International Journal of Minerals, Metallurgy and Materials
Volume 19, Number 7, July 2012, Page 00
DOI: 00

Computer modeling of high-pressure leaching process of nickel laterite by design of experiments and neural networks

Milovan Milivojevic[1)], *Srecko Stopic*[2)], *Bernd Friedrich*[2)], *Boban Stojanovic*[3)], *and Dragoljub Drndarevic*[1)]

1) Technical and Business College, Uzice, Serbia

2) IME Process Metallurgy and Metal Recycling, RWTH Aachen University, Germany

3) Department of Mathematics and Informatics, Faculty of Science, University of Kragujevac, Serbia

(Received: 4 August 2011; revised: 31 December 2011; accepted: 4 January 2012)

Abstract: Due to the complex chemical composition of nickel ores, the requests for the decrease of production costs, and the increase of nickel extraction in the existing depletion of high-grade sulfide ores around the world, computer modeling of nickel ore leaching process became a need and a challenge. In this paper, the design of experiments (DOE) theory was used to determine the optimal experimental design plan matrix based on the D optimality criterion. In the high-pressure sulfuric acid leaching (HPSAL) process for nickel laterite in "Rudjinci" ore in Serbia, the temperature, the sulfuric acid to ore ratio, the stirring speed, and the leaching time as the predictor variables, and the degree of nickel extraction as the response have been considered. To model the process, the multiple linear regression (MLR) and response surface method (RSM), together with the two-level and four-factor full factorial central composite design (CCD) plan, were used. The proposed regression models have not been proven adequate. Therefore, the artificial neural network (ANN) approach with the same experimental plan was used in order to reduce operational costs, give a better modeling accuracy, and provide a more successful process optimization. The model is based on the multi-layer neural networks with the back-propagation (BP) learning algorithm and the bipolar sigmoid activation function.

Keywords: nickel laterite; leaching; computer simulation; design of experiments (DOE); response surface method (RSM); neural networks

1. Introduction

As a result of the permanent increase of nickel production costs related to traditional pyrometallurgical methods and the depletion of high-grade sulfide ores, a renewed interest has developed concerning the production of nickel and cobalt by high-pressure sulfuric acid leaching (HPSAL) of nickel laterite [1]. In early 1970s, Habashi [2-3] made some studies about whether pressure hydrometallurgy was the key to better and nonpolluting processes. In the second of the two-article series [3], he discussed how pressure hydrometallurgy was being applied in the leaching of nickel oxide, sulfide, and arsenide and pointed out the importance of the pressure reactors also known as autoclaves. Nowadays, laboratory autoclaves for hydrometallurgical investigation are available in a variety of sizes, models, and materials for their construction [4].

HPSAL is successfully applied for nickel laterite projects [5-10]. The behavior of the differing minerals from tropical and arid laterite in Western Australia during leaching is being widely examined.

In the Tindall dissertation [5], the fundamentals of HPSAL of iron oxides were examined using synthetic goethite as a model ore. It was found that goethite transformed to hematite by a dissolution-reprecipitation mechanism. The leaching rates were dependent on acid concentration, slurry oxidation potential, and cations in solution. It was also confirmed that nontronite reacted more readily than iron oxides.

Because of a very expensive investigation and expensive corrosion-resistant materials which the autoclave is made of (*e.g.*, titanium), a new challenge is to make a better design

Corresponding author: Srecko Stopic E-mail: sstopic@ime-aachen.de

of experimental investigations and the result prediction in order to minimize the costs. Recently, the different software (Minitab and FactSage) have been used for the design of experiments (DOE), thermochemical analysis, and evaluation of the obtained results in HPSAL of nickel laterite. This approach gives the results that show the main effects plots and interaction plots, providing at the same time the information on the reliability of empirical methods [11-12]. A standard method of using multiple linear regression (MLR) for processing experimental data has been applied. It is described in every textbook on statistics and implemented in every data processing software package [13]. For example, to optimize nickel extraction and iron dissolution from nickeliferous laterite by a process of sulfation-roasting-leaching, Li *et al*. [14] used a response surface method (RSM), which employed a two-level and two-factor full factorial central composite design (CCD) experimental plan. The influence of roasting temperature and reaction time on nickel recovery was studied. The experiments were carried out for fitting two non-linear suitable regression models of nickel extraction and iron dissolution. The obtained models were helpful in predicting the results by performing only a limited set of experiments [14].

Recently, neural networks are included in the newest investigation in order to predict experimental results in hydrometallurgy. In the last decade, this modeling framework has been used by many authors in various fields [15-17]. However, to the knowledge of the authors, the application of this framework to determine HPSAL for nickel laterite is scarce.

In this paper, a computer data-driven modeling of the high-pressure leaching of nickel laterite is studied. The effects of possible processing variables such as reaction temperature, reaction time, mol ratios of sulfuric acid and nickel lateritic ore, and stirring rates are considered. Here, we focus on the DOE theory, MLR, RSM, and back-propagation (BP) neural networks, a class of feed-forward artificial neural network (ANN).

2. Experimental procedure

2.1. Material (Ore "Rudjinci", Serbia)

"Rudjinci" deposits, near Vrnjacka Banja, are the most abundant ones in Serbia. The ore has a low level of metal components and high level of SiO_2. "Rudjinci" nickel ore deposits belong to a group of exogenous nickel deposits, a subgroup of laterite-silicate deposits. The sample of "Rudjinci" ore was previously homogenized with the following composition (wt%): 54.20 SiO_2, 14.90 Fe, 4.00 Al_2O_3, 1.09 Cr_2O_3, 1.13 Ni, 0.06 CoO, 1.40 CaO, 3.22 Mg, 0.48 MnO, 0.05 Na_2O, and 0.05 K_2O.

2.2. Experimental design (DOE)

The plan of experiments investigated four leaching variables associated with four factors for treatment of "Rudjinci" ore [9] (see Table 1). The experiments were carried out using a batch high-pressure leaching method at the IME Process Metallurgy and Metal Recycling of the RWTH Aachen University.

According to literature values of high-pressure leaching variables in industrial conditions [1, 5], the leaching variables given in Table 1 were tested. Regarding the fact that selective leaching is possible in a small interval of temperature, the temperature range of 230 to 250°C was chosen. A further increase of temperature above 250°C leads to an increased pressure and a dangerous oxidation of titanium material, creating a possibility of explosion. A decrease of temperatures leads to the increased leaching of iron, which was not the aim of our research. At temperatures below 230°C, iron and aluminium (in the trivalent state) dissolve and precipitate forming solid products. An acid to ore ratio of about 0.2 was found adequate for Ni and Co leaching of a limonite by Chou *et al.* [18]. Different types of lateritic ores require different acid to ore ratio. Georgiou *et al.* [19] used an acid to ore ratio between 0.15 and 0.35 in a temperature range of 230 to 270°C. Increasing the ratio of sulfuric acid to ore above 0.5 can cause the formation of basic sulfate as a solid product in the system and thus significantly slow the process of diffusion of sulfuric acid in the ore. Because of a reported negligible effect of stirring speed between 400 and 600 r/min, our study proposes an interval between 250 and 1500 r/min in order to investigate the leaching process under the conditions of higher mass transfer and milling effect in the system.

Table 1. Investigated leaching variables of HPSAL for nickel laterite

Variable	Symbol	Minimum	Mean	Maximum
Temperature	T / °C	230	240	250
Acid to ore mass ratio (m_{H2SO4}/m_{ore})	c	0.24	0.32	0.40
Stirring speed	v / (r·min^{-1})	250	875	1500
Leaching time	t / min	30	45	60

Determining the number and arrangement of measuring points in the experimental hyperspace is an essential part of

the modern experiment theory (DOE). The optimal experimental design is a key advantage of this methodology in comparison to other methods of modeling objects, processes, and systems. The studies made by Fisher [20] and Box and Wilson [21] are the theoretical bases for this mathematical modeling.

The criterion used in this paper is the criterion of regression coefficient accuracy, called "D optimality criterion". In general, the design matrix denoted by X, formed on the basis of D optimality criterion, contains a set of boundary points of the input factors' interval, which lie on the border of the hyper sphere, forming a hypercube. Also, the used experimental plans are rotatable and composable and the points are chosen randomly [22-23].

In order to reduce the costs, an experimental design was directed to create linear models. The leaching variables associated with factors X_i (referred to as predictor variables in the context of regression) are transformed into the encoded domain, which is easy for analysis.

Transformation equations have the form:

$$x_i = \frac{X_i - X_{0i}}{W_i}, i = 1, 2, \cdots, k \quad (1)$$

$$X_{0i} = \frac{X_{i\max} + X_{i\min}}{2}, \quad W_i = \frac{X_{\max} - X_{\min}}{2} \quad (2)$$

where k is the number of significant factors, X_i is an actual value of the i-th factor, X_{0i} is the mean for factor X_i (basic level), and W_i is the interval of its variation. In the new scale, the top level of the factors corresponds to +1, the lowest level to −1, and the basic level to 0.

The layout of the experimental points for k≥4 leads to a hypercube and cannot be shown in a 3D space but is mathematically easily described using the matrix. As example, an experimental plan with three factors (k=3) in schematic form is given in Fig. 1. Plans formed in this manner are symmetrical and orthogonal [23].

In order to carry out a dispersion analysis (determining the statistical significance of regression coefficients, the model adequacy, the reliability of the DOE model, *etc.*), there should be n_0 replicates in the central point of the plan, with the assumption of dispersion equivalence ($s^2(y_1) = s^2(y_2) = \cdots = s^2(y_n)$) at every point of the plan.

The plan of leaching experiments was determined by using "2^4 design", generated by the authors of "Fidija" software [24] and additionally verified by "MODeling and DEsign" (MODDE) software package [25]. According to the above, 23 experiments were performed, comprising 16 cube points treatments and 7 replicates in central points. Coordinates of experimental points (T, c, v, t) take coded values −1, 0, or 1.

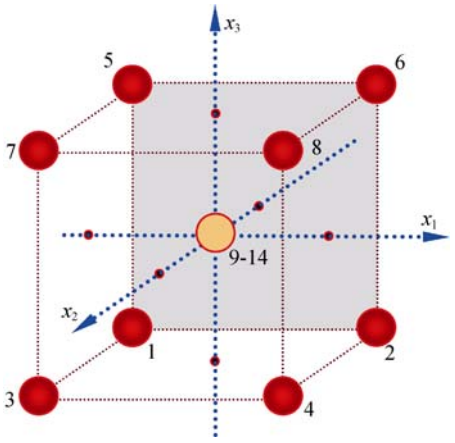

Fig. 1. Experimental layout in a 2^k space (k = 3).

Polynomial regression models require additional treatment. CCD is a successful factorial DOE, which is accomplished by adding experimental points (superposition) along each coordinate axis at opposite sides of the origin and at the distance equal to the semi-diagonal of the hypercube of the factorial design. As an easy way to estimate the response surface, factorial design is the most useful scheme for the optimization of variables with a limited number of experiments [14]. However, additional experiments on each of the coordinate axes would increase the investigation costs.

2.3. Experimental treatment

Under high-pressure leaching conditions, the tests were performed in an autoclave (volume 2 L) manufactured by the Autoclave Engineers, USA.

The temperature was controlled within ±1°C by a temperature control system, manipulated by both an electrical heating mantle and a water cooling system. Agitation was provided by a titanium-made impeller that was magnetically driven. The autoclave was equipped with an acid injection device and a system to withdraw the sample designed by IME. This allows an exact definition of the reaction starting point. A certain amount of laterite was mixed with a pre-calculated amount of deionized water and placed in an autoclave. The slurry was then heated up to a predetermined temperature in the range of 230 to 250°C (corresponding to pressures of 2.8 MPa and 4.0 MPa respectively) under continuous agitation. Upon temperature stabilization, a certain amount of concentrated sulfuric acid (96wt%) correspond-

ing to different acid-to-ore ratios was injected into the autoclave under pressure using the injection device made at IME. Using the sampling system, 20 mL of almost clear liquid was periodically withdrawn through a dip tube and then rapidly cooled. After the end of experiments, solutions aliquots were filtered and analyzed, aiming at Ni, Co, Mg, Al, Fe, and Si by inductively coupled plasma-optical emission spectrometry (ICP-OES).

2.4. Measurements: nickel extraction during leaching of "Rudjinci" ore

A series of leaching experiments were performed using various parameter sets shown in Table 2 (actual leaching variable values and their coded values). The percentages of obtained nickel are related to the maximum content of 1.13% in "Rudjinci" ore.

Table 2. DOE plan matrix: coded and actual predictor variable values and obtained results

Data set		No.	T / °C (230,240,250)		c (0.24,0.32,0.40)		v / (r·min^{-1}) 250,875,1500		t / min (30,45,60)		Extraction of Ni / %
			c.v.	a.v.	c.v.	a.v.	c.v.	a.v.	c.v.	a.v.	a.v.
Data set for linear regressions modeling	Data set for neural networks modeling	1	1	250	1	0.40	–1	250	–1	30	87.50
		2	–1	230	–1	0.24	1	1500	1	60	86.70
		3	–1	230	–1	0.24	1	1500	0	30	75.69
		4	1	250	1	0.40	1	1500	1	60	99.94
		5	–1	230	1	0.40	–1	250	1	60	92.76
		6	1	250	1	0.40	1	1500	–1	30	96.54
		7	1	250	–1	0.24	–1	250	–1	30	54.08
		8	1	250	–1	0.24	1	1500	–1	30	56.78
		9	–1	230	1	0.40	1	1500	–1	30	92.54
		10*	0	240	0	0.32	0	875	0	45	61.65
		11	–1	230	1	0.40	1	1500	1	60	94.02
		12	1	250	–1	0.24	1	1500	1	60	72.63
		13	1	250	–1	0.24	–1	250	1	60	58.84
		14	1	250	1	0.40	–1	250	1	60	81.87
		15	–1	230	1	0.40	–1	250	–1	30	80.73
		16	–1	230	–1	0.24	–1	250	–1	30	72.82
		17	–1	230	–1	0.24	–1	250	1	60	79.85
	Replicates in central point	10*	0	240	0	0.32	0	875	0	45	62.00
		10*	0	240	0	0.32	0	875	0	45	63.00
		10*	0	240	0	0.32	0	875	0	45	60.00
		10*	0	240	0	0.32	0	875	0	45	60.50
		10*	0	240	0	0.32	0	875	0	45	62.00
		10*	0	240	0	0.32	0	875	0	45	62.50

Note: c.v., coded value; a.v., actual value; * repeated experiments in the central point of the experimental plan.

3. Model setup

3.1. Process modeling

A process model is the mathematical structure that maps process variables associated with factors X to the process multivariate output variable Y:

$$\boldsymbol{Y} = \eta(\boldsymbol{X}) \tag{3}$$

For a multivariate input, X is a vector of independent variables (k is a number of significant factors),

$$\boldsymbol{X} = [x_1, x_2, \cdots, x_k]^{\mathrm{T}} \tag{4}$$

where superscript T stands for the transposed matrix (vector).

For a multivariate output, Y would also be a vector. Function η is a response surface.

In this study, the process variables were $x_1 \equiv T$ (temperature), $x_2 \equiv c$ (sulfuric acid to ore ratio), $x_3 \equiv v$ (stirring speed), and $x_4 \equiv t$ (time). Ni_{ext} (degree of nickel extraction) was the model response Y.

Each data set consists of a pair of input values (T, c, v, t) and an output value (Ni_{ext}).

Data-driven modeling is very varied, ranging from MLR and RSM [14], cube and B-splines, fuzzy logic systems, and neural networks [26-28] to heuristic algorithms such as a genetic algorithm or a support vector machine [29]. Some of these models have an analytical form of a response surface η (MLR, B-spline, *etc.*), but others do not (neural networks, genetic algorithms, *etc.*). Here, we focus on MLR and ANN with a BP learning algorithm.

The main goal in MLR is to determine the functional dependence between input and output values in analytical form. ANN is the generalization of the idea of regression models. The main advantage of ANN versus classic statistical regression analysis is that networks are more general functional forms. This is especially important in a multidimensional space where regression techniques often cannot produce adequate approximation. Also, ANN allows simultaneous modeling multivariate output vectors and minimization of errors for each individual response. Neural network models have the potential to also represent craggy response surfaces provided that enough mesh points are disposable [13]. However, these two methods are complementary in many aspects.

3.2. Linear regression models

MLR is a standard method to process experimental data. Linearity refers to the unknown parameters, not to the regressors. Linear regression models can result in a non-linear response surface. The response surface represents a linear combination of regression coefficients β_i and function forms $f_i(\boldsymbol{X}^{\mathrm{T}})$, named "basis functions":

$$\eta = \eta(\boldsymbol{\beta}, \boldsymbol{X}) = \sum_{i=0}^{d} \beta_i \cdot f_i(\boldsymbol{X}^{\mathrm{T}}) \tag{5}$$

where d is the number of basis functions.

In MLR regression, the unknown coefficients β_i ($i = 0,1,\cdots,d$) are estimated by values b_i ($i = 0,1,\cdots,d$) obtained by using experimental results. Coefficients b_i determine the empirical regression model:

$$\hat{y} = \hat{\eta}(x_1, x_2, \cdots, x_k, b_1, b_2, \cdots, b_d) \tag{6}$$

In this work, we used the following models in linear polynomial forms.

a) First-order MLR,

$$y = b_0 + \sum_{i=1}^{k} b_i x_i \tag{7}$$

b) First-order MLR with interaction effects,

$$y = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} b_{ij} x_i x_j + \sum_{i=1}^{k-2} \sum_{j=k+1}^{k-1} \sum_{l=j+1}^{k} b_{ijl} x_i x_j x_k \tag{8}$$

where k is the number of input factors ($k = 4$).

Regression analysis determines parameters of empirical mathematical models. The essence of the method is the minimum of sum squared residuals. Dispersion analysis focuses on reliability. Fisher's test was used for testing the adequacy of the model (F-test) [20-23].

Statistical data analyses in this work were performed by Fidija regression modeling software.

For every model from Table 3, the best estimation of the regression parameters is given together with its 95% confidence intervals. Also, there are values of F-tests that show significant differences between theoretical and calculated values of the F-statistics. Results in Table 3 show that none of the developed linear models has satisfying accuracy, *i.e.*, they are not adequate.

Second-order terms of polynomial functions are not used since they would require the expansion of the experimental plan with new experiments, according to the CCD strategy.

3.3. Neural network model

In order to reduce further investigation costs for HPSAL of nickel laterite, a neural network approach in process modeling (based on the same experimental plan) was used.

The ANN is a simplified mathematical model of natural neural network analogous to biological neurons. Neurons (nodes) are interconnected to weighted links. The number of layers and neurons determines the network complexity (see Fig. 2). Weights are usually adjustable and can be trained through a learning process and training example. The neural feed-forward networks in general and neural networks with a BP algorithm in particular are described in detail in Refs. [26-28].

Typically, the first challenge when designing ANN is to determine the appropriate architecture, especially the number of hidden neurons. If this number is too small, there are

Table 3. Linear regression model structures, values of the regression parameters and their 95% confidence interval, values of *F*-tests, as well as the standard deviation of residuals σ_{res}

First-order linear regressors			
Model: $y=\beta_0+\beta_1 \cdot x_1+\beta_2 \cdot x_2+\beta_3 \cdot x_3+\beta_4 \cdot x_4$			
β_0	0.74541	[0.7375 … 0.7533]	$F_t(f_R,f_E,\alpha)=F_t(12,6,0.05)=4$,
β_1	–0.04183	[–0.0513 …–0.0320]	$F_{rLF}=79.64$,
β_2	0.10532	[0.0959 … 0.1148]	$s^2(y)=0.0002382$,
β_3	0.04149	[0.0320 … 0.0509]	$s_{LF}^2=0.018971$,
β_4	0.03122	[0.0218 … 0.0407]	$F_{rLF}>F_t$ is not adequate,
			$\sigma_{res}=0.1377353$
Multiple linear regressors with interaction effects (significant terms only)			
Model: $y=\beta_0+\beta_1 \cdot x_1+\beta_2 \cdot x_2+\beta_3 \cdot x_3+\beta_4 \cdot x_4+\beta_5 \cdot x_1 \cdot x_2+\beta_6 \cdot x_1 \cdot x_3+\beta_7 x_1 \cdot x_4+$ $\beta_8 \cdot x_2 \cdot x_3+\beta_9 \cdot x_2 \cdot x_4+\beta_{10} \cdot x_3 \cdot x_4+\beta_{11} \cdot x_1 \cdot x_2 \cdot x_3+\beta_{12} \cdot x_1 \cdot x_2 \cdot x_4+\beta_{13} x_1 \cdot x_3 \cdot x_4+\beta_{14} \cdot x_2 \cdot x_3 \cdot x_4$			
β_0	0.74541	[0.7375 ... 0.7533]	$F_t(f_R,f_E,\alpha)=F_t(5,6,0.05)=4.39$,
β_1	–0.04183	[–0.0513 ... –0.0320]	$F_{rLF}=145.63$,
β_2	0.10532	[0.0959 ... 0.1148]	$s^2(y)=0.0002382$,
β_3	0.04149	[0.0320 ... 0.0509]	$s_{LF}^2=0.03459$,
β_4	0.03121	[0.0218 ... 0.0407]	$F_{rLF}>F_t$ is not adequate,
β_5	0.04908	[0.0396 ... 0.0585]	β_7, β_8, β_{10}, β_{11} are not significant,
β_6	0.01301	[0.0036 ... 0.0225]	$\sigma_{res}=0.1862538$
β_9	–0.01711	[–0.0266 ... –0.0077]	
β_{12}	–0.01144	[–0.0209 ... –0.0020]	
β_{13}	0.01668	[0.0072 ... 0.0261]	
β_{14}	–0.01037	[–0.0198 ... –0.0009]	
Linear regressors with minimum F_{rLF} (significant terms only)			
Model: $y=\beta_0+\beta_1 \cdot x_1+\beta_2 \cdot x_2+\beta_3 \cdot x_3+\beta_4 \cdot x_4+\beta_5 \cdot x_1 \cdot x_2$			
β_0	0.74541	[0.7375 ... 0.7533]	$F_t(f_R,f_E,\alpha)=F_t(11,6,0.05)=4.035$,
β_1	–0.04183	[–0.0513 ... –0.032]	$F_{rLF}=72.17$,
β_2	0.10532	[0.0959 ... 0.1148]	$s^2(y)=0.0002382$,
β_3	0.04149	[0.0320 ... 0.0509]	$s_{LF}^2=0.017192$,
β_4	0.03121	[0.0218 ... 0.0407]	$F_{rLF}>F_t$ is not adequate,
β_5	0.04908	[0.0396 ... 0.0585]	$\sigma_{res}=0.1311171$

not enough degrees of freedom to represent data. If it is too high, overfitting occurs [28]. In this study, based on Kolomogorov's theorem, a network with one hidden layer and nine neurons was chosen. Fig. 2 describes its structure with the following parameters: X is the vector matrix of predictor variables, Y is the response, w_{ih} is the hidden layer weight matrix, and w_{ho} is the output layer weight matrix.

The learning process is performed in iterative cycles. Each cycle consists of two stages: a step forward and a step backward. The step forward implies calculating the output value of the network (y_{1calc}) for the given inputs. It starts from the first layer that does not perform any processing of input values (x_1, x_2, x_3, and x_4) but only links the nodes from the first layer to all process elements in the hidden layer. One process element with a value of 1 is added (bias) into the input and hidden layers. Weighing coefficients w_{ij} are initialized in a random manner from the interval [-0.5, 0.5]. Total input α_j in process element j (in the hidden layer) is a weighted sum of all inputs x_i:

$$\alpha_j=\sum_{i=0}^{N_i} x_i w_{ij}, j=1,2,\cdots,N_h$$

(in our model, $N_i=4, N_h=9$) (9)

where w_{ij} is the weight of the link between neurons i and j, N_i is the number of input neurons, and N_h is the number of hidden neurons.

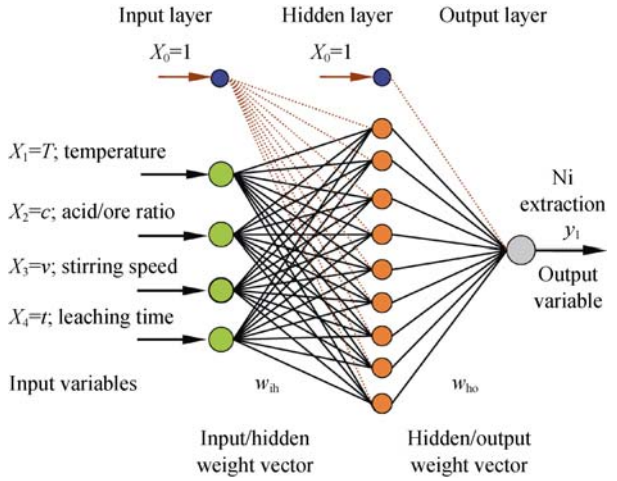

Fig. 2. Signal flow in BP neural networks with 4 input, 9 hidden, and 1 output neurons.

The output from process element *j* in the hidden layer is obtained using the corresponding activation (transfer) function *g*(*x*).

In this study, "bipolar sigmoid transfer function" is used (see Fig. 3).

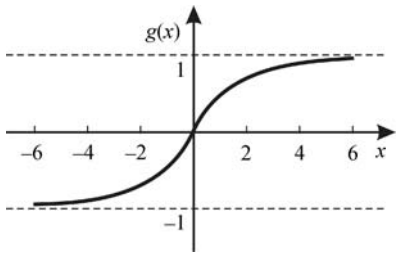

Fig. 3. Bipolar sigmoid transfer function.

Accordingly, the output of process element *j* in the hidden layer will be

$$x_j = g(\alpha_j) = \frac{1-\mathrm{e}^{-\alpha_j}}{1+\mathrm{e}^{-\alpha_j}} \tag{10}$$

Finally, we have the output from process element *r*, which is the output of neural networks:

$$y_r = g(\alpha_r) = \frac{1-\mathrm{e}^{-\alpha_r}}{1+\mathrm{e}^{-\alpha_r}} \quad \text{(in our model } r\text{=1)} \tag{11}$$

Then, the mean sum squared error (MSSE) of the output vector with reference to measurements is calculated:

$$E_r = \frac{1}{P}\sum_{t=1}^{P}\frac{1}{2}\left(y_{rt}^{\text{meas}} - y_{rt}^{\text{calc}}\right)^2 \tag{12}$$

where t = 1, 2,⋯, P, and P is the number of input/output pairs.

In order to perform the weight correction, we need to determine the error vector, successively for all layers and all nodes. According to the evidence detailed in Ref. [27], the error for output layer will be

$$\delta_r = y_r^{\text{calc}} \cdot \left(1 - y_r^{\text{calc}}\right) \cdot \left(y_r^{\text{meas}} - y_r^{\text{calc}}\right) \tag{13}$$

where δ_r is the error vector of output nodes, y_r^{calc} is the output value calculated by ANN, and y_r^{meas} is the target of the output layer. Derivates of the complex functions give error in the hidden layer (in the following equation the index *h* stands for the hidden layer):

$$\delta_q^{(h)} = x_q^{(h)} \cdot (1 - x_q^{(h)}) \cdot \sum_{r=1}^{m} \delta_r w_{qr}(t) \tag{14}$$

where index *q* denotes any of hidden nodes, *m* is the number of output neurons, $\delta_q^{(h)}$ is the error vector of hidden nodes, $x_q^{(h)}$ is the output of each hidden node, δ_r is the error vector of output nodes, and $w_{qr}(t)$ is the weight vector of the output layer, which is taken from previous iteration, according to Ref. [27].

The correction of weights is performed in the error propagation, back from the output to the input layer. The correction of the weight w_{ij} for process elements *i* and *j*, between two layers in iteration *t*+1, is done as follows:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \cdot (1-\mu) \cdot \delta_j \cdot x_i + \eta \cdot \mu \cdot (w_{ij}(t) - w_{ij}(t-1)) \tag{15}$$

where $w_{ij}(t)$ is the weight before modification, $w_{ij}(t+1)$ is the weight after modification, η is the learning rate (rate of convergence between the current solution and the global minimum), δ_i is the error for process element *j*, x_i is the output value of process element *i*, μ is the momentum that helps the network to overcome the local minima, and $w_{ij}(t-1)$ is the weight of two iterations before the current iteration. The speed of convergence cannot be significantly affected by increasing the learning rate η because it can lead to "skipping" the error minimum and the oscillation of the system.

In this work, the learning rate η takes the value of 0.8, while momentum constant μ takes the value of 0.9. All inputs and outputs (see Fig. 2) were normalized in the interval [–1,1] using linear scaling.

The process of weight correction is performed in a series of successive learning cycles. The training process can be terminated by any of the following criteria: total time of learning, total number of learning cycles, or criteria defined by learning error. In this article, the criterion for interrupting the training process is defined by the maximal number of learning cycles (50000 cycles). At the end of the learning

process, neural network knowledge is encoded in the weighing coefficients.

The BP algorithm and the learning method described above are the basis of authors' neural modeling software Neuro2010 [30].

Since the performance of BP can depend on initial values of weights and on the details of the training process, in this study, the modeling process is repeated more than 30 times. At the end of the repeated training procedure, the most successful network (with the lowest learning error) is chosen as the relevant. The obtained value of MSSE is E=1.7×10^{-6}. The comparison of experimental and calculated values obtained for 17 experimental points is given in Table 4.

Table 4. Experimental and ANN model results in 17 points

No.	1	2	3	4	5	6	7	8	9
T/°C	250	230	230	250	230	250	250	250	230
c	0.4	0.24	0.24	0.4	0.4	0.4	0.24	0.24	0.4
v/(r·min^{-1})	250	1500	1500	1500	250	1500	250	1500	1500
t/min	30	60	30	60	60	30	30	30	30
$y_{measured}$	0.875	0.867	0.7569	0.9994	0.9276	0.9654	0.5408	0.5678	0.9254
$y_{calculated}$	0.875000	0.867001	0.756899	0.998373	0.927605	0.965414	0.541480	0.567802	0.925399
No.	10	11	12	13	14	15	16	17	
T/°C	240	230	250	250	250	230	230	230	
c	0.32	0.4	0.24	0.24	0.4	0.4	0.24	0.24	
v/(r·min^{-1})	875	1500	1500	250	250	250	250	250	
t/min	45	60	60	60	60	30	30	60	
$y_{measured}$	0.6165	0.9402	0.7263	0.5884	0.8187	0.8073	0.7282	0.7985	
$y_{calculated}$	0.616490	0.940207	0.726299	0.588395	0.818699	0.807301	0.728199	0.798502	

Fig. 4 shows the MMSE during the learning process as a function of the number of cycles. Fig. 5 gives a comparison between calculated and measured values in several learning steps.

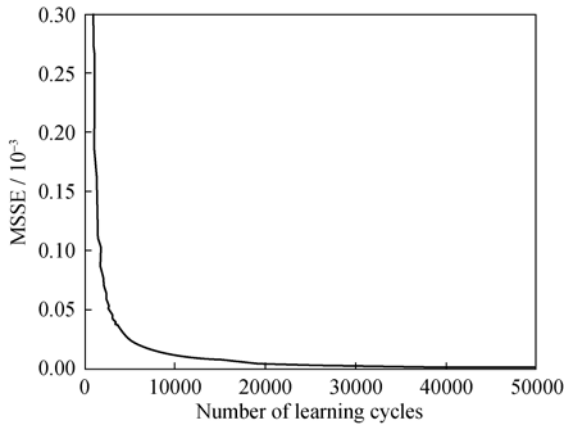

Fig. 4. MSSE during the learning process.

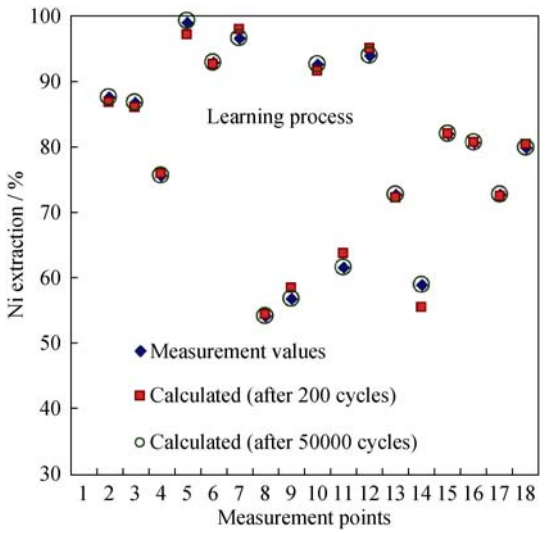

Fig. 5. Comparison of measured and ANN values after 200 and 50000 learning cycles.

4. Results and discussion

4.1. Model analysis

Since the objective function (Ni extraction) depends on four parameters, 3D representation of the response surface is not possible. However, there are 3D plots of the response surface obtained by using neural network models for some selected constant factor values. Some typical diagrams are shown in Figs. 6-8.

As shown in Fig. 6, at a constant temperature of 250°C and at a constant leaching time of 60 min, increasing the c

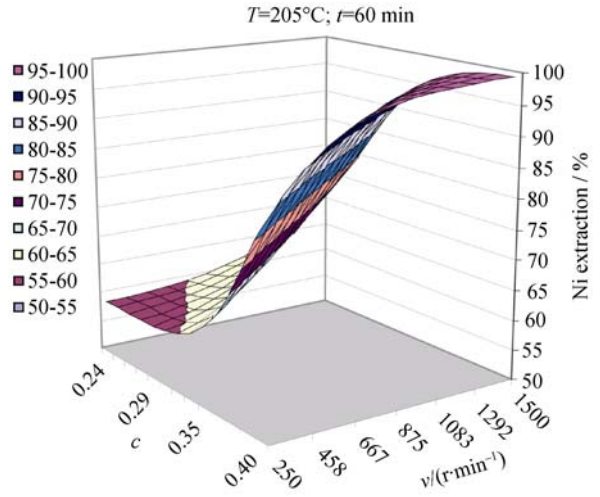

Fig. 6. Response surface of leaching experiment at T=250°C and t=60 min.

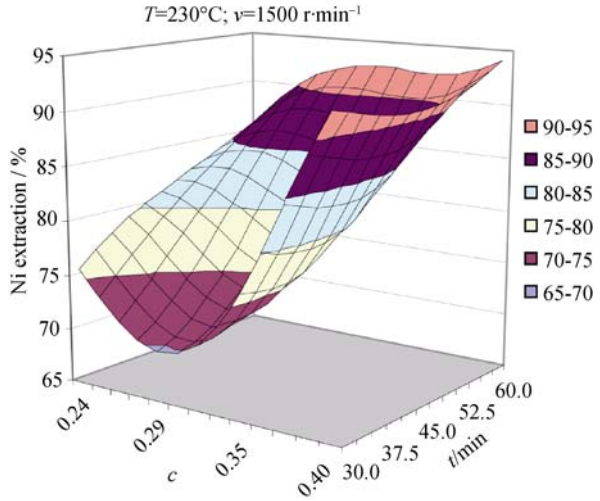

Fig. 7. Response surface of leaching experiment at T=230°C and v=1500 r·min^{-1}.

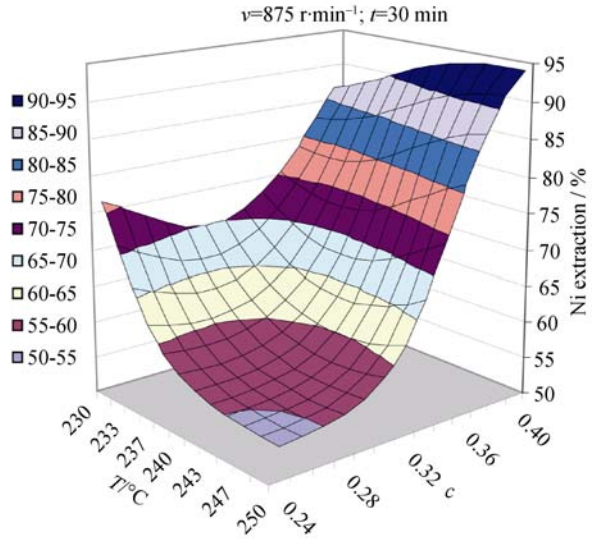

Fig. 8. Response surface of leaching experiment at v=875 r·min^{-1} and t=30 min.

ratio and stirring speed can increases the extraction of nickel. At a lower c ratio up to 0.30 and a stirring speed between 250 and 500 r/min, the extraction of nickel is insufficient. A further increase of c up to 0.4 and stirring speed to 1500 r/min leads to the increasing of nickel extraction. In contrast to an experimental work that gives only one result, this computer model of HPSAL offers results for arbitrary combinations of reaction parameters. It helps us to understand the changes in the system and the influence of various experimental conditions without performing new experiments.

At the minimum chosen temperature and maximum stirring speed (see Fig. 7), the nickel extraction decreases by increasing the c ratio between 0.2 and 0.3. This is because of an expected formation of basic sulfate during leaching. These results should be tested in further experimental work. A further increase of c leads to an increase of nickel extraction.

At a constant stirring speed of 875 r/min in 30 min (see Fig. 8) and a minimum c ratio of 0.24, the increase of temperature between 230 and 250°C leads to the decrease of nickel extraction. If the c ratio is increased from 0.24 to 0.30, the trend of decrease of nickel extraction stays without changes. This behavior could be explained by insufficient stirring in the system. Under the same conditions and a highest c ratio about 0.35 to 0.4, the increase of temperature leads to an increasing nickel extraction.

In order to obtain a maximum percentage of nickel regardless of costs, the regimens in Fig. 6 should be used. However, it can be seen from Fig. 8 that the following parameters give a sufficient percentage of nickel (93%-95% from 1.13% available in nickel ore) with less costs (shorter reaction time and lower temperature): temperature of 240°C, sulfuric acid to ore mass ratio of 0.35 to 0.4, stirring speed of 875 r/min, and leaching time of 30 min,. Fig. 8 suggests that the further increasing of temperature and acid concentration would increase the extraction of nickel additionally. However, as explained earlier (section 2.2), designing experiments in this direction would probably lead to dangerous oxidation and explosion. Some researchers have worked at a temperature interval of 270 to 280°C, but that is the subject of further technical analysis in terms of work safety and process costs.

4.2. Sensitivity analysis

The software generated in this work can test different combinations of reaction parameters in order to investigate their influence on nickel extraction and propose an optimal experimental setup. Using ANN, sensitivity analysis was

carried out to determine the relative significance of each of the input parameters. The sensitivity analysis method is described by Rashidi *et al.* [17].

The sensitivity for each input factor was calculated using the following equation:

$$\overline{s}_j = \frac{1}{N}\sum_{i=1}^{N} s_j^{(i)}, j = 1,2,\cdots,k \tag{16}$$

where k is the number of input factors and N is the number of measurements. Sensitivity $s_j^{(i)}$ is the absolute change of the percent of Ni when the j-th input factor is varied while other factors remain constant and equal to those in the i-th observation.

Values of $s_j^{(i)}$ are determined by the following relationship:

$$s_j^{(i)} = f_{\mathrm{ANN}}\left[x_j^{(i)} + \frac{\Delta}{100}\cdot(x_j^{(\max)} - x_j^{(\min)})\right] \tag{17}$$

where $x_j^{(i)}$ is the i-th value of the input parameter x_j, Δ is the amount of the input parameter change (in percents), $x_j^{(\max)} - x_j^{(\min)}$ is the range of the j-th input factor and f_{ANN} is the "mapping function" defined by weights of the developed ANN model of HPSAL for nickel laterite.

The mean absolute changes of obtained Ni for Δ equal to –20%, –10%, 10% and 20% are presented in Fig. 9.

As shown in Fig. 9, sulfuric acid to ore mass ratio and stirring speed have a big influence on the leaching process. The influence of temperature between 230 and 250°C is not of big importance. However, it should be noted that the influence of temperature is inverse. After 30 min of starting the process, the high amount of nickel is dissolved. Under these conditions, temperature increasing leads to forming a sulfated compound that probably contains nickel, which can be the explanation for such an inverse influence.

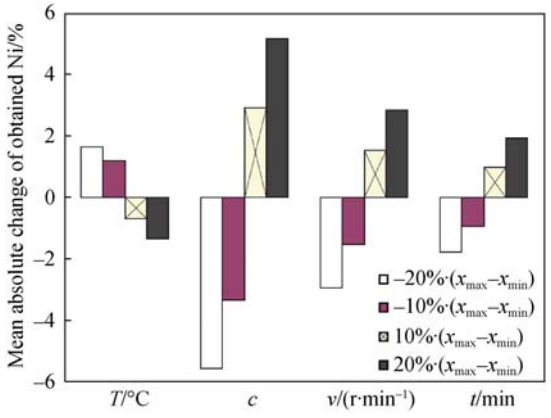

Fig. 9. Sensitivity analysis of HPSAL for nickel laterite.

4.3. Comparison of regression models and neural network

It can be concluded (see Table 3) that none of linear regression models with linear regressors are appropriate to model the process. The comparison between a standard deviation of residuals for the best linear regression model and a neural network model based on the BP algorithm is shown in Table 5.

Table 5. Comparison of σ_{res} for the best linear regression model and neural BP model

Estimated measurement error	$s^2(y)$=0.0002382
Linear regressors with minimum F_{rLF} (significant terms only) Model: $y=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+\beta_4 x_4+\beta_5 x_1 x_2$	σ_{res}=0.1311171, $F_t(f_R,f_E,\alpha)=F_t(11,6,0.05)$=4.035, F_{rLF}=72.17
BP neural network model with 9 hidden neurons after 50000 learning cycles	$E = 0.0017\times10^{-3}$, σ_{res}=0.000411

The given residual dispersion of the applied neural network model ($\sigma_{\mathrm{res\text{-}ANN}}$=0.000411) compared with the residual dispersion of the regression model ($\sigma_{\mathrm{res\text{-}MLR}}$=0.1311171) indicates a significantly higher accuracy of the model based on ANN.

5. Conclusion

Based on the DOE theory, MODDE and Fidija software packages have been successfully used in planning experiments in the improvement and development of methods related to the production of nickel by HSPAL. RSM and CCD strategies to model the process with MLR were used. The complexity of relations in the leaching process with four factors causes the proposed linear regression models not to be adequate. Therefore, the analysis with neural networks came as a subsequent step. The model is based on ANN with a BP learning algorithm and a bipolar sigmoid transfer function. Using software Neuro2010, the process was modeled with high accuracy (E=0.0017×10^{-3}, σ_{res}= 0.000411). Sensitivity analysis was carried out to determine the relative significance of each of the input variables. It was found that sulfuric acid to ore ratio and stirring speed are the most important variables in the system. The results also show an indirect influence of temperature in the chosen interval.

The developed software can test different combinations of predictor variables. The analyses lead to the following conclusions. (1) The maximum percentage of nickel, regardless of operational costs, can be obtained at the following values: $T \approx$247-250°C, $c \approx 0.4$, $v \approx$1437-1500 r/min, and $t \approx$58.5-60 min. (2) If safety and operational costs are taken into account, a sufficient percentage of nickel (93%-95% from 1.13% available in nickel ore) can be obtained using values of factors: $T \approx 240$°C, $c \approx$ 0.35-0.4, $v \approx 875$ r/min, and $t \approx 30$ min.

These values correspond to the actual parameter domain and suggest that the resulting model based on a BP neural network is successfully applied. The results of modeling would have been even better if the number of experimental data had been made larger. The lack of data testing set can be overcome in further research in order to improve the presented ANN model for HPSAL of nickel laterite; especially, the self-interactive of parameters using recurrent and different kinds of networks will be investigated. A detailed analysis of operational costs and environmental impact can be the subject of our further techno-economical research.

References

[1] B.I. Whittington and D. Muir, Pressure acid leaching of nickel laterites, *Miner. Process. Extr. Metall. Rev.*, 21(2000), p.527.

[2] F. Habashi, Pressure hydrometallurgy: Key to better and non-polluting process, Part 1, *Eng. Min. J.*, 172(1971), No.2, p.96.

[3] F. Habashi, Pressure hydrometallurgy: Key to better and non-polluting process, Part 2, *Eng. Min. J.*, 172(1971), No.5, p.88.

[4] F. Habashi, *Laboratory Autoclaves for Hydrometallurgical Research, Extraction & Processing Division*, P.R. Taylor, ed., The Minerals, Metals, & Materials Society, Nashville, 2000, p.411.

[5] G.P. Tindall, *High Temperature Acid Leaching of Western Australian Laterites* [Dissertation], Murdoch University, 1998, p.214.

[6] D.R. Vucurovic, *Study of Possibility of Nickel Valorization from its Ore in Serbia* [Report], Faculty of Technology and Metallurgy, Belgrade, 1983, p.271.

[7] S. Stopic, B. Friedrich, and R. Fuchs, Kinetics of leaching of nickel lateritic ore at an atmospheric pressure, *J. Metall.*, 7(2003), p.235.

[8] S. Stopic, B. Friedrich, and R. Fuchs, Sulphuric acid leaching of the Serbian nickel lateritic ore, *Erzmetall*, 56(2003), p.204.

[9] S. Stopic, High pressure leaching of nickel lateritic ore, *Humboldt Kosmos*, 82(2003), p.34.

[10] S. Stopic, B. Friedrich, and N. Anastasijevic, Kinetics of high pressure leaching of the Serbian nickel lateritic ore, [in] *Proceeding of EMC 2003, Volume 1, Copper and Nickel*, Hanover, 2003, p.189.

[11] S. Stopic, B. Friedrich, N. Anastasijevic, and A. Onija, Experimental design approach regarding kinetics of high pressure leaching processes, *J. Metall.*, 9(2003), p.273.

[12] R.V. Lenth, Quick and easy analysis of unreplicated factorials, *Technometrics*, 31(1989), p.469.

[13] M. Leparoux, M. Loher, C. Schreuders, and S. Siegmann, Neural network modeling of the inductively coupled RF plasma synthesis of silicon nanoparticles, *Powder Technol.*, 185(2008), p.109.

[14] D. Li, K.H. Park, Z. Wu, and X.Y. Guo, Response surface design for nickel recovery from laterite by sulfation-roasting-leaching process, *Trans. Nonferrous Met. Soc. China*, 19(2010), Suppl. 1, p.92.

[15] E.O. Ezugwi, D.A. Fadore, J. Bonney, K.B. Da Silva, and W.F. Sales, Modeling the correlation between cutting and process parameters in high-speed machining of Inconel 718 alloy using artificial neural network, *Int. J. Mach. Tool. Manuf.*, 45(2005), p.1375.

[16] M. Younesi, M.E. Bahrololoom, and M. Ahmadzadeh, Prediction of wear behaviors of nickel free stainless steel-hydroxyapatite bio-composites using artificial neural networks, *Comput. Mater. Sci.*, 47(2010), p.645.

[17] A.M. Rashidi, A.R. Eivani, and A. Amadeh, Application of artificial neural networks to predict the grain size of nanocrystalline nickel coating, *Comput. Mater. Sci.*, 45(2009), p.499.

[18] E.C. Chou, P.B. Queneau, and R.S. Rickard, Sulfuric acid pressure leaching of nickeliferous limonites, *Metall. Trans. B,* 8(1977), p.547.

[19] D. Georgiou and V.G. Papangelakis, Sulphuric acid pressure leaching of a limonitic laterite: chemistry and kinetics, *Hydrometallurgy*, 49(1998), p.23.

[20] R. Fisher, *Design of Experiments,* 9th ed., Reprinted, Hafner, New York, 1974.

[21] G.E.P. Box and K.B. Wilson, On the experimental attainment of optimum conditions, *J. Roy. Statist. Soc. Ser. B*, 13(1951), p.1.

[22] J. Stanic, *Methods of Engineering Measurements* [Dissertation], Faculty of Mechanical Engineering, University of Belgrade, 1990.

[23] M. Milivojevic, *Mathematical Modeling of Real Objects, Processes and Systems by Means of Central Compositional Symmetric Experimental Plans* [Dissertation], Faculty of Mechanical Engineering, University of Belgrade, 1996.

[24] M. Milivojevic and D. Drndarevic, *Package FIDIJA, Design of Experiments & Regression and Dispersion Analyses Package*, High Technical and Business School, Uzice, Serbia, 1996.

[25] Software MODDE 5.0 (Version 5.0, 9 Nov. 1999), Umetrics AB, USA, 1999.

[26] J.L. Mc Clelland and D.E. Rumelhalt, *Explorations in Parallel Distributed Processing*, MIT Press, Cambridge, 1988.

[27] D. Drndarevic, *Modeling and Optimization of Powder Metallurgy Process by Neural Networks* [Dissertation], Faculty of Mechanical Engineering, University of Belgrade, 1996.

[28] Z.S. Yu, *Feed-Forward Neural Network and Their Application in Forecasting*, Beyond Dream, 2001.

[29] J.A. Cruz and D.S. Wishart, Applications of machine learning in cancer prediction and prognosis, *Cancer Inf*, 2(2006), p.59.