

Janusz Rusek*

Computational Intelligence Methods in the Problem of Modelling Technical Wear of Buildings in Mining Areas**

1. Introduction

As it may be concluded from the conducted research [e.g. 5, 7, 10, 25, 26] the technical wear of buildings, in addition to determining the technical condition, can form the basis for the valuation of mining damage. In the subject literature numerous formulas may be found which allow us to estimate the value of technical wear. These formulas have been presented synthetically in [5]. The most accurate method of determining the value of the technical wear of buildings is the so-called weighted average method:

$$s_z = \sum_{i=1}^n \frac{u_i \cdot s_{ei}}{100} \quad [\%] \quad (1)$$

where:

- u_i – percentage share of an individual element reconstruction cost in the reconstruction cost structure of the object, acting as a weight,
- s_{ei} – degree of wear of an individual element [%],
- n – number of the evaluated elements in the object.

Unfortunately, determining the value of technical wear using weighted average method is symptomatic and disqualifies this method in forecasting of problems, since the evaluation of wear of individual elements is carried out *post factum*, which is e.g. mining shock or subsidence of the ground surface as a result of mining activities. In addition, prognostic or approximating issues defined in this paper, first of all, should be described in relation to the parameters characterizing the effects of underground mining activities, visible on the surface of the mining area and constituting a direct threat to buildings (e.g. terrain deformation indexes or shock characteristics), secondly, it should be generalized within a homogeneous group of buildings, which are buildings of the mining area potentially endangered with the occurrence of mining damage.

* AGH University of Science and Technology, Faculty of Mining Surveying and Environmental Engineering, Krakow, Poland

** The paper was prepared within the scope of the AGH statutory research no. 11.11.150.005

The research carried out at the Department of Surveying Engineering and Construction at AGH University of Science and Technology proved the existence of a statistically justified relationships between the degree of technical wear of buildings and the parameters characterizing the potential influence of mining, occurring on the surface of the mining area. Thus, the decision to apply methods allowing us to create a model of the process of technical wear of buildings was also justified, whose domain will be the potential causes (deformation indexes or mining shock characteristics) and not separately inventoried consequences for individual elements of a building, as it is in the weighted average method. Such a model, which is the approximator of the wear degree value, was to forecast the possible deterioration of the utility value of a given building, caused by underground mining activity, which a deterioration could be identified with a range of mining damage.

The relationships between technical wear and impacts of mining activity obtained by earlier studies [e.g. 5, 7, 27], were represented in relation to isolated mining factors (one-parameter regression models) and were of a linear character. However, the extension of analyses by one-parameter non-linear regression models and the results obtained there allowed us to conclude the existence of non-linear relationships between mining factors and the course of technical wear of buildings [5].

Such an approach to the problem as well as the conclusions drawn from the conducted analyses allowed to establish criteria for the selection of methods to create a model of the course of technical wear of buildings for the purpose of forecasting in the context of the deterioration of the utility value of buildings caused by us underground mining activity.

2. Criteria for the Selection of the Methodology

The main criteria for the selection of a method which would allow to solve that problem have been presented here. The first approach focuses on the one-way methods of artificial neural networks.

The choice of this family of methods was based on such properties as [e.g. 4, 15, 20]:

- the possibility of presenting the course of technical wear in the multi-dimensional domain of explanatory variables,
- the possibility of implementing a non-linear mapping,
- the lack of a necessity to select a starting form of mapping,
- the lack of guidelines in relation to the quality of the collected model data (requirements of the classical approximation methods in statistics),
- the possibility to generalize the knowledge acquired during the learning process.

These features prove the universality of neural networks in the context of solving non-linear multidimensional problems [15].

Taking into account that the family of neural networks is very wide, a clarifying step was taken in order to determine which method is most effective in the context of the problem being solved.

The leading criteria in that part of the research were:

- the degree of generalization of the acquired knowledge,
- the clarity of the learning process,
- the possibility of conducting a sensitivity analysis with respect to the created model.

The scope of the search was extended by the SVM method (Support Vector Machine [e.g. 2, 21, 24]) related to neural networks, which consequently were selected as the most effective in the context of the assumed criteria.

3. Comparison of Neural Networks and the SVM Method

The regression approach of the method Support Vector Machine (SVM), also called ε -SVR [e.g. 2, 21, 24], in the context of the structure and operation of the given system is becoming very similar to the neural networks with radial kernel functions (RBF Radial Basis Function Neural Network) [e.g. 13, 15]. The difference is revealed both in the approach associated with the learning of such systems and the determination of their structure. The attached diagrams (Figs 1, 2), in a simplified manner, present the operation of neural networks for the approximation case.

In typical neural networks, the learning process, which is tuning of the synaptic weights between neurons, is an iterative minimization of a certain measure of fit, which is usually the mean square error (MSE).

In order to increase the degree of generalization of these systems, a wide range of actions are taken, which come down either to redefining the function of error, leading to a reduction in the value of the particular weights of the system, or they impose certain conditions on the complexity of their structure [9, 15, 23]. In general, actions modifying the objective functions introduce a compromise between the quality of matching and generalization of the system. They are based on the development of the error function with the so-called regularization factors [15]. These factors can be identified with the so-called penalty functions, which occur in the optimization issues [11].

On the other hand, the methods to reduce the complexity of the neural networks structure are based on sensitivity analysis of individual connections represented by weight values to the final value of the defined measures of fit [4, 9].

Most often, however, determining the structure complexity of the artificial neural networks is assumed arbitrarily and the optimal selection of the number of neurons in the hidden layers requires a number of independent simulations.

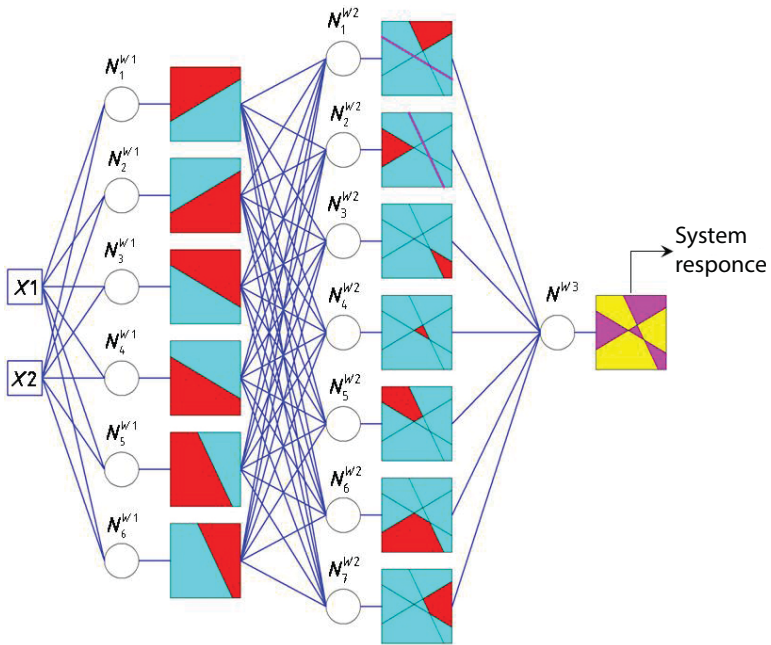


Fig. 1. Operation diagram of a non-linear neural network for the approximation case

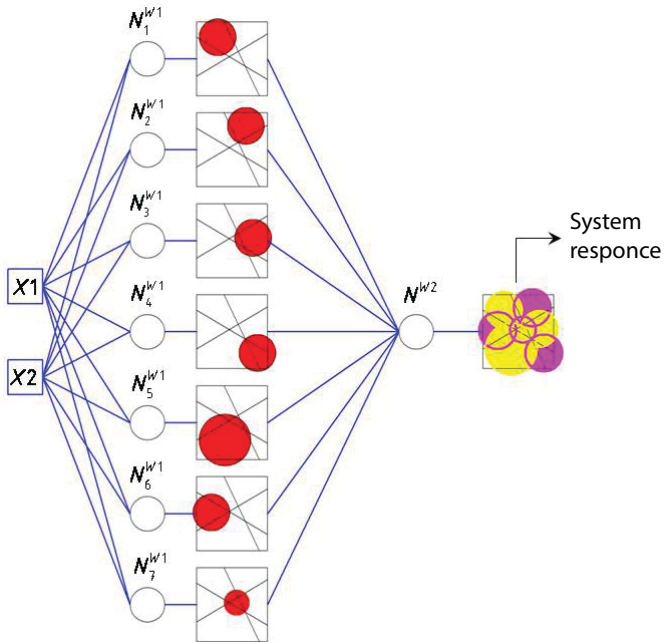


Fig. 2. Operation diagram of a radial neural network for the approximation case

It should also be noted that the form of the objective function expressed in the weight domain of the system is multi-modal. In order to minimize this task, choosing a starting point in the optimization process, a problem arises which is related to the convergence of solutions to local minima, thus the oscillation (the stall of an iteration minimizing process) in the area of these minima. In order to eliminate such a situation, a certain inertia is included in the process of tuning weight values of the system by introducing the *momentum* factor [4, 23]. The introduction of such a factor protects the optimization process both from the stall in the local minimum area, as well as it protects from leaving the global solution area.

The qualified for the study SVM method differs in several issues from the approach associated with the structure of typical neural networks.

The first difference is the spontaneous development of the complexity of the system, which in the case of radial kernels depends upon three parameters designating the appropriate characteristics of the kernel functions (parameter – σ), the degree of compromise between fitting and generalization of the system (parameter – C) and the width of the tolerance band (parameter – ε) [e.g. 24].

These constants are selected arbitrarily but there are methods of the optimal selection of their values [e.g. 1, 3, 9, 12]. The possibility of selecting the optimal parameters C , ε , and is the second major advantage characteristic of the SVM method.

Another advantage of the support vectors methods is getting rid of a multimodal form of objective function hypersurface [13, 15]. This feature eliminates the possibility of oscillation around local minima, which is the case with conventional neural networks. Thus, the high sensitivity of these systems to the selection of the starting point in the optimization process is eliminated. In relation to the SVM method, the surface of the objective function is one-modal, which is a result of bringing the learning procedure to a quadratic programming task [8, 11]. There are even formulations of the SVM method, which can bring the error surface to the surface of the first degree (Least Square Support Vector Machine – LSSVM method [22]). The final advantage, characteristic of the support vectors methods is their low sensitivity to the number of learning patterns. These methods are, therefore, less susceptible to the phenomenon of the so-called “curse of dimensionality”, which limits the operation of typical neural networks [15]. In other words, the number of learning patterns needed to map the modelled phenomenon, with an appropriate extension of the margin of separation, becomes independent of the complexity of the system [13, 15].

Summarizing, both the artificial neural networks and support vectors methods are universal approximation methods of functions of several variables. In numerous cases provided in the literature, the accuracy of mapping for the two groups of methods is comparable. However, the very process of building such systems is characterized by a greater clarity eliminating external interference and control of the designer.

4. Description of the SVM Method in ε -SVR Regression Terms

In the regression approach in the SVM method, the record of the original sought function approximating model data takes the following form [15]:

$$y(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}) + b \quad (2)$$

where:

$\mathbf{x} \in R^n$ – a vector of input data in n-dimensional space,
 $\varphi(\cdot): R^n \rightarrow R^{n_h}$ – a certain transformation converting raw input data into the so-called space of system characteristics.

The mapping $\varphi(\cdot): R^n \rightarrow R^{n_h}$ is given implicitly, and it is the result of the application of the kernel function of a specific type [2, 24].

In the final record, apart from the structural details of the method adopted in his way, which were accurately described, inter alia, in [2, 18, 24], the sought function takes the following form:

$$y(\mathbf{x}) = \sum_{k=1}^{N_{SV}} (\alpha_k - \alpha_k^*) K(\mathbf{x}, \mathbf{x}_k) + b \quad (3)$$

where:

$K(\mathbf{x}_k, \mathbf{x}_j)$ – the kernel of the system that is inflicted in an explicit manner and is a result of the submission of implicit functions,
 $\alpha_k - \alpha_k^*$ – Lagrange multipliers [11].

5. The Applied Method of Optimal Scaling of ε -SVR Model

The main problem in the ε -SVR approach was to determine the optimal values of the parameters C , ε and σ . The necessity of their determination results directly from the adopted objective function, whose minimization is a fundamental stage in tuning the structure of the system [e.g. 15, 18]. Additional parameter σ is the result of the adoption of the radial kernel functions for the built approximator and, according to the relation (3), it defines its width [e.g. 13]:

$$K(\mathbf{x}_k, \mathbf{x}_j) = \exp\left(-\frac{(\mathbf{x}_k - \mathbf{x}_j)^2}{\gamma^2}\right)^{\sigma = \frac{1}{\gamma^2}} = \exp(-\sigma(\mathbf{x}_k - \mathbf{x}_j)^2) \quad (4)$$

An attempt to determine the sought parameters, hereinafter referred to as hyperparameters, was discussed in [18], in which, based on the concept of Meta-SVM [9], the error FPE (Final Prediction Error) was additionally used [22]. In the further course of the study it was decided, however, to apply a more efficient, according to

the author, method based on the concept described in [3]. The primary stage of this method is the n -fold oblique validation conducted on pre-sets: training and testing. For each iteration of validation, a range of considered parameters C , ε and σ expressed on a logarithmic scale is assumed. Then, according to the proposed optimization algorithm grid search [1, 3], the minimization the objective function is carried out, adopted as an MSE error averaged of all n test sets used in the validation. As a result, the optimum to the adopted strategy, set of sought hyperparameters C , ε and σ is achieved. Algorithm grid search is a no-gradient method of global minimization [16]. It has, however, a disadvantage of a necessity to provide ranges for the searched area as well as a specific starting point. Therefore, in the paper, instead of the algorithm grid search, an optimization method was used based on the genetic algorithm GA. The applied method is also a no-gradient algorithm allowing us to identify the global minimum. The description of the operation of genetic algorithms in the problems of minimizing a function of several variables can be found for example in [14, 16, 19, 20].

6. Summary and Conclusions

The idea behind the methodology adopted in this paper was the creation of a course model of technical wear of buildings in mining areas. In addition, a model created through sensitivity analysis allowed us to identify both the quantitative and qualitative contribution, which the mining and general construction factors affect the course with. In addition, having such a model as well as the results on the impact of different variables on the course of the modelled process, it was possible to transfer the obtained information to the formalism of fuzzy logic and fuzzy reasoning. Such a transfer allowed for the creation of a rule-based fuzzy inference system, which in addition to the approximation of the degree of technical wear of buildings, serves as a classifier of the technical condition given in the form of inaccurate linguistic expressions [17]. The obtained results may also support the prediction of the degree of damage proposed in [6].

In further research plans it is justified to use probabilistic conditional inference formalisms and to implement Bayesian networks or a specific type of neural networks PNN (Probabilistic Neural Networks). They will, in the author's intention, combine formal uncertainty with the probability of occurrence of the variables that affect the course of the technical wear of buildings. However, the used in the work, sensitivity analysis related to the SVM method will be transferred to the structures of the so-created models. In addition, the use of a probabilistic approach offered by Bayesian networks will allow us to link the existing research with the domain of structural reliability, in which there has long functioned techniques based on well-defined probability distributions.

References

- [1] Bi J., Bennet K., Enbrechts M.J., Breneman C.M., Song M.: *Dimensionality Reduction via Sparse Support Vector Machines*. 2003.
- [2] Burges C.J.C.: *A Tutorial on Support Vector Machines for Pattern Recognition*. *Data Mining and Knowledge Discovery*, vol. 2, 1998, pp. 121–167.
- [3] Chang C-C., Lin C-J.: *LIBSVM: A Library for Support Vector Machine*. Software available at: <http://www.csie.nyu.edu.tw/~cjlin/libsvm>, 2008.
- [4] Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R.: *Sieci neuronowe*. [in:] Nałęcz M. (red.), *Biocybernetyka i inżynieria biomedyczna*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2000.
- [5] Firek K.: *Badanie wpływu czynników górniczych i budowlanych na zużycie techniczne tradycyjnej zabudowy terenu górniczego LGOM*. Akademia Górniczo-Hutnicza, Kraków 2005 (Ph.D. Thesis, unpublished).
- [6] Firek K.: *Proposal for classification of prefabricated panel building damage intensity rate in mining areas*. *Archives of Mining Sciences*, vol. 54, no. 3, 2009, pp. 467–479.
- [7] Firek K., Wodyński A.: *Assessment of surface deformation impacts on technical wear of masonry buildings located in the Legnica-Głogów Copper District*. *Schriftenreihe des Institutes für Markscheidewesen und Geodäsie an der Technischen Universität Bergakademie Freiberg*, Heft 2006-1, 8 Geokinematischer Tag, Freiberg 2007.
- [8] Górecki H.: *Optymalizacja i sterowanie systemów dynamicznych*. Uczelniane Wydawnictwa Naukowo-Dydaktyczne AGH, Kraków 2006.
- [9] Jankowski N., Grabczewski K.: *Toward optimal SVM*. The third IASTED International Conference on Artificial Intelligence and Applications, Anaheim, Calgary, Zurich. The International Association of Science and Technology for Development, ACTA Press, 2003.
- [10] Kocot W.: *Diagnozowanie stanu technicznego budynków o tradycyjnej konstrukcji na terenach górniczych*. Akademia Górniczo-Hutnicza, Kraków 1998 (Ph.D. Thesis, unpublished).
- [11] Kusiak J., Danielewska-Tulecka A., Oprocha P.: *Optymalizacja*. Wydawnictwo Naukowe PWN, Warszawa 2009.
- [12] Lahiri S.K., Ghanta K.C.: *The Support Vector regression with the parameter tuning assisted by a differential evolution technique: study of the critical velocity of a slurry flow in a pipeline*. *Chemical Industry & Chemical Engineering Quarterly*, West Bengal, India, 2008.
- [13] Łęski J.: *Systemy neuronowo-rozmyte*. WNT, Warszawa 2008.
- [14] Matlab: *Genetic Algorithm and Direct Search Toolbox*. 1984–2007: The MathWorks, Inc.

-
- [15] Osowski S.: *Sieci neuronowe do przetwarzania informacji*. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2006.
- [16] Ostanin A.: *Informatyka z matlabem*. Wydawnictwo Politechniki Białostockiej, Białystok 2007.
- [17] Rusek J.: *Modelowanie stopnia zużycia technicznego budynków na terenach górniczych z wykorzystaniem wybranych metod sztucznej inteligencji*. Akademia Górniczo-Hutnicza, Kraków 2010 (Ph.D. Thesis, unpublished).
- [18] Rusek J.: *Creating a model of technical wear of buildings in mining areas with utilization of regressive SVM approach*. Archives of Mining Sciences, vol. 54, no. 3, 2009, pp. 455–466.
- [19] Rutkowska D., Piliński M., Rutkowski L.: *Sieci neuronowe algorytmy genetyczne i systemy rozmyte*. Wydawnictwo Naukowe PWN, Warszawa 1999.
- [20] Rutkowski L.: *Metody i techniki sztucznej inteligencji*. Wydawnictwo Naukowe PWN, Warszawa 2005.
- [21] Schölkopf B., Smola A.J.: *Learning with Kernels*. MIT Press, Cambridge, Massachusetts, 2002.
- [22] Suykens J.A.K., Van Gestem T., De Brabanter J., De Moor B., Vandewalle J.: *Least squares support vector machines*. World Scientific, Singapore, 2002.
- [23] Tadeusiewicz R.: *Wprowadzenie do sieci neuronowych*. StatSoft Polska, 2001.
- [24] Vapnik V.: *Statistical learning theory*. John Wiley & Sons, New York 1998.
- [25] Wodyński A., Barycz S.: *Analiza porównawcza stanu technicznego i zakresu uszkodzeń budynków na terenie górniczym LGOM*. Materiały Szkoły Eksploatacji Podziemnej 2002, seria: Sympozja i Konferencje, nr 54, PAN/AGH, Kraków–Szczyrk 2002.
- [26] Wodyński A.: *Zużycie techniczne budynków na terenach górniczych*. Uczelniane Wydawnictwa Naukowo-Dydaktyczne AGH, Kraków 2007.
- [27] Wodyński A., Kocot W.: *Analysis of technical wear of traditionally structured buildings on mining areas*. Proceedings of 11th International Congress of the International Society for Mine Surveying, Kraków 2000.