

Multifactor Mathematical Model of Geometry of the Rotor Impeller of Artificial Heart

Katalazhnova I.N.*

Department of Higher Mathematics, Komsomolsk-on-Amur State University, Komsomolsk-on-Amur, Russia

*Corresponding author. Email: 777irnik@mail.ru

ABSTRACT

The problem of creating auxiliary circulatory systems (ACS) for extreme medicine is crucial. In areas of natural disaster or in conditions of hostilities, it is sometimes necessary to replace the heart and organize ancillary circulation. The paper analyzes the optimization of energy parameters of small-sized centrifugal pumps used in auxiliary blood supply systems. The calculation and analytical base for designing miniature rotary pumps is borrowed from the field of designing industrial centrifugal pumps. Calculation methods are replete with a large number of empirical coefficients that do not have physical justification and do not ensure qualitative results at the design stage of auxiliary circulatory systems.

Keywords: auxiliary blood supply systems, small centrifugal pump, statistical processing of experimental data

1. INTRODUCTION

The problem of creating auxiliary circulatory systems (ACS) for extreme medicine is crucial.

In areas of natural disaster or in conditions of hostilities, it is sometimes necessary to replace the heart and organize ancillary circulation. Figure 1 shows the diagram of connection of the hydraulic pump implanted in the abdominal cavity below the diaphragm and connected to the cardiovascular system in such a way as to perform the functions of the left ventricle. The probability of heart failures can be reduced.

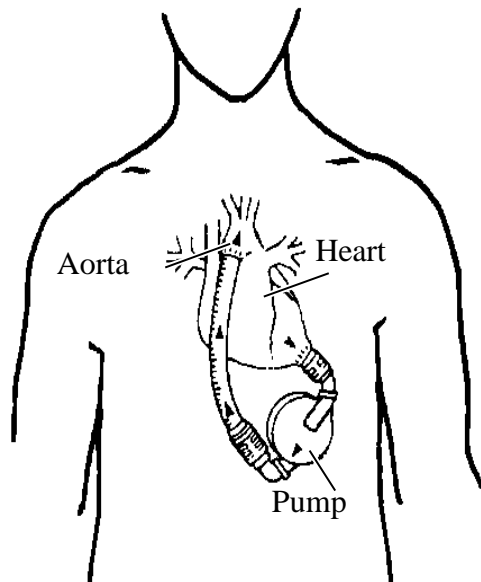


Figure 1 Scheme of connection of the hydraulic pump to the human cardiovascular system

There are several types of hydraulic pumps which can be used as an artificial heart. The best weight, size and performance parameters have miniature rotary pumps. "... this pump is the size of a conventional Duracell battery, its capacity is 5–6 liters of blood per minute. There is only one moving part – an impeller which makes 10 thousand revolutions per minute [1].

The existing calculation methods are based on a two-dimensional potential flow model that does not take into account the spatial nature of the flow in the miniature impeller. As a result, the design results are ineffective in terms of the energy transfer whose optimization is required.

From a hydrodynamic point of view, the miniaturization of the impellers is accompanied by an increase in the relative thickness of the boundary layer in the interscapular canals. As a result, the deflecting properties of the impeller are reduced, the pressure quality of the artificial heart of the rotor type is deteriorated, the radial dimensions of the structure are increased. One of the main advantages of using such hearts in auxiliary circulatory systems is minimum dimensions and weight.

This article is devoted to the development of a regression multifactorial mathematical model of the rotor impeller of an artificial heart.

2. CHARACTERISTICS OF TARGET FUNCTIONS OF THE MODEL

The model of the impeller geometry allows us to transform the flow part and increase the energy efficiency of the ACS pump.

Mathematical modeling was carried out according to an algorithm in which the results of hydraulic tests of the AICS pump were subjected to the correlation analysis. Variables that have the greatest impact on the response function were selected. The composition of the target

functions of 3 energy parameters of the centrifugal pump of the ACS and 6 variables in the form of geometric parameters of the impeller are presented in Table 1.

Hydraulic tests were carried out using different versions of the impellers for a given number of revolutions and a change in the flow rate of the working fluid.

Table 1 Designations of target functions and variables

Statistical Analysis Parameters	
Designation of target functions and variables	Composition of target functions and variables
Target function Y	1) Pressure ratio \bar{H} ; 2) hydraulic parameter $k_z \eta_r$; 3) Pump efficiency η
Variable X_j	1) relative impeller diameter (ratio of inlet and outlet diameters) $\bar{D}_1 = D_1/D_2$; 2) geometric degree of diffusivity of the impeller channels (the ratio of areas of the impeller channels at the inlet and outlet) $\bar{F} = F_2/F_1$; 3) the relative width of the channels of the impeller at the exit (the ratio of the width of the channels to the diameter at the exit of the impeller) $\bar{b}_2 = b_2/D_2$; 4) relative width (the ratio of the width of the channels at the exit of the impeller to the length of the blades) b_2/L ; 5) outlet blade angle β_{2n} , degrees; 6) speed factor n_s

3. STATISTICAL PROCESSING ALGORITHM

The algorithm for statistical processing of experimental data includes seven stages.

Stage I “Stratification of the initial experimental data”. At this stage, before the statistical processing of the results, gross systematic and random errors made during the experiments are identified and eliminated [2–4].

The target design parameter is the mean square error $\bar{\sigma}(y)$ of the response function values.

Under an unknown value of $\bar{\sigma}(y)$ its assessment is made approximately using the ratio:

$$\bar{\sigma}(y) = \left[\frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y}_j)^2 \right]^{1/2},$$

where y_i – measurements of energy parameters: pressure and efficiency η of centrifugal pumps obtained during the experiment; \bar{Y}_j – average value of energy parameters: pressure and efficiency η of the centrifugal pump.

The absolute difference $|\tilde{y} - \bar{Y}_j|$ between the value \tilde{y} and the mean value \bar{Y}_j other acceptable results is divided by the mean square error $\bar{\sigma}(y)$. The determined value

$z = \frac{|\tilde{y} - \bar{Y}_j|}{\bar{\sigma}(y)}$ is compared with $t_N(P)$, the set level of the value $\gamma = 0.05$ from the ratio:

$$P \left[\frac{|\tilde{y} - \bar{Y}_j|}{\bar{\sigma}(y)} \leq z_N(P) \right] = 1 - \gamma$$

where $z_N(P)$ – P-quantile for normal distribution $t(f)$ c $f = N$; N – the number of acceptable measurements; $z_N(P) = t_{1-\gamma/2}$ ($f=N$); $P = 1 - \gamma/2$, critical values $z_N(P)$ are presented in statistical tables.

If, for a given number of acceptable results N , the inequality: $\frac{|\tilde{y} - \bar{Y}_j|}{\bar{\sigma}(y)} > z_N(P)$, we can assume that a gross

error lies at the basis of the “popping up” result. This result is excluded from further statistical data processing.

At the final stage, after eliminating the gross errors, the normalization of variables is carried out using the equations:

$$y_i^\circ = \frac{y_{ji} - \bar{Y}_j}{\sigma_y}, \quad x_{ji}^\circ = \frac{x_{ji} - \bar{x}_j}{\sigma_{x_j}},$$

where $y_{ji}^\circ, x_{ji}^\circ$ – normalized experimental data on the response function and variables;

σ_Y, σ_{x_j} – standard deviations of the response function and variables. To determine the average values of the response function and variables, the expressions are used:

$$\bar{Y}_j = \frac{\sum_{i=1}^N y_{ji}}{N}, \quad \bar{x}_j = \frac{\sum_{i=1}^N x_{ji}}{N}.$$

The standard deviations are calculated by formulas:

$$\sigma_{Y_j} = \sqrt{\frac{\sum_{i=1}^N (y_{ji} - \bar{Y}_j)^2}{N-1}}, \quad \sigma_{x_j} = \sqrt{\frac{\sum_{i=1}^N (x_{ji} - \bar{x}_j)^2}{N-1}}.$$

Stage 2. "Correlation analysis." The correlation analysis allows for a qualitative assessment of the dependence or independence of energy parameters on the input geometric and operational parameters of the centrifugal pump. The energy parameters of the centrifugal pump are pressure and efficiency η ; the geometric and operating parameters are the following variables: relative diameter, diffusivity and relative channel width of the impeller at the outlet, $b_2, b_2/L$, working fluid flow rate, speed coefficient n_s .

Sample correlation coefficients between the output and input parameters are calculated; their significance is assessed; the one-factor regression model $\hat{Y}(\alpha_0, \alpha_1, x_1) = \alpha_0 + \alpha_1 x_1$ is built.

The one-factor model is expanded by selecting a new variable according to the highest value of the coefficient of partial correlation between the remainder of the regression and variables not included in the regression equation.

Selected correlation coefficients are calculated by formulas:

$$corr(Y_j^\circ, x_j^\circ) = \frac{\sum_{i=1}^N (x_{ij}^\circ - \bar{x}_j^\circ) \cdot (y_{ji}^\circ - \bar{Y}_j^\circ)}{\sqrt{\sum_{i=1}^N (x_{ij}^\circ - \bar{x}_j^\circ)^2 \cdot \sum_{i=1}^N (y_{ji}^\circ - \bar{Y}_j^\circ)^2}}.$$

The most correlated variable X_{k_1} with the response function is selected by selecting the element with the maximum absolute value and the two-factor regression model is built:

$$\hat{Y}(\alpha_0, \alpha_1, \alpha_2) = \alpha_0 + \sum_{j=1}^2 \alpha_j x_j$$

The coefficients of the regression equation α_j are determined by the least squares method:

$$\sum_{i=1}^N (y_{ji}^\circ - \bar{Y}_j^\circ)^2 \rightarrow \min$$

Having built the two-factor model, we proceed to stage 3, assessing the adequacy of the resulting model.

Stage III. "Variance analysis." It is designed for a comprehensive two-level assessment: the adequacy of the model is determined using the Fisher F-test and the significance of the variable based on the sequential F – Fisher criterion.

The adequacy criterion $Y_{k_1 \dots k_m}$ is the inequality for the two-factor model $m=2$:

$$F_{k_1 \dots k_m} > (F_{k_1 \dots k_m})_{ma\bar{\sigma}},$$

where $F_{k_1 \dots k_m}$ – estimated value of the criterion, a $(F_{k_1 \dots k_m})_{ma\bar{\sigma}}$ – tabular value of the F – Fisher test.

The table value of the Fisher test at a significance level of $\gamma = 0.05$ is calculated by formula:

$$(F_{k_1 \dots k_m})_{ma\bar{\sigma}} = F_{1-\gamma}(p(k_1 \dots k_m), N - p(k_1 \dots k_m) - 1)$$

To determine $F_{k_1 \dots k_m}$ it is necessary to calculate:

- residual variation of the impact of the response function of unaccounted factors – $RSS_{k_1 \dots k_m}$ by formula:

$$RSS_{k_1 \dots k_m} = (Y_j^\circ - ((A_{k_1 \dots k_m})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T \times (Y_j^\circ - ((A_{k_1 \dots k_m})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T$$

- variation $SS_{k_1 \dots k_m}$, characterizing the impact $(k_1 \dots k_m)$ of the factors on the response function:

$$SS_{k_1 \dots k_m} = [(Y_j^\circ - \bar{Y}_j^\circ)^T \cdot (Y_j^\circ - \bar{Y}_j^\circ)] \times [(Y_j^\circ - ((A_{k_1 \dots k_m})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T \times (Y_j^\circ - ((A_{k_1 \dots k_m})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T]$$

- the number of freedom degrees $p(k_1 \dots k_m)$, equal to the number of variables, i.e. $p(k_1 \dots k_m) = m$.

$F_{k_1 \dots k_m}$ – the Fisher's criterion is determined by:

$$F_{k_1 \dots k_m} = \frac{(Y_j^\circ - \bar{Y}_j^\circ)^T \cdot (Y_j^\circ - \bar{Y}_j^\circ) - [(Y_j^\circ - ((A_{k_1 \dots k_2})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T \times (Y_j^\circ - ((A_{k_1 \dots k_2})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T]}{(N - p(k_1 \dots k_m) - 1) \cdot [(Y_j^\circ - ((A_{k_1 \dots k_2})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T \times (Y_j^\circ - ((A_{k_1 \dots k_2})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T] \cdot p(k_1 \dots k_m) \times (Y_j^\circ - ((A_{k_1 \dots k_2})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T \times (Y_j^\circ - ((A_{k_1 \dots k_2})^T \cdot (X1_{k_1 \dots k_m})^T)^T)^T \cdot p(k_1 \dots k_m)}$$

As a result of checking the model for adequacy, there can be four outcomes.

1. If the adequacy condition is satisfied for the one-factor model, the regression equation is regarded as a satisfactory description of the response function. The transition to the model description accuracy is being carried out.

2. If the adequacy condition for the one-factor model is not fulfilled, the model and the included variable are excluded and are not considered for further analysis. They return to the stage of selective correlation analysis in order to build a one-factor model.

3. If the adequacy condition for the m – factor model is not fulfilled, the model is inadequate. The last included variable is excluded from the model and is not used for further analysis. The next step is the transition to the private correlation analysis in order to expand the mathematical model.

4. If the adequacy condition is fulfilled, the accuracy of the response function description is significant. One can assess the significance of variable k_m , included at the last step.

The significance of variable k_m is studied by F_{k_m} – Fisher criterion. The significance of variable k_m is studied by the inequality $F_{k_m} > (F_{k_m})_{ma\sigma}$.

$(F_{k_m})_{ma\sigma}$ is determined by the statistical tables. At a given level of significance $\gamma = 0.05$.

The table value is found from the equation:

$$(F_{k_m})_{ma\sigma} = F_{1-\gamma}(1, N - p(k_1 \dots k_m) - 1).$$

To determine F_{k_m} – successive inclusion criterion used for assessing the significance of variable k_m , it is necessary to calculate the components:

- variation of the impact of factors $k_1 \dots k_m$ on the response function – $SS_{k_1 \dots k_m}$;
- the number of degrees of freedom of the m-factor regression model $p(k_1, k_m)$, equal to the number of factors $p(k_1, k_m) = m$.

F – Fisher criterion for sequential inclusion of a variable is determined using the expression:

$$F_{k_m} = \frac{(SS_{k_1 \dots k_m} - SS_{k_1 \dots k_{m-1}}) \cdot (N - p(k_1 \dots k_m) - 1)}{(Y_j^2 - (A_{k_1 \dots k_m})^T \cdot (X1_{k_1 \dots k_m})^T)^T \cdot (Y_j^2 - (A_{k_1 \dots k_m})^T \cdot (X1_{k_1 \dots k_m})^T)^T}$$

As a result of checking the variable by the sequential inclusion criterion there can be four outcomes:

1. If the condition of significance of a variable is not satisfied in the one-factor model, the variable and the resulting model are excluded and are not considered for further analysis. There is a return to the stage of selective correlation analysis in order to build a one-factor mathematical model.
2. If the condition of significance of the variable in the one-factor model is fulfilled, the regression equation adequately describes the object under study. One can proceed to the stage of model description accuracy.
3. If the significance condition for the variable k_m is not satisfied in the m – factor model, it is excluded from the regression equation. The least-squares method is used to reconstruct a mathematical model with the remaining variables without including new ones. The resulting model is evaluated for adequacy by the F – Fisher criterion. With an adequate model for the analysis of significance of factors, they proceed to the stage of regression analysis.
4. If the condition of significance of the variable k_m is satisfied, one can proceed to the stage of regression analysis to assess the significance of previously included variables.

Stage 4. "Regression analysis". The regression analysis checks the impact and significance of variables k_j on the value of the multiple correlation coefficient [5].

Assessment of the significance of variables included before the last step is checked by the Student t-test, where $j = 1 \dots (m - 1)$

The condition of significance of the Student's t-test is inequality $t_j > (t_j)_{ma\sigma}$.

The Student t-test is checked by

$$t_j = \frac{|\alpha_{j+1}|}{\sqrt{\left(\left[(X1_{k_1 \dots k_m})^T \cdot (X1_{k_1 \dots k_m}) \right]^{-1} \right)_{j+1, j+1} \cdot \frac{1}{N} \cdot RSS_{k_1 \dots k_m}}}$$

The value determined empirically is compared with the tabular value of the criterion. The number of degrees of freedom at the selected level of significance $\gamma = 0.05$ is equal to $p(t_j) = (N - p(k_1 \dots k_m) - 1)$.

When checking the significance assessment of previously included variables, two options are possible.

1. The condition of significance $t_j > (t_j)_{ma\sigma}$ is fulfilled.

This means that the coefficient α_j is statically significant and the variable is stored in the regression equation. The transition to the stage of model description accuracy is being carried out.

2. The condition of significance $t_j > (t_j)_{ma\sigma}$ is not fulfilled. The contribution to the regression of this variable is considered negligible. With respect to this variable, the question about eliminating or storing it in the equation arises.

Stage 5. "Accounting for pair interaction." The analysis of assessment of the correlation coefficient between the variables is carried out.

If the variable does not satisfy the condition of the Student criterion, it is necessary to study the pairwise interaction of this variable with the included variables.

The correlation coefficient of pair interaction is checked for the condition of pair interaction characterized by the inequality:

$$0.1 < |\text{corr}(Y_{\text{ocr}})_{k_1 \dots k_m}, x_m| < 0.85$$

The fulfillment of this condition means the presence of a linear relationship between the variables. Failure to meet the conditions of pair interaction means the presence of a multicollinear connection, as a result of which the variable is excluded from the current model.

The analysis of the correlation coefficient has two options:

1. If the condition of pair interaction is violated, the corresponding variable is excluded from the mathematical model and is not considered for further analysis. The transition to the correlation analysis is carried out in order to expand the description of the response function.

2. When the conditions for pair interaction are fulfilled, the constructed model is checked for adequacy.

Stage 6. "Calculation of the model description accuracy assessment." The current and previous models are compared to select one of them as an adequate model. The criterion for the accuracy of description is an increase in the absolute value of the coefficients of the relative importance of the variables:

- fraction of the variation explained (%):

$$R_{k_1 \dots k_m} = \left(1 - \frac{RSS_{k_1 \dots k_m}}{RSS_{k_1 \dots k_m} + SS_{k_1 \dots k_m}} \right) \cdot 100\%$$

- multiple correlation coefficient $KMK_{k_1 \dots k_m}$, reflecting tightness of the relationship between the response function and all included variables;

the value of the multiple correlation coefficient lies in the range

$$0 \leq KMK_{k_1 \dots k_m} \leq 1$$

$$KMK_{k_1 \dots k_m} = \sqrt{\frac{\Delta^\circ}{\Delta}}$$

$$\Delta^\circ = \begin{vmatrix} \text{corr}(Y_j, x_1) & \text{corr}(Y_j, x_2) & \text{corr}(Y_j, x_3) & \text{corr}(Y_j, x_p) & 0 \\ 1 & \text{corr}(x_1, x_2) & \text{corr}(x_2, x_3) & \text{corr}(x_1, x_p) & \text{corr}(x_1, Y_j) \\ \text{corr}(x_2, x_1) & 1 & \dots & \text{corr}(x_2, x_p) & \text{corr}(x_2, Y_j) \\ \dots & \dots & 1 & \dots & \dots \\ \text{corr}(x_p, x_1) & \text{corr}(x_p, x_2) & \text{corr}(x_p, x_3) & 1 & \text{corr}(x_p, Y_j) \end{vmatrix}$$

$$\Delta = \begin{vmatrix} 1 & \text{corr}(x_1, x_2) & \dots & \text{corr}(x_1, x_p) \\ \text{corr}(x_2, x_1) & 1 & \dots & \text{corr}(x_2, x_p) \\ \dots & \dots & \dots & \dots \\ \text{corr}(x_p, x_1) & \text{corr}(x_p, x_2) & \dots & 1 \end{vmatrix}$$

The fraction of the explained variation increases when the variables are included in the regression equation; however, they have a weak impact on the response function. Many variables can be included in the equation with a slight increase in accuracy and an increase in the model dimension. To overcome this drawback, systematic errors are corrected. The corrected multiple correlation coefficient is calculated using the following equation [6, 7]:

$$K_{k_1 \dots k_m} = \sqrt{1 - \left(1 - K_{k_1 \dots k_m}^2\right) \cdot \frac{N-1}{N-p(k_1 \dots k_m)-1}}$$

Evaluation of the standard regression error shows what is the average error when evaluating the value of the response function of the model; it is calculated by formula [6, 8, 9]:

$$\sigma_{Y_j, k_1 \dots k_m}^2 = \sqrt{\frac{\sum_{i=1}^m (y_{ji} - \hat{y}_{ji})^2}{N-p(k_1 \dots k_m)-1}}$$

The variance of the standard error characterizes the model error due to unaccounted factors.

Stage 7. "The adoption of the final version of the Mathematical Model." from several models differing in the composition of variables; the option with maximum values of the fraction of the explained variation and the coefficient of multiple correlation is selected as optimal.

The regression model helps identify the optimal flow part of the central nervous system, at which the output energy parameters, pressure and efficiency coefficients reach maximum values.

Having accepted the model as adequate, the transition to Stage 2 is carried out expanding the model with new factors, thus repeating the cycle.

Table 2 Regression dependencies

β_{2n}	Target function	$SS_{k_1 \dots k_m}$ (%)	MKK
	Pressure ratio \bar{H}		
30°	$\bar{H} = 0.538 + 0.363 \bar{D}_1 - 0.827 \bar{D}_1^2$	97.8	0.988
n_s	$\bar{H} = 1.33442 - 0.0125n_s$	92.7	0.962
Hydraulic parameter $k_z \eta_r$			
30°	$k_z \eta_r = 0.567 + 0.405 \bar{D}_1 - 0.905 \bar{D}_1^2$	98.8	0.994
n_s	$k_z \eta_r = 1.345 - 0.012n_s$	93.1	0.965

4. RESULTS

During the processing of experimental data, the following regression dependencies were constructed. $\bar{H} = f(\bar{D}_1)$, $k_z \eta_r = f(\bar{D}_1)$, $\eta = f(\bar{D}_1)$, $\bar{H} = f(n_s)$, $k_z \eta_r = f(n_s)$ (Table 2). The fraction of the explained variation and the MCC characterize the accuracy of description of the target functions.

5. CONCLUSION

The analysis of the nature of changes in the energy indices of ACS centrifugal pumps with D_1 showed that an increase in the relative diameter of the impeller decreases the pressure qualities of the pump characterized by the value

of the pressure coefficient H and the hydraulic parameter $k_z \eta_r$. This trend corresponds to the generally accepted classification of the central nervous system according to the speed coefficient n_s , according to which the growth of \bar{D}_1 increasing the speed of the pump, increases the flow rate of the working fluid and reduces the degree of flow swirling in the interscapular channels of the impeller.

Using one-factor models of MCC energy indicators, the optimal combinations of the main geometric parameters of the impeller were revealed: the angle of the blades at the outlet and the intervals of the relative diameter, with the recommended [10] relative width of the blades at the exit ≥ 0.1 . The optimality criterion was the area of variation of the pressure coefficient and the hydraulic parameter limited by the range of $\pm 5\%$ of the extreme value of the corresponding indicator.

REFERENCES

- [1] “Academician Michael DeBaakey: I prefer to be an ordinary doctor than a venerable scientist”, *Izvestia*, vol. 72, pp. 7, 2000
- [2] N.V. Kuprienko, *Statistical methods for studying relationships. correlation and regression analysis*. St. Petersburg: Publishing House of St. Petersburg, 2008, pp. 34–78.
- [3] N. Yu Enatskaya, E.R. Khakimullin, “Probability theory and mathematical statistics for engineering areas: a textbook and workshop for applied undergraduate, 2015, pp. 56–178.
- [4] N.S. Kravchenko, *Methods of processing measurement results and error estimation in a training laboratory workshop: a training manual*. Tomsk: Tomsk Polytechnic University, 2011, pp. 1–88.
- [5] V.S. Pugachev. *Theory of Probability and Mathematical Statistics*. Tutorial, 2002, pp. 351–496.
- [6] P.Ya. Bushin, *Statistical methods*, Khabarovsk state. Acad. Economics and Law, 2002, pp. 5–79.
- [7] A.A. Zharkovsky, S.N. Shkarbul, S.A. Dmitrieva, *Mathematical models for paddle hydraulic machines. Correlation and regression analysis*. St. Petersburg: St. Petersburg. state tech. Univ., 1999, pp. 3–48.
- [8] F.A.Painters, V.A. Ivanov, *Regression analysis in experimental physics*. Moscow: Moscow State University, 2001, pp. 56–167.
- [9] G.V. Gorelova, I.A. Katsko, *Probability theory and mathematical statistics in examples and tasks using Excel*. Phoenix, 2002, pp. 34–189.
- [10] A.V. Bobkov, “Improving the efficiency of small-sized centrifugal pumps of aircraft power plants and thermal control systems”, *Dis. ... Dr. tech. Sciences*, 2004, 399 p.