

Janusz RUSEK

## APPLICATION OF SUPPORT VECTOR MACHINE IN THE ANALYSIS OF THE TECHNICAL STATE OF DEVELOPMENT IN THE LGOM MINING AREA

### ZASTOSOWANIE METODY SUPPORT VECTOR MACHINE W ANALIZIE STANU TECHNICZNEGO ZABUDOWY TERENU GÓRNICZEGO LGOM\*

*The paper presents the results of the analysis of technical wear of buildings located within impact of mining plant in the Legnica - Głogów Copper District (LGOM). The study used method related to neural networks, support vector (Support Vector Machine) in regression approach  $\varepsilon$ -SVR (Support Vector Regression). The aim of the study was to assess the impact of variables describing the structural protection and renovations on the course modeled phenomenon. The basis for the analysis was created model of technical wear of buildings in the form of a network  $\varepsilon$ -SVR. In addition to the variables determining the level of structural protection and renovations in the model included variables describing: terrain deformation, mining intensity tremors and the age of the buildings. The choice of model parameters were performed using, as gradientlessness optimization method, genetic algorithm. Based on the established model  $\varepsilon$ -SVR two types of sensitivity analysis were applied. Assessing the impact of the structural protections have been studying by the analysis of variability of the gradient vector for the modeled hypersurface. The analysis of the impact of renovations on the course modeled process was carried out based on the comparator simulation results of  $\varepsilon$ -SVR model. The results confirmed the usefulness of the methodology of research and allowed to draw important conclusions on the impact of analyzed factors on the technical wear traditional buildings LGOM.*

**Keywords:** Support Vector Machine, influence of mining, structure resistance, technical wear, technical condition.

*W pracy przedstawiono wyniki analizy zużycia technicznego budynków zlokalizowanych w zasięgu wpływów eksploatacji górniczej na terenie Legnicko-Głogowskiego Okręgu Miedziowego (LGOM). W badaniach zastosowano pokrewną sieciom neuronowym metodę wektorów podpierających (Support Vector Machine) w podejściu regresyjnym  $\varepsilon$ -SVR (Support Vector Regression). Celem badań było uzyskanie oceny wpływu zmiennych opisujących zabezpieczenia konstrukcyjne i remonty na przebieg modelowanego zjawiska. Podstawą do analiz był utworzony model zużycia technicznego budynków w postaci sieci  $\varepsilon$ -SVR. Oprócz zmiennych określających poziom zabezpieczeń konstrukcyjnych i remontów, w modelu uwzględniono zmienne opisujące: deformacje terenu pochodzenia górniczego, intensywność wstrząsów oraz wiek budynków. Dobór parametrów modelu przeprowadzono z wykorzystaniem, jako bezgradientowej metody optymalizacyjnej, algorytmu genetycznego. Bazując na utworzonym modelu  $\varepsilon$ -SVR przeprowadzono dwurodzajową analizę wrażliwości. Oceny wpływu zabezpieczeń konstrukcyjnych dokonano badając zmienność wektora gradientu modelowanej hiperpowierzchni. Natomiast analiza wpływu remontów na przebieg modelowanego procesu została przeprowadzona na bazie komparacji wyników symulacji modelu  $\varepsilon$ -SVR. Wyniki badań potwierdziły przydatność przyjętej metodyki badań oraz pozwoliły na sformułowanie istotnych wniosków dotyczących wpływu analizowanych czynników na zużycie techniczne tradycyjnej zabudowy LGOM.*

**Słowa kluczowe:** Support Vector Machine, wpływy górnicze, odporność budynków, zużycie techniczne, stan techniczny.

#### 1. Introduction

Resistance of building structures to the impacts of mining depends on their technical condition, defined in terms of technical wear. In recent years, the results of the research studies [6, 7] confirmed the significant influence of mining exploitation on the technical wear of building structures, both in the form of ground deformations and mining tremors. They also demonstrated the importance of structural preventive measures and of current repairs. These relationships, however, remained implicit in the context of mathematical functional form. Therefore, the use of more complex analytical methods, taking into account the multidimensionality and non-linearity of the modeled process, was justified in the further studies. Such a model should allow for a more effective assessment of the technical condition of building structures and for the analysis of the influence of individual variables included in the description of the modeled phenomenon. Taking this

into consideration, a model of the course of technical wear as the SVM (Support Vector Machine) network in the  $\varepsilon$ -SVR regression approach (Support Vector Regression) was used in the studies which were presented in the article.

The SVM method [8, 25], as well as its regression approach  $\varepsilon$ -SVR [4, 21], comprise a subgroup of the methods belonging to Machine Learning. The structure of these systems is very similar to Artificial Neural Networks (ANN) [17]. The choice of the SVM method to carry out the research studies presented in this paper was dictated the fact that:

- in the  $\varepsilon$ -SVR method, it is not necessary to predetermine the mapping function, which allows to create a model for the multidimensional process in which the relationship between the variables are non-linear [6, 7] but without explicit strictly functional form,

(\*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie [www.ein.org.pl](http://www.ein.org.pl)

- in contrast to the artificial neural networks, the procedure of building the  $\varepsilon$ -SVR network allows for the optimal, in the generalizing sense, expansion of its structure, which takes place during the calibration process of the parameters  $C$ ,  $\varepsilon$  and  $\gamma$ . These parameters, described in Chapter 3 of this article, are contained in the raw formulation of the objective function for the  $\varepsilon$ -SVR method [21.25]
- the final mathematical form of the  $\varepsilon$ -SVR network allows for a relatively simple determination of the values of the partial derivatives, which form the basis for the sensitivity analysis of the model with respect to the specified continuous input variables.

The model of technical wear presented in this paper was created by the optimal choice of the parameters  $C$ ,  $\varepsilon$  and  $\gamma$  using the LIBSVM package [6] and the genetic algorithm GA [19]. In order to identify the effect of the variable describing the level of structural preventive measures, ( $w_{ZAB}$ ), the sensitivity analysis of the model was carried out, involving the study of the course of the gradient component for such a defined index. On the other hand, in order to determine the influence of the categorical variable  $w_{REM}$  describing the extent of the repair works, the values of the model prediction obtained by means of its simulation were compared.

## 2. Description of the database of building structures

A group of 930 masonry residential buildings located within the impacts of the Mining Plant of KGHM "Polska Miedz" S.A., comprised the database for the analysis. The database of buildings was created as part of the examination of their technical condition. At the stage of detailed inspection, each building was cataloged using 93 variables. These variables describe the location of the structure, geometric data, type of development, the data on the structural components and finishing elements, the age of the buildings, technical wear determined by the method of weighted average [23], the scope of the repair works, the level of preventive measures against the impacts of mining and the current, at the time of the inspection, mining impacts in the form of deformation indices and characteristics of mining tremors.

This development has been subjected to mining impacts for more than 40 years. Initially, they resulted from the formation of the subsidence trough over the mining excavations and a large trough associated with rock mass drainage [23]. In the mining area of LGOM, also mining tremors have been occurring for 30 years [24].

The building structures located in LGOM, and erected after 1970, are protected against the effects of ground deformation caused by mining activities already during the construction stage. Older buildings are also regularly subjected to preventive measures (anchoring, reinforced concrete ties, etc.). Since the late 80's of the twentieth century, structural protection against parasismic effects has also been introduced in newly constructed objects.

In order to assess the degree of wear of the studied buildings, the so-called method of weighted average was used [e.g. 23]. It involves an individual assessment of the degree of wear of individual elements, and then – by assigning appropriate weights to them – determining the weighted average degree of technical wear of the whole building structure.

To specify and describe quantitatively the indices affecting the technical condition of the development, two additional indices  $w_{REM}$  and  $w_{ZAB}$  were defined.

Protection index  $w_{ZAB}$  was defined using the category of structure resistance ( $KO$ ), which is the structural characteristics of the building, and the category of hazards to area from mining impacts ( $KT$ ).

Mining area categories ( $KT$ ) describe the intensity of continuous surface deformation, expressed by assigning the characteristic values of the slopes ( $T$ ), curves ( $R$ ), and horizontal deformation ( $\varepsilon$ ) to the specific ranges of these indices (from 0 to V). On the other hand, the category of structure resistance ( $KO$ ) is understood as the resistance

of a building structure to the horizontal strains and curvatures of the land, adapted to the ranges of the values of these indices in the mining area categories ( $KT$ ) – (from 0 to 4) [e.g. 9]. A building structure is considered to be resistant to the effects of mining when the category of its resistance  $KO$  is not smaller than the mining area category  $KT$ . In practice, it is assumed that the implementation of protective measures in a building structure aims at increasing the category of structure resistance ( $KO$ ), at least to the level of the hazard category occurring in a given area ( $KT$ ).

Depending on the scope of these protective measures, the total number of points ( $p$ ) was established for each building structure according to the empirical point method for assessing the resistance of building structures [9], in the range from  $p=0$  (a building protected at the level of foundations and all floor structures) to  $p=15$  (no protection). As a result, for each building structure, the value of the index  $w_{ZAB}$  was determined in accordance with the dependence (1):

$$w_{ZAB} = \begin{cases} \frac{(15-p)}{15} \cdot (|KT - KO|) + KO & \text{dla } KO < KT \\ \frac{(15-p)}{15} + KO & \text{dla } KO \geq KT \end{cases} \quad (1)$$

where:  $KT$  – mining area category, for which additional protective measures were implemented,  
 $KO$  – category of structure resistance, predetermined during the design stage,  
 $p$  – number of points from the point method for evaluating the resistance of building structures for a given scope of protection.

On the other hand, the index  $w_{REM}$  reflects the extent of repair works for each building. This is a categorical variable, described by the following values: 1 - for building structures with no repairs implemented, and 2 - for building structures undergoing comprehensive repairs.

## 3. Research Methodology

This part of the study presents:

- mathematical interlude in the scope of basic formulations of the SVM method in regression approach,
- method of calibration of the parameters  $C$ ,  $\varepsilon$  and  $\gamma$  determining the final form of the SVM network structure, which was used in the study,
- procedure of determining the gradient components in the sensitivity analysis.

### 3.1. Description of the SVM method in regression approach

In the regression approach in the SVM method, the sought approximating function is as follows [16]:

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

where:  $\mathbf{x} \in R^n$  – is the input variable vector in  $n$ -dimensional space,

$\phi(\cdot): R^n \rightarrow R^{n_h}$  – is a certain transformation transforming the raw input variables into the so-called feature space,

$\mathbf{w}^T$  – is the vector of weights,

$b$  – free component (bias)

Mapping  $\phi(\cdot): R^n \rightarrow R^{n_h}$  is implicit, and it is the result of the use of the kernel function of a specific type (e.g. [25]).

The basis for the formulation of the problem of learning of the SVM system in regression approach is the definition of the error function, minimized in the adaptation process. With a certain set of reference data  $\{\mathbf{x}_k, y_k\}_{k=1}^N$ , the loss function for a single reference is expressed as (eg. [25]):

$$L_\varepsilon(y_k, y(\mathbf{x}_k)) = \begin{cases} |y_k - y(\mathbf{x}_k)| - \varepsilon & \text{dla } |y_k - y(\mathbf{x}_k)| \geq \varepsilon \\ 0 & \text{dla } |y_k - y(\mathbf{x}_k)| \leq \varepsilon \end{cases} \quad (3)$$

This formula is called  $\varepsilon$ -insensitive loss function [10]. As a result of using the above function, imposing a certain tolerance margin  $\varepsilon$ , the regression approach using SVM is usually called  $\varepsilon$ -SVR ( $\varepsilon$ -Support Vector Regression (e.g. [21])).

The introduction of such an error function is typical of the so-called robust methods resistant to interference in the data, or the outliers [14]. It also allows for the subsequent formulation of the problem of learning as the task of quadratic programming [2], which is one of the main advantages of the  $\varepsilon$ -SVR method.

With a predetermined error function, for a single reference (3), the objective function is written, the minimization of which will be followed by the learning process:

$$\min_{\mathbf{w}, b, \xi_k^*, \zeta_k} J(\mathbf{w}, \xi_k^*, \zeta_k) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{k=1}^N (\xi_k^* + \zeta_k), k = 1 \dots N \quad (4)$$

With inequality constraints:

$$y_k - \mathbf{w}^T \phi(\mathbf{x}_k) - b \leq \varepsilon + \xi \quad (5)$$

$$\mathbf{w}^T \phi(\mathbf{x}_k) + b - y_k \leq \varepsilon + \zeta \quad (6)$$

$$\xi_k, \zeta_k \geq 0 \quad (7)$$

The above record is the so-called original minimization problem with inequality constraints (5), (6) and (7) (e.g. [2]). The problem presented in this way compromises between the generalization and the quality of fitting the approximator expressed as (2).

The component  $\sum_{k=1}^N (\xi_k^* + \zeta_k)$  in the equation (4), together with

the constraints (5), (6) and (7) is responsible for the minimization of the global error function for all the learning references:

$$L_\varepsilon^N = \frac{1}{N} \sum_{k=1}^N L_\varepsilon(y_k, y(\mathbf{x}_k)) \quad (8)$$

The values  $\xi_k^*, \zeta_k$  are the deviations of mapping each reference beyond the predetermined error tolerance  $\varepsilon$ . The learning process attempts to minimize them while complying with all the constraints of the function (5), (6) and (7). On the other hand, the component  $\frac{1}{2} \mathbf{w}^T \mathbf{w}$  is responsible for the so-called maximizing the margin of separation [16]. In the regression approach, implemented in the feature space, minimization of that component leads to the optimal determination of the approximator hyperplane within the predetermined error tolerance  $\varepsilon$ .

All in all, both components of the equation (4) are opposing, and a compromise is determined by introducing a regularization constant  $C$ . The higher the value, the better fitting of the system, and the lower the generalization properties, and vice versa.

To solve the so defined problem of minimization, Lagrange functions are created (e.g. [2, 14]). Thus, the dependence (4) is obtained, extended by the set of all constraints (5), (6) and (7) controlled by Lagrange multipliers:

$$L(\mathbf{w}, b, \xi_k, \zeta_k^*, \alpha_k, \alpha_k^*, \mu_k, \mu_k^*) = J(\mathbf{w}, \xi_k^*, \zeta_k) - \sum_{k=1}^N \alpha_k (\mathbf{w}^T \phi(\mathbf{x}_k) + b - y_k + \varepsilon + \xi_k) + \sum_{k=1}^N \alpha_k^* (y_k - \mathbf{w}^T \phi(\mathbf{x}_k) - b + \varepsilon + \zeta_k^*) - \sum_{k=1}^N (\mu_k \xi_k^* + \mu_k^* \zeta_k), k = 1 \dots N \quad (9)$$

Then, using Fermat's theorem and Karush-Kuhn-Tucker conditions [25], dual formulation of the quadratic programming is obtained, expressed in the field of Lagrange multipliers in the following form [14]:

$$\max_{\alpha_k, \alpha_k^*} Q(\alpha_k, \alpha_k^*) = \sum_{k=1}^N y_k (\alpha_k - \alpha_k^*) - \varepsilon \sum_{k=1}^N (\alpha_k + \alpha_k^*) - \sum_{k=1}^N \sum_{j=1}^N (\alpha_k - \alpha_k^*) (\alpha_j - \alpha_j^*) K(\mathbf{x}_k, \mathbf{x}_j) \quad (10)$$

With inequality constraints:

$$0 \leq \alpha_k^* \leq C \quad (11)$$

$$0 \leq \alpha_k \leq C \quad (12)$$

$$\sum_{k=1}^N (\alpha_k - \alpha_k^*) \geq 0 \quad (13)$$

This formulation is subject to further maximization using Lagrange multipliers.

The factor  $K(\mathbf{x}_k, \mathbf{x}_j)$  appearing in the equation (10) is the kernel of the system, which is given explicitly, and is the result of the equation of the implicit functions  $\phi(\cdot)$  [14]:

$$K(\mathbf{x}_k, \mathbf{x}_j) = \phi(\mathbf{x}_k) \phi(\mathbf{x}_j) \quad (14)$$

The form of the kernel is selected arbitrarily from all the functions meeting the assumptions of Mercer's Theorem [14].

As a result of such a learning procedure, having determined the values of Lagrange multipliers, the weight vector is determined from the relationship:

$$\mathbf{w} = \sum_{k=1}^{N_{SV}} (\alpha_k - \alpha_k^*) \phi(\mathbf{x}_k) \quad (15)$$

Which, in turn, allows to write the final form of the approximator:

$$y(\mathbf{x}) = \sum_{k=1}^{N_{SV}} (\alpha_k - \alpha_k^*) K(\mathbf{x}, \mathbf{x}_k) + b \quad (16)$$

3.2. Calibration of the parameters C, ε and γ

An attempt to determine the sought parameters, hereinafter referred to as *hyperparameters*, was made in the paper [19], in which additionally, based on the concept of *Meta-SVM* [10], *FPE (Final Prediction Error)* was used [19]. In the course of further research, however, the author decided to use a more efficient method based on the concept described in [4]. The main stage in this method is the *n-fold cross-validation* carried out on preliminary sets: training set and testing set. For each iteration of the validation, a certain range of the discussed parameters C, ε and γ is examined, which are expressed in the logarithmic scale. Then, according to the proposed optimization algorithm *GS (Grid Search*, e.g. [4]), the minimization of the objective function is carried out, which was adopted as *MSE (Mean Squared Error)*, averaged from all n testing sets used in the validation. As a result, an optimal set of the sought hyperparameters C, ε and γ is obtained. The *GS* algorithm is a gradientless global minimization method (e.g. [2]). It has one drawback, though. There is a necessity to define ranges for the searched area and a specific starting point. Therefore, instead of the *GS* algorithm, the optimization method was used, based on the *Genetic Algorithm (GA)*. The applied *GA* method is also a gradientless algorithm, allowing to identify the global minimum, e.g. [2, 15].

3.3. Sensitivity analysis - study of the course of gradient components

The created *SVM* model, as in the case of artificial neural networks (*ANN*) is the so-called “*Black box*” and is primarily used to estimate the input variables. In recent years, more and more attempts have been made to analyze the internal structure of such models [26]. These activities aim to clarify the information about the interactions between the input variables and the output variable. The primary, and at the same time the simplest procedure, allowing to determine the quantitative influence of each of the dependent variables on the explanation of the raw variability contained in the output variables, is the so-called *variable elimination*. Basing on such a procedure, it is possible to make the selection of the most important variables in the model, the so-called *feature Selection* [3, 13]. Another approach, although rarely undertaken, is the analysis of the course of the gradient components relative to each of the input variables included in the model. Such an approach, used in the research studies presented in the paper, allows for the qualitative and quantitative assessment of the influence of each variable on the monotonicity of the estimated process [5].

This method was used here to study the influence of continuous variable such as the  $w_{ZAB}$  index. In order to carry out the analysis, the analogy of the structure of the  $\epsilon$ -*SVR* model to the *Radial Basis Function Neural Networks* was used ([16]). Thus, the sensitivity analysis was performed in accordance with the procedures proposed in [5, 18, 22]. Considering the fact that the record of the  $\epsilon$ -*SVR* approximator can be expressed in the form of a deterministic function in accordance with the equation (16), for any component of the vector of input variables  $i$  at a given point  $x_p$ , it is possible to determine the gradient component in the following form:

$$\frac{dy(x_p)}{dx_{pi}} = 2\sigma \sum_{k=1}^{NSV} (\alpha_k - \alpha_k^*) K(x_p, x_k)(x_{pi}, x_{ki}) \quad (17)$$

In the equation (17), the component  $K(x_k, x_j)$  is the kernel of the system, and in the selection of radial basis functions, it takes the form:

$$K(x_k, x_j) = \exp\left(-\frac{(x_k - x_j)^2}{\gamma^2}\right)^{\sigma = \frac{1}{\gamma^2}} = \exp(-\sigma(x_k - x_j)^2) \quad (18)$$

4. Test results

4.1. The model approximating the course of technical wear of building structures

The basis for the inference on the influence of the indices  $w_{REM}$  and  $w_{ZAB}$  was the created predictive model in the form of the  $\epsilon$ -*SVR* network structure. The general characteristics of the model were illustrated in Table 1 and Fig. 1. The average of the absolute values of the error between the model mapping and the actual variables  $V_{er}$ , and the so-called *Success Ratio SR* [%] as a function of the relative error  $ep$  [%], were used as measures specifying the assessment of the quality of the resulting model [11]. These results were presented in Tables 2, 3 and 4.

In addition to the examined indices, the set of the input variables included as well: the variable describing the age of the building, the index reflecting the impact of continuous deformation in the mining area  $\epsilon$  (+), and the mining tremors intensity index  $a_{sg}$  [e.g. 24]. A detailed analysis confirming the necessity to include these variables in the model have been discussed in detail in [6, 7, 23].

All the analyses were performed in the *Matlab* environment using *LIBSVM* package [4] and *Genetic Algorithm Toolbox* [15].

Basing on the analysis of the results contained in Table 1, it is possible to conclude that the developed model is characterized by a high level of both fitting and generalization of the acquired knowledge, as evidenced by error values for training and testing sets, respectively. The confirmation of generalization features is also the number which has emerged, of 200 support vectors representing 33% of all references used for learning from the training set. An additional information on the “*smoothness*” of the resulting approximator is a relatively large width of the basis function  $\gamma = 0.82$ , which for the standardized variables  $N(0.1)$  allows to continuously cover the ranges of all the analyzed variables.

Table 1. Basic parameters summarizing the  $\epsilon$ -*SVR* model

Training set <i>MSE</i>	Testing set <i>MSE</i>	Number of support vectors <i>nSV</i>	Regularization parameter C	Width of radial basis functions $\gamma$	Tolerance margin $\epsilon$
0,0026	0,0027	200	0,790	0,820	0,036

Table 2. Values of the  $V_{er}$  error for the training set and the testing set, expressed in units of technical wear  $s_z$  [%]

Training set	Testing set
6,05 [%]	6,28 [%]

Analysis of the value of the  $V_{er}$  error (cf. Tab. 2) allows to evaluate the model for fitting and generalization. Values of the  $V_{er}$  error, expressed in units of the degree of technical wear, both for the training and testing sets, oscillate around  $\pm 6\%$ . This is a good result in terms of fitting the model prediction to the reference data. Moreover, the values of the  $V_{er}$  error for the training and testing sets are very similar, which confirms good generalization properties.

Similar conclusions are drawn from the analysis of the distribution of the *Success Ratio* measure with respect to the relative error  $ep$ . In Table 3, for the accepted limit of the relative error  $ep$  of 30%, the

Table 3. Summary of the SR values [%] in the domain of the relative error  $ep$  [%] for the training and testing sets

Training set		Testing set	
$ep$ [%]	SR [%]- cumulated for $ep < ep_i$	$ep$ [%]	SR [%]- cumulated for $ep < ep_i$
0,0	10,2	0	14,4
5,0	34,0	5	40,4
10,0	52,6	10	51,4
15,0	65,0	15	68,5
20,0	74,8	20	76,7
25,0	80,2	25	82,2
30,0	83,4	30	86,3
35,0	87,3	35	93,2
40,0	89,5	40	95,2
45,0	91,7	45	95,2
50,0	92,6	50	96,6
55,0	93,7	55	97,3
60,0	95,3	60	97,3
65,0	95,4	65	97,3
70,0	95,8	70	97,9
75,0	95,9	75	97,9
80,0	96,4	80	97,9
85,0	96,6	85	97,9
90,0	97,0	90	98,6
95,0	97,5	95	99,3
100,0	100,0	100	100,0

accuracy of prediction, expressed by the SR measure, is 83.4% for the training set and 86.3% for the testing set, respectively. The level of accuracy proves good fitting of the model to the reference data. On the other hand, the approximate value of the course of the SR measure (cf. Tab. 4) for the training and testing sets demonstrates that the generalizing properties of the model are retained.

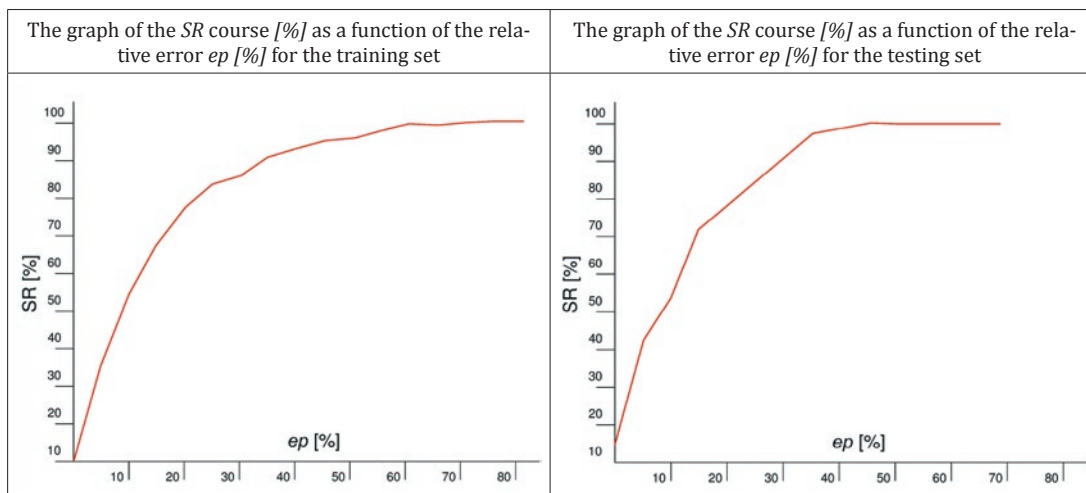
The comparison of the approximated values relative to the source data (Fig. 1) demonstrates that the correlation coefficient between them is close to 1.

#### 4.2. Studying the effect of the protection index $w_{ZAB}$ on the technical wear of building structures

In this paper, according to the description of the adopted methodology contained in Section 3.3, the problem of analyzing the influence of the protection index  $w_{ZAB}$  was solved by the simulation of the model for uniformly distributed grid of points in the space of input variables. The scale of each of the standardized variables was divided into 20 equal parts. Given that the repair index  $w_{REM}$  adopted only two categorical values, simulation space equal to 388,962 points was obtained. At each point, according to the dependence (17), the values of the partial derivatives with respect to each continuous variable contained in

the model was determined. The obtained values of the derivatives, calculated relative to the protection index  $w_{ZAB}$  were illustrated in Fig. 2 and 3. In Fig. 2, the value of the derivative calculated relative to the variable  $w_{ZAB}$  demonstrates that the modeled value of the technical wear of buildings increases in  $w_{ZAB} = 0 \div 0.64$ . Beyond this range, the degree of technical wear gradually starts to decrease.

On the other hand, Fig. 3 illustrates the course of the derivative calculated relative to the protection index  $w_{ZAB}$  in the domain of

Table 4. The graphs of the SR course [%] in the domain of the relative error  $ep$  [%] for the training and testing sets

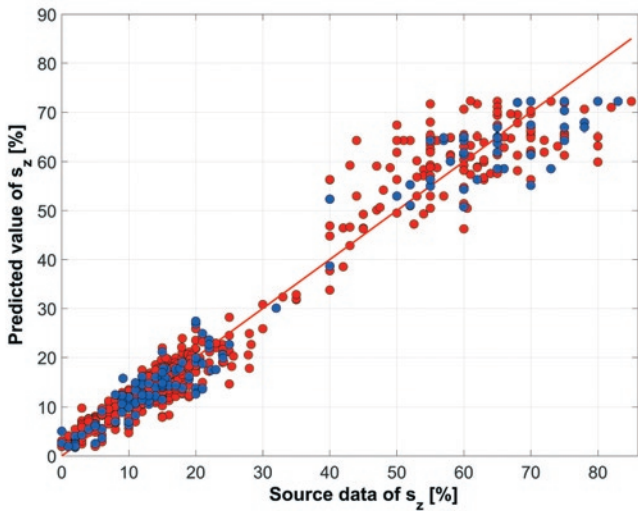


Fig. 1. Spread of the model prediction values relative to the reference data – the training set (red), the testing set (blue)

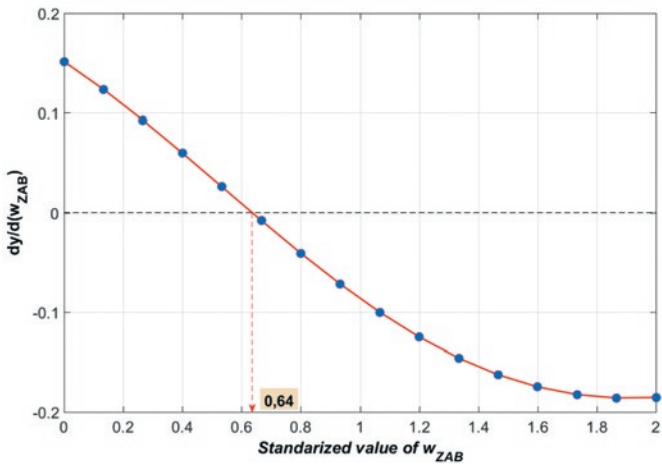


Fig. 2. Distribution of the gradient component for the variable  $w_{ZAB}$

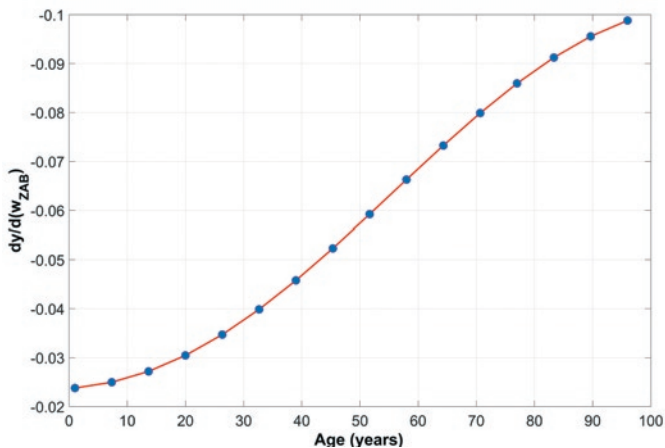


Fig. 3. Distribution of the gradient component of the model for the variable  $w_{ZAB}$ , illustrated relative to the variable describing the age of building structures

the variable describing the age of building structures. Apparently, the significance of the influence of protection increases with the age of buildings, causing a decrease in the modeled value of the degree of wear.

The analysis of the results leads to the following conclusions:

- noticeable reduction in the approximated values of the technical wear of buildings takes place at the time when the scope of the applied protective measures raises the category of structure resistance ( $KO$ ) above the given mining area category ( $KT$ ) by the relative value  $w_{ZAB} = 0.64$  (cf. Fig. 2).
- the influence of protective measures on reducing the approximated value of technical wear becomes clearer with time, that is, with an increase in its natural wear (cf. Fig. 3).

### 4.3. Studying the effect of the repair index $w_{REM}$

The repair index  $w_{REM}$  used in this study is a categorical variable. To evaluate the influence of this index, the developed model was subjected to a simulation using two sets of variables. One included non-renovated building structures, the second one – renovated ones. The remaining values of the variables in both cases were the same, which allowed for a comparison of the course of the technical wear and for drawing conclusions. The measure of the influence was adopted as the value of the difference between the obtained prediction for the first set of variables (non-renovated building structures) and the prediction for the second set. The results of the differences were summarized in the domain of the variable describing the age of the buildings, and they were illustrated in Fig. 4.

The analysis of the results presented in Fig. 4 shows that repair works clearly contribute to the decrease in the approximated value of the degree of technical wear. It is also noticeable that this influence is associated with the age of the building. The change in the values of the differences between the prediction of the model for non-renovated and renovated building structures reaches up to 8% for the buildings aged more than 90 years (cf. Fig. 4).

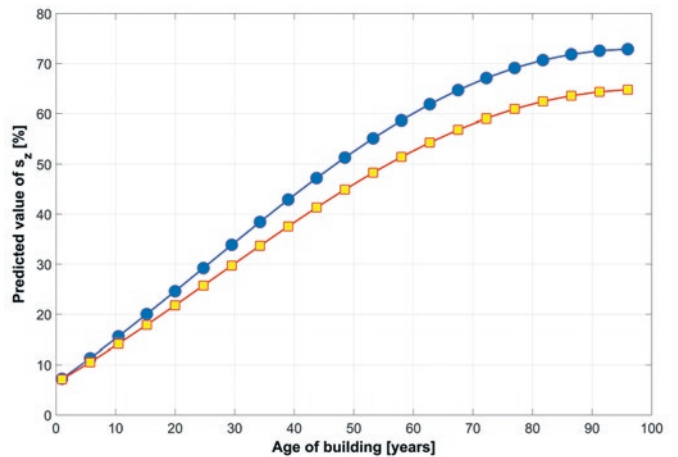


Fig. 4. Approximated values of the technical wear determined for a set of reference variables, separately for the renovated (yellow) and non-renovated (blue) building structures

## 5. Summary and conclusions

The results of the conducted research studies presented in this paper allow to draw conclusions both regarding the applicability of the *SVM* method for modeling technical wear of building structures, as well as the influence of different variables on the course of the modeled phenomenon.

In the case of the applicability of the *SVM* method, it was found that it allows to develop a model implementing nonlinear mapping of the multidimensional domain of input variables to the approximated value of technical wear. This model maintains the required level of accuracy and generalization. In addition, despite the lack of raw approximating function, the mathematical form, which is representative

for the model, allows to carry out the sensitivity analysis with respect to continuous variables.

As a result of the analysis of the influence of structural protective measures on the technical wear of the study group of building structures, the limit of significant influence of protective measures was established equal to  $w_{ZAB} = 0.64$ . The obtained value, according to the equation (1), demonstrates that the protection level becomes significant when, having applied the protective measures, the resistance category of a given building structure reaches the value of *KO* higher than the value of *KT* by 0.64. Generally speaking, it can be stated that protective measures contribute to the improvement of the technical condition of a building structure when they result in the resistance category *KO* exceeding hazard to mining area category *KT* by at least 1 category. This information can be very useful in evaluating the so-called *mining damage*.

The use of the SVM model simulation also showed the influence level of the repair index  $w_{REM}$  on the technical wear of the study group of buildings. The repair works, depending on the age of the buildings, have resulted in an improved technical condition ranging from 0 to 8%.

The analysis of the subject literature demonstrates that the *SVM* method may also be used in the case of approximating the response surface of a structure in the reliability analysis [1]. Therefore, it is planned to use this methodology for the risk assessment of building structures in mining areas subject to the reliability theory.

*The article was prepared as part of the AGH statutory research No. 11.11.150.005.*

## References

1. Alibrandi, U., Alani, A. M., & Ricciardi, G. - A new sampling strategy for SVM-based response surface for structural reliability analysis. *Probabilistic Engineering Mechanics* 2015; 41: 1-12, <http://dx.doi.org/10.1016/j.probenmech.2015.04.001>.
2. Bierlaire M. - Optimization: principles and algorithms. EPFL Press 2015.
3. Chandrashekar, G., & Sahin, F. - A survey on feature selection methods. *Computers & Electrical Engineering* 2014; 40(1): 16-28, <http://dx.doi.org/10.1016/j.compeleceng.2013.11.024>.
4. Chang C-C., Lin C-J. - LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2011; 2(3): 1-27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
5. Cortez, P., & Embrechts, M. J. - Using sensitivity analysis and visualization techniques to open black box data mining models. *Information Sciences* 2013; 225: 1-17, <http://dx.doi.org/10.1016/j.ins.2012.10.039>.
6. Firek K., Wodyński A. - Wpływ zabezpieczeń profilaktycznych oraz remontów na zużycie tradycyjnej zabudowy terenów górniczych (The influence of preventive security measures and renovation works on the wear of traditional buildings in mining areas). The scientific research problems of building. Pt. 3, pracazbiorowa pod red. M. Broniewicza, J. Prusiel. Wydawnictwo Politechniki Białostockiej 2007, ISBN 978-83-60200-33-9.
7. Firek K., Rusek J., Wodyński A. - Assessing the influence of mining impacts on technical wear of masonry buildings using multiple regression analysis. *International Multidisciplinary Scientific GeoConference SGEM*. Albena, Bulgaria, 2015.
8. Guo Y-M., Wang X-T., Chong L., Zheng Y-F., Cai X-B. - Diagnostyka uszkodzeń układu elektronicznego z wykorzystaniem Wielojądrowej Maszyny Wektorów Nośnych (SVM) zoptymalizowanej przy pomocy poprawionego algorytmu CPSO. *Electronic System Fault Diagnosis with Optimized Multi-kernel SVM by Improved CPSO*. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2014; 16(1): 85-91
9. Instrukcja GIG nr 12 2000: Zasady oceny możliwości prowadzenia podziemnej eksploatacji górniczej z uwagi na ochronę obiektów budowlanych (Guidline GIG 12 2000: Principles for Evaluating opportunities for underground mining for the protection of buildings. Katowice 2000.
10. 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.
11. Kuźniar, K. - Analiza drgań budynków ścianowych o średniej wysokości podlegających wstrząsom górniczym z wykorzystaniem sieci neuronowych, 2004.
12. Liao, P., Zhang, X., Li, K., Fu, Y., Wang, M., & Wang, S. - Parameter Optimization for Support Vector Machine Based on Nested Genetic Algorithms. *Journal of Automation and Control Engineering* 2015; 3(6): 78-83, <http://dx.doi.org/10.12720/joace.4.1.78-83>.
13. Lin, X., Yang, F., Zhou, L., Yin, P., Kong, H., Xing, W., ... & Xu, G. - A support vector machine-recursive feature elimination feature selection method based on artificial contrast variables and mutual information. *Journal of Chromatography B*, 2012; 910: 149-155, <http://dx.doi.org/10.1016/j.jchromb.2012.05.020>.
14. Łęski J. - Systemy neuronowo-rozmyte. WNT, Warszawa 2008.
15. MATLAB and Genetic Toolbox 2015b, The MathWorks, Inc., Natick, Massachusetts, United States
16. Osowski S. - Metody i narzędzia eksploracji danych. BTC. Legionowo, 2015.
17. Rojek I., Studziński J. - Porównanie różnych typów sieci neuronowych do lokalizacji awarii w sieciach wodociągowych. Comparison of different types of neuronal nets for failures location within water-supply networks. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2014; 16(1): 42-47
18. Rueda I.E.A., Arciniegas F.A. - SVM Sensitivity Analysis: An Application to currency Crises Aftermarks. *IEEE trans. Systems, Man and Cybernetic*, 2004; 34: 3.
19. Rusek J. - Developing a model of the course of technical wear of building structures in mining areas using the method of Support Vector Machine in regression approach. X Days of Mining Metrology and Protection of Mining Areas. Cracow, 28-29.05.2009. Academic Scientific Publishing House of AGH University of Science and Technology in Cracow. Cracow, 2009.
20. Russell S., Norvig P. - Artificial Intelligence: A Modern Approach (3rd ed.). Upper Saddle River: Prentice Hall 2010. ISBN 9780136042594.
21. Schölkopf B., Smola A.J. - Learning with Kernels. MIT Press. Cambridge, Massachusetts 2002.
22. Tikka J., Hollmen J. - Selection of important input variables for RBF network using partial derivatives - ESANN'2008 proceedings, European Symposium on Artificial Neural Networks-Advances In Computational Intelligence and Learning. Bruges (Belgium), 2008.

23. Wodyński A. - Zużycie techniczne budynków na terenach górniczych (The process of technical wear of buildings in mining areas). Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, Kraków 2007.
24. Wodyński A., Lasocki S. - Conception Of Assessment Of Mining Tremor Effects On Technical Wear Of Buildings. Acta Montana, series A, 2002; 21.
25. Vapnik V. - Statistical learning theory. John Wiley & Sons, New York, 1998.
26. Yeh, I. C., & Cheng, W. L. - First and second order sensitivity analysis of MLP. Neurocomputing 2010; 73(10): 2225-2233, <http://dx.doi.org/10.1016/j.neucom.2010.01.011>.

---

**Janusz RUSEK**

AGH University of Science and Technology  
Al. A. Mickiewicza 30, 30-059 Kraków, Poland  
E-mail: rusek@agh.edu.pl

---