

Regularization of the problems of planning and processing the test results of pipeline systems for hydraulic losses

Oksana Grebneva^{1,*}

¹Melentiev Energy Systems Institute of the Siberian Branch of the Russian Academy of Sciences, 664033, Lermontova str., 130, Irkutsk, Russia

Abstract. The lack of reliable information on actual characteristics and parameters is the main deterrent to the effective application of mathematical and computer modelling methods for solving problems of optimal reconstruction, adjustment, and development of operational problems. The lack of such information is solved on the basis of applying the identification problems that require expert (approximate) specification of a priori parameter values. This leads to the sensitivity of the solution of identifying a problem to the errors in setting the initial information. One way to solve this problem is to apply regularization methods. In this paper, a numerical example proposes a rule for expert assignment of a priori information on the parameters of elements, which makes it possible to minimize its influence on the results of planning active experiments. In addition, the paper shows the independence of the solution of the identification problem from the composition of the vector of independent parameters.

1 Introduction

The lack of reliable information on actual characteristics and parameters is the main deterrent to the effective application of mathematical and computer modeling methods for solving problems of optimal reconstruction, adjustment, development of operational modes, dispatching control of pipeline systems (PS) modes. For PS, this problem is solved by performing full-scale tests [1-3] or by solving identification problems. To solve such problems, methods developed both in Russia [4-5] and abroad [6-8] are applicable. In the Melentiev Energy Systems Institute of the Siberian Branch of the Russian Academy of Sciences, within the framework of the theory of hydraulic circuits [9], the method of active identification of PS [10-14] was developed, which assumes optimal planning of special experiments in order to extract the maximum of necessary information from them with the least costs for their implementation. The proposed active identification technique [10-14] is a sequential (step-by-step) planning strategy, when the next experiment is planned taking into account the information obtained after processing the results of the previous one. Unlike the traditional linear regression models for the theory of experimental planning [15],

* Corresponding author: oksana@isem.irk.ru

when the experimental design does not depend on the values of the required parameters, the TPS models are nonlinear, which significantly complicates the identification problem. The priori assignment of the unknown parameters of the elements to ensure the solvability of the task does not guarantee the optimality of the experiment and, accordingly, the achievement of the required accuracy of identification. Therefore, along with the identification problem, it is necessary to solve the related problem of the influence of the accuracy of a priori information on the accuracy of the obtained estimates of the model parameters. One way to solve this problem is to apply regularization methods [16-19].

2 The initial model and the basic design relations

The initial model is the isothermal flow distribution model [9]

$$U(\mathbf{Z}) = U(\mathbf{x}, \bar{\mathbf{Q}}, \bar{\mathbf{P}}, \mathbf{s}) = \left\{ \begin{array}{l} \bar{\mathbf{A}}\mathbf{x} - \bar{\mathbf{Q}} \\ \bar{\mathbf{A}}^T \bar{\mathbf{P}} - f(\mathbf{s}, \mathbf{x}) \end{array} \right\} = 0, \quad (1)$$

where m is the number of nodes of the circuit; n is the number of branches of the circuit; $\bar{\mathbf{A}}^T - m \times n$ – the complete matrix of incidents of nodes and branches; $\bar{\mathbf{A}} - m \times n$ – the transposed complete matrix of incidents of nodes and branches; x is the n -dimensional vector of expenditure on branches; $x_i \geq 0$ if the flow direction coincides with the orientation of the branch i , and $x_i < 0$ otherwise; s is the n -dimensional resistance vector at the branches; $\bar{\mathbf{P}} - m$ – the dimensional vector of nodal pressures; $\bar{\mathbf{Q}} - m$ – the dimensional nodal flow vector ($Q_j > 0$ in the case of inflow, $Q_j < 0$ in the case of selection (drain), $Q_j = 0$ if node j is a simple branch junction).

For model (1), it is possible to form a complete vector of parameters and represent it in the form $\mathbf{Z} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}$, where $\mathbf{X} = [\bar{\mathbf{P}} \quad \boldsymbol{\alpha}]$ is the vector of independent parameters of the model, which ensures the solvability of the model equations relative to the vector of the dependent parameters $\mathbf{Y} = [\bar{\mathbf{Q}} \quad \mathbf{x}]^T$. On the other hand, the complete vector of the model parameters can be represented as follows $\mathbf{Z} = [\mathbf{R} \quad \boldsymbol{\alpha}]^T$. Thus, in addition to the parameters of the regime (model variables) \mathbf{R} , the vector of the parameters of the elements (the coefficients of the model) $\boldsymbol{\alpha}$ also forms part of the vector \mathbf{Z} .

Also, the vector \mathbf{Z} includes both measured \mathbf{Z}_1 and non-measurable parameters \mathbf{Z}_2 . The components of the vector \mathbf{Z}_1 are the measured values of the total state vector \mathbf{Z} , namely $\mathbf{Z}_1 = \mathbf{I}_Z \mathbf{Z}$, where \mathbf{I}_Z is the matrix of correspondence between the measured parameters and the model parameters.

Taking into account the partitioning of the total state vector into the mode parameter and parameter parameters vectors, one can write instead $\mathbf{Z}_1 = \begin{bmatrix} \mathbf{R}_1 \\ \boldsymbol{\alpha}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_R & 0 \\ 0 & \mathbf{I}_\alpha \end{bmatrix} \times \begin{bmatrix} \mathbf{R} \\ \boldsymbol{\alpha} \end{bmatrix}$, where \mathbf{R}_1 is the vector of the measured values of the mode parameters; $\boldsymbol{\alpha}_1$ – a vector of a priori values of the parameters of the elements; $\mathbf{I}_R, \mathbf{I}_\alpha$ – matrices of correspondence for the vector of the parameters of the regime and the vector of the parameters of the elements, respectively.

The vector \mathbf{Z}_1 is characterized by measurement errors determined by the accuracy of the measuring instruments used. The presence of inhomogeneous measurement errors determines the necessity of using measurement errors in the problems of identification of

the covariance matrix $\mathbf{C}_{Z_1} = \begin{bmatrix} \mathbf{C}_{R_1} & 0 \\ 0 & \mathbf{C}_{\alpha, N-1} \end{bmatrix}$, where $\mathbf{C}_{R_1} = \text{diag}(\sigma_i^2, i = \overline{1, l_R})$ covariance matrix of measurement errors of the parameters of the regime; σ_i^2 - the dispersion of the measurement according to the accuracy class of the corresponding device; l_R - the number of measured parameters; $\mathbf{C}_{\alpha, N-1}$ - the a priori covariance matrix of the parameters of the elements, which is determined from the data of the previous identification or according to normative documents and has the form $\mathbf{C}_{\alpha, N-1} = \text{diag}(\sigma_i^2, i = \overline{1, l_\alpha})$, where l_α is the number of a priori specified parameters of the elements.

Obtaining information about the parameters of the elements is related to finding the a posteriori covariance matrix of the parameters of the elements $\hat{\mathbf{C}}_{\alpha, N}$, which is the submatrix of the covariance matrix of independent parameters \mathbf{C}_X [20]:

$$\mathbf{C}_X = \begin{bmatrix} \hat{\mathbf{C}}_{X_R}^{(N)} & \hat{\mathbf{C}}_{X_R, \alpha}^{(N)} \\ (\hat{\mathbf{C}}_{X_R, \alpha}^{(N)})^T & \hat{\mathbf{C}}_{\alpha, N} \end{bmatrix} = \mathbf{F}^{-1}. \quad (2)$$

The information matrix \mathbf{F} in (2) is as follows [20]:

$$\mathbf{F} = \bar{\mathbf{J}}^T \mathbf{C}_{Z_1}^{-1} \bar{\mathbf{J}} = \begin{bmatrix} (\bar{\mathbf{J}}_R^{(N)})^T \mathbf{C}_{R_1}^{-1} \bar{\mathbf{J}}_R^{(N)} & (\bar{\mathbf{J}}_R^{(N)})^T \mathbf{C}_{R_1}^{-1} \bar{\mathbf{J}}_\alpha^{(N)} \\ (\bar{\mathbf{J}}_\alpha^{(N)})^T \mathbf{C}_{R_1}^{-1} \bar{\mathbf{J}}_R^{(N)} & (\bar{\mathbf{J}}_\alpha^{(N)})^T \mathbf{C}_{R_1}^{-1} \bar{\mathbf{J}}_\alpha^{(N)} + \hat{\mathbf{C}}_{\alpha, N-1}^{-1} \end{bmatrix}. \quad (3)$$

It is clear from (3) that the matrix $\hat{\mathbf{C}}_{\alpha, N}$ is a function of the a priori matrix $\mathbf{C}_{\alpha, N-1}$ and, in turn, is a priori for subsequent identification. The Jacobi matrix in (3) $\bar{\mathbf{J}}$ is defined as $\bar{\mathbf{J}} = \begin{pmatrix} \partial \mathbf{Z}_1 \\ \partial \mathbf{X} \end{pmatrix} = \mathbf{I}_Z \times \begin{bmatrix} \mathbf{E} \\ \mathbf{F}' \end{bmatrix}$, where \mathbf{E} is the identity matrix; $\mathbf{F}' = \frac{\partial \mathbf{Y}}{\partial \mathbf{X}}$ - matrix of derivatives, which can be found from the relations, $\frac{\partial \mathbf{Y}}{\partial \mathbf{X}} = \begin{pmatrix} \partial \mathbf{U} \\ \partial \mathbf{Y} \end{pmatrix}^{-1} \begin{pmatrix} \partial \mathbf{U} \\ \partial \mathbf{X} \end{pmatrix} = \begin{bmatrix} \bar{\mathbf{A}}(\mathbf{f}'_x)^{-1} \bar{\mathbf{A}}^T & \bar{\mathbf{A}}(\mathbf{f}'_x)^{-1} \mathbf{f}'_\alpha \\ (\mathbf{f}'_x)^{-1} \bar{\mathbf{A}}^T & (\mathbf{f}'_x)^{-1} \mathbf{f}'_\alpha \end{bmatrix}$, $\frac{\partial \mathbf{U}}{\partial \mathbf{X}} = \begin{vmatrix} 0 & 0 \\ \bar{\mathbf{A}}^T & \mathbf{f}'_\alpha \end{vmatrix}$, $\frac{\partial \mathbf{U}}{\partial \mathbf{Y}} = \begin{vmatrix} -\mathbf{E}_m & \bar{\mathbf{A}} \\ 0 & \mathbf{f}'_x \end{vmatrix}$, where $\mathbf{f}'_x = \text{diag}(2s_i |x_i|)$; $\mathbf{f}'_\alpha = \text{diag}(|x_i| |x_i|)$.

From (2), (3), one can obtain a recurrence relation for calculating the a posteriori covariance matrix of the parameters of the elements

$$\hat{\mathbf{C}}_{\alpha, N}^{-1} = \hat{\mathbf{C}}_{\alpha, N-1}^{-1} + (\bar{\mathbf{J}}_\alpha^{(N)})^T \mathbf{G}_{R_1}^{(N)} \bar{\mathbf{J}}_\alpha^{(N)}. \quad (4)$$

Relation (4) establishes the relationship between the value of the covariance matrix of the parameters of the elements obtained as a result of the planning and the initial information about the parameters of the elements.

3 Regularization of identification tasks

After the first experiment, you can level the variance values during the data processing σ_α . The objective function has the following form:

$$\Phi = (\tilde{\mathbf{R}} - \hat{\mathbf{R}})^T \mathbf{C}_{R_1}^{-1} (\tilde{\mathbf{R}} - \hat{\mathbf{R}}) + (\tilde{\mathbf{a}} - \hat{\mathbf{a}})^T \mathbf{C}_\alpha^{-1} (\tilde{\mathbf{a}} - \hat{\mathbf{a}}), \quad (5)$$

where $\tilde{\mathbf{R}}, \tilde{\mathbf{a}}$ - the vector of the measured parameters of the regime and the parameters of the elements, respectively; $\hat{\mathbf{R}}, \hat{\mathbf{a}}$ - the vector of estimates of the parameters of the regime and the parameters of the elements, respectively.

Substituting $\mathbf{C}_\alpha = \text{diag}(\sigma_i^2, i = 1, \overline{l_\alpha})$, where the values σ_i are proportional to the a priori uncertainty of specifying the parameters of the elements $\tilde{\mathbf{a}}$, one can write into expression (5)

$$\Phi = (\tilde{\mathbf{R}} - \hat{\mathbf{R}})^T \mathbf{C}_{R_1}^{-1} (\tilde{\mathbf{R}} - \hat{\mathbf{R}}) + \frac{1}{\sigma_\alpha^2} (\tilde{\mathbf{a}} - \hat{\mathbf{a}})^T \mathbf{C}_\alpha^{-1} (\tilde{\mathbf{a}} - \hat{\mathbf{a}}) . \quad (6)$$

It can be seen from expression (6) that for σ_α a very small task, the estimates obtained are shifted relative to the a priori given, so the model becomes inadequate. If set σ_α too large, the task becomes ill-conditioned and the effect of using a priori information is reduced to none. Therefore, it is necessary to find σ_α a value that would ensure, on the one hand, the adequacy of the model, on the other hand, the solvability of the identification problem.

The following approach is proposed in [20]. Find the greatest value σ_α that would ensure the adequacy of the model.

In Fig. 1 shows the dependence of the value of the objective function and the mean error on the standard deviation value.

In addition, there is a line corresponding to the threshold value of the objective function, which corresponds to χ^2 for degrees of freedom, taking into account the real measurements.

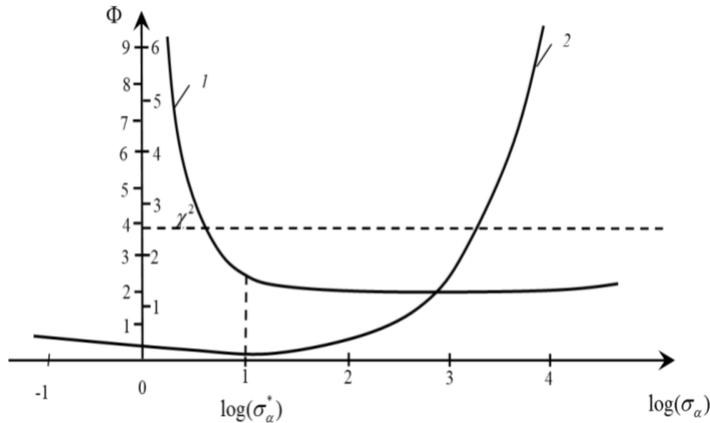


Fig. 1. Change in the minimum of the objective function and the relative average error, depending δ_α on the value σ_α : 1 - the curve for Φ ; 2 - the curve for δ_α .

From Fig. 1. that the minimum value corresponds δ_α to the inflection point of the objective function that lies below the threshold value

4 Calculation research

For convenience, the results of the studies will be demonstrated using the conditional TPS scheme shown in Fig. 2.

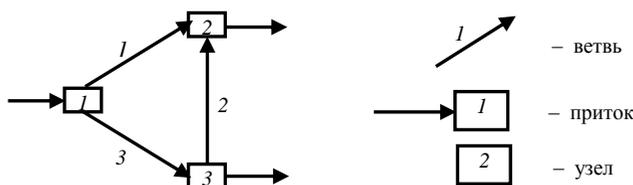


Fig. 2. The conditional scheme of PS.

4.1 Investigation of the influence of initial information on the quality of identification

The purpose of numerical studies is to reveal the nature of the effect of setting the accuracy of the parameters of elements on the information criterion - the determinant of the covariance matrix of the parameters of elements $\det(C_\alpha)$, which is used in the task of planning conditions for active identification as an information criterion. The results of the investigations are shown in Fig. 3.

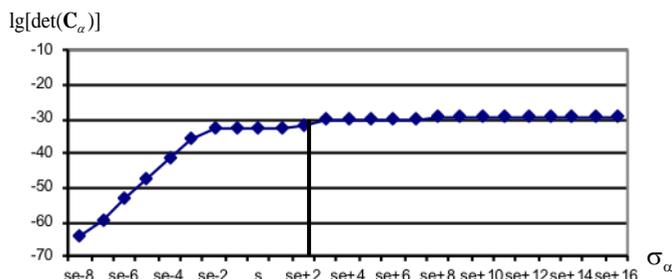


Fig. 3. Dependence of the information criterion on the initial information.

Fig. 3 shows that up to a value whose order corresponds to the order of resistance, there is a region of distortion of the value of the determinant of the covariance matrix of the parameters of the elements. After this, the alignment occurs and, reaching σ_α a value of 3 orders of magnitude higher than the resistance, the curve becomes smoothed out.

This can be interpreted as follows. If the standard deviation exceeds the absolute value of the parameter itself, then this estimate is statistically insignificant. Therefore, it is necessary to specify such a variance so that the estimation of the parameters does not have a big influence on the chosen information criterion.

Based on the results of the experiments on the conditional example of a PS using an isothermal flow distribution model, a threshold value is obtained σ_α , with further increase of which there is no distortion of the information criterion. Thus, in the future should be set such a value of the standard deviation, which would be a maximum of 2 orders of magnitude less than the value itself.

4.2 Dependence of the solution on the composition of the vector of independent parameters of the model

To calculate the information criterion, in addition to the accuracy of specifying the parameters of the elements, it is necessary to know the composition of the vector of independent parameters of the model X , which along with the parameter vector of the elements includes the vector of independent parameters of the X_R regime. The question

arises as to the influence of the composition of the independent parameters of the model on the optimality of the experiment and on the value of the threshold value σ_α .

For the comparative analysis, two variants with different composition of the vector X were chosen. A comparative analysis of the results of the studies showed the complete coincidence of the curves of the information criterion variation depending on the initial accuracy of the parameters of the elements for the two selected variants of the composition of the vector X (Figure 4).

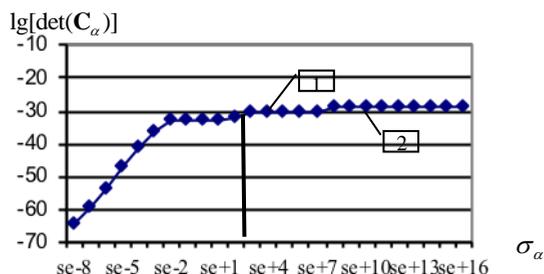


Fig. 4. Changing the value of the information criterion: 1 - 1-st variant of the composition of the vector of independent parameters; 2 - 2-nd version of the composition of the vector of independent parameters.

5 Conclusion

A numerical investigation of the influence of the accuracy of a priori information setting on the quality of PS active identification is performed in the work. Based on the results of the experiments performed on the conditional example of a PS using an isothermal flow distribution model, a threshold value of the accuracy of setting the initial values of the parameters was obtained, with further increase of which there is no distortion of the information criterion. The rule of the expert task of a priori information on the parameters of the elements is proposed, which makes it possible to minimize its influence on the results of planning active experiments.

Numerical examples show the independence of the optimality of the experiment and the threshold value of the accuracy of the initial information from the composition of the vector of independent parameters.

The research was carried out within the framework of the III.17.4.3 project of the Basic Research Program of the Siberian Branch of the Russian Academy of Sciences (AAAA-A17-117030310437-4) with the financial support of the Russian Foundation for Basic Research and the Government of the Irkutsk Region within the framework of the scientific project No. 17-48-380021.

References

1. RD34.20.519-97 (Service of Excellence, Moscow, 1997)
2. RD34.09.255-97 (Service of Excellence, Moscow, 1998)
3. RD153.34.1-20.526-00 (OJSC "VTI", 2001)
4. N. N. Novitsky, *Estimation of the parameters of hydraulic circuits* (Nauka, Novosibirsk, 1998)
5. N. N. Novitskii, V. G. Sidler, Bulletin of the AS USSR: Energy and Transport, **4**,

- 15(1984)
6. Srinivasa Lingireddy, Lindell E. Ormsbee. *Civil Engineering & Environmental Sys*, **19(1)**, 13 (2012)
 7. S. Alvisi, E. Creaco, M. Franchini. *Urban Water Journal*, **8 (4)**, 203 (2011)
 8. Yu. Guoping, Roger S. Powell. *Intern. Journal of Systems Science*, **25(12)**, 2155 (1994)
 9. A. P. Merenkov, V. Ya. Hasil, *Theory of hydraulic circuits* (Nauka, Moscow, 1985)
 10. N. N. Novitsky, O. A. Grebneva, *Pipeline energy systems: models, applications, information technologies* (Publishing House "Oil and Gas", Moscow, 2000)
 11. *Pipeline Energy Systems: Methods of mathematical modeling and optimization* (Nauka, Novosibirsk, 2007)
 12. O. A. Grebneva, N. N. Novitsky, *Thermal Engineering*, **61 (9)**, 685 (2014)
 13. O. A. Grebneva, N. N. Novitsky, *Thermal Engineering*, **61 (10)**, 754 (2014)
 14. *Pipeline systems of power engineering* (Nauka, Novosibirsk, 2015)
 15. Yu. P. Adler, E. V. Markova, Yu. V. Granovsky, *Planning of Experiments* (Nauka, Moscow, 1976)
 16. A. N. Tikhonov, V. Ya. Arsenin, *Methods for solving ill-posed problems* (Nauka, Moscow, 1974)
 17. A. N. Tikhonov, N. V. Ufimtsev, *Statistical processing of experimental results* (Moscow State University Press, Moscow, 1988)
 18. O'Leary D.P. *SIAM J. Sci. Comput*, **23 (4)**, 1161 (2001)
 19. N. S. Hoang, A. G. Ramm *Inverse Problems in Science and Engineering*, 16 (5), 616 (2008)
 20. N. N. Novitsky, *Author's abstract. dis. ... cand. tech. sciences* (ISEM SB RAS, Irkutsk, 1986)