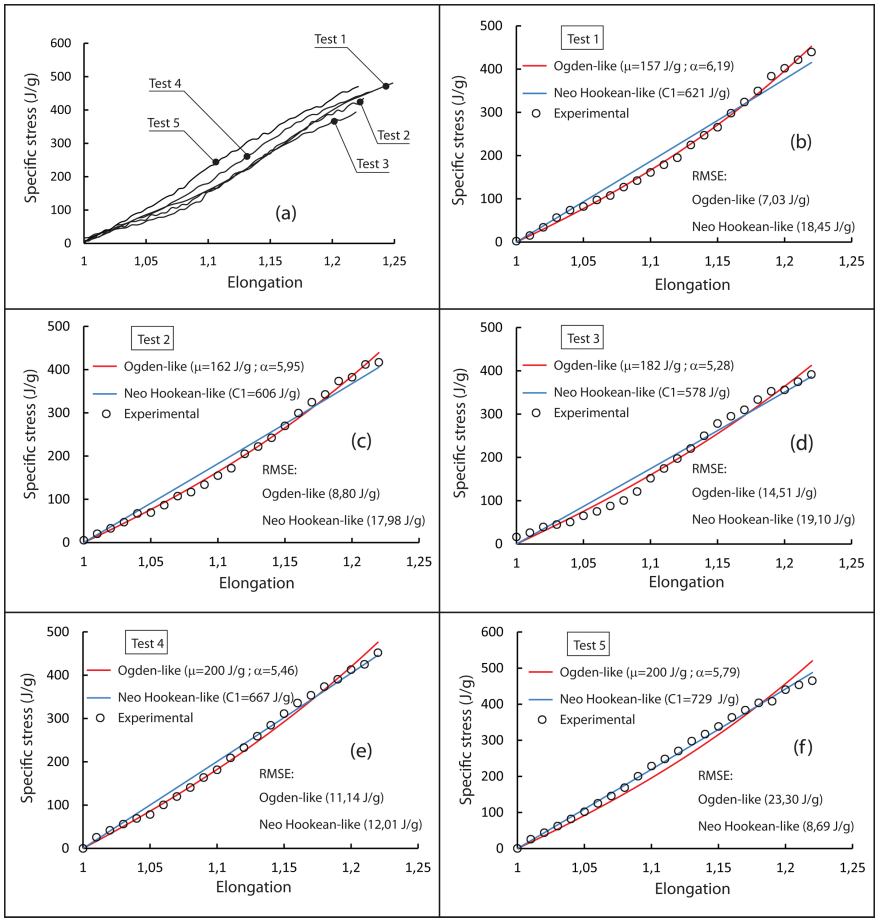


Fig. 4. Characterization of the yarn used in our study: (a) Specific stress-elongation curves of experimental tests 1, 2, 3, 4 and 5; (b), (c), (d), (e) and (f) Fitting of each curve by the Ogden-like model and the Neo Hookean-like model with the least square method

It is important to note that the model parameters are identified using a least-squares approach over the entire loading range. This strategy ensures robustness and allows for a consistent comparison between the Ogden-like and the Neo Hookean-like models but it may smooth localized features of the experimental data and introduce bias depending on the selected elongation range. The Neo Hookean-like model, characterized by a linear specific stress–elongation relationship, does not account for nonlinear effects at elongation ranges where they occur. Even if the Ogden-like model provides a better flexibility, its predictive capability remains dependent on the identification strategy and the range of data that is considered. Therefore, the results should be interpreted with caution, for example, when extrapolating beyond the elongation zone presented in this paper or when considering local features that are not highlighted by the global adjustment procedure.

4. Results

4.1. Bounded validity zone

In the context of this study, the chosen values of the parameters limits of the two models $\mu_{\min}$, $\mu_{\max}$, $\alpha_{\min}$, $\alpha_{\max}$, $C1_{\min}$ and $C1_{\max}$, are given in Appendix A and Fig. 5 shows specific validity zones of the Ogden-like model (Fig. 5a), of the Neo

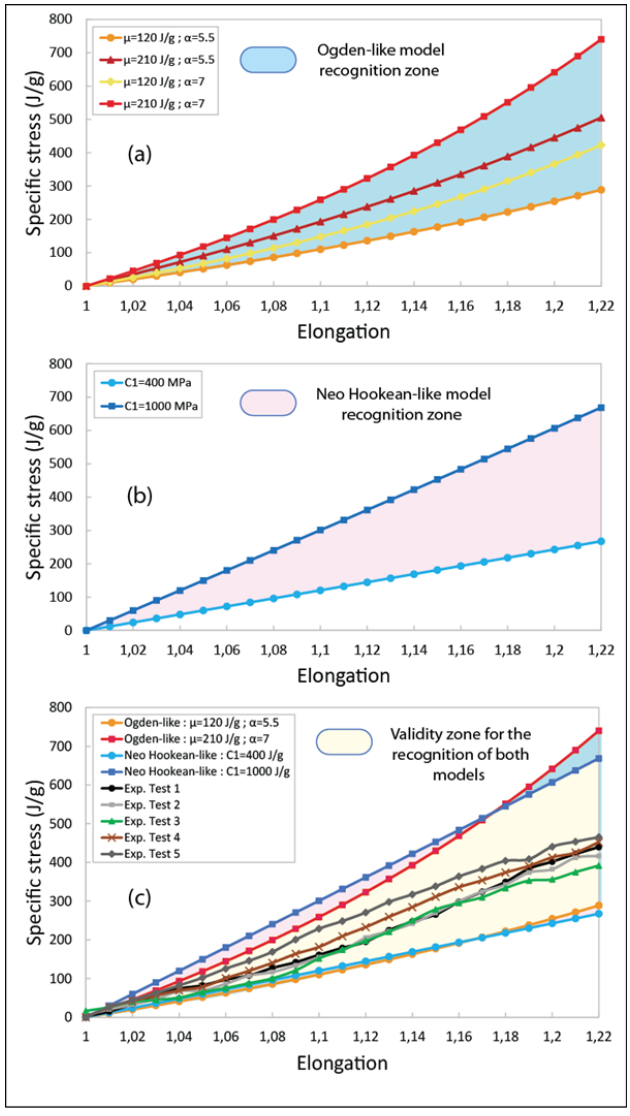


Fig. 5. (a) and (b) Specific validity zone associated to respectively the Ogden-like model and the Neo Hookean-like model. (c) Definition of the validity zone as the intersection of the both models specific validity zones where the experimental points of the tests 1, 2, 3, 4 and 5 are included

Hookean-like model (Fig. 5b) and the final bounded validity zone in specific stress where we wish to ensure recognition of the two models, which is represented by the yellow zone in Fig. 5c. We check that the experimental points of tests 1, 2, 3, 4 and 5 are included in this bounded validity zone. Let us note that we have chosen the parameters limits to form a bounded validity zone large enough to show the neural networks recognition capability over a large range of parameters values and despite the small size of the training set. This choice also leads to reducing calculation times to reasonable values.

4.2. Results of recognition on labelled training data

We applied the learning process several times in the frame of a sensitivity study of the recognition rates R_O, R_{NH} and R, to the parameters b and n. The parameter T is always set up to 10 000. Let us note that the greater n, b and T are, the longer are computation times to get $R = 1$, so it is important for the user to assess the possible values for n, b and T that give satisfaction and do not imply too long computation times.

Results with method A: Four neural network learning campaigns are considered: campaign A-0-12 ($b = 0$ and $n = 12$), campaign A-0-23 ($b = 0$ and $n = 23$), campaign A-5-12 ($b = 5$ and $n = 12$) and campaign A-5-23 ($b = 5$ and $n = 23$). For each campaign, 100 learning processes are carried out. When b was set to 0 in campaigns A-0-12 and A-0-23, the learning process allowed, by applying a single loop (phase 1, phase 2) of the selection process, to determine a matrix W such that $R = R_O = R_{\mathrm{NH}} = 1$ whatever the values of n adopted. When $b = 5$, learning rate remains quite high and we also obtain $R = R_O = R_{\mathrm{NH}} = 1$ after reasonable computation times (for one learning process, they are from a few minutes to a few tens of minutes in the case where a standard commercial personal computer is used).

Results with method B: We set $n = 12$ and we consider three neural network learning campaigns. For each, 100 learning processes are still carried out. We note campaigns with the code B-5 ($b = 5$), B-10 ($b = 10$) and B-15 ($b = 15$). We notice that the greater is the value of b, the lower is the learning rate. Consequently, the number of maximal loops S is set up to 100 (if $b \leq 10$) and to 200 (if $b = 15$) which enables us to get values of R that are as high as possible and prevents too long computation times (for one learning process, computation times are about 5 minutes if $S = 100$ or 10 minutes if $S = 200$ with a standard commercial personal computer). For each campaign, the mean value of R for 100 different learning processes is listed in Table 2. Let us note that the noise around the average value

Table 2. Mean value of R obtained for 100 learning processes (campaigns with method B)

Campaign	B-5	B-10	B-15
R	0.985	0.929	0.854

applied on labelled input data for the campaign B-15 appears to be higher than the one observed on the external experimental data. Although the mean value of R is lower, because training is more difficult in this case, it is still conceivable that the trained neural networks can obtain good recognition results on the external experimental data presented in this paper.

4.3. Results of recognition on external experimental data

Once matrix W is determined by the neural network learning method (see Section 2.2.2), we can proceed to validate this learning by confronting the neural network with external experimental data that it has never seen. We therefore introduce the experimental specific stress data from tests 1, 2, 3, 4 and 5 (see Section 3.2, Fig. 4) in the input layer and we retrieve the values of o_1 and o_2 at the output layer. If $o_1 > o_2$, the Ogden-like model is recognized whereas if $o_2 > o_1$, the Neo Hookean-like model is recognized. We check whether the Ogden-like model that better represents the experimental data from tests 1, 2, 3 and 4 is recognized or not. Same for the Neo Hookean-like model concerning test 5.

• Results with method A

Tables 3a, 3b, 3c and 3d show the confusion matrices obtained concerning the tests of recognition of Ogden-like and Neo Hookean-like models on experimental tests 1, 2, 3, 4 and 5, from the sets of matrices W obtained from the different neural network trainings from the campaigns A-0-12, A-0-23, A-5-12 and A-5-23. If the recognition is perfect for the 100 neural network trainings of a campaign, we should get respectively 400 and 100 in the case corresponding to Ogden-like expected / Ogden-like recognized (tests 1, 2, 3, and 4) and in the case corresponding to Neo Hookean-like expected / Neo Hookean-like recognized (test 5). In addition, we should get 0 in the other cases. In addition, we define in the Appendix C, a performance index Q that provides additional information beyond the number of times the correct model was recognized out of the total number of recognition tests performed in a campaign. Indeed, we are also interested in its ability to clearly rule out the model that is less adaptable. The index Q is expressed in % in Tables 3a, 3b, 3c, 3d. What's more, it is close to 100%, the better the performance is. An untrained neural network would give a Q value close to 50%. First, we note that recognition results are very sensitive to the value of b. The campaigns where $b = 5$, in one hand, clearly leaded to a better recognition of the most adapted model for an experimental test and, on the other hand, could better rule out the other one. Moreover, for all the campaigns, the tests 1, 2, 3 and 5 were always recognized and the errors always occurred for the test 4, which was a consequence of the difficulty of recognition for this test (see Fig. 4e). For the campaign A-5-12, only 17 errors out of 500 are detected which represent 3.4%. Besides, for the test 4, the campaigns A-5-12 and A-5-23 lead to the right recognition in more than 80% of the cases and the performance index Q greater than 97%. Sensitivity to the parameter n is not obvious and it is difficult to draw a conclusion despite the fact that we observe

that for both cases where b is fixed at 0 and then at 5 the Q value is slightly higher for $n = 23$. It is also interesting to notice that despite the numerous errors for campaigns when b is set to 0, the Q value remains high, which means that the errors are rather made with a narrow margin.

Table 3. Confusion matrix and value of Q in % for campaigns: (a) A-0-12; (b) A-0-23; (c) A-5-12 and (d) A-5-23

(a)

A-0-12 ($Q = 95.15$)	Ogden-like recognized	Neo Hookean-like recognized
Ogden-like expected	333	67
Neo Hookean-like expected	0	100

(b)

A-0-23 ($Q = 95.71$)	Ogden-like recognized	Neo Hookean-like recognized
Ogden-like expected	339	61
Neo Hookean-like expected	0	100

(d)

A-5-12 ($Q = 97.06$)	Ogden-like recognized	Neo Hookean-like recognized
Ogden-like expected	383	17
Neo Hookean-like expected	0	100

(d)

A-5-23 ($Q = 98.11$)	Ogden-like recognized	Neo Hookean-like recognized
Ogden-like expected	381	19
Neo Hookean-like expected	0	100

- Results with method B

Tables 4a, 4b and 4c, respectively, show the confusion matrices and the value of the performance index Q in %, concerning the tests of recognition of Ogden-like and Neo Hookean-like models in experimental tests 1, 2, 3, 4 and 5, from the campaigns B-5, B-10 and B-15. Recognition results are also here, sensitive to the value of b. The best results are given by the campaign B-15 ($b = 15$ and $n = 12$) with a 3.6% error rate (the most adapted model has been identified 482 times on 500) and a Q value of 98.52%. Also, for experimental test 4, the Ogden-like model has been recognized more than 4 times out of 5. In the cases presented here, using method B with $b = 15$, we were able to equal the performance given by the bests results obtained in using method A.

Table 4. Confusion matrix and index Q in % for campaigns: (a) B-5; (b) B-10, and (c) B-15 (n = 12, b = 15)

(a)

B-5 (Q = 95.47)	Ogden-like recognized	Neo Hookean-like recognized
Ogden-like expected	346	54
Neo Hookean-like expected	3	97

(b)

B-10 (Q = 97.07)	Ogden-like recognized	Neo Hookean-like recognized
Ogden-like expected	364	36
Neo Hookean-like expected	0	100

(c)

B-15 (n = 12, b = 15) (Q = 98.52)	Ogden-like recognized	Neo Hookean-like recognized
Ogden-like expected	382/400 (Test 1: 100/100) (Test 2: 100/100) (Test 3: 98/100) (Test 4: 84/100)	18/400 (Test 1: 0/100) (Test 2: 0/100) (Test 3: 2/100) (Test 4 : 16/100)
Neo Hookean-like expected	0/100	100/100

4.4. Concluding remarks about the results

We note that the experimental signals to be recognized have relatively fluctuating specific stress values (see Fig. 4a), which suggests, at first glance, to preferentially set up training of neural networks with noisy labelled data. It is clearly seen that the value of the noise parameter b applied on input data significantly affects the recognition and setting this value to 5 for method A and to 15 for method B, makes it possible to significantly reduce the overfitting problem (which reflects the good performance on the training data, but the errors observed on new external data that the neural network had not yet encountered). For both methods A and B, computation times and the best results obtained are similar: perfect or near-perfect recognition of the most suitable model on experimental tests 1, 2, 3 and 5 and more than 80% recognition of the Ogden-like model for test 4 where recognition is difficult because the curves standing for the both models are very close to each other. For the recognition of the most suitable model on the experimental data presented in the papier, the obtained results allow us to recommend the choice of a very simple neural network such as that shown in Fig. 1, (i) used as a feedforward neural network (no backpropagation algorithm) and its training with parameters such as: n = 12, b = 5 and T = 10 000, or (ii) with a backpropagation algorithm and

its training with parameters such as: $n = 12$, $b = 15$ and $T = 10\,000$. According to the results given by the least square method corroborated by the recognition results given by the neural network, we finally recommend to model the experimental data with the Ogden-like model. In conclusion, these results finally make it possible to meet the set objective and to validate the proposed method in the case of the recognition between two models of mechanical behaviour laws that are Ogden-like and Neo Hookean-like applied to uniaxial traction tests on the PA66 material presented in this paper.

5. Conclusion

In this paper, we have proposed methods that use a classifier to recognize automatically the model best suited to data in the form of experimental elongation-stress curves, among two models proposed as the Ogden-like and the Neo Hookean-like allowing to model the mechanical behaviour in specific stress, in uniaxial tension of PA66 yarns, thanks to an Artificial Neural Network (ANN) with only one input layer and one output layer (no hidden layer). We carried out various supervised learning by varying the learning method (method A: without backpropagation algorithm or method B: with backpropagation algorithm) and by varying parameters, like the discretization parameter n and the noise parameter b (in % around the mean value, uniform distribution) applied on input labelled data in specific stresses as a function of elongation. We noticed that the noise parameter b has a great influence on the learning process of neural networks and the ability of these trained networks to recognize the model best suited to experimental data. Noising the labelled input data lead to better recognition results but also resulted in longer computation times for the learning process. In the case of this study which presents a low-dimensional problem, both methods A and B can be used to reach the objective. Furthermore, a strong point to note is that the neural network was trained using training sets based on labelled data that do not require the introduction of external data. This allows us to avoid the need for massive data, which is a recurring drawback of techniques using machine learning and deep learning. The study put forward in this paper showed a certain robustness of the method of recognition of a model of mechanical behaviour law on elongation-stress curves. It meets the classification objective consisting in distinguishing the cases even if some of them result in curves that can be very close. The methods detailed in this article, intended for members of the mechanical engineering community wishing to begin using artificial intelligence in their research, could be applied to various fields of materials mechanics. For example, these methods could be adapted: (i) in imaging for recognizing stress or strain states in a material in relation to images that have undergone prior processing to reduce the input data for the neural network; or (ii) for studies in a two-dimensional context, such as identifying biaxial tensile laws in the framework of modelling a relationship between stress states (σ_x, σ_y) and strain states $(\varepsilon_x, \varepsilon_y)$.

Those prospects could be the subject of future works carried out by our research team in the frame of studies of heterogeneous materials.

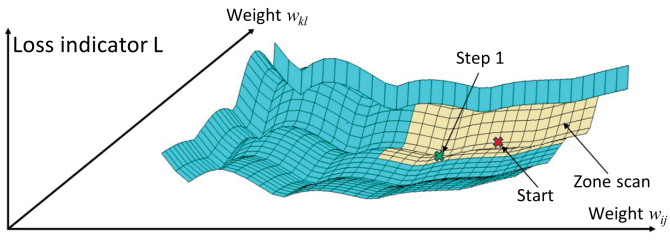
A. Values of the parameters used in the learning process

Parameter		Learning process with method A	Learning process with method B
n		12 or 23	12
V		400 J/g	
b		0 or 5	5, 10 or 15
S		1000	100 (if $b \leq 10$) 200 (if $b = 15$)
P		Randomly chosen between 300 and 10 000 for each Phase1-Phase2 loop	
T		10 000	
W_A		10	100
Ogden-like	$\mu_{\min}$	120 J/g	
	$\mu_{\max}$	210 J/g	
	$\alpha_{\min}$	5.5	
	$\alpha_{\max}$	7	
Neo Hookean-like	$C_{1_{\min}}$	400 J/g	
	$C_{1_{\max}}$	1000 J/g	

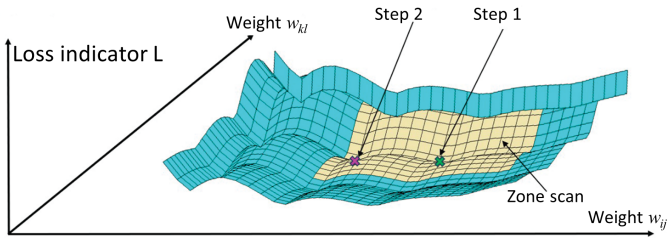
B. Details about the backpropagation algorithm

Learning process explained in Section 2.2.2 involves a method B that puts forward a backpropagation algorithm to find a weights matrix W that minimizes a loss indicator L defined in the Appendix C. This algorithm uses a Stochastic Hill Climbing type optimization (SHC). Fig. 6 shows a schematic case of a 3D graph representation of the loss indicator L as a function of two elements w_{ij} and w_{kl} of the matrix W. In reality, the algorithm acts on all elements of matrix W. SHC explores a zone around a starting point. If a point located in the zone is found to better minimize L, this point is selected (step 1 in the schematic Fig. 6a). A new zone is defined around this point and a new point that better minimizes L is sought (step 2 in the schematic Fig. 6b). SHC continues the exploration until it cannot find a better solution for a specified number of iterations. The main advantages of SHC are: (i) the implementation of this algorithm is quite easy and (ii) the searching

method can potentially escape a local minimum located near the starting point. In that case, it can be more efficient than a gradient algorithm often used in the frame of backpropagation for neural networks training.



(a)



(b)

Fig. 6. (a) Initialization of SHC algorithm from a starting point and research of a point in the zone that minimizes the loss indicator L (step 1). (b) From the point found at step 1 a new zone is defined and a new point that minimizes L is sought (step 2)

C. Definition of the loss indicator *L*

We introduce as input to the neural network, turn by turn, the first part of the training set containing the specific stresses Σ_O (Eq. (8)) calculated from the Ogden-like model whose parameters are defined by the four pairs: ($\mu_{\min}$, $\alpha_{\min}$), ($\mu_{\min}$, $\alpha_{\max}$), ($\mu_{\max}$, $\alpha_{\min}$) and ($\mu_{\max}$, $\alpha_{\max}$). Then, we introduce again, turn by turn, the second part of the training set containing the specific stresses Σ_{NH} (Eq. (9)) of the Neo Hookean-like model calculated for the 2 parameters: $C_{1_{\min}}$ and $C_{1_{\max}}$. We note each time the values of o_1 and o_2 obtained at the output of the network, which we index: o_1^k and o_2^k with $k \in [\![1,4]\!]$ when we introduce the Ogden-like model and $k \in [\![5,6]\!]$ when we operate with the Neo Hookean-like model. For $k \in [\![1,4]\!]$, the recognition of the Ogden-like model is expected and we wish to obtain in an ideal case: $o^k_{1_{\text{expected}}} = 1$ and $o^k_{2_{\text{expected}}} = 0$. For $k \in [\![5,6]\!]$, the Neo Hookean-like model should be recognized and we wish to get in an ideal case: $o^k_{1_{\text{expected}}} = 0$ and $o^k_{2_{\text{expected}}} = 1$.

Cross-entropy, also known as logarithmic loss, is a loss function commonly used in machine learning to measure the performance of the classification realized by the neural network. We then calculate loss indicators L_{O} and L_{NH} as follows:

$$\begin{aligned} L_{\mathrm{O}} &= -\sum_{k=1}^{k=4}\left[o^k_{1_{\mathrm{expected}}}\log\left(o^k_1\right)+o^k_{2_{\mathrm{expected}}}\log\left(o^k_2\right)\right] \\ &= -\sum_{k=1}^{k=4}\left[\log\left(o^k_1\right)\right], \end{aligned} \tag{C1}$$

$$\begin{aligned} L_{\mathrm{NH}} &= -\sum_{k=5}^{k=6}\left[o^k_{1_{\mathrm{expected}}}\log\left(o^k_1\right)+o^k_{2_{\mathrm{expected}}}\log\left(o^k_2\right)\right] \\ &= -\sum_{k=5}^{k=6}\left[\log\left(o^k_2\right)\right]. \end{aligned} \tag{C2}$$

We then define a loss indicator L considering the two previous contributions, weighted so that they have the same weight. We thus propose the following formulation:

$$L = \frac{1}{3}L_{\mathrm{O}} + \frac{2}{3}L_{\mathrm{NH}}. \tag{C3}$$

D. Definition of the performance index Q

We consider all the weights matrices W obtained after learning processes for a whole campaign of neural network trainings. For each campaign, depending on the number of the experimental test j (ranging from 1 to 4 in the case where the Ogden-like model is the most suitable one) and the learning i carried out (ranging from 1 to 100), we can index each value obtained for o_1 and o_2 by the notations: $o_1(i,j)$ and $o_2(i,j)$.

We now define quantities d_1, d_2 and D such as:

$$d_1 = \sum_{j=1}^{j=4}\left(\sum_{i=1}^{i=100}\left((o_1(i,j)-1)^2+(o_2(i,j)-0)^2\right)\right), \tag{D1}$$

$$d_2 = \sum_{i=1}^{i=100}\left((o_1(i,5)-0)^2+(o_2(i,5)-1)^2\right), \tag{D2}$$

$$D = d_1 + d_2\,. \tag{D3}$$

If we consider an ideal case of perfect recognition performance on all experimental tests for a given campaign, we must obtain $o_1(i,1)$, $o_1(i,2)$, $o_1(i,3)$, $o_1(i,4)$, $o_2(i,5)$ equal to 1 and $o_2(i,1)$, $o_2(i,2)$, $o_2(i,3)$, $o_2(i,4)$, $o_1(i,5)$ equal to

0. In this case, D is null. In the strictly opposite case where all the values $o_1(i, j)$ and $o_2(i, j)$ are reversed, D takes the maximum value 1 000. Thus, to quantitatively evaluate the quality of recognition over the entirety of a campaign, we define a performance index, noted Q and expressed in %, defined as:

$$Q = 100\sqrt{\frac{1000 - D}{1000}}\,. \tag{D4}$$

References

[1] R. Liu, B. Yang, E. Zio, and X. Chen. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, 108:33–47, 2018. doi: 10.1016/j.ymssp.2018.02.016.

[2] Z. Heda. Fault diagnosis and life prediction of mechanical equipment based on artificial intelligence. *Journal of Intelligent and Fuzzy Systems*, 37(3):3535–3544, 2019. doi: 10.3233/JIFS-179157.

[3] L. Wang, Y. Li, L. Zhou, Y. Lou, S. Liu, D. Zheng, and M. Yi. Progress in additive manufacturing, additive repair and fatigue evaluation of aviation titanium alloy blades. *Materials Research Letters*, 11(12):973–1012, 2023. doi: 10.1080/21663831.2023.2275599.

[4] D.V. Dao, H.B. Ly, S.H. Trinh, T.T. Le, and B.T. Pham. Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete. *Materials*, 12(6):983, 2019. doi: 10.3390/ma12060983.

[5] N. Nawafleh and F.M. AL-Oqla. Artificial neural network for predicting the mechanical performance of additive manufacturing thermoset carbon fiber composite materials. *Journal of the Mechanical Behavior of Materials*, 31(1):501–513, 2022, https://www.degruyterbrill.com/document/doi/10.1515/jmbm-2022-0054/html.

[6] J. Tang, Z. Wang, Q. Tan, J. Cao, and X. Ling. Determining the twist angle of stacked MoS2 layers using machine learning-assisted low-frequency interlayer Raman fingerprints. *Journal of Raman Spectroscopy*, 54(9):1021–1029, 2023. doi: 10.1002/jrs.6577.

[7] C. Ma, L. An, A. Di Donna and D. Dias. Application of machine learning technique to predict the energy performance of energy tunnels. *Computers and Geotechnics*, 166, 2024, 106010. doi: 10.1016/j.compgeo.2023.106010.

[8] X. Peng, J. Zhou, C. Zhang, R. Li, Y. Xu, and D. Chen. An intelligent optimization method for vortex-induced vibration reducing and performance improving in a large Francis turbine. *Energies*, 10(11):1901, 2017. doi: 10.3390/en10111901.

[9] Y. Kabaldin and D. Shatagin. Artificial intelligence and cyber-physical mechanical processing systems in digital industry. In MATEC Web of Conferences, Vol. 226, p. 05001. EDP Sciences, 2018.

[10] G.M. Uddin, S.M. Arafat, A.H. Kazim, M. Farhan, S.G. Niazi, N. Hayat ... and S. Kamarthi. Artificial intelligence-based Monte-Carlo numerical simulation of aerodynamics of tire grooves using computational fluid dynamics. *Ai Edam*, 33(3):302–316, 2019. doi: 10.1017/S0890060419000039.

[11] A. Arbaoui, A. Ouahabi, S. Jacques, and M. Hamiane. Concrete cracks detection and monitoring using deep learning-based multiresolution analysis. *Electronics*, 10(15):1772, 2021. doi: 10.3390/electronics10151772.

[12] T. Gantala and K. Balasubramaniam. Automated defect recognition for welds using simulation assisted TFM imaging with artificial intelligence. *Journal of Nondestructive Evaluation*, 40(1):28, 2021, https://link.springer.com/article/10.1007/s10921-021-00761-1.

[13] T. Zhang, D. Shi, Z. Wang, P. Zhang, S. Wang, and X. Ding. Vibration-based structural damage detection via phase-based motion estimation using convolutional neural networks. *Mechanical Systems and Signal Processing*, 178, 109320, 2022. doi: 10.1016/j.ymssp.2022.109320.

[14] J. Koutsoupakis, P. Seventekidis, and D. Giagopoulos. Machine learning based condition monitoring for gear transmission systems using data generated by optimal multibody dynamics models. *Mechanical Systems and Signal Processing*, 190, 2023, 110130. doi: 10.1016/j.ymssp.2023.110130.

[15] C. Yang, Y. Kim, S. Ryu, and G.X. Gu. Prediction of composite microstructure stress-strain curves using convolutional neural networks. *Materials and Design*, 189, 2020, 108509. doi: 10.1016/j.matdes.2020.108509.

[16] B. Caglar, G. Broggi, M.A. Ali, L. Orgéas, and V. Michaud. Deep learning accelerated prediction of the permeability of fibrous microstructures. *Composites Part A: Applied Science and Manufacturing*, 158, 2022, 106973. doi: 10.1016/j.compositesa.2022.106973.

[17] J.N. Heidenreich, C. Bonatti, and D. Mohr. Transfer learning of recurrent neural network-based plasticity models. *International Journal for Numerical Methods in Engineering*, 125(1), e7357, 2024. doi: 10.1002/nme.7357.

[18] A. Danoun, E. Prulière and Y. Chemisky. FE-LSTM: A hybrid approach to accelerate multiscale simulations of architectured materials using Recurrent Neural Networks and Finite Element Analysis. *Computer Methods in Applied Mechanics and Engineering*, 429, 2024, 117192. doi: 10.1016/j.cma.2024.117192.

[19] T.N.K. Nguyen, T. Dairay, R. Meunier and M. Mougeot. Physics-informed neural networks for non-Newtonian fluid thermo-mechanical problems: An application to rubber calendering process. *Engineering Applications of Artificial Intelligence*, 114, 2022, 105176. doi: 10.1016/j.engappai.2022.105176.

[20] G. Basso Della Mea, C. Ovalle, L. Laiarinandrasana, E. Decencière, and P. Dokladal, Physics Informed Self-Supervised Segmentation of Composite Materials. Available at SSRN 4807639. doi: 10.1016/j.cma.2024.117355.

[21] M. Zhou and G. Mei. Transfer Learning-Based Coupling of Smoothed Finite Element Method and Physics-Informed Neural Network for Solving Elastoplastic Inverse Problems. *Mathematics*, 11(11):2529, 2023. doi: 10.3390/math11112529.

[22] F. Masi, I. Stefanou, P. Vannucci, and V. Maffi-Berthier. Thermodynamics-based artificial neural networks for constitutive modeling. *Journal of the Mechanics and Physics of Solids*, 147, 2021. 104277. doi: 10.1016/j.jmps.2020.104277.

[23] N. Pistenon, P. Kerfriden, J.L. Bouvard, D.P. Muñoz, and S. Cantournet. Un réseau thermodynamique récurrent (RTAN) avec encodage de Maxwell pour modéliser la réponse de matériaux viscoélastiques. In 16ème Colloque National en Calcul de Structures. 2024, https://hal.science/hal-04611057/.

[24] L. Civier, Y. Chevillotte, G. Bles, G. Damblans, F. Montel, P. Davies and Y. Marco. Visco-elasto-plastic characterization and modeling of a wet polyamide laid-strand sub-rope for floating offshore wind turbine moorings. *Ocean Engineering*, 303, 2024, 117722. doi: 10.1016/j.oceaneng.2024.117722.

[25] W. Dib, G. Blès, A. Blaise, and A. Tourabi. Modelling of cyclic visco-elasto-plastic behaviour of coated woven fabrics under biaxial loading and finite strain. *International Journal of Solids and Structures*, 154:147–167, 2018. doi: 10.1016/j.ijsolstr.2017.08.002.

[26] F. Fer. Thermodynamique macroscopique 2 (Macrocopic thermodynamics 2), 1971.

[27] J. Mandel, Introduction à la mécanique des milieux continus déformables, 1974.

[28] P. Rougée. Mécanique des grandes transformations (Vol. 25). Springer Science and Business Media, 1997.

[29] J. Alfthan. The effect of humidity cycle amplitude on accelerated tensile creep of paper. Mechanics of Time-Dependent Materials, 8:289–302, 2004, https://link.springer.com/article/10.1007/s11043-004-0536-0.

[30] G. Leclère, A. Nême, J.Y. Cognard, and F. Berger. Rupture simulation of 3D elastoplastic structures under dynamic loading. *Computers and structures*, 82(23-26):2049–2059, 2004. doi: 10.1016/j.compstruc.2004.03.073.

[31] A. Curnier. Mécanique des solides déformables: Cinématique, dynamique, énergétique (Vol. 1). PPUR presses polytechniques, 2005.

[32] J.E. Fitzgerald. A tensorial Hencky measure of strain and strain rate for finite deformations. *J. Appl. Phys.*, 51:5111–5115, 2008. doi: 10.1063/1.327428.

[33] E. Bele, B.A. Bouwhuis, and G.D. Hibbard Microstructural design in work hardenable perforation stretch formed micro-truss cores. *Composites Part A: Applied Science and Manufacturing*, 40(8):1158–1166, 2009. doi: 10.1016/j.compositesa.2009.05.005.

[34] L.R.G. Treloar. The elasticity of a network of long-chain molecules – II. *Transactions of the Faraday Society*, 39:241–246, 1943. doi: 10.1039/TF9433900241.

[35] R.W. Ogden. Large deformation isotropic elasticity–on the correlation of theory and experiment for incompressible rubberlike solids. Proceedings of the Royal Society of London. A. *Mathematical and Physical Sciences*, 326(1567):565–584, 1972. doi: 10.1098/rspa.1972.0026.