TEKA. SEMI-ANNUAL JOURNAL OF AGRI-FOOD INDUSTRY

TEKA. Semi-Annual Journal of Agri-Food Industry, 2021, 21(2), 26–35
https://doi.org/10.12912/27197050/139346
ISSN 2657-9537, License CC-BY 4.0

Received: 2021.05.28
Accepted: 2021.12.28
Published: 2021.12.30

METHOD OF SEARCHING FOR EXTREME OF MULTIDIMENSIONAL FUNCTIONS IN SOLVING ENGINEERING PROBLEMS OF COMBINE HARVESTERS

Titova L. L.[a]

National University of Life and Environmental Sciences of Ukraine, Heroiv Oborony Street, 15, Kyiv, 03041, Ukraine

Corresponding author's e-mail: l_titova@nubip.edu.ua[a]

Abstract
This article is devoted to the analysis of the most common optimization methods used in practical engineering problems of finding the extremum of multidimensional functions and forming on the basis of the identified properties of recommendations for choosing the best on different data sets. In the process of analysis, various implementations of gradient descent methods, impulse methods, adaptive methods and quasi-Newtonian methods were considered, the advantages and problems of each of the methods in their use were generalized. A computer program has been developed that implements all the considered methods. A computational experiment on three functions showed that the most effective methods were zero-order - Rosenbrock and zero-order - Powell.

Key words:
optimization methods, gradient descent method, stochastic gradient, quasi-Newtonian methods, objective function.

Introduction

In the process of designing intelligent control systems, the task often arises to determine the best values of parameters or structure of objects (Masek *et al.* 2017). This task is called optimization (Rogovskii *et al.* 2021j). Today, optimization problems and decision-making problems are modelled and solved in various fields of technology (Dubbini *et al.* 2017). Skills of mathematical substantiation of decision-making include skills of mathematical modelling of optimization problems (Palamarchuk *et al.* 2021), selection of adequate mathematical software (method, algorithm, software system) with the necessary justification, analysis of results and their interpretation in terms of subject area (Rogovskii *et al.* 2021g).

To estimate the approximation of the local extremum obtained using the methods of classical and stochastic gradient descent, pulse, adaptive and quasi-Newtonian, narrowing the neighbourhood and the decay vector (Viba & Lavendelis, 2006). Determine the dependence of the dimension of the problem and time costs in finding the global extremum of the goal function (Rogovskii *et al.* 2021e).

Methods for minimizing the function under nonlinear constraints can be divided into two classes (Rogovskii *et al.* 2021a). The first class includes those in which the search for a conditional minimum is reduced to an unconditional minimization of the function (Rogovskii *et al.* 2021h), resulting in the addition of a penalty for inconsistency of restrictions on the objective function (Kuzmich *et al.* 2021). In the methods of the second class the constraints are taken into account directly (Rogovskii *et al.* 2021b), and the search is on the admissible points with monotonically decreasing values of the objective function (Novotny, 2016). The first class includes methods of barrier and penalty functions (Zagurskiy *et al.* 2018). The second class includes methods of direct and random search to solve problems with limitations (Rogovskii *et al.* 2021d).

Formulation of problem

In the agro-industrial complex there are a large number of problems that can be solved using combinatorial optimization methods: the task of forming complexes of agricultural machinery, construction of crop rotations, the task of determining the profitability of agricultural enterprises, etc. (Pinzi *et al.* 2016). This is due to the discrete nature of the sets on which the optimization is performed (Rogovskii *et al.* 2021f).

For example, - discrete sets of crops, fields, permitted and unacceptable crop rotations, sets of forestry machinery and technological operations to perform agricultural work, and so on (Sergejeva *et al.* 2018). The main purpose of the technological process is to find a combination of required parameters (Rogovskii *et al.* 2021c), which provides the extreme of one of the quality criteria – the criterion of increasing the biological potential of plants (Yata *et al.* 2018), reducing energy consumption or reducing the impact on the ecosystem (Rogovskii, 2020). Choosing a fast and reliable algorithm for finding such an extreme is relevant (Drga *et al.* 2016).

Purpose of research

The aim of the study was to analyse multidimensional methods to optimize the search and find the optimal rate of convergence (number of iterations), the accuracy of finding the minimum function and the speed of the algorithm. It was planned to conduct a computational experiment for three functions with the help of a developed computer program.

The purpose of this computational experiment, on the one hand, is important self-importance to obtain conditions for the application of methods to minimize functions to solve practical engineering problems, and on the other hand it will search for a global extremum for the real (largest) dimension of the problem.

Research results and discussion

The task of optimization in general is reduced to the task of finding the extremum (minimum or maximum) of the objective function with given constraints. Its mathematical formulation is as follows: it is necessary to determine the value of the vector of variables $x = (x_1, x_2 \dots x_n)$, which satisfy the constraints of the form:

$$f_i(x_1, x_2 \dots x_n) \leq b_i, \quad (1)$$

for all in which the maximum or minimum of the objective function is achieved:$i = 1 \dots k$

$$f_i(x_1, x_2 \dots x_n) \to (min).$$

An admissible solution of the problem will be a solution that satisfies its constraints (1). The set of valid solutions to the problem is called the area of acceptable solutions (AAS). The final solution of the problem is a pair (x^{opt}, $f^{opt}(x^{opt})$), which consists of the optimal solution and the optimal value of the objective function (Nazarenko *et al.* 2021).

Methods of mathematical programming give a great variety of algorithms for solving this problem. In general, search algorithms implement methods of descent to the extremum, in which the value of the objective function is consistently improved until the extremum is reached (Titova, 2021). Depending on the possibility of finding the algorithm of local or global extremum, they are divided into algorithms of local and global search.

Algorithms in which the objective function takes the maximum or minimum value are intended for search of a local extremum or one of local extremums on set of admissible decisions. In their construction can be used as a deterministic descent into the region of extremum, and random search. Among the deterministic methods there are zero-order and gradient methods (1st and 2nd order). The first calculates the value of the function being optimized. The latter use private derivatives of the appropriate order. To find the extremum in cases where the type of optimized function is not fully known, or its structure is too complex, methods of stochastic programming or neural networks are used. The efficiency of the optimum search procedure - the ability to find a solution and convergence to a solution by speed depend on the type of function and the method used for it. Consider the strategy of each method in more detail, examining the minimization of the objective function for certainty (Luo & Guo, 2013).

Direct methods (zero order). Of the direct methods, the most well-known methods are: coordinate descent - alternate optimization of parameters along the axes by one of the known one-dimensional methods; spiral coordinate descent; rotating coordinates (Rosenbrock method); simplex search; Hook-Jeeves with a search for a sample; Rosenbrock; Powell, etc. (Astashev & Krupenin, 2017).

The method of coordinate descent is that as the directions of the descent trajectory from the previous search point x^{k-1} to the next x^k are taken in turn the directions of the coordinate axes x_i $(i = 1 \dots n)$. After descending one step on the coordinate x_1 there is a transition to the descent one step on the coordinate x_2 and so on, until you find the next search point x^k with coordinates x_1^k, $x_2^k \dots x_n^k$. The movement along the descent trajectory continues until the vicinity of the minimum point x^{opt} of the objective function, which is determined by the accuracy of calculations, is reached. To find the coordinates of the point x^k at each step of the iteration, you can use any of the methods of one-dimensional minimization: the method of golden section, the method of dividing the segment in half, the method of interpolation-extrapolation, and so on.

The method of spiral coordinate descent differs from that discussed above only in that the step h changes each time you go from finding the minimum for one variable to finding the minimum for another variable. In three-dimensional space, it resembles a descent to a depression in a spiral. Usually this method gives some reduction in search time, although this method is less effective in the presence of surfaces with "ravines". Attempting to move in any direction can cause "deterioration" of the target function. At the same time, advancing along the "ravine" can give "improvement" of the target function (Rogovskii *et al.* 2019).

Rosen Brock method aimed at eliminating one of the disadvantages of the method of coordinate descent - high sensitivity to the choice of coordinate system. In the process of searching by the Rosen Brock method, the coordinate axes are rotated so that one of the axes is directed along the direction of the "ravine". Consider the algorithm of the method in the case of one-dimensional minimization. At each iteration, the procedure performs an iterative search along n linearly independent and orthogonal directions. When a new point is obtained at the end of the iteration, a new set of orthogonal vectors is constructed (Fig. 1).

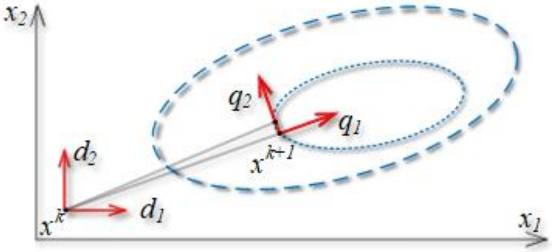


Figure 1. The scheme construction a new coordinate system by the method of Rosen Brock.

Construction of search directions. Let $d_1 \dots d_n$ be linearly independent vectors, normally equal to one. Assume that these vectors are mutually orthogonal, i.e. $d_i \cdot d_j = 0$ for $i \neq j$. Starting from the current point x^k, the objective function is successively minimized along each of the directions, resulting in a point x^{k+1}. A new set of directions $q_1 \dots q_n$ is built using the Gram-Schmidt procedure:

$$a_j = \begin{cases} d_j, IF\ \lambda_i = 0 \\ \sum_{i=j}^{n} d_i \cdot \lambda_i, IF\ \lambda_i \neq 0' \end{cases} \quad q_i = \frac{b_j}{|b_j|}. \qquad (2)$$
$$b_j = \begin{cases} a_j, IF\ j = 0 \\ a_j - \sum_{i=1}^{j-1} (a_i \cdot q_i) \cdot q_i, IF\ j \geq 2 \end{cases}$$

The new directions, constructed as described, are linearly independent and orthogonal. And although Rosen Brock's method eliminates the problems associated with obtaining a solution of a given accuracy, but this approach relatively increases the search time, which is a relative disadvantage of this method.

The Nelder-Mead method (simplex search) uses a geometric configuration called simplex. A simplex is a convex polyhedron with the number of vertices equal to $n + 1$, where n is the dimension of space. Its important feature is the ability to build a new simplex on any face of the source, by moving the selected vertex to some distance along the line connecting this vertex with the center of gravity of other vertices of the simplex.

The algorithm begins with the construction of a regular simplex in the space of independent variables of the problem and estimating the value of the objective function at its vertices. Then the point x_i with the largest value of the function is reflected through the center of gravity of other points:

$$x_c = \frac{1}{N} \sum_{i \neq j, i=0}^{N} x_i. \qquad (3)$$

The new point is used as the vertex of the new simplex. Iterations continue until either the minimum point is reached or cyclic motion of two or more simplexes begins.

Among the methods of deformation of the original simplex (which occurs after the rejection of the worst vertex and subsequent search for a new suitable vertex) are: vertex reflection (vertex is simply reflected on one side of the simplex), reduction (simplex retains its shape but decreases). However, this method, despite its simplicity, has certain disadvantages: there are difficulties associated with scaling the task (in real problems, different variables are often not comparable in value); the algorithm works slowly (the information of previous iterations is not used); there is no easy way to resize a simplex without recalculating all the values of the objective function.

The Hooke-Jeeves method is a combination of two types of search: the search study and the sample search. The first is focused on identifying the nature of the local behavior of the objective function and determining the directions along the "ravines". The size of the step is set, which can be different for different coordinate directions and change in the search process.

If the value of the target function at the test point does not exceed the value at the original, the search step is considered successful. Otherwise, you need to go back to the previous point and take a step in the opposite direction. After searching all N coordinates, the search ends. The resulting point is called the base.

The sample search is to implement a single step from the obtained base point along the line connecting it to the previous base point. And the new point is built on the formula:

$$x_b^{k+1} = x^k + \lambda \cdot (x^k - x^{k-1}), \qquad (4)$$

де x^k – поточна базова точка; x^{k-1} – попередня базова точка; x_b^{k+1} – точка, побудована при русі за зразком; λ – параметр алгоритму.

If the movement of the sample does not lead to a decrease in the objective function, the point x_b^{k+1} is fixed as a temporary base point and again the search is performed from this point. If the result is a point with the value of the function less than x^k, it is considered as a new base point x^{k+1}. If the researched search is unsuccessful, then there is a return to x^k and the search is performed in the opposite direction. If it also does not lead to success, the magnitude of the step is reduced and resumes the search. The search ends when the step size becomes small enough. The advantages of this method are a simple search strategy and a small amount of memory required. However, the algorithm is based on cyclic motion in coordinates, and this can lead to the degeneration of the algorithm into an infinite sequence. To prevent this, an iteration limiter is set, after which the algorithm stops.

These are not all zero-order methods, but we have considered the main ones.

Gradient methods. A group of methods whose iterative processes for solving unconditional optimization problems coincide with the antigradient of a function at each step are called gradient methods, or descent methods. They are also called first-order methods, or descent methods. These methods use both the values of the function and the values of the first-order partial derivatives, so they can be used to minimize the functions that are differentiated. First-order methods converge faster than direct search methods, as they consider the derivatives that characterize the direction of the most rapid decline of the function. Consider some of them.

The fastest descent algorithm implements an iterative procedure of moving to a minimum from an arbitrarily selected starting point in the direction of the strongest reduction of the function, determined in the vicinity of the current value of the argument of the function being minimized. This direction is opposite to the direction given by the gradient vector $grad\ f(x) = \nabla f(x)$ of the minimized function $f(x)$. The general formula for finding the value of the argument x^{k+1} by the value of x^k found in the k-th step of the fastest descent algorithm:

$$x^{k+1} = x^k + \lambda^k \cdot s^k,$$

where s^k is a vector of unit length in the direction opposite to the direction of the gradient $\nabla f(x^k)$, defined at the point x^k; λ^k is the step of the gradient procedure.

$$s^k = \frac{-\nabla f(x^k)}{|\nabla f(x^k)|}, \quad (5)$$

where $|\nabla f(x^k)|$ - norm of the gradient vector.

The fastest descent algorithms differ in the method of determining the step λ^k. If the step λ^k does not depend on k (is constant), then in the vicinity of the extremum there are inextinguishable oscillations, the amplitude of which depends on the value of λ and the shape of the function, which is minimized. Using a constant step allows you to build the simplest version of the algorithm; at large values of λ provides rapid movement to the extremum, but leads to noticeable changes on the outskirts of the extremum; at small values of λ leads to a low rate of convergence to the extreme; information about the acceptable step size λ is obtained during the debugging of the algorithm. If far from the extremum the function $f(x)$ has a small gradient, the rate of convergence may be unacceptably slow. This problem is solved by modifying the algorithm.

The method of combined gradients of Fletcher and Reeves, the direction of descent in which deviates from the direction of the antigradient by adding to it the vector of the direction used in the previous step, multiplied by some positive number. The search directions on each iteration are determined by the formula:

$$s^k = -\nabla f(x^k) + \sum_{i=0}^{k-1} \alpha^i \cdot s^i. \quad (6)$$

Parameter values are selected so that the direction is associated with all previously constructed search directions. This is possible when the following condition is met:

$$\alpha_{k-1} = \frac{\left(\nabla f(X_k)\right)^2}{\left(\nabla f(X_{k-1})\right)^2}. \quad (7)$$

Practical studies [13] have shown that this method converges faster than the method of the fastest descent, and its effectiveness increases in the final stages of finding the minimum function. In addition, it should be noted that this method can be used to minimize functions with discontinuous derivatives. The search "does not hang on the fracture» but goes along the line connecting the breakpoints of the level lines, which usually passes through the minimum point.

Gradient methods are quite effective, but at the initial stage of minimization. In the next stages, when the search points are near the minimum point, it is necessary to use methods that have a higher rate of convergence. These methods are second-order methods, which include Newton's method and related quasi-Newtonian methods.

Newton's method is based on the quadratic approximation of a function that is minimized about the point x^k. The minimum of a quadratic function is easy to find by equating its gradient to zero. You can immediately calculate the position of the extremum and select it as the next approximation to the minimum point. Calculating the point of the new approximation by the formula:

$x^{k+1} = x^k + \Delta x^k$, and decomposing $f(x^{k+1})$ into a Taylor series, we obtain a quadratic approximation.

$$f_{sq}(x^{k+1}) = f(x^k) + \\ +\left(\nabla f(x^k)\right)^T x^k + \\ +\frac{1}{2!}(x^k)^T \nabla^2 f(x^k) x^k. \quad (8)$$

Under the condition of a minimum, look for the length of the step x^k:

$$x^k = -[\![\nabla^2 f(x^k)]\!]^{-1} \cdot \nabla f(x^k). \quad (9)$$

Advantages of Newton's method:

- if the minimized function is quadratic, the method will allow to find at least one step;
- if the function belongs to the class of surfaces of rotation (i.e. has symmetry), the method also provides convergence in one step;
- if the function is asymmetric, then the method does not provide convergence for a finite number of steps. But for many functions, a much higher rate of convergence is achieved than with other modifications of the fastest descent method.

The disadvantages of Newton's method are related to the need for calculations and inverse matrices of the second derivatives. This not only wastes machine time, but as significant computational errors can also occur if the matrix $\nabla^2 f(x^k)$is poorly determined.

The Davidon-Fletcher-Powell method, also called the variable metric method, falls into the general class of quasi-Newtonian procedures in which the search directions on the k-th iteration are given as: $s^k = -h^k \nabla f(x^k)$. The direction of the gradient deviates due to multiplication by h^k, which is a positively defined symmetric matrix of order $n \times n$, approximating the inverse Hesse matrix. At each step, the matrix is updated, ie takes the form: $h^{k+1} = h^k + a^k + b^k$.

$$a^k = \frac{\Delta x^k \cdot (\Delta x^k)^T}{(\Delta^k)^T \cdot \Delta g^k}, \qquad \Delta x^k = x^k + x^{k+1},$$
$$b^k = -\frac{h^k \cdot (\Delta^k)^T \cdot \Delta g^k \cdot h^k}{\Delta x^k \cdot h^k \cdot (\Delta x^k)^T}, \qquad \Delta g^k = \nabla f(x^{k+1}) - \nabla f(x^k). \quad (10)$$

The Davidson-Fletcher-Powell algorithm can sometimes lead to a situation where the matrix becomes poorly conditioned, or the condition of positive certainty is violated. The reason for this is the poor choice of the initial approximation, as well as the presence of rounding errors. To overcome these difficulties, it is necessary to increase the accuracy of calculations and periodically update the iterative process.

The Davidon-Fletcher-Powell method is widely used to solve a variety of problems and is high.

Thus, although there is no one-size-fits-all method that allows you to successfully solve all problems, some methods are better suited to solve certain types of problems. Careful selection of the appropriate algorithm often saves both machine time and effort spent by the engineer to solve the problem. The section provides recommendations on which algorithms to prefer, for which direct search methods, gradient methods and second-order methods are considered.

When creating and building decision-making systems there is a need to solve a variety of optimization problems. The considered methods of multidimensional optimization as methods of solving the general problem of finding the local extremum, and the comparisons of their efficiency did not give an unambiguous answer to the question of when and which method should be used. The result of the analysis of the applicability of the considered research algorithms led to a deeper numerical experiment aimed at identifying the fastest algorithm with maximum plausibility.

Numerical experiments (practical implementation). For practical use, a program in the Delphi programming language was developed (Fig. 2).

The developed interface of the program allows you to quickly enter the target function and search for a solution in one of 15 methods. The result of such a search contains the time spent on the calculation and the numerical values of the intermediate values at each step (iteration) of the calculation (Table 1).

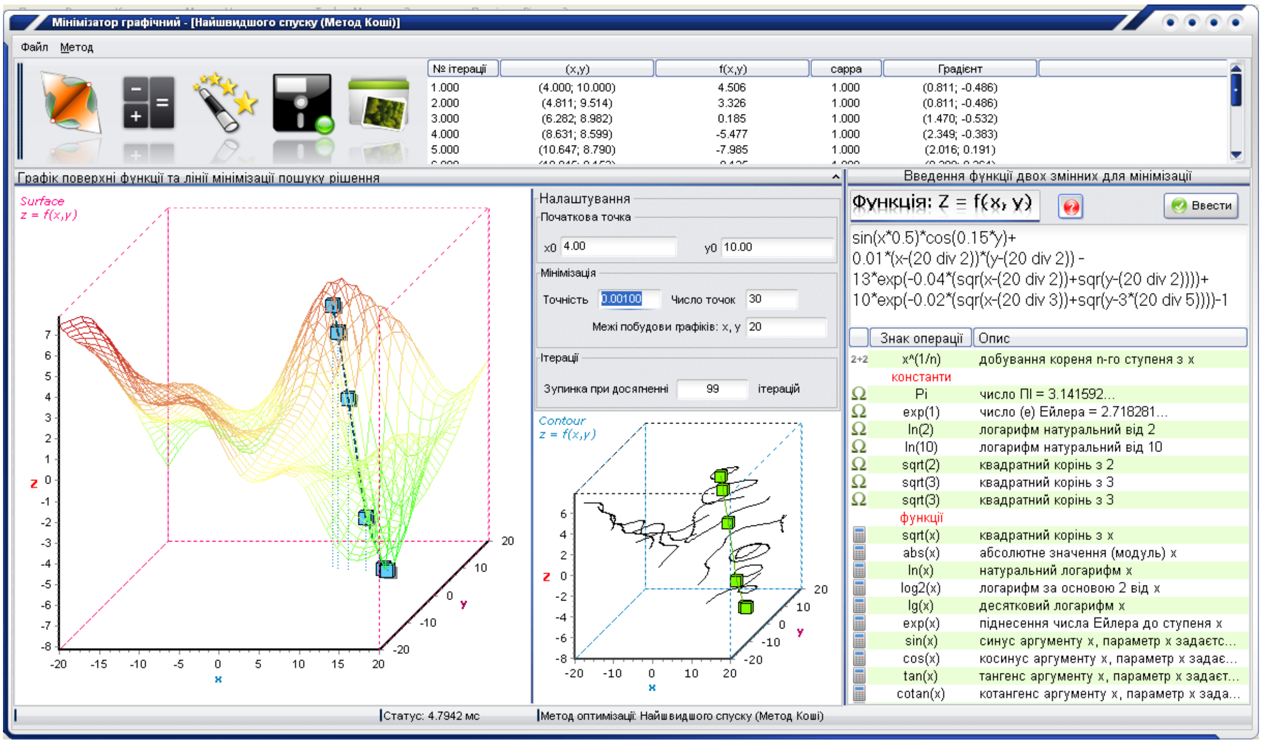


Figure 2. View of the main window of the *minGraph* program (graphic minimizer).

Table 1. The result of generating intermediate values of the calculation program minGraph.

$Z(x) = 8 \cdot x^2 - 4 \cdot x \cdot y + 5 \cdot y^2 + 8 \cdot \sqrt{5} \cdot (x + 2 \cdot y) + 64$
Method: Coordinate descent (golden ratio) - Gauss - Seidel method
Calculation time: 12.2292 ms

Itep.	(x, y)	Z = *f(x, y)*	scalar argument*	gradient
1	(-15.000; 10.000)	3053.443	0.000	(-15.000; 10.000)
2	(-4.868; -5.525)	13.869	1.000	(-4.868; -5.525)
3	(-2.499; -4.578)	-35.501	1.000	(-2.499; -4.578)
4	(-2.263; -4.482)	-35.995	1.000	(-2.263; -4.482)
5	(-2.239; -4.473)	-36.000	1.000	(-2.239; -4.473)
6	(-2.237; -4.473)	-36.000	1.000	(-2.237; -4.473)
7	(-2.236; -4.472)	-36.000	1.000	(-2.236; -4.472)
8	(-2.236; -4.472)	-36.000	1.000	(-2.236; -4.472)

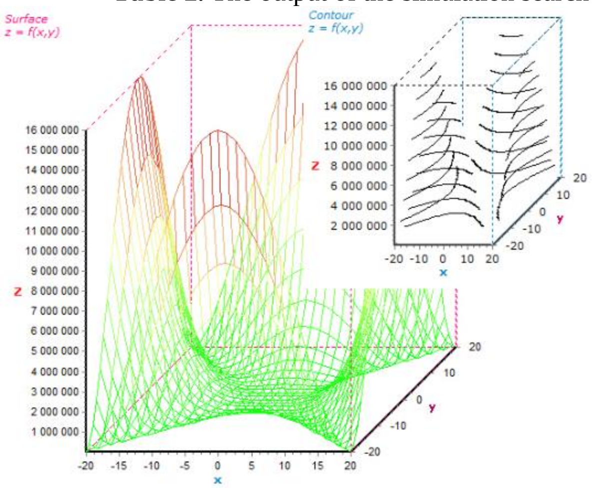
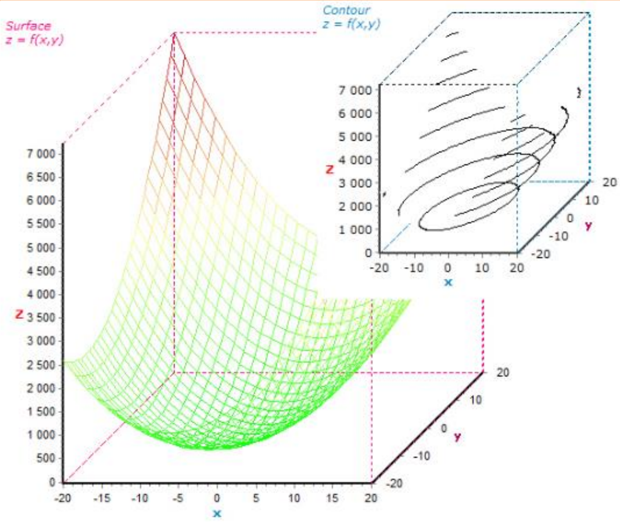
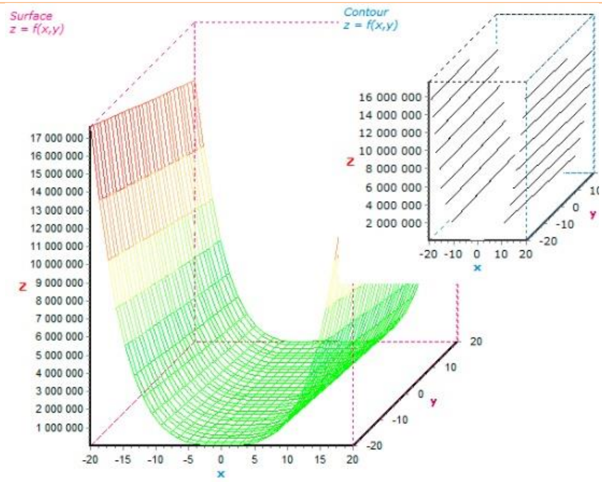
Note * is a scalar argument k of a scalar function of the form: $\Psi(k) = f\left(\overrightarrow{x^k} + k \cdot \overrightarrow{y^k}\right)$.

The method of numerical experiment was as follows: for three target functions under certain initial conditions (Table 2) the speed and accuracy of ascent were studied using 15 minimum search methods: 1 – coordinate descent (golden ratio) – Gauss-Seidel method; 2 – random search; 3 – gradient descent with step crushing; 4 – the fastest descent (Cauchy method); 5 – related areas; 6 – Fletcher-Reeves; 7 – DFP (Davidon-

Fletcher-Powell); 8 – cyclic coordinate descent; 9 – zero order – Hook-Jeeves; 10 – zero order – Rosenbrock; 11 – zero order – Powell; 12 – irregular simplex – Nelder-Mead; 13 – Newton's method; 14 – conditional gradient; 15 – projections of the antigradient. The following parameters were recorded: the number of iterations, the time spent searching for a solution, the found value of the function and the subsequent comparison with the exact value by finding the absolute error (Table 3).

The finding the minimum of functions by iterative methods, one of the decisive factors should be considered the stability of the results. This means that small deviations from the initial values of the desired functions should not lead to a significant change in the final result. For clarity, the type of surfaces and the computational process of minimization for two variables is shown in Fig. 3.

Table 2. The output of the simulation search for the best algorithm for the three functions.

 	***Function 1 -:f_1*** $\boldsymbol{f(x,y) = 100 \cdot (y^2 - x^2)^2 + (1 - x)^2}$ Initial data: Starting point -$(x0 = -19.00; y0 = 0.00)$ Accuracy -$e = 0.001$ Number of points - 30 The boundaries of the graph - x, y = 20
 	***Function 2 -:f_2*** $f(x,y) = 8x^2 - 4xy + 5y^2 + 8\sqrt{5} \cdot (x + 2y) + 64$ Initial data: Starting point -$(x0 = -15.00; y0 = 10.00)$ Accuracy -$e = 0.001$ Number of points - 30 The boundaries of the graph - x, y = 20
 	***Function 2 -:f_3*** $f(x,y) = 100 \cdot (y - x^2)^2 + (y - 1)^2$ Initial data: Starting point -$(x0 = -19.00; y0 = 0.00)$ Accuracy -$e = 0.001$ Number of points - 30 The boundaries of the graph - x, y = 20

In this computational experiment, deliberately complex operating conditions were used. To do this, the starting points were chosen as far as possible from the probable minimum.

Table 3. Estimated minimum values of functions for different algorithms.

Method	Found value	Iteration	Time, ms	Δ
***function* – f_1 ($min = 0$)**				
1.	323,997	3	6,4131	-323,997
2.	0,059	74	116,97	-0,059
3.	3592310	2	4,2522	-3592310
4.	0,771	5	6,4882	-0,771
5.	0,771	4	5,5555	-0,771
6.	0,771	8	21,8611	-0,771
7.	0,557	98	159,5394	-0,557
8.	99,994	3	202,1436	-99,994
9.	110,249	4	7,2663	-110,249
10.	0	5	6,6056	0
11.	0	5	6,6534	0
12.	0,001	2	4,4209	-0,001
13.	0,771	16	29,3393	-0,771
14.	0,049	143	223,6698	-0,049
15.	4	308	480,2124	-4
***function* – f_2 ($min = -36$)**				
1.	-36	8	15,32	0
2.	-35,745	79	142,667	-0,255
3.	-36	29	58,9249	0
4.	-36	9	22,9714	0
5.	-36	3	8,9138	0
6.	-36	3	5,9635	0
7.	-36	3	4,7089	0
8.	-36	8	214,511	0
9.	-36	12	20,3669	0
10.	-36	7	24,5134	0
11.	-36	3	11,0239	0
12.	-36	2	4,3565	0
13.	-36	3	4,8218	0
14.	-20,22	3	4,9149	-15,78
15.	-34,997	7	25,2023	-1,003
***function* – f_3 ($min = 0$)**				
1.	0,15	99	180,254	-0,15
2.	4,977	74	134,857	-4,977
3.	3586954	2	14,4432	-3586954
4.	0,361	98	160,222	-0,361
5.	0,478	49	104,287	-0,478
6.	43,643	99	162,982	-43,643
7.	0	92	156,589	0
8.	0	987	7604	0
9.	0,044	480	1867,2	-0,044
10.	0	135	248,126	0
11.	0	6	30,9801	0
12.	0	2	8,6935	0
13.	0	22	43,0731	0
14.	0	17	39,3458	0
15.	0	371	570,204	0

After finding the optimum of the functions (Fig. 3), the results of the calculation were compared by methods for three functions on the rate of ascent, the accuracy of ascent.

A computational experiment showed that the first function turned out to be quite complex for almost all algorithms (minimum at: $x = 1; y = 1$). Methods 7, 8 and 14 and 15 showed the worst rate of ascent. These methods are also unsatisfactory. The optimal solution was not found, and the values of approaching the minimum ranged from 0.049… 99.994. Methods 10, 11 and 11 proved to be the most effective. The accuracy of these methods was 100%. It is inadmissible to use to find the minimum of this function 3 method (gradient descent with crushing step). Of course, if we take into account the value of the function at the starting point (13032500), the value found is 0.049…0.771 (this is almost the minimum, with some assumptions can be considered a sufficient solution by methods 2,4,5,6,7, 13 and 14).

The second function (minimum = -36) turned out to be easier to find a solution. This is due to the more pronounced peak of the minimum (table 2). Only three methods 2, 14 and 15 failed the task. All the others showed high accuracy. Regarding the rate of ascent, 2 and 3 methods performed work for 79 and 29 iterations, respectively. The running time of the algorithm depended not so much on the number of iterations as on the complexity of the algorithm.

The dependences of the rate of convergence of the three functions for 15 algorithms (Fig. 4) showed the best values for 11, 12 and partially 3 methods. However, given the current speed of computers, the number of iterations close to 100 is not so important, although the time difference (Fig. 5) is more striking (8 and 9 methods - 2000, 8000 ms). More decisive for the methods is the accuracy of finding the minimum.

The third experimental function in finding the minimum showed the excellent work of methods 7-15. And this is due to the complexity of Rosen Brock's function. This function is a classic optimization problem, also known as the banana function. It has a large slowly descending plateau. Finding a plateau is a trivial task, but convergence to the global optimum is difficult. This function is used to evaluate the operation of optimization algorithms, and therefore was chosen by us for research. All other methods stumbled on this function and showed questionable results, especially method 3. Of particular interest in the work of 8 and 9 methods (cyclic coordinate descent and zero order - Hook-Jeeves). There is a rather poor ascent and a lot of time for the algorithm. However, the result was excellent for method 8 and sufficient for method 9.

Computational experiments to minimize three functions (Table 2) for 15 optimization methods showed that two methods can be recommended for all studied functions: 10 – zero order – Rosen Brock and 11 – zero order – Powell (Fig. 3).

This is since these algorithms immediately move to the local minimum and have the widest range of gravity, so it is more likely that at least one of the iterations falls into the specified area and has the maximum impact on further behavior of the search algorithm.

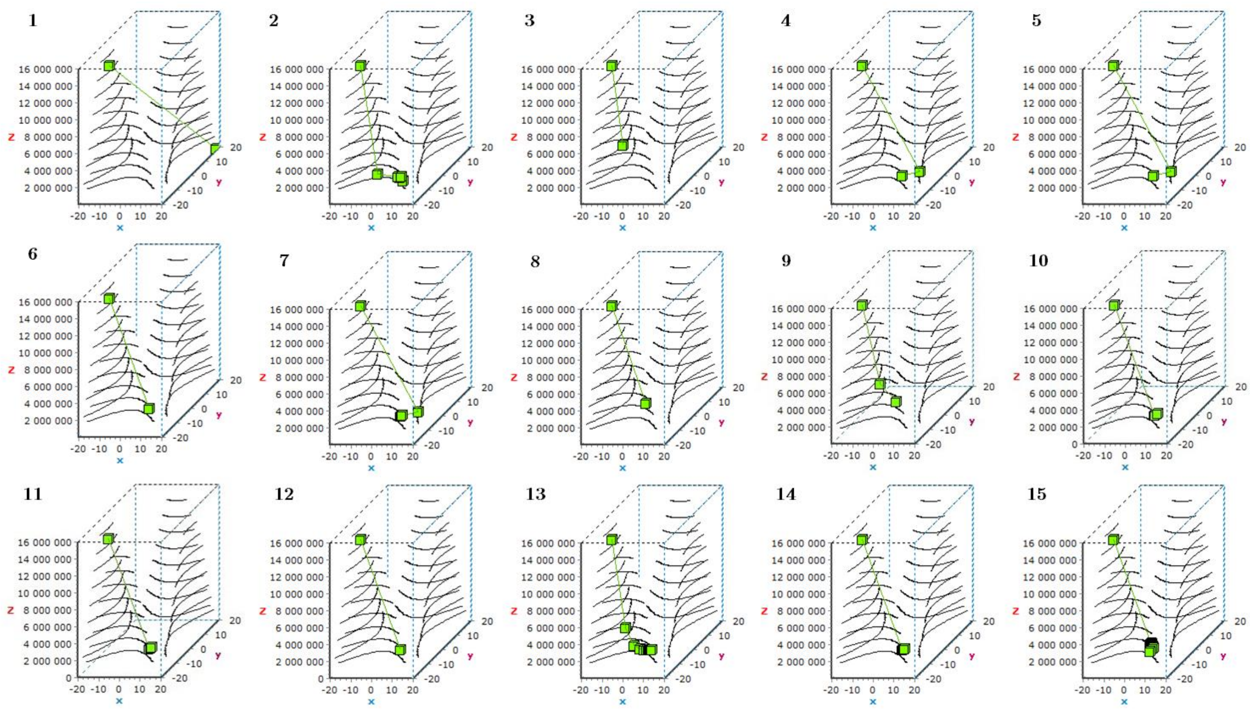


Figure 3. A set of minimum search images for a function $\boldsymbol{f_1}$ by different methods.

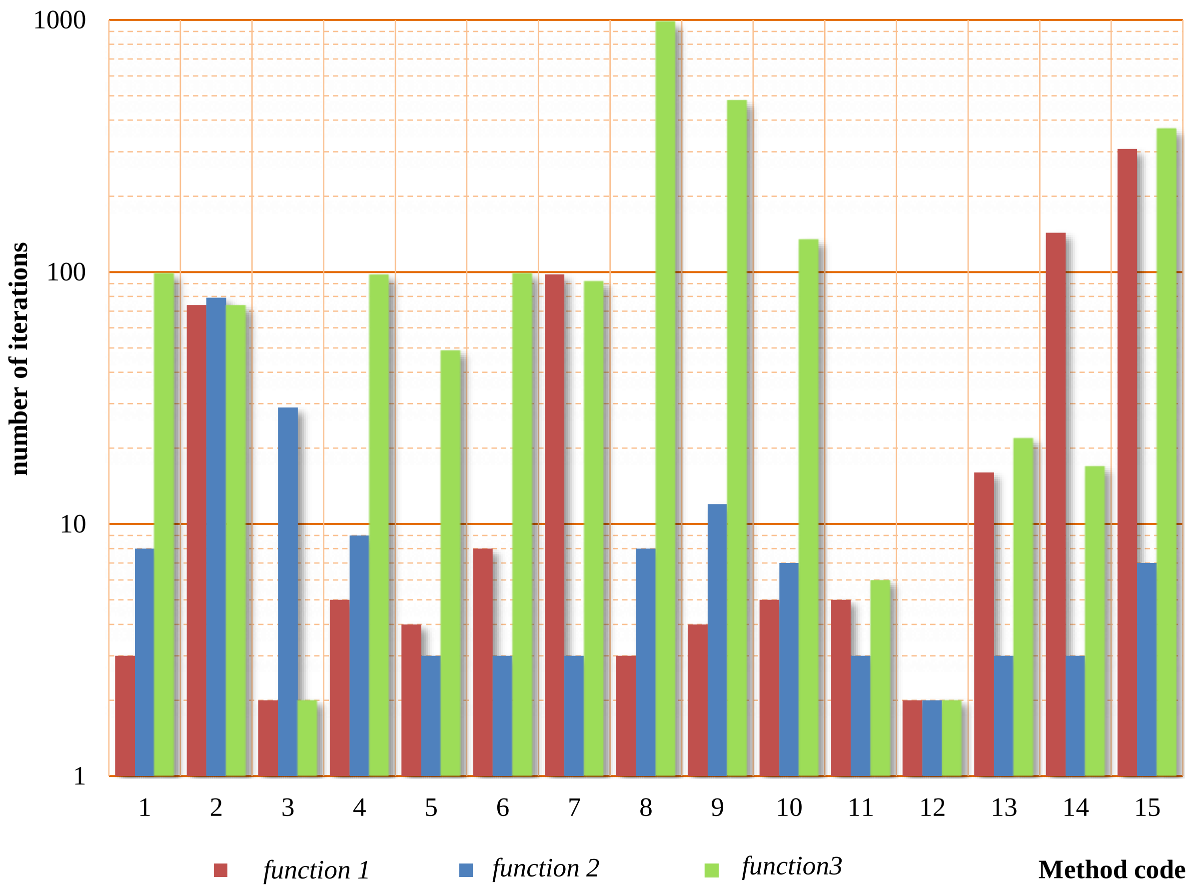


Figure 4. Histograms of convergence of experimental functions for 15 methods.

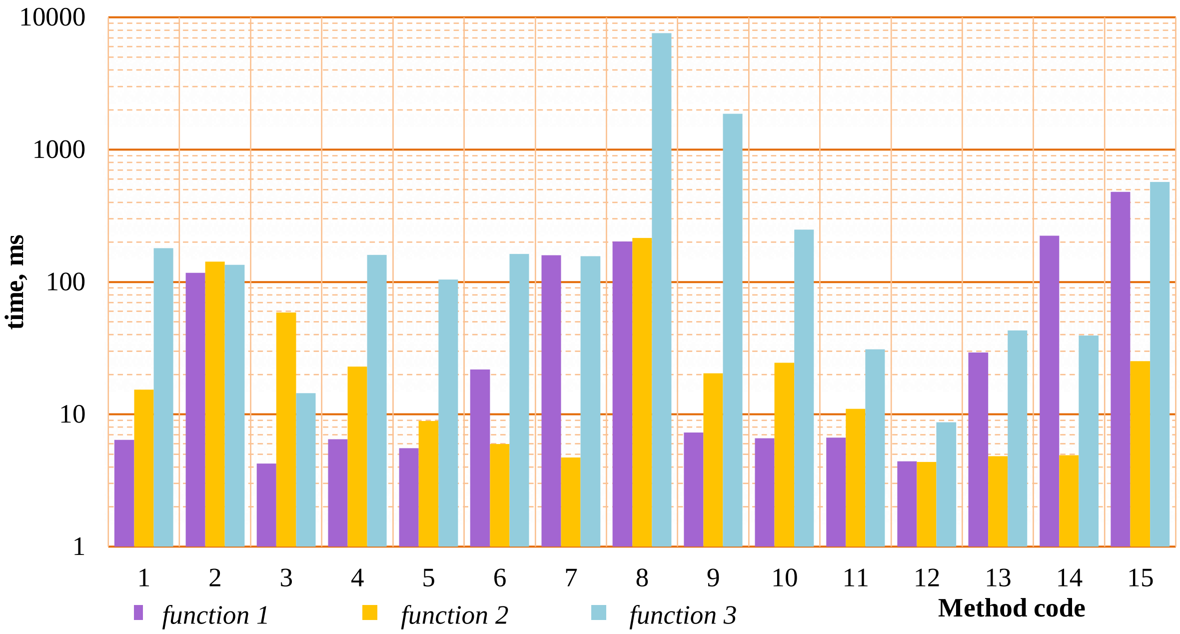


Figure 5. Histograms of convergence of experimental functions for 15 methods.

Only these methods showed the best convergence and calculation accuracy for all functions. Other methods have partially shown excellent results for some functions, while for other functions not at all.

It should be noted that it is advisable to use these methods on more complex functions, which in addition to the global minimum contain several local minima, which may be of interest in solving applied engineering problems. The use of the developed tool will speed up the analysis of engineering functions, and in general will increase the stability of the results.

Conclusions

In the course of this work the main optimization methods used in modern engineering problems and methods of learning neural networks were considered. In the course of the research the analysis of properties and features of the considered methods was carried out, and also conditions and the substantiation of their most optimum application in various practical tasks from the point of view of their convergence and accuracy were formulated.

The analysis data are accompanied by practical results, which confirm the formulated recommendations for the use of the considered methods of classical and stochastic gradient descent, pulse, adaptive, and quasi-Newtonian optimization algorithms in engineering problems.

Areas of further research. The considered multidimensional optimization methods can be used in the future for training neural networks.

References

Astashev V., Krupenin V. 2017. Efficiency of vibration machines. *Proceedings of 16th International Scientific Conference "Engineering for rural development"*. Jelgava, Latvia, May 24-26, Latvia University of Agriculture. Faculty of Engineering. Vol. 16. P. 108-113.

Drga R., Janacova D., Charvatova H. 2016. Simulation of the PIR detector active function. *Proceedings of 20th International conference on Circuits, Systems, Communications and Computers* (CSCC 2016), July 14-17, 2016, E D P Sciences, 17 Ave Du Hoggar Parc D Activites Coutaboeuf Bp 112, F-91944 Cedex A, France, 76, UNSP 04036.

Dubbini M., Pezzuolo A., De Giglio M., Gattelli M., Curzio L., Covi D., Yezekyan T., Marinello F. 2017. Last generation instrument for agriculture multispectral data collection. *CIGR Journal*. Vol. 19. P. 158-163.

Kuzmich I. M., Rogovskii I. L., Titova L. L., Nadtochiy O. V. 2021. Research of passage capacity of combine harvesters depending on agrobiological state of bread mass. *IOP Conference Series: Earth and Environmental Science*. Vol. 677. P. 052002. https://doi.org/10.1088/1755-1315/677/5/052002.

Luo A. C. J., Guo Y. 2013. Vibro-impact Dynamics. Berlin: Springer-Verlag: 213 p.

Masek J., Novak P., Jasinskas A. 2017. Evaluation of combine harvester operation costs in different working conditions. *Proceedings of 16th International Scientific Conference "Engineering for rural development"*. Jelgava, Latvia, May 24-26, Latvia University of Agriculture. Faculty of Engineering. Vol. 16. P. 1180-1185.

Nazarenko I., Mishchuk Y., Mishchuk D., Ruchynskyi M., Rogovskii I., Mikhailova L., Titova L., Berezovyi M., Shatrov R. 2021. Determiantion of energy characteristics of material destruction in the crushing chamber of the vibration crusher. *Eastern-European Journal of Enterprise Technologies*. Vol. 4(7(112). P. 41–49. https://doi.org/10.15587/1729-4061.2021. 239292.

Novotny J. 2016. Technical and natural sciences teaching at engineering faculty of FPTM UJEP. *Proceedings of 15th International Scientific Conference "Engineering for rural development"*. Jelgava, Latvia, May 23-25, Latvia University of Agriculture. Faculty of Engineering. Vol. 15. P. 16-20.

Palamarchuk I., Rogogvskii I., Titova L., Omelyanov O. 2021. Experimental evaluation of energy parameters of volumetric vibroseparation of bulk feed from grain. *Engineering for Rural Development*. Vol. 20. P. 1761-1767. https://doi.org/10.22616/ ERDev.2021.20.TF386.

Pinzi S., Cubero-Atienza A. J., Dorado M. P. 2016. Vibro-acoustic analysis procedures for the evaluation of the sound insulation characteristics of agricultural machinery. *Journal of Sound and Vibration*. Vol. 266 (3). P. 407-441.

Rogovskii I. L., Titova L. L. 2021a. Change of technical condition and productivity of grain harvesters depending on term of operation. *IOP Conference Series: Earth and Environmental Science*. Vol. 720. P. 012110. https://doi.org/10.1088/1755-1315/720/1/012110.

Rogovskii I. L., Titova L. L. 2021b. Design of landing of assembly machine building units with circulating load rolling bearing rings. *Journal of Physics: Conference Series*. Vol. 1889. P. 042004. https://doi.org/10.1088/1742-6596/1889/4/042004.

Rogovskii I. L., Titova L. L. 2021c. Modeling of normativity of criteria of technical level of forage harvesters combines. *IOP Conference Series: Earth and Environmental Science*. Vol. 720. P. 012109. https://doi.org/ 10.1088/1755-1315/720/1/012109.

Rogovskii I. L., Titova L. L. 2021d. Modeling the weight of criteria for determining the technical level of agricultural machines. *IOP Conference Series: Earth and Environmental Science*. Vol. 677. P. 022100. https://doi.org/ 10.1088/1755-1315/677/2/022100.

Rogovskii I. L., Titova L. L., Berezova L. V. 2021e. Conceptual bases of system technology of designing of logistic schemes of harvesting and transportation of grain crops. *IOP Conference Series: Earth and Environmental Science*. Vol. 723. P. 032032. https://doi.org/10.1088/1755-1315/723/3/032032.

Rogovskii I. L., Titova L. L., Gumenyuk Yu. O., Nadtochiy O. V. 2021f. Technological effectiveness of formation of planting furrow by working body of passive type of orchard planting machine. *IOP Conference Series: Earth and Environmental Science*. Vol. 839. P. 052055. https://doi.org/10.1088/1755-1315/839/5/052055.

Rogovskii I. L., Titova L. L., Ohiienko M. M. 2021g. Research of garden sprayer machines of near-stem and inter-stem strips of orchards. *IOP Conference Series: Earth and Environmental Science*. Vol. 723. P. 022035. https://doi.org/10.1088/1755-1315/723/2/022035.

Rogovskii I. L., Titova L. L., Sivak I. M. 2021h. Research of machines for mulching near-trunk strips in perennial fruit plantations. *IOP Conference Series: Earth and Environmental Science*. Vol. 723. P. 042041. https://doi.org/ 10.1088/1755-1315/723/4/042041.

Rogovskii I. L., Titova L. L., Solomka O. V. 2021j. Research of sliding bearings with reverse friction pair and inlaid liners made of thermoplastic composite materials. *Journal of Physics: Conference Series*. Vol. 1889. P. 042010. https://doi.org/10.1088/1742-6596/1889/4/04 2010.

Rogovskii I. L. 2020. Model of stochastic process of restoration of working capacity of agricultural machine in inertial systems with delay. *Machinery & Energetics. Journal of Rural Production Research*. Kyiv. Ukraine. Vol. 11(3). P. 143-150.

Rogovskii I., Titova L., Novitskii A., Rebenko V. 2019. Research of vibroacoustic diagnostics of fuel system of engines of combine harvesters. *Engineering for Rural Development*. Vol. 18. P. 291-298.

Sergejeva N., Aboltins A., Strupule L., Aboltina B. 2018. Mathematical knowledge in elementary school and for future engineers. *Proceedings of 17th International Scientific Conference "Engineering for rural development"*. Jelgava, Latvia, May 23-25, 2018, Latvia University of Agriculture. Faculty of Engineering. Vol. 17. P. 1166-1172.

Titova L. L. 2021. Simulation of transitional processes of dynamic model of power energy installation of machines for forestry works. *Machinery & Energetics. Journal of Rural Production Research*. Kyiv. Ukraine. Vol. 12(1). P. 147-154.

Viba J., Lavendelis E. 2006. Algorithm of synthesis of strongly non-linear mechanical systems. *Industrial Engineering – Innovation as Competitive Edge for SME*, 22 April 2006. Tallinn, Estonia. P. 95-98.

Yata V. K., Tiwari B. C., Ahmad I. 2018. Nanoscience in food and agriculture: research, industries and patents. *Environmental Chemistry Letters*. Vol. 16. P. 79-84.

Zagurskiy O., Ohiienko M., Rogach S., Pokusa T., Titova L., Rogovskii I. 2018. Global supply chain in context of new model of economic growth. Conceptual bases and trends for development of social-economic processes. Monograph. Opole. Poland. P. 64-74.