Research of Neural Network Models Regression Depth of Technological Parameters Identification for Magnetite Quartzite Beneficiation Process

Peter Zholondiyevsky, Andrey Kupin, Kuztetsov Denis, Bogdan Poddubny

Abstract— **Based on autoregressive models of neural network identification the equation of one-step and multi-step predictor variables showed on the basis of NNARX, NNOE models. Research of the influence of the depth of regression on the quality of the identification and on the speed of convergence was performed. The results of identification modelling on real performance of beneficiation process are given below.**

Index Terms— **identification; NNARX; NNOE; beneficiation process, magnetite quartzite.**

I. INTRODUCTION

Opening question regarding the identification of technological process (TP) parameters of magnetite quartzite (iron ore) beneficiation process in industrial environments for next intelligent regulator synthesis for all production stages coordinated control. The relevance and the statement of this problem are given in the articles of the authors [1-2].

Previous researches results showed that neural networks approach can be chosen for identification of technological parameters of beneficiation process under conditions of ore concentration plants. Required quality and convergence of the model can be reached, in case of correct selection of modelling structure basis. There is also an open issue of some parameters of neural network models selection, which directly influence on the quality and speed of the system training (parametrization). Resulting from the described above, the purpose of current article is the definition of optimal time delay values of input and output signals for autoregressive models of NNARX, NNOE [2-3] type.

II. MODELS OF IDENTIFICATION AND PROBLEM STATEMENT

In accordance with [4-6] TP of beneficiation process considered as multidimensional discrete system of the form:

$$Z^P = \{[u(k),\ y(k)],\quad k = \overline{1, N}\}, \quad (1)$$

where P is order of system; u(k) is vector of input control signals; y(k) is vector of output signals of system; N is maximal quantity of discrete supervision readouts; k is discrete time.

Autoregressive NNARX and NNOE models are mostly widespread now. According to [3], they are recommended for use in case of a limited set of input statistics (e.g. creating an automated process control system).

Generally, the mathematical models (in other way, approximators, predictors or regressors) based on artificial neural networks are represented by nonlinear equation of the form:

$$\hat{y}(k+1) = NN\begin{pmatrix} y(k), y(k-1),..., y(k-l_1), \\ u(k), u(k-1),..., u(k-l_2) \end{pmatrix}, \quad (2)$$

where ŷ is output signals vector of identificational model; NN(·) is resulting transformation of "input-output" type, which is performed by the neural network; l1 is regression depth (the number of delay signals) of feedback on the output of the model; l2 is input delay depth.

This is one-step predictor or NNARX model (Neural Network based AutoRegressive eXogenous signal). Such predictor allows forecast one step forward and can be successfully used for identification.

To solve this problem, neural covariates models, which allow to perform multi-step forecasts (short-term predictors) [6-7] are quite successfully used too.

NNOE (Neural Network Output Error) is one of such models, which defined by recursive equation of the form:

$$\hat{y}(k+1) = NN\begin{pmatrix} \hat{y}(k), \hat{y}(k-1),..., \hat{y}(k-l_1), u(k), \\ u(k-1),..., u(k-l_2) \end{pmatrix}. \quad (3)$$

As model structure of approximator, architecture based on the recurrent network with memory is used (Figure 1). According to the assumptions, stated in [8-11], these particular neural network structures are quite effective for modelling of beneficiation TP.

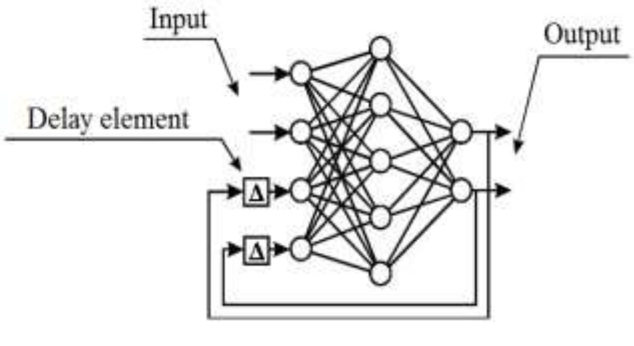

Figure 1. Analyzed structures of neural networks

III. TECHNIQUE OF RESEARCHES

For parameterization (training) of models types (2)-(3) certain amount of statistical information is needed, namely data templates of "input-output" type. These data templates represented as multitude:

Peter Zholondiyevsky, Kryvyi Rih National University (Master Student), Ukraine, 50027, V.Matusevicha str., 11
Andrey Kupin, Kryvyi Rih National University (Department of Computer Systems and Networks), Ukraine, 50027, V.Matusevicha str., 11
Kuztetsov Denis, Kryvyi Rih National University (Department of Computer Systems and Networks), Ukraine, 50027, V.Matusevicha str., 11
Bogdan Poddubny, Kryvyi Rih National University (Post Graduated Student), Ukraine, 50027, V. Matusevicha str., 11

$$\Xi\{z_i, d_i\}, \quad i = \overline{1, M}, \qquad (4)$$

where z is input template; d is output template; M is specified number of templates.

Hence, according to (1-4), vector of input signals of the system is defined as:

$$z(k) = \begin{pmatrix} y(k), y(k-1), ..., y(k-l_1), u(k), \\ u(k-1), ..., u(k-l_2) \end{pmatrix},$$

and the resulting output vector $d = y(k+1)$.

Best regressor model must provide an acceptable training rate, generalizing properties, and on the other hand requires proper prediction accuracy.

To assess the adequacy of the resulting neural network model it is reasonable to use expression:

$$\|\hat{y} - y\| \le \varepsilon,$$

where $\|\cdot\|$ is determined norm; ε is measure of approximation inaccuracy.

Traditional criterion here is average quadratic error of prediction (quadratic norm) of the form:

$$MSE = \frac{1}{2M}\sum_{t=1}^{M}\left(y(k) - \hat{y}(k)\right)^2 = \frac{1}{2M}\sum_{t=1}^{M}\varepsilon^2(k).$$

To assess the depth of regression, method of Lipschitz coefficients is used, which described in [12]. Evaluation of delayed input and output signals quantity is made on the basis of such criteria:

$$\overline{q}^{(l)} = \left(\prod_{k=1}^{p}\sqrt{l}\,q^{(l)}(k)\right)^{1/p}, \qquad (5)$$

where $\overline{q}^{(l)}$, $q^{(l)}(k)$ are respectively current average Lipschitz coefficient value under delay depth in l-signals (considered that $l_1 = l_2 = l$); p is the quantity of largest selected coefficients out of all possible "input-output" type combinations of pairs ($p = 0{,}01N \div 0{,}02N$).

Current Lipschitz coefficients values are calculated according to the formula:

$$q_{ij}^{(l)} = \left|\frac{y(k_i) - y(k_j)}{u(k_i) - u(k_j)}\right|, \quad i \ne j,$$

where $|\bullet|$ denotes the Euclidean norm, and the values of the regression input and output vector are determined based on equations (2) and (3).

We perform simulations of TP identification for the magnetite quartzite beneficiation, using covariates of NNARX (2) and NNOE (3) types and perform selection of delay signals optimal quantity by the criterion (5).

According to [1] during the progress of magnetite quartzite beneficiation process, under conditions of modern ore concentration plants it is necessary to control next set of factors:

$$Z^P = \begin{Bmatrix} \overline{\alpha}, \overline{\xi}, \overline{g}, d_0, Q_0, \overline{Q}, \overline{d}, \overline{P}_m, \overline{\rho}_\kappa, \overline{\rho}_c, \\ \overline{\beta}_{nn}, \overline{\beta}_{x\theta}, \beta_\kappa, \overline{\gamma}, \gamma_\kappa, \overline{\varepsilon}, \varepsilon_\kappa \end{Bmatrix}, \qquad (6)$$

where $\overline{\alpha} = \{\alpha_i\}$, $i = 1...N_r$ is useful component content value in raw ore (in magnetite quartzite batch); N_r is quantity of industrial species in the batch; $\overline{\xi} = \{\xi_i\}$ is density of the respective species; $\overline{g} = \{g_i\}$ is a proportion (ratio) of species in the batch; d_0 is averaged ore coarseness before beneficiation process; Q_0 is ore feed at the first stage of beneficiation process; $\overline{Q} = \{Q_j\}$, $j = 1...N_s$ is processing at each stage of beneficiation process; N_S is quantity of beneficiation process stages; $\overline{d} = \{d_j\}$ is averaged ore coarseness of a product after each stage of beneficiation process; $\overline{P}_m = \{P_{mj}\}$ is content of solid material during degradation process; $\overline{\rho}_\kappa = \{\rho_{\kappa j}\}$ is a density of pulp before classification process; $\overline{\rho}_c = \{\rho_{cj}\}$ is a density of pulp before magnetic separation process; $\overline{\beta}_{pp} = \{\beta_{pp\,j}\}$ is content of useful component in industrial product after each stage; $\overline{\beta}_x = \{\beta_{xj}\}$ is losses of useful component in tailings; β_k is quality of concentrate; $\overline{\gamma} = \{\gamma_j\}$ is an output of useful component in final middling product; γ_K is an output of useful component in concentrate; $\overline{\varepsilon} = \{\varepsilon_j\}$ is extraction of useful component in final middling product; ε_K is extraction of useful component in concentrate.

IV. Modelling Results and Conclusions

Training of neural networks requires a statistical representative sample of TP. According to these requirements, during 2006-2016, authors collected statistics of basic technological performance indicators of all mining complexes concentration plants in the city of Kryvyi Rih, Ukraine (Tsentralnyi, Inguletskyi, Pivnichnyi, Pivdennyi, ArcelorMittal Kryvyi Rih mining complexes). The list of indicators was defined according to (6). Farther on, according to the method of passive experiment, preliminary data processing was done. The individual "emissions" (incorrect data) was removed. Then, according to the requirements of (4), the set of indicators for each plant was divided into two subsets:

1) subset for training (parameterization) models (up to 2000 templates);

2) subset of the tests (verification of adequacy) (up to 500 templates).

Subsequent modelling was based on the described methodology by using Neuro Solution package. The recurrent network model was taken as basis with the usage of direct signals and elements with memory distribution technology (Figure 1). For parameterization of models, Levenberg-Marquardt training method (LM - method) was used.

Table 1 and Figure 2 show main results of simulation the identification of magnetite quartzite TP beneficiation process and calculation of Lipschitz coefficients for different models.

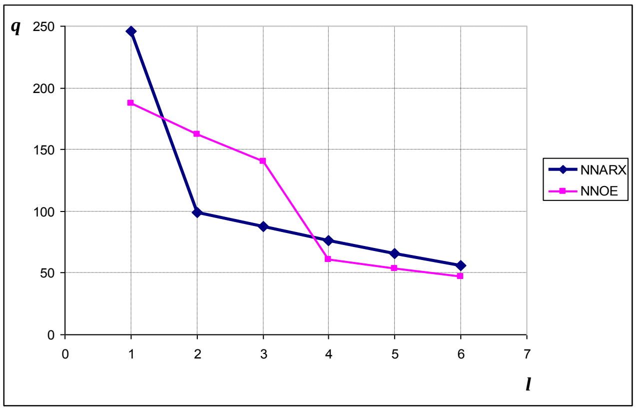

Figure 2. Dependence of the Lipschitz coefficients values(*q*) on the depth of regression (*l*) for NNARX, NNOE types of models

TABLE I. QUALITY AND CONVERGENCE INDICATORS OF MATHEMATICAL MODELS

№	Indicators and types of models		
	Indicators name	***NNARX***	***NNOE***
1	MSE (maximal)	2,62	2,58
2	Training speed (in epochs)	588	893
3	Training method	LM	LM

V. CONCLUSIONS

On the basis of modelling results, conclusions were drawn.

1. For identification of TP of iron ore magnetic beneficiation process, usage of NNARX and NNOE type of neural network predictor models is sufficiently effective.

2. In the process of identification with usage of NNARX type of model sufficient depth of delay is 2-3 signals (based on point of fracture position).

3. Usage of NNOE type of model requires a marginally greater depth of regression (3-4 signals).

REFERENCES

[1] A.Kupin Neural identification of techno-logical process of iron ore beneficiation // Proceedings of 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems Technology and Applications (IDAACS'2007).– *Dortmund, Germany*.–2007.– pp.225–227. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] A.Kupin, Y.Kumchenko Usage of training methods to parameterization of multilayer neural computing structures for technological processes // Radioelectronic and Computer systems. - *Vol. 5(69).- Kharkiv «KhAI»*, 2014.- pp.100-104.

[3] R.Jozefowicz An Empirical Exploration of Recurrent Network Architectures / R. Jozefowicz, W. Zaremba, I. Sutskever // *Proc. of the 32nd International Conference on Machine Learning "ICML-2015", (Lille, France, 6–11 July 2015). – Lille, France, 2015*. – pp. 2342 – 2350.

[4] A.Kupin Research of properties of conditionality of task to optimization of processes of concentrating technology is on the basis of application of neural networks // *Metallurgical and Mining Industry, 2014, No4*, pp. 51-55.

[5] A.Kupin Application of neurocontrol principles and classification optimization in conditions of sophisticated technological processes of beneficiation complexes // *Metallurgical and Mining Industry, 2014, No6*, pp. 16-24.

[6] A.Kupin, A.Senko Principles of intellectual control and classification optimization in conditions of technological processes of beneficiation complexes // *Proceedings of the 11th International Conference on ICT in Education, Research and Industrial Applications: Integration, Harmonization and Knowledge Transfer (ICTERI-2015), Lviv, Ukraine, May 14-16, 2015*, pp.153 160.

[7] Y.Bengio Advances in optimizing recurrent networks / Y. Bengio, N. Boulanger-Lewandowski, R. Pascanu *// Proc. of 2013 IEEE International Conference on Acoustics, Speech and Signal Proc., ICASSP-2013, (Vancouver, Canada, 26–31 May 2013).* – Vancouver, Canada, 2013. – P. 8624 – 8628.

[8] A.Kupin, B.Mysko The analysis of possibilities of modern neural network simulator software for realization of local intellectual regulators // *Computer Science, Information Technology, Automation, 2016, No.2*, pp. 36-40.

[9] Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups / G. Hinton, L. Deng, D. Yu [et al.] // *IEEE Signal Processing Magazine. – 2012. – Vol. 29.* – pp. 82 – 97.

[10] Chernodub A. Training Dynamic Neural Networks Using the Extended Kalman Filter for Multi-Step-Ahead Predictions / A. Chernodub // Artificial Neural Networks Methods and Applications in Bio-Neuroinformatics / P. Koprinkova-Hristova, V. Mladenov, N.K. Kasabov (eds.). – *Springer International Publishing Switzerland, 2015.– Vol. 4*. – pp. 221 – 244.

[11] Long-term Recurrent Convolutional Networks for Visual Recognition and Description / J. Donahue, L. A. Hendricks, S. Guadarrama [et al.] // *Proc. of 2015 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, (Boston, USA, 7–12 June 2015)*. – Boston, USA, 2015. – pp. 2625 – 2634.

[12] On the Stochastic Maximum Principle in Optimal Control of Degenerate Diffusions with Lipschitz Coefficients / K.Bahlali, B.Djehiche, B.Mezerdi // *APPLIED MATHEMATICS AND OPTIMIZATION, 2007.- Vol. 56 Issue: 3*.- pp. 364-378.