ANN constitutive material model in the shakedown analysis of an aluminum structure

Beata Potrzeszcz-Sut

Department of Mechanics, Metal Structures and Computer Methods,
Kielce University of Technology
Al. Tysiąclecia Państwa Polskiego 7, 25-314 Kielce, Poland
e-mail: beatap@tu.kielce.pl

Ewa Pabisek

Institute for Computational Civil Engineering
Cracow University of Technology
Warszawska 24, 31-155 Kraków, Poland
e-mail: E.Pabisek@L5.pk.edu.pl

The paper presents the application of artificial neural networks (ANN) for description of the Ramberg-Osgood (RO) material model, representing the non linear strain-stress relationship of $\varepsilon(\sigma)$. A neural model of material (NMM) is a feed-forward layered neural network (FLNN) whose parameters were determined using the penalized least squares (PLS) method. A FLNN performing the inverse problem: $\sigma(\varepsilon)$, using pseudo empirical patterns, was developed. Two models of NMM were developed, i.e. a standard model (SNN) and a model based on Bayesian inference (BNN). The properties of the models were compared on the example of a reference truss structure. The computations were performed by means of the hybrid FEM/NMM program, in which NMM developed previously described the current model of the material, and made it possible to explicitly build a tangent operator $E_t = \mathrm{d}\sigma/\mathrm{d}\varepsilon$. The neural model of material was applied to the analysis of the shakedown of load carrying capacity of an aluminum truss.

Keywords: artificial neural network, inverse problem, material modeling, finite element method, hybrid program, shakedown analysis.

1. Introduction to material modeling

The main task of material modeling is appropriate formulation of constitutive relations to predict the physical behaviour of a material. Constructing general constitutive relations is a complex issue. The process of formulating a nonlinear material model consists of several phase, which include, see [1]:

1. development of an initial concept of the model,

2. performance of laboratory tests,

3. mathematical expression of the structure of the model,

4. identification of the model parameters,

5. checking the quality of the model.

The purpose of the conceptual phase of model formulation is to juxtapose all the requirements of the model resulting from its future application, e.g., range of loading of benchmark structure, scale of observation (micro or macro) as well as specification of the variety of factors including the material behaviour, such as isotropic reinforcement, Bauschinger's effect, cracking, etc.

The main objective of laboratory tests is to investigate the relationships between control variables and physical permissible (measurable) variables of state.

A mathematical model of the material is a set of information about its physical properties expressed using formalised mathematical description. In a general case, the material parameters are unknown therefore they must be determined in the process of calibration by means of the data obtained from experiment.

The identification of the material model parameters or the material model is based on solving an inverse problem and it can be done by means of ANN.

2. Application of neural networks for modeling of materials

Neural networks enable solving various inverse problems, allowing simultaneous calculation of many material parameters, see e.g. [2–4].

This approach uses a neural network FLNN, (also known as a multilayer perceptron (MLP)) developed in the off-line mode using the data sets derived from laboratory experiments. Neural networks also enable taking into account the knowledge of experts from the scope of constitutive relationship modeling.

The papers [5, 6] show that neural networks, treated as universal approximators of non-linear functions, may be used to describe the constitutive equations of material. In this case, the material stiffness matrix results from the neural model of the material (NMM), in the form of the mapping realized by the neural network:

$$\mathbf{y} = \text{NMM}(\mathbf{x}) \tag{1}$$

where $\mathbf{x}$, $\mathbf{y}$ – input and output vectors.

The paper presents and compares the properties of two types of ANN: 1) standard neural network (SNN) and 2) FLNN with Bayesian inference (BNN). Both networks were used to identify the material model of an aluminum structure.

3. Ramberg-osgood material model

Ramberg-Osgood (RO) model of material allows one to extend the linear stress-strain relationship for the scope of plastic deformation. This law assumes that plastic deformation develops from the beginning of the load process. The nonlinear relationship between stress and strain $\varepsilon(\sigma)$ is described by an exponential coefficient of strengthening. Referring to the Eurocode [7], in case of materials without explicit yield stress, such as aluminum alloys, the exponential law of Ramberg-Osgood (RO) can be applied. In the example discussed in the paper, a modified RO model [5] was adopted, wherein the relationship $\varepsilon(\sigma)$ during the load process is described by a skeleton curve, see Fig. 1:

$$\varepsilon = \frac{\sigma}{E} + \frac{2.5\sigma_0}{3E}\left(\frac{\sigma}{\sigma_0}\right)^n, \tag{2}$$

where ε, σ – strain and stress, E – Young's modulus, σ_0 – yield strength, n – index exponent. The process of unloading and reloading is described by a family of hysteresis loops, cf. [8]:

$$\varepsilon - \varepsilon_R = \frac{\sigma - \sigma_R}{E} + \frac{4\sigma_0}{3E}\left(\frac{\sigma - \sigma_R}{2\sigma_0}\right)^n, \tag{3}$$

where ε_R, σ_R – strain and stress at the start of the process of unloading/reloading. The curves presented in Fig. 1 were prepared for the following data: $\sigma_0 = 71.6$ MPa, $n = 5$, $E = 59.5$ GPa.
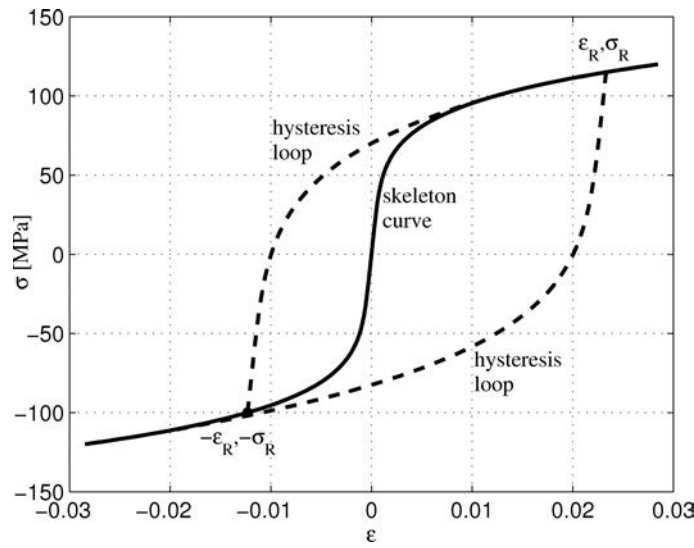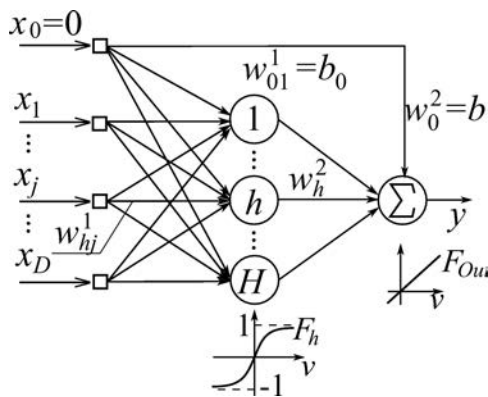
Fig. 1. Skeleton curve and hysteresis loop for RO material model.

4. Neural network for regression

In the paper, two types of neural networks were applied: 1) SNN (standard neural network), 2) BNN (Bayesian neural network). The adopted networks were used to reverse the power law of RO relationship $\varepsilon(\sigma)$, i.e., they implemented the mapping $\sigma(\varepsilon)$ during the cyclic loading of the truss. Because of the lack of experimental data, the stress and strain sets used to formulate ANN were determined using numerical simulation. For this purpose Eqs. (2) and (3), describing a skeleton curve and hysteresis loops, were used. In each case, the stress $t = \sigma$ was the output of the network, however the input $\mathbf{x}$ depended on the type of load process:

1. $\mathbf{x} = \{\varepsilon\}$ – for the skeleton curve,

2. $\mathbf{x} = \{\varepsilon, \sigma_R\}$ – for the hysteresis loop.

(4)

NMM was formulated using the Netlab toolbox for MATLAB [9] and MATLAB Neural Toolbox [10].
    Two FLNN architectures (see Fig. 2) corresponding to (4) have been adopted for further analysis.

Two FLNN architecture: D-H-1

1) $\text{FLNN}_s : D = 1, H = 5$
2) $\text{FLNN}_h : D = 2, H = 20$

Fig. 2. Architecture of FLNN.

Implemented FLNN, with H neurons in the hidden layer may be described by equation [11]:

$$y(\mathbf{x}; \mathbf{w}) = \sum_{h=1}^{H} w_h^{(2)} F_h \left(\sum_{j=1}^{D} w_{hj}^{(1)} x_j + w_{0j}^{(1)} \right) + w_0^{(2)},$$

(5)

where F_h – activation function for the neurons of the hidden layer of the FLNN network, see Fig. 2, H – the number of neurons in the hidden layer, D – number of input variables, $\mathbf{w}$ – vector of network parameters (weights and biases).

In the case of one-dimensional problem, the neural material model describes the relationship between stress and strain (5), from which the neural tangent operator E_t can be determined [12]:

$$E_t = \frac{S_\sigma}{S_\varepsilon} \sum_{h=1}^{H} w_h^{(2)} \left[1 - F_h \left(\sum_{j=1}^{D} w_{hj}^{(1)} x_j + w_{h0}^{(1)} \right)^2 \right] w_{h1}^{(1)}, \tag{6}$$

where S_σ, S_ε – scaling factors.

It is worth emphasizing that the value of neural tangent operator E_t is explicitly defined by the network parameters and the known values of input and output. In this sense FLNN representing the neural material model is no longer a "black box".

Network parameters $\mathbf{w}$ in the (5) and (6) can be determined by minimising the sum-of-squares error function with the penalty term that avoids over-fitting:

$$E(\mathbf{w}) = \frac{\beta}{2} \sum_{p=1}^{P} (y(\mathbf{x}_p; \mathbf{w}) - t_p)^2 + \frac{\alpha}{2} ||\mathbf{w}||^2, \tag{7}$$

where P is the size of the data set and α, β are hyperparameters. In the case of deterministic network, including FLNN, the parameter of regularization $\gamma = \alpha/\beta$ is applied. Such a classical (non-Bayesian) approach allows prediction of the new t value by calculating $y(\mathbf{x}; \mathbf{w})$ by means of parameter γ.

4.1. Bayesian neural network (BNN)

The Bayesian approach to neural networks learning and prediction is based on Bayesian inference. In this approach all parameters of neural model are treated as random values.

Initially, we start with a priori probability distribution $p(\mathbf{w}|\alpha)$ which expresses our knowledge of the network parameters $\mathbf{w}$ before data is observed.

Once we observe the data set $\mathbf{t}$, Bayes' theorem can be used to update our beliefs and we obtain the posterior probability density in the form:

$$p(\mathbf{w}|\mathbf{t}, \alpha, \beta) = \frac{p(\mathbf{t}|\mathbf{w}, \beta)\, p(\mathbf{w}|\alpha)}{p(\mathbf{t}|\alpha, \beta)}, \tag{8}$$

where $p(\mathbf{t}|\mathbf{w}, \beta)$ is the likelihood function, $p(\mathbf{t}|\alpha, \beta)$ is the marginal likelihood function (ML) (also called evidence for hyperparameters) obtained by integrating overweight parameters:

$$p(\mathbf{t}|\alpha, \beta) = \int p(\mathbf{t}|\mathbf{w}, \beta)\, p(\mathbf{t}|\mathbf{w}, \alpha)\, d\mathbf{w}. \tag{9}$$

Using Gaussian approximation the following expression for ML function is obtained:

$$p(\mathbf{t}|\alpha, \beta) = \left(\frac{\beta}{2\pi} \right)^{N/2} \left(\frac{\alpha}{2\pi} \right)^{W/2} \int \exp\{-E(\mathbf{w})\}\, d\mathbf{w}, \tag{10}$$

where W is the dimensionality of $\mathbf{w}$ and $E(\mathbf{w})$ is the regularized error function defined as

$$E(\mathbf{w}) \approx E(\mathbf{w}_{\text{MAP}}) + \frac{1}{2} (\mathbf{w} - \mathbf{w}_{\text{MAP}})^{\text{T}} \mathbf{A} (\mathbf{w} - \mathbf{w}_{\text{MAP}}), \tag{11}$$

$$E(\mathbf{w}_{\mathrm{MAP}}) = \beta E_D(\mathbf{w}_{\mathrm{MAP}}) + \beta E_W(\mathbf{w}_{\mathrm{MAP}}) = \frac{\beta}{2}\sum_{p=1}^{P}\{t_p - y(x_P;\mathbf{w})\}^2 + \frac{\alpha}{2}\mathbf{w}_{\mathrm{MAP}}^{\mathrm{T}}\mathbf{w}_{\mathrm{MAP}}, \qquad (12)$$

where matrix $\mathbf{A}$ is the Hessian matrix.

The optimal weights vector $\mathbf{w}_{\mathrm{MAP}}$ is computed applying the criterion of penalized least square error, i.e., $\min_{\mathbf{w}} E(\mathbf{w})$.

Taking the logarithm of the ML function (10), we have:

$$\ln\mathrm{ML} \equiv \ln p(\mathbf{t}|\mathbf{w},\alpha,\beta) \approx -E(\mathbf{w}_{\mathrm{MAP}}) - \frac{1}{2}\ln|\mathbf{A}| + \frac{W}{2}\ln\alpha + \frac{N}{2}\left(\ln\beta\ln(2\pi)\right). \qquad (13)$$

Using MML criterion which corresponds to $\max_{\mathbf{w}}\ln\mathrm{ML}$ the optimal model with H_{opt} neurons in hidden layer of BNN is chosen.

The optimal hyperparameters α and β and weights $\mathbf{w}$ can be determined in an iterative way using the evidence procedure that leads to the following formulae:

$$\alpha^{\mathrm{new}} = \frac{\gamma}{2E_W(\mathbf{w}_{\mathrm{MAP}}^{\mathrm{old}})}, \qquad \beta^{\mathrm{new}} = \frac{N-\gamma}{2E_D(\mathbf{w}_{\mathrm{MAP}}^{\mathrm{old}})}, \qquad (14)$$

where γ is related to the eigenvalues of the Hessian matrix, E_D and E_W are defined by (11). The details of the evidence approximation approach can be found in [11].

4.2. Formulation of neural material models

Two approaches to NMM selection and learning FLNN were compared using pseudo-experimental data sets.

1. In the first approach the process of formulating SNN_s: 1-5-1 was completed after $S = 20$ epochs of learning. The least squares (LS) method and the Levenberg-Marquart algorithm were applied. The values of mean square error (MSE) and average percentage error (APE) were respectively: $MSE_s = 4.08e^{-5}, APE_s = 0.01\%$. SNN_h: 2-20-1 was formulated after $S = 600$ epochs for $MSE_h = 5.68e^{-4}$, $APE_h = 0.18\%$.

   Optimal model SNN_s was obtained using $L_s = V_s = 200$ training and validating patterns whereas SNN_h model was formulated by means of $L_h = V_h = 3600$ patterns.

   Figure 3 presents the reference and predicted values of stress $\sigma(\varepsilon)$ and the tangent operators $E_t(\varepsilon)$.

2. In the second approach the process of formulating BNN_s: 1-5-1 and BNN_h: 2-20-1 was completed after 10 re-estimation steps with 900 epochs of scaled conjugate gradient algorithm and penalized least squares (PLS) method. In this case, the learning errors were respectively: $MSE_s = 1.77e^{-6}$, $APE_s = 0.11\%$ and $MSE_h = 2.60e^{-3}$, $APE_h = 2.9\%$. Both models were formulated by means of $L_s = 200$ and $L_h = 3600$ learning patterns.

   In Fig. 4 the reference and predicted values of stress $\sigma(\varepsilon)$ and tangent operators E_t are shown.

In the light of the combined results it can be concluded that both formulated networks: SNN and BNN correctly reproduce the physical RO law and E_t modulus. Tangent modulus E_t was determined directly using parameters of FLNN (8).

To formulate FLNN a large number of pseudo-experimental patterns was used. Therefore, the regularization used in Bayesian approach does not play a significant role. The formulated networks SNN and BNN have the same architecture and similar prediction properties.
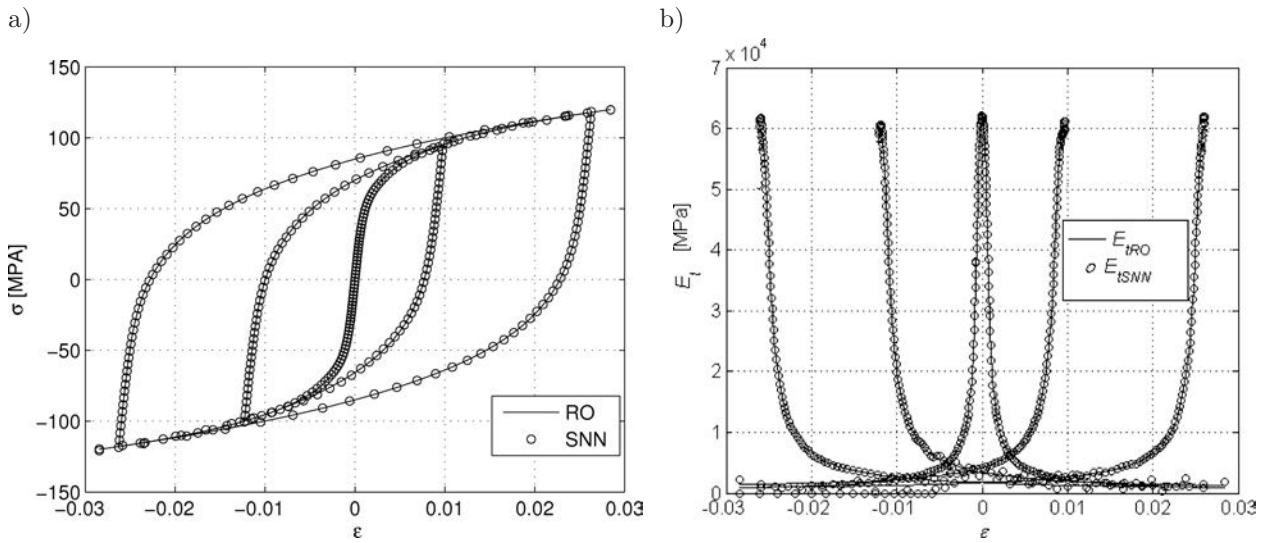
a)   b)

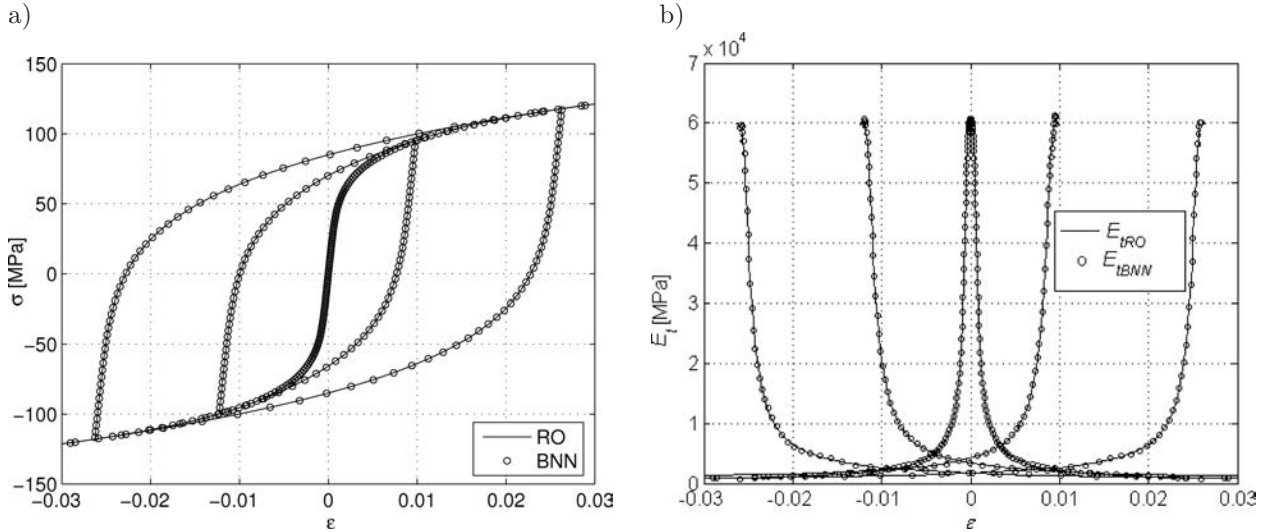**Fig. 3.** Reference and predicted values: a) of stress $\sigma(\varepsilon)$, b) of tangent operators E_t.

a)   b)

Fig. 4. Reference and predicted values: a) of stress $\sigma(\varepsilon)$, b) of tangent operators E_t.
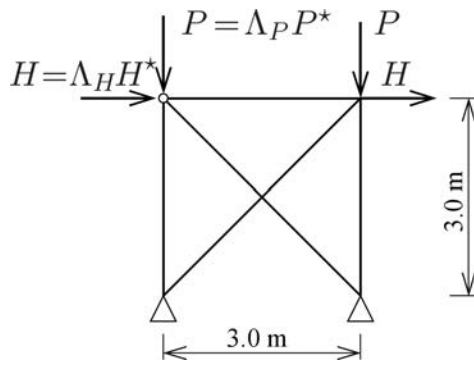
5. APPLICATION OF THE NEURAL MATERIAL MODEL IN CASE STUDIES

The neural material models SNN and BNN formulated previously have been integrated with a FEM program. This FEM/NMM hybrid program was used for the analysis of two boundary problems for a two-parameter, cyclic loading of plane trusses.

5.1. Case study 1

The comparison of the effectiveness of the formulated above NMMs has been made on the example of a five element reference truss. A static scheme of the truss is shown in Fig. 5.

The truss was subjected to loads shown in Fig. 5. It was assumed that the vertical load $\mathbf{P} = \Lambda_P \, \mathbf{P}^\star$ and horizontal load $\mathbf{H} = \Lambda_H \mathbf{H}^\star$ are treated as independent ones. The process of loading is controlled by two load factors: Λ_P, Λ_H, whose values are calculated by means of increments $\Delta_n \Lambda_P$, $\Delta_m^c \Lambda_H$, where n, m are numbers of load levels (step), number of cycle c is understood as a next

Areas of truss members:
verticals:
$A = 77.064 \cdot 10^{-4}$ m^2
horizontal and diagonals:
$A = 25.688 \cdot 10^{-4}$ m^2

Loads:
$P^\star = 50$ kN, $\quad H^\star = 25$ kN

Fig. 5. Data of reference truss structure.

repetition of the load program. The loading program can be illustrated by means of the following scheme:

$$
\begin{array}{llll}
\text{(I)} & \mathbf{P}^0 = \Lambda_P \mathbf{P}^\star & \text{for} & \Lambda_P : \; 0.0 \to 10.0 \\
\text{(II)} & \mathbf{P}^0 + (\Lambda_H \mathbf{H}^\star)^c & \text{for} & \Lambda_H^c : \; (0.0 \to 4.0 \to 0.0)^c
\end{array}
\tag{15}
$$

where $c = 1, 2, \ldots, 10$.

Determined by the test, a little more numerically effective NMM$_{\mathrm{BNN}}$ network was used for subsequent calculations. Tangent modulus E_t derived from BNN was calculated with a smaller error, thus the number of iterations in the Newton-Raphson algorithm was smaller than in the FEM/NMM$_{\mathrm{SNN}}$ program.

In Fig. 6 the equilibrium paths $\Lambda(v_2)$, $\Lambda(u_2)$ determined for node number 2 of the truss are shown. The paths plotted with solid lines were computed using the program FEM/RO, and the paths plotted with points were computed using hybrid program FEM/NMM$_{\mathrm{SNN}}$. The two paths are very close to each other.
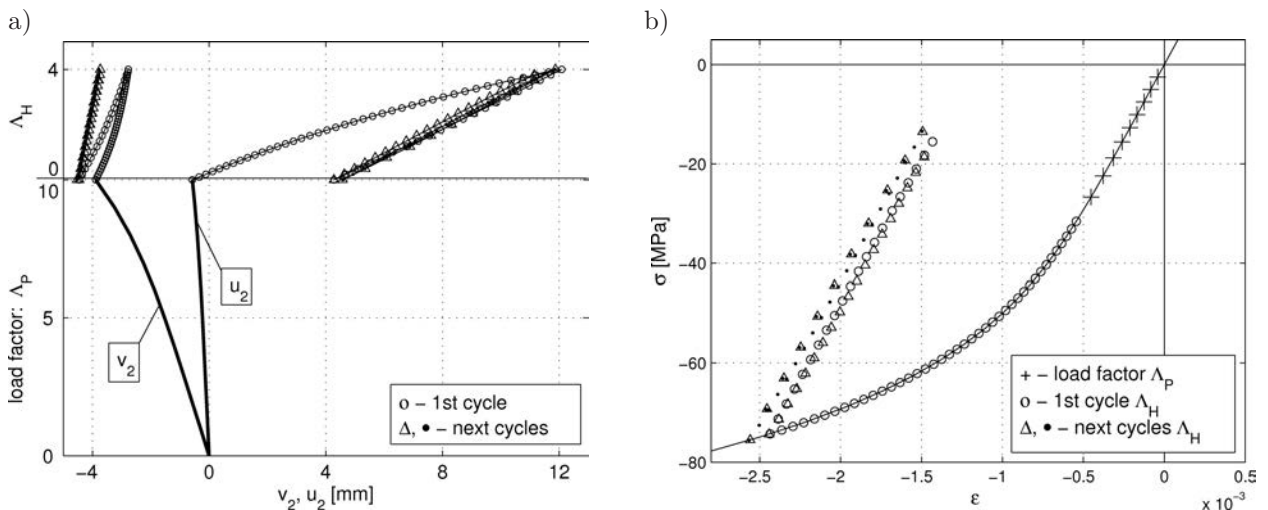

Fig. 6. a) Equilibrium paths $\Lambda(v_2), \Lambda(u_2)$, b) development of strain and stress in element number 3 in $c = 2$ cycles of load.

The analysis points out that after performing $c = 2$ load cycles, the field of plastic strain was stabilized: the truss exhibits plastic shakedown effects [14]. This means that no subsequent change of load (according to the program) will cause the increase of the plastic strain.

5.2. Case study 2

NMM_{SNN} used in the previous analysis, was used again in the analysis of a more complex problem of the shakedown analysis. A truss structure made of the same material as the reference truss was analyzed.

The geometry of the truss was taken from the real design of a tower with a tank, cf. [15]. The load was assumed on the level to fulfil ultimately the limit state and serviceability limit state conditions. In Fig. 7 the static scheme, loads and cross-sections of the profiles of the analyzed truss are shown.
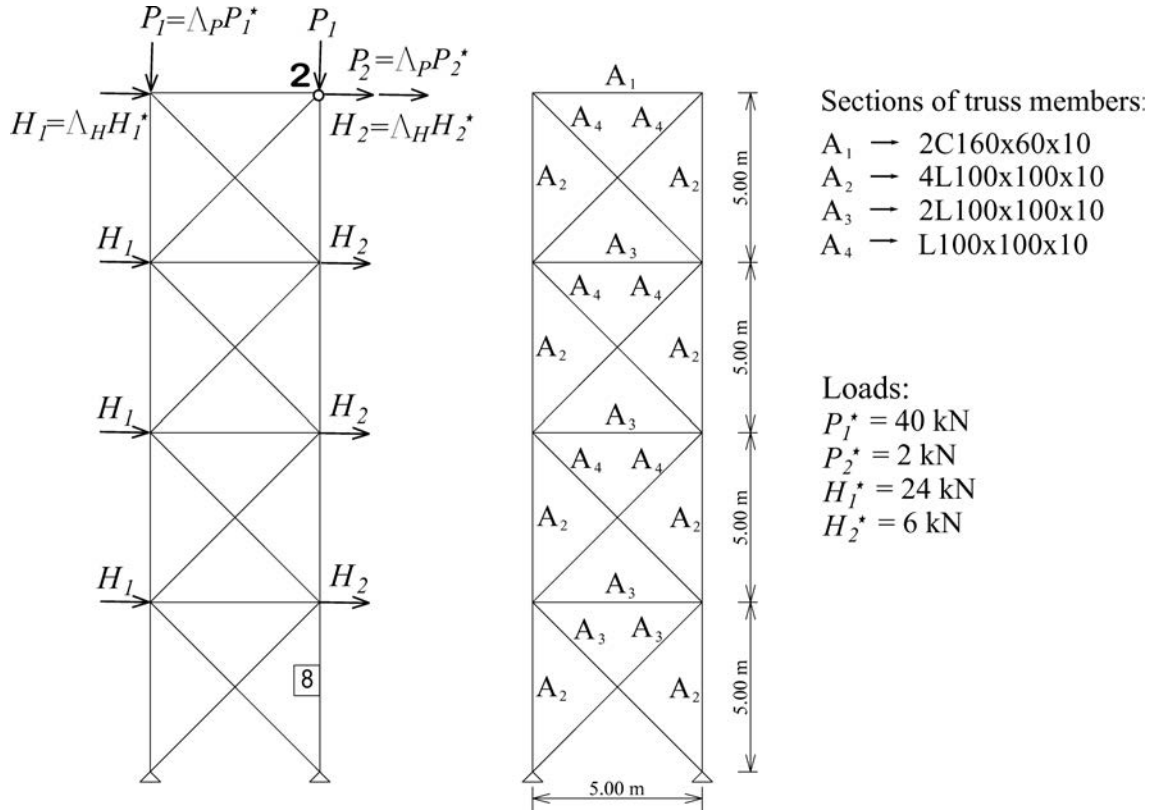

Fig. 7. Data of truss structure from [15].

In this case the following program of the cyclic loading was adopted:

$$
\begin{array}{llll}
\text{(I)} & \mathbf{P}^0 = \Lambda_P \mathbf{P}^\star & \text{for} & \Lambda_P: \quad 0.0 \to 15.0; \\
\text{(II)} & \mathbf{P}^0 + (\Lambda_{H_r} \mathbf{H_r}^\star)^c & \text{for} & \Lambda_{H_r}^c: \quad (0.0 \to 10.0 \to 0.0)^c; \\
\text{(III)} & \mathbf{P}^0 + (\Lambda_{H_l} \mathbf{H_l}^\star)^c & \text{for} & \Lambda_{H_l}^c: \quad (0.0 \to 10.0 \to 0.0)^c,
\end{array}
\tag{16}
$$

where $\mathbf{H}_l^\star = -\mathbf{H}_r^\star$ – horizontal loads, $\mathbf{P}^\star$ – vertical load. Indices r, l indicate the direction of forces $\mathbf{H}$. The loads listed above can be interpreted according to the civil engineering interpretation as dead load $\mathbf{P}$ and wind action $\mathbf{H}$. In this way the shakedown phenomenon can be analyzed.

As in the previous example, the analysis was performed twice: by means of the integrated system FEM/NMM_{SNN} and the FEM/RO program. In Fig. 8 the equilibrium paths determined for the horizontal and vertical displacements of the node number 2 and the stress distribution in the 8th element of the truss are presented.

In this case, too, the equilibrium points found by FEM/NMM_{SNN} program are very close to the equilibrium paths computed by means of the standard program FEM/RO. After four loading cycles
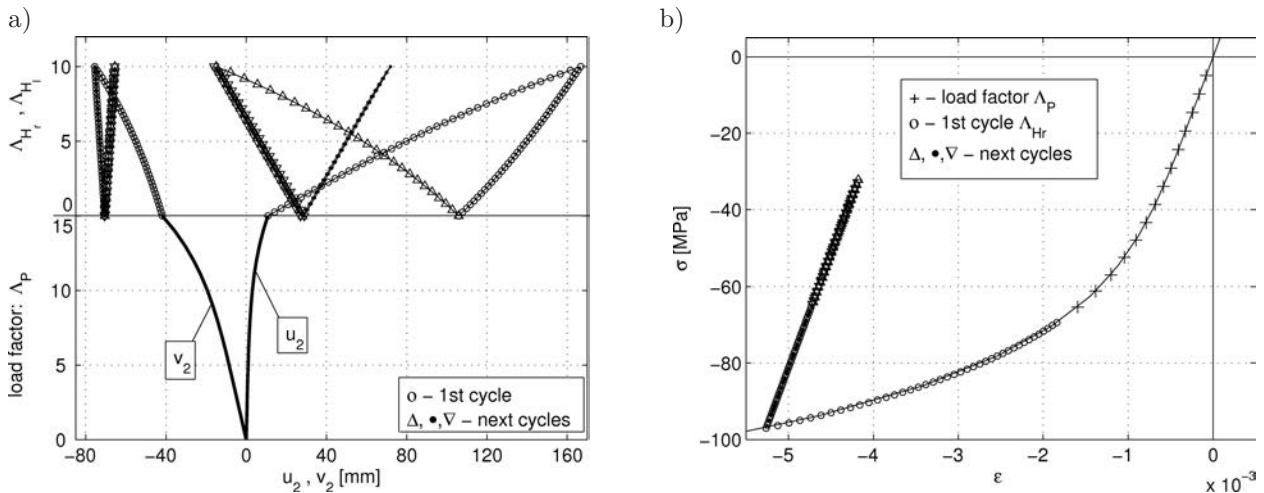
a)

b)

Fig. 8. a) Equilibrium paths $\Lambda(v_2)$, $\Lambda(u_2)$, b) development of strain and stress state in element number 8, during four load cycles.

it was proved that the shakedown load capacity was achieved. This means the structure behavior after adding the next horizontal loads **H** is elastic. This corresponds to the lack of plastic strains.

6. FINAL REMARKS AND CONCLUSIONS

1. In the paper it was shown that ANN can be used as a neural model of the material describing a physical law of the material in the aluminum structures. The *off line* training technique, which allows the use of experimental data, was applied to formulate the NMM. Two types of neural networks: SNN and BNN were used for solving the inverse problem. The developed networks have been integrated with the FEM program, thus forming a hybrid program FEM/NMM.

2. Developed ANN can be successfully applied to the shakedown analysis of aluminum truss.

3. The main aim of this study was to demonstrate the possibility of development of a hybrid program FEM/NMM$_{ANN}$ for the analysis of various boundary problems in the uniaxial stress state.

ACKNOWLEDGEMENT

REFERENCES

[1] J. Lemaitre, J.L. Chaboche. *Mechanics of solid materials*. University Press, Cambridge, 1994.
[2] T. Furukawa, T. Sugata, S. Yoshimura, M. Hoffman. An automated system for simulation and parameter identification of inelastic constitutive models. *Comput. Methods Appl. Mech. Eng.*, **191**: 2235–2260, 2002.
[3] M. Lefik, B.A. Schrefler. Artificial neural network for parameter identifications for an elasto-plastic model of superconducting cable under cyclic loading. *Comp & Struct.*, **80**: 1699–1713, 2002.
[4] G. Bolzon, V. Buliak, G. Maier, B. Miller. Assessment of elastic-plastic material parameters comparatively by three procedures based on indentation test and inverse analysis. *Inverse Problems in Science and Eng.*, **19**: 815–837, 2011.
[5] J. Ghaboussi, D.A. Pecknold, M. Zhang, R.M. Haj-Ali. Autoprogressive training of neural network constitutive models. *Int. J. Num. Meth. Eng.*, **42**: 105–126, 1998.

[6] E. Pabisek. Identification of an equivalent model for granular soils by FEM/NMM/p-EMP hybrid system. *Computer Assisted Mech. Eng. Sci.*, **18:** 283–290, 2011.

[7] PN-EN 1999-1-1:2011, Design of aluminum structures. Part 1-1: General rules [in Polish: Projektowanie konstrukcji aluminiowych. Część 1-1: Reguły ogólne].

[8] T. Akazawa, M. Nakashima, O. Sakaguchi. Simple model for simulating hysteretic behavior involving significant strain hardening. *11th World Conf. on Earthquake Eng.*, 264, 1996.

[9] I.T. Nabney. *Netlab: algorithms for pattern recognition*, London, Springer, 2006.

[10] *Neural network toolbox for use with MATLAB*. The MathWorks. Inc., 2006.

[11] C.M. Bishop. *Neural networks for pattern recognition*. Oxford, Oxford University Press, 1995.

[12] Y.M. Hashash, S. Jung, J. Ghaboussi. Numerical implementation of a neural network based material model in finite element analysis. *Int. J. Num. Meth. Eng.*, **59**: 989–1005, 2004.

[13] A.P. Pérez-Foguet, A. Rodriguez-Ferran, A. Huerta. Numerical differentiation for local and global tangent operators in computational plasticity. *Comput. Methods Appl. Mech. Engrg.*, **189**: 277–296, 2000.

[14] Z. Waszczyszyn, E. Pabisek. Elastoplastic analysis of plane steel frames by a new superelement. *Archives of Civ. Eng.*, **48**: 159–181, 2002.

[15] *20 Meter Water Tower*. U.S. Army Corps of Engineers, Engineering and Construction Division. US ARMY Engineer District, Afghanistan. APO AE, 09356, 2009.