# Diagnostics of leaks with unknown amplitudes against the background of interference caused by accidental consumption in the hydraulic system for the forest complex

*Svetlana* Sazonova[1*], *Konstantin* Zolnikov[1], *Tatyana* Skvortsova[1], *Andrey* Kravchenko[1] and *Anton* Zarevich[1]

[1]Voronezh State Forestry University named after G. F. Morozov, 394087 Voronezh, Russia

**Abstract.** The formation of algorithms for leak detection is considered. It is shown that the solution to the problem of specific leak detection may vary depending on the type of hydraulic system and the technologies used. Hydraulic fire extinguishing systems for buildings and structures are used in the forest complex. It is shown that main pipelines with combustible substances pass through the forest complex, leaks in which can cause fires, explosions and environmental damage. Mathematical models and methods for diagnosing leaks in hydraulic systems are considered, including identifying the facts of leakage based on the use of mathematical models to determine the location and size of such leaks. In this paper the research focuses on the detection of leaks with unknown amplitudes based on the verification of a two-alternative hypothesis for a hydraulic system, taking into account interference from stochastic consumption. A single leak is considered as a significant amplitude jump against the background of noise arising in the hydraulic pipeline system due to the selection of the target product by consumers. It is noted that the transition from the task of verifying a two-alternative hypothesis to the task of verifying multi-alternative hypotheses or pattern recognition provides a deeper and more comprehensive analysis of leaks.

## 1 Introduction

Various hydraulic pipeline systems can be used or transited in the forest complex. For example, hydraulic fire extinguishing systems are used in the forest complex, which contribute to the rapid extinguishing of fires near populated areas. Leaks of various nature may occur in transit pipeline systems of gas supply and in oil pipelines, which can cause environmental pollution and be an environmental hazard. Sabotage actions can lead to the partial destruction of such systems, cause explosions and fires in forests. Explosions and fires can also occur in the event of a fire of the target product as a result of leaks. Such leaks not only cause serious material damage to the forest complex, but can also lead to human casualties.

* Corresponding author: ss-vrn@mail.ru

A mathematical model and a method for implementing the problem of leak detection in hydraulic systems (HS) represent an important study to ensure the safety and reliability of such systems. The mathematical model and leak detection methodology have the potential to significantly improve the efficiency and accuracy of leak detection in HS, ensuring the safety and stability of such systems.

When creating an information system for processing information about the state of a control object, it is necessary to take into account not only specific criteria for its optimality, but also to ensure sufficient accuracy and reliability of the information received [1]. The process of data collection, processing and analysis should be optimized to ensure the efficient operation of the management system.

Information systems used to manage facilities can be used in various fields such as manufacturing, medicine, transportation, etc. They allow you to quickly receive information about the state of objects, analyze them and make necessary management decisions [2, 3].

Therefore, the development of an information system is an important and integral part of the management process, which allows to increase its efficiency and effectiveness [4, 5].

The choice of an information system category depends on which aspects of facility management need to be monitored and what information is needed to make decisions. When a leak is detected, the system must be able to detect the leak quickly and accurately and take the necessary measures to prevent leakage [6, 7].

It is also important to take into account that the choice of an information system category must be adapted to the specific conditions and requirements of the organization that will use it. They must consider the structure and characteristics of the management object, possible threats and risks, as well as available resources that will be used to create and maintain the system [8].

The application of the two-alternative hypothesis assumes that in the conditions of a conventional leak, it is not a prerequisite, but significantly simplifies the task without losing the solution of common problems.

## 2 Leak detection algorithm

The task of combining unknown parameters, such as leakage amplitude and noise intensity, is really more difficult and more important for practical implementation. In this case, it is necessary to solve the problem of simultaneous estimation of two parameters, which requires more complex methods and models.

Various statistical methods, such as the maximum likelihood method or the Bayesian statistical method, can be used to solve common leak diagnosis problems. It is also important to consider the possible relationship between the scattering amplitude and the noise intensity, if such a relationship exists.

The method of solving complex problems requires a deeper analysis of data and a more careful selection of models and evaluation methods. However, a successful solution to such a problem can bring significant practical benefits in the diagnosis of leaks in the hydraulic system.

The least squares method can be used to estimate the amplitude of a signal in such a sample. To do this, select a model, for example, a linear model. Then the loss function is minimized, defined as the sum of the squares of the deviation between the observed value and the predicted value of the model.

The obtained estimates of the model parameters allow us to determine the amplitude of the signal. It is important to remember that increasing the randomness of consumption increases the error of the estimated parameters of the model. Therefore, to increase accuracy, it is recommended to increase the number of observations in the sample.

Information about the leak will be formulated in the following form [1]

$$X_n = \lambda\alpha S_n + \Xi_n, \tag{1}$$

in which $\lambda = 1$ for $p_1$; $\lambda = 0$ for $p_2 (p_1 + p_2 = 1)$; $S_n$ is the specified vector $(s_i = s(t))$; $K_\xi = \sigma^2 \|\delta_{ij}\|$, $\sigma^2 = 2\Delta f N_0$.

To detect a signal in a noisy environment, the optimal method may be to compare the corresponding threshold value of the probability ratio. To do this, calculate the ratio of the probability of a possible signal to the probability of a non-existent signal.

The threshold value is set in such a way as to ensure the maximum balance between signal detection and minimizing false alarms. If the probability ratio exceeds the set threshold value, the signal is considered detected.

This method is optimal from the point of view of detection theory, since it provides the most efficient use of observational information for signal detection and provides a well-known signal and noise model.

$$\Lambda(X_n \mid \alpha^*, \sigma_1^{2*}, \sigma_2^{2*}) = \frac{P_1(X_n \mid \alpha^*, \sigma_1^{2*})}{P_2(X_n \mid \sigma_2^{2*})}. \tag{2}$$

$P_2(X_n \mid \sigma^2)$ and $\sigma_2^{2*}$ we determine as

$$\sigma_2^{2*} = X_n^T X_n / n, \tag{3}$$

$$P_2(X_n \mid \sigma_2^{2*}) = \frac{1}{(2\pi\sigma_2^{2*})^{n/2}} \exp\left(-\frac{n}{2}\right). \tag{4}$$

$P_1(X_n \mid \alpha, \sigma^2)$ we determine as

$$P_1(X_n \mid \alpha, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}(X_n - \alpha S_n)^T (X_n - \alpha S_n)\right]. \tag{5}$$

Determine $\alpha^*$ and $\sigma_1^{2*}$

$$\frac{\partial \ln P_1(X_n \mid \alpha, \sigma^2)}{\partial \alpha} = 0, \quad \frac{\partial \ln P_1(X_n \mid \alpha, \sigma^2)}{\partial(\sigma^2)} = 0. \tag{6}$$

Next, we transform the equation to the form

$$(1/\sigma^2)(X_n - \alpha S_n)^T S_n = 0, \quad (1/2\sigma^2)[n - (X_n - \alpha S_n)^T (X_n - \alpha S_n)/\sigma^2] = 0, \tag{7}$$

therefore

$$\alpha^* = \frac{X_n{}^T S_n}{S_n{}^T S_n} = \frac{1}{E}\sum_{i=1}^{n} x_i s_i, \tag{8}$$

$$\sigma_1^{2*}(\alpha^*) = \frac{1}{n}(X_n - \alpha^* S_n)^T (X_n - \alpha^* S_n) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \alpha^* s_i)^2. \tag{9}$$

Transform (2) to the form:

$$\Lambda(X_n \mid \alpha^*, \sigma_1^{2*}, \sigma_2^{2*}) = \left[\frac{\sigma_2^{2*}}{\sigma_1^{2*}(\alpha^*)}\right]^{n/2} = \left\{\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2 - \frac{1}{E}\left(\sum_{i=1}^{n} x_i s_i\right)^2}\right\}^{n/2}, \tag{10}$$

In which $E = \sum_{i=1}^{n} s_i^2.$

Compare the expression (10) with the threshold

$$C_{12} = \frac{g_{21} - g_{22}}{g_{12} - g_{11}} \frac{p_2 a_2(X_n)}{p_1 a_1(X_n)}. \tag{11}$$

Determine $a_2(X_n)$

$$a_2(X_n) = \frac{1}{\Delta(\sigma^2)} \frac{(2\pi)^{1/2}}{(n^3/2)^{1/2}} B_2, \tag{12}$$

In which $B_2 = X_n^T X_n.$

Determine $a_1(X_n)$

$$a_1(X_n) = \omega(\alpha^*, \sigma_1^{2*}) \frac{2\pi}{\det^{1/2} D_1}. \tag{13}$$

For $\alpha$ and $\sigma^2$ [1], $\omega(\alpha^*, \sigma_1^{2*}) = (1/\Delta\alpha)(1/(\Delta\sigma^2)).$ For the matrix $D_1$ the form is valid

$$D_1 = \left\| \begin{matrix} -\frac{\partial^2 \ln P_1}{\partial\alpha^2} & -\frac{\partial^2 \ln P_1}{\partial\alpha\partial(\sigma^2)} \\ -\frac{\partial^2 \ln P_1}{\partial(\sigma^2)\partial\alpha} & -\frac{\partial^2 \ln P_1}{\partial(\sigma^2)^2} \end{matrix} \right\|_{\substack{\alpha=\alpha^* \\ \sigma^2=\sigma_1^{2*}}} \tag{14}$$

Based on (7)

$$\left.\frac{\partial^2 \ln P_1}{\partial\alpha^2}\right|_{\substack{\alpha=\alpha^* \\ \sigma^2=\sigma_1^{2*}}} = -\left.\frac{E}{\sigma^2}\right|_{\substack{\alpha=\alpha^* \\ \sigma^2=\sigma_1^{2*}}} = -\frac{En}{B_1(\alpha^*)}, \tag{15}$$

in which

$$B_1(\alpha) = (X_n - \alpha S_n)^T (X_n - \alpha S_n); \quad \left.\frac{\partial^2 \ln P_1}{\partial (\sigma^2)^2}\right|_{\substack{\alpha=\alpha^* \\ \sigma^2=\sigma_1^{2*}}} = -\frac{1}{2\sigma^4}\left(n - \frac{2B_1(\alpha)}{\sigma^2}\right)\Bigg|_{\substack{\alpha=\alpha^* \\ \sigma^2=\sigma_1^{2*}}} = -\frac{n^3}{2}\frac{1}{B_1^2(\alpha^*)};$$

$$\left.\frac{\partial^2 \ln P_1}{\partial\alpha\partial(\sigma^2)}\right|_{\substack{\alpha=\alpha^* \\ \sigma^2=\sigma_1^{2*}}} = -\frac{1}{\sigma^4}(X_n - \alpha S_n)^T S_n\Big|_{\substack{\alpha=\alpha^* \\ \sigma^2=\sigma_1^{2*}}} = 0. \qquad (16)$$

Therefore

$$\det D_1 = \frac{n^4}{2}\frac{E}{B_1^3(\alpha^*)}, \qquad (17)$$

$$a_1(X_n) = \frac{2\pi\sqrt{2}B_1^{3/2}(\alpha^*)}{\Delta\alpha\Delta(\sigma^2)n^2E^{1/2}}. \qquad (18)$$

The threshold $C_{12}$ have the form

$$C_{12} = \frac{g_{21} - g_{22}}{g_{12} - g_{11}}\frac{p_2}{p_1}\frac{\Delta\alpha}{\sqrt{\pi}}\sqrt{\frac{E}{2\sigma_1^{2*}(\alpha^*)}}\frac{\sigma_1^{2*}}{\sigma_1^{2*}(\alpha^*)}. \qquad (19)$$

If

$$\sigma_2^{2*} / \sigma_1^{2*}(\alpha^*) \geq C(X_n), \qquad (20)$$

the threshold $C(X_n)$ is determined this way

$$C(X_n) = \left(\frac{g_{21} - g_{22}}{g_{12} - g_{11}}\frac{p_2}{p_1}\frac{\Delta\alpha}{\sqrt{\pi}}\sqrt{h_0^*}\right)^{\frac{2}{n-2}}, \qquad (21)$$

in which $h_0^* = E / 2\sigma_1^{2*}(\alpha^*)$ is the estimated value of the normalized signal-to-noise ratio $h_0 = E / 2\sigma^2$. If we assume $\alpha_1 = 0$,

$$C(X_n) = \left(\frac{g_{21} - g_{22}}{g_{12} - g_{11}}\frac{p_2}{p_1}\sqrt{\frac{\Delta h^*}{\pi}}\right)^{\frac{2}{n-2}}. \qquad (22)$$

The algorithm (20) is converted to the form

$$\frac{\int_0^T x^2(t)dt}{\left\{\int_0^T x^2(t)dt - \frac{1}{E_0}\left[\int_0^T x(t)s(t)dt\right]^2\right\}} \geq C[x(t)], \tag{23}$$

in which

$$E_0 = \int_0^T s^2(t)dt. \tag{24}$$

Determine the threshold $C[x(t)]$

$$C[x(t)] = C\left\{\int_0^T x^2(t)dt - \frac{1}{E_0}\left[\int_0^T x(t)s(t)dt\right]^2\right\}^{-\frac{1}{2\Delta fT - 2}}, \tag{25}$$

in which

$$C = \left(\frac{g_{21} - g_{22}}{g_{12} - g_{11}} \frac{p_2}{p_1} \sqrt{\frac{\Delta fTE_m}{\pi}}\right)^{\frac{1}{\Delta fT - 1}}, \tag{26}$$

$E_m = \alpha_2^2 E_0$ is the maximum signal energy.
The threshold $C$ is determined this way

$$C = \frac{g_{21} - g_{22}}{g_{12} - g_{11}} \left(\frac{p_2}{p_1}\right)^{2/n}. \tag{27}$$

The algorithm disregarding the threshold value

$$\frac{\sum_{i=1}^{n} x_i^2}{\left(\sum_{i=1}^{n} x_i^2 - y_n^2\right)} \geq C, \tag{28}$$

In which

$$y_n = \frac{1}{\sqrt{E}} \sum_{i=1}^{n} x_i s_i. \tag{29}$$

After the introduction of the linear transformation

$$Y_n = GX_n \tag{30}$$

With the matrixes $G$ and $G^T = G^{-1}$ we have

$$X_n = G^{-1}Y_n = G^T Y_n; X_n^t = Y_n^T G; \quad X_n^T X_n = Y_n^T G G^{-1} Y_n = Y_n^T Y_n, \tag{31}$$

that is

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i^2. \tag{32}$$

The correlation function

$$\overline{(y_i - \bar{y}_i)(y_j - \bar{y}_j)} = \sum_{k,l=1}^{n} G_{ik} G_{jl} \overline{(x_k - \bar{x}_k)(x_l - \bar{x}_l)} = \sum_{k=1}^{n} \sigma^2 G_{ik} G_{kj}^{-1} = \sigma^2 \delta_{ij}. \tag{33}$$

It is taken into account that according to (29) and (30)

$$G_{ni} = s_i / \sqrt{E}. \tag{34}$$

The problem situation $G^{-1} = G^T$, $\sum_{k=1}^{n} G_{ik} G_{jk} = \delta_{ij}$, means that

$$\sum_{k=1}^{n} G_{ik} G_{nk} = \sum_{k=1}^{n} G_{ik} \frac{s_k}{\sqrt{E}} = \delta_{in}. \tag{35}$$

Therefore with $i \neq n$ $\sum_{k=1}^{n} G_{ik} s_k = 0,$ so with $\lambda = 1$

$$\bar{y}_i = \sum_{k=1}^{n} G_{ik} \bar{x}_k = \alpha \sum_{k=1}^{n} G_{ik} s_k = 0. \tag{36}$$

with $\lambda = 0$ the mathematical expectation $\bar{x}_k = 0$, therefore $\bar{y}_i = 0$.

Based on (29) we have

$$m_n = \bar{y}_n = \begin{cases} 0 & whith \;\; \lambda = 0, \\ \alpha\sqrt{E} & whith \;\; \lambda = 1. \end{cases} \tag{37}$$

Therefore,

$$\frac{\sum_{k=1}^{n} x_i^2}{\left(\sum_{k=1}^{n} x_i^2 - y_n^2\right)} = 1 + \frac{y_n^2}{\sum_{i=1}^{n-1} y_i^2}, \tag{38}$$

The detection algorithm

$$\varsigma^2 = \frac{y_n^2}{\sum_{i=1}^{n-1} y_i^2}. \tag{39}$$

Find the probability $P(\varsigma^2 \geq C_0)$. It is obvious that

$$P(\varsigma^2 \geq C_0) = 1 - F_\varsigma\left(\sqrt{C^0}\right) + F_\varsigma\left(-\sqrt{C^0}\right), \tag{40}$$

in which $F_\varsigma$ is the integral distribution law for $\varsigma$.

Introduce

$$\eta = \sqrt{n-1}\varsigma = \frac{y_n}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1} y_i^2}}, \tag{41}$$

So

$$P(\varsigma^2 \geq C_0) = 1 - F_\eta\left(\sqrt{(n-1)C^0}\right) + F_\eta\left(-\sqrt{(n-1)C^0}\right), \tag{42}$$

in which $F_\eta$ is the Student's distribution.

When $n$ are large enouth

$$\sum_{i=1}^{n-1} y_i^2 \approx (n-1)\sigma^2. \tag{43}$$

A decision on the presence of a leak is made if

$$y_n^2 > (n-1)\sigma^2 C^0 = K^0. \tag{44}$$

The algorithm reduces to a comparison at each $n$ -th step of the observations of the ratio

$$\sigma_2^{2*}(n)/\sigma_1^{2*}(n,\alpha^*(n)) = \frac{\sum_{i=1}^{n} x_i^2}{\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{E_n}\left(\sum_{i=1}^{n} x_i s_i\right)^2\right]} \tag{45}$$

with two thresholds

$$C_1(n) = \left[\frac{p_2}{p_1}\frac{g_2(n)}{\varphi_1(n)}\sqrt{\frac{\Delta h^*(n)}{\pi}}\right]^{\frac{2}{n-2}}, \quad C_2(n) = \left[\frac{p_2}{p_1}\frac{\varphi_2(n)}{g_1(n)}\sqrt{\frac{\Delta h^*(n)}{\pi}}\right]^{\frac{2}{n-2}}, \tag{46}$$

in which $\Delta h^*(n) = \alpha_{max}^2 E_n / 2\sigma_1^{2*}(n,\alpha^*(n))$.

Imagine the rule for deciding whether there is a leak at the $n$-th observation step in the form

$$\ln\frac{\sigma_2^{2*}}{\sigma_1^{2*}(n,\alpha^*(n))} \geq \frac{2}{n-2}\ln\left[\frac{g_{21}-g_{22}}{g_{12}-g_{11}}\frac{p_2}{p_1}\sqrt{\frac{\Delta h^*(n)}{\pi}}\right]. \tag{47}$$

Substituting (45) into (47), we get

$$\ln\frac{\sigma_2^{2*}}{\sigma_1^{2*}(n,\alpha^*(n))} \approx \ln\frac{\sigma_2^{2*}(n-1)}{\sigma_1^{2*}(n-1,\alpha^*(n-1))} + \\ + \ln\frac{1+x_n^2/(n-1)\sigma_2^{2*}(n-1)}{1+(x_n-\alpha_i^*(n-1))^2/(n-1)\sigma_1^{2*}(n-1,\alpha^*(n-1))}. \tag{48}$$

The values $\alpha^*(n)$, $\sigma_1^{2*}(n,\alpha^*(n))$ and $\sigma_1^{2*}(n)$ are formulated based on (9) and (3) and have the form taking into account (48)

$$\alpha^*(n) = \alpha^*(n-1) + (s_n/E_n)(x_n - \alpha^*(n-1)s_n),$$
$$\sigma_1^{2*}(n,\alpha^*(n)) \approx \sigma_1^{2*}(n-1,\alpha^*(n-1)) + [(x_n-\alpha^*(n-1)s_n)^2 - \sigma_1^{2*}(n-1,\alpha^*(n-1))]/n,$$
$$\sigma_2^{2*}(n) = \sigma_2^{2*}(n-1) + [x_n^2 - \sigma_2^{2*}(n-1)]/n. \tag{49}$$

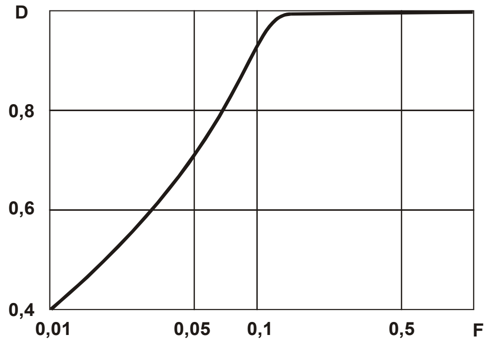


**Fig. 1.** Characteristics of a signal with unknown amplitude when detected in noise of unknown intensity

With this method, leak detection will be obtained by comparing statistical indicators such as pressure in the target product or other parameters measured in the system with a certain threshold value.

Mathematical and statistical methods such as hypothesis testing, maximum likelihood methods, Bayesian methods, etc. are used to model such a system. The problem of leak detection must be implemented in conjunction with the tasks of current distribution analysis and static assessment of the state of the HS.

As a result, based on the verification of a two-alternative hypothesis, a comprehensive leak diagnostics model was developed in the HS, which effectively identifies leaks and takes the necessary measures to eliminate them.

## 3 Conclusion

The development of a method based on the verification of a two-alternative hypothesis is a common method in statistics and signal processing.

In the context of detecting leaks in the HS, the process may look like this. Two alternative hypotheses are considered: $H_0$ (null hypothesis) and $H_1$ (alternative hypothesis). In this case, $H_0$ may represent the absence of leaks, and $H_1$ may represent the presence of leaks.

Data related to HS parameters such as pressure, gas flow and other relevant variables are collected. Based on the collected data, a statistical model is built, which may include the probability distribution, expected values and variances of the parameters. A statistical criterion is selected that allows you to compare data with null and alternative hypotheses.

The remote leak detection system requires regular monitoring and maintenance to ensure reliable operation. The solution to the problem of specific leak detection may vary depending on the type of HS and the technologies used.

The results of the research will contribute to improving the reliability and safety of HS systems by providing more effective detection and control of leaks and other anomalies in the operation of pipeline systems.

## References

1. V.G. Repin, G.P. Tartakovsky, Statistical synthesis under a priori uncertainty and adaptation of information systems (Sovetskoye radio, Moscow, 1977)
2. I. S. Kvasov, M. Ya. Panov, V. G. Stogney, News of universities. Energy **1(2)**, 76-78 (1995)

3. B. Ch. Meskhi, Yu. I. Bulygin, V. V. Maslensky, I. N. Loskutnikova, Bezopasnost' Truda v Promyshlennosti **2021(2)**, 7–14 (2021)
4. S. Sazonova, V. Asminin, T. Zyazina, D. Sysoev, O. Sokolova, A. Osipov, A. Lemeshkin, AIP Conference Proceedings **3021(1)**, 020028 (2024)
5. I. Yaitskov, A. Chukarin, Akustika **32,** 92-96 (2019)
6. A. Beskopylny, A. Chukarin, P. Kurchenko, A. Isaev, Transportation Research Procedia **63,** 2529-34 (2022)
7. A. Beskopylny, A. Chukarin, B. Meskhi, A. Isaev, Transportation Research Procedia **54,** 39-46 (2021)
8. S. Sazonova, V. Asminin, V. Zherdev, E. Epifanov, A. Venevitin, E. Druzhinina, S. Korablin. AIP Conference Proceedings **3021(1)**, 060013 (2024)