JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. 19 No. 1 (2011), pp. 161–169

The Application of Artificial Neural Networks in Optimization of Heat Treatment Processes of Steel

Emilia Wołowiec, Piotr Kula

Technical University of Łódź
Institute of Materials Science and Engineering
Stefanowskiego 1/15, 90-924 Łódź, Poland
emilia.wolowiec@p.lodz.pl, piotr.kula@p.lodz.pl

Abstract. *This article is dedicated to processes of vacuum carburization, which constitute an alternative method to industrial multi-segment carburization. Special attention has been paid to the possibility of using artificial neural networks to design processes of such type. The following subchapters deal with the essence and purposes of vacuum carburization, the course of research on processes and the possibility of using neural networks to design such processes, and the architecture of a sample neural network which achieves this goal.*
Keywords: *neural network, heat treatment, vacuum carburizing.*

1. Introduction

Today's is rapid technical progress of civilization makes devising new technologies a multi-faceted issue. Nowadays, machine and software designers focus on simultaneously selecting construction characteristics and elements of machines, specifying their technological processing, optimal software and the most suitable materials. As a result it is possible to cater to human needs in the best of ways

and at the lowest cost [1]. New possibilities for designing machines and devices are offered by tools which support designer work, including those using artificial intelligence. The advantages and efficiency of those tools make them increasingly popular in this field.

This article gives an account of the research on an instrument for simulating and optimizing vacuum carburizing of steel. Carburization itself is well-researched. However, research on the concomitant carbide forming processes, common for this type of processing, is not satisfactory. The use of neural networking has made it possible to skip the stage of creating a mathematical and physical model of the phenomenon (for which there is no account of exact mathematical relations to be found in literature) and directly pass on to the stage of simulating the material properties on the basis of carburization parameters.

The following subchapters briefly deal with the essence and purposes of vacuum carburization, the research on possible neural network uses and the architecture of a given neural network which is an actual application of the design.

2. Vacuum carburizing

Vacuum carburizing is one of the thermochemical processing methods [2, 3]. Carburization consists in saturating steel with carbon at high temperatures. As a result, an appropriate carbon concentration profile is formed in the surface layer. Normally, it consists of the saturation stage – during which atmospheric carbon is applied to a steel surface, and the diffusion stage – during which the carbon from the steel surface is distributed inside the steel element. The course of the process depends on the treated element (material type, dimensions, condition of the surface), the type and flow of the carbonising atmosphere, the process temperature and pressure, as well as the duration of the saturation and diffusion stages.

The purpose of carburizing is to impart certain mechanical, physical and chemical properties to steal, optimal for a given application. On the surface of the machine a protective layer is formed that is resistant to abrasive wear and contact fatigue, while preserving the necessary core ductility [4].

During experiments with vacuum carburizing it was observed that an inappropriately managed treatment causes the appearance of carbide clusters on the edge of steel grains (the so-called cementite network). A machine element made of such material is automatically rejected as faulty and is not admitted for use, since a cementite network is the most frequent reason for material cracking and the element needs to be quickly replaced.

Of course, thermal treatment can be conducted in a "safe" way that guarantees that there are no carbides. However, this is much more costly in terms of time and energy expenditure than the treatment of the boundary area where the carbide forming begins. Hardening plants all over the world offer safe treatment but they also show a vivid interest in any technology that might reduce the cost of high-volume production.

The above situation seemed to offer a perfect scope for testing neural network possibilities. The idea is to both simulate the poorly researched phenomenon in mathematics and optimise their duration of the whole treatment.

3. The user of neural networks in vacuum carburizing treatment

Several carburizing treatments were performed which produced steel samples that contained or did not contain the disqualifying carbide structures. Next, the samples were metallographically analysed in order to determine the concentration of carbides in the material, depending on the duration and intensity of carburization. Simultaneously, the carbon profile distribution in the samples was analysed. The complete results were collected in a set (approx. 5000) of standards intended for neural network training.

3.1. Determination of the form of the training standards

The research on neural networks began with an analysis of vacuum carburizing and the identification of significant parameters of the process. Based on that, it was determined that the neural network should receive the following input signals: treatment temperature (Temp), carbon saturation stage time (Carb), diffusion stage time (Diff), chemical composition of the material (C, Si, Mn, Cr, Ni, Mo, Al, V, Cu), as well as the distance of the analysed place in the sample from the sample surface (x). The output signals should be the values that describe the carbonised sample: Carbon concentration on the sample surface in percentage terms (Cp), carbon and carbide contents at the analysed depth below the sample surface (Cx and MeC).

The set of all the standards was randomly divided into sets of training, testing and validating standards in the following proportion: 75% – the training set, 15% – the testing set, 15% – the validating set.

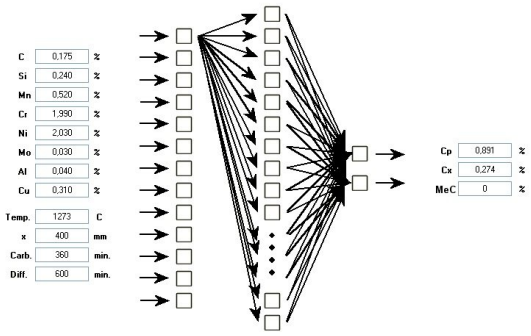

Figure 1. A diagram of the neural network

3.2. Neural network structure and training

Devising the neural network, it was noticed that the network should have extrapolating properties, i.e. it should correctly predict cases that are not contained in the training set. The above condition suggested the choice of the MLP network, considering its possibilities within the scope[4, 5]. For the same reason, the RBF network was rejected. Recurrent networks were not taken into consideration, since during the analysis of carbonisation no connections were noticed that might justify the use of networks involving signal feedback.

Another stage in the network construction was to determine the correct number of layers and neurons in the layers, as well as the choice of the correct activation functions. On the basis of the Kolmogorov and Cybenko theorems [6, 7], it was assumed that in order to solve the problem one hidden layer would suffice.

It was arbitrarily assumed that the hidden layer should not contain fewer than 5 and more than 20 neurons. With this purpose in mind, a few dozen architectures with varying numbers of hidden neurons and different activation functions were tested (the testing concerned the linear, sigmoidal, tangensoidal and exponential functions). The network weights were initiated with random values and the networks were then trained using the steepest descent method, the BFGS method (the Broydena-Fletchera-Goldfarba-Shanno method) and the conjugate gradient method. The training was carried out both with and without weight reduction. However, in the majority of cases, weight reduction caused a serious deterioration of network quality. The error function was assumed to be the sum of squares differences.

Table 1. A summary of the neural network training process

	MLP 12-10-3
SSE (learning)	0.772
SSE (testing)	0.938
SSE (validation)	0.772
Quality (learning)	0.977
Quality (testing)	0.965
Quality (validation)	0.951

Finally, based on the quantity of the error function, the MLP network with 12 input neurons, 10 hidden neurons and three output neurons was chosen (Fig.1.).

The activation function for the hidden neurons was the sinusoidal function and for the output neurons – the exponential function. The best training results were obtained with the BFGS algorithm.

4. Experimental results

The chosen network was trained with the BFGS algorithm and after 138 epochs it virtually reached a steady global error level (after 20 successive epochs the error decreased by a less than 0,0000001). The values that summer the training process are given in Tab.1.

The most significant input signals turned out to be: the distance from the surface (x), the diffusion duration (Diff) and the carbon content in the material before treatment (%C) The network correctly predicted the steel properties on the basis of the process parameters which were provided at the input. (Tab.2).

It was observed that the network acquired the best command of the processes in which the carbide structures were not present or were not numerous. This phenomenon can be explained with the fact that among the training standards there was a small proportion of those with a large number of carbides. Focusing on the numerous cases and a certain downplaying of the untypical cases allow the network to minimise the global training error.

In Fig.2, Fig.3 and Fig.4 are presented the relations acquired by the neural network during its training. The tendencies coincide with the tendencies observed in practice. In fact, the highest numbers of carbides were observed on the surface

Table 2. A juxtaposition of the responses generated by the network and those expected

No.	Sample	Network output	Required answer	Network answer
1.	Testing	% Cp	1,75	1,74
		% Cx	1,39	1,62
		% MeC	4,40	7,02
2.	Validation	% Cp	1,75	1,72
		% Cx	0,20	0,19
		% MeC	0,00	0,00
3.	Validation	% Cp	1,94	2,00
		% Cx	0,26	0,28
		% MeC	0,00	0,00
4.	Testing	% Cp	1,85	1,82
		% Cx	0,20	0,20
		% MeC	0,00	0,00
5.	Learning	% Cp	2,03	1,97
		% Cx	0,54	0,57
		% MeC	0,00	0,00
6.	Learning	% Cp	2,28	2,17
		% Cx	0,18	0,20
		% MeC	0,00	0,00
7.	Testing	% Cp	2,28	2,22
		% Cx	1,49	1,66
		% MeC	6,55	6,37
8.	Validation	% Cp	1,46	1,46
		% Cx	0,18	0,20
		% MeC	0,00	0,00
9.	Testing	% Cp	1,46	1,46
		% Cx	0,18	0,19
		% MeC	0,00	0,00
10.	Learning	% Cp	1,46	1,45
		% Cx	0,53	0,54
		% MeC	0,00	0,01

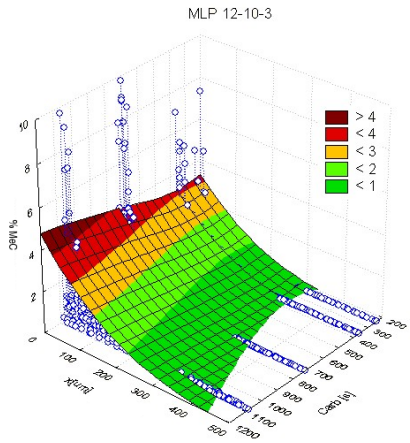

Figure 2. A diagram of the relation of the carbide content to the distance from the surface and the saturation stage duration

and directly under the sample surface. The increase in the number of the carbide clusters corresponded with the duration of the saturation stage (cf. Fig.2.).

The carbon molecules in the material (Cx) are the most numerous close to the surface and as the distance from the surface grows the carbon contents decreases and, as a result, the material carbon profile becomes a bell-shaped curve. However, the longer the diffusion stage lasts, the more carbon penetrates inside the material and the carbon profile assumes a gentler course. In Fig.3. one can see a certain irregularity in the diagram. For these parameters of the process, it is the moment when the carbon concentration in the material falls below the maximal carbon concentration in austenite. The carbide structures become dissolved and the carbon released from the carbides diffuses inside their material causing the lowering of the profile.

Fig.4. depicts the role of the carbonisation and diffusion stages of the vacuum carbonising treatment. A long saturation stage and a short diffusion stage cause carbon accumulation (after exceeding the concentration limit for carbon, carbide cluster formation occurs). A short saturation stage and a long diffusion stage cause a lowering and extension of the whole carbon profile. A prolongation of the diffusion stage, while keeping the same saturation stage duration causes an extension of the carbonised layer thickness (and simultaneously, a fall in the carbon concentration on the surface).

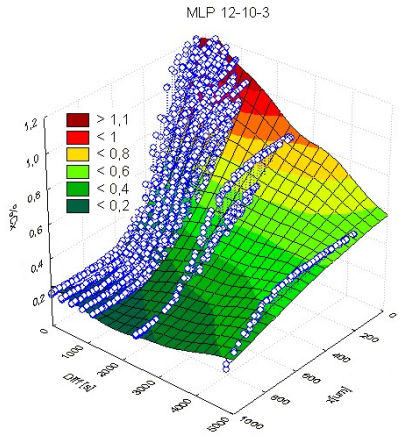

Figure 3. A diagram of the relation of the carbon content to the distance from the surface and the diffusion stage duration

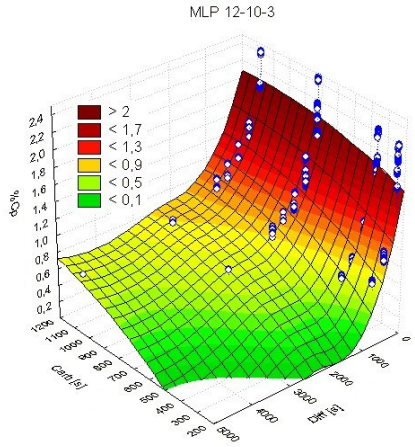

Figure 4. A diagram of the relation of the surface carbon content to the duration of the saturation stage and diffusion stages

5. Conclusions

The devised neural network correctly simulates the real relations and can be applied for the design of the vacuum carburizing treatment processes. The phe-

nomenon of carbide formation and dissolution can also be simulated with an appropriate neural network, thus skipping the creation of a mathematical and physical standard. Unfortunately, when creating a precise simulator, a certain disadvantage is the necessity to carry out numerous experimental processes. Metallographic analysis is also time consuming and expensive. However, it should be stressed that this is a one-time effort.

The use of computer vacuum carburizing treatment simulators enables the design and optimisation of industrial processes without actual technological trials. In practice, it means a reduction in the duration and cost of the treatment during mass production.

References

[1] Sitek, W. and Dobrzański, L., *Computer aiding of material selection on machines parts*, 3rd Scientific Conference on Materials, Mechanical and Manufacturing Engineering. Gliwice, 2005, (in Polish).

[2] Kula, P., *Progress in carburizing technology*, Material Engineering 3 (116). Lodz. 2000.

[3] Kula, P., Olejnik, J., *Some technological aspects of vacuum carburizing*, IProc. XIIth IFMT & SE Congress Vol.3, pp. 195-220. Melbourne, 2000.

[4] Szczepaniak, P., *Intelligent computations, fast transforms and classifiers*, EXIT Academic Publishing House. Warsaw, 2004, (in Polish).

[5] Korbicz, J., Obuchowicz, A., Uciński D., *Artificial neural network. Basis and applications*, PLJ Academic Publishing House. Warsaw, 1994, (in Polish).

[6] Hornik, K., Stinchcombe, M., White, H., *Multi-layer feed forward networks are universal approximations*, Neural Networks, Vol.2. pp 359-366, 1989.

[7] Cybenko, G., *IContinuous valued neural networks with two hidden layers are sufficient*, Technical Report. Tufts University, Medford, MA: Dept of Computer Science, 1988.