

New computer technologies, analysis and interpretation of geodata

Valery Taranchuk*

Belarusian State University, Nezavisimosti avenue, 4, 220030, Minsk, Republic of Belarus

Abstract. The article discusses the problems of development, tool filling, and usages of the integrated program complex of the composer of digital geocological models. Possibilities of adapting of digital models, interactive graphics visualization, and comparison of results are marked.

1 Introduction

The construction of digital geocological models nowadays is a mandatory part of expert's decisions in several domains, particularly while monitoring the state of environment, subsoil, while solving the problems of rational usage of mineral-resources base, in the projects about the protective measures related to the description of relief and engineering-geological structure of terrain, while planning underground site characterization, including geologic setting, rock and soil properties, and existing underground infrastructure. Geological models have a special role to substantiate the design and construction of underground structures, and will continue to be critical to the delivery of underground facilities with lower initial costs and risk, and better lifecycle performance.

Geological simulation is an independent stream, which includes the progressing of mathematical methods and algorithms; development of computer programs, which provide the cycle of models' construction, database creation, provisioning and maintenance. The corresponding software includes the loading from different sources and data preprocessing, correlation, creation of digital cubes of reservoir properties, interactive data analysis, visualization with the help of any type graphics, mapping. Despite the rich traditions in USSR's geological science, the construction of digital geological models is rather young stream among the CIS countries, it's developing rapidly, but still large state-owned and private companies, consumers of corresponding software systems, prefer the developments of global leaders. Such software systems even with the minimal equipment functionally complete are very expensive. Also the foreign systems are focused on special expensive computing equipment. That's why development and implementation of alternative computer geocological models are relevant. One of the most important part is an estimation of adequacy and accuracy of proposed digital models, but automation of installation, adaptation the models according to constantly incoming additional data, as well as the audit of results of processing initial data within the new methods of interpretation are especially relevant. Coding, implementation of visualization algorithms and adaptation of

* Corresponding author: valery.taranchuk@gmail.com

geological models are time consuming, involves the usage of unique mathematic methods. For example based on the experience of the development, maintenance and implementation of system «GeoBazaDannych» ([1], [2]), we can see that keeping it in up-to-date state, reflecting to constantly increasing software capabilities requires a large team of experienced programmers.

It seems that nowadays more effective and not as time-consuming will be an approach based on combination and integration into one software system the modules of modern versions of computer algebra system (CAS) and geographical information system (GIS). Moreover it should be considered that to solve the problem of processing initial data there's no specific GIS to be the full set of space-analytical methods and analysis tools. In many cases it's necessary to combine the tools provided by GIS with programs for static data analysis, tools for mathematically complex computations which include implementations of modern methods and algorithms of analysis and interpretation of spatial data.

2 Basic concepts

In the proposed approach while solving problems of mathematical modeling of the objects of geology there is a basic idea that the introduction of a digital field of the distribution of every site characterization parameter (for example, for describing the boundary surface) is a core and basis for building a computer geological model. However, the basic stage is building generalized surfaces which describe the topology of object, sequence of occurrence of geological bodies, layers, so called "set of shelves" ([2, 3]). For structured by layers 3D geological model there's an approach to build a model in a "constructor" mode, when building and editing of the model are made piece by piece, where the pieces are separate geological elements. Distributions of studied parameters for the layers are included in the description. For example, stress-strain behavior, abrasivity, thermal conductivity, permeability, geologic and human-development histories of a site can be considered as parameters. Initial data for these descriptions is usually the values of observed parameter in points with known geometrical locus, in the points which are placed on the area not regularly (for example, measurement data of seismic profile).

Let's mark out proposed and implemented software components, which are included into integrated computer system "The generator of the geological model of deposit" (GGMD). System is assigned for creation and estimated accuracy of configurable geological model based on the usage of CAS and GIS, "smart" methods of model adaptation while in service, "self-tuning" of models considering additional data from the actual development of processes. Development platform is computer algebra system Mathematica (www.wolfram.com/mathematica), language is Wolfram Language [4], geographical information system is Golden Software Surfer (www.goldensoftware.com). While programming in Wolfram Language technical solutions, described in [5], were implemented, moreover software system in a particular configuration can be used after it's built and saved in computable document format (CDF – www.wolfram.com/cdf). Calculations, user work with CDF version of application are possible on every personal computer. When viewing CDF version, hosted on webserver, viewer is automatically loaded in the form of browser plugin. Offline work is possible after the installation of free distributed CDF Player. Alternative additional configurations, which provide an interactivity of CDF version, are described in [6, 7].

2.1 The components of computer system GGMD

Let's make some clarifications for the illustrations of implementations marked above, the results of the usage of program modules which are selected and modified for the problems

we're solving. Below we mention the components, which are actually standalone program modules. They can be also considered as parts of automated workstation of specialist, who during the computational experiments works out techniques of adaptation of digital fields. We should specifically note an important technical solution – all the work steps with the complex of modules are provided with the possibility of import and export of obtained result with several configurations of output format. It provides the user with additional possibilities for performing similar calculations in different (including the others) applications, the comparison of results.

It is necessary to understand that creation and maintenance of geocological model don't expect to have unique solution to a mathematical problem. Subjective opinion, the qualification of an expert – are the factors that always take place in such activity. While working a user has to operate with data of different accuracy, some initial data is even conflicting; data density with measurements differs on different parts. That's why for construction digital models it's important to have tools for interactive data processing, simulation of possible situations of receiving and correction of input data. All the steps of working with data in GGMD include various options of graphic visualization, logging and comparison of incoming and placed to archive results. Complex's tools give a user possibility to "play" with initial data and compare the results with prepared etalons, what is more it's allowed to import and export the data and images and to scale them. Extensive data exchange possibilities are important for simultaneous work in several software environments.

In computer system GGMD the following tools are implemented:

- tools and patterns for preparation of reference model of digital field, which corresponds to the specified properties ("Digital field constructor");
- tools and several options of "distortion" of reference model;
- tools for data capture simulation, which are used in simulation practice ("Generator of profile observer");
- modules for calculation, visualization, comparison of digital fields approximation by several different methods ("Approximation component");
- tools and adaptation modules for digital model being formed ("Adaptation component").

The main idea and purpose of the development of this computer system is to choose the method of processing the original data by comparing the reference digital field and reconstructed by "observations". The reference distribution for a rectangular area is formed using mathematical descriptions. Each expert determines and includes own typical fragments in the model. Then the instruments of the GGMD perform "observations", simulate the measurement of the reference distribution, and the geometry of the measurement points and their accuracy are also determined by the user of the system and should approximately correspond to the initial data of the of the field of interest. The layout of points with measurements should not be regular. As a result of this stage, the user receives a set of data "observations", the main ones being the coordinates of the point and the value in it. The next step is to select the algorithm for processing the resulting set ("reproducing" the digital field) by performing interpolation and extrapolation. Comparison of the results of "playback" and the reference model will prompt the expert the method of processing, the geometry of the observation points

Digital field constructor (DFC). Software components from this group provide in interactive mode the construction of the model's surface from standard elements with accompanied visualization of mathematical description (analytic function), model's surface is interpreted as a relief – set of surface shapes. The construction is made in the module which is programmed in system Mathematica and includes the generation of surface equation – function of two arguments x and y which is continuous (or piecewise

continuous) and defined in the rectangle. User defines the boundaries of domain $xMin$ and $xMax$, $yMin$ and $yMax$ and surface height limits $zMin$ and $zMax$. Let's mark out that all the notations in DFC are given in format of InputForm (string format), that is accepted specially as some users can use application written in Excel, Delphi, C or others, where mathematics notation isn't supported inside the program code.

There are mathematical expressions (elements) which allow for the user of the GGMD to reproduce the behavior of the areas, which are typical for relief, in the set (library) of components of the function being formed. User at the first stage of reference model formation sets up a piecewise-defined function $zBasic(x)$. In the terminology of the complex, this is the basic profile – tape of specified width and length, which imitates the types of relief with the elements of plateau, slope, cliff. Then user can add perturbations of different shapes, sizes and orientations to the base surface. The construction with DFC module of basic profile from the fragments is possible with continuous transition “fragment – added fragment”, smooth transition, a jump (imitation of split). In case of continuous, smooth transitions the “connection” parameters of piecewise-defined functions are defined automatically by program module. Different variants of determining of the base surface are given in [8-10]. We note separately that the user of the system obtains an analytic expression.

The next step of construction with DFC module is usage of program module's tools to add perturbations, fragments of typical elements of relief to a basic profile. Template (patterns) library includes elements which correspond to perturbations (areas of distortion of basic surface) of different geometrical shape. While connecting the patterns it's possible to set interactively their position and size. Described mathematical elements, which imitate the following shapes of relief: hill, embankment, pit, excavation, trench, canal, quarry, ravine, hollow (vug), are included in basic package. It should be noted that all the elements listed above are specified by analytic expressions, such as $z = fFrgm(x, y)$. Besides it, those are written for the square $[-1,1] \times [-1,1]$, and then in the final function the arguments are scaled. System Mathematica includes a big amount of spatial graphic primitives of which cone, ball, cuboid, cylinder are used in DFC module, also different pyramids are included to the library. An example of reference surface model formation, which is obtained from the basic surface, by adding elements of the listed types (2 pyramids, 2 hills, pit, trench, embankment) is shown below. It's important that in the resulting equation the coefficients in the formulas of perturbation elements $fHill$, $fPyramid$, $fTrench$ are chosen by user while visual construction. When programming the module DFC, the following functions of the system Mathematica were used: Solve, Piecewise, Cases, Table, Plot, Plot3D, Manipulate. While visual examinations of the plots in DFC module, user have a possibility to define the coefficients of function expression by moving sliders or setting specific values on the panels which are the part of the interface of Manipulate function (more in [6, 7]). At this stage of constructing of the surface, the user of the DFC module also obtains an analytic expression. Different variants of determining of the reference surface are given in [8-10].

Tools and options of “distortion” of reference model (TDRM). The implemented methods and developed program modules of this group will be illustrated in the interactive mode during the report. Separate attention will be paid to demonstrating techniques for configuring neural networks for data processing, in which there are various kinds of measurement errors. When programming the module TDRM, the following functions of the system Mathematica 11 were used: RandomChoice, RandomVariate, SeedRandom, RandomSample, Interpolation, Plot, ListLinePlot, NetChain (NetEvaluationMode, TargetDevice, BatchSize, DotPlusLayer, ElementwiseLayer), NetModel (EvaluationFunction, EvaluationNet, SourceMetadata, TrainingSetData, TaskType), NetTrain (Method, BatchSize, MaxTrainingRounds, LearningRateMultipliers, LossFunction, Quantity, ValidationSet, TrainingProgressFunction). Specially selected

examples will explain the effects of changing options: ValidationSet, Scaled, NetEvaluationMode, BatchSize, TakeDrop.

2.2 Model visualization tools

More than ten options of 1D, 2D and 3D plots were implemented in GGMD system, including modules for obtaining and designing maps and inserts on them, plots on profiles, 3D visualizations of possible incisions (vertical, horizontal), clipping (simple and complex). The illustrations of the usage of several visualization tools are given below (examples and explanations are detailed in [8-10]).

Profile visualization. Geological profile is a graphic image in vertical plain of subsoil structure and deposits contained in it. In case of GGMD, a profile is a line, which is obtained at the intersection of analyzed surface and vertical surface in a given direction. Software module gives for user a possibility to choose the direction and set a specific position of the cutting plane with slider or by setting correct value. Wherein the section line is shown on 3D plot of surface and profile is displayed at the left part of the window.

Visualization of incisions, cross-sections. There are interactive modules with coordinates dashboard for 3D visualizations of possible incisions (vertical, horizontal), clipping (simple and complex).

Of course, isoline and zone (density) maps are part of tools for models' visualization, their examples are given below.

2.3 Tools for data capture simulation and mapping

While creating geological models the main data sources are the results of remote sensing, ground geophysical techniques that provide data from which subsurface conditions may be inferred. Basic site characterization tools such as remote sensing include the definition and ranking of satellite and terrestrial light, digital photogrammetry, radiometric technologies and interferometry method. Ground geophysical techniques include seismic refraction and reflection, spectral analysis of surface waves, crosshole tomography, geoelectric, electromagnetic, and potential field methods (gravity, magnetic). In practice, all the data mentioned are usually ambiguous, unstable, they have various kinds of inaccuracies. From the other side, for example the direction of improvement of seismic profiling technology is the development and justifications of the method of address compaction of seismic profiles. A part of this technology is designing of network of seismic profiles with certain density and relative uniformity.

The tools of observation profile generator (OPG) of GGMD that can be used in estimating the optimization of the network of seismic profiles provide two modes: manual and automatic. During the "manual" work, the expert interactively determines the scheme. In the automatic mode option, when the "Adaptation Component" is executed while processing initial data, the program module offers the choice of several variants of the schemes. Several results are mentioned and illustrated below.

The algorithm of preparing given below illustrations is the following. The user, using mentioned above interactive tools, draws profiles scheme, setting for each the number of nodes. In a separate software module of GGMD for each node a triple of numbers is formed: coordinates x and y , calculated value of surface function in point-node. The numbers of profile, node on profile and mentioned triple of numbers are exported in a file, the user chooses a format from offered options; it can also be exported in .XLS (XLSX, CSV). Let's notice, that such simulation of data obtaining made by surface levels in a measurement point gives an accurate value, but there are separate tools in GGMD for

“distortion of observation results” (on each separate profile) by the imposition of noise using sets of different functions with oscillations and random number generator.

Then the user in available application from a scattered set of measurement points, using one of selected algorithms, builds an approximating digital field. For example in program Golden Software Surfer any of the following methods can be used for interpolation, extrapolation: Triangulation with Linear Interpolation, Minimum Curvature, Polynomial Regression, Inverse Distance to a Power, Shepard's Method, Radial Basis Functions, Kriging.

3 Examples of results

Some results of seismic survey data preprocessing and designing of network of seismic profiles are discussed below. Examples are shown on figures 1 and 2. Let's mark out, that illustrations were prepared not to show qualitative result of model creation but to explain user work. Another important aspect – is that fragments of maps are illustrated in isolines, which is 2D graphics, but not three dimensional views, since discussed details are hardly distinguished on 3D images.

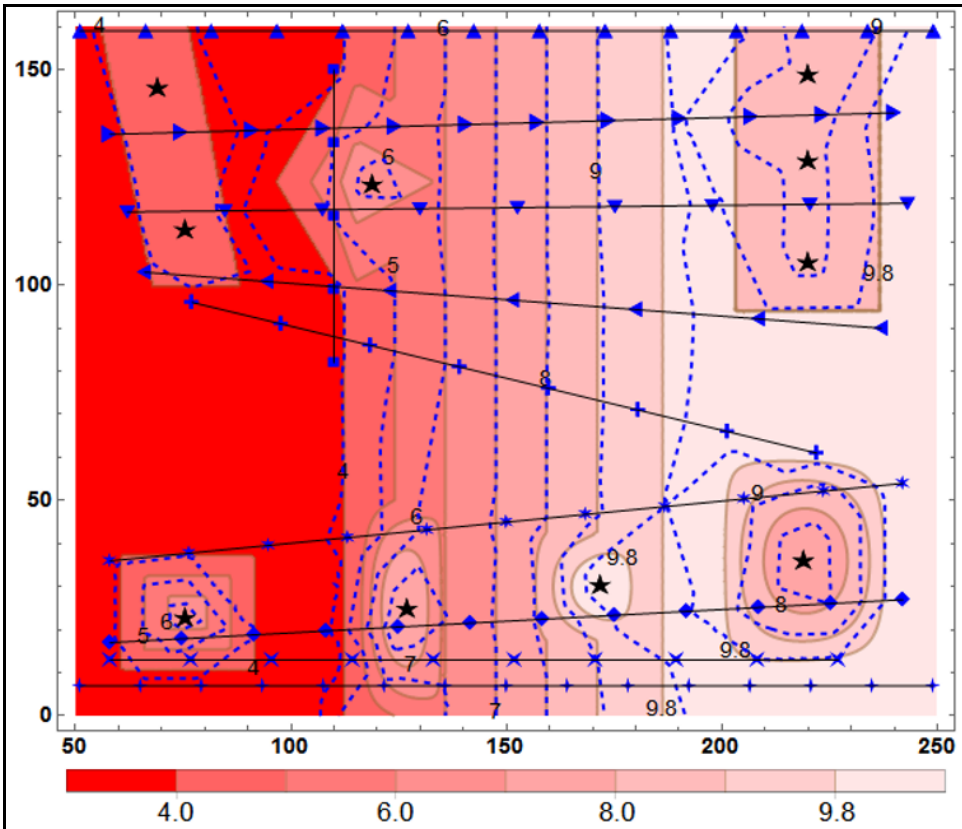


Fig. 1. Isolines of reference and reconstructed by 113 nodes fields

Figure 1 shows the results of comparison of reference digital field and reconstructed by the set of points on profiles field. The nodes are marked with various blue primitives. The scheme of profiles – one of offered by the expert to whom the information on the form of a

base surface and position of fragments-indignations has been given. Black stars note representative control points of fragments-indignations which also are used at approximation. The analytical expression for the surface formed and shown in the figures is the following:

$$\begin{aligned}
 z_{Surf}(x, y) &= f_{Origin}(x, y) + 3 \times f_{Pyramid1}[0.05 \times (x - 115), 0.03 \times (y - 124)] + \\
 &+ 3 \times f_{Pyramid2}[0.06 \times (x - 76), 0.07 \times (y - 24)] + 3.1 \times f_{Hill}[0.1 \times (x - 125), 0.04 \times (y - 32)] - \\
 &- 1.3 \times f_{Trench1}[0.05 \times (x - 68), 0.02 \times (y - 150)] + 1.4 \times f_{Hill}[0.08 \times (x - 168), 0.08 \times (y - 30)] + \\
 &+ 1.5 \times f_{Trench2}[0.04 \times (x - 220), 0.02 \times (y - 144)] - 3 \times f_{Hill}[0.05 \times (x - 220), 0.04 \times (y - 36)], \\
 f_{Origin}(x, y) &= z_{Basi}(x), \\
 z_{Basi}(x) &= \text{If}[x_{Min} \leq x \leq x_{Otk1}, f_{Plt}(x, 0), 0] + \text{If}[x_{Otk1} < x \leq x_{Otk2}, f_{Otk}(x, 0), 0] + \\
 &+ \text{If}[x_{Otk2} \leq x \leq x_{Sk12}, f_{Plh}(x, 0), 0] + \text{If}[x_{Sk12} \leq x \leq x_{Max}, f_{Sk1}(x, 0), 0]; \\
 f_{Plt}(x, y) &= z_{Plt}; \quad f_{Plh}(x, y) = f_{Otk}(x, y) - \text{perkoef} \times (x - x_{Otk2})^2; \\
 f_{Otk}(x, y) &= z_{Plt} + \text{Tan}(ug_{Otk}) \times (x - x_{Otk1}); \quad z_{Otk2} = f_{Otk}(x_{Otk2}, 0); \\
 f_{Sk1}(x, y) &= z_{Sk12} + \text{Tan}(ug_{Sk1}) \times (x - x_{Sk12}); \quad z_{Sk12} = f_{Plh}(x_{Sk12}, 0),
 \end{aligned} \tag{1}$$

where $x_{Min}=50.0$, $x_{Max}=250.0$, $y_{Min}=-0.0$, $y_{Max}=160.0$, $z_{Min}=0.0$, $z_{Max}=10.0$, $x_{Otk1}=110.0$, $x_{Otk2}=170.0$, $x_{Sk12}=190.0$, $z_{Plt}=1.0$, $ug_{Otk}=0.08$, $ug_{Sk1}=0.01$, expressions f_{Hill} , $f_{Pyramid1}$, $f_{Pyramid2}$, $f_{Trench1}$, $f_{Trench2}$ are the same as in [6-8].

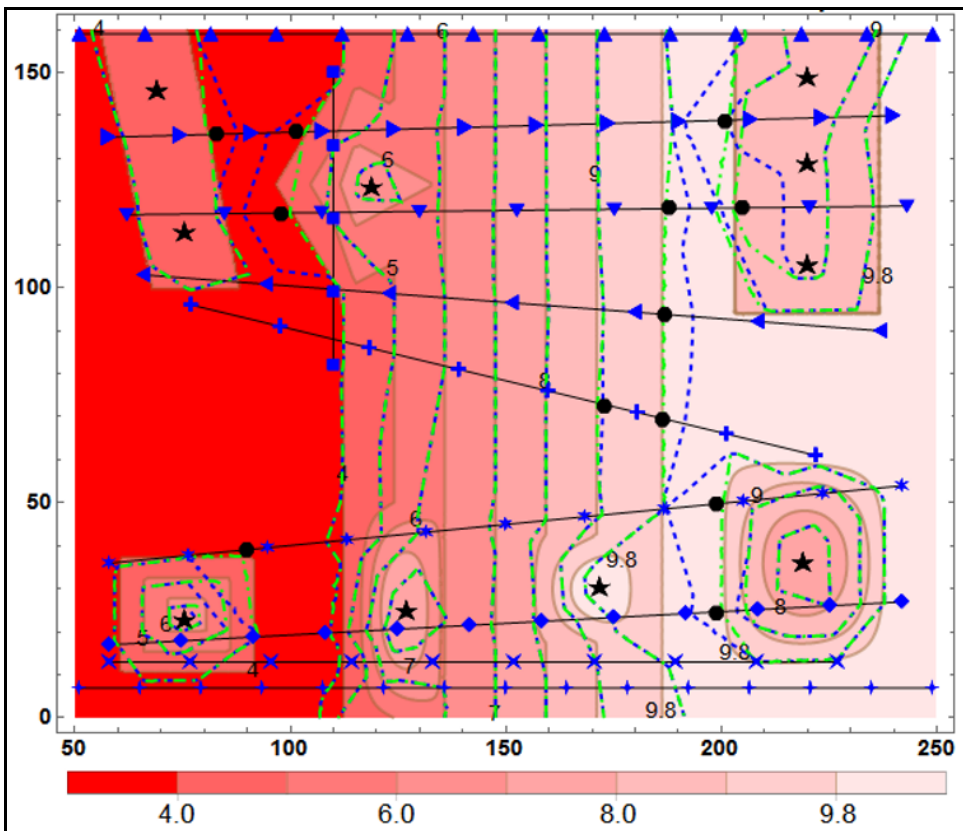


Fig. 2. The isolines of the reference and restored fields for 113 and 125 nodes

Solid brown lines mark isolines of reference digital field, which is built by analytically specified distribution $z_{\text{Surf}}(x,y)$, the levels of isolines were specified and selected so that the non-plateau area is covered with isolines relatively uniformly. Blue dashed lines mark isolines of the same levels, but for the functions, formed with first-order interpolation method in the Mathematica system. In the given example in calculation 113 nodes were considered.

An example is given for demonstration the tools of GGMD and options for user actions, not the best interpolation method was chosen specifically. For example, on such set of nodes of initial data Kriging method provides much closer to the reference results.

The results of proposed and programmed algorithm of address compaction of seismic profiles from the "Adaptation component" are shown on figure 2. The nodes of the second profiles scheme are marked with black primitives. It's important that previously used nodes are saved, the new ones supplemented and all of them are on the same profiles.

The complex module in which the algorithm of a method of address consolidation is realized, offers the user for acceptance some variants, one of which is resulted. As in example in figure 1, but by the second set of points with measurements isolines of a function, formed by the same method in the Mathematica system, are shown with green dot-dash lines. The comparison of configurations of isolines shows that new digital field reproduces reference field much more accurate.

4 Conclusion

The tools of automated workstation of specialist, who during the computational experiments works out techniques of adaptation of digital fields, which applied to problem of geoecological models formation, are described and illustrated. The developed integrated computer system allows using the possibilities of manipulating the initial data, provides analysis and comparison of interpretations and expert options, provides different methods for approaching and comparing the approximate and reference models.

References

1. S.A. Barvenov, M.F Kibash., V.B. Taranchuk, Vybranyya navukovyya pracy BDU, t. "Matematika", Minsk, 34–65 (2001)
2. M.G. Yasoveev, V.B. Taranchuk, *Geoekologiya Belarusi* (Pravo i ekonomika, Minsk, 2006)
3. D.V. Bulygin, V.Ya. Bulygin, *Geologiya i imitaciya razrabotki zalejei nefti*, (Nedra, Moskva, 1996)
4. V.B. Taranchuk, *Osnovy programmirovaniya na yazyke Wolfram* (Minsk, BGU, 2015)
5. V.B. Taranchuk, Vestnik Samarskogo gosudarstvennogo universiteta. Estestvennonauchnaya seriya, razdel Matematika, **128**, 6 (2015)
6. V.B. Taranchuk, Nauchnye vedomosti Belgorodskogo gosudarstvennogo universiteta, Seriya: Ekonomika. Informatika, **33**, 1 (2015)
7. V.B. Taranchuk, V.A. Kulinkovich, Informatizaciya obrazovaniya, **76**, 2 (2015)
8. V.B. Taranchuk, Problemy fiziki, matematiki i tehniki, **32**, 3 (2017)
9. V.B. Taranchuk, Nauchnye vedomosti Belgorodskogo gosudarstvennogo universiteta. Seriya: Ekonomika. Informatika, **43**, 16 (2017)
10. V.B. Taranchuk, Studia i Materialy, **14**, 2 (2017)