

Mode diagnostics of pipeline systems of the energetics

Mikhail Sukharev.¹

¹National University of Oil and Gas «Gubkin University»

Abstract. The paper proposes a method for diagnosing gradual failures in pipeline power systems, based on tracking the dynamics of flow regime parameters. The method also makes it possible to promptly adjust the coefficients of a mathematical model of the system objects. Conclusions are made based on the analysis of the entire set of measurements, which are considered random variables due to measurement errors. Conclusions are made based on the analysis of the entire set of measurements, which are considered random variables due to instrumental errors. Examples of a gas pumping unit and a complex looped gas pipeline system are given. Calculations are performed using standard software.

1 Introduction

The information base for monitoring the state of gas and oil pipeline transportation systems is constantly being improved. However, its increased potential is actually not fully used. Mode diagnostics (MD) is an effective way to recognize gradual failures and also to quickly assess the technical condition of equipment. It, as a rule, does not require significant investment. The paper proposes the MD methodology. The constructed models are based on the maximum likelihood method; their computer implementation does not require the development of specialized digital procedures. To test the methods, computational experiments were carried out under the conditions of both stationary and unsteady fluid flow.

2 Operational diagnostics of the gas compressor (GC)

Standard devices measure the following mode parameters: p_{in}, p_{out} – pressure at the inlet and outlet of the unit, T_{in}, T_{out} – temperature at the inlet and outlet and of the blower, n speed of the centrifugal pump. If the GC has an electric drive, the power N is measured. The volumetric flow rate can be estimated using measurements of the general plant parameters of the mode, in particular the flow rate. The mathematical model of the GC connects these parameters. It includes 3 thermodynamic relationships [1] for power, compression ratio and temperature. Let's write them

$$N = \left[55,6 p_{in} Q (\varepsilon^{0,3} - 1) \right] / \eta(Q)n, \quad (1)$$

$$p_{out} / p_{in} = \varepsilon(Q,n), \quad T_{out} = T_{in} \varepsilon^{(m-1)/m}$$

Here $\eta(Q)$ is the efficiency, $\varepsilon(Q,n)$ – is the pressure characteristic, ε is the compression ratio, n is the number of revolutions of the centrifugal blower, m is the polytropic index. In addition to (1), the mathematical model of the compressor includes a number of technological limitations in the form of inequalities: maximum pressure, maximum discharge temperature and power.

The operating rules provide for various operations to monitor its technical condition based on diagnostic indicators, the dynamics of which allows one to assess the state of various GC subsystems. In contrast to these particular tasks of technological control, the problem considered in this paper refers to the diagnostics of GC. It takes into account all the information on standard measurements of the mode parameters, all physical (thermodynamic) relations (1) connecting them, and limitations determined by the design features of the unit.

The purpose of the MD is to recognize gradual failures [2]. For GCs, such failures are manifested in the subsidence of one or several characteristics, for example, for the pressure characteristic $p_{out} / p_{in} = k_{\varepsilon} \varepsilon(\hat{Q}, \hat{n})$. The adaptation coefficient k_N is introduced in a similar way into the power characteristic (the 1st formula in (1)). Measurements are always associated with possible errors. The

¹ Corresponding author: Sukharev.mgsukharev@mail.ru

measurement result is the sum of the true value and the error for example $p_{in}^* = p_{in} + \delta p_{in}$. Acting within the framework of the theory of errors, errors are considered to be normally distributed quantities $\delta p_{in} \in N(0, \sigma_p^2), \dots$. Dispersions σ_p^2, \dots are determined by the accuracy class of the measuring instrument. The set of measured values will be designated as a vector $\mathbf{Y}^* = \mathbf{Y} + \delta \mathbf{Y}$. The coefficients k_e, k_N are estimated using the maximum likelihood principle. The maximum of the likelihood function corresponds to the minimum of the function $(\mathbf{Y}^* - \mathbf{Y})^T \mathbf{C}(\mathbf{Y}^* - \mathbf{Y})$,

$\mathbf{C} = \mathbf{R}^{-1}$ the inverse matrix of the covariance matrix \mathbf{R} . The task of the MD is to estimate one or two (depending on the formulation of the problem) adaptation coefficients k_e, k_N on the basis of a one-time series of measurements and, if the coefficient is significantly less than 1, to conclude that the equipment is worn out and the GC model needs to be adjusted. MD formalization has the form of a mathematical programming problem $(\mathbf{Y}^* - \mathbf{Y})^T \mathbf{C}(\mathbf{Y}^* - \mathbf{Y}) \rightarrow \min_{\mathbf{K}}$ (2)

with conditions in the form of equalities (1) and inequalities due to technological restrictions. Here $\mathbf{K} = \|k_e, k_N\|^T$ is the vector of the estimated parameters. Typically, errors are independent and the matrix \mathbf{R} is diagonal. An iterative process with Newtonian linearization at each iteration step was used for the solution. It is very important to check on factual material how the iterative process converges depending on the initial approximation. You can get an idea of this by conducting a computational experiment. Such an experiment was carried out for two types of GCs NTs – 16/76 (with a power of 16 MW) and 2N - 25 - 76 (with a power of 25 MW). The main conclusion based on the experiment: the computational procedure adequately estimates the adaptation coefficients of the mathematical model. In the case of 2 estimated coefficients, the computational procedure has resulted to the correct, one and the same, result when varying the initial approximation and variances that determine the quality of the initial information for simulating.

3 Mode diagnostics GPS at steady flow

Let us now consider the pipeline system of the main gas supply (Fig. 1).

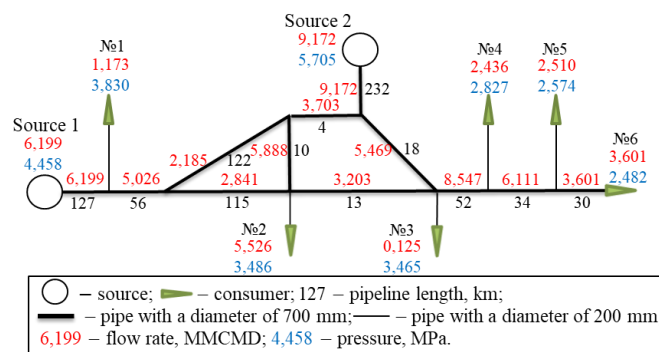


Figure 1. Schematic diagram of the GPS.

The example will illustrate the MD methods applied to a large-scale gas pipeline system (GPS). GPS in Fig. 1 is a fragment of the UGSS RF, a complex gas pipeline branch with metrological support, the real modes of which are sufficient for testing diagnostic methods. The actual problem of the operational control of the GPSs is the timely recognition of the formation of hydrates or condensate deposits. These phenomena cause an increase of hydraulic resistance of the pipes and lead to a partial failure of the pipeline. To take them into account, the actual coefficient of hydraulic resistance λ_{fact} , or the coefficient of efficiency $K = \sqrt{\lambda_{theor} / \lambda_{fact}}$, is entered into the mathematical models of the pipeline. In design calculations, it is assumed $K = 0,95$. A significant decrease in K – a signal to clean the pipe. To control the resistance of pipelines it is advisable to monitor the dynamics of the entire set of measured operating parameters of the GPS. Despite the great interest to the problems of identification and routine diagnostics of GPS [1, 3–5], the practical implementation of the research results leaves much to be desired. In this paper, the MD is applied to the GPS (Fig. 1) for both stationary and unsteady flow regimes. To analyze the mode, the methods of the theory of hydraulic circuits are used [1, 6]. Identification problems constitute one of the sections of this theory [7]. In the accepted notation, the stationary flow model is written in the form

$$\mathbf{Ax} = \mathbf{Q}, \bar{\mathbf{A}}^T \mathbf{P} = \mathbf{Ax} \quad (3)$$

where $\mathbf{x} = (x_j); j=1, K, n$ is the vector of flow rates along the arcs of the GPS graph (n is the number of graph arcs), $\mathbf{P} = (P_i) = (p_i^2); i=1, K, m$ is the vector of potentials (m is the number of graph vertices), $\mathbf{Q} = (Q_i); i=1, K, m$ is the vector of external inflows / withdrawals, $\bar{\mathbf{A}}$ is the complete matrix of incidents, \mathbf{A} is the matrix of incidents with linearly independent rows, $\mathbf{\Lambda}$ is the diagonal matrix with elements equal to the generalized resistance coefficients of pipelines, \mathbf{X} – diagonal matrix with elements $X_j = |x_j|; j=1, K, n$.

Stationary flow identification can be carried out according to various principles. The fewer parameters

are identified, the greater the chances of successfully solving the problem. To study the features of the computational process on the considered GPS, a number of experiments were carried out. From the measured values of inflows / withdrawals in sources and consumers, the flow distribution (vector \mathbf{x}) was determined according to the equations of the 1st Kirchhoff's law. This flow distribution is named \mathbf{x}^* . The values of the efficiency coefficients were evaluated according to the criterion arising from the principle of maximum likelihood (the minimum of the sum of the squares of the discrepancies between the solution and the observed values of the operating parameters). Measurement errors are again assumed to be distributed according to the normal law. If \mathbf{Y}^* the vector of measured parameters, and \mathbf{R} its correlation matrix, then the criterion is written in the form (2). To estimate the efficiency coefficients $\mathbf{K} = (K_j); j = 1, K, n$, the problem of conditional minimization of function (2) is obtained, where constraints (3) in the form of equalities are imposed on the variables. For the solution, it is again advisable to choose an iterative procedure with linearization according to Newton. The success of this method largely depends on the initial estimate.

When choosing an initial estimate, information that is not included in the formalization of the mathematical programming problem should be used as much as possible. We first made an assumption, based on technological assumptions, that the efficiency ratios of all sites are equal $K_j = K; j = 1, K, n$. Four values $K^1 = 1; K^2 = 0,95; K^3 = 0,92; K^4 = 0,90$ of the quantity K were preliminarily outlined. For each of them, solutions $\mathbf{x}^k, \mathbf{P}^k, \mathbf{Q}^k; k = 1, K, 4$ have been found. Table 1 contains the values $\|\mathbf{x}^* - \mathbf{x}^k\|, k = 1, K, 4$ and

$$\|\mathbf{Q}^* - \mathbf{Q}^k\| = \sqrt{\sum_{j \in \Theta} (Q_j^* - Q_j^k)^2}, k = 1, K, 4$$

characterizing the degree of proximity of the calculated modes to the fact. Both criteria clearly indicate that the efficiency coefficients $K_j = 0,92; j = 1, K, n$ with an accuracy quite sufficient for practice is conditional minimization problem solution.

Table 1. Analysis of expertly outlined options for identifying efficiency coefficients.

Criterion	\mathbf{x}^1	\mathbf{x}^2	\mathbf{x}^3	\mathbf{x}^4
$\sqrt{\sum_{j=1}^n (x_j^* - x_j^k)^2}$	3,7714	0,5525	0,0005	0,2199
$\sqrt{\sum_{j \in \Theta} (Q_j^* - Q_j^k)^2}$	1,5187	0,2201	0,0002	0,0906

In this case, heuristic selection turned out to be extremely successful. The example shows that, knowing the features of the functioning of the GPS, one can expertly outline several options for the set of identified coefficients and expect that among the compared options there will be one (or several) that will be a good approximation to the desired solution and can be used as initial estimate for more precise procedures.

In the case of a less successful initial selection, the following steps should be taken to consistently improve the initial estimate. From the assumption that the efficiency factors for all pipelines are equal K , go to two values K', K'' , that is, divide all pipelines into 2 groups with the same efficiency factors in the group. It is expedient to divide the partition according to the contribution of the pipelines to the residual sum of the squares of the criterion $(\mathbf{Y}^* - \hat{\mathbf{Y}})^T \mathbf{C} (\mathbf{Y}^* - \hat{\mathbf{Y}})$, where $\hat{\mathbf{Y}}$ is the solution obtained with the current approximation of the desired vector. Then the procedure can be continued, sequentially increasing the number of different values of the efficiency coefficients, that is, dividing the pipelines into 3, 4, etc. groups.

4 Mode diagnostics of GPS at unsteady flow regime

The operating parameters of the gas flow are constantly changing over time. Taking into account the nonstationarity in the MD model increases the degree of its adequacy. For unsteady gas flows in pipes, it is advisable to use a lumped parameter model (LPM), which describes well the normal mode of gas flow [1, 5]. In this model, each arc j of the graph is characterized not by one, but by two values of the flow rate: at the beginning q_j^s and end q_j^f of the arc. In this case, the 1st equation in system (3) is replaced by $\mathbf{A}^s \mathbf{q}_i^s + \mathbf{A}^f \mathbf{q}_i^f = \mathbf{Q}_i$, the 2nd one – by a finite-difference scheme for solving ordinary equations of the LPM [1, 6]. These two vector equations play the role of equality constraints in the mathematical programming problem for evaluating the vector \mathbf{K} of efficiency coefficients. The likelihood function in this case will depend on measurements over the entire observation period. The set of data acquisition times will be denoted by T^* . Due to the independence of measurements at different times, the covariance matrix \mathbf{R} for all observations is equal $\mathbf{R} = \prod_{t \in T^*} \mathbf{R}_t$, and, therefore, $\mathbf{C} = \prod_{t \in T^*} \mathbf{C}_t$. To estimate \mathbf{K} , we obtain the conditional minimization problem (3) with equality constraints. To solve it, an iterative process with Newtonian linearization is used.

To test the method, the de facto daily graphs of the GPS (Fig. 1) parameters were used. The previously obtained result for the stationary regime was used as an initial estimate $K_j = 0,92; j = 1, K, n$. The solution was carried out in 2 stages. At the 1st stage, under the assumption that all efficiency coefficients are equal, an

estimate $\hat{K}'_j = 0,918; j=1, K, n$ is obtained. In fig. 2 of out all solution components 2 functions are represented. It can be seen that the solutions of the 1st stage (curves 2) differ markedly from the measurements (curves 1) for some t .

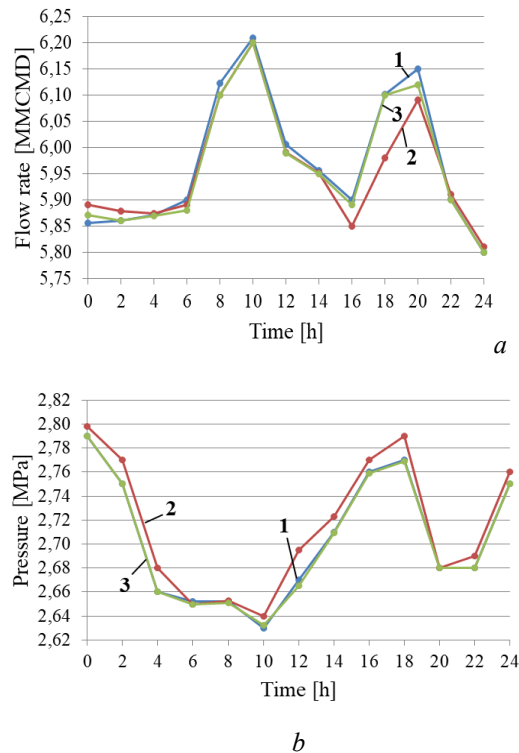


Figure 2. Flow rate at source 1 (a), pressure at consumer No. 5 (b). Legend: 1 – measurements, 2 – solution of the 1st stage, 3 – solution of the 2nd stage.

To determine what corrections it is advisable to make in the assumption about the estimated vector \mathbf{K} , let us analyze the residual sum of squares $\Phi(\mathbf{Y}^*, \hat{\mathbf{Y}}')$

$$= \sum_{t \in T^*} \sum_{j=1}^r (\hat{Q}_{j,t} - Q_{j,t}^*)^2 \sigma_{Q_{j,t}}^{-2} + \sum_{t \in T^*} \sum_{j=1}^s (\hat{P}_{j,t} - P_{j,t}^*)^2 \sigma_{P_{j,t}}^{-2},$$

which is obtained when solving the optimization problem (3). The largest contribution to the function $\Phi(\mathbf{Y}^*, \hat{\mathbf{Y}}')$ is made by nodes No. 5 and No. 6. Hence,

it is expedient to introduce the efficiency coefficients of the pipelines supplying gas to these consumers into the number of estimated coefficients of the model (Fig. 1, pipelines 34 and 30 km long). Let's designate these values K_{11}, K_{12} . At the 2nd stage of the computational experiment, when evaluating 3 unknowns, the following result is obtained: $\hat{K}_{11}'' = 0,917; \hat{K}_{12}'' = 0,925$, the remaining coefficients are equal $\hat{K}_{1-10}'' = 0,918$.

The obtained solutions – curves 3 (Fig. 2) – visually almost do not differ from measurements (curves 1). For a quantitative characteristic of the mismatch "fact – calculation", we introduce the distance between $\mathbf{Y}^*(t)$ and $\mathbf{P}(t)$ using the formula

$$\rho(\mathbf{Y}^*, \mathbf{Y}) = (1/24) \int_0^{24} \sum_j (|P_j^*(t) - P_j(t)| / P_j^*(t))^2 dt.$$

Calculations give the result $\rho(\mathbf{Y}^*, \mathbf{Y}') = 2,71 \cdot 10^{-5}$

and $\rho(\mathbf{Y}^*, \mathbf{Y}'') = 4,33 \cdot 10^{-7}$. The distance decreased by 2 orders of magnitude, from a practical point of view, the adequacy of the model is achieved. The computational experiment convincingly shows that the proposed procedures lead to the required result with high accuracy if we start the iterative procedures with a suitable initial approximation and divide the solution process into stages, providing for the correction of the set of parameters at each stage.

5 Dynamic problems

MD and model identification are carried out by the same methods and are needed in operational control, but to perform different production functions. Both problems are more significant for practice in the case when not a one-step series of measurements is considered, but a sequence of series. Let's consider this case. Let there be N^* series of measurements and $\mathbf{Y}_t^* = (\mathbf{Q}_t^*, \mathbf{P}_t^*)^T, t \in T^*$, where $T^* = \{t_1^*, t_2^*, \dots, t_{N^*}^*\}$. In each series r flow rates and s pressures are measured. Object i of the GPS is characterized by the indicator $K_i(t), i \in \hat{E}^c$ by which the diagnostics are carried out, \hat{E}^c is a set of diagnosed objects. Along with the indicators $K_i(t)$, sometimes it is necessary to evaluate some parameters of the modes. All estimates of both indicators $K_i(t)$ and parameters of the mode should be obtained based on the processing of measurements. The estimation problem is optimization and is written

$$\begin{aligned} & \sum_{t \in T^*} \sum_{j=1}^r (\hat{Q}_{j,t} - Q_{j,t}^*)^2 \sigma_{Q_{j,t}}^{-2} + \\ & \text{as } \sum_{t \in T^*} \sum_{j=1}^s (\hat{P}_{j,t} - P_{j,t}^*)^2 \sigma_{P_{j,t}}^{-2} \rightarrow \min_{\mathbf{Y}, \mathbf{K}} \quad (4) \\ & \text{under conditions } \mathbf{A}_t \mathbf{x}_t = \mathbf{Q}_t, \\ & \bar{\mathbf{A}}_t^T \mathbf{P}_t = \mathbf{A} \mathbf{K}_t \mathbf{X}_t \mathbf{x}_t, t \in T^* \end{aligned}$$

Here \mathbf{K}_t is a diagonal matrix, element of the i th row of this matrix is equal to 1 if the coefficient K_i for the arc i cannot be estimated $i \notin \hat{E}^c$, otherwise $i \in \hat{E}^c$ the coefficient K_i is included in the number of estimated values of problem (4). The condition $t \in T^*$ included in the formulation of problem (4) provides for the adjustment of the coefficients after each session of taking information about measurements. With another strategy, the adjustment is carried out not after each session, but only with significant changes in the mode parameters. The formalization of such a strategy is written in the form (4) upon replacement $T^* \Rightarrow T'$, where $T' \subseteq T^*$ is the set of moments when the adjustment is made. Let's denote this problem (4'). After each communication session, problem (4) is solved in a simplified setting, in which matrices $\mathbf{K}'_t, \mathbf{K}'$ are replaced by unity matrices. In its full form, the problem (4') is solved in those moments when the reevaluation of the coefficients is envisaged.

When solving problems (4), (4'), along with formal methods, it is recommended to use heuristic approaches.

1. It is desirable to split the solution into stages, minimizing the number of estimated coefficients from set \hat{L}^c at each stage. An example of such a technique is given in section 4.

2. Time decomposition. Consider the problem (4) for estimating parameters for each communication session. The solution to this sequence of N^* tasks will be the sequence of estimates $\hat{\mathbf{K}}_1, \hat{\mathbf{K}}_2, \dots, \hat{\mathbf{K}}_{N^*}$. Let us denote the minimum of the objective function by $\hat{W}(t) = \Phi(\hat{\mathbf{Q}}_t, \hat{\mathbf{P}}_t)$. If N^* is large enough, then time series methods can be applied to study the sequence of estimates. Question number 1: can this series be considered white noise? If possible, problem (4') splits into N^* independent problems. If not, then the correlation dependences of the vectors $\hat{\mathbf{K}}_1, \hat{\mathbf{K}}_2, \dots, \hat{\mathbf{K}}_{N^*}$ are essential. The function $\hat{W}(t)$ can be called "fact-model distance" $\hat{W}(t) = \rho(\hat{\mathbf{Y}}_t - \mathbf{Y}_t^*)$. The constancy of this function, starting from the moment of the last correction of the vector of coefficients of the model, is an indicator of the stability of the technical state of the GPS.

3. Question number 2: can a time series be considered stationary? Mathematical statistics offers different methods for answering it. a). Search for change points [9, 10]. This apparatus has several modifications, the user must choose one of them based on the results of experiments on suitable examples of time series. b). Two-sample criteria [11, 12]. A segment of a time series is divided into 2 parts – the beginning and the end. The challenge is to determine if there is a meaningful difference between these parts. The two-sample Wilcoxon test is very effective in solving practical problems. c) The method of successive differences [11, 12] serves to transform a non-stationary time series into a stationary one by passing from the original series to a series of differences. If the transition to at least the first differences is statistically justified, then the hypothesis about the stability of the situation should be discarded. d). A regression analysis method applied to a moving horizon made up of the last members of the series $\hat{W}(t)$ under consideration. The tail part of the sample $\hat{W}(t), \hat{W}(t-1), \dots, \hat{W}(t-u)$, adjacent to the current moment of time, is considered. Regression dependences of increasing degrees are constructed for it. The smallest degree is determined, after which a further increase in the degree is impractical. If the transition from degree zero to degree 1 is justified, then there is a trend: the series is nonstationary. Regression analysis in this case is carried out using orthogonal polynomials, which simplifies the computational procedure. Of the listed techniques a) – d) for practical use, first of all, the Wilcoxon test, the

method of successive differences for time series, and regression analysis for a moving sample using orthogonal polynomials are recommended. They are easy to implement and interpret the results.

4. The procedures a) – d) recommended for the analysis of a random sequence $\hat{W}(t)$ should also be applied to the estimates of model coefficients – vector components $\hat{\mathbf{K}}_t; t \in T^*$. This gives another, along with $\hat{W}(t)$, indicator of the mismatch "fact – model", which also allows you to determine, due to the change in which coefficients of the model it happened.

5. Estimation errors $\hat{\mathbf{K}}_t$ can be reduced using the information accumulated during the functioning of the GPS. If during a certain period of time according to criteria a) – d) the model coefficients were not adjusted, then each component of the estimates $\hat{\mathbf{K}}_t$ can be considered as a stationary time series and its current value can be corrected using one or several previous ones with weighting factors, for example, $\hat{\mathbf{K}}'_t = \hat{\mathbf{K}}_t + \alpha_1 \hat{\mathbf{K}}_{t-1} + \alpha_2 \hat{\mathbf{K}}_{t-2}$. Here $\hat{\mathbf{K}}_t$ is the current estimate, $\hat{\mathbf{K}}_{t-1}$ the estimate obtained from the measurements at the previous moment, etc., $\hat{\mathbf{K}}'_t$ is the adjusted estimate, α_1 and α_2 are the weight coefficients that are determined from the history.

6 Conclusions

Equipping the facilities of pipeline systems of the energetics with telemetry systems, control and measuring instruments, the growing level of computerization allow collecting and processing large amounts of information on the parameters of the fluid flow (pressure, flow rate, temperature, etc.). However, the growing information potential is actually far from being fully utilized. The paper proposes stochastic models of mode diagnostics and methodically adjacent adaptation models focused on use of standard devices measure. The models combine formalized methods of mathematical programming and heuristic techniques for finding an initial approximation of iterative procedures. Models can be implemented using standard software packages.

The developed methodological support can form the basis for the digitalization of the operational control of the energy pipeline systems at the highest hierarchical level.

References

1. M.G. Sukharev, R.V. Samoylov, Analysis and control of stationary and non-stationary conditions of gas transport (Gubkin State University of Oil and Gas, Moscow, 2016. 396 p. [in Russian].

2. Reliability of energy systems: a collection of recommended terms. Resp. ed. N.I. Voropai - M.: Energy. 2007. 191 c.
3. M.G. Sukharev, K.O. Kosova, Method and computer experiments to identification of gas supply systems // Materials of XIV All-Russian scientific workshop "Mathematical models and methods of analysis and optimal synthesis of developing pipeline and hydraulic systems." Irkutsk: ISEM SB RAS, 2014. P. 319–330. [in Russian]
4. M. G. Sukharev, K. O. Kosova, A Parameter Identification Method for Natural Gas Supply Systems under Unsteady Gas Flow. Automation and Remote Control, 2017, Vol. 78, No. 5, pp. 882–890. Original Russian Text published in Avtomatika i Telemekhanika, 2017, No. 5, pp. 141–151.
5. M.G. Sukharev, K.O. Kosova, R.V. Popov, Mathematical and computer models for identification and optimal control of large-scale gas supply systems. Energy. 2019. T. 184. C. 113-122.
6. A.P. Merenkov, V.Ya. Khasilev, Theory of hydraulic circuits. M.: Nauka, 1985. – 278 p. [in Russian]
Merenkov A.P., Hasilev V.YA. Teoriya gidravlicheskih cepej. M.: Nauka
7. N.N. Novitsky, Estimation of hydraulic circuit parameters. Novosibirsk: Science. Sib. RAS enterprise. 1998.-- 214 p. [in Russian]
9. I.V. Nikiforov, Sequential detection of changes in properties of time series. - Moscow: Nauka, 1983
10. B. E. Brodsky, B. S. Darkhovsky, Nonparametric Methods in Change-Point Problems. The Netherlands: Kluwer Academic Publishers, 1993. 209 p.
11. M. Kendall, A. Stuart, The Advanced Theory of Statistics. Vol. 3. C. Griffin. 1973. 736 p.
12. M.G. Sukharev, Forecasting methods. Russian State University of Oil and Gas. 2009.298 p.