Neural networks for predicting the temperature-dependent viscoelastic response of PEEK under constant stress rate loading

Alexandra Aulova *, Alen Oseli, Marko Bek

Faculty of Mechanical Engineering, University of Ljubljana, Ljubljana, Slovenia, Aškerčeva Ulica 6, Ljubljana, 1000; Slovenia

ARTICLE INFO

Keywords: 
PEEK
Temperature
Neural network
Multilayer perceptron
Constant stress rate
Prediction
Modeling
High-performance polymers
Composites

ABSTRACT

High-performance polymer composites are used in demanding applications in civil and aerospace engineering. Often, structures made from such composites are monitored using structural health monitoring systems. This investigation aims to use a multilayer perceptron neural network to model polymer response to a non-standard excitation under different temperature conditions. Model could be implemented into health monitoring systems. Specifically, the neural network was used to model PEEK material’s creep behavior under constant shear stress rate excitation at different temperatures. Optimal neural network topology, the effect of the amount of training data and its distribution in a temperature range on prediction quality were investigated.

The results showed that based on the proposed optimization criterion, a properly trained neural network can predict polymeric material behavior within the experimental error. The neural network also enabled good prediction at temperatures where stress-strain behavior was not experimentally determined.

1. Introduction

High-performance polymer composites (HPC) can be found in almost any engineering field, including aerospace, defense, energy and automotive. Comprised of a polymer resin and a filler, these materials demonstrate high strength and stiffness at low weight and represent a strong alternative to traditional manufacturing materials, including steel and aluminum.

Frequently, thermostet resins are used to make HPC materials. However, lately, more and more HPC materials are made using high-performance thermoplastic resins. There are several advantages to using thermoplastic-based composites: damaged structures or parts are easily repaired, manufacturing of parts is relatively simple, high production process automation is possible, and parts can be recycled and used again. High-performance thermoplastic polymers such as polyethersulphone (PES) and polyetheretherketone (PEEK) in certain areas even surpass thermostets in both temperature and humidity resistance [1].

One of the challenges with HPC materials is their complex mechanical behavior and lack of reliable methods for fatigue and failure prediction that would include time, temperature and humidity effects. All these factors result in very demanding HPC failure prediction and analysis.

Often, changes in material properties and structural changes need to be monitored over time for complex HPC-based constructions. To track the state of a construction, structural health monitoring systems (SHM) are used. SHM was developed to monitor structures made of metals, which are time-independent materials. Therefore, they are not designed to distinguish between the geometrical changes of structures (caused by cracks or delamination) and change related to composites’ viscoelasticity and the influence of temperature or humidity on viscoelasticity. The signal obtained from the structure contains combined information of geometrical and material changes and SHM cannot detect changes in the signal caused by environmental effects (temperature and humidity).

Including viscoelastic material response in SHM systems is not a trivial task since the SHM should solve in real-time the problem of interrelating excitation, material properties (time-, temperature- and humidity-dependent) and structure response. To predict which changes in structural response are caused by viscoelastic effects analytically, a complete description of viscoelastic material properties needs to be known (for example, creep compliance). Determination of viscoelastic function from excitation and structural response is an inverse problem, well known in viscoelasticity [2–6]. Its analytical solution is known only for simple geometries and standard excitations (step and harmonic). While the analytical approach for complex geometries, non-standard excitations that would also consider influences of temperature-,
humidity- and time do not exist. The inevitable presence of experimental errors at material characterization or noise in signal during the structural health monitoring of real structures makes this inverse problem additionally ill-posed [7,8].

Classical methods for solving ill-posed inverse problems, such as fitting [2] and regularization methods [8,9] provide solutions in a single point and require additional utilization of finite element models to determine the mechanical state of the whole structure. Thus, using numerical procedures for monitoring complex structures in real-time is computationally very demanding.

To deal with modelling of complex structure response made from polymeric materials to non-standard excitation with consideration of viscoelasticity and temperature-, humidity- and time-dependence neural networks (NN) could be used. NN are ideal for modeling processes that are complex and hard to describe mathematically. They are also ideal for modeling processes where we do not entirely understand the physical interactions between different processes. The advantage of NN is also that once trained, networks deliver results fast with short computational times. Another positive feature of NN is their resistance to noise in the signals (robustness) and generalization capabilities [10-12].

Due to their advantages, NNs are utilized for SHM application already for some time. The review article of Hany El Kady [13] provides a comprehensive overview of publications on the modelling of mechanical behavior of fiber-reinforced composites using neural networks before 2005. The review paper shows that artificial NN can lead to predictions as accurate as, if not better, than those obtained by the conventional methods.

In more recent papers, researchers investigated NN’s applicability to detect delamination of composites based on different input signals. Zhang et al. [14,15] determined the location and severity of composite damage based on measured changes in frequency response using three inverse algorithms, including NN. Delamination detection based on noise polluted natural frequencies measurements were done by Ihesiulu for the composites using SHM [21]. Authors propose to weight neural inverse algorithms, including NN. Delamination detection based on damage based on measured changes in frequency response using three and size prediction performance.

To detect vibration-based damage, one-dimensional convolution NN was utilized for the real-time health monitoring systems [18]. Verified experimentally on the simulator 1D convolution NN showed superior ability to extract optimal features from the accelerometer data. This research was limited to detect only slight damage.

Convolution NN was used in civil engineering applications more extensively in the past years to determine the structures’ mechanical state. NNs were utilized for denoising the vibration signals from complex civil constructions [19] and for damage detection [20]. Effects of uncertainties and defects in the manufacturing process were also addressed for the composites using SHM [21]. Authors propose to weight neural network inputs with signal-to-noise ratios to improve damage location and size prediction performance.

Even though several authors mention and emphasize the importance of the environmental conditions and their nonlinear effect on composite structural behavior, only a few studies were published that take this into account. Among recent studies, Hsu and Loh [22] report on nonlinear environmental effects and emphasize the importance of temperature effects. Xia et al. showed a 10% difference in the reinforced concrete slab’s natural frequencies with temperature [23]. They propose the utilization of nonlinear principal component analysis using autoassociative neural network to extract the environmental factors. Another example is given by Zhou and Ko [24] where the temperature-caused vibration modal variability of a bridge is investigated using NN.

While these studies show the importance of environmental factors in civil engineering, this is also very important to consider, for example, in aerospace engineering structures made of composite material. In such applications, parts and elements are exposed to extreme temperature and humidity changes during their use. For example, it is reported that temperatures of the planes outer layer can change for up to 100 °C for a short period [25].

In this respect, this paper aims to investigate the applicability of NN for the SHM system and predict the viscoelastic material response at different temperatures. Specifically, a multilayer perceptron neural network was used for modelling the creep behavior of PEEK material under constant shear stress rate excitation at different temperatures. Within the paper, the optimal NN topology, the effect of the amount of training data and its distribution in a temperature range on prediction quality were investigated.

In our previous work [7], we demonstrated that the neural network could solve the inverse problem and obtain the material transfer function based on excitation and response in the presence of noise. Additionally, we have shown that NN could outperform methods that are usually used to solve the inverse problem. However, we did not consider temperature effects, and the analysis was made using artificially generated data. In contrast, this research aims at the prediction of material response at different temperatures. It is determined directly by assuming that the neural network represents the material function. The analysis is entirely based on experimental data.

2. Materials and methods

2.1. Material

For the investigation, PEEK material was used (CENTROPEEK, Centroplast Engineering Plastics GmbH, Germany). The material was supplied in 2 mm thick sheets machined into rectangular samples with dimensions 2 × 10 × 50 mm. Before measurements, samples were annealed to remove residual stresses and avoid physical ageing during testing. Annealing was performed at 250 °C for 9 h. Afterwards, samples were cooled to the ambient temperature with a cooling rate of 0.1 °C/min.

The DSC analysis measured at 10 °C/min was conducted on the annealed samples, where the first heating cycle was used to determine the material’s thermal properties. The material showed the glass transition onset temperature at 148.19 °C and glass transition temperature Tg = 152.32 °C. The material exhibited a double melting peak. The first one, a smaller one at Tm1 = 263.50 °C (∆Hm1 = 8.59 J/g) and at Tm2 = 338.97 °C (∆Hm2 = 38.93 J/g). The degree of crystallinity of the samples was χ = 36.56% (100% ∆Hm = 130 J/g [26]).

2.2. Experimental material characterisation

For shear stress-strain measurements which were used for NN training and prediction, constant shear stress rate experiments were performed at different temperatures on a rotational rheometer (MCR 702, Anton Paar, Austria). Constant shear stress rate excitation was selected as one of the possible (non-standard) loading profiles to which composite structures are exposed, for example, in aerospace applications. The constant stress rate was set to 1000 Pa/s, and experiments

### Table 1

| Temperatures [°C] | 130 | 140 | 145 | 150 | 155 | 160 | 165 | 170 | 180 | 190 | 200 |
|-------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Values            |     |     |     |     |     |     |     |     |     |     |     |
were limited to 300 s (maximal shear stress 0.3 MPa). Within this time, 1000 points were recorded. For material characterization, 11 different temperatures were selected, Table 1.

Three repetitions of stress-strain measurements at different temperatures were made on three different samples. The samples were first tempered at 130 °C for 5 h. Afterwards, tests at each temperature consisted of 1 h of temperature stabilization, followed by shear stress-strain measurement (300 s). After each loading, samples were unloaded (the applied moment on the samples was set to 0), and the temperature was raised following the temperature profile in Table 1. During the temperature stabilization time (1h) residual stresses caused by loading were minimized, so there was no effect of the previous loading in the next loading cycle (at the next temperature). During the complete measuring procedure, the normal force was set to 0 N to prevent normal stresses in the sample because of thermal expansion or shear loading.

Shear stress was calculated following equation (1):

\[
\tau(t, T) = \frac{M}{W_t},
\]

(1)

where \(M\) is applied moment and \(W_t\) is the torsional section modulus. The shear strain was determined using equation (2):

\[
\gamma(t, T) = \frac{I_t \phi}{W_t L},
\]

(2)

where \(I_t\) is the torsional moment of inertia, \(\phi\) is measured deflection angle, and \(L\) is the sample length.

### 2.3. Neural network design

Neural networks are distributed parallel processors with high computational power. Their correct use and proper training can be challenging since there is no well-established procedure for designing a network for a particular case. The procedure depends on the amount of available data, quality of the data, number of inputs and outputs, training algorithm and its parameters, as well as desired result quality.

Additionally, there are also no standardized methodologies to determine the topology of the network, which includes an optimal number of neurons in each layer and the number of layers. Therefore, the developed algorithm based on general guidelines for NN design is presented in continuation for the selected Multilayer Perceptron (MPL) NN. It consists of two main parts: training procedure and topology selection and optimization.

#### 2.3.1. Training procedure

Before the training procedure, experimental data were pre-processed for implementation to NN. Pre-processing of the data includes rescaling and normalization to mean to increase fitting precision. Rescaling was performed by the neural network function `fitnet`, while normalization was done by mapstd MATLAB function. The pre-processing of input data is required to avoid the saturation of activation functions operating with a too large or too small value. Output data is transformed back to the original scale in the post-processing step.

The goal of the training procedure is to determine optimal weights for NN. For MPL NN, the feedforward supervised learning is typically used with the backpropagation algorithm [27]. Starting with some initial values, the algorithm iteratively determines optimal weights by propagating error between the current result and target value from the output layer to input.

Based on the amount of experimental data and recommendations from the literature [28], function `trainbr` was used for training. The `trainbr` is based on the Levenberg-Marquardt algorithm procedure and is upgraded with Bayesian regularization. This training algorithm has shown good results for solving inverse problems [7] and is appropriate for the data comprised of multiple data points, as it has an embedded mechanism to prevent overfitting. During training, the Bayesian regularization algorithm minimizes squared error and the sum of weights and prevents overfitting [29], which is shown in the low generalization performance of NN. `Trainbr` uses Jacobian for computation and assumes performance in the form of the mean or sum of squared errors. `Trainbr` algorithm was selected for the investigation with mse performance functions. Default training parameters were utilized.

Proper selection of training data distribution is crucial for NN training quality. In our case, experimental stress-strain curves at different temperatures can be used for training. Considering outstanding interpolation and weak extrapolation MLP properties, the first and last temperature was always included in the training data set. Other training data sets can be distributed differently. Moreover, since one of the sub-goals is to check how many data sets and the distribution of data sets in temperature range are necessary for a successful prediction of material behavior, several different data combinations were used.

Generally, NN requires data for training, validation and testing. Validation data is typically used for stopping the training algorithm. Test data is used to check NN performance after training and determines the generalization abilities of NN. The training function used in our investigation, `trainbr` does not require validation data set as it uses a built-in regularization procedure (to prevent overfitting). Additional inclusion of validation step could prevent the training algorithm from exploring large weights [28].

By default, the MATLAB training function splits the input data randomly in a 70%-15%-15% ratio (training-validation-test). However, as there is no need for validation, the data was split into different sets corresponding to a particular temperature using the function `divideind`. This function allows setting particular data for each group. By varying the ratio between training (tr) and testing data (test), the effect of training data on the test date (i.e. prediction capabilities) was verified. Different data sets (Training set) used in this investigation are presented in Table 2.

The following logic was used for selecting data sets:

- **Set 0** used all available experimental data for training and was selected as an extreme (reference) case.
- **Set 1** excludes two temperatures from training.
- **Set 2** excludes every second temperature from training.
- **Set 3** uses more training data from the central part of the experimental data (in the middle of the testing temperature range).

### Table 2

| Training Set | Temperature [ °C] |
|---------------|-------------------|
| 0 (ref)       | 130 140 145 150 155 |
| 1             | 160 165 170 180 190 200 |
| 2             | 0 (ref) 130 140 145 150 155 160 165 170 180 190 200 |
| 3             | 0 (ref) 130 140 145 150 155 160 165 170 180 190 200 |
| 4             | 0 (ref) 130 140 145 150 155 160 165 170 180 190 200 |
| 5             | 0 (ref) 130 140 145 150 155 160 165 170 180 190 200 |
| 6             | 0 (ref) 130 140 145 150 155 160 165 170 180 190 200 |
The first four sets were selected for the systematic investigation of the effect of the number of training data. Simultaneously, the selection of the last three sets was based on the physical nature of the PEEK material. As the creep process is faster above the glass transition temperature, training data sets were selected in relation to the $T_g$ location.

For NN implementation, neural network fitting function fitnet was selected as it is used for regression and curve fitting applications [28].

### 2.3.2. Neural network topology selection and optimization

The topology of a neural network determines inputs and outputs and the number of variable parameters in the system. The number of layers in NN, how they are interconnected, and the number of neurons in each layer affect the total number of parameters representing the network computation ability.

The input-output structure was defined as stress values for inputs and strain values for outputs for our investigation. Input stress data was implemented as vectors “in series”, where each data point in time was provided to the NN one after the other (Fig. 1). Additionally, the temperature was set as another input parameter as a time-varying variable. However, in our case, the temperature at the individual measurement was practically constant (max. temperature deviation at individual temperature was about $\Delta T = 0.1 ^\circ C$). By not considering temperature as a constant parameter, we mimic realistic conditions where rapid temperature changes are likely to happen.

The topology of the MLP neural network typically consists of at least 3 layers. Input layer that introduces input data into the system. One or more hidden layers consisting of neurons with sigmoidal activation functions (hyperbolic tangent or logistics function) that determine the computation power of NN. And the output layer with linear activation function neurons to provide the output. According to Universal Approximation Theorem, MLP with one hidden layer is capable of approximation of any continuous function for inputs within a defined range. Therefore, only one hidden layer was considered. A number of neurons in the hidden layer was varied to determine the optimal topology for each training data set.

Mathematically, the output at a particular time of the selected topology can be represented as:

$$\gamma = \sum_{i=1}^{z} w_{ik} f(w_{ij} \cdot \tau + w_{ij} \cdot T + c) ,$$

where $z$ is the number of neurons in the hidden layer indexed by $i$. Synaptic weights are represented by $w$ with two indices, where the first one presents the neuron accepting the signal for and the second – the neuron sending the signal. $c$ denotes biases. Function $f(x) = \frac{1}{1 + e^{-x}}$ is an activation function of argument $x$.

The topology optimization methodology proposed within this research is an upgrade of the methodology presented in our previous work [7]. While before the initialization of NN, weights was fixed to one value to avoid variations during multiple NN training runs, now the effect of initialization randomness was considered by running the training procedure 10 times for each tested topology. Initialization was done by the Nguyen-Widrow initialization algorithm [30].

Topology variations are limited to 1 hidden layer and maximum $z = 50$ neurons in it. For each topology, training algorithms is started 10 times, and the following performance parameters are recorded:

1. Mean square error, $MSE$, [Pa$^2$] is calculated as follows:

$$MSE = \frac{\sum_{j=1}^{n} (\gamma_j - \gamma_{j\text{target}})^2}{n} .$$

2. $R_{99.05}$ is defined as a percentage of the number $n_{99.05}$ of data points of the $\gamma_j(t)$ curve that is estimated with more than 5% relative error. This value of 5% relative error was selected as a maximal acceptable error for engineering purposes. The data points of a curve $\gamma_j(t)$ which are predicted with more than 5% relative error satisfy the following condition:

$$\gamma_i: \frac{(\gamma_j - \gamma_{j\text{target}})}{\gamma_{j\text{target}}} > 0.05 , i = 1..n$$

The performance parameter $R_{99.05}$ of modelled specific $\gamma_j(t)$ curve is determined as:

$$R_{99.05} = \frac{n_{99.05}}{n} \cdot 100\% ,$$

where $n$ is the number of data representing the $\gamma_j(t)$ curve. The modelling performance of a particular NN is then characterized as the average value $R_{99.05}$ of the $R_{99.05}$ with respect to the number of all analyzed testing temperatures. The lower $R_{99.05}$ is, the more data points are predicted inside of 10% error tube.

3. $E_{\text{max}}$ [%] is a maximal relative error calculated for every strain curve (every temperature), which is determined for the $j$-th curve as:

$$E_{\text{max}} = \max_i \left( \frac{(\gamma_j - \gamma_{j\text{target}})}{\gamma_{j\text{target}}} \right) \cdot 100\% , i = 1..n$$

All three selected performance parameters characterize the reconstructed strain curve from different points of view: MSE provides the average error level throughout the curve, $R_{99.05}$ determines the number of outliers and $E_{\text{max}}$ defines the amplitude of those outliers.

The optimization criterion $C$ combines all three performance parameters as median normalized values average over temperatures (marked with the apostrophe):
The optimization criterion considers the effect of the initialization randomness of NN. The criterion itself represents the geometrical distance between the specific NN topology (topology n) and the ultimate topology calculated in the 3D Euclidean space with performance parameters as the axis. The schematic is shown in Fig. 2.

3. Results and discussion

Fig. 3 shows the averaged values of stress-strain data for each temperature. The error bar in the diagram indicates the maximal deviation from the average value for all temperatures. The average relative experimental error ranged between 0.5 and 4% depending on temperature (excluding the initial transient response). Only each 20th data point is shown in Fig. 3 for clarity.

As mentioned in the methodology and following Table 2, part of the measured segments was used for NN training and part as test data to check the generalization capabilities of NN.

Fig. 4 demonstrates normalized input and output data in the form of a 3D diagram to show how data is distributed in the space and forms a nonlinear surface (again, only every 20th point is plotted for clarity). It is visible that the curves are intentionally more densely distributed in the central part of testing temperatures – around glass transition temperature \( T_g = 152.32 \) °C, since more rapid changes in material behavior are expected in that region.

The optimal topology for every training data set was determined as described in the methodology to evaluate the NN performance and its dependence on the training data number and distribution. Results are shown in Table 3. The values of performance parameters are average values obtained from NN output from the whole data set provided as input. Therefore, it considers both training performance on the training data sets and generalization on the test data sets.

The smallest value of optimization criterion, \( C \), which combines all three performance parameters is observed for the data set 0. The data set 0 was selected as an extreme case since all the data was used for training. In this respect, the result is expected as the maximal available amount of training data was provided, and overfitting was avoided using the Bayesian regularization training algorithm. Therefore, this data set will not be investigated further but considered as a reference for comparison with other data sets.

The largest value of optimization criterion \( C \) can be seen for data set 4, which covers more data segments below \( T_g \). It also required the largest number of hidden neurons for fitting the data. This could be related to the training data concentration in one part of the training data space and, therefore, less successful training compared to other data sets. On the other hands, the smallest value of \( C \) corresponds to the data set 6. In the case of data set 6, more data in \( T_g \)'s vicinity was used for training.

While optimization criterion unites all NN performance parameters, different data sets correspond to different data distribution in the temperature domain, the number of training data points directly affects the training quality, and therefore all these parameters should be analyzed together. In practice, the smallest amount of training data that still provides reliable results is beneficial as it leads to fewer experiments.

Both the optimization criterion, \( C \) and number of training data points are...
MSE but cause more outliers (lower required (6 segments, 1000 points/temperature), while at the same shown in Fig. 5 for each training set. In data set 6, 6000 data points are shown in Fig. 5. Performance parameters of optimal topologies.

| Training set | 0 (ref) | 1 | 2 | 3 | 4 | 5 | 6 |
|--------------|--------|---|---|---|---|---|---|
| Topology     | [7]    | [6] | [8] | [8] | [27] | [5] | [6] |
| Criterion C², [−] | 0.0221 | 0.0384 | 0.0528 | 0.0357 | 0.0907 | 0.0745 | 0.0240 |
| Average MSE, [−] | 1.70·10⁻¹² | 1.82·10⁻⁹ | 4.71·10⁻¹⁰ | 1.66·10⁻¹⁰ | 2.32·10⁻¹⁰ | 1.04·10⁻¹⁰ | 1.80·10⁻⁹ |
| Average R₀,95, [%] | 14.97 | 24.86 | 27.92 | 20.01 | 21.43 | 47.56 | 34.15 |
| Average Eₘax, [%] | 697.62 | 172.34 | 1979.36 | 533.35 | 434.32 | 310.52 | 362.42 |
| No. training data points | 11000 | 9000 | 6000 | 7000 | 6000 | 5000 | 6000 |

Fig. 5. Optimization criterion values, C (left axis) and a number of training data points (right axis) for each data set.

shown in Fig. 5 for each training set. In data set 6, 6000 data points are required (6 segments, 1000 points/temperature), while at the same number of data points, data sets 2 and 4 did not result in such small optimization criterion values. Less training data points have been used only for data set 5, however, it showed a higher optimization criterion value, C.

Comparing data sets 6 and 2 with the same number of training data points shows that the distribution of data in temperature domain plays an important role in NN performance. Alternating the training and test segments with temperature (set 2) provides a lower average value of MSE but cause more outliers (lower R₀,95) and in total results in a higher value of optimization criterium. Data set 6 provides more training data condensed around glass transition temperature, where more pronounced physical changes to material start to appear. The results indicate that this phenomenon should be considered while designing NN.

Generally, relatively high values of R₀,95 shown in Table 3 are related to the outliers. Detailed inspection of reconstructed curves shows that those outliers occur mainly at the beginning of the stress-strain curve and result in high values of Eₘax. With prior knowledge of the investigated system, these errors could be avoided by setting first data points to zero automatically but within this investigation this was not considered.

3.2. Effect of the number of training data and topology

NN performance is affected not just by the training data set corresponding to the total number of data points but also by the number of neurons in the particular topology. A number of neurons in the single hidden layer of MLP that was calculated by the optimization algorithm determines the computational power of NN as it directly affects the number of NN adjustable parameters (weights and biases). In principle, the higher the number of neurons, the better the network performance is, unless overfitting phenomena affects the NN performance.

Therefore, in Fig. 7, MSE and R₀,95 in dependence on the number of training data points, and the number of neurons in optimal topology is presented. Maximal error is not considered for this analysis. The surface in Fig. 7 was obtained by "natural" interpolation using gridfit MATLAB function combined with surf function. Markers in Fig. 7 are the values from Table 3 and are used for interpolation.

As expected, the smallest number of training data points shows the worst performance as the MSE and R₀,95 values are the highest. On the other hand, using the whole data set for training yields the best performance without overfitting. Fig. 7 also demonstrates the general trend that both performance parameters decrease (improving performance) with the increase of number of training data points and the number of hidden neurons. The difference between performance parameters can be observed at the smaller number of neurons. While MSE decreases faster as the number of neurons is smaller, the R₀,95 shows a more gradual training (test data). In all parts of Fig. 6, values obtained from data set 0 are shown with the dashed line as a reference value obtained when all experimental data was used for training.

Following conclusions can be made:

- Data set 0 delivers the smallest MSE with the smallest variations. The error does not significantly change with temperature. This result is expected since all the experimental data was used for training.
- Generally, data sets with a smaller amount of training data show better performance for the training data, while worse results for tests data. The error increases with the temperature; this could be attributed to a higher rate of material properties change at higher temperatures.
- The most homogeneous distribution of MSE in respect to temperature was shown for Data set 2, where training and tests data sets were alternating with each temperature. The largest error corresponds to the highest test temperature.
- The smallest MSE values regarding temperature are shown for data set 4, where most data corresponding to temperatures below Tg was used for training. In this case, errors can be up to 2 orders of magnitude smaller than reference test 0. Non-homogeneous training data distribution resulted in a maximal number of neurons in the hidden layer (compared to other data sets). However, as discussed, data set 4 resulted in the largest optimization criterion, C and required the maximal number of hidden neurons.
- Data sets 5 and 6 correspond to the cases where training data is condensed above and in Tg’s vicinity. In both cases, the MSE values for training data are comparable with the reference 0 data set. The generalization errors for data sets 5 and 6 are comparable.
Fig. 6. Average MSE value with respect to temperature.

Fig. 7. Performance parameters a) log MSE and b) $R_{0.95}$ as a function of the number of training data points and the number of hidden neurons of optimal MLP topology.
3.3. Prediction quality

To demonstrate how well the NN predicts the material behavior a direct comparison was made between the experimentally determined stress-strain curves at different temperatures and predicted curves. The analysis showed that the best prediction in terms of optimization criterion, $C$ is given by the Data set 6. Therefore, Fig. 8 depicts the predicted stress-strain curves using Data set 6. Data used for training is shown in a) and for testing in the Fig. 8 b). The lines in the figure showed averaged measured values, while the symbols show the NN predicted values (for clarity, only every 20th point is shown).

From Fig. 8 it can be seen that NN can predict the stress-strain behavior of a material at different temperatures within the experimental error.

4. Conclusions

Within this paper, the prediction of stress-strain curves of PEEK material at different temperatures using the neural network was addressed. For best prediction quality, the optimal NN topology, effect of the amount of training data and training data distribution (or temperatures used for training) were analyzed. The prediction quality was evaluated using three performance parameters: mean square error, $MSE$; parameter $R_{0.95}$; and maximal relative error, $E_{max} [\%]$, combined into optimization criterion $C$.

We have shown that once properly trained, the MLP neural network can predict the stress-strain behavior of polymeric material within the experimental error ranging from 0.5 to 4 \%. NN also demonstrated good generalization capabilities and enabled good prediction at temperatures where stress-strain behavior was not experimentally determined (within experimental error).

In respect to the NN performance, the following specific conclusions can be made:

1) Training data distribution

The best performance in terms of the introduced optimization criterion was shown by the data set 6, which utilized more training data concentrated in the vicinity of the material’s glass transition temperature. This indicates that the physical background of the modelled process should be considered while selecting training data for NN.

It was also shown that as the temperature increased, the prediction quality worsened for all data sets. For example, the MSE parameter for data set 6 increased from minimal value of approximately $10^{-12}$ to maximum of approximately $10^{-8} \text{MPa}^2$. This might be attributed to the higher rate of change at higher temperatures (especially above $T_g$). A higher rate of changes in response is harder to model using NN; therefore, while planning experiments, more data should be available at temperatures corresponding to thermal transitions.

2) Amount of training data and NN topology

All the data sets except one required less than 10 neurons in the hidden layer while trained with the Bayesian regularization method to avoid overfitting. The data set with a maximal number of hidden neurons (27) can be correlated with inhomogeneous training data distribution covering more data below $T_g$ and the highest among the tested value of optimization criterion requires further investigation for further training procedure adjustment. A small number of hidden neurons, in general, is an encouraging fact since no overfitting signs were detected. From a practical point of view, smaller training data sets are desirable. However, a detailed analysis shows that better performance is related to the maximal number of hidden neurons and larger training data set as evident from Fig. 7. Thus, the balance between extended topology and training data provided to the network should be considered.

The findings of this research indicate promising applications of NN generalization capabilities. For example, to reduce the number of needed experiments for material characterization and long-term viscoelastic material prediction following the procedure elaborated, for example, in [2]. The methodology can also be expanded to study environmental effects and the influence of (macro and nano) fillers.

In another example NN can be used for structural health monitoring, since it as a model unites material properties, geometry, and environmental effects. A network upgraded with harmonic loading could be utilized by existing SHM systems for high performance composite structures at data fusion, cleansing, and partially at feature extraction stages of the SHM process. Since the model does not require the development of a specific measurement sensor system, the existing ones can be utilized. For example, fiber Bragg grating (OFBG) [31], polyvinyl fluoride (PVDF) sensors [32], self-sensing carbon fiber and nano-composites [33], vibration sensors or piezo-electric wafer sensors (PWAS).

Data availability

The raw data required to reproduce these findings are available to download from: https://doi.org/10.17632/79387sp2mg.1.

Declaration of competing interest

The authors declare that they have no known competing financial
interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was financially supported by the Slovenian Research Agency in a frame of postdoctoral project Z2-1865 and research core funding No. P2-0264.

References

[1] P.E. Irving, C. Souissi (Eds.), Polymer Composites in the Aerospace Industry, Woodhead Publishing, 2015.
[2] L. Sapunov, M. Gergesova, I. Emri, Prediction of viscoelastic material functions from constant stress- or strain-rate experiments, Mech. Time-Dependent Mater. 18 (2014) 349–372, https://doi.org/10.1016/j.mtdm.2013.09.003.
[3] W.G. Knauss, J. Zhao, Improved relaxation time coverage in ramp-strain histories, Mech. Time-Dependent Mater. 11 (2007) 199–216, https://doi.org/10.1007/s11043-007-9055-3.
[4] H. Lu, B. Wang, J. Ma, G. Huang, H. Viswanathan, Measurement of creep compliance of solid polymers by nanoinstrumentation, Mech. Time-Dependent Mater. 7 (2003) 189–207, https://doi.org/10.1016/S1476-3816(03)00014-4.
[5] Y.G. Yanovsky, Y.A. Basistov, A new method of calculation of polymer media approximation by neural networks, Neural Network. 16 (2003) 79–90, https://doi.org/10.1016/S0893-6080(03)00167-5.
[6] Pardo, P. Tomba, J. Carella, A solution for an Ill-conditioned inverse problem. A approximation by neural networks, Mech. Time-Dependent Mater. 21 (2017), https://doi.org/10.1007/s11043-016-9332-x.
[7] A. Aulova, E. Govekar, I. Emri, Determination of relaxation modulus of time-dependent materials using neural networks, Mech. Time-Dependent Mater. 21 (2017), https://doi.org/10.1007/s11043-016-9332-x.
[8] A. Tikhonov, V. Arzenin, Solutions of Inverse Problems, John Wiley & Sons, Washington D.C., 1977.
[9] M. Burger, A. Neubauer, Analysis of Tikhonov regularization for function approximation by neural networks, Neural Network. 16 (2003) 79–90, https://doi.org/10.1016/S0893-6080(03)00167-5.
[10] J. Zhong, C. Yang, W. Ma, Z. Zhang, Long-term creep behavior prediction of polymethacrylimide foams using artificial neural networks, Polym. Test. 93 (2021) 69–80, https://doi.org/10.1016/j.polymertesting.2020.106893.
[11] L.F. Simões Hoffmann, F.C. Parquet Bizarria, J.W. Parquet Bizarria, Detection of liner surface defects in solid rocket motors using multilayer perceptron neural networks, Polym. Test. 88 (2020) 106559, https://doi.org/10.1016/j.polymertesting.2020.106559.
[12] F. Lo Savio, G. Capizzi, G. La Rosa, G. Lo Sciuto, Creep assessment in hyperelastic material by a 3D Neural Network Reconstructor using bulge testing, Polym. Test. 63 (2021) 65–72, https://doi.org/10.1016/j.polymertesting.2021.08.009.
[13] H. El Kadi, Modeling the mechanical behavior of fiber-reinforced polymeric composite materials using artificial neural networks - a review, Compos. Struct. 73 (2006) 1–23, https://doi.org/10.1016/j.composites.2005.01.020.
[14] Z. Zhang, K. Shankar, T. Ray, E.V. Morozov, M. Tahtali, Vibration-based inverse algorithms for detection of delamination in composites, Compos. Struct. 102 (2013) 226–236, https://doi.org/10.1016/j.composites.2013.03.012.
[15] Z. Zhang, M. He, A. Liu, H.K. Singh, K.R. Ramakrishnan, D. Hui, K. Shankar, E.V. Morozov, Vibration-based assessment of delaminations in FRP composite plates, Compos. B Eng. 144 (2018) 254–266, https://doi.org/10.1016/j.compositesb.2018.03.003.
[16] O.K. Besluice, K. Shankar, Z. Zhang, T. Ray, Delamination detection with error and noise polluted natural frequencies using computational intelligence concepts, Compos. B Eng. 56 (2014) 906–925, https://doi.org/10.1016/j.compositesb.2013.09.032.
[17] G. Sarasangapani, R. Ganguli, Detection of local matrix cracks in composite beam using modal data and modular radial basis neural networks, ISSS J. Micro Smart Syst. 6 (2017) 127–140, https://doi.org/10.1016/j.iims.2017.01.015.
[18] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, D.J. Inman, Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks, J. Sound Vib. 388 (2017) 154–170, https://doi.org/10.1016/j.jsv.2016.10.045.
[19] G. Fan, J. Li, H. Hao, Vibration signal denoising for structural health monitoring by residual convolutional neural networks, Meas. J. Int. Meas. Confed. 157 (2020) 107651, https://doi.org/10.1016/j.measurement.2020.107651.
[20] J. Yang, L. Zhang, C. Chen, Y. Li, R. Li, G. Wang, S. Jiang, Z. Zeng, A hierarchical deep convolutional neural network and gated recurrent unit framework for structural damage detection, Inf. Sci. (Ny) 540 (2020) 117–130, https://doi.org/10.1016/j.ins.2020.05.090.
[21] H. Teimouri, A.S. Milani, R. Seethaler, A. Heidarzadeh, On the Impact of Manufacturing Uncertainty in Structural Health Monitoring of Composite Structures : A Signal to Noise Weighed Neural Network Process, 2016, pp. 26–39, https://doi.org/10.4236/ijst.2016.61004.
[22] T.-Y. Hsu, C.-H. Lo, Damage detection accommodating nonlinear environmental effects by nonlinear principal component analysis, Struct. Constr. Health Monit. 17 (2010) 338–354, https://doi.org/10.1002/stc.320.
[23] Y. Xia, H. Hao, G. Zanardo, A. Deeks, Long term vibration monitoring of an RC slab: temperature and humidity effect, Eng. Struct. 28 (2006) 441–452, https://doi.org/10.1016/j.enganstruct.2005.09.001.
[24] H.F. Zhou, Y.Q. Ni, J.M. Ko, Constructing input to neural networks for modeling temperature-caused modal variability: mean temperatures, effective temperatures, and principal components of temperatures, Eng. Struct. 32 (2010) 1747–1759, https://doi.org/10.1016/j.enganstruct.2010.02.026.
[25] J.A. Hough, S.K. Karad, F.R. Jones, The effect of thermal spiking on moisture absorption, mechanical and viscoelastic properties of carbon fibre reinforced epoxy laminates, Compos. Sci. Technol. 65 (2005) 1299–1305, https://doi.org/10.1016/j.compscitech.2005.01.009.
[26] D.J. Blundell, B.N. Osohm, The morphology of poly(aryl-ether-ether-ketone), Polymer 24 (1983) 953–958, https://doi.org/10.1016/0032-6797(83)90144-1.
[27] P. Werbos, The Roots of Backpropagation: from Ordered Derivatives to Neural Networks and Political Forecasting, John Wiley & Sons, Inc., 1994.
[28] M. Hudson, B. Martin, T. Hagan, H.B. Demuth, Deep learning Toolbox™ user’s guide, n.d. www.mathworks.com (accessed March 3, 2021).
[29] D.J.C. MacKay, A practical bayesian framework for backpropagation networks, Neural Comput 4 (1992) 448–472.
[30] D. Nguyen, B. Widrow, Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights, in: ICCN Int. Jt. Conf. Neural Networks vol. 3, 1990, pp. 21–26.
[31] N. Gupta, M.J. Augustin, S. Sathy, S. Jain, S.R. Viswamurthy, K.M. Gaddikeri, R. Sundaram, Structural health monitoring of composite aircraft structures using fiber bragg grating sensors, J. Indian Inst. Sci. 93 (2013) 735–750, https://doi.org/10.1007/s10934-013-0051-6.
[32] A. Chukwujevsku Okafor, K. Chandranekhar, Y.P. Jiang, Delamination prediction in composite beams with built-in pieoelectric devices using modal analysis and neural network, Smart Mater. Struct. 5 (1996) 338–347, https://doi.org/10.1088/0964-1726/5/3/012.
[33] Y. Li, V. Liao, Z. Su, Graphene-functionalized polymer composites for self-sensing of ultrasonic waves: an initiative towards “sensor-free” structural health monitoring, Compos. Sci. Technol. 168 (2018) 203–213, https://doi.org/10.1016/j.compscitech.2018.09.021.