Wiesław Piwowarski
ORCID: 0000-0003-2395-2664

PROJECTION OF REAL MASS DEFORMATION SCENES – INTELLIGENT SUPPORT

AGH University of Science and Technology, Kraków, Poland
piwowar@agh.edu.pl

Keywords: computational intelligence, modeling, differential and stochastic equations, the process of deformation of the medium

Abstract

The paper analyses the process of post-mining displacements generated by underground mining. Innovative mathematical structures for the modeling of hazard field emission were developed as strong solutions to partial differential equations in R^{3+1}. Moreover, a stochastic equation in $L^2(\Omega)$ (probabilistic space) was defined and applied as a model that takes into account the randomness of the process. Monitoring of a mining area based on solutions in the GNSS technology and classical geodesy supports the analysis of topological transformations of a given subspace. The data was archived and stored in digital form and then analyzed in many ways. The quality of the representation (measurements and modeling) was estimated with the use of incremental statistics. Thus, obtained distributions of density function are not ranked as normal distribution. The performed analyses make it possible to predict the optimal scenarios for post-mining environmental hazards.

ODWZOROWANIE SCEN RZECZYWISTYCH DESTRUKCJI GÓROTWORU – INTELIGENTNE WSPOMAGANIE

Słowa kluczowe: inteligencja obliczeniowa, modelowanie, równania różniczkowe i stochastyczne, proces deformacji górotworu

Abstrakt

W artykule przeanalizowano proces przemieszczeń pogórniczych generowanych przez górnictwo podziemne. Innowacyjne struktury matematyczne do modelowania emisji pola zagrożenia opracowano jako silne rozwiązania równań różniczkowych cząstkowych w R^{3+1}. Ponadto zdefiniowano i zastosowano równanie stochastyczne w $L^2(\Omega)$ (przestrzeni probabilistycznej) jako model uwzględniający losowość procesu. Monitoring obszaru górniczego, w oparciu o rozwiązania w technologii GNSS i klasycznej geodezji, wspomaga analizę przekształceń topologicznych danej podprzestrzeni. Dane archiwizowano i przechowywano w formie cyfrowej, a następnie analizowano na wiele sposobów. Jakość reprezentacji (pomiary i modelowanie) oszacowano za pomocą statystyk przyrostowych. Tak uzyskane rozkłady funkcji gęstości nie są klasyfikowane jako rozkład normalny. Przeprowadzone analizy pozwalają przewidzieć optymalne scenariusze zagrożeń dla środowiska pogórniczego.

1. INTRODUCTION

The mappings of the analyzed process are most often solutions to differential equations. On the other hand, the registration of the subspace topology, as a rule, does not coincide with the eigenstate of the operator of the examined feature – the knowledge about the process is basically statistical knowledge.

The article characterizes the description of the rock mass deformation process generated by mining in terms

of qualitative and deterministic modeling. Moreover, an attempt was made to define and describe the phenomenon as a stochastic process and to optimize the model in relation to the measurement.

The problem is formulated as follows:

For the real X process with càdlàg trajectories, the problem of defining the Y model was considered, the trajectories of which are uniformly close to the trajectories of the X process and have the smallest possible oscillation.

Optimization of the stochastic description was carried out with intelligent support. Registration of subspace topology is a procedure carried out in hardware systems. It was assumed that the representations and observations of the process would allow for the optimal creation of scenarios related to environmental hazards. The conducted empirical analyzes and mathematical modeling of the considered process are part of the so-called intelligent procedures

In research projects, even the best methodology does not prevent errors [1], [14] – there is a cognitive gap.

2. THE METHOD OF THE PROCESS REPRESENTATION AND ITS GENERALIZATION

The existing models of post-mining deformation description generate smooth macroscopic projections. The results of the observation of the real scene indicate that both the basins and the trajectories in the distinguished subspace are not smooth – often described as chaotic – which will be shown. This subspace is here a topological subspace.

Adequate descriptive knowledge was presented in relation to the examined fragment of reality. The determining ways of using knowledge in this area were indicated.

2.1. Evolutionary representations of mining area deformation – differential equations

The subspace of the rock mass in the vicinity of mining operations is a physical structure composed of a sequence of geological layers that make up the pile – Fig. 1.

The scene of the subspace of the rock mass in the vicinity of the mining excavation.
(www source. images. mining damage)

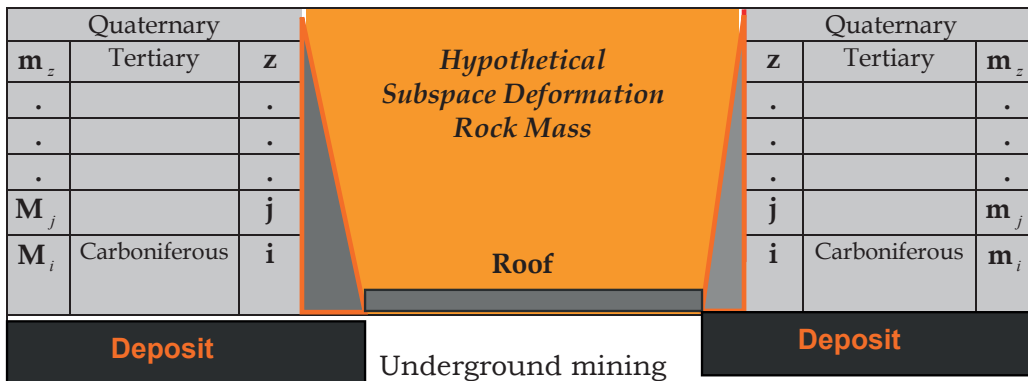Scena podprzestrzeni górotworu w otoczeniu wyrobiska górniczego

Quaternary			Hypothetical Subspace Deformation Rock Mass		Quaternary		
$\mathbf{m}_z$	Tertiary	$\mathbf{z}$		$\mathbf{z}$	Tertiary	$\mathbf{m}_z$
.		.		.		.
.		.		.		.
.		.		.		.
$\mathbf{M}_j$		$\mathbf{j}$		$\mathbf{j}$		$\mathbf{m}_j$
$\mathbf{M}_i$	Carboniferous	$\mathbf{i}$	Roof	$\mathbf{i}$	Carboniferous	$\mathbf{m}_i$
Deposit			Underground mining		**Deposit**	

Fig. 1. Hypothetical configuration of the physical structure of the rock mass
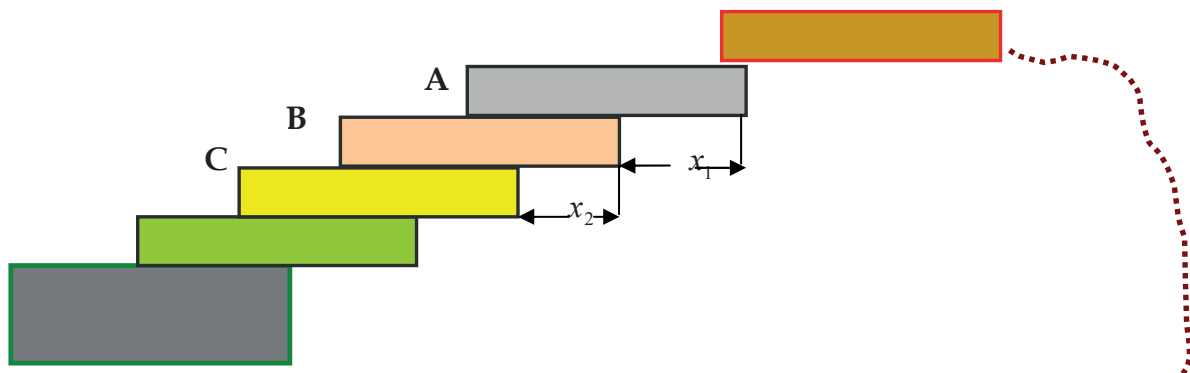**Rys. 1.** Hipotetyczna konfiguracja budowy fizykalnej górotworu

Fig. 2. Scheme presenting the discordance (packing) of rock blocks above the exploitation void
Rys. 2. Schemat ułożenia warstw skalnych (bloków) powyżej pustki ekspolatacyjnej

For each subset of the "*L*" layers, an attempt was made to arrange the layers of the stack in the area of underground mining so that they would not be destroyed. For this, consider a stack and some two layers from that stack "i" and "j". Let "i" be lower in the stack and M denote the weight of the layers over "j" and "o" the load on the stack. So we can write down:

$$o_i + m_i \geq m_i + M + m_i > o_j + M_j o_i + m_i$$
– layer coefficient "i"

Hence, it follows that the adjacent layers should not be transposed only when the higher layer has a lower coefficient – otherwise we replace them, obtaining a sorted set as an optimal structure (Fig. 1, 2). If the actual arrangement of the rock mass layers can be identified with a sorted pile, there is a chance that the rock mass layers will deflect with most minimal destruction – which can be described as quasi-stability.

It was assumed that the rock mass above the excavation in space R^2 is a stack of thin homogeneous blocks of unit length – see: Fig. 2. Single blocks are shifted in relation to each other in individual layers by x_i units. Such a construction – if no other forces act – is in equilibrium. The threat to the structure (loss of stability) appears when the concentrated force acts here.

The segments x_1 ,..., x_n form a certain sequence belonging to the real numbers, so we can introduce a function $M(x_i, x_{i+1})$ – such that the inequality holds:

$$\min(x_i, x_{i+1}) \leq M(x_i, x_{i+1}) \leq \max(x_i, x_{i+1}) \qquad (*)$$

where: $x_i, x_{i+1} \in \Delta$; Δ – interval,

which makes it possible to determine the following terms of the sequence.

A well-organized cognitive process in science consists in the proper generalization of experimental data and constructing appropriate theories. The interpretation of experimental data should be unambiguous – then it firmly shapes the framework of scientific theories.

The article considers the following problems: *modelling of the process of the medium destruction, observation of the process, visualization and analysis of representations* in the sense of minimizing deformation.

Due to the transitivity of the conjunction of logical equivalence, the quality of the representation is not well defined in terms of theory. This means that the theoretical tools at our disposal for proving do not ensure that in general we characterize the problem well.

The dominant majority of the processes taking place in nature are space-time phenomena. In general, the models of the processes are partial differential equations that are very difficult to solve and usually require simplifying assumptions about the geometry of space-time [3], [6]. The solutions to these equations then take different forms and lead to different interpretations of the same observational data and the proper characteristics of the problem in general.

Currently, the most common indicators that measure the state of post-mining deformation are the point displacement vectors (*u, v, w*) and the appropriate differential operations on these vectors (Euclidean geometry) – the representation according to S. Knothe [5], [6]:

The model of the displacement field description in R^1 – according to S. Knothe:

$$w(x, S) = \frac{W_{\max}}{r} \int_a^b e^{-\pi \frac{(x-\lambda)^2}{r^2}} d\lambda \qquad (2.1)$$

The formation of period subsidence
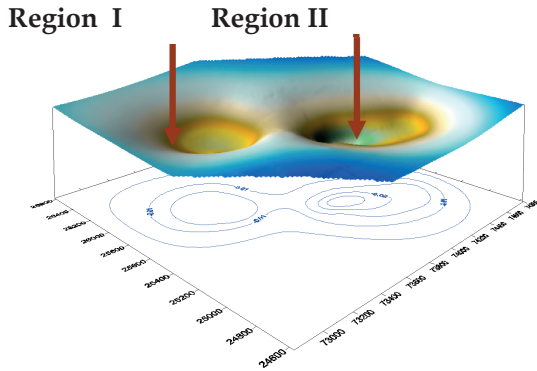
Region I   **Region II**

REGION I	REGION II
$V_{subsidence} = 16200m^3$	$V_{subsidence} = 36870m^3$
$V_{bed} = 110510m^3$	$V_{bed} = 70900m^3$
$a = \dfrac{V_{subsidence}}{V_{bed}} = 0.15$	$a = \dfrac{V_{subsidence}}{V_{bed}} = 0.52$

Fig. 3. Distribution of the subsiding troughs $\subset R^3$; $V_{subsidence} \leftarrow$ *technology* InSAR
(*) Characteristic parameters of the deformation subspace: Region I ^ II
Rys. 3. Rozkłady niecek obniżeniowych $\subset R^3$; $V_{niecka} \leftarrow$ *technologia* InSAR
(*) Charakterystyczne parametry podprzestrzeni deformacji (rejon I i II)

Solution (2.1) has the following property: all trajectories attain practically the same distribution – which results from the reference to constraints depending on time. In reality we do not obtain exactly the same distributions. The observed field of post-mining displacements is randomly disturbed by the heterogeneity of the rock mass structure and the random distribution of stresses.

In turn, the modeling of the displacement vector in R^{n+1} (space-time) is illustrated by differential equations (2.3), [3], [5].

where: λ – integration variable; r – parameter; $a \leq S \leq b$; W_{max} – process constant

If $b = a + \delta$ then $w(x, S) > 0$, in practice if δ is small then $w(x, S) = 0$

The subarea of surface deformation of the mining ground can be defined as $S_p \subset [a - r \leq S_p \leq b + r]$.

So, we can assume that $w(x, S)$ is a random variable in such subarea. Now, if we introduce an event A such that $P(A) > 0$ and if we replace the ordinary probability with a conditional probability, we get [13]:

$$E(W|A) = \sum_{k=1}^{n} w_k P(W = w_k|A) \qquad (2.2)$$

Solution (2.1) has the following property: all trajectories attain practically the same distribution – which results from the reference to constraints depending on time. In reality we do not obtain exactly the same distributions. The observed field of post-mining displace-

ments is randomly disturbed by the heterogeneity of the rock mass structure and the random distribution of stresses.

In turn, the modeling of the displacement vector in R^{n+1} (space-time) is illustrated by differential equations (2.3) [3], [5].

$$\left. \begin{array}{l} \dfrac{\partial W(x,t)}{\partial t} = c\big[W^k - W(x,t)\big] \\[2mm] W(x,0) = W_0(x_0,0) \\[2mm] W(x,0) = 0\big|\partial\Omega \end{array} \right\} \xrightarrow{\;Solution\;}$$

$$W(x,t) = W^k\left(1 - e^{-c \cdot t}\right) \qquad (2.3)$$

Solution (2.3) estimates the kinetics of the displacement process without taking into account perturbations.

3. INTELLIGENT VISUALIZATION OF THE REAL SUBSPACE

Intelligent visualization of the scene of terrain threats is an innovative concept in terms of theory and application, which involves modeling process propagated in the physical space. The model of the process streamlines and supports multiple analyses and optimizes the management of the discussed processes. Environmental threat processes are subject to constant structural reorganization. Hence, the modeling procedures need to be adapted to changing internal and external conditions. Dipolar procedure (visualization and mathematical modeling) as well as quaternions illustrate the actual

state of the representation of image sampling as implemented in computer animation.

Intelligent support makes it possible to connect, communicate with each other and facilitates optimal management of the distributed network elements equipped with intelligent process recorders.

Optimal *registration* (monitoring) of the geometry of the observed scene in **3D** is based on the approximation of the vector in the area of the actual scene, the identification of a motion vector and the procedure computationally efficient and acceptable in terms of measurement.

Layer visualization of 3D scenes

The state of deformation of the central element is a determination of the displacement of a subspace points – usually on the basis of distance measuring and leveling.

Optimal distribution of the points of observation network – in the sense of maximizing the accuracy of estimates of the derived parameter representation – includes:

Optimal distribution network observation points – in the sense of maximizing the accuracy of estimates derived parameter representation – includes:

- stationary points – measurement in a finite number of locations $l_1, l_2 ..., l_k$
- dynamic points – measurement of a finite number of moments. $t_1, t_2 ..., t_M$
- incomplete data supplemented by two objects classes similarities.

Advanced Technology Registration (e.g., receiver *GNSS*) and digital signal processing, provide high levels of resolution, accuracy and repeatability and protocol (*I P*).

Application layer

Performed procedures (illustrated by Fig. 4):

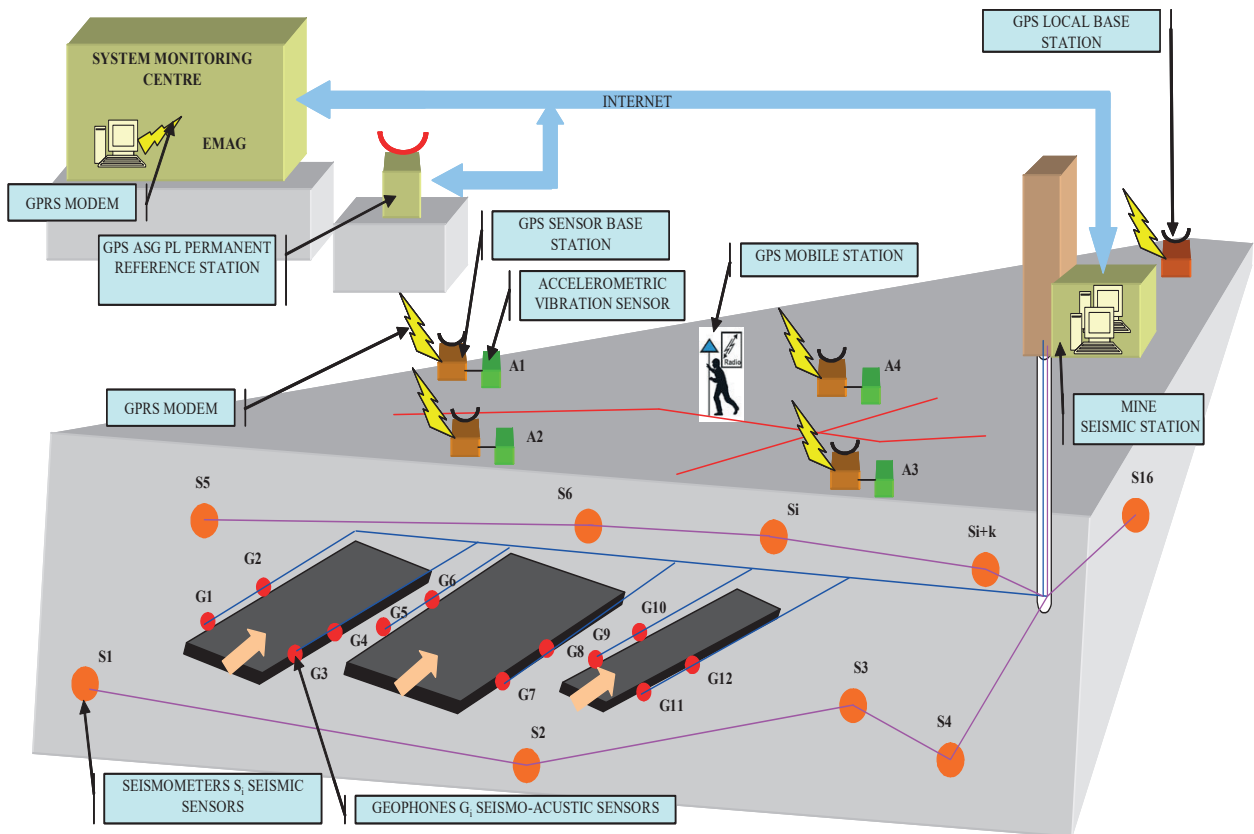- numerical modeling of the emissions of the generated hazard fields,

Fig. 4. The conceptual structure concerning the measurement part of the monitoring system in the mining area
Rys. 4. Struktura ideowa części pomiarowej systemu monitoringu na terenie górniczym

- process registration system in intelligent config-uration,
- optimization of the technology of core-shell struc-tures and formal analysis,
- application solutions in the decision-making area.

The modular structure of the measurement part of the system facilitates the control of large areas and the gradual increase of the system capabilities – as present-ed in Fig. 4. [10]. Stations equipped with measuring concentrators make it possible to collect and store the recorded parameters characterizing the destruction of subspace. Permanently collected data concerning the location of reference stations located in such way can be used at the post-processing stage in the monitored area, which should significantly improve the quality of the obtained observation results.

Optimization of the location of measuring sensors

The properties of the measurement track in the case of the observation of an active mining area are especial-ly important as regards the visualization, archiving the measurement results and object identification in order to construct an optimal mathematical model of the ana-lyzed process – especially a random one.

The application based on a map associated with the monitored area appears to be a layered structure com-prising:

- the layer of mining fields,
- the layer of sub-area deformation,
- sensor layer,
- the layer of object topography,
- the layer of the process representation $H(u, v, w)$, based on the model.

Let $\Theta \subset (w_1, w_1, ... w_n)$ – be a set of observation points monitoring the dislocation in the Ω area.

Each value of the expected reduction in the Ω set can be characterized as

$$w_i \subset \left(w_i^o, \sigma_i\right)$$

where:
w_i^o – the expected value of the displacement at point i,
ς_i – the deviation from the expected value of the dis-placement.

Covering the observation area O with the set of points $(x_i \leftarrow w_i)$ is acceptable if: $\forall_{i=1,2,...,n} \left|w_i - w_i^o\right| \leq \Delta$; where: Δ – permissible deviation.

For a highly diversified surface of the mining area (altitude), a semivariogram analysis is used. The loca-tion of the sensors is associated with scattered points located in a quasi-optimal manner. The process of rock mass displacement with state coordinates is based on the transformation of one state into another through the destruction of the medium.

4. THE PROCESS OF SCENE DISPLACEMENT IN 3D

Processes of rock mass displacement with state co-ordinates are based on the transformation of one state into another through the destruction of the medium.

This knowledge makes it possible to predict the course and effects of the phenomenon in various condi-tions, either controlled or uncontrolled by man.

Global properties are expressed as integrals stretched over the volume occupied by the system. The material derivative of the global quantity is given by the so-called general theorem of transporting extensive quantities.

All information concerning the vector field is car-ried by n functions X_i whose value encodes the value of the vector's coordinates. If the vector field X is suit-ably smooth (e.g. Class C^1), then the initial problem is locally well positioned. This means that for each point $x \in M$ there is one local integral curve, $f_x(t)$ where $\langle -\varepsilon, \varepsilon \rangle \supset t \rightarrow f_x(t) \in M$ starting from this point $f_x(0) = $ **x** and satisfying the initial condition $f_x(0) = $ **x** $X \not\subset$ Class C^1 As a rule, a vector field $X \not\subset$ Class C^1

4.1. The process of scene displacement in 3D

Percolations – Probabilistic model

Classical mathematics uses smooth projections. The observation results indicate that the process trajectories are irregular, often chaotic. Chaos theory eliminates La-place's statement concerning the predictability of a phe-nomenon.

Consider the rock mass fragment $X \subset Z^2$ as a set of elements $\{-n, -n + 1, ..., n-1, n\}$ (Fig. 5)

If each edge of the set X is open, deleting the element x_i will cause the dislocation of element w_i, otherwise a path of open edges connecting w_i to the element x is necessary in order that w_i may appear (see Fig. 5).
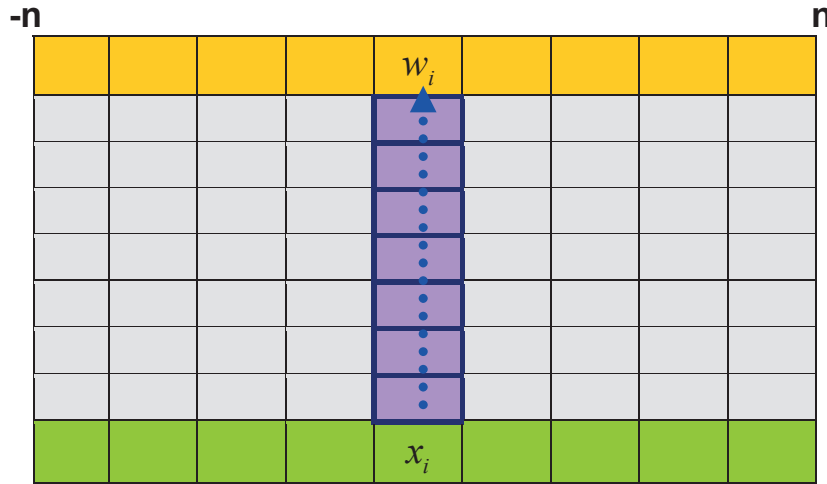
Fig. 5. The conceptual structure of a random displacement process; w_i – event
Rys. 5. Struktura ideowa losowego procesu przemieszczeń; w_i – zdarzenie

There is a principle in physics assuming that it is impossible to determine the exact state (event) of the process. However, it can be assumed that the existing states are equally probable (L. Boltzmann, an Austrian physicist) [10]. In the physical system – at each time step – it is possible to distinguish certain quantities that are measurable. These values make it possible to determine the probability distribution of an event.

Probabilistic interpretation

Let (Ω, Σ, P) – probabilistic space; $P(\Omega) = 1$.
   (X, A, μ) – measurable space.

If mapping: $\xi_i : \Omega \rightarrow X$ if for this $\xi^{-1}(A) \in \Sigma$,

   $A \in A$ and $\Phi_\xi(A) = P(\xi^{-1}\{A\})$ and

   $\Phi_\xi(X) = P(\xi^{-1}\{X\})$

Continuous decomposition: $f : X \rightarrow R^+$ this is:

   $\Phi_\xi(X) = P(\xi^{-1}\{X\})$; Whereas the measure:

   $\mu(A_1 \times A_2) = \mu_1(A_1) \cap \mu(A_2)$ $A := A_1 \times A_2$

   $\Phi_\xi(A_1 \times A_2) = P\{(\xi_1, \xi_2) \in A_1 \times A_2 =$

   $= P(\xi_1 \in A_1 \cap \xi_2 \in A_2)$

If then $\iint\limits_{A_1 \times A_{21}} f_1(x_1) f_2(x_2) = \int\limits_{A_1} f(x_1) dx_1 \cdot \int\limits_{A_2} f(x_2) dx_2$

The macroscopic states of the process representation are important, there is no exact solution to the problem (sometimes heuristic description is applied).

Definition:

The probability that starting with x_i $X \not\subset Class$ C^1 we reach point w_i for the first time in n steps – where $(n \geq 1)$ and $\{\xi_t\}_{t \triangleright 0}$ – can be written as (i):

$$P(\xi_{t_{n+1}} = i_{n+1} | \xi_{t_n} = i_n, ..., \xi_{t_1} = i_1) = P(\xi_{t_{n+1}} = i_{n+1} | \xi_{t_n} = i_n) \quad (i)$$

We designate a random variable here, where τ – the stopping moment, $\tau : \Omega \rightarrow [0, +\infty)$, the random variable being such that the following condition is satisfied:

$$\{\tau < t\} \in \sum_t \Rightarrow \{\tau < t\} = \bigcap_{n \in N} \tau < t + \frac{1}{n} \in \sum_t \quad (ii)$$

Thus, the probabilistic space $\Omega = \{\omega_1, \omega_2, ..., \omega_k\}$ is a finite set x_i, the family of events $\Sigma = 2^\Omega$, while the probabilistic measure is such that:

$$P(\{\omega_i\}) > 0 \quad i = 1, 2, ..., k \quad (iii)$$

We introduce σ – field Σ t $t \in \{0, 1, ..., T\}$, which we interpret as a set of knowledge about the process obtained up to the moment t. We establish that any $\sigma \subseteq \sum$ as well as any non-empty set U and σ – field constitute a set of events observable in the analyzed subspace up to the moment $(t-1)$ (or t, respectively). The modeling of the transformations of post-mining center (a stochastic process) [11], [14] is understood as a family of random vectors $\left\{\left(X_t^1, X_t^2, ..., X_t^m\right)\right\}$ $t \in T$ representing type: $\int\limits_0^T \omega_t d\omega_i$; $\left(\Omega, \sum, P\right)$ – probability space

Approximation Lemma:

1) $\forall n \quad \tau_n(\Omega)$ is a discrete set

2) $\tau_n^{(\omega)} \geq \tau_{n+1}^{(\omega)} \quad \forall \omega \in \Omega \quad \forall n \in N$

3) $\tau_n(\omega) \to \tau(\omega) \quad \forall \omega \in \Omega$

We call these moments predictable.

If the process $\{X_t\}_\tau$ is uniformly integrable then it is possible to approximate the process for the moment (t_{i+1}). The family $\{X_t\}_\tau$ is uniformly integrable $\Leftrightarrow$ when:

– $\sup_t E(|X_t|) < +\infty \leftarrow$ (by definition)

– $\forall \xi > 0 \quad \exists \delta > 0 \quad \forall A \in \Sigma \quad P(A) < \delta \Rightarrow$

$\Rightarrow \sup_J \int_A |X_J dP| < \varepsilon$

The measure of the set over which we are integrating:

$$P(\{|X| > M\}) \leq \frac{C}{M} < \delta$$

Theorem

If $\sup_t E(|X_t|) < +\infty$ then the sequence $\{X_t\}$ – almost certainly converges to integrable random variable $E(X_t) = E(X_t^+) + E(X_t^-)$.

A stochastic process $X_T := (X(t))$ is a collection of random variables $X(t)$, indexed with the same t values, defined in the same probability space. For two different states of exploitation we can distinguish two such probability spaces: $(\Omega_1, \Sigma_1, P_1)$ or $(\Omega_2, \Sigma_2, P_2)$, where Ω – the space of elementary events, Σ – σ-algebra, (Σt) – filtration, P – measure. The stochastic process is also a random variable.

Contamination of models (stacking, Stacked Generalization) is used in predictive data mining – treated as relationships of predictions concerning various models. The assignment model was implemented by the loss function (with reference to the following methods: quasi-Newton method, Simplex, Hooke-Jeeves method and Rosenbrock method [7], [8]). The adoption of a particular form of qualified model was based on the distribution of residues at both the modeling and verification stages – under the following condition: the expected value of about zero residues.

It cannot be unequivocally proved that the model is true. There are no reliable conclusions about the underlying relationships integrating the models.

Measurements of free surfaces are performed on a set of points of the space, setting coordinates for those points $(x_i, y_i, z_i)_{i=1}^n$ – monitoring. Observations of the mining area – as an undertaking of control – measuring character – are applied for the verification of a deformation model of a particular mining subarea. Those observations are also used as the basis for the analysis of the topological transformation of the given region. Convergence of the process $Y_i \to Y$ in topology L^2 is the convergence in relation to probability, i.e. $\exists \varepsilon > 0$, such that (4.1) holds:

$$\forall \varepsilon > 0 \quad \lim_{i \to +\infty} P(\{\omega : |Y_i - Y| \geq \varepsilon\}) = 0 \qquad (4.1)$$

Estimators of the parameters

$$\hat{\Theta} = \arg\min_{\theta \in \Theta} \sum_{j=1}^m \sum_{i=1}^n \left[y_i^j - \hat{u}(x^j, t_i; \theta) \right]^2$$

– discrete measurement (4.2)

$$\hat{\Theta} = \arg\min_{\theta \in \Theta} \sum_{j=1}^m \int_0^t \left[y_i^j - \hat{u}(x^j(t), t_i; \theta) \right]^2 dt$$

– continuous measurement (4.3)

The measurement points are usually stationary, while the observations are carried out in a finite number of measurement sessions $t_1, ..., t_k$, / as a result the following relationship (4.4) holds:

$$u_k^j = u\left(x^j, t_k : \hat{\theta}\right) + \varepsilon(x^j, t_k) \quad {}^{j=1,...,n}_{k=1,...K}; \quad \varepsilon(x,t)$$

– random noise. (4.4)

The conditions of the non-contradictory theory should be fulfilled here.

Algorithmically, the modeling procedures state that:

– each state of exploitation development is conjugate to the defining process,

– there is a transition between states,

– previous state $\xleftarrow{\text{operation (prefiksu)}}$ state of the next event,

– initial state l_0 – is an instance of the process in the specification clause,

– final state $l_N \to$ END without defining further actions.

The probability of generating the observation vector O_t by the M model is given as follows:

$$P(O_t|M) = P(O_{t1}|M) \cdot P(O_{t2}|M), ..., \cdot P(O_{ts}|M) \qquad (4.5)$$

Note that the final process (deformation) takes place after the destruction of the structure of all layers above the exploited seam.

The *state of the occurrence* of the event (deformation) can be written as (4.6):

$$\Omega_a = \{x: f(x) \leq .0\} \tag{4.6}$$

f (x) – resistance to destruction of the structure of a part of the rock mass layer

The certainty (uncertainty) of the hypothesis A is represented by the interval

belief ~ plausibility $Bel(A) \leftrightarrow Pl(A)$

$Bel(A) = \sum_{B \subseteq A} m(B)$ – belief function;

$Pl(A) = \sum_{B \cap A \neq 0} m(B)$ – credibility function

Accurate determination of the process states is practically impossible – hence the analyses are carried out in accordance with the stochastic approach. The orders determined by the stochastic dominance, the risk measure (if it maintains the first-order stochastic dominance) and the probability measure concerning the product seem to be particularly significant here in terms of the theory. The analysis of the measurement results shows that the trajectories of the process of medium deformation are usually characterized by limited fluctuations.

5. STOCHASTIC PROCESS

The real process must have a "non-zero" output, because through the output the existence of the process is revealed. There is a causal relationship between the concepts of "input" and "output", in the sense that knowing (observing) the input one can usually determine the output. The reverse correlation is generally ambiguous (does not hold).

Conceptual Structure: *Stochastic process, conditional expected value*

Mathematics Modules:

– Probabilistic: *Random variables, Distributions, Probability, Probability axioms, Bayes' rule, Chain rule, Random functions.*

– Parameter estimation: *Distance minimization, Parameter value estimation,* MLE (maximum probability).

5.1. Active mining area – characteristics

The processes of mining area deformation, observed based on the measurements, are finite sets characterized by irregularity – often referred to as chaotic phenomena. According to Devaney's mathematical interpretation [1], [6], [10] chaotic phenomena are characterized by the following properties:

– *they are transitive,*
– *they have a dense set of periodic points,*
– *they are sensitive to changes in initial conditions*

Transitivity means that the trajectories in any open set *A* are sent to each open set **B**. When pairs of non-empty open sets *A, B* $\subset$ [0,1] and $S^n(A) \cap B \neq 0$ then $S : [0,1] \rightarrow [0,1]$. This means that for each open set *B* $\subset$ [0,1] the trajectory started in *A* will be generated.

Density of periodic points means that periodic points form such dense set that: $\forall x \in [0,1] \wedge k \geq 1$ $S^k(x) = x$. The orbits of periodic points are finite sets, where k – the base period of the point *x*.

The property of sensitivity – $\forall$ representation O in the neighbourhood of the point *x* there is a point *y* and a natural number *n* that: $\left| S^n(x) - S^n(y) \right| \geq \delta$, where δ – the so-called constant of transformation *S*, sensitive to changes in initial conditions. It means that each subset of the open set [0, 1] contains arbitrarily close points (*x, y*), the trajectories of which differ significantly after some time. A slight disturbance is enough for the trajectory to assume a different projection.

Transitivity and the existence of a dense set of periodic points generate sensitivity to initial conditions. In addition: transitivity and density of periodic points $\subset$ topology. The initial conditions are measured in nature. Hence the implication that: chaos has a topological property (example – the Lorenz attractor). Deformation processes (components of the displacement field) are also phenomena of a measuring nature. The measures of the structure transformation are not additive.

Algorithmic modeling procedures state:
– each state of exploitation development is projected into the defining process
– there is a transition between states
– the previous state is an instance of the process in the specification clause
– final state END without defining further actions

The state procedure makes it possible to describe the probability $P(x_i, x_j)$ of being in a state at moment t state x_i and the probability $P(x_i, x_j)$ of transition from state x_i to state x_j which we can write:

$$\frac{dP(x_i, t)}{dt} = \sum_{j=1}^{n} P(x_j, t) \cdot P(x_j \rightarrow x_i) -$$
$$- \sum_{j=1}^{n} P(x_i, t) \cdot P(x_i \rightarrow x_j) \qquad (5.1)$$

The above equation can be interpreted as follows: the progress of time depends only on the current state of the system and is fully defined thanks to the knowledge of the transition probabilities. The first sum in equation (5.1) represents all transitions entering the state x_i, while the second sum – all transitions coming from this state.

The model is trained by rules and patterns, thus creating the autoassociator.

Differential equation and important concepts

If the process contains a random factor, the model of the process can be written as (5.2)

$$x'(t) = f(t, x(t)) + \xi(t)$$
$$x(0) = x_0 \qquad\qquad \xi(t) \equiv \xi_t$$
$$\text{– random quantity} \qquad (5.2)$$

For (5.2) we define a probabilistic space $(\Omega, \Sigma\{\Sigma_t\}_{t \geq 0}; P)$, where $\{\xi\}_{t \geq 0}$; and $P : \Sigma \rightarrow [0, 1]$, then yields

$P(0) = 0$, $P(\Omega) = 1$; $\sigma - algebra$ Σ on Ω is family $\{\Sigma\}_{t \geq 0}$ of subsets of Ω with the following properties: $0 \in \Sigma$

Let us consider the general stochastic process [8], [10], [12].

The stochastic process $\xi(t)$ is formulated as (5.3) [8]:

$$d\xi(t) = b(t; \xi(t))dt + \sigma(t; \xi(t))dW(t)$$
$$\xi(0) = \xi_0 \qquad (5.3)$$

- $t_1 \neq t_2 \Rightarrow W_{t_1}$ *and* W_{t_2} are independent; $E[W_t]$ for all t
- if $E[W_t^2] = 1$ then the function $(t, \omega) \rightarrow W_t(\omega)$; W_t – the Wiener process

Let (5.3) be recorded as an integral formula $\xi(t)$ over trajectories and as Itô's integral:

$$\xi(t) = \xi_0 + \underbrace{\int_0^t b(s; \xi(s))ds}_{\text{integral of time}} + \underbrace{\int_0^t \sigma(s; \xi(s))dW(s)}_{|integral\ of\ Ito} \quad (5.4)$$

where: b, σ – meet Lipschitz conditions.

Uniqueness of the formula (5.4)

Let: ξ_t or $\hat{\xi}_t$ be the solution of (5.4) then yields:

$$\xi_t - \hat{\xi} = \int_0^t [b(s; \xi(s)) - b\left(s; \hat{\xi}(s)\right)]ds +$$

$$+ \int_0^t [\sigma(s; \xi(s)) - \sigma(s; \hat{\xi}(s))]dW(s) \rightarrow$$

Probability:

$$P\left(\left|\xi_t - \hat{\xi}_t\right|\right) > 0 \leq P\left[\sup_{0 < s < 1} \geq \left(\xi_s^2 - \hat{\xi}_s^2\right)\right] > N_{\rightarrow 0}$$

N- arbitrary

Estimation (5.5): If N $\rightarrow$ 0 then: $P\left[\sup_{0 < s < 1} \geq \left(\xi_s^2 - \hat{\xi}_s^2\right)\right] \rightarrow 0$

$$\left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \end{array}\right\} (5.5)$$

Stochastic integral

$$\int_a^b \xi_t dW_t := \lim_{n \to \infty} \int_a^b \xi_n(t)dW_t \xrightarrow{\quad int\,egral \quad relatively \quad} .$$

$$E\left[\left(\int_a^b \xi(t)dW_t\right)^2\right] = E\left[\int_a^b \xi^2(t)dW_t\right] \qquad (*)$$

It is essential for the analysis of the process $\{W\}_t$ that for $t \geq 0$ the function should be right-continuous.

Observations and representation of the displacement process:

Let us define two spaces in general sense:

Spaces: (Ω, Σ, P) – probability space; (X, A, μ) – measurable space

where: Ω – the space of elementary events, $\Sigma - \sigma -$ algebra, (Σ_t) – filtration, P – probabilistic measure $X \in R$ (R^n) – state variable; A – min $(\sigma - algebra)$; μ – measure.

Let representation $\xi_i : \Omega \rightarrow X$ be the random quantity and let $\xi^{-1}(A) \in \Sigma$ then: $P(\Omega) = P(\xi^{-1}(x))$. If $f : X \rightarrow R$ then: $P(\xi \in A) = \int_A f(x)\mu(dx)$, hence: $f : X \rightarrow [0, \infty)$ and $\int_X f(x)\mu(dx) = 1$; moreover, the measure is defined as: $\mu_f(A) = \int f(x)\mu(dx)$, for $X = \{x_i\}_{i=1}^k$ $\rightarrow X = \sum_{i=1}^k x_i \cdot \mathbf{1}$

When there is a downward – sloping function, then the integral equation (5.4) may be written as (5.6).

The calculations in formula (5.5) are performed with the use of numerical methods. Solutions to partial differential equations are characterized by chaotic behaviour – which is related to the existence of an invariant measure $\mu(E)$; $= \mu(T^{-1}(E))$; where: $T : X \rightarrow X$; $E \in \sum$; $\Sigma - \sigma -$ set algebra .

The random measures of the processes:

$$P\left(\sup_{t \in T} \xi_t \geq \lambda\right) \leq \frac{1}{\lambda}\left[E\xi_a + E\xi_b\right]$$

$$\text{and } P\left(\inf_{t \in T} \xi_t \leq -\lambda\right) \leq \frac{1}{\lambda}\left[E|\xi_b|\right]$$

Let $\{p$ and $q\}$ – be conjugate exponents $\left([\frac{1}{p}+\frac{1}{q}]=1\right)$ then $E\left(\left|\sup_{t \in T} \xi_t\right|^p\right)^{\frac{1}{p}} \leq q\left(E|\xi_k|^p\right)^{\frac{1}{p}}$

Using (5.3) we can write a multivariate stochastic differential equation (5.6):

$$\left.\begin{array}{l} d\xi^i(t) = b_i(t;\xi(t))dt + \sum_{j=1}^{p} \sigma_{i,j}(t;\xi(t))dW^i(t) \\ \\ \xi_0^i = \eta^i; \quad \xi_t = u\left(\overline{t-t};\xi_t\right) \end{array}\right\} \quad (5.6)$$

Stochastic differential equation (5.6) represents the process $\xi(x,t) \in R^{2+1}$. Based on the multivariate Itô formula (5.6) can be written as (5.7):

$$\left.\begin{array}{l} d\xi_t = u\left(\bar{t};\bar{x}\right) - \int_0^t \frac{\partial u}{\partial t}\left(\bar{t}-s;\xi_s\right)ds + \\ \\ +\int_0^t \sum_{i=1}^{n} \frac{\partial u}{\partial x}\left(\bar{t}-s;\xi_s\right) \cdot b_i\left(\xi_0\right)ds + \\ \\ +\int_0^t \sum_{i=1}^{n}\sum_{j=1}^{p} \frac{\partial u}{\partial x_i}\left(\bar{t}-s;\xi_s\right) \cdot \sigma_{ij}dW_s + \\ \\ +\frac{1}{2}\left(\int_0^t \sum_{i=1}^{n}\sum_{j=1}^{p} \frac{\partial^2 u}{\partial x_i \partial x_j}\left(\bar{t}-s;\xi_s\right) \cdot \sigma_{ij}ds\right) \end{array}\right\} \quad (5.7)$$

Then the density of the invariant measure will be represented by the Frobenius – Perron operator [8].

The Wiener process may be expressed as an integral formula (5.8):

$$\int_0^T W_t dW_t = \lim_{n \to \infty} I(f_n) = \frac{1}{2}W_T^2 - \frac{1}{2} \cdot T \quad H \to L^2(\Omega) \quad (5.8)$$

The methods of representing knowledge in the form of probabilities are *artificial intelligence techniques* [12], [14].

An exact analytical solution of the equation (Fokker-Planck) describing the evolution of the system is impossible, it is possible to approximate that: the system can evolve from state (A) to state (B) with probabilities (P_a, P_b). If there has been a destruction of the structure of several sub-areas which make up a given center then the sub-area generating the event and the probability of occurrence of this event can be determined as follows:

$$\Omega_z = \bigcap_{i=1}^{m}\Omega_{z_i} \text{ ; } P[\Omega_z] = P\left[\bigcap_{i=1}^{m}\left\{\Omega_{z_i}\right\}\right] \Rightarrow$$

$$\Rightarrow P\left(X_B|X_A\right) = \frac{P\left(X_B \cap X_A\right)}{P\left(X_A\right)} \text{ etc.} \quad (5.9)$$

Ω_{z_i} – sub-areas of destruction of the original structure of the medium, $P(X_A,...)$ – probabilistic measure.

Knowledge about the process, its environment, cause-effect relationships and conditions regulating the analyzed events states that a probability measure cannot be precisely defined.

Therefore, in the article the support for the mapping of process states was based on neural network. The neural network model is mainly a *processing function*: compounding the scalar product of the input vector with the weight vector and the so-called activation function. Input weights are optimized in the course of network learning process.

Many different techniques were used during the implementation of the algorithms, with particular regard to genetic approach that should be distinguished here.

6. INTELLIGENT SUPPORT REPRESENTATION

The future cannot be accurately predicted as an event that is certain. Uncertainty results from the conditions of a specific event (process). Confidence regarding the event would mean that the outcome of the prediction is a determinant, and this is against the rules of physics. Therefore, the forecast result should be provided with a (probability) measure. Currently, almost every phenomenon is recorded. Datasets contain a lot of interesting information but are difficult to obtain with standard methods.

Artificial intelligence algorithms – using databases from the past, generate a forecast result with a max. probability.

The knowledge base consists of "n" rules, each containing "m" premises.

The model is trained using rules and patterns and creates a self-associator.

Mathematics Modules AI (Artificial Intelligence – tab. 1)

- **Linear algebra**: computational tool in artificial intelligence,
- **Numerical methods**: Optimization of weights in OR,
  Deep Learning Functions, Mapping Optimization → Error minimization
- **Training neural network** in propagation mode
  Estimators: $\underbrace{\textit{Single Value Estimator}}_{MLE}$;

  $\underbrace{\textit{Multiple Value Estimator}}_{MAP}$

Table 1. Intelligent Calculations (Artificial Intelligence Algorithms)
Tabela 1. Obliczenia inteligentne (Algorytmy sztucznej inteligencji)

Neurat networks	Fuzzy logic	Evolutionary algorithms
Rough sets	Uncertain variables	Probabilistic methods

Basic Approximation Theorem (Cybenko, 1989)

Let *F* be a continuous sigmoid function then the finite sum matrix assumes the following form:

$$G(x) = \sum_{j=1}^{N} \alpha_j \cdot F\left(w_j^T \cdot x + \theta_j\right)$$

Finite sums are dense in the space of continuous functions with *n* – variables on the unitary cube.

Conclusion: A single-layer network with any continuous activation function can approximate a continuous function with fixed accuracy (Unacceptable acronym DIDO).

A natural neuron works by producing an output only when the input signal exceeds a certain threshold. The bias value, like the input weights, is optimized in the network training process. The metric, as a parameter of the effectiveness of the feature detectors, carries information on the frequency of repetitions of the selected feature in the images of the training dataset in various scenes. The result of input data detection is a task scenario for neural networks in the training process [13], [14] (Fig. 6). Models initially learned on the set (data set) constitute the process of knowledge transfer and model tuning (transfer learning).

A "learning model" is any model that specifies the learning rules used by individual players and their interaction in the course of the training process when this process is optimized iteratively.
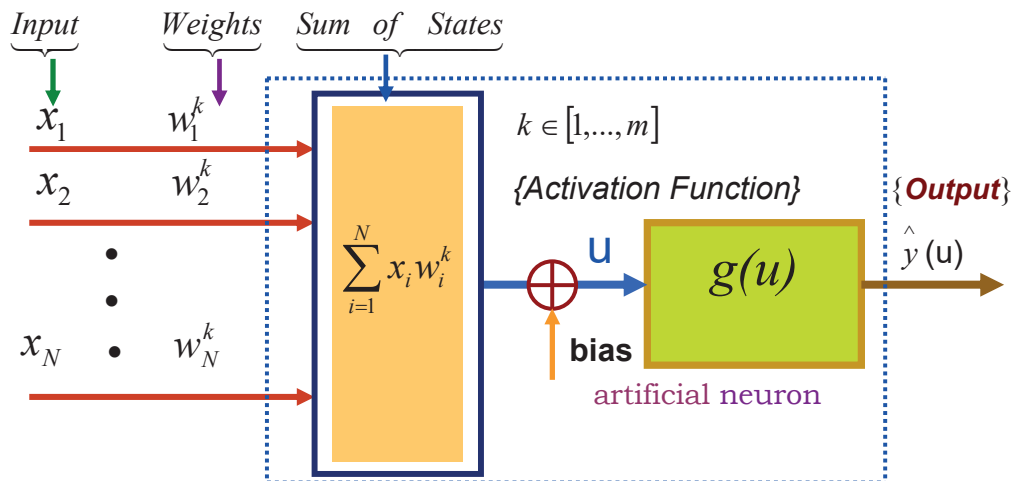

Fig. 6. The Structure (model) of Artificial Neural Network
Rys. 6. Model sztucznej sieci neuronowej

Optimization of the power of input set

Note that we should consider the observation resulting from the technical acronym: GIGO when preparing the input data (set X).

Input file X: $X = [x_1, x_2,...,x_N]$ is the basis of the output file representation Y^{Model}

$$[x_1, x_2,...,x_N] \cdot \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ w_N \end{bmatrix} = \Gamma(X,W) \sim Y^{Model}$$

The specified input data does not affect uniformly the representation Y^{Model}. Hence, it is important to correct the set $(x_i)_{i=1}^N \xrightarrow{selection} \left(\overset{o}{x} \right)_{j=1}^K$; $K \subset N$. The correction of the set X is necessary to ensure the optimal power of the input set. The optimal power of the input set X is: the power of a selected subset of the set X $\left(\# \overset{o}{X} \right)$ such that the following condition is satisfied:

$$\left(Y^{Measurement} - Y^{Model} \right)^2 = \text{minimum; where:}$$

$$\overset{o}{X} = \begin{bmatrix} \overset{o}{x}_1, \overset{o}{x}_2,..., \overset{o}{x}_K \end{bmatrix} \quad \Gamma\left(\overset{o}{X}, W \right) \sim Y^{Model}$$

Selected subset $\overset{o}{X} := K$. For $K \subset N$ we define a function $v(K)$ where the input set satisfies the following conditions:

$$\left. \begin{array}{l} v(K) \approx Y; \quad Y = Y^{Measurement} \\ v(K) = a \cdot (\#K) + (\#K) \cdot (\#K - 1) \cdot b; \quad v(1) = a \\ \text{For any subsets of } K, L \subset N \text{ if } K \cap L \neq 0 \\ \text{we obtain: } v_\Gamma(K \cap L) \geq v_\Gamma(K) + v_\Gamma(L) \\ \text{If the selected input subset is:} \\ v_\Gamma(K) + v_\Gamma(N|K) = v_\Gamma(K) + v_\Gamma(L) \\ \text{then the activation function } v_\Gamma(N) \text{ is additive} \Leftrightarrow \\ \Leftrightarrow v(N) = \sum_{i \in N} v_\Gamma(i) \end{array} \right\} \quad (°)$$

If K is a singleton {i} then for $K \subset I$ $v_\Gamma(K) = \max$. If there are series and the configuration $\langle I, v \rangle$ is an imputation and if there exist natural numbers $\{\alpha_1, \alpha_2, ..., \alpha_n\}$, then
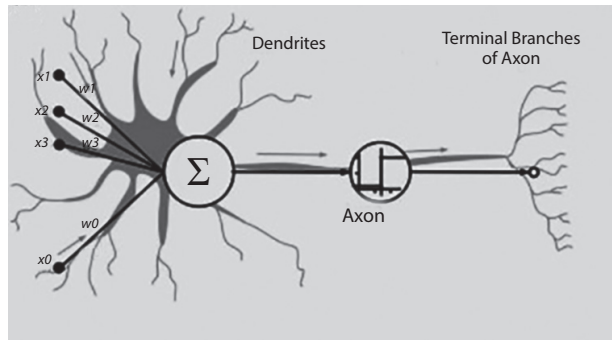

Fig. 7. Biological Model of the Neural Network (BNN)
Rys. 7. Biologiczny model sieci neuronowej
The speed of the Human Brain ~ 1017 [operations/sec]

$$\left. \begin{array}{l} \sum_{i=1}^n \alpha_i = v(I) - \sum_{i=1}^n v(i) \\ \text{If } \{x_1 < x_2 < x_3\} \text{ then for } \lambda \in (0, 1) \text{ we have:} \\ x_2 = x_1 + (1 - \lambda) \cdot x_3 \end{array} \right\} \quad (°°)$$

The conditions (°) and (°°) make it possible to select a subset $\left\{ \overset{o}{X} \right\}$

A neural network (SN) is a set of single processing units (so-called neurons) connected with each other, which from a mathematical point of view generate a description of a specific process – universal approximations [3], [12, [14] (Fig. 7). By the terms of Stone-Weierstrass theorem, it can be shown that a network with tangensoidal nonlinear processing units can approximate any continuous function with arbitrarily small error – Fig. 7 (PACKETS: Matlab (MATrix LABoratory), Statistica, SAS [8])

It should be noted that in the process of representation the weight matrix: $w_i = n^{(i)}, n^{(i-1)}$ and $\partial w_i = (n^{(i)}, n^{(i-1)})$ are particularly significant as regards the phenomenon modeling.

Equivalence of the stochastic process

Processes $\{y_i\}_{i=1}^n$ and $\left\{ \hat{y}_i \right\}_{i=1}^n$ are stochastically equivalent if: $\forall i \in I$ $P\left(y_i \neq \hat{y}_i \right) = 0$ (*)

The sequence $\left\{ \hat{y}_i \right\}$ converges stochastically to some random variable y $\Leftrightarrow$

$$\forall \varepsilon > 0 \quad \lim_{i,j \to \infty} P\left(\left|y_i - \hat{y}_j\right| > \varepsilon\right) = 0 \qquad (**)$$

Artificial Neural Network – mathematical conversion units

Model of the Neural Network:

$$\begin{bmatrix} x_{11}, x_{12}, \ldots, x_{1m} \\ x_{21}, x_{22}, \ldots, x_{2m} \\ . \\ . \\ . \\ x_{n1}, x_{n2}, \ldots, x_{nm} \end{bmatrix} \cdot \begin{bmatrix} W_1 \\ W_2 \\ . \\ . \\ . \\ W_m \end{bmatrix} = \begin{bmatrix} u_1^k \\ u_2^k \\ . \\ . \\ . \\ u_m^k \end{bmatrix} \Leftrightarrow \hat{Y} = g[U] \quad (6.1)$$

Updating weight values

$$\varepsilon_i \Big|_{i=1}^n = [y_i - \hat{y}_i] \Big|_{i=1}^n \qquad (6.2)$$

$$\left. \begin{array}{l} IF \cdot \varepsilon_i < 0 \Rightarrow w_i = w_{i-1} + \eta \\ IF \cdot \varepsilon_i > 0 \Rightarrow w_i = w_{i-1} - \eta \end{array} \right\} \qquad (6.3)$$

Learning parameter η: $\eta(t) = \dfrac{1}{t+1}$

Estimation error: $\varepsilon_i = y_i - \hat{y}_i$ $\qquad (6.4)$

$y\Big|_{i=1}^m$ – Measurement results; $\hat{y}\Big|_{i=1}^m$ – Modeling results

Activation function – artificial neural networks [1, 8, 14]

• sigmoid   • tangensioidal

$$g(u) = \frac{1}{1+\exp(-u)} \qquad g(u) = \frac{1-\exp(-2u)}{1+\exp(-2u)} \quad (6.5)$$

Input weights are optimized in the network learning process.

Updating the value of the weights:

$$w_{i+1} = w_i + \eta \cdot \nabla f(x_i, .)$$

where: $f(x, .)$ – multivalent differentiable function

η – correction factor signifying the length of the successive measuring points $\eta \to 0$. In that case (i.e., if $\eta \to 0$) the representation assumes the form $\left[\hat{y}_i\right]$, otherwise the weight correction is repeated and then the representation $\left[\hat{y}_i\right]$ proceeds.

Weighted neural sum

$$Y = [y_i]_{i=1}^n$$

Output of radial networks

$$\hat{y}_i \Big|_{i=1}^n = \delta_0^i + \left[\sum_{j=1}^m w_j^i \cdot g(u_j^i)\right]_{i=1}^m \quad (6.6)$$

w_j^i – weights of connections between the jth output and the ith input.

Over – parametrized networks may be able to interpolate any training data. The qualitative behavior of adaptive optimization algorithms is illustrated by formula (6.7):

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ . \\ . \\ . \\ \hat{y}_m \end{bmatrix} = \begin{bmatrix} \delta_0^1 + \left\{\sum_{j=1}^m g(u_j^1)\right\} \\ \delta_0^2 + \left\{\sum_{j=1}^m g(u_j^2)\right\} \\ . \\ . \\ . \\ \delta_0^m + \left\{\sum_{j=1}^m g(u_j^m)\right\} \end{bmatrix} \Leftrightarrow \hat{Y} = \delta_0^l + g(U^l) \quad (6.7)$$

Determining the response vector

The Method of Maximum Likelihood Estimation

$f(Y|\theta)$ – Plausibility function.

$$\overset{\wedge}{\Theta} = arg \min_{\theta \in R^n} \sum_{i=1}^n \log\left(f\left(Y^{Measurment}|\theta\right)\right) =$$

$$= arg \min_{\theta \in R^n} \left\{\sum_{k=1}^n \left(y_k^{Measurement} - y^{Model}(x_k, \theta)\right)\right\} =$$

$$= \left\|Y^{Measur.} - Y^{model}(., \theta)\right\|^2 \qquad (6.8)$$

The estimator has stochastic character.

Optimization Model

We have: $\displaystyle\sum_{j=1}^m \left\{\sum_{i=1}^n w_i^j \cdot x_i\right\} := U = \sum_{i=1}^n w_i^j \cdot x_i$;

$$f : U \times Y \to R ; \qquad (6.9)$$

As follows from the distribution of product functions when certain boundary properties are reached, we get:
$\forall u \in U; \quad \forall y \in Y \quad f(u, y) \le \sup f(u, y)$

Consequently, we have:

$$\sup_{u\in U}\inf_{y\in Y} f(u,y) \leq \inf_{y\in Y}\sup_{u\in U} f(u,y) \;\Rightarrow$$

$$\Rightarrow\; \max_{u\in U}\inf_{y\in Y} f(u,y) \leq \min_{u\in U}\sup_{y\in Y} f(u,y) \qquad (6.10)$$

On the assumption that $(u^*, y^*) \in U \times Y$ – we obtain the optimal solution:

$$= \min_{y\in Y}\; \sup_{u\in U} f(u,y) = f(u^*,y^*) \qquad (6.11)$$

Relative extremum: We have: $X \cdot W \cdot Y^T$ – as the base space.

Let us assume that for a fixed ε we have an optimal solution $(x_\varepsilon, y_\varepsilon)$, moreover

$$\forall x \in X \;,\quad \forall y \in Y: \quad f(x, y_\varepsilon) - \varepsilon \leq$$

$$-\varepsilon \leq f(x_\varepsilon, y_\varepsilon) \leq f(x_\varepsilon, y) + \varepsilon \qquad (6.12)$$

$\Rightarrow$ quasi – solution, where function $f(x, y)$ concludes the above model as a corresponding singleton.

Education of Network

– range of values – $\left\{(x_i)_{i=1}^k, (y_i)_{i=1}^k\right\}$
  – an education set

– weights – $\left\{(w_i^0)_{i=1}^k\right\}$ – a random set   } *Input Data*

– approximate result

$$y_j\Big|_{j=1}^m = \left[w_j^0 + \left(\sum_{i=1}^K w_i^j \cdot g(x_i^j)\right) \right]_{j=1}^m$$

where: $g(x)$ – *the function of activation of neural networks*

The Weights of Neural Network – Optimization

Modification weights – $\mathbf{w}_i^j\Big|_{i=1}^n = \left[-\eta \cdot w_{i-1}^j\right]_{j=1}^m$ } *Education*

where: $\eta > 0$ – education factor

$\mathbf{y}(x) = \Delta\mathbf{w} * g(x)$

PROCEDURE

$$A := \int_0^t b(s; \xi(s))ds\;; \qquad B := Y_{model}\;;$$

$$\xi(t) = \xi_0 + A + B\;;$$

Iterative Learning

We draw the weights of the network and set the input values "**x**".

Mathematical Model Algorithm :
– the reproduction procedure,
– crossing procedure.
– mutation procedure.

The initial representation is a chromosome with randomly assigned weights. Network learning applies the method of error backpropagation. Reproduction is the means generating successive representations – chromosomes. The iteration of weights can be presented as follows:

$$w_i^k = w_i^{k-1} + \left(1 - w_i^{k-1}\right)\cdot \overset{\vee}{\Delta}$$
$$w_i^k = w_i^{k-1} - \left(1 - w_i^{k-1}\right)\cdot \overset{\vee}{\Delta}$$
} $\overset{\vee}{\Delta}$ – parameter; w_i^k gen

$$\mathbf{X} = \left[x_{i,1}, x_{i,2}, \dots x_{i,k}\right]_{i=1}^n\;; \qquad \mathbf{Y}^{Measurement} = \left[y_i^M\right]_{i=1}^n$$

$$W^0 = \begin{bmatrix} w_1^0 \\ w_2^0 \\ \cdot \\ \cdot \\ \cdot \\ w_k^0 \end{bmatrix} \Rightarrow \mathbf{Y}^{Measurement} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} \to \varepsilon^0 =$$

$$= \frac{1}{m}\cdot\left\{\sum_{j=1}^m\left(y_j^M - \hat{y}_j\right)^2\right\}\;; \quad \hat{y}_j^{\,0} = \sum_{j=1}^k f\left(x_{i,j}, w_j^0, \vartheta\right)\;; \; \mathbf{k:=0}$$

The Iteration of Weights:

$$W^1 = \left\{\; w_i^0 + \left(1 - w_i^0\right)\cdot\overset{\vee}{\Delta}\;\right\}_{i=1}^n$$
$$OR$$
$$W^1 = \left\{\; w_i^0 - \left(1 - w_i^0\right)\cdot\overset{\vee}{\Delta}\;\right\}_{i=1}^n$$
} $k := k+1;$

$$W^{Measurement} = \begin{bmatrix} y_1^k \\ y_2^k \\ \cdot \\ \cdot \\ \cdot \\ y_n^k \end{bmatrix} \to \varepsilon^k = \frac{1}{m}\left(\sum_{j=1}^m [y_{j=1}^M - \hat{y}_{j=1}]\right)\; \hat{y}_i^{\,0} =$$

$$= \sum_{j=1}^m f\left(x_{j}, w_j^0, \gamma\right)$$

$$\left\{ IF\left(\ \varepsilon^k \le \varepsilon^{k-1} \ ; \varepsilon^k \ge \varepsilon^{acceptable} \ \right) \right\} \quad \textit{Then go to Weights Iteration}$$

If $\varepsilon^k \le \varepsilon^{acceptable}$ then

    For j:=1 step 1 to N   (t = 0)
    For i:=1 step 1 to M
    BEGIN

$$\hat{y}[i,j] = f\left(x_j^i, w_j^k, \gamma\right)$$

$$Y^{Model}[i,j] = \hat{y}\left(x_j^i, w_j^k, \gamma\right)$$

    END

If $\hat{y}^k(t) = \hat{y}^k(t+1)$ then it means the end of education

Covariance as a measure of dependence of random variables is not always correctly evaluated in terms of dependence (independence). Moreover, the covariance characterizes the pairs of random variables. As a rule, in the article we consider random vectors – hence the reference to a measure based on a supermodular function $\phi\left(Y^{Model}, (Y^{Measur})\right)$.

If $(Y^{Model}, (Y^{Measur})$ random vectors are the measure of the maximum approximation of the process, then we get: $E(\phi(Y^{Mode}) \ge E(\phi(Y^{Measur})$

The neuron activation function determines the output value of the "$Y \leftarrow y$" model.

Propagation $\varepsilon \xrightarrow{\textit{proof} - \textit{reading}}$ of the reverse error concerns weight identification.

Algorithm of the Process Analysis (see: tab. 2)

– Each state of the process is projected into the procedure action,
– State transitions constitute correlations between the actions of the prefix operator,
– The distinction between states is realized by application of the selection operator,
– The initial state is an instance of the process in the specification clause,
– The final state is a "Finish" process with no further action defined.

Table 2. Correlation: Measurement $\Leftrightarrow$ Model
Tabela 2. Korelacje: Pomiar $\Leftrightarrow$ Model

Variables	Correlation (Measurement $\Leftrightarrow$ Model) Significance level $\alpha \le 0{,}05000$ Observations: N=23					
	Y^I_{measur}	$\hat{Y}^I_{model}$	Y^{II}_{measur}	$\hat{Y}^{II}_{model}$	Y^{III}_{measur}	$\hat{Y}^{III}_{model}$
Y^I_{measur}	1.0	0,993834				
$\hat{Y}^I_{model}$	0,993834	1,0				
Y^{II}_{measur}			1,0	0,990443		
$\hat{Y}^{II}_{model}$			0,990443	1,0		
Y^{III}_{measur}					1,0	0,993616
$\hat{Y}^{III}_{model}$					0,993616	1,0

Where:

$$Y^I_{Measurement} = \left\{ y_1^I, y_2^I, ..., y_{23}^I \right\}; \qquad \hat{Y}^I_{Model} = \left\{ \hat{y}_1^I, \hat{y}_2^I ..., \hat{y}_{23}^I \right\}$$

$$Y^{II}_{Measurement} = \left\{ y_1^{II}, y_2^{II}, ..., y_{23}^{II} \right\}; \qquad \hat{Y}^{II}_{Model} = \left\{ \hat{y}_1^{II}, \hat{y}_2^{II} ..., \hat{y}_{23}^{II} \right\} \quad \textit{Sets} \ (\text{Measurement, Model})$$

$$Y^{III}_{Measurement} = \left\{ y_1^{III}, y_2^{III}, ..., y_{23}^{III} \right\}; \qquad \hat{Y}^{III}_{Model} = \left\{ \hat{y}_1^{III}, \hat{y}_2^{III} ..., \hat{y}_{23}^{III} \right\}$$

Generation $Y(\text{t},...)_{Measurement}$ Vectors $\{\hat{Y}(\text{t},...)_{Model} := \xi(t)\}\}$
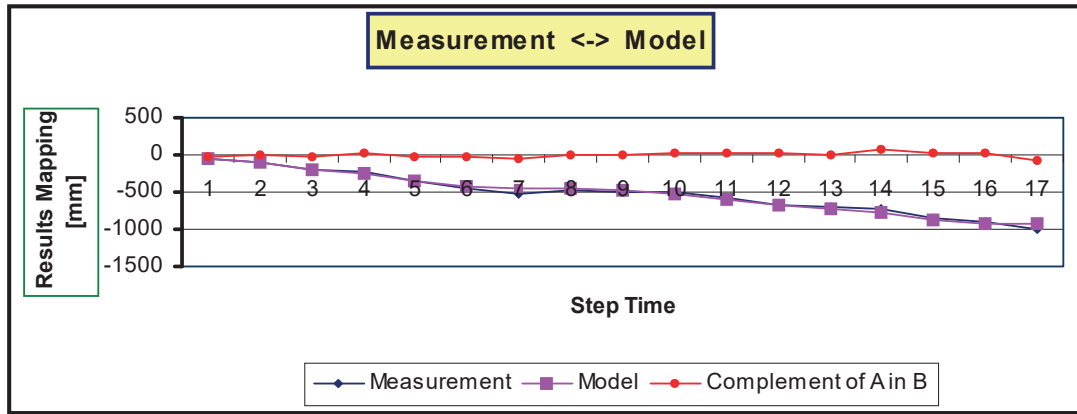
Fig. 8. Distribution of vertical displacements in time along the measurement line $\subset$ *2D*
Rys. 8. Rozkład przemieszczeń pionowych w funkcji czasu wzdłuż linii pomiarowej $\subset$ *2D*
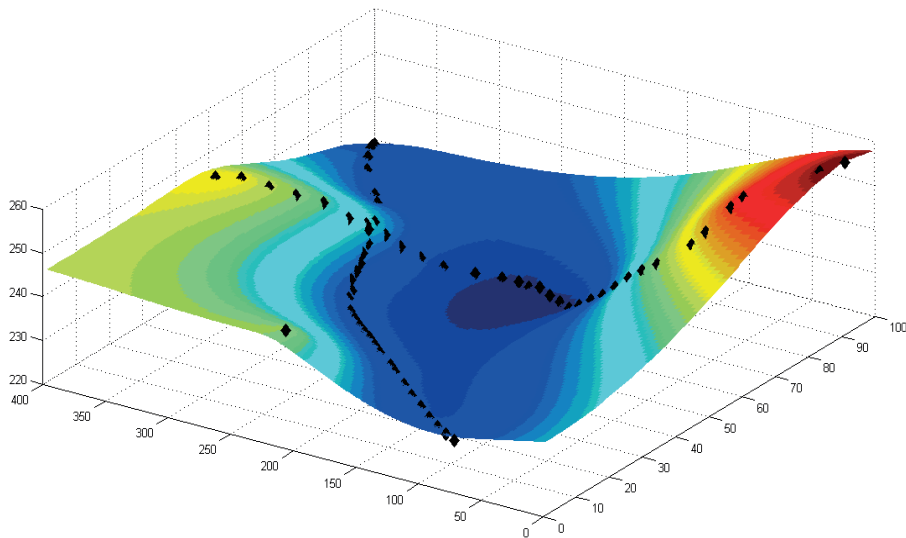Pomiar: $Y(t,…)$; Model: $Y^{\wedge}(t,…)$

Fig. 9. Subsiding trough $\subset R^3$
Visualization in **3D**, data: *measurement, model* $\rightarrow$ applicator Matlab – Mapping-Toolbox: Kriging.
Rys. 9. Niecka obniżeń $\subset R^3$ (Wizualizacja w **3D**, aplikacja: Matlab)

State procedure makes it possible to describe the probability of being found at time t in state x_i and the probability of transition from state x_i to state x_j $P(x_i, x_j)$, which can be written down as:

$$\frac{dP(x_i,t)}{dt} = \sum_{j=1}^{n} P(x_j,t) \cdot P(x_j \rightarrow x_i) -$$

$$- \sum_{i=1}^{n} P(x_i,t) \cdot P(x_i \rightarrow x_j) \qquad (6.13)$$

The first sum in the above equation represents all transitions entering the state (input source – source),

while the second sum represents all transitions that come out of that state (input sink – outflow). The model is trained by rules and patterns and creates a self-associator. The given dependencies in the information system describe the cause and effect relationships between the sets of attributes.

The correlations between the sets are important here. However, the distributions of the density functions do not belong to the normal distribution.

The test of independence of sets of observations of depressions and horizontal displacement is shown in tab. 3.
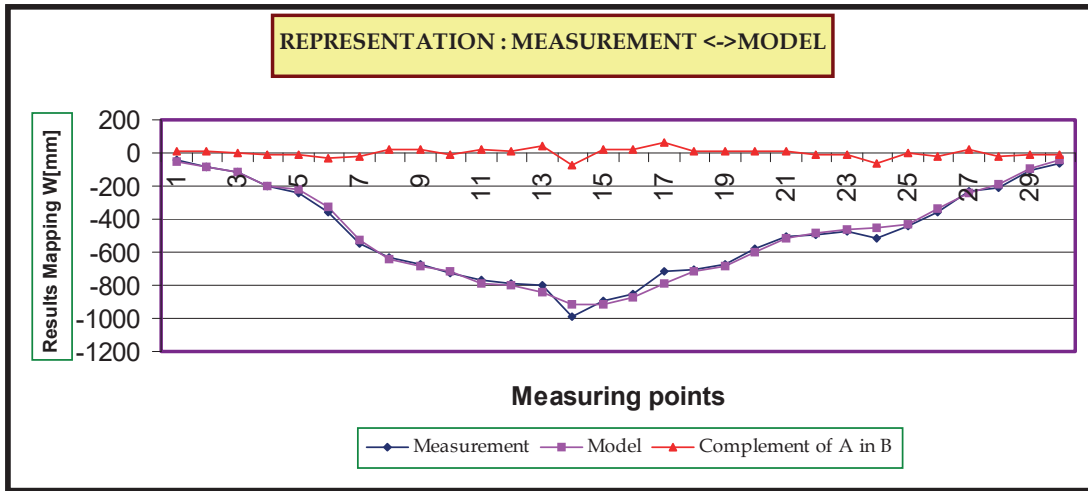
Fig.10. Visualization of the results of the MEASUREMENT W^{Measurement} $(X_t,…)$
^ MODELLING $\hat{W}^{Model}$ $(X,.)$
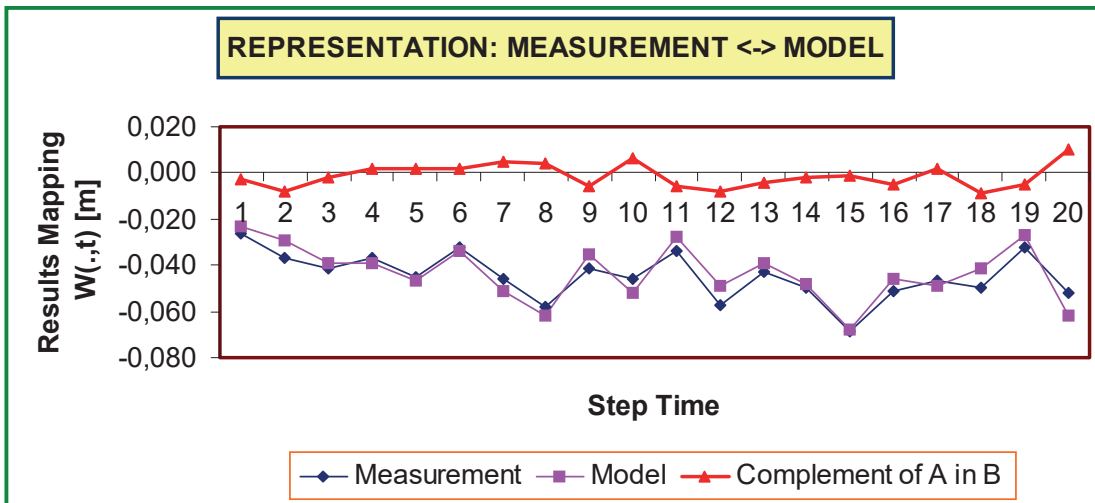**Rys. 10.** Wizualizacja wyników pomiaru i modelowania trajektorii $(W_t, – pomiar, W_t – model)$

Fig. 11. Visualization of the Subsidence Increment $W^{Measurement}(, t) \wedge W^{Model}(, t)$
Rys. 11. Wizualizacja przyrostu obniżeń $W^{Measurement}(, t) \wedge W^{Model}(, t)$

Estymatory statystyczne : (Tab. 2 i 3) potwierdzają zgodność odwzorowań: Pomiar ⇔ Model

The Statistic Estimators – as presented in (Tab. 2 and 3) – acknowledge the representation conformity: Measurement ⇔ Model

Comparing the results of numerical simulations with the results of experimental measurements constitutes a basic step in the process of fine-tuning mathematical models. Factor MAC (Modal Assurance Criterion) determines the degree of correlation between two vectors:

$$MAC_{X,Y} = \frac{\left|X^T \cdot Y\right|^2}{\left(Y^T \cdot Y\right) \cdot \left(X^T \cdot X\right)} \qquad (6.14)$$

where: X, Y – vectors to be compared.

The given dependencies in the information system describe the cause and effect relationships between the sets of attributes.

A set of non-trivial methods (algorithms, technological solutions, artificial intelligence tools, etc.) as well
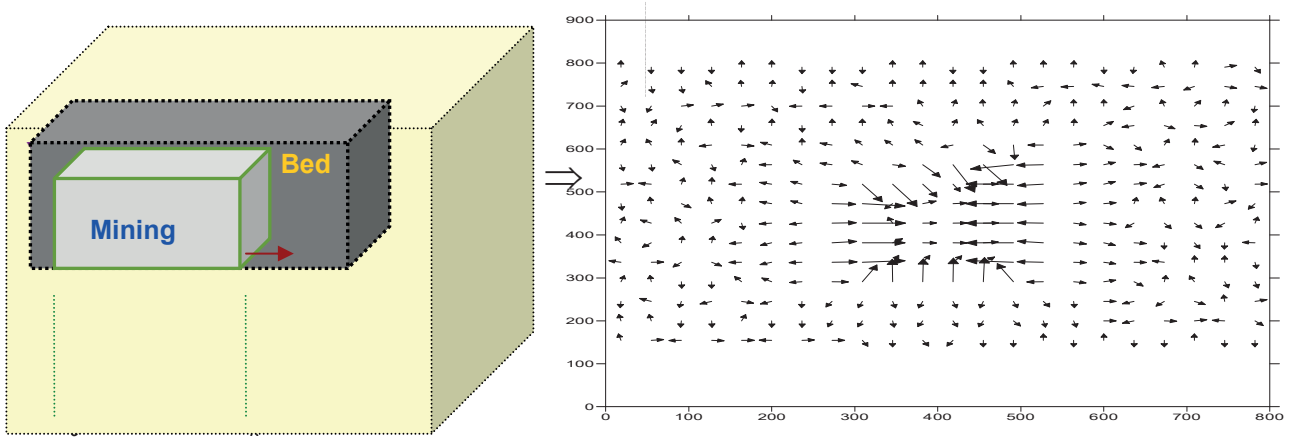
Fig. 12. Randomly disturbed subspace of the dislocation field $Y(X; t = t_k)$
Rys. 12. Podprzestrzeń pola przemieszczeń o losowym rozkładzie $Y(X; t = t_k)$

Table 3. The Statistic Estimators of Distribution : $Y_{MEASUREMENT} \Leftrightarrow Y_{MODEL}$ (Standard deviation)
Tabela 3. Statystyczne estymatory rozkładów : $Y_{POMIAR} \Leftrightarrow Y_{MODEL}$
Średnia (MEAN), Statystyka T (P) (STAT: T, P); Odchylenie Standardowe

VARIABLE	MEAN CLASTER 1	MEAN CLASTER 2	STAT T	STATP	STAND.-DEVIAT. CLASTER 1	STAND.-DEVIAT. CLASTER 2	QUATIENT VAR	STAT P VAR
$Y^I_{MEAS.} : \hat{Y}^I_{MODEL}$	−0,3698	−0,3692	0,000	1,000	0,065851	0,065851	1,000000	1,0000
$Y^I_{MEAS.} : \hat{Y}^I_{MODEL}$	−0,3698	−0,6473	26,36	0,000	0,065851	0,043137	2,330385	0,0028
$Y^I_{MEAS.} : \hat{Y}^{III}_{MODEL}$	−0,3698	−0,4686	10,65	0,000	0,065851	0,022162	8,829180	0,0000

as the detection of features related to pathology (CAD – Computer-Aided Detection) determine the quality of representation [1], [14]. The ontology systems used for semantic description and, above all, interpretation of changes in objects in an image were used here. The ontology is here a means of standardizing the description and reducing variability in the interpretation of an image with the corresponding mathematical descriptors and processing methods. The essence of the method is to estimate the point displacement gradient in the analyzed period of time. The determined total density of the distribution of the process probability is understood as the value of displacements at a point x and as a function of time t. The final result of such processing is a semantic response to the user.

Observation

The thick tails of the distribution of actual fluctuations in the occurrence of threats in a certain subspace of the environment (non extensive fluctuations) suggest that a conventional approach may not be appropriate for describing and modeling such phenomena. The combination of modeling tools directly with measuring devices and intelligent support to control real-time numerical simulation of processes are utilitarian very important.

Now that we have learned to fly through the air like birds, to swim underwater like fish, we only need one thing: learn to live on earth like humans
(G. B. Shaw)

CONCLUSION

Analytical considerations, monitoring of post-mining processes and numerical experiments presented in the article lead to the following conclusions:

1. The considered processes are modeled here through deterministic representations and as stochastic processes, with particular reference to statistical approach. The probability of a critical event occurrence is estimated based on the monitoring results.

2. An attempt was made to detect the structure-time patterns in the period of activity of the analyzed processes, applying such structures as: irregularities, statistical scattering of measurement data and the chaotic course of the analyzed process. The nature of the recorded data was assessed by means of statistical procedures.

3. The use of various types of algorithms describing spatial filtration – aimed at improving the quality of the process of detection with the use of computational intelligence made it possible to optimize the accuracy of the process representation.

4. Comparing the modeling results with the measurement results is an important step in process analysis. The main difficulty of the procedure is related to the fact that the mathematical model is usually marked by a large number of degrees of freedom, while the number of sensors for real measurements is limited and the measurement is affected by errors. Therefore, the strict equality of σ-bodies cannot be ensured in all cases. The application criterion of uniqueness means the condition: that the considered classes of events should differ from each other at least with the accuracy close to the measure of sets 0.

5. A measurement system composed of appropriate hardware sets is a tool for recording and transmitting data on the condition of the center in the form of sets of measurements carried out in the observed area. That is, in a sense, an algorithm for generating information about the condition of the monitored object – such as, for example, a mining area.

6. The significance test $\alpha = 0.05$ did not reject the hypothesis about the compatibility of distributions: $W^{Measurement}(x_i, \cdot) \sim W^{Model}(x_i, \cdot)$. This means that in the light of the test, the presented results of the modeling of the process $W^{Model}(x_i, \cdot)$ are uniformly close in relation to the process $W^{Measurement}(x_i, \cdot)$. Those processes are marked by as little fluctuation as possible.

7. Detailed knowledge of the process and its environment, cause-effect relationships and conditions regulating given events cannot be precisely specified.

REFERENCES

[1] Aliev R.A., Alizadeh A.V., Huseynov O.H., Jabbarova K.I.: *Z-Number-Based Linear Programming. "International Journal of Intelligent Systems",* Vol. 30, No. 5, 2015.

[2] Brockwell P. J., Davis R. A.: *Introduction to time series and forecasting.* Springer – Verlag, New York 1996.

[3] Bruna J., Mallat S.: *Invariant scattering convolution networks.* IEEE transsections on pattern analysis and machine intelligence 35(8) 1872 –1886. 2013.

[4] Evans C,: *Równania różniczkowe cząstkowe.* PWN. Warszawa 2002.

[5] Knothe S. : *Równanie profilu ostatecznie wykształconej niecki osiadania.* Archiwum Górnictwa i Hutnictwa, t. I, z. 1, Warszawa 1953.

[6] Knothe S.: *Wpływ czasu na kształtowanie się niecki osiadania.* Archiwum Górnictwa i Hutnictwa, t. I, z. 1, Warszawa 1953.

[7] La Valle S. M.: *Planning Algorithms.* Cambridge University Press, 2006.

[8] Luszniewicz A., Słaby T.: *Statystyka z pakietem komputerowym STATISTICA ™ PL. Teoria i zastosowania.* Wydawnictwo C.H. Beck, Warszawa 2001.

[9] Oksendal B.: *Stochastic Differential Equations. An Introduction with Applications.* Springer-Verlag, Berlin Heidelberg 2000.

[10] Piwowarski W.: *Estimation of the missing results of the surveys concerning a non-stationary post- mining dislocations field.* Geodezja i Kartografia t. LII, z. 3, Wydawnictwo PAN, Warszawa 2003.

[11] Piwowarski W., Juzwa J., Kuciara I., Siciński K.: *Analyses of parabolic processes to assess mapping stability of mining area ground dislocations in the INGEO system.* Mining – Informatics, Automation and Electrical Engineering, 04(528), pp 15-25, Katowice 2016.

[12] Shamir O.: *Distribution – specific hardness of learning neural networks.* The Journal of Machine Learning Research 19: 1135 – 1163. 2018.

[13] R.Sztencel, J. Jakubowski: *Wstęp do teorii prawdopodobieństwa.* Wydawnictwo SCRIPT, Warszawa 1980.

[14] Tadeusiewicz R.: *Sieci neuronowe.* Akademicka Oficyna Wydawnicza RM Warszawa 1993.

[15] Zastawniak T., Brzeźniak Z.: *Basic Stochastic Processess. A Course Through Exercises,* Springer-Verlag, Londyn 1999.