Transaction of Azerbaijan National Academy of Sciences, Series of Physical-Technical and Mathematical Sciences: Informatics and Control Problems, Vol. XXXV, No.3, 2015
www.icp.az/2015/3-01.pdf

T.A. ALIEV, A.H. RZAYEV, G.A. GULUYEV, T.A. ALIZADA, N.E. RZAYEVA

ROBUST NOISE TECHNOLOGY AND SYSTEM FOR OIL WELL SRPU DIAGNOSTICS AND MANAGEMENT

We have analysed noise-induced difficulties in the diagnostics of sucker rod pumping units (SRPUs) related to the peculiarities of their field operation. We have demonstrated that the existing diagnostic methods based on the interpretation of dynamometer cards built from signals received from force and stroke sensors do not allow diagnostic and management problems to be solved in real time. A technology is therefore proposed for determining robust normalized correlation functions, which are used to form the combinations of informative attributes that correspond to the possible emergency states of SRPU. The identification of those states is duplicated by determining and forming the combinations that correspond to the noise characteristics of the force signal, which enhances the reliability of the results. The simplicity of the realization of the processing algorithms makes it possible to solve the problem of force signal identification by means of inexpensive controllers in real-time mode. The application of the technology to more than 100 real objects has demonstrated that the profitability of oil wells increases significantly due to energy savings and an increase in the overhaul period.

Keywords: noise technology, dynamometer card, sucker rod pumping, oil well, noise variance, cross-correlation function

1. Introduction. SRPUs are known as the primary method of oil lifting; currently, sucker rod pumping is a widespread practice in world oil production, covering over two-thirds of the total active well stock [1]. The method is popular due to its simplicity, reliability and applicability to a wide variety of operating conditions.

However, in the face of decreasing oil reserves, increased reservoir flooding and well shutdowns caused by delayed equipment diagnostics, the profitability of oil production by SRPU has decreased considerably. Therefore, improvements in the identification and diagnosis of SRPU faults are a key issue in the problem of oilfield profitability over long-term operation. The timely detection of SRPU faults and implementing preventive measures to eliminate those faults can ensure the necessary stabilisation of oil production. To increase the overhaul period and create the most favourable conditions for oil production management, various methods and tools for SRPU technical control and diagnostics have been developed over the last several decades [2-11]. Those works demonstrate that the information concerning the force on the rod hanger centre contains comprehensive and least-distorted data about the condition of the underground pumping equipment. Therefore, dynamometry, i.e., the reading and analysis of the force curve $U_p(t)$ received from the force sensor in the rod hanger centre $P(S)$, is considered the customary way to control and diagnose SRPU.

The authors of [2] demonstrate the possibility of recognizing force curves $U_p(t)$ with the use of infra-low frequency spectrum analysers. The possibility of obtaining the amplitude spectrum of the dynamometer card is considered an advantage of t his method. For instance, it was revealed that the dynamometer cards of normal operating pumps have no even harmonics, while the dynamometer cards of leaky pumps have even harmonics whose amplitudes are heavily dependent on the scale of the leakage; however, only four types of faults could be recognized using this method [2]. The statistical method was also used for fault recognition by dynamometer cards. This method compares favourably with other methods of separating dynamometer cards into classes due to the small amounts of computation and memory required by its application. The authors of [2, 3, 5-8] provide detailed descriptions of the results of numerous studies that have been carried out in this field over many years. The above-mentioned diagnostic methods have been used at different SRPU control stations at real oilfields for many years, and a scientific foundation has been formed based on these works and various systems for SRPU diagnostics and control by means of dynamometer cards at the well head.

Based on the operational results of those systems, dynamometer card-based diagnostic methods have been categorized as follows [2]:

- Diagnostics based directly on the characteristics of the ground dynamometer card;
- Diagnostics based on the secondary characteristics of the dynamometer card (spectral characteristics such as variance, correlation and regression of the force sensor signal, coefficients of Fourier series expansion for the dynamometer card, etc.);
- Diagnostics based on the typical characteristics of the ground dynamometer card's shape;
- Diagnostics comparing the shape of the dynamometer card under consideration with a reference card selected immediately after well repair and stored in the device memory;
- Diagnostics based on the characteristics of the plunger dynamometer card calculated from the data of the ground dynamometer card and well design; and
- Diagnostics based on the typical characteristics of the shape of the plunger dynamometer card.

The shortcoming of these methods is that they do not allow the automatic identification of a dynamometer card in real time with sufficient ability. For this reason, the identification of dynamometer cards in real life is mostly performed via semi-automatic interpretation, which eventually leads to a visual analysis of the obtained dynamometer information by a technologist, who makes the final decision concerning the presence of a fault in SRPU. The results of the diagnostics depend on the qualifications of the technologist, and running diagnostics on all wells takes some time. Further, even a highly qualified specialist sometimes cannot determine precisely the type of pump fault visually using only dynamometer cards, particularly for deep wells. Therefore, new technologies for the analysis and identification of dynamometer cards with the use of modern controllers in real-time mode must be developed. It is reasonable to ensure the monitoring of the latent period of the origin of SRPU transition into the emergency state by identifying the rod hanger force per pumping cycle signal of the pumping unit. Our research has demonstrated that one of the most efficient ways to solve this problem is to create a hybrid technology for the analysis and identification of signals of hanger force combined with the technologies of correlation and noise analysis [7-10].

2. Problem statement. In the known SRPU control systems, the dynamometer card information comes from force and stroke sensors in the form of the electric signal of force $U_p(t)$ and stroke $U_S(t)$ via the communication channel. Using the combination of these two variables $U_p(t)$ and $U_S(t)$, the dynamometer card $U_p(t) = f(U_S)$, whose form is described by a parallelogram (Fig. 1), is formed. A skilled technologist can identify more than 20 types of SRPU faults through a visual analysis of distortions in different sections of its shape [2]. Performing this operation for a hundred wells, however, is a complicated task.

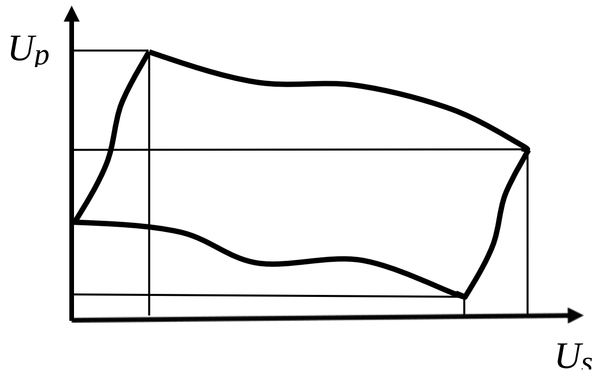

Fig. 1 Typical dynamometer card of a normally operating SRPU

It should be noted that in the case that an object's state is identified through a hardware implementation, there is no need to use $U_S(t)$, as SRPU can be diagnosed by analysing the stress curve $U_p(t)$ only. The main challenge in this problem is currently associated with the lack of a technology that can ensure the adequate identification of the force signal $U_p(t)$ in real time. For instance, when the correlation analysis technology is used for this purpose, the condition of robustness is not satisfied because the error of the obtained estimates of the correlation functions caused by the effects of the noise $\varepsilon_1(t)$ accompanying the useful force signal $U_p(t)$ under operation changes across a rather broad range. This is related to the fact that the control object, i.e., SRPU, operates in the field environment (subject to temperature and humidity extremes, etc.). Additionally, during pump operation, various faults can also cause the formation of a noise $\varepsilon_2(t)$ that correlates with the signal $U_p(t)$ [12]. Therefore, the noise $\varepsilon(t)$ accompanying the force signal $U_p(t)$ forms under the influence of the following two factors:

$$\varepsilon(t) = \varepsilon_1(t) + \varepsilon_2(t)$$

$\varepsilon_1(t)$ forms due to the changes in the environment (temperature and humidity differences, etc.); $\varepsilon_2(t)$ forms during the operation of the object due to the changes in the technical condition of the mechanical components of the pump, such as wear and tear, bends, cracks, fatigue, etc.

Thus, during operation, the signal contaminated with the noise $\varepsilon(t)$ is input into the system instead of the signal $U_p(t)$. The analysed signal looks like the following in the analogue form:

$$g(t) = U_p(t) + \varepsilon(t),$$

and like the following in the digital form:

$$g(i\Delta t) = U_p(i\Delta t) + \varepsilon(i\Delta t). \tag{2.1}$$

For the above reasons, both the amplitude and the spectrum of the noise $\varepsilon(i\Delta t)$ vary across a fairly wide range. For the same reasons, the errors of the obtained estimates of the correlation functions $R_{gg}(i\Delta t)$ of the measuring signal $g(i\Delta t) = U_p(i\Delta t) + \varepsilon(i\Delta t)$ also vary across a wide time range. Therefore, we fail to provide the condition of robustness for the estimates of the correlation function in real-time mode, i.e., to rule out the dependence of the obtained results on the effects of the noise $\varepsilon(i\Delta t)$. This, first of all, complicates solving the problem of the identification of the dynamometer card using correlation methods. Consequently, ensuring an adequate identification requires that the conditions of robustness hold, i.e., the elimination of the effects of the said factors on the error of the estimates $R_{gg}(i\Delta t)$. To that end, it is appropriate to first reduce the estimates

$R_{gg}(i\Delta t)$ to a single dimensionless value by applying the normalization procedure [2, 12-14]. Our analysis, however, demonstrates that the application of traditional methods introduces additional error in the normalized estimates of the correlation functions $r_{gg}(i\Delta t)$, which, in turn, complicates the attempt to ensure adequate results for the above-mentioned problems. This issue will thus be considered in greater detail in the following paragraphs.

It is known that normalized correlation functions of the useful signal $U_p(i\Delta t)$ are calculated from the following formula [12-14]:

$$r_{UU}(\mu) = R_{UU}(\mu)/D_U, \tag{2.2}$$

where the estimate of the variance $R_{UU}(\mu)$ at $\mu = 0$ is determined from the expression:

$$R_{UU}(\mu = 0) = D_U = 1/N\sum_{i=1}^{N} U(i\Delta t)U(i\Delta t). \tag{2.3}$$

The estimates of the correlation function $R_{UU}(\mu)$ of the useful signal $U_p(i\Delta t)$ at $\mu \neq 0$ are determined from the formula:

$$R_{UU}(\mu) = 1/N\sum_{i=1}^{N} U(i\Delta t)U((i+\mu)\Delta t),\ \ \mu = 0,1,2,3... \tag{2.4}$$

It is also known that the estimates of the normalized correlation functions $r_{gg}(\mu)$ of the noisy signal $g(i\Delta t)$, consisting of the force signal $U_p(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$, are calculated from the formula:

$$r_{gg}(\mu) = R_{gg}(\mu)/D_g, \tag{2.5}$$

Where:

$$R_{gg}(\mu = 0) = 1/N\sum_{i=1}^{N} g(i\Delta t)g(i\Delta t) = 1/N\sum_{1}^{N}[U(i\Delta t)+\varepsilon(i\Delta t)][U(i\Delta t)+\varepsilon(i\Delta t)] = 1/N\sum_{i=1}^{N}[U(i\Delta t)U(i\Delta t)+$$
$$+[U(i\Delta t)\varepsilon(i\Delta t)+\varepsilon(i\Delta t)U(i\Delta t)+\varepsilon(i\Delta t)\varepsilon(i\Delta t)]] \tag{2.6}$$

Expression (2.6) clearly shows that the equality $R_{gg}(\mu = 0) = D_g$ holds only at $\mu = 0$, as the correlation of $U_p(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$ disappears at $\mu \neq 0$, i.e.:

$$1/N\sum_{i=1}^{N}[U(i\Delta t)\varepsilon(i\Delta t)+\varepsilon(i\Delta t)U(i\Delta t)+\varepsilon(i\Delta t)\varepsilon(i\Delta t)] \approx 0.$$

Therefore, the following equality will hold for estimates $R_{gg}(\mu)$ at $\mu \neq 0$:

$$R_{gg}(\mu) = R_{UU}(\mu).$$

For this reason, when formula (2.5) is applied, the correct result of the division of $R_{gg}(\mu)$ by D_g is obtained only at $\mu = 0$. For all other cases, when $\mu \neq 0$, an additional error is introduced into the result, which is one of the factors that violates the adequacy of the identification results.

To eliminate the stated error in the traditional normalization procedure, estimates $r_{gg}(\mu \neq 0)$ should be normalized by means of the following expressions:

$$r_{gg}(\mu) = R_{gg}(\mu)/R_{UU}(\mu = 0). \tag{2.7}$$

As noted earlier, however, it is practically impossible to determine estimate $R_{UU}(\mu = 0)$ due to the influence of noise $\varepsilon(i\Delta t)$.

Thus, the problem we set in this paper is developing technologies that attempt to eliminate the effects of noise $\varepsilon(i\Delta t)$ on the estimates of the normalized correlation functions $r_{gg}(\mu \neq 0)$ to ensure the adequate identification of the technical condition of SRPU by means of correlation functions. We must also obtain estimates of the normalized correlation functions $r^R{}_{gg}(\mu)$ of the noisy signals $g(i\Delta t)$ that are sufficiently robust to ensure that the following equality holds:

$$r^R{}_{gg}(\mu) \approx r_{UU}(\mu).$$

As was indicated earlier, when the latent period of fault origin begins in SRPU, a correlation between the useful signal $U_p(i\Delta t)$ and the noise $\varepsilon(i\Delta t) = \varepsilon_1(i\Delta t) + \varepsilon_2(i\Delta t)$ arises. Our experimental research has demonstrated that the effects of the noise $\varepsilon(i\Delta t)$ on the estimates of correlation functions are practically non-existent, but only in the normal state of SRPU. With different types of SRPU faults because of the influence of the noise the estimates R_{gg} change; in other words, a certain range of the noise variance D_ε, cross-correlation function $R_{U\varepsilon}(\mu = 0)$ between the force signal $U_p(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$ and the estimate of the sum noise $D_{\varepsilon\!\varepsilon}$ corresponds to every possible fault.

3. Technology for correcting the errors of normalization of correlation functions. As noted earlier, calculating the estimates of normalized correlation functions is associated with certain difficulties.

Considering that the error of the estimate $R_{gg}(\mu = 0)$ at $(\mu = 0)$ is affected by the presence of a correlation between the useful signal $U_p(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$, as well as formula (2.1) and equality (2.6) can be written as follows:

$$R_{gg}(\mu = 0) = \sum_{i=1}^{N} g(i\Delta t) g(i\Delta t) = 1/N \sum_{i=1}^{N} U^2(i\Delta t) + 2 \cdot 1/N \sum_{i=1}^{N} U(i\Delta t)\varepsilon(i\Delta t) + 1/N \sum_{i=1}^{N} \varepsilon^2(i\Delta t) =$$
$$= R_{UU}(\mu = 0) + 2R_{U\varepsilon}(\mu = 0) + D_\varepsilon \tag{3.1}$$

The estimate of the variance D_ε of the noise $\varepsilon(i\Delta t)$, [12-14] is determined from the formula:

$$D_\varepsilon = 1/N \sum_{i=1}^{N} \varepsilon(i\Delta t)\varepsilon(i\Delta t). \tag{3.2}$$

Therefore, considering equalities (3.1) and (3.2), the formula for calculating $R_{UU}(\mu = 0)$ can be represented as follows:

$$R_{UU}(\mu = 0) = R_{gg}(\mu = 0) - [2R_{U\varepsilon}(\mu = 0) + D_\varepsilon] = D_g - D_{\varepsilon\!\varepsilon}. \tag{3.3}$$

From expression (3.3), it follows that in the presence of a correlation between the useful signal and the noise, the estimate of the autocorrelation function $R_{gg}(\mu = 0)$ contains an error equal to the value of the sum variance $D_{\varepsilon\!\varepsilon}$, i.e.:

$$D_{\varepsilon\!\varepsilon} = 2R_{U\varepsilon}(\mu = 0) + D_\varepsilon. \tag{3.4}$$

At $(\mu \neq 0)$, expression (2.6) can also be represented in the following form:

$$R_{gg}(\mu \neq 0) = 1/N \sum_{i=1}^{N} g(i\Delta t) g((i+\mu)\Delta t) = 1/N \sum_{i=1}^{N} [U(i\Delta t) + \varepsilon(i\Delta t)][U(i+\mu)\Delta t + \varepsilon(i+\mu)\Delta t] =$$
$$= 1/N \sum_{i=1}^{N} U(i\Delta t)U((i+\mu)\Delta t) + 1/N \sum_{i=1}^{N} U(i\Delta t)\varepsilon((i+\mu)\Delta t) + 1/N \sum_{i=1}^{N} \varepsilon(i\Delta t)U((i+\mu)\Delta t) + \tag{3.5}$$
$$+ 1/N \sum_{i=1}^{N} \varepsilon(i\Delta t)\varepsilon((i+\mu)\Delta t).$$

As the following equalities hold at $\mu \neq 0$:

$$\begin{cases} R_{U\varepsilon}(\mu)1/N\sum_{i=1}^{N}U(i\Delta t)\varepsilon(i+\mu)\Delta t \approx 0 \\ R_{\varepsilon U}(\mu)1/N\sum_{i=1}^{N}\varepsilon(i\Delta t)U(i+\mu)\Delta t \approx 0 \\ R_{\varepsilon\varepsilon}(\mu)1/N\sum_{i=1}^{N}\varepsilon(i\Delta t)\varepsilon(i+\mu)\Delta t \approx 0 \end{cases} \tag{3.6}$$

the following can be written:

$$R_{gg}(\mu \neq 0) = 1/N\sum_{i=1}^{N}U(i\Delta t)U(i+\mu)\Delta t = R_{UU}(\mu \neq 0). \tag{3.7}$$

Therefore, considering (3.3)-(3.7), formula (2.7) for calculating the estimate of the normalized autocorrelation function of the noisy signal $g(i\Delta t)$ can be represented as follows:

$$r_{gg}^{R}(\mu = 0) = \frac{R_{gg}(\mu=0)}{D_g} = 1 \tag{3.8}$$

$$r_{gg}^{R}(\mu \neq 0) = \frac{R_{gg}(\mu\neq 0)}{R_{UU}(\mu=0)} \approx \frac{R_{gg}(\mu\neq 0)}{D_g - 2R_{U\varepsilon}(\mu=0) - D_\varepsilon}. \tag{3.9}$$

Thus, the additional error of the normalized correlation functions can be eliminated using expressions (3.8) and (3.9) if one can estimate the noise variance D_ε and the correlation function $R_{U\varepsilon}(\mu = 0)$ of the useful signal $U(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$.

4. Algorithm for calculating the estimates of the noise variance and the cross-correlation function between the useful signal and the noise. As shown above, the problem of identifying dynamometer cards can be solved by using the estimates of the noise variance D_ε and cross-correlation function $R_{U\varepsilon}(\mu)$ between the useful signal, the noise and the variance of the sum noise $D_{\varepsilon\varepsilon}$. Therefore, let us consider one possible way to determine these estimates.

We know that the formula for calculating the estimate of the autocorrelation function $R_{gg}(\mu)$ of the centred sampled random signal $g(i\Delta t)$ with allowance for the effects of the noise $\varepsilon(i\Delta t)$ can be represented in the following form:

$$\begin{aligned} R_{gg}(\mu) &= 1/N\sum_{i=1}^{N}g(i\Delta t)g((i+\mu)\Delta t) = 1/N\sum_{i=1}^{N}[U(i\Delta t)+\varepsilon(i\Delta t)][U((i+\mu)\Delta t)+\varepsilon((i+\mu)\Delta t)] = \\ &= 1/N\sum_{i=1}^{N}[U(i\Delta t)U((i+\mu)\Delta t)+U(i\Delta t)\varepsilon((i+\mu)\Delta t)+\varepsilon(i\Delta t)U((i+\mu)\Delta t)+\varepsilon(i\Delta t)\varepsilon((i+\mu)\Delta t)] = \\ &= R_{UU}(\mu)+\lambda_1(\mu) \end{aligned} \tag{4.1}$$

Where:

$$R_{UU}(\mu) = 1/N\sum_{i=1}^{N}U(i\Delta t)U((i+\mu)\Delta t). \tag{4.2}$$

We also know that when the samples of the useful signal $U(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$ and the samples of the noise $\varepsilon(i\Delta t)$ and $\varepsilon((i+\mu)\Delta t)$, do not correlate the equalities (3.6) hold.

Taking into account (3.2) and (3.6), equality (4.1) can be written as follows:

$$R_{gg}(\mu=0) = 1/N\sum_{i=1}^{N}[U(i\Delta t)U(i\Delta t)+U(i\Delta t)\varepsilon(i\Delta t)+\varepsilon(i\Delta t)U(i\Delta t)]+D_\varepsilon \approx R_{UU}(\mu=0)+D_\varepsilon \tag{4.3}$$

$$R_{gg}(\mu \neq 0) = 1/N\sum_{i=1}^{N}[U(i\Delta t)U((i+\mu)\Delta t) + U(i\Delta t)U((i+\mu)\Delta t)] + \varepsilon(i\Delta t)U((i+\mu)\Delta t) \approx R_{UU}(\mu \neq 0)$$

It is also a fact [12-15] that if estimates $R_{UU}(\mu = 0)$, $R_{UU}(\mu = 1)$, and $R_{UU}(\mu = 2)$ are determined from expression (4.2), their differences are close values:

$$\lim_{\substack{T\to\infty \\ \Delta t\to 0}} R_{UU}(\mu = 0) - R_{UU}(\mu = 1) \approx \lim_{\substack{T\to\infty \\ \Delta t\to 0}} R_{UU}(\mu = 1) - R_{UU}(\mu = 2). \tag{4.4}$$

Therefore, the following equality can be regarded as true:

$$R_{UU}(\mu = 0) - R_{UU}(\mu = 1) \approx R_{UU}(\mu = 1) - R_{UU}(\mu = 2). \tag{4.5}$$

Hence, the following can be written:

$$R_{UU}(\mu = 0) \approx 2R_{UU}(\mu = 1) - R_{UU}(\mu = 2). \tag{4.6}$$

Considering the equalities:

$$\left.\begin{aligned} R_{gg}(\mu = 1) \approx R_{UU}(\mu = 1) \\ R_{gg}(\mu = 2) \approx R_{UU}(\mu = 2) \end{aligned}\right\}, \tag{4.7}$$

and (4.1), (4.3), (4.5)-(4.7), the following will hold:

$$D_{æ} \approx R_{gg}(\mu = 0) + R_{gg}(\mu = 2) - 2R_{gg}(\mu = 1). \tag{4.8}$$

In other words, if conditions (3.6), (4.4)–(4.7) are satisfied, the expression for determining the estimate of the sum noise $D_{æ}$ can be represented as [7,12,15]:

$$D_{æ} \approx 1/N\sum_{i=1}^{N}[g^2(i\Delta t) + g(i\Delta t)g((i+2)\Delta t) - 2g(i\Delta t)g((i+1)\Delta t)]. \tag{4.9}$$

It is clear that when estimate $D_{æ}$ is known, the formula for calculating the variance of the useful signal can be written in the following form:

$$D_U = D_g - D_{æ}.$$

The estimate $R_{gg}(\mu = 0)$ in expression (4.1) can be represented as follows:

$$R_{gg}(\mu = 0) = 1/N\sum_{i=1}^{N}[U(i\Delta t) + \varepsilon(i\Delta t)]^2 = 1/N\sum_{i=1}^{N}U^2(i\Delta t) + 1/N\sum_{i=1}^{N}2[U(i\Delta t)\varepsilon(i\Delta t)] +$$
$$+1/N\sum_{i=1}^{N}\varepsilon^2(i\Delta t) \tag{4.10}$$

If we assume the following notations:

$$1/N\sum_{i=1}^{N}U^2(i\Delta t) = R_{UU}(\mu = 0),$$

$$1/N\sum_{i=1}^{N}2[U(i\Delta t)\varepsilon(i\Delta t)] = 2R_{U\varepsilon}(\mu = 0),$$

$$1/N\sum_{i=1}^{N}\varepsilon^2(i\Delta t) = R_{æ}(\mu = 0) = D_{\varepsilon},$$

expression (4.10) will take the following form:

$$R_{gg}(\mu = 0) = R_{UU}(\mu = 0) + 2R_{U\varepsilon}(\mu = 0) + R_{æ}(\mu = 0).$$

Hence, expression (4.8) can be transformed as follows:

$$D_{æ} \approx R_{gg}(\mu = 2) - 2R_{gg}(\mu = 1) + R_{UU}(\mu = 0) + 2R_{U\varepsilon}(\mu = 0) + R_{æ}(\mu = 0).$$

Considering that:

$$R_{gg}(\mu = 1) = R_{UU}(\mu = 1),$$

$$R_{gg}(\mu=2)=R_{UU}(\mu=2),$$

$$R_{UU}(\mu=0)+R_{UU}(\mu=2)=2R_{UU}(\mu=1)$$

we can assume that the following equality is true:

$$R_{gg}(\mu=2)-2R_{gg}(\mu=1)+R_{UU}(\mu=0)=0.$$

Therefore, formulas (4.8) and (4.9) for the calculation of $D_{æ}$ can be transformed as follows:

$$D_{æ}\approx 2R_{U\varepsilon}(\mu=0)+R_{æ}(\mu=0)=2R_{U\varepsilon}(\mu=0)+D_{\varepsilon}. \tag{4.11}$$

$$D_{\varepsilon}=D_{æ}-2R_{U\varepsilon}(\mu=0).$$

Estimate $D_{æ}$ can be calculated from expression (4.9); however, the use of this formula does not allow us to determine estimates $R_{U\varepsilon}(\mu=0)$ and D_{ε}. In this regard, let us consider in more detail the possibility of determining it by means of the technologies for calculating the estimates of the relay correlation function $R^{*}_{gg}(\mu=0)$. To that end, we assume the following notations:

$$\operatorname{sgn} g(i\Delta t)=\operatorname{sgn} U(i\Delta t)=\begin{cases}+1 \ at \ g(i\Delta t)>0\\ 0 \ at \ g(i\Delta t)=0\\ -1 \ at \ g(i\Delta t)<0\end{cases} \tag{4.12}$$

and the validity of the equalities:

$$\begin{cases}1/N\sum_{i=1}^{N}\operatorname{sgn}\varepsilon(i\Delta t)\varepsilon((i+\mu)\Delta t)=0 \ at \ \mu\neq 0\\ 1/N\sum_{i=1}^{N}\operatorname{sgn}\varepsilon(i\Delta t)\varepsilon(i\Delta t)\neq 0 \ at \ \mu=0\end{cases} \tag{4.13}$$

Thus, the formula for determining the estimates of the relay correlation function $R^{*}_{gg}(\mu=0)$ will be represented as follows:

$$\begin{aligned}R^{*}_{gg}(\mu=0)&=1/N\sum_{i=1}^{N}\operatorname{sgn} g(i\Delta t)g(i\Delta t)=1/N\sum_{i=1}^{N}\operatorname{sgn} g(i\Delta t)\cdot[U(i\Delta t)+\varepsilon(i\Delta t)]=\\ &=1/N\sum_{i=1}^{N}\operatorname{sgn} g(i\Delta t)\cdot[U(i\Delta t)+\varepsilon(i\Delta t)]=1/N\sum_{i=1}^{N}[[\operatorname{sgn} g(i\Delta t)\cdot U(i\Delta t)]+[\operatorname{sgn}\ g(i\Delta t)\cdot\varepsilon(i\Delta t)]]=\\ &=1/N\sum_{i=1}^{N}\operatorname{sgn} g(i\Delta t)U(i\Delta t)+1/N\sum_{i=1}^{N}\operatorname{sgn} g(i\Delta t)\cdot\varepsilon(i\Delta t)=1/N\sum_{i=1}^{N}\operatorname{sgn}U(i\Delta t)U(i\Delta t)+1/N\sum_{i=1}^{N}\operatorname{sgn}U(i\Delta t)\varepsilon(i\Delta t)=\\ &=R^{*}{}_{UU}(\mu=0)+R^{*}{}_{U\varepsilon}(\mu=0)\end{aligned}$$

$$R^{*}{}_{gg}(\mu=0)=1/N\sum_{i=1}^{N}\operatorname{sgn} g(i\Delta t)g(i\Delta t)=1/N\sum_{i=1}^{N}\operatorname{sgn} g(i\Delta t)[U(i\Delta t)+\varepsilon(i\Delta t)]. \tag{4.14}$$

As we know [7,12,15] in the absence of a correlation between $U(i\Delta t)$ and $\varepsilon(i\Delta t)$ under the conditions (4.12), (4.13) the following approximate equations can be regarded as true for the estimates of the relay correlation function:

$$\begin{aligned}&R^{*}{}_{gg}(\mu=0)-R^{*}{}_{gg}(\mu=1)\approx R^{*}{}_{gg}(\mu=1)-R^{*}{}_{gg}(\mu=2)\approx R^{*}{}_{gg}(\mu=2)-R^{*}{}_{gg}(\mu=3)\approx\\ &\approx R^{*}{}_{gg}(\mu=3)-R^{*}{}_{gg}(\mu=4)\end{aligned} ,\tag{4.15}$$

$$\begin{aligned}&R^{*}{}_{UU}(\mu=0)-R^{*}{}_{UU}(\mu=1)\approx R^{*}{}_{UU}(\mu=1)-R^{*}{}_{UU}(\mu=2)\approx R^{*}{}_{UU}(\mu=2)-R^{*}{}_{UU}(\mu=3)\approx\\ &\approx R^{*}{}_{UU}(\mu=3)-R^{*}{}_{UU}(\mu=4)\end{aligned} ,\tag{4.16}$$

$$\Delta R^{*}{}_{gg}(\mu=0)\approx\Delta R^{*}{}_{gg}(\mu=1)\approx\Delta R^{*}{}_{gg}(\mu=2)\approx\Delta R^{*}{}_{gg}(\mu=3), \tag{4.17}$$

$$\Delta R^*_{UU}(\mu=0) \approx \Delta R^*_{UU}(\mu=1) \approx \Delta R^*_{UU}(\mu=2) \approx \Delta R^*_{UU}(\mu=3)\,. \qquad (4.18)$$

In the presence of a correlation between $U(i\Delta t)$ and $\varepsilon(i\Delta t)$, the following expressions can be regarded as true:

$$\Delta R^*_{gg}(\mu=0) - \Delta R^*_{gg}(\mu=1) \neq \Delta R^*_{gg}(\mu=1) - \Delta R^*_{gg}(\mu=2), \qquad (4.19)$$

$$\left.\begin{array}{l} \Delta R^*_{gg}(\mu=1) - \Delta R^*_{gg}(\mu=2) \approx \Delta R^*_{gg}(\mu=2) - \Delta R^*_{gg}(\mu=3) \approx \Delta R^*_{gg}(\mu=3) - \Delta R^*_{gg}(\mu=4) \approx 0 \\ \Delta R^*_{UU}(\mu=1) - \Delta R^*_{UU}(\mu=2) \approx \Delta R^*_{UU}(\mu=2) - \Delta R^*_{UU}(\mu=3) \approx \Delta R^*_{UU}(\mu=3) - \Delta R^*_{UU}(\mu=4) \approx 0 \end{array}\right\}$$

From equation (4.14), it follows that the estimate of the relay cross-correlation function $\Delta R^*_{U\varepsilon}(\mu=0)$ can be calculated from the formula:

$$\Delta R^*_{gg}(\mu=0) \approx \Delta R^*_{UU}(\mu=0) + \Delta R^*_{U\varepsilon}(\mu=1)\,,$$

$$\Delta R^*_{U\varepsilon}(\mu=0) - R^*_{gg}(\mu=0) - R^*_{UU}(\mu=0)\,. \qquad (4.20)$$

Hence, calculating $\Delta R^*_{U\varepsilon}(\mu=0)$ from expression (4.20) requires that $R^*_{UU}(\mu=0)$ be determined. It follows from equalities (4.14)–(4.19) that estimate $R^*_{UU}(\mu=0)$ can be calculated by the following expression:

$$R^*_{UU}(\mu=0) \approx R^*_{UU}(\mu=1) + \Delta R^*_{UU}(\mu=1) \approx R^*_{gg}(\mu=1) + \Delta R^*_{gg}(\mu=1) \approx R^*_{gg}(\mu=1) +$$
$$+[R^*_{gg}(\mu=1) - R^*_{gg}(\mu=2)] = 2R^*_{gg}(\mu=1) - R^*_{gg}(\mu=2)$$

Thus, expression (4.20) can be represented as follows:

$$R^*_{U\varepsilon}(\mu=0) = R^*_{gg}(\mu=0) - [2R^*_{gg}(\mu=0) - R^*_{gg}(\mu=2)] = R^*_{gg}(\mu=0) - 2R^*_{gg}(\mu=1) + R^*_{gg}(\mu=0)\,.$$

Therefore, the expression for calculating the estimate of the relay cross-correlation function $R^*_{U\varepsilon}(\mu=0)$ between the useful signal $U(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$ can be represented as follows:

$$R^*_{U\varepsilon}(\mu=0) \approx 1/N\sum_{i=1}^{N}[\mathrm{sgn}\,g(i\Delta t)g(i\Delta t) - 2\,\mathrm{sgn}\,g(i\Delta t)g(i+1)\Delta t) + \mathrm{sgn}\,g(i\Delta t)g(i+2)\Delta t)]\,. \qquad (4.21)$$

As we can see, with estimate $R^*_{U\varepsilon}(\mu=0)$ available, it is quite easy to calculate estimate $R_{U\varepsilon}(\mu=0)$. We know from [12-15] that knowing the estimates $R^*_{U\varepsilon}(\mu=0)$, $R^*_{gg}(\mu=1)$ and $R_{gg}(\mu=1)$ and using the relationship between $R^*_{gg}(\mu=1)$ and $R_{gg}(\mu=1)$, we can assume that the following equality holds:

$$R^*_{gg}(\mu=1)/R_{gg}(\mu=1) = R^*_{U\varepsilon}(\mu=0)/R_{U\varepsilon}(\mu=0)\,,$$

Hence, estimate $R_{U\varepsilon}(\mu=0)$ can be calculated by the formula:

$$R_{U\varepsilon}(\mu=0) = R_{gg}(\mu=1)R^*_{U\varepsilon}(\mu=0)/R^*_{gg}(\mu=1) \qquad (4.22)$$

It is obvious that with a knowledge of estimate $R_{U\varepsilon}(\mu=0)$, estimate D_ε can be determined from the expression (4.11):

$$D_\varepsilon = D_{æ} - 2R_{U\varepsilon}(\mu=0)\,.$$

Thus, estimates D_ε and $R_{U\varepsilon}(\mu=0)$ are determined from formulas (4.11) and (4.22) based on estimate $D_{æ}$ from formula (4.9), $R_{gg}(\mu=1)$ from formula (4.1), $R^*_{gg}(\mu=1)$ from formula (4.14) and $R^*_{U\varepsilon}(\mu=0)$ from formula (4.21).

By inserting the determined values of the noise variance D_ε and correlation function $R_{U\varepsilon}(\mu=0)$ of the useful signal and the noise in expressions (3.8) and (3.9), we will obtain the robust normalized estimates of the autocorrelation function $r^R_{gg}(\mu)$ of the noisy signal $g(i\Delta t)$.

5. Robust normalisation procedure for estimating correlation functions. The technology for calculating estimates $D_{æ}$, D_{ε} , $R^{*}{}_{U\varepsilon}(\mu=0)$ and $R_{U\varepsilon}(\mu=0)$ makes it possible to ensure the robustness of the estimates of the normalized correlation functions by eliminating the additional noise-induced error.

Proposed below is the appropriate sequence of calculations, the totality of which makes up the procedure of normalizing correlation functions.

1) The estimates of the autocorrelation function of the noise signal are calculated from expression (4.1):

$$R_{gg}(\mu)=1/N\sum_{i=1}^{N}g(i\Delta t)g((i+1)\Delta t$$

2) Based on expression (4.14), the estimate of the relay autocorrelation function of the noisy signal at $\mu=1$ is calculated:

$$R^{*}{}_{gg}(\mu=1)=1/N\sum_{i=1}^{N}g(i\Delta t)g((i+1)\Delta t$$

3) The estimate of the relay cross-correlation function between the useful signal and the noise at $\mu=0$ is calculated from expression (4.21):

$$R^{*}{}_{U\varepsilon}(\mu=0)\approx 1/N\sum_{i=1}^{N}[\operatorname{sgn} g(i\Delta t)g(i\Delta t)-2\operatorname{sgn} g(i\Delta t)g((i+1)\Delta t+\operatorname{sgn} g(i\Delta t)g((i+2)\Delta t].$$

4) The estimate of the cross-correlation function between the useful signal and the noise at $\mu=0$ is calculated from expression (4.22):

$$R_{U\varepsilon}(\mu=0)=R_{gg}(\mu=1)R^{*}{}_{U\varepsilon}(\mu=0)/R^{*}{}_{gg}(\mu=1)$$

5) The estimate of the variance of the sum noise $D_{æ}$ and the variance of the noise D_{ε} are calculated from expressions (4.9) and (4.11):

$$D_{æ}\approx 1/N\sum_{i=1}^{N}[g^{2}(i\Delta t)+g(i\Delta t)g((i+2)\Delta t)-2g(i\Delta t)g((i+1)\Delta t)],$$

$$D_{\varepsilon}=D_{æ}-2R_{U\varepsilon}(\mu=0).$$

6) The robust estimates of the normalized autocorrelation function of the noisy signal at $\mu=0,1,2,3,...$ are calculated from expressions (3.8) and (3.9):

$$r^{R}{}_{gg}(\mu)=\begin{cases}R_{gg}(\mu=0)/D_{g}=1 & at\ \mu=0\\ R_{gg}(\mu)/D_{g}-2R_{U\varepsilon}(0)-D_{\varepsilon} & at\ \mu\neq 0\end{cases}$$

Thus, the proposed procedure allows us to calculate the robust estimates of the normalized autocorrelation function $r^{R}{}_{gg}(\mu)$ of the noisy signal $g(i\Delta t)$ such that it is free from noise-induced errors.

The simplicity, accessibility and authenticity of the proposed normalization procedure based on the technology described above are demonstrated below through a computing experiment.

1. Assume that condition $D_{æ}\approx 0.1D_{U}$ holds for the noisy signal $g(i\Delta t)$ (Fig.2, blue graph) consisting of the useful component in the form of a sinusoid
$U(i\Delta t)=A\sin(i\Delta t)$ (Fig. 2, green graph) and noise $\varepsilon(i\Delta t)$ (Fig. 2, red graph).

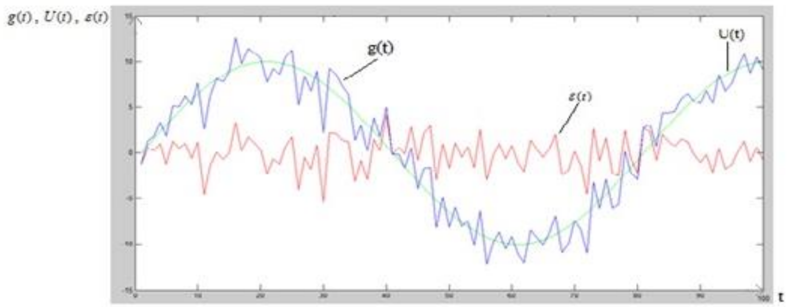

Fig. 2 Behaviour of the useful signal $U(i\Delta t)$, the noisy signal $g(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$ in time

Fig. 2 shows the first 100 samples of the signals $g(i\Delta t)(A = 10;\ \mathrm{t = 0 : pi/40 : 120 * pi - pi/40;})$, $U(i\Delta t)$, $\varepsilon(i\Delta t)$ $(\varepsilon = normrnd(0,2,1,4800))$.

2. Using formulas (3.1) and (3.2):

$$D_g = 1/N\sum_{i=1}^{N} g(i\Delta t)g(i\Delta t)$$

$$D_\varepsilon = 1/N\sum_{i=1}^{N} \varepsilon(i\Delta t)\varepsilon(i\Delta t)$$

we calculate the estimates of the variance of the noisy signal $D_g = 55.229$ and the noise variance $D_\varepsilon = 3.9064$.

3. The estimate $D_{æ} = 3.8508$ is determined from expression (4.9).

4. Using formulas (2.2), (2.3) and (2.4), we carry out the normalization procedure:

$$r_{UU}(\mu) = R_{UU}(\mu)/D_U$$

and plot its relationship with $\mu = 0,1,2,3,\ldots$ (Fig. 3, green graph).

5. Using formula (2.5), we carry out the traditional normalization procedure:

$$r_{gg}(\mu) = R_{gg}(\mu)/D_g$$

and create its plot $r_{gg}(\mu)$ (Fig. 3, red graph).

6. Using formula (3.9), with allowance for (4.11), we carry out the proposed normalization procedure:

$$r^R{}_{gg}(\mu) = R_{gg}(\mu \neq 0)/D_g - 2R_{U\varepsilon}(\mu = 0) - D_\varepsilon = R_{gg}(\mu)/D_g - D_{æ}$$

and create its plot (Fig. 3, blue graph).

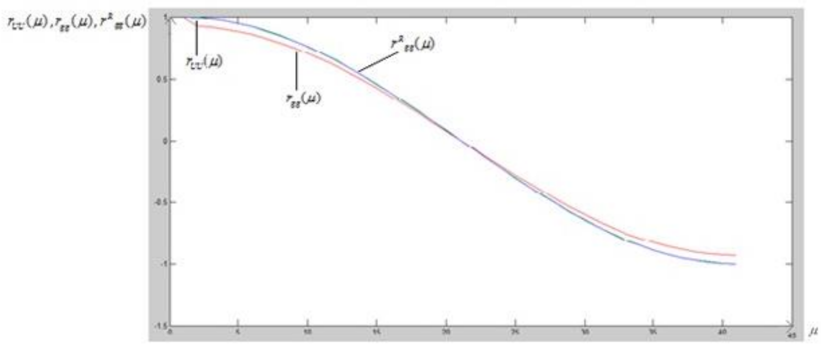

Fig. 3 Charts of the normalized correlation functions of the noisy signal $r_{gg}(\mu)$ and the useful signal $r_{UU}(\mu)$, as well as the robust correlation function of the noisy signal $r^R{}_{gg}(\mu)$

Table 1 shows the results of the obtained estimates $r_{UU}(\mu)$, $r_{gg}(\mu)$, $r^R{}_{gg}(\mu)$ in relation to the value of μ.

Table 1

The results of the estimates $r_{UU}(\mu)$, $r_{gg}(\mu)$, $r^R{}_{gg}(\mu)$ in relation to the value of μ

μ	$r_{UU}(\mu)$	$r_{VV}(\mu)$	$r^R_{VV}(\mu)$	μ	$r_{UU}(\mu)$	$r_{VV}(\mu)$	$r^R_{VV}(\mu)$
0	1	1	1	21	-0.0731	-0.0663	-0.0707
1	0.9973	0.9261	0.9873	22	-0.1512	-0.1395	-0.1488
2	0.9885	0.9145	0.9749	23	-0.2282	-0.2111	-0.2250
3	0.9736	0.9027	0.9623	24	-0.3039	-0.2807	-0.2992
4	0.9527	0.8827	0.9410	25	-0.3777	-0.3507	-0.3739
5	0.9259	0.8549	0.9114	26	-0.4492	-0.4158	-0.4433
6	0.8934	0.8266	0.8812	27	-0.5179	-0.4775	-0.5090
7	0.8554	0.7913	0.8436	28	-0.5835	-0.5378	-0.5733
8	0.8121	0.7495	0.7990	29	-0.6454	-0.5961	-0.6355
9	0.7639	0.7096	0.7564	30	-0.7033	-0.6499	-0.6929
10	0.7109	0.6575	0.7009	31	-0.7569	-0.7017	-0.7480
11	0.6535	0.6053	0.6452	32	-0.8059	-0.7465	-0.7958
12	0.5921	0.5484	0.5846	33	-0.8498	-0.7843	-0.8361
13	0.5270	0.4872	0.5194	34	-0.8886	-0.8220	-0.8763
14	0.4587	0.4232	0.4511	35	-0.9218	-0.8524	-0.9087
15	0.3876	0.3575	0.3811	36	-0.9494	-0.8786	-0.9366
16	0.3141	0.2897	0.3089	37	-0.9711	-0.8963	-0.9555
17	0.2386	0.2198	0.2343	38	-0.9868	-0.9121	-0.9724
18	0.1617	0.1509	0.1609	39	-0.9965	-0.9202	-0.9810
19	0.0838	0.0784	0.0836	40	-1.0000	-0.9252	-0.9863
20	0.0053	0.0062	0.0066				

The obtained results demonstrate that the graphs of $r_{UU}(\mu)$ and $r^R{}_{gg}(\mu)$ practically coincide, while the graph of $r_{gg}(\mu)$ shows a significant deviation from the true values of $r_{UU}(\mu)$ due to the non-robustness of the traditional normalization procedure. Thus, the suitability of the practical application of the proposed normalization procedure is obvious.

In the appendix, we provide a listing of the experiment conducted in the MATLAB computing environment. The charts in Fig. 3 and the data in Table 1 are the results of that experiment.

6. Determining the informative attributes for the diagnosis of the technical condition of SRPU. Our experiments on real oilfield objects have demonstrated that it is possible to form up to 12 informative attributes using the normalized correlation functions of the force signal $U_p(i\Delta t)$, which are determined from expressions (3.8) and (3.9). Moreover, it has been found that the informative attributes obtained by means of the noise technology from expressions (4.9), (4.11) and (4.22) also allow us to identify the technical condition of SRPU. Below, we provide the procedure for determining the informative attributes by means of the normalized correlation functions. First, the value of the pumping cycle T_{ST} is divided into 8 equal time intervals T_{ST}, which are determined using the following formula:

$$\Delta T_{ST} = T_{ST} / 8 .$$

(E.g., iff $T_{ST} = 104$ s, then $\Delta T_{ST} = 13$ s).

Further, using formulas (3.8) and (3.9), estimates $r_{gg}^{R}(\mu = 0)$, $r_{gg}^{R}(\mu = 1\Delta T_{ST})$, $r_{gg}^{R}(\mu = 2\Delta T_{ST})$, …, $r_{gg}^{R}(\mu = 7\Delta T_{ST})$ are determined and using formulas

$$\Delta r_{gg}^{R}(\mu = 1\Delta T_{ST}) = r_{gg}^{R}(\mu = 0) - r_{gg}^{R}(\mu = 1\Delta T_{ST}),$$

$$\Delta r_{gg}^{R}(\mu = 3\Delta T_{ST}) = r_{gg}^{R}(\mu = 2\Delta T_{ST}) - r_{gg}^{R}(\mu = 3\Delta T_{ST}),$$

$$\Delta r_{gg}^{R}(\mu = 5\Delta T_{ST}) = r_{gg}^{R}(\mu = 4\Delta T_{ST}) - r_{gg}^{R}(\mu = 5\Delta T_{ST}),$$

$$\Delta r_{gg}^{R}(\mu = 7\Delta T_{ST}) = r_{gg}^{R}(\mu = 6\Delta T_{ST}) - r_{gg}^{R}(\mu = 7\Delta T_{ST})$$

are calculated their difference.

The next step is to calculate the minimum value of the normalized correlation function of the noisy signal $r_{gg}^{min}(\mu)$ and the corresponding value of $\mu_{\min}$, after which the informative attributes are determined in the form of the following coefficients:

$$K_{N1} = \frac{\Delta r_{gg}^{R}(\mu = 1\Delta T_{ST})}{\Delta T_{ST}}; \quad K_{N2} = \frac{\Delta r_{gg}^{R}(\mu = 3\Delta T_{ST})}{\Delta T_{ST}};$$

$$K_{N3} = \frac{\Delta r_{gg}^{R}(\mu = 5\Delta T_{ST})}{\Delta T_{ST}}; \quad K_{N4} = \frac{\Delta r_{gg}^{R}(\mu = 7\Delta T_{ST})}{\Delta T_{ST}};$$

$$K_{N5} = \frac{\Delta r_{gg}^{R}(\mu = 1\Delta T_{ST})}{\Delta r_{gg}^{R}(\mu = 3\Delta T_{ST})}; \quad K_{N6} = \frac{\Delta r_{gg}^{R}(\mu = 1\Delta T_{ST})}{\Delta r_{gg}^{R}(\mu = 5\Delta T_{ST})};$$

$$K_{N7} = \frac{\Delta r_{gg}^{R}(\mu = 1\Delta T_{ST})}{\Delta r_{gg}^{R}(\mu = 7\Delta T_{ST})}; \quad K_{N8} = \frac{\Delta r_{gg}^{R}(\mu = 3\Delta T_{ST})}{\Delta r_{gg}^{R}(\mu = 5\Delta T_{ST})};$$

$$K_{N9} = \frac{\Delta r_{gg}^{R}(\mu = 3\Delta T_{ST})}{\Delta r_{gg}^{R}(\mu = 7\Delta T_{ST})}; \quad K_{N10} = \frac{\Delta r_{gg}^{R}(\mu = 5\Delta T_{ST})}{\Delta r_{gg}^{R}(\mu = 7\Delta T_{ST})};$$

$$K_{N11} = r_{gg}^{min}(\mu); \quad K_{N12} = \mu_{min}.$$

Furthermore, using the estimates of the noise variance D_{ε}, the variance of the sum signal D_g, the variance of the sun noise $D_{æ}$ and the cross-correlation function $R_{U\varepsilon}(\mu)$, the informative attributes are determined in the form of coefficients:

$$K_{\varepsilon 1} = D_{æ} / D_g ; \; K_{\varepsilon 2} = D_{\varepsilon} / D_{æ} ; K_{\varepsilon 3} = R_{U\varepsilon} / D_{æ} .$$

Thus, the informative attributes $K_{N1} - K_{N12}$ of the normalized correlation function allow us to identify the technical condition of SRPU. Coefficients $K_{\varepsilon 1} - K_{\varepsilon 3}$ also allow us to duplicate the identification of the technical condition of SRPU.

7. Identifying and diagnosing the technical condition of SRPU in real time. As indicated earlier, the efficient operation of SRPU requires maintaining continuous control of the parameters and diagnostics of the object's technical condition in real time.

Fig. 4 shows the block diagram of the system for SRPU diagnostics and management. The diagram comprises three levels:

1. The deep-well pumping unit level, consisting of plunger pump 1; plunger 2; tubing 3; rods 4; polished rod 5; horsehead 6; walking beam 7; pitman 8; crank counterweight 9; reducer 10; multiple V-belt drive 11; prime mover 12; equalizer 13; force sensor 14; wellhead pressure sensor 15; rotation angle sensor 16; and crank of the beam-pumping unit 17.
2. The robust management station level (RMS, Fig. 5 and Fig. 6), consisting of the controller for data acquisition from force sensors 15; well-head pressure sensor 15 and rotation angle sensor 16; frequency converter for controlling the speed of the prime mover; a wireless modem equipped with an antenna to provide information exchange for RMS installed directly by the well with the centralized control station; and RS 485 interconnecting cables.
3. The oil field's centralized control station level, which serves up to 200 wells and consists of an industrial computer and a wireless modem with an antenna.

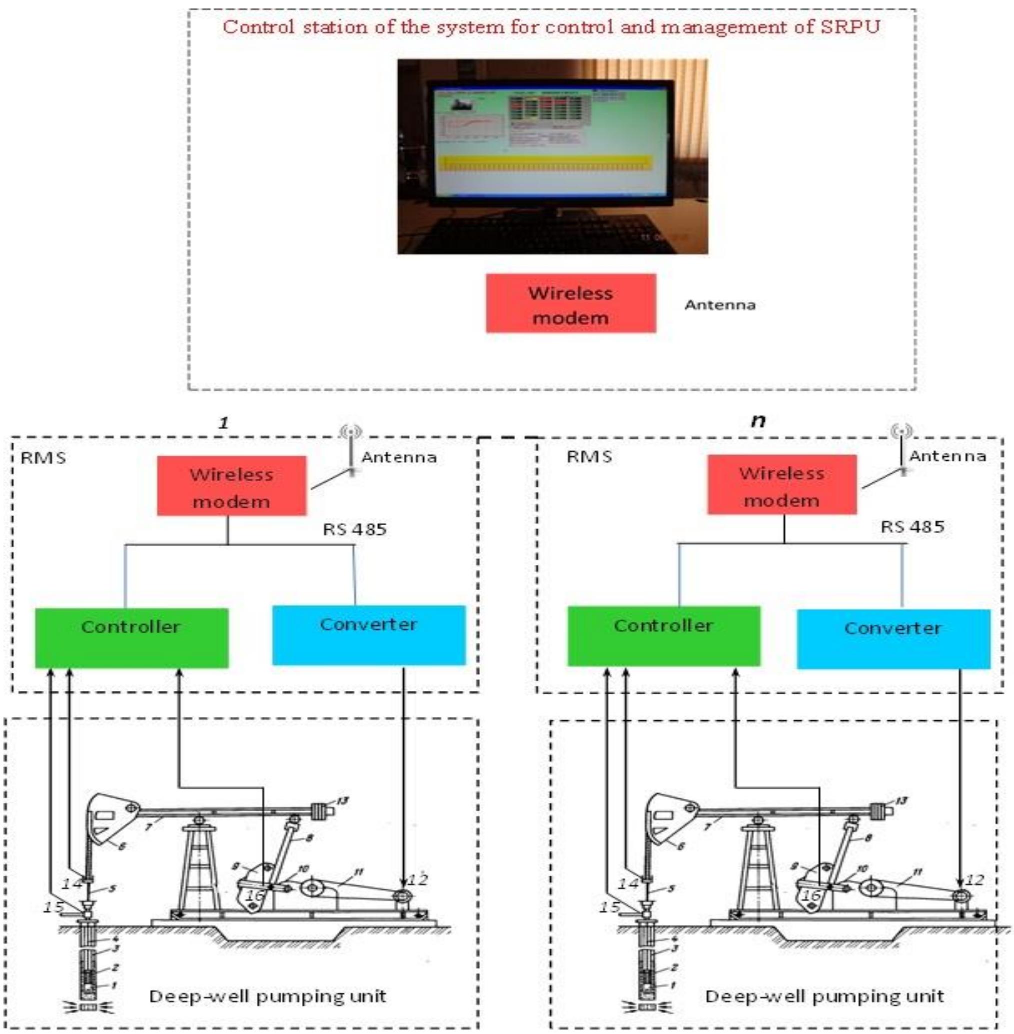

Fig. 4 Block diagram of the system for SRPU diagnostics and management

Fig. 5 External appearance of the RMS of the system for SRPU diagnostics and management

Fig. 6 RMS on the beam-pumping unit

For diagnostics of the state of SRPU, the system employs combinations of the technologies for identifying the dynamometer card based on the estimates of normalized correlation functions and noise characteristics of the force signal. To implement them in a robust system for SRPU diagnostics and management, the sampling interval of the force signal $U_p(t)$ was determined first, based on the duration of the pumping cycle T_{ST}. For most oil wells, the duration of T_{ST} varies within a range of 5-20 seconds. We have deduced from experiments that to obtain the estimate of the correlation function $r^R{}_{gg}(\mu)$ with the required accuracy, we need only to sample the force signal at the frequency $f = 500 - 1000$ Hz. It has also been established that robust normalized estimates of $r^R{}_{gg}(\mu)$ can be obtained from a number of samples of $U_p(t)$ $n \geq 1024$. Our experiments have proven that any minute change in the technical condition of SRPU during the pumping cycle T_{ST} affects the estimate of the normalized correlation functions of the force signal $U_p(i\Delta t)$. This, in turn, leads to a change in the informative attributes $K_{N1} - K_{N12}$ and $K_{\varepsilon 1} - K_{\varepsilon 3}$. Thus, during the operation of SRPU, combinations of those coefficients form easily based on its various technical conditions (Table 2), which is why they allow for the reliable identification of the force signal $U_p(i\Delta t)$, i.e., the technical condition of SRPU in real-time mode (Table 2).

Therefore the identification of SRPU faults comes down to the search for relevant combinations of the normalized coefficients $K_{N1} - K_{N12}$,which makes a visual interpretation of the dynamometer card unnecessary in determining the fault in SRPU. To demonstrate the possibility of the considered identification option in real industrial practice, we provide the 8 most common fault types in Fig. 7 below. Table 2 contains combinations of relevant estimates of normalized coefficients.

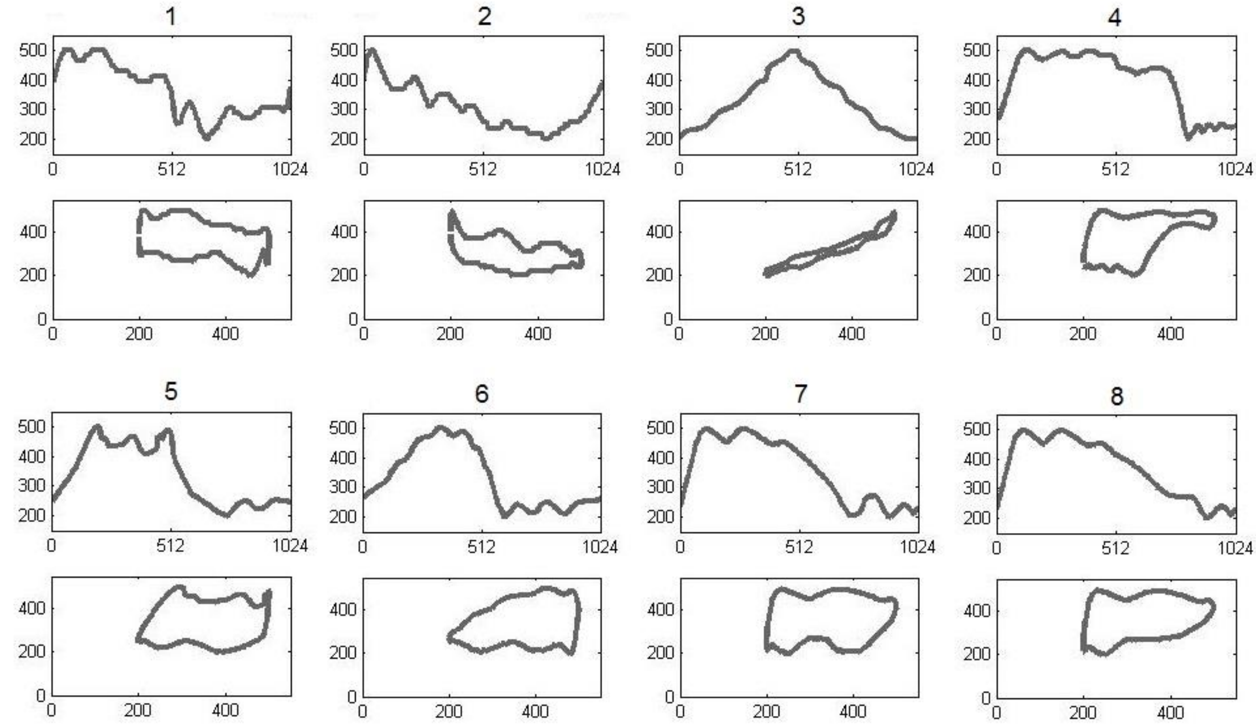

Fig. 7 Template force curves and dynamometer cards

Table 2

Combinations of estimates of normalized coefficients

Attributes / Fault type	**K$_{N1}$** $r^R{}_{gg}(\mu=1)/\Delta T_{ST}$	**K$_{N2}$** $r^R{}_{gg}(\mu=3)/\Delta T_{ST}$	**K$_{N3}$** $r^R{}_{gg}(\mu=5)/\Delta T_{ST}$	**K$_{N4}$** $r^R{}_{gg}(\mu=7)/\Delta T_{ST}$	**K$_{N5}$** $r^R{}_{gg}(\mu=1)/r^R{}_{gg}(\mu=3)$	**K$_{N6}$** $r^R{}_{gg}(\mu=1)/r^R{}_{gg}(\mu=5)$
Breakdown of the discharge line	0.058302	0.021427	-0.020932	-0.062865	2.7209	-2.7853
Breakdown of the suction line	0.042181	0.02277	-0.0067786	-0.043884	1.8525	-6.2226
Plunger sticking	0.072276	0.024799	-0.037473	-0.071132	2.9145	-1.9288
Pump down of the level	0.056666	0.043252	0.0013712	-0.053839	1.3101	41.325
High plunger fit	0.055614	0.011688	-0.025684	-0.06399	4.7581	-2.1653
High percentage of leakage in the discharge line	0.064761	0.0021131	-0.04916	-0.056753	30.647	-1.3173
Leakage in the suction line	0.061186	0.032006	-0.014403	-0.058085	1.9117	-4.2481
High percentage of leakage	0.058415	0.034291	-0.0071802	-0.050712	1.7035	-8.1357

Attributes / Fault type	**K$_{N7}$** $r^R{}_{gg}(\mu=1)/r^R{}_{gg}(\mu=7)$	**K$_{N8}$** $r^R{}_{gg}(\mu=3)/r^R{}_{gg}(\mu=5)$	**K$_{N9}$** $r^R{}_{gg}(\mu=3)/r^R{}_{gg}(\mu=7)$	**K$_{N10}$** $r^R{}_{gg}(\mu=5)/r^R{}_{gg}(\mu=7)$	**K$_{N11}$** $r^R{}_{gg\min}$	**K$_{N12}$** $\mu_{\min}$
Breakdown of the discharge line	-0.92741	-1.0236	-0.34084	0.33297	-0.9189	99
Breakdown of the suction line	-0.96119	-3.3591	-0.51887	0.15447	-0.8189	104
Plunger sticking	-1.0161	-0.66178	-0.34863	0.52681	-0.92938	94
Level pumpdown	-1.0525	31.543	-0.80337	-0.025469	-0.90529	104
High plunger fit	-0.8691	-0.45508	-0.18266	0.40137	-0.83869	94
High percentage of leakage in the discharge line	-1.1411	-0.042985	-0.037234	0.86621	-0.7803	81
Leakage in the suction line	-1.0534	-2.2221	-0.55101	0.24796	-0.89677	104
High percentage of leakage	-1.1519	-4.7758	-0.6762	0.14159	-0.84584	104

The identification of the technical condition of SRPU is duplicated using noise technology to enhance authenticity. Here, we use the technology to calculate the cross-correlation function $R_{U\varepsilon}(\mu=0)$ between the useful signal $U_p(i\Delta t)$ and the noise $\varepsilon(i\Delta t)$ from expressions (4.9), (4.11), (4.21), (4.22). According to our numerous experiments, in the normal state, the estimate of the cross-correlation function $R_{U\varepsilon}(\mu=0)$ between the useful signal and the noise is equal to zero. When any fault arises, the estimate of $R_{U\varepsilon}(\mu=0)$ is different from zero. As a result, when the technical condition of SRPU changes due to the effects of the noise that forms when a fault arises, estimates $K_{\varepsilon1}-K_{\varepsilon3}$ change. A certain combination of the values of those estimates corresponds to each of the 8 fault types (see Table 2).

Under field conditions when real objects are operating, the signals received from force sensor 144 (Fig. 4) are used by the system to determine combinations of the normalized coefficients $K_{N1},K_{N2},...,K_{N12}$ and $K_{\varepsilon1}-K_{\varepsilon3}$ in real time during the SRPU pumping cycle. They are compared to the reference coefficients, which were determined in advance (e.g.,, they are given for one object in Table 2). The most likely fault in the SRPU is identified based on the closest reference. The simplicity of calculating those coefficients makes it possible to implement the proposed diagnostic technique by means of any modern controller (in our case, LPC 2148 FBD64). Thus, during the operation of SRPU, control commands for the object are formed based on the results of identification, e.g., to change the duration of the pumping cycle. At the same time, the information concerning the state of the object is transmitted to the control station via an RMD 400 SP4 wireless modem (Fig. 4). When the object is in the normal state, this information is displayed on the screen at the central control station in green; when the initial stage of a fault is detected, this information is displayed in yellow, and an emergency state is indicated by red.

If the proposed technologies are introduced in oilfields with a large number of wells, combinations of reference coefficients are determined for each of well successively as various faults arise in SRPU. This is implemented with the participation of a technologist, who identifies the nature of the fault by interpreting the dynamometer card and registers the corresponding combinations of coefficients $K_{N1},K_{N2},...,K_{N12}$ and $K_{\varepsilon1}-K_{\varepsilon3}$ in the system's identification block. Thus, during the operation, as certain faults arise at each well, a combination of corresponding reference estimates of coefficients is formed in the identification block of each system. As a result, after a certain period of operation, combinations of reference coefficients for corresponding faults are formed and saved in the identification units of the SRPU diagnostic and management system at all wells. When all possible SRPU faults have been diagnosed and saved in the identification blocks the system moves into the automated diagnostic and management mode.

Fig. 7 shows template force curves $U_p(i\Delta t)$ and corresponding dynamometer cards for eight typical SRPU faults. Table 2 shows the combinations of the normalized coefficients $K_{N1}-K_{N12}$. It is clear that only one combination of estimates $K_{N1}-K_{N12}$ corresponds to each dynamometer card. It is also clear that if those combinations of coefficients $K_{N1},K_{N2},...,K_{N12}$ are available, we can definitely determine the same fault types that are usually determined visually by a technologist from a dynamometer card. Thus, combinations of these coefficients make it possible to perform automated identification of the technical condition of SRPU. Consequently, we no longer have to use dynamometer cards to visually identify the state of SRPU in the semi-automated mode. To enhance the authenticity and reliability of the identification results, the obtained results are compared with the combination of coefficients $K_{\varepsilon1}$ and $K_{\varepsilon3}$, which, alongside the informative attributes, allow for the identification of the technical condition of SRPU. The estimates are also determined and saved in advance for each fault type. Due to the use of these coefficients, the process of identification is duplicated, which enhances the authenticity and reliability of the system operation.

8. Conclusion. SRPU is widely used for oil lifting at old oilfields; however, as it is impossible to diagnose the technical condition of SRPU in real-time mode, the profitability of oilfields in long-term operation decreases. Therefore, the problem of developing new, efficient technologies and systems to diagnose the technical condition of SRPU is of both academic and economic interest.

The commonly used semi-automated method based on the interpretation of the shape of the dynamometer card does not yield desired results for true control of the technical condition of SRPU. This is because the human factor required by this method limits the time available for identification, on the one hand, and implies a dependence on the qualifications of the technologist, on the other. Moreover, the deeper the well, the more difficult it is to identify certain fault types from the shape of the dynamometer card alone. Analytically, correlation and spectral methods are the most suitable for identifying the technical condition of SRPU by analysing the force signal $U_p(i\Delta t)$. However, control objects actually operate under field conditions and the wide variation in climatic conditions worldwide leads to major errors in the obtained results. Further, mechanical processes associated with an object's transition to the emergency state also cause noises that vary over a wide range, and for these reasons, adequate estimates cannot be obtained either by correlation or by spectral methods. At first glance, the effects of the errors on the results of the identification of dynamometer cards can be eliminated by filtering the noise accompanying the useful signal $U_p(i\Delta t)$. If the noise spectrum is stable, the application of filtration usually produces satisfactory results. Under the field conditions, however, the spectrum of the noise varies widely due to abrupt changes in the factors of its formation. Additionally, the spectrum of the noise that forms due to changes in SRPU mechanical processes changes also varies widely, frequently overlapping the range of the useful signal spectrum. For these reason, we cannot get the desired effect using the technology for filtering the force signal, in which case the spectrum of the useful signal can be significantly distorted. Thus, we cannot always achieve satisfactory results through correlation or spectral analysis of the force signal by using filtration. Therefore, solving the problem under consideration first requires developing the technologies that can determine estimates of correlation and spectral characteristics and will remain practically unaffected by changes in said noises.

This paper establishes both theoretically and experimentally that it is possible to determine the normalized coefficients $K_{N1}, K_{N2},..., K_{N12}$, which are practically unaffected by the above-mentioned noises, by means of normalized correlation functions. The advantage of using these coefficients is that they are easy to calculate on modern controllers (for instance, LPC 2148 FBD64). Thus, it becomes possible to perform diagnostics in real time. When SRPU diagnostics are duplicated with the application of coefficients $K_{\varepsilon 1} - K_{\varepsilon 3}$ obtained from the estimates of the noise variance and cross-correlation functions between the useful signal and the noise, as well as the sum variance, the reliability and authenticity of the results is enhanced. The simplicity of realization of these technologies allows the development of a simple, reliable and inexpensive system for SRPU diagnostics and management, which we have introduced to working oil fields in Azerbaijan. The service experience of the system at 35 wells in the Bibi Heybat oilfield and 190 wells of the Azerbaijan-British company "Shirvan Oil", among others, has demonstrated the reliability of these systems. The enhanced diagnostics and management of SRPU made it possible to operate the wells in an adequate mode and increase their profitability due to savings in energy and prolonged overhaul periods. For instance, at the Bibi Heyat oilfield, energy savings reached 50 per cent and overhaul periods increased by 30 per cent.

It should be noted that the operational experience of these systems have established that the values of the coefficients $K_{N1} - K_{N12}$ and $K_{\varepsilon 1} - K_{\varepsilon 3}$ vary insignificantly (no more than 5-10 per cent) for wells of the same depth at every oilfield. Therefore, if we determine combinations of these coefficients at corresponding faults for one well, they can be used in the control systems for other wells of similar depth. Considering that the pump is placed at approximately the same depth at most old deposits, it is clear that forming the reference base of coefficient combinations in the diagnostics and management systems will not take much time.

Appendix:

```
A=10;
t=0:pi/40:120*pi-pi/40;
U=A*sin(t);
randn('state',0)
randn('state',0)
e=normrnd(0,2,1,4800);
V=U+e;
N=2400;
RVV0_s=0;
for i=1:N
RVV0_t(i)=V(i)*V(i);
RVV0_s=RVV0_s+RVV0_t(i);
end
RVV0=RVV0_s/N;
D_V=RVV0;
Ree0_s=0;
for i=1:N
Ree0_t(i)=e(i)*e(i);
Ree0_s=Ree0_s+Ree0_t(i);
end
Ree0=Ree0_s/N;
D_epsilon=Ree0;
RVV1_s=0;
for i=1:(N-1)
RVV1_t(i)=V(i)*V(i+1);
RVV1_s=RVV1_s+ RVV1_t(i);
end
RVV1=RVV1_s/(N-1);
RVV2_s=0;
for i=1:(N-2)
RVV2_t(i)=V(i)*V(i+2);
RVV2_s=RVV2_s+RVV2_t(i);
end
RVV2=RVV2_s/(N-1);
Dee=RVV0-2*RVV1+RVV2;
RUU0_s=0;
for i=1:N
RUU0_t(i)=U(i)*U(i);
RUU0_s=RUU0_s+RUU0_t(i);
```

```
end
RUU0=RUU0_s/N;
for k=1:(N-1)
for f=1:(N-k)
RUU_t=U(f)*U(f+k);
if f==1
RUU_s=RUU_t;
else
RUU_s=RUU_s+RUU_t;
end
end
RUU(k)=RUU_s/(N-k);
end
RUU_n_0=RUU0/RUU0;
for i=1:40
RUU_n(i)=RUU(i)/RUU0;
end
for k=1:(N-1)
for f=1:(N-k)
RVV_t=V(f)*V(f+k);
if f==1
RVV_s=RVV_t;
else
RVV_s=RVV_s+RVV_t;
end
end
RVV(k)=RVV_s/(N-k);
end
RVV_n_0=RVV0/RVV0;
for i=1:40
RVV_n(i)=RVV(i)/RVV0;
end
RVV_n_new_0=RVV0/RVV0;
for i=1:40
RVV_n_new(i)=RVV(i)/(D_V-Dee);
end
for i=1:100
U1(i)=U(i);
e1(i)=e(i);
V1(i)=V(i);
end
plot(U1,'g')
hold on
plot(e1,'r')
hold on
plot(V1)
figure
```

```
plot([RUU_n_0 RUU_n],'g')
hold on
plot([RVV_n_0 RVV_n],'r')
hold on
plot([RVV_n_new_0 RVV_n_new])
```

References

1. V.V. Andreyev, K.R. Urazakov, V.U. Dalimov et al. Oil extraction handbook (in Russian), M.: Nedra-Biznestsentr LLC, Moscow, 2000.
2. T.M. Aliev et al. Automated control and diagnostics of SRPU (in Russian), M.: Nedra, Moscow, 1988.
3. A.B. Neely and H.E. Tolbert. Experience with Pump-Off Control in the Permian Basin, SPE Technical Conference and Journal of Petroleum Technology. 40 №5 (1988) 645-648.
4. O.K. Nusratov, G.A. Guluyev, A.G. Rzayev, F.G. Pashayev, I.B. Yusifov, M.G. Rizvanov. Information measuring, monitoring, diagnostics and management complex for oil wells operated by sucker rod pumping units (in Russian), News of ANAS. 29 №6 (2009) 176-183.
5. A.B. Neely and S.G. Gibbs. Computer Diagnosis of Downhole Conditions in Sucker Rod Pumping Wells, Journal of Petroleum Technology. 18 №1 (1966) 91-98.
6. T.A. Aliev, O.K. Nusratov. Algorithms for the analysis of cyclic signals (in Russian), Avtomatika i Vichislitelnaya Tekhnika. 2 (1998) 68-74.
7. T.A. Aliev, G. A. Guluyev, F. H. Pashayev, A. B. Sadygov. Noise monitoring technology for objects in transition to the emergency state, Mechanical Systems and Signal Processing. 27 (2012) 755-762.
8. T.A.Aliev, A.M. Abbasov, G.A. Guluyev, F.H. Pashayev, U.E. Sattarova. System of robust noise monitoring of anomalous seismic processes, Soil Dynamics and Earthquake Engineering. 53 (2013) 11-15.
9. A.A. Abdullayev, A.A. Javadov, A.A. Levin et al. Telemechanic complexes for oil industry (in Russian), M. Nedra, Moscow, 1982.
10. T.R. Fasel, M.D. Todd. Chaotic insonification for health monitoring of an adhesively bonded composite stiffened panel, Mechanical Systems and Signal Processing. 24 № 5 (2010) 1420-1430.
11. Hao Tang, Y.H. Liao, J.Y. Cao, Hang Xie. Fault diagnosis approach based on Volterra models, Mechanical Systems and Signal Processing. 24 №. 4 (2010) 1099-1113.
12. T.A. Aliev. Digital Noise Monitoring of Defect Origin, Springer-Verlag, London, 2007.
13. Bendat Julius S. and Piersol Allan G. Random Data, Analysis & Measurement Procedures, fourth ed., Wiley-Interscience, New Jersey, 2000.
14. J.Bendat, A. Pierson. Application of correlation and spectral analysis (in Russian), M.: Mir, Moscow, 1983.
15. T.A. Aliev. Robust Technology with Analysis of Interference in Signal Processing, Kluwer Academic/Plenum Publishers, New York, 2003.

UOT 681.58

T.A. Əliyev, A.H. Rzayev, Q.A. Quluyev, T.Ə. Əlizadə, N.E. Rzayeva

Robust Noise texnologiya və neft quyularının ştanqlı dərinlik nasos qurğusunun diaqnostika və idarəetmə sistemi

Çöl şəraitində istismar olunan, küylərin təsirinə məruz qalan ştanqlı dərinlik nasos qurğularının (ŞDNQ) vəziyyətinin diaqnoz olunmasının çətinlikləri analiz olunur. Qüvvə və gediş vericilərindən alınmış siqnallar əsasında qurulmuş dinamoqramların interpretasiyasına əsaslanan mövcud diaqnostika üsullarının real vaxt miqyasında diaqnozlaşdırma və idarəetmə məsələlərini həll edə bilməməsi göstərilmişdir. ŞDNQ – in mümkün qəza vəziyyətlərinə uyğun gələn informatik əlamətlər kombinasiyası formalaşdırmaq məqsədi ilə robast, normaya salınmış korrelyasiya funksiyalarının təyin edilmə texnologiyası təklif olunmuşdur. Eyni zamanda ŞDNQ – in bu vəziyyətlərinin identifikasiyası qüvvə siqnalının noise xarakteristikalarına uyğun kombinasiyaların təyin edilməsi və formalaşdirilması yolu ilə təkrarlanir, bu isə diaqnozlaşdırmanın nəticələrinin etibarlılıq dərəcəsini yüksəldir.

İşləmə alqoritmlərinin sadəliyi qüvvə siqnalının identifikasiya məsələsini sadə və ucuzlu kontroller vasitəsi ilə real vaxt miqyasında həll etməyə imkan verir.

Yaradılmış alqoritmlərin 100- dən çox real obyektdə tətbiqi elektrik enerjisinə qənaət və təmirlər arası vaxtın artırılması hesabına neft quyularının rentabelliyinin kifayət qədər yüksəlməsini göstərmişdir.

Açar sözləri: noise texnologiya, dinamoqramma, ştanqlı nasos qurğusu, neft quyusu, küyün dispersiyası, qarşılıqlı korrelyasiya funksiyası

УДК 681.58

Т.А. Алиев, А.Г. Рзаев, Г.А. Гулуев, Т.А. Ализаде, Н.Э. Рзаева

Робастная Noise технология и система диагностики и управления ШГНУ нефтяных скважин

Анализированы трудности диагностики штанговых глубинно насосных установок (ШГНУ), вызванные влиянием помехи, связанные спецификой его эксплуатации в полевых условиях. Показано, что существующие методы диагностики, основанные на интерпретации динамограмм, построенных по сигналам, получаемым от датчиков усилия и хода, не позволяют решить задачи диагностики и управления в реальном масштабе времени. Предложена технология определения робастных нормированных корреляционных функций, при помощи которых формируются комбинации информативных признаков, которые соответствуют возможным аварийным состояниям ШГНУ. Идентификация этих состояний ШГНУ дублируется путем определения и формирования комбинаций, соответствующих noise характеристик сигнала усилия, что повышает степень надежности результатов диагностики. Простота реализации алгоритмов обработки позволяет решить задачу идентификации сигнала усилия при помощи недорогих контроллеров в реальном масштабе времени. Применение на более 100 реальных объектах показало, что при этом за счет экономии электроэнергии и увеличения межремонтного периода, рентабельность нефтяных скважин значительно повышается.

Ключевые слова: noise технология, динамограмма, штанговая насосная установка, нефтяная скважина, дисперсия шума, взаимнокорреляционная функция

Institute of Control Systems of the Azerbaijan National Academy of Sciences Presented 06.04.2015