Publishing House
AKAPIT

COMPUTER METHODS IN MATERIALS SCIENCE
Informatyka w Technologii Materiałów

Vol. 14, 2014, No. 1

METAMODEL DRIVEN OPTIMIZATION OF THERMOMECHANICAL INDUSTRIAL PROCESSES

JAN KUSIAK, ŁUKASZ SZTANGRET*, ŁUKASZ RAUCH, MACIEJ PIETRZYK

AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków, Poland
**Corresponding author: szt@agh.edu.pl*

Abstract

The main objective of the paper is presentation of the metamodel driven optimization (MDO) strategy of thermomechanical industrial processes. In this approach, the model of considered industrial process is replaced by a metamodel, which allows obtaining a significant reduction of the simulation time. Low simulation time of the analyzed process gives the ability to use heuristic optimization methods, which increase the probability of finding the global optimum. The paper discusses the idea of metamodelling and presents results of metamodelling and optimization of a selected metal forming process.

Key words: metamodel driven optimization, artificial neural networks, cooling of rails

1. INTRODUCTION

Optimization problems encountered in various fields of science, require searching for new methods, which allows to find optimal solutions. Most of these methods are based on the principle of iterative search for the optimal solution. Selection of an appropriate method depends on the considered optimization problem, in particular, on the form of optimized objective function (multimodality, differentiability, continuity) and the cost of objective function evaluation. In most real industrial problems, theassumption that the objective function is unimodal is false. Moreover, values of the objective function are calculated using computer simulations based on time-consuming algorithms. These algorithms, in the case of metal forming processes, are mainly based on the Finite Element Method (FEM). FEM simulations give proper response, however they are usually time consuming. This results in arisingtwo main obstacles which must be overcome:

– the multimodality of the objective function,
– long time of computer simulations which must be performed to compute the values of objective function.

The first obstacle can be overcome by using heuristic optimization methods, that search for a global optimum not only within a local region but also, over the whole domain of objective function. This advantageresults from the fact that most of heuristic methods do not operate on a single solution but on a whole population of solutions.On the other hand increasing the number of analyzedsolutions causes extension of the computing time, which leads toexceeding of the acceptable time. This problem can be overcome by the use of much faster metamodel instead of slow model (figure 1). Model is only used to generate the large enough training data set.

The paper presents the optimization problem of the controlled cooling of rails, while considered process was metamodelled using the artificial neural networks.

 ISSN 1641-8581

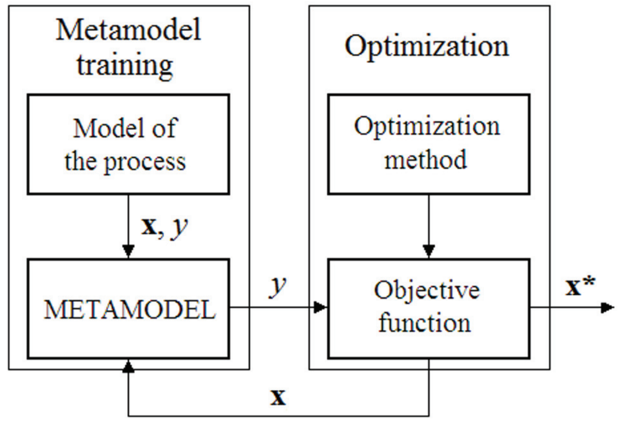

***Fig. 1.** Metamodel driven optimization flowchart.*

2. PROCESS DESCRIPTION

Problem of the controlled cooling of rails is thoroughly discusses by Kuziak et al. (2014) and optimization of this process is presented by Rauch et al. (2014). In the former paper the system for cooling of rails developed at the IMŻ Gliwice is described in detail. Briefly, the idea of the system is to enforce that the pearlitic transformation occurs at the lowest possible temperature. On the other hand, since the bainite is a disadvantageous component in rail steels, the temperature cannot be too low so that bainitic transformation can be avoided. The general requirements concerning microstructure of the rail are as follows:

– purely pearlitic microstructure at the whole cross section of the rail, without bainite or martensite,
– as small as possible interlamellar spacing in the rail head,
– uniform distribution of hardness at the cross section of the rail head.

On the basis of tests performed on Gleeble 3800 it was shown by Kuziak and Zygmunt (2013) that a very good relation between strength and ductility for eutectoid steels was obtained after isothermal holding at 550ºC. The pearlite microstructure after holding at this temperature was composed of fine pearlite having interlamellar spacing around 0.11 μm and colony size of 8 μm. When the pearlitic transformation occurs at lower temperature range, the structure contains, besides lamellar pearlite, undesirable degenerated pearlite and bainite. Increase in the volume fraction of those components in rail structure reduces its ductility.

On the basis of these observations Kuziak et al. (2014) proposed a cyclic immersion hardening of the rail head, which enables more homogenous hardness to be achieved in this head, without the necessity of accurate control of the total time of heat treatment. Rauch et al. (2014) presented a specialized computer system integrating various external programs, optimization library, metamodels and material database to perform optimization of industrial processes and cooling of rails was one of the case studies in that work. In the present paper the idea of development of metamodels for optimization is described and metamodel for the cooling of rails, which is used in the computer system for optimization, is presented.

3. METAMODELLING

The general idea of metamodelling relates to postulate that metamodel approximates the model of considered process. Metamodel must correctly correspond to the model and the metamodel output value has to be evaluated with a radically lower computing time than using the original model.

Metamodelling is a process of construction of an approximation of the analyzed model, on the basis of different techniques. In other words, the metamodel is *a model of the model*. The accuracy of the metamodel is usually verified using the statistical methods. The accuracy of metamodel depends on the used metamodelling technique and on the number of the points generated by the model. Usually, the higher number of points gives the better metamodel accuracy. This paper is focused on one of the most common metamodelling technique, i.e. technique based on artificial neural network (ANN). Examples of successful application of artificial neural networks in optimization can be found in (Sztangret et al., 2011; Kusiak et al., 2012). Sztangret et al. (2012) applied metamodel to solve optimization task in the inverse analysis of plastometric tests.

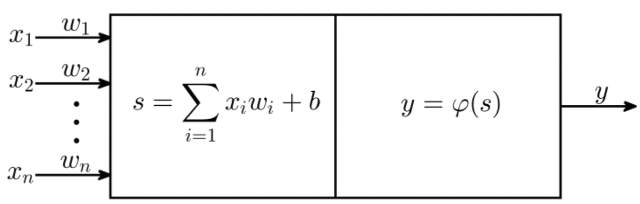

***Fig. 2.** Artificial neuron.*

ANN is an information processing system composed of a given number of artificial neurons. Each neuron has a number of inputs $x_1, \dots, x_n$. All inputs values are multiplied by a synaptic weight coefficients $w_i$and the sum is calculated (figure 2). The weight parameters allow changing the influence of the inputs on the neurons output. Additional constant value named bias is also added. Next, the calculated sum becomesthe argument of the activation function.

Whereas the activation function is selected in the first step of the design of the neuron, the weights coefficients and bias undergo variation during training process. The aim of the training process is to estimate the values of these parameters that enable trained neuron to predict the proper value, the closest to the model output. Supervised training process requires the training data set composed with pairs (X, Y), where $X = \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ is the set of input vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$, and Y is a set of corresponding model outputs. The difference between the proper values y_i and related values of neuron output is minimized with respect to weights w_i and the bias value during the training stage.

Metamodelling capabilities of a single neuron are limited, therefore, neural networks which are combination of several neurons are usedin practice. Thus, artificial neural network is composed of the defined number of neurons interconnected in a specific way. The most commonlyused network in metamodelling is MLP (Multi Layer Perceptron) feedforward network. Topology of MLP network is presented in figure 3. Neurons are collected into layers (input, one or two hidden and output layers). The number of neurons in input layer is equal to the number of metamodel inputs. There is usually one neuron in output layer. The number of hidden layers and the number of neurons are defined on thebasis ofresearcher's experience. More detailed description of MLP network can be found in Tadeusiewicz (1993), Bishop (1995), Haykin (1999).

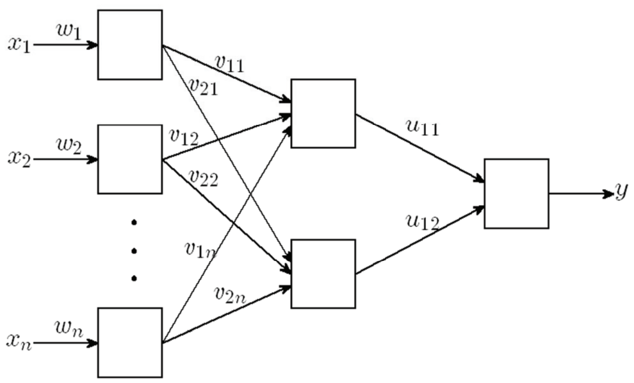

Fig. 3. *MLP feedforward network.*

MLP is a static network, which means that the output is calculated directly from the input through feedforward connections. Static network is not able to learn dynamic relations which may be present in the model. In dynamic network, the output depends not only on the current input to the network, but also on the current and formerinputs and outputs.It can be achieved by introduction of feedback and delay elements. It is expected that application of the dynamic neural network as metamodel of the industrial process should improve quality of the optimization, therefore, a brief description of this network is given below.

The relation between input and output is described by the formula:

$$y(k) = f\big(x(k), x(k-1), \dots, x(k-p), \\ y(k-1), \dots, y(k-q)\big) \quad (1)$$

where: p,q - the number of formerinputs and outputs which influence the current output value.

The nonlinear autoregressive network with exogenous inputs (NARX) is a recurrent dynamic network, with delays elements and feedback connections enclosing several layers of the network. The topology of NARX network is presented in figure 4. The "D" blocks represent unitary delay.

The training of NARX network is similar to training of MLP network, but training data set consists of values of discrete input and output signal.

In order to verify the possibility of application of NARX network in metamodelling of dynamical system a following test was performed. Simple time-variant SISO (Single Input, Single Output) system given by equation (2) was used as the model.

$$\begin{cases} u(k+1) = A\big(y(k)\big) \cdot u(k) + B \cdot x(k) \\ y(k) = C \cdot u(k) + D \cdot x(k) \end{cases} \quad (2)$$

where: $u(k), x(k), y(k) \in \mathbb{R}$- state, input and output vectors, respectivelyA, B, C, D- state, input, output and feedforward matrices, respectively

Value of state matrix A varies during the simulation as the output value is changing. Training signal was generated by the model built using MATLAB/Simulink software. Figure 5 shows the scheme of the model described by equation (2).

The generated training signal is shown in figure 6. It consists of 50 step responses.

The network composed of 4 layers (1, 10, 3 and 1 neurons respectively) has been tested. The sigmoid function was chosen as the activation function of input and hidden layers and linearactivation function in the output layer. The number of delays was equal to 5 for input signal and 3 for feedback signal.

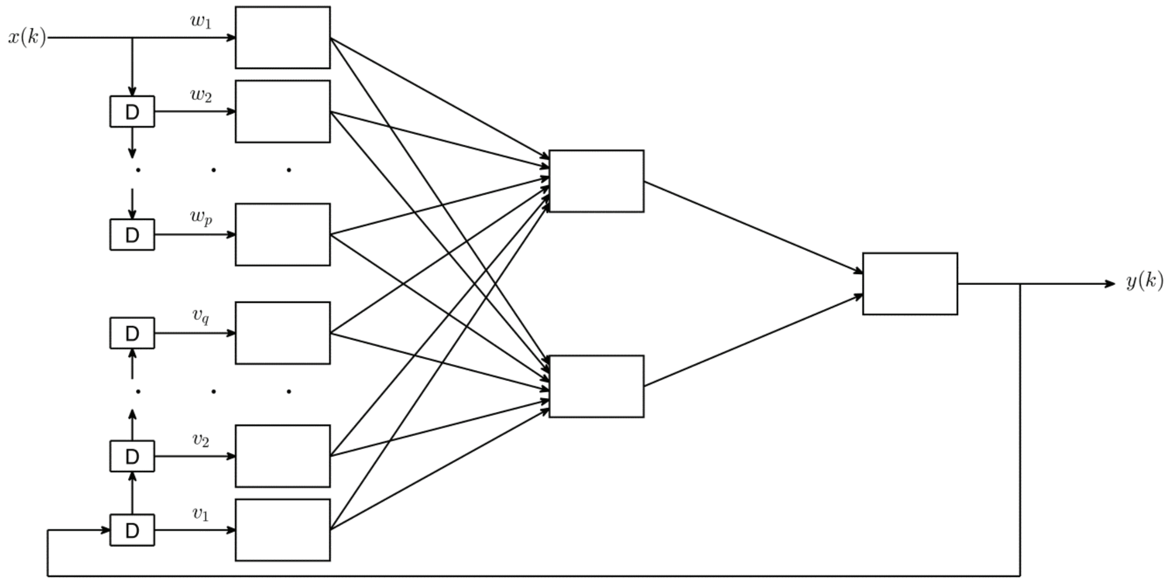

Fig. 4. NARX network (to avoid illegibility not all weights areshownin the figure).

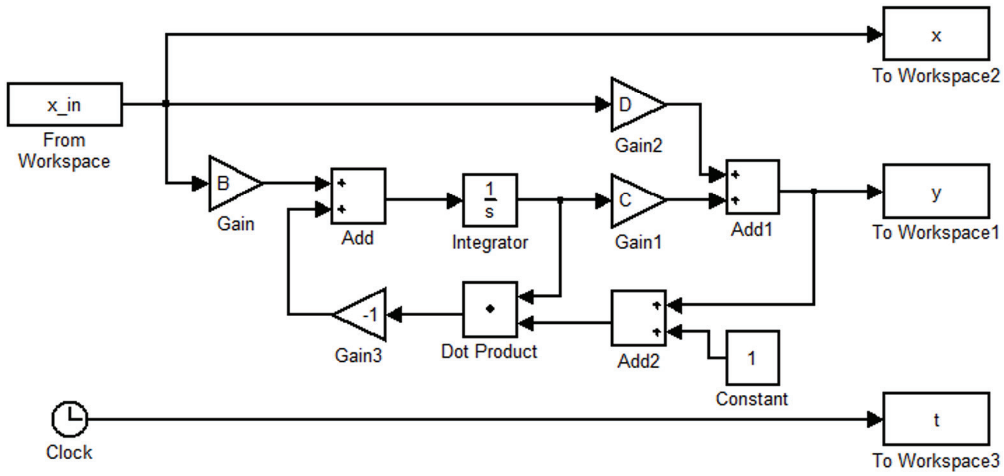

Fig. 5. Dynamic system built in MATLAB/Simulink.

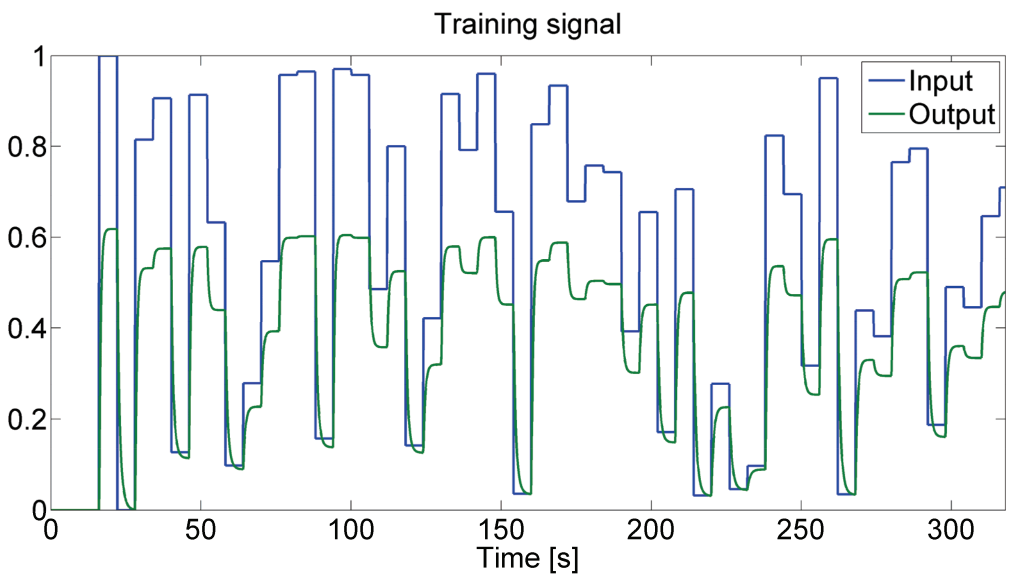

Fig. 6. Training signal for NARX network.

After training, test was performed using different signals, generated using the same scheme. The error was calculated usingfollowing equation:

$$\varepsilon = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i^{ANN}\right| \tag{3}$$

where: n- the number of signal samples.

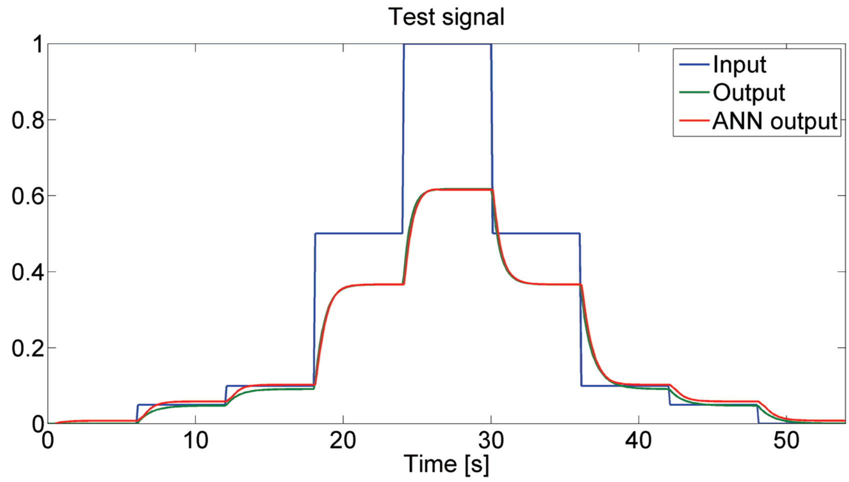

Fig. 7. *Test signal for NARX network.*

The calculated error was equal to 0.0056. Figure 7 presents input test signal as well as proper output and ANN's output.

4. OPTIMIZATION OF CONTROLLED COOLING OF RAILS BASED ON METAMODELS

Generally, models of the controlled cooling of rails built on the basis of Finite Element analysis are computationally time consuming, so their application in the optimization iterative procedures is practically impossible. Therefore, the metamodelling approach was tested in the present work.

Three MLP networks were used to build metamodel, which predicts properties of rail after cooling process. These properties are: hardness of the rail head (HB), volume fraction of bainite (F_b) and interlameller spacing (S). Input vector was composed of four variables: heat transfer coefficient (htc) and three parameters of cooling sequence (time of first immersion (t_{I1}), time of the air cooling (t_{ac}) and time of second immersion (t_{I2})). Gathereddata set of 2000 records was divided into training set (80% of records) and testing set (20% of records). The accuracy of metamodels was evaluated using the following equation:

$$\varepsilon = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - y_i^{ANN}}{\max\{y_i\} - \min\{y_i\}}\right)^2} \cdot 100\% \tag{4}$$

where: y_i- the required value and y_i^{ANN}- the value predicted by ANN.

Data for training the neutral network were generated by the model of accelerated cooling of rails. Detailed description of this model is given in (Pietrzyk & Kuziak, 2012). In this model equations describing microstructure and properties of eutectoid steels were implemented in the finite element program, which calculates temperature distribution in the rail.

Obtained accuracy as well as ANN's topology are presented in table 1. The objective of the optimization was to obtain as low as possible inhomogeneity of the hardness (ΔHB), volume fraction of bainite F_b and interlamellar spacing S. Inhomogeneity of the hardness was defined as:

$$\Delta HB = \sqrt{\sum_{i=1}^{m}\left[\frac{1}{m}\left(\frac{HB_i - HB_{ave}}{HB_{ave}}\right)^2\right]} \tag{5}$$

where: m - the number of the Gauss integration points at the cross section of the rail head, HB_{ave} - average hardness at the cross section of the rail head.

Optimization of the cooling of rails process was performed based on the described metamodel. Particle swarm optimization (PSO) method was used. This method is based on the behavior of populations of individuals. Particles traverse the decision space by following the particle representing the best solution found so far, while remembering the best position at which they have been so far. Each particle is described by two vectors: position **x** and velocity **v**. New velocity vector is evaluatedin each iteration. It leads to the change of the position of the particle. The velocity vector changes according to the following relationship:

$$\mathbf{v}_i^{(t+1)} = w\mathbf{v}_i^{(t)} + c_1 r_1\left(\mathbf{p}^g - \mathbf{x}_i^{(t)}\right) + c_2 r_2\left(\mathbf{p}_i - \mathbf{x}_i^{(t)}\right) \tag{6}$$

where $\mathbf{x}_i^{(t)}$ and $\mathbf{v}_i^{(t)}$ are the position and velocity of the i-th particle in the t-th iteration, respectively, $\mathbf{p}^g$ defines the best position found so far by the swarm, $\mathbf{p}_i$ is the best solution found so far by the i-th particle; w is defined as the inertia coefficient, c_1 and c_2 are acceleration coefficients (called also training coefficients), r_1 and r_2 are random numbers from the

[0,1] interval of the uniform distribution. The new particle's position is defined as follows:

$$\mathbf{x}_i^{(t+1)} = \mathbf{x}_i^{(t)} + \mathbf{v}_i^{(t+1)} \quad (7)$$

Table 1. *Error values and topologies of the networks.*

Property of the rail	ANN's topology	Activation function	Error
HB	4-10-7-1	sigmoid;log-sigmoid;radial;linear	8.5379%
F_b	4-15-1	sigmoid;log-sigmoid;linear	9.5067%
S	4-2-12-1	sigmoid; sigmoid; log-sigmoid; linear	7.5214%

After displacement of all particles to their new positions, they are subjected to an assessment and the swarm leader is selected. Coefficients values affect the swarm behaviour. The value of the inertia coefficient is usually selected from the [0,1] interval. A higher value is favourable for the global searching of the solution space, and a lower value for the local searching. Usually, the value is constant throughout the entire optimization process. However, it also may vary. Then, at the beginning, a high value is assumed, enabling global searching, and while approaching to the maximum that is sought, it gradually decreases. Acceleration coefficients are usually equal and selected from the [0,2] interval. The values are selected with the assumptions of the maximum particle velocities. The stop criterion is defined as the maximum number of iterations or the satisfactory value of the solution.

The goal of optimization was the determination of the values of heat transfer coefficient and parameters of cooling sequence ensuring the minimal values of the hardness inhomogeneity in the rail head, volume fraction of bainite and interlameller spacing. The objective function was given by the equation:

$$\Phi = \sqrt{\Delta HB + F_b^2 + 100S^2} \quad (8)$$

The following block constraints were imposed on decision variables:

– heat transfer coefficient (in the optimization it was understood as a decrease or increase of the heat transfer coefficient with respect to its basic value for the polymer solution) $hct \in [-500, 500]\,{}^{\mathrm{W}}/_{\mathrm{m^2K}}$
– parameters of cooling sequence $t_{l1}, t_{ac}, t_{l2} \in [1, 60]\mathrm{s}$

Minimum value of the objective function was found for $hct = -389.6\,{}^{\mathrm{W}}/_{\mathrm{m^2K}}, t_{l1} = 56.3\mathrm{s}, t_{ac} = 60\mathrm{s}, t_{l2} = 60\mathrm{s}$. These optimal process parameters result in the following rail head properties $\Delta HB = 0.126, Fb = 0.028, S = 0.132$ µm.

5. CONCLUSIONS

The use of metamodel driven optimization approach to the problem of reduction of optimization procedure computing time of complex industrial problem was discussed in the paper. The idea of metamodelling was presented as well as results of application of this approach to optimization of controlled cooling of rails (cooling by two immersions). Comparison of the obtained optimal process parameters with these presented in Rauch et al. (2014) confirms that the accuracy of created metamodel was sufficient for optimization purposes. However, the obtained optimal values of rail head properties are not acceptable from the practical point of view (practical requirements are: $Fb < 0.02$; $S < 0.12$ µm). It could be overcome by increasing the number of immersions and/or changing of the constraints.

Another way to solve that problem can be foreseen in the use of the dynamic neural network (NARX) in metamodelling of the rail cooling process.Application of the dynamic neural network gives the opportunity to model the variation of the temperature on a rail surface in function of time (during immersion and/or air cooling). Such model gives better prediction of the relationship between the temperature on the surface of the rail head and the evolution of considered parameters (ΔHB, F_b and S). Thus, it can be used in more precise optimization of their values. It can be expected, that the curve of the evolution of the temperature of the rail surface during immersion has a shape similar to that presented in figure 7. Optimization of the variation of the temperature on a rail surface can be seen as a first step in optimization of the cooling process. The second one would be a searchof a cooling technology following the optimal waveform of the temperature evolution. However, elaboration of the dynamic metamodel of the cooling of rails requires further research (measurements of the temperature evolution in function of time), which will be the goal of Authors future research.

ACKNOWLEDGEMENTS

Financial support of the NCN, Poland, project no. R07 0006 10 is acknowledged.

REFERENCES

Bishop, C.M., 1995, *Neural Networks for Pattern Recognition*, Oxford UniversityPress, Oxford.

Haykin, S., 1999, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, New York.

Kusiak, J., Szeliga, D., Sztangret, Ł., 2012, Modellingtechniques for optimizing metal forming processes, in: *Microstructure evolution in metal forming processes*,eds. Lin J., Balint D., Pietrzyk M., Woodhead Publishing Limited, 35-66.

Kuziak, R., Zygmunt, T., 2013, A new method of rail head hardening of standard gauge rails for improved wear and damage resistance, *Steel Research International*, 84, 13-19.

Kuziak, R., Molenda, R., Wrożyna, A., Kusiak, J., Pietrzyk M., 2014, Experimental verification and validation of the phase transformation model used for optimization of heat treatment of rails, *Computer Methods in Materials Science*, 14, (1), 53-63.

Pietrzyk, M., Kuziak, R., 2012, Numerical simulation of controlled cooling of rails as a tool for optimal design of this process, *Computer Methods in Materials Science*, 12, 233-243.

Rauch, Ł., Skiba, M., Kusiak, J., 2014, Computer system dedicated to optimization of production processes and cycles in metal forming industry, *Computer Methods in Materials Science*, 14, (1), 3-12.

Sztangret, Ł., Rauch, Ł., Kusiak, J., Jarosz, P., Małecki, S., 2011, Modeling of the oxidizing roasting process of zinc sulphide concentrates using artificial neural networks, *Computer Methods in Materials Science*, 11, 122-127

Sztangret, Ł., Szeliga, D., Kusiak, J., Pietrzyk, M., 2012, Application of inverse analysis with metamodelling for identification of metal flowstress, *Canadian Metallurgical Quarterly*, 51, 440-446.

Tadeusiewicz, R., 1993, *Sieci neuronowe*, Akademicka Oficyna Wydawnicza, Warszawa, (in Polish).

OPTYMALIZACJA TERMOMECHANICZNYCH PROCESÓW PRZEMYSŁOWYCH WSPOMAGANA METAMODELOWANIEM

Streszczenie

Głównym celem artykułu jest przedstawienie strategii opartej na metamodelu w zastosowaniu do optymalizacji termomechanicznych procesów przetwórstwa metali. Przedstawione podejście polega na zastąpieniu modelu rozważanego procesu poprzez metamodel, co pozwala na osiągnięcie znaczącej redukcji czasu obliczeń. Pomijalnie niski czas obliczeń umożliwia zastosowanie heurystycznych metod optymalizacji, które zwiększają prawdopodobieństwo znalezienia optimum globalnego. W artykule przedstawiono ideę metamodelowania oraz wyniki optymalizacji wybranego procesu przemysłowego.

Received: December 29, 2013
Received in a revised form: January 3, 2014
Accepted: January 5, 2014