BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 71(4), 2023, Article number: e145681
DOI: 10.24425/bpasts.2023.145681

ARTIFICIAL AND COMPUTATIONAL INTELLIGENCE

Influence of modeling phase transformations with the use of LSTM network on the accuracy of computations of residual stresses for the hardening process

Joanna WRÓBEL* and Adam KULAWIK

Department of Computer Science, Czestochowa University of Technology, Dąbrowskiego 73, 42-201 Częstochowa, Poland

Abstract. Replacing mathematical models with artificial intelligence tools can play an important role in numerical models. This paper analyses the modeling of the hardening process in terms of temperature, phase transformations in the solid state and stresses in the elastic-plastic range. Currently, the use of artificial intelligence tools is increasing, both to make greater generalizations and to reduce possible errors in the numerical simulation process. It is possible to replace the mathematical model of phase transformations in the solid state with an artificial neural network (ANN). Such a substitution requires an ANN network that converts time series (temperature curves) into shares of phase transformations with a small training error. With an insufficient training level of the network, significant differences in stress values will occur due to the existing couplings. Long-Short-Term Memory (LSTM) networks were chosen for the analysis. The paper compares the differences in stress levels with two coupled models using a macroscopic model based on CCT diagram analysis and using the Johnson-Mehl-Avrami-Kolmogorov (JMAK) and Koistinen-Marburger (KM) equations, against the model memorized by the LSTM network. In addition, two levels of network training accuracy were also compared. Considering the results obtained from the model based on LSTM networks, it can be concluded that it is possible to effectively replace the classical model in modeling the phenomena of the heat treatment process.

Key words: RNN network; hardening process; temperature; phase transformations in the solid state; effective stresses; numerical modeling.

1. INTRODUCTION

Modeling stresses for the hardening process of steel elements is complex and requires consideration of many factors. This area of modeling is currently strongly developed. For example, printing metal elements (directed energy deposition (DED) technology) is of great importance. Since successive layers of the element are produced in the process of melting the powder with a heat source, we have to deal with a complicated state of stress [1]. Modeling, for example, of welding processes in the production of car bodies has a similar impact on the quality of manufactured parts [2]. Papers on phenomena during heat treatment are also in stress modeling. Recursive networks are also successfully applied, e.g., for the determination of non-linear plastic response under multiaxial loading [3]. Control over phase transformations, temperatures and stresses is key to achieve product quality. It is crucial to achieving a high level of modeling accuracy before manufacturing parts.

It can be summarized that each process with a heat-affected zone, especially with a local heat-affected area, requires initial control using an appropriate simulation model. If this is not done, hot or cold cracking is possible [4, 5].

*e-mail: joanna.wrobel@pcz.pl

Manuscript submitted 2022-12-30, revised 2022-12-30, initially accepted for publication 2023-04-20, published in August 2023.

To obtain a satisfactory level of calculation accuracy, it is necessary to consider a minimum of three main elements of the process – temperature changes, phase transformations in the solid state, and stresses in the elastic-plastic range [6]. Currently, macroscopic models based on empirical equations such as Johnson-Mehl-Avrami-Kolmogorov (JMAK) [7–9] are still prevalent in determining phase transformations in the solid state. The authors of this paper decided to replace the continuous model based on the analysis of CTPc diagrams with the use of JMAK and Koistinen-Marburger [10] equations (macroscopic model of phase transformations in the solid state) with the results obtained from the Long-Short-Term Memory (LSTM) network representing the continuous model. Recurrent neural networks of the LSTM type can model tasks with multiple input variables. This is a great advantage, especially for time series prediction, where classical linear methods can be difficult to adapt. It is possible to replace time series analysis with recursive networks, e.g., Gated Recurrent Unit (GRU) or LSTM [11]. However, whether the resulting model inaccuracies will significantly affect changes in stress states is essential.

In the numerical model of the hardening process, the artificial neural network was used as a black box giving information on the transformation of austenite to ferrite, pearlite, bainite and martensite based on a series of temperature levels representing the cooling process in the nodes of the considered geometry of the mesh. The results obtained from the LSTM network are subject to some errors, but this approach allows us to obtain faster results with greater immunity to errors in diagram analysis. Errors in the obtained phase transformations refer to temporary underestimation or overestimation of values at the grid nodes. Different levels of deviations can be obtained depending on the selected geometry of the neural network or the chosen method and learning parameters. In this paper, two models of phase transformation based on LSTM with two error levels were analysed.

The level of accuracy of the results obtained for the phase transformation model strongly influences the other elements of the modeling of the hardening process. This is related to couplings of parts of the model, such as the effect of: the kinetics of transformations on structural and transformation strains, the level of phase transformations on the yield point of the material (weighted sum of yield points of individual phases), the latent heat of phase transformations on temperature changes during the cooling process, or the level of phase transformations on other material properties [12, 13]. The couplings mentioned above show a high dependence on residual stress levels in the elastic-plastic model due to changes in phase transformations. Considering that the obtained differences in temporary underestimation or overestimation of values may concern the neighboring nodes or mesh elements of the considered geometry, there may be conditions for the formation of temporary errors in the level of residual stresses. Accumulation of errors can occur. This is especially important for areas of the geometry of hardened components with a tendency to develop large stress gradients – notches, heat treatment boundaries, inclusions in the material, etc. It is also possible that after averaging over a finite element from nodal values, error levels will decrease.

The presented models take into account the results obtained and presented in the authors' paper entitled "Algorithm for determining time series of phase transformations in the solid state using Long-Short-Term Memory Neural Network" [14], this applies especially to the geometry of the network. The results presented in this paper are an attempt at error analysis for the case of replacing a continuous mathematical model of phase transformations with an LSTM network that allows analysis of time series representing temperature changes during cooling.

2. NUMERICAL MODEL AND SIMULATION

The numerical model, based on a model of the associated phenomena of the heat treatment process (Fig. 1), takes into account three of its most essential elements: temperature, phase transformations in the solid state, stresses.

To verify the learned neural network, stress state calculations were performed for the test geometry. An element with a cross-section of 0.1×0.1 m with indentations 0.005 m wide and 0.02 and 0.015 m deep, respectively, was analysed (Fig. 2). Due to the adopted geometry of the angle and the locations of the indentations, which were expected to cause the appearance of stress concentrations. It was decided to discretise the cross-sectional space of the steel element in quite a detailed manner. Due to the cooling curves occurring in the task (their range of rates), to obtain correct results, from all three basic elements of the model, a time step was adopted in the temperature region below A_{c3} at $\Delta t = 0.02$ s. This resulted in calculations for over 2000 time steps to cool the material.

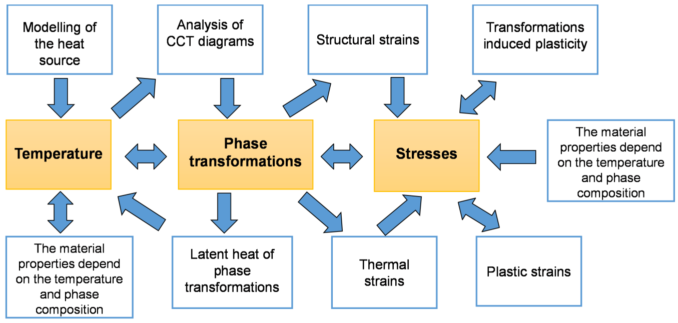

Fig. 1. Modeling scheme for heat treatment phenomena

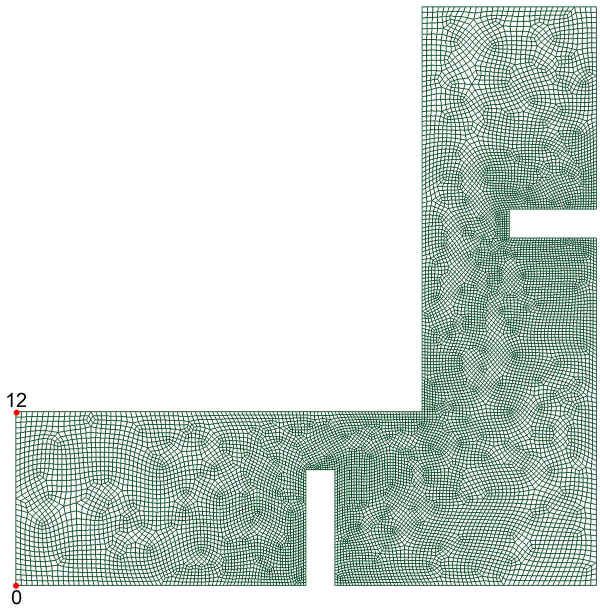

Fig. 2. Discretisation of the cross-sectional analysis space

2.1. Temperature

A numerical model solving the differential heat transfer equation for non-stationary flow in Lagrange coordinates was used to determine temperature changes. The choice of coordinates was dictated by the need to model an exemplary steel element with a constant cross-section without considering free convection in the surrounding coolant. The heat treatment process was limited to cooling only. Thus, it was assumed that a structurally homogeneous element without initial stresses with an initial temperature of 1200 K is cooled in water at 293 K. The differential heat conduction equation was solved by the finite element method using the discretisation of space shown in Fig. 2. The basic finite element was a quadrangular element with bi-linear shape functions (number of nodes – 14393). Due to the adopted cooling, a boundary condition of the third type was used for modeling in the temperature range with the heat transfer coefficient from the environment depending on the values at the boundary nodes of the geometry [15]. The thermophysical values depended on the temperature [16], and their average value in a single finite element was assumed.

2.2. Phase transformation

The purpose of this paper is to present the impact of typical errors occurring during the prediction of values by artificial intelligence tools on the values determined in the numerical model and analysed by technologists. In the discussed coupled model of heat treatment process phenomena, the model of phase transformations in the solid state was selected as the element to be replaced by the LSTM network. Based on the classical model of macroscopic transformations [16, 17], the stresses were determined as the data most sensitive to errors in the preceding calculations. The same range of stresses was generated by the model including the LSTM network. This model was trained using data obtained from the macroscopic model considering cooling rates in the range of 10–80 K/s with a rate-of-change step of $\Delta V_c = 0.02$ K/s. As a result, 3600 cooling lines were obtained, which were used to memorize the functioning of the model of phase transformations in the solid state by an artificial neural network. Each cooling line was represented by up to 83 temperature changes along with changes in phase shares. Because the range of cooling rates is quite large some of the data in the form of time series had to be supplemented (for the rate of 80 K/s, only 20 elements were carrying the information). Since the paths of the cooling lines occurring during the modeling of cooling in the water of the steel element do not correspond to the cooling curves during learning, for each point (time, temperature) on the cooling line the average cooling velocity $(T - T_{A_{c3}})/(time - time_{A_{c3}})$ was determined (Fig. 3). This average cooling rate determined from which cooling curve we would determine the levels of phase transformations. The difference between the values obtained from successive values of the average rates and the current level of temperature and time allowed us to determine successive gains or losses of phase transformations in the solid state.

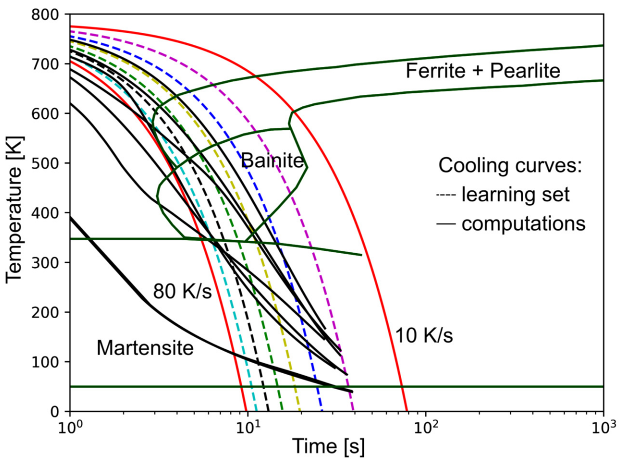

Fig. 3. CCT diagram with cooling curves

2.3. Stresses

The stress state was determined by solving a differential equation defining the equilibrium equations in incremental form. The relationship between strains and stresses in incremental form takes into account temperature-dependent changes in the elasticity matrix. It was assumed that elastic strains result from the difference between total strains and the sum of thermal, structural, plastic and transformational strains. The model built concerned the elastic-plastic range, taking into account the material model with isotropic hardening. The Huber-Mises-Hencky yield condition was used. The obtained equations were solved using the iterative method. The yield point depended on the phase composition and temperature [16]. Degrees of freedom were taken away at two points of the considered geometry to ensure static determinability and so that restraints would not generate additional stresses ($U_x[0] = U_y[0] = 0$; $U_x[12] = 0$).

2.4. LSTM network

A model predicting the distributions of phase transformations using the RNN network, trained on the classical macroscopic model of phase transformations for constant cooling rates, was developed in the authors' previous paper [14]. A description of the model architecture is included in Table 1. Five network configurations were considered, which could be used to extract other phases or groups of phases based on sums or differences. It was not possible for one network to learn all the phase transformations. Finally, three configurations were chosen in the paper: ferrite and pearlite because they have similar thermal and structural expansion coefficients and they can be treated as a homogeneous mixture. On the other hand, bainite and martensite are significantly different, so it was decided to consider them separately. The learning process was carried out using Adam's optimization method. The rectified linear unit activation function (ReLU) was used at the output of the network.

Table 1
Geometry of the analysed RNN network

Layer (Type)	Output Shape	Param
lstm_1 (LSTM)	(100, 83, 83)	28,220
lstm_2 (LSTM)	(100, 83, 83)	55,444
lstm_3 (LSTM)	(100, 83, 83)	55,444
time_distributed_1 (TimeDist	(100, 83, 1)	84
activation_1 (Activation)	(100, 83, 1)	0
Total params: 139,192 Trainable params: 139,192 Non-trainable params: 0		

In the paper, the training process was carried out for 50 and 2000 epochs. Figure 4 shows the training and validation process. Despite using the random state of the weights as the initial state, the accuracy obtained after 50 epochs in both training cases is comparable.

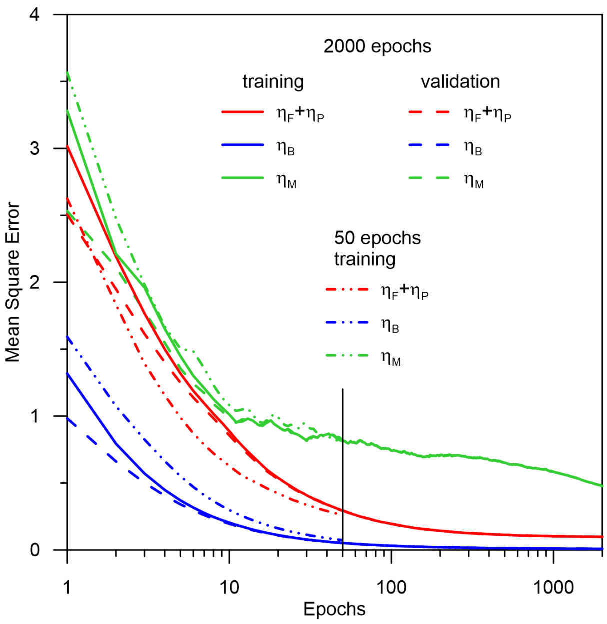

Fig. 4. Training and validation error history for different network architectures

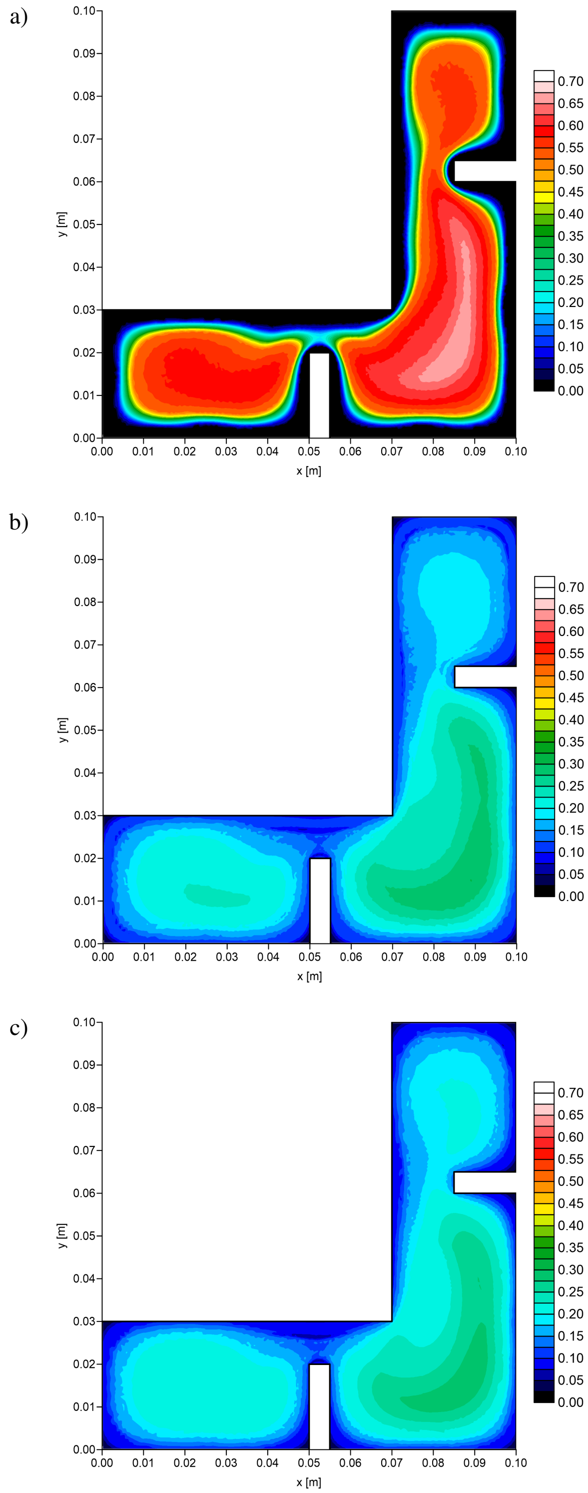

Fig. 5. Sum of ferrite and pearlite fraction (after cooling process): a) numerical model, b) RNN 2000 epochs, c) RNN 50 epochs

3. RESULTS AND CONCLUSIONS

The use of training for constant rates resulted in a model that behaves well, and the results obtained are continuous, but it can be seen that they differ from the classical model. To create a model based on LSTM layers, only the analysis of cases with constant cooling rates was used. Because of the above, the results obtained for variable cooling speeds from the RNN model deviate and differ significantly (Figs. 5–7). Probably, the error level would be reduced if the analysed cooling speeds were included, however, it cannot be assumed that the input data will be known temperature profiles. Analysing the values of the differences, it can be concluded that there was a large overestimation of the value of martensite inside the area (Fig. 7), while underestimating the ferrite–pearlite transformation (Fig. 5). On the other hand, in the boundary layer we have an underestimation of the value of martensite and an overestimation of the proportion of soft structures. The most complex nature of the difference is in the bainite structure (Fig. 6), which in the area close to the edge is strongly underestimated while inside it is overestimated.

As part of the article, the second type of analysis was also performed - how much the accuracy of the learning process affects the differences in stress values (Fig. 9). When analysing the distribution of strains or effective stresses, one can assume a high correlation and quite insignificant differences between the considered cases. Such an analysis was dictated by the need to complete the learning process when the model, instead of memorizing, was also used for generalizations. Stopping the learning process at the moment of overfitting (in this case only hypothetical) is a typical procedure when teaching a neural network. Unfortunately, the presented model does not generalize but remembers, but it is still possible to analyse the results taking into account certain levels of training inaccuracy. The results regarding the training process for the validation set suggest that there is no overfitting (Fig. 4). However, these results are for validation data that are very close to the training data. Splitting the input set of 3600 samples randomly selected into the training, test and validation sets with such small differences in velocity values gives us a false picture of the lack of overfitting.

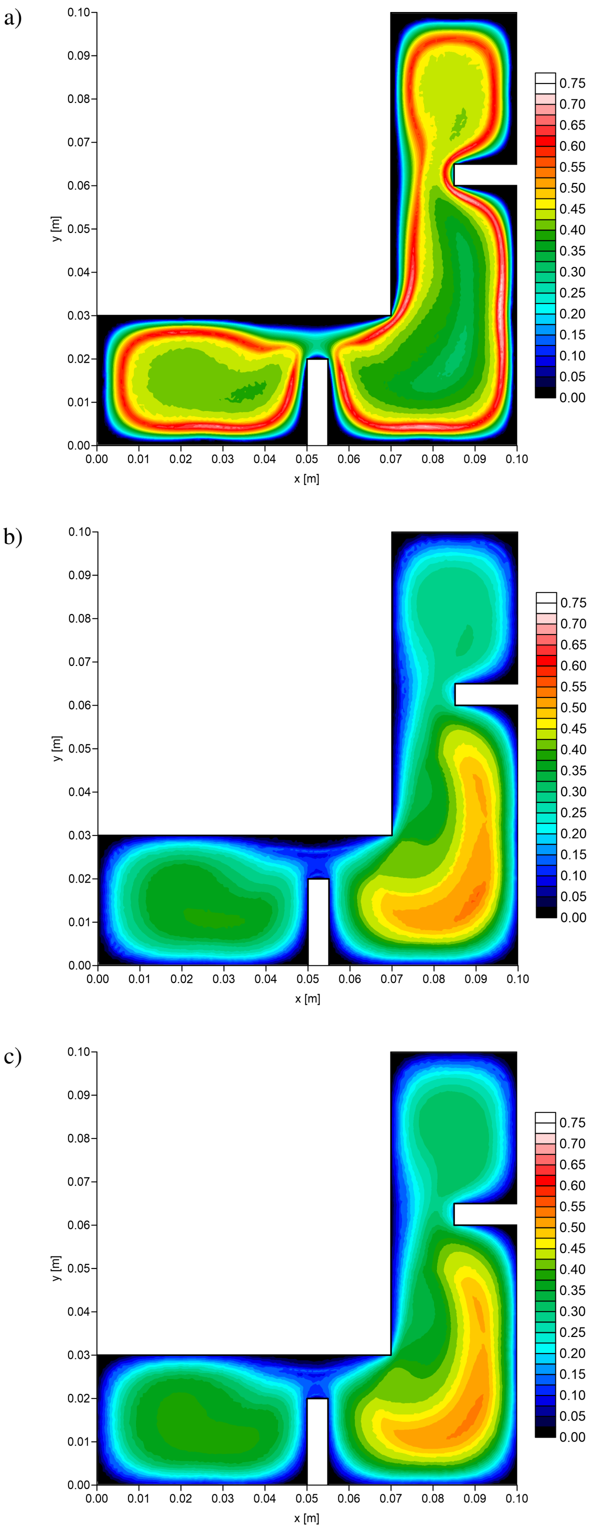

Fig. 6. Bainite fraction (after cooling process): a) numerical model, b) RNN 2000 epochs, c) RNN 50 epochs

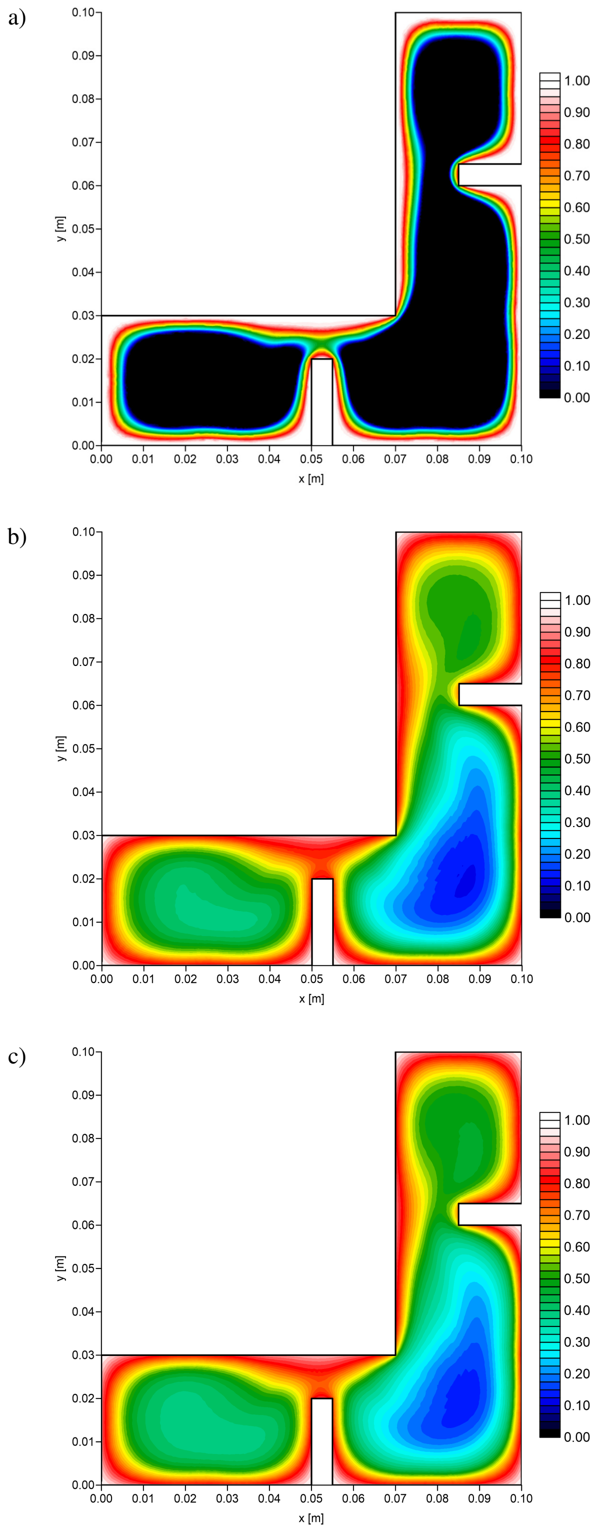

Fig. 7. Martensite fraction (after cooling process): a) numerical model, b) RNN 2000 epochs, c) RNN 50 epochs

The training error in terms of standard deviation between stopping at 50 and 2000 epochs decreases more than twice. However, on the scale of the whole error, the difference is only a few percent. These few percents are also noticed in the values of the obtained phase transformations (the differences between the shares are no more than 4.5%) (Fig. 8). Even such a small difference can result in local differences between the obtained stresses of more than 30%, and the average difference in the whole area is not more than 6% (Fig. 10). On the other hand, these differences apply to areas with low effective

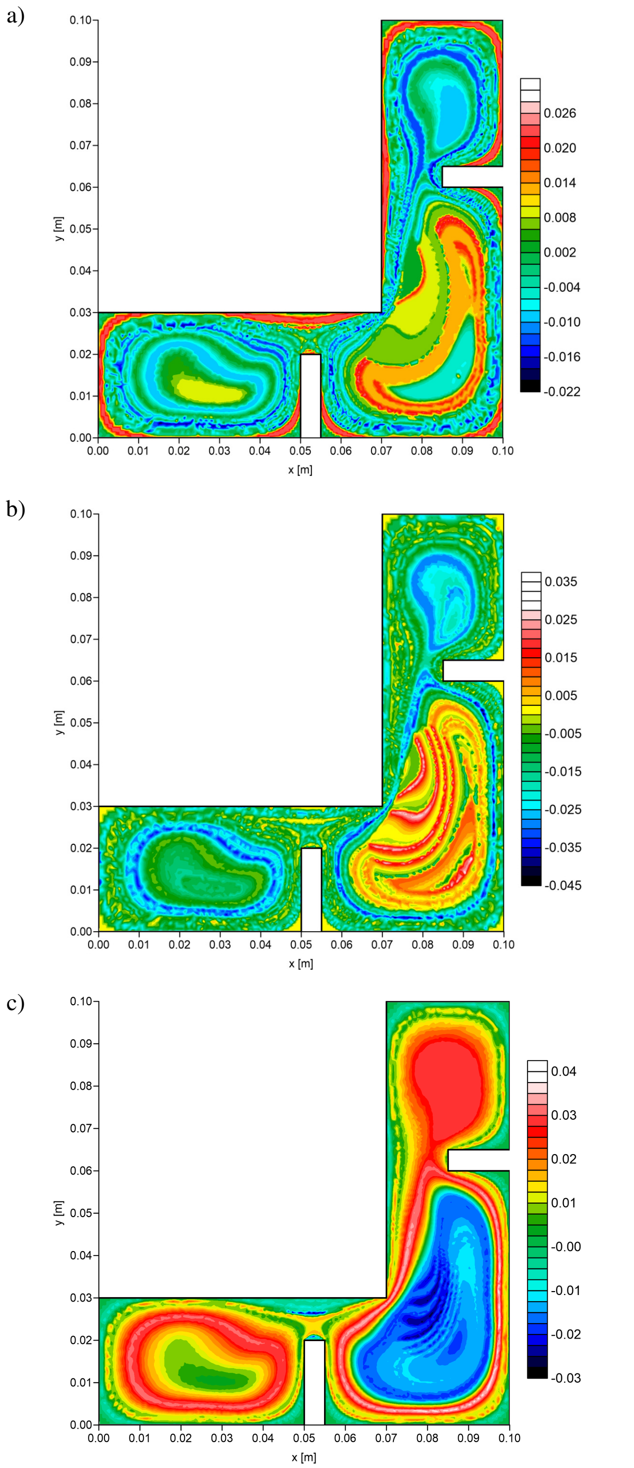

Fig. 8. Difference between RNN models of 2000 and 50 epochs: a) sum of ferrite and pearlite, b) bainite, c) martensite

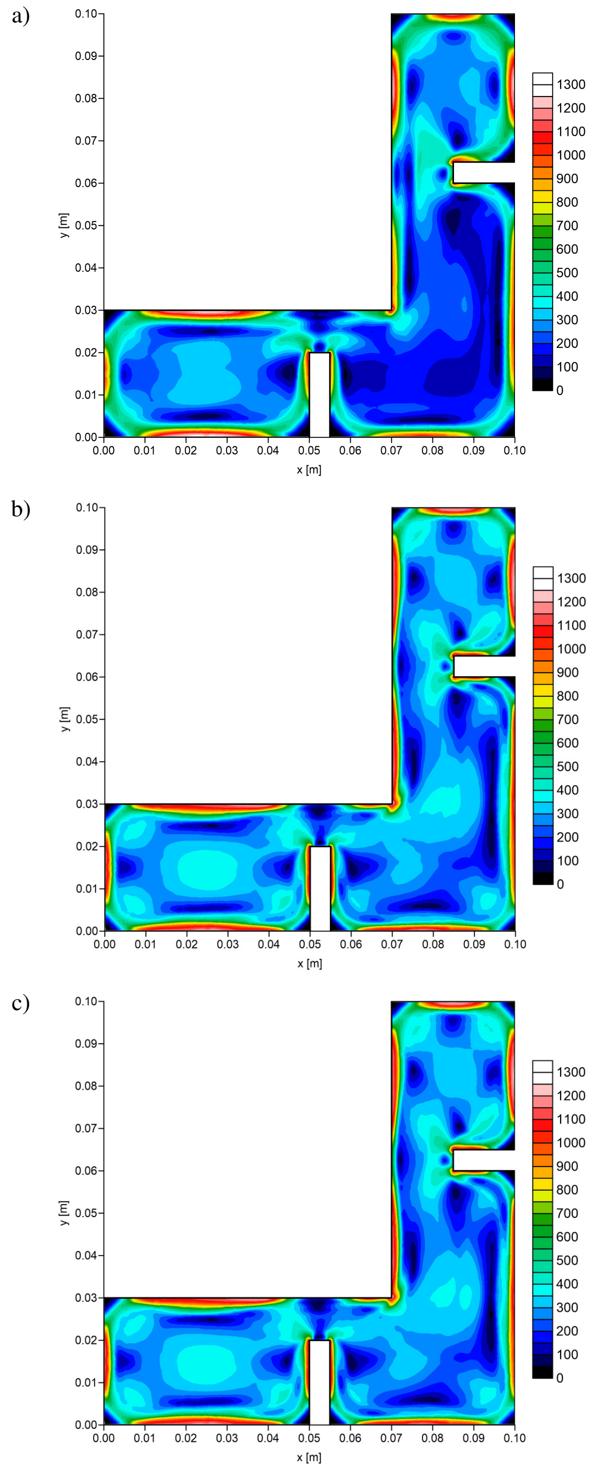

Fig. 9. Residual effective stresses distribution [MPa]: a) numerical model, b) RNN 2000 epochs, c) RNN 50 epochs

stresses. Considering that the spread of effective stress distribution reaches the level of 1300 MPa, the differences from −40 to 50 MPa can be considered insignificant.

Considering the results obtained in previous paper [14] regarding the comparison of the same models only for constant velocities, it can be concluded that the proposed solution is correct only in cases where the artificial network gets as training data the same family of cooling curves, which occur in the later considered computational case. Confirmation of this thesis is evident, especially when comparing the values of the obtained effective stresses. Nevertheless, the results obtained allow for much better accuracy of calculations than in the case of not considering phase transformations when modeling stresses in the heat treatment process. The critical element linking stress lev-

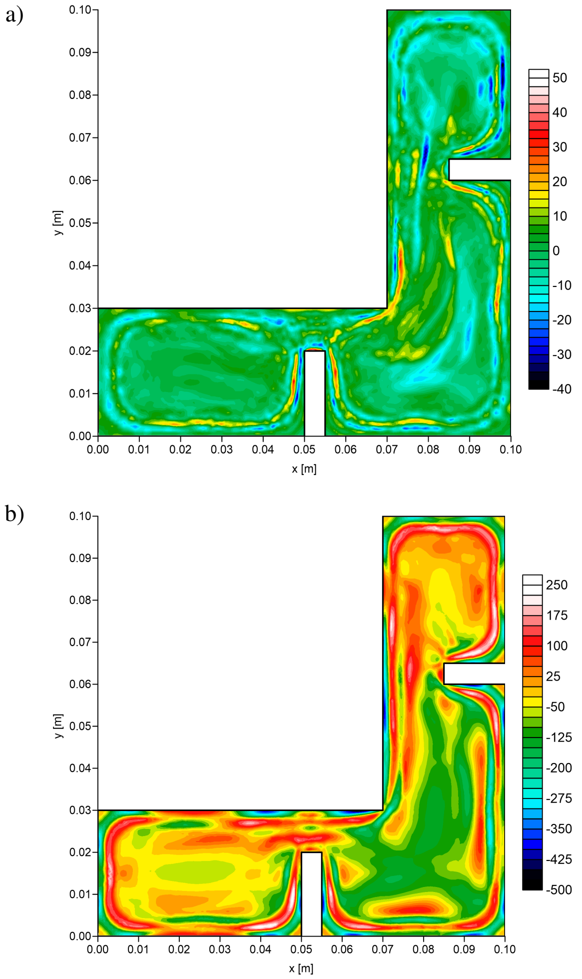

Fig. 10. Residual effective stresses distribution [MPa]: a) difference between RNN models, b) difference between numerical model and RNN models with 2000 epochs

els and phase transformations is probably taking into account the yield point of individual (according to [16]) phases in the calculations of the elastic-plastic model.

REFERENCES

[1] W. Chen, L. Xu, Y. Han, L. Zhao, and H. Jing, "Control of residual stress in metal addit. manuf. by low-temperature solid-state phase transformation: An experimental and numerical study," *Addit. Manuf.*, vol. 42, p. 102016, 2021, doi: 10.1016/j.addma.2021.102016.

[2] B. Feujofack Kemda, N. Barka, M. Jahazi, and D. Osmani, "Modeling of phase transformation kinetics in resistance spot welding and investigation of effect of post weld heat treatment on weld microstructure," *Met. Mater. Int.*, vol. 27, pp. 1205–1223, 2021, doi: 10.1007/s12540-019-00486-x.

[3] L. Borkowski, C. Sorini, and A. Chattopadhyay, "Recurrent neural network-based multiaxial plasticity model with regularization for physics-informed constraints," *Comput. Struct.*, vol. 258, p. 106678, 2022, doi: 10.1016/j.compstruc.2021.106678.

[4] M. Lindroos, T. Pinomaa, A. Antikainen, J. Lagerbom, J. Reijonen, T. Lindroos, T. Andersson, and A. Laukkanen, "Micromechanical modeling approach to single track deformation, phase transformation and residual stress evolution during selective laser melting using crystal plasticity," *Addit. Manuf.*, vol. 38, p. 101819, 2021, doi: 10.1016/j.addma.2020.101819.

[5] P. Taraphdar, R. Kumar, C. Pandey, and M. Mahapatra, "Significance of finite element models and solid-state phase transformation on the evaluation of weld induced residual stresses," *Met. Mater. Int.*, vol. 27, pp. 3478–3492, 2021, doi: 10.1007/s12540-020-00921-4.

[6] K. Zhang, W. Dong, and S. Lu, "Transformation plasticity of af1410 steel and its influences on the welding residual stress and distortion: Experimental and numerical study," *Mater. Sci. Eng. A*, vol. 821, p. 141628, 2021, doi: 10.1016/j.msea.2021.141628.

[7] M. Avrami, "Kinetics of phase change. i general theory," *J. Chem. Phys.*, vol. 7, no. 12, pp. 1103–1112, 1939, doi: 10.1063/1.1750380.

[8] W. Sun, F. Shan, N. Zong, H. Dong, and T. Jing, "A simulation and experiment study on phase transformations of ti-6al-4v in wire laser addit. manuf." *Mater. Des.*, vol. 207, p. 109843, 2021, doi: 10.1016/j.matdes.2021.109843.

[9] S. Chen, K. Bandyopadhyay, S. Basak, B. Hwang, J. Shim, J. Lee, and M. Lee, "Predictive integrated numerical approach for modeling spatio-temporal microstructure evolutions and grain size dependent phase transformations in steels," *Int. J. Plast.*, vol. 139, p. 102952, 2021, doi: 10.1016/j.ijplas.2021.102952.

[10] D. Koistinen and R. Marburger, "A general equation prescribing the extent of the autenite-martensite transformation in pure iron-carbon alloys and plain carbon steels," *Acta Metallurgica*, vol. 7, no. 1, pp. 59–60, 1959, doi: 10.1016/0001-6160(59)90170-1.

[11] Y. Li and S. Li, "Deep learning based phase transformation model for the prediction of microstructure and mechanical properties of hot-stamped parts," *Int. J. Mech. Sci.*, vol. 220, p. 107134, 2022, doi: 10.1016/j.ijmecsci.2022.107134.

[12] J. Guo, X. Deng, H. Wang, L. Zhou, Y. Xu, and D. Ju, "Modeling and simulation of vacuum low pressure carburizing process in gear steel," *Coatings*, vol. 11, p. 1003, 2021, doi: 10.3390/coatings11081003.

[13] M. Yaakoubi, M. Kchaou, and F. Dammak, "Simulation of the thermomechanical and metallurgical behavior of steels by using abaqus software," *Comput. Mater. Sci.*, vol. 68, pp. 297–306, 2013, doi: 10.1016/j.commatsci.2012.10.001.

[14] J. Wróbel and A. Kulawik, "Algorithm for determining time series of phase transformations in the solid state using long-short-term memory neural network," *Materials*, vol. 15, p. 3792, 2022, doi: 10.3390/ma15113792.

[15] C. Gür and A. Tekkaya, "Numerical investigation of non-homogeneous plastic deformation in quenching process," *Mater. Sci. Eng. A*, vol. 319–321, pp. 164–169, 2001, doi: 10.1016/S0921-5093(01)01064-4.

[16] A. Kulawik, "Modeling of thermomechanical phenomena of welding process of steel pipe," *Arch. Metall. Mater.*, vol. 57, no. 4, pp. 1229–1238, 2012, doi: 10.2478/V10172-012-0137-X.

[17] A. Bokota and S. Iskierka, "Effect of phase transformation on stress states in surface layer of laser hardened carbon steel," *ISIJ Int.*, vol. 36, no. 11, pp. 1383–1391.