STATISTICAL ANALYSIS OF THE RELATIVE POSITION OF THE ROD HANGER AND THE WELLHEAD

Ibrahim Habibov \boxtimes *Department of Industrial Machines¹ h.ibo@mail.ru*

Sevinc Abasova Department of Industrial Machines¹

Vusala Huseynova Department of Industrial Machines¹

1 Azerbaijan State Oil and Industry University 20 Azadlig ave., Baku, Azerbaijan, AZ 1010

***Corresponding author**

Abstract

One of the leading methods of exploitation of oil fields is oil production with the help of downhole rod pumping units (DRPU). Over 80 % of the operating well stock of Azneft PA is equipped with deep well pumps and about 30 % of oil is produced in the country with their help.

The widespread use of DRPU is associated with a fairly high maturity of installations, simplicity of its design and maintenance, repair in field conditions, ease of adjustment, the possibility of servicing the installation by unskilled workers, a small effect on the operation of DRPU of the physical and chemical properties of the pumped liquid, as well as high efficiency.

However, along with the high efficiency of the applied DRPU, there are also complaints regarding the need to increase the reliability and resource of wellhead equipment, including in order to improve the environmental situation in the oil fields.

One of the conditions for ensuring high reliability of the ground equipment of the DRPU is to ensure the tightness of the wellhead rod-wellhead stuffing box assembly, the violation of which is not only a failure of the installation, but also leads to environmental pollution.

This is facilitated by inaccuracies in the assembly and installation of DRPU at the wellhead. When mounting the pumping unit, for many reasons, the tolerance of the wellhead rod with the suspension point of the rod string to the balancer head is not ensured.

In this regard, in the requirements for the accuracy of mounting the pumping unit at the point of application, a certain mismatch of the axes within the circular coordinates is allowed. So, for widely used pumping units of the CK8 type, the permissible mismatch between the axis of the wellhead rod and the suspension point of the rods is determined by the conditions under which the projection of the suspension point of the rods onto the plane of the base of the pumping unit at any position of the balancer is allowed within a circle with a diameter of 25 mm.

Keywords: downhole rod pumping unit, balancer, eccentricity, Rayleigh distribution.

DOI: 10.21303/2461-4262.2022.002267

1. Introduction

It is known that of the mechanized methods of oil production, the most widely used are downhole rod pumping units (DRPU) [1, 2]. However, despite the high degree of development of the DRPU designs, there are some difficulties associated with the installation of the unit at the wellhead. So, for pumping units (type SK 8 and SK10), which have found the greatest use at the oil and gas production departments (OGPD) located on the Absheron Peninsula of Azerbaijan, the eccentricity of the location of the suspension point of the rods relative to the axis of the wellhead rod (in the plane of the base of the pumping unit at any location of the balancer) should not exceed 25 mm [3, 4]. Naturally, ensuring an acceptable eccentricity presents certain difficulties.

The specified eccentricity not only affects the performance of the wellhead rod, but also increases the likelihood of failure of the gland of the wellhead rod, leads to leaks of the extracted products, and, consequently, worsens the environmental situation in the fields [5, 6].

Engineering

2. Materials and methods of research

To identify the true state of affairs, at 80 wells in Absheron, where pumping units of the SK 8 type are operated, measurements were made and the eccentricity of the location of the suspension point of the rods relative to the wellhead was determined with an accuracy of 1 mm. The measurement results are presented in **Table 1**.

Table 1

The determination of the eccentric location of the suspension point of the rods relative to the wellhead was carried out in accordance with the requirements of GOST 8.051-81.

3. Results and discussion

The experiments were carried out in five oil and gas production departments (OGPD of Azerbaijan: named after N. Narimanov (in 24 wells), named after Amirov (in 16 wells), Apsheronneft (in 18 wells), Oil Rocks (in 12 wells), and on the 28th May (in 20 wells). According to statistical data characterizing the eccentricity of the location of the suspension point of the rods relative to the wellhead and presented in the form of **Table 1**, it is required to establish the law of distribution of eccentricity and determine the accuracy of the relative position of the axes.

According to the grouped data in **Table 1**, given with a step $h = 4$, let's construct a histogram of the sample distribution of eccentricity (**Fig. 1**).

Fig. 1. Histogram of the selected eccentricity distribution

Initially, the hypothesis of the distribution of eccentricity according to the Rayleigh law was tested. Most often, the random value *P* of the deviation of the eccentricity of the axes has a distribution according to the Rayleigh law [7, 8]. It is one-parameter, and the distribution function has the following form:

$$
\varphi(R) = \frac{R}{\sigma^2} e^{-\frac{R^2}{2\sigma^2}},
$$
\n(1)

where *R* – variable value of the eccentricity, and $R = \sqrt{x^2 + y^2}$, and *x* and *y* are the coordinates of the point *P* (**Fig.** 2); σ – standard deviation of *x* and *y* coordinate values having the same distribution, therefore $\sigma = \sigma_x = \sigma_y$.

Fig. 2. Eccentricity (*R*) between the suspension axis of the sucker rods and the wellhead

The interval distribution law has the expression [9]:

$$
F(R) = \frac{1}{\sigma^2} \int_0^R R \cdot e^{-\frac{R^2}{2\sigma^2}} dR = 1 - e^{-\frac{R^2}{2\sigma^2}}.
$$
 (2)

This distribution is different in that it is based on the normal distribution. Here the *x* and *y* coordinates of the point *R* are normally distributed, but the distribution itself is not normal. To calculate the distribution function (2) of the random variable *R*, it is necessary to know only one parameter σ_R by the relation:

$$
\sigma = \frac{\sigma_R}{\sqrt{2 - \frac{\pi}{2}}} \,. \tag{3}
$$

Using **Table 1**, let's determine from the sample values of the eccentricity mean X_R and variance S_R , for which $\sum R_i \cdot P_i$ $\cdot P_i =$ $\sum_{i=1}$ 13 16.6, $S_R = 11.456$.

i On average \bar{X}_R , the value of the sample, does not change, and the variance increases by the Sheppard value (the difference between the calculated and actual variance) [10].

Taking into account the Sheppard correction, let's obtain $S_R = \sqrt{(131.24 - 1.3333)} = 11.3977$. Assuming $\sigma_R \approx S_R$ according to (3) let's obtain $\sigma = \sigma_R / \sqrt{2 - \pi/2} = 11.3977/0.655 = 17.4$.

Using the table in Appendix 11 [8], for each observed value of the eccentricity, let's calculate the theoretical values $F(R)$, and by $F(R)$ determine the theoretical values of the frequency. In this case, when calculating R_i/σ , the upper value $R_{H\sigma}$ of the intervals of *R* values should be taken as the eccentricity.

Table 2 (columns $3-5$) shows the data for calculating the theoretical frequencies f' . To complete column 5, it is necessary to subtract the previous value $F_{i-1}(R)$ from each subsequent value of $F_i(R)$. For example, for the second row $f'/n = 0.1004 - 0.0261 = 0.0743$; for the third row $f'/n = 0.2188 - 0.1004 = 0.1184$, etc. For the first row, taking into account the value $F(0) = 0$, let's obtain $f'/n = 0.0261$. Column 6 is filled in by multiplying the data in column 5 by $n = 80$.

Fig. 3 shows the theoretical and empirical distribution curves according to the Rayleigh law.

Table for calculating the frequency according to the Rayleigh distribution law

Fig. 3. Distribution curves according to the Rayleigh law: 1 – theoretical, 2 – empirical

Fig. 3 shows a significant discrepancy between the empirical and theoretical distribution functions (*DF*) under the hypothesis of the theoretical Rayleigh distribution law, which most likely indicates the illegitimacy of subordinating the put forward hypothesis $H: F(x)$ of Rayleigh. To verify this, let's use the statistical criterion of A. N. Kolmogorov – λ [11].

To calculate the influence of λ , one should first determine the values of the empirical $F_n(x)$ and theoretical $F(x)$ distribution functions (assuming that $F(x)$ is the Rayleigh *DF*) for each observed value of the random variable x ($x = R$). Then, the maximum difference of these functions is determined using the following formula

$$
\lambda = \left| F(x) - F_n(x) \right|_{\max} \cdot \sqrt{n} = D\sqrt{n}.\tag{4}
$$

Since $F(x) = N'_x/n$ and $F_n = N_x/n$, where the indicators in the numerator are, respectively, the accumulated theoretical and empirical frequencies, and the indicator in the denominator represents the sample size, instead of formula (4), it is also possible to use the following formula:

$$
\lambda = \frac{|N_x - N_x|_{\text{max}}}{n} \cdot \sqrt{n}.\tag{5}
$$

The cumulative frequency of any such x_m value is the sum of the frequencies of all previous values of x_i , including the frequency of x_i itself, i.e.:

$$
N_{xm} = \sum_{i=1}^{m} f_i,\tag{6}
$$

where m – the number of eccentricity values; f_i – frequency of the current value of the eccentricity.

It is known [11] that for continuous random variables:

$$
P(D\sqrt{n}\le\lambda)=K(\lambda),\tag{7}
$$

where $K(\lambda) = \sum_{\gamma = -\infty}^{+\infty} (-1)^{\gamma} l^{-2\gamma \lambda^2}$. For large *n* and any $\lambda > 0$:

$$
P(\lambda) = 1 - K(\lambda). \tag{8}
$$

The function of continuous random variables is tabulated, and using a table of their values, a table of values of the eccentricity deviation value is compiled, which is given in [12]. According to the calculated value of A. N. Kolmogorov's statistical criterion, according to (5) and Appendix 12, the probability of continuous random variables is determined. If this probability turns out to be very small, in practice, when $P(\lambda) \le 0.05$, then the discrepancy between the empirical and theoretical distribution functions is considered significant, and not random, and the hypothesis about the proposed law of the distribution of x is rejected. If the probability of the distribution function is large enough (practically, when >0.05), then the hypothesis is accepted. For the convenience of calculating λ , an auxiliary table is compiled 3.

From **Table 3** let's find $|N'_x - N_x|_{\text{max}} = 17.536$ and calculating λ according to formula (5), let's obtain $(17.536/13)\sqrt{13} = 4.86$. Further from [12] let's find that already at $\lambda = 2.50$ $P(\lambda) = 0$ since $P(\lambda)$ is a decreasing function, then $P(4.86) = 0$. Therefore, the hypothesis *H* is not true.

It should be noted that the use of the criterion λ assumes the continuity of $F(x)$ and, in addition, it is assumed that the empirical *DF* $F_n(x)$ is built on the values of the random variable *x* not grouped into intervals. However, when the grouping intervals are small, the criterion gives, although approximate, but quite acceptable for practical purposes, an estimate of the proximity of the empirical *DF* to the theoretical *DF*.

Subsequently, the hypothesis of the distribution of eccentricity according to the Weibull law was tested. The distribution function of a random variable obeying the Weibull law has the following expression:

$$
F(x) = 1 - \exp\left(-\left(\frac{x}{x_0}\right)^m\right),\tag{9}
$$

and depends on two parameters m and x_0 .

Let a and σ_x be sample values of the mean and standard deviation of a random variable *x* with distribution (9). Then the relation:

$$
\upsilon_x = \frac{\sigma_x}{a},\tag{10}
$$

is a function of *m* whose values are tabulated in [12]. In the example $a = 16.6$ and $\sigma_x = 11.3971$ (taking into account Sheppard's correction). Substituting the values of a and σ into (10), let's obtain υ*x* = 0.6866 and [12] find *m* = 1.5.

To estimate the accuracy of calculating the parameter *m*, let's write (9) in the notation $x_0 = \theta$, $m = \beta$ in the form of the expression:

$$
F(x) = 1 - \exp(-\left(x/\theta\right)^{\beta}).
$$
\n(11)

Distribution (11) has a one-to-one correspondence with the distribution:

$$
F(x) = 1 - \exp(-\exp((x - \xi)/\psi)),\tag{12}
$$

which is called the limit distribution of the 1st type with parameters ξ and Ψ [13].

The distribution parameters (11), (12) are related by the relations:

$$
\xi = \ln \theta, \ \psi = \frac{1}{\beta},\tag{13}
$$

In [14], the following method for approximate calculation of the confidence interval for the parameter ψ and, thus, for the parameter $\beta = 1/\psi$ is proposed. The $(h\hat{\psi})/\psi$ distribution, where $h = 2/D^2(\hat{\psi}/\psi)$ (*D*-dispersion sign) is replaced by χ^2 -distribution with *h* degrees of freedom. The value *h* is chosen so that the first two moments of the true distribution and the distribution approximating χ^2 coincide. For non-integer values of *h*>3 and proved 0.01 <γ <0.99, the Wilson-Hilferty approximation is recommended [15]:

$$
\frac{\chi_{\gamma}^{2}h}{h} = \left(1 - \frac{2}{9h} + \frac{2}{9h} \cdot z_{\gamma}\right)^3,\tag{14}
$$

where $\chi^2(\hbar) - \gamma$ – the $\chi^2(\hbar)$ -distribution quantile and the *z*_γ- γ quantile of the standard normal distribution. For $γ = 0.1$:

$$
P\left\{\frac{h\hat{\psi}}{\psi} > \chi^2_{0.1}(h)\right\} = 0.1 \Leftrightarrow P\left\{\frac{h\hat{\psi}}{\psi} \le \chi^2_{0.1}(h)\right\} = 0.9 \Leftrightarrow P\left\{\frac{m}{\hat{m}} \le \frac{\chi^2_{0.1}(h)}{h}\right\} = 0.9. \tag{15}
$$

Let a random variable obeying the standard normal law $N(0.1)$ and $z_{0.1}$ – be its level quantile γ = 0.1 i.e.:

$$
P\big|\{Z > z_{0.1}\} = 0.1.\tag{16}
$$

So,

$$
\Phi(z_{0.1}) \equiv P\{Z \le z_{0.1}\} = 0.9,\tag{17}
$$

where $\Phi(z)$ is the distribution function *Z*. From **Table 2** of Appendix II [16] let's find $z_{0,1} = 1.28$.

Substituting this value $z_{0,1}$ into (14) and setting in (14):

$$
1 - \frac{2}{9h} + \frac{2}{9h} \cdot 1.28 = 1.01.
$$
 (18)

Let's obtain:

$$
\chi_{0.1}^2(h)/h = 1.03. \tag{19}
$$

From (18) let's find *h* = 6.22. Thus, when found *h* from (15), (19), let's obtain the upper confidence interval for the parameter *m*:

$$
P\{m < 1.55\} = 0.9.\tag{20}
$$

Assuming now $\gamma = 0.9$ there is:

$$
P\left\{\frac{h\hat{\psi}}{\psi} > \chi^2_{0.9}(h)\right\} = 0.9 \Longleftrightarrow P\left\{\frac{\hat{\psi}}{\psi} \le \frac{\chi^2_{0.9}(h)}{h}\right\} = 0.1 \Longleftrightarrow P\left\{\frac{m}{\hat{m}} > \frac{\chi^2_{0.9}(h)}{h}\right\} = 0.9. \tag{21}
$$

From **Table 2** of Appendix II [8] let's find the value of the Laplace function. Substituting this value into formula (14) let's obtain $\chi_{0.9}^2(h)/h = 0.0017$. From (21) let's find the lower confidence interval for the eccentricity value:

$$
P\{m > 0.026\} = 0.9.\tag{22}
$$

To estimate the distribution parameter ξ (13), let's use a simple linear unbiased estimate [17]:

$$
\hat{\xi} = \sum_{i=1}^{13} P_i l_n x_i + c^* \hat{\psi},
$$
\n(23)

where $c^* \approx 0.5772$ – Euler constant.

Based on the confidence intervals (20), (22), it is possible to take $\dot{m} = 1.5$ as an estimate of the parameter with a confidence probability $P_q = 0.9$ and, therefore, $\hat{\mathbf{y}} = 1/\hat{m} = 0.66$. From (23) Let's find: $\hat{\xi} = 2.93252$ and find the parameter estimate θ :

$$
\hat{\theta} = \ell^{2.93} = 18.8. \tag{24}
$$

Let's accept the hypothesis:

$$
H: F(x) = 1 - \exp\left(-\left(\frac{x}{18.8}\right)^{1.5}\right).
$$
\n
$$
(25)
$$

Let's compare the empirical distribution $f_n(x)$ with the theoretical distribution $F(x)$ defined by expression (25). As an empirical distribution, let's take:

$$
F_n(x) = \begin{cases} 0, & x \le x_1^*, \\ N_{X_i^*}/n, & x_{i-1}^* \le x_i^*, \end{cases}
$$
 (26)

Here, $x_1^* = 4$, $x_2^* = 8$, ..., $x_{13}^* = 52$ the right parts of the intervals [0, 4], [4, 8], ..., [48, 52], which form the variational series:

$$
x_1^* < x_2^* < \dots < x_m^* \ (m = 13), \tag{28}
$$

where N_{X_i} – accumulated point x_i^* ($i = 1, 13$), frequencies determined by formula (6).

To calculate the criterion λ, let's compile an auxiliary **Table 4**. When filling in column 4, let's assume $x_i = x_i^*$.

Let's apply the λ Kolmogorov criterion. Calculating λ by formula (4) let's obtain $\lambda = 0.038x\sqrt{80} \approx 0.34.$

Next, let's find the probability value -0.9997 ; hypothesis (26) is accepted. According to this hypothesis $P(x < 25) = 1 - e^{-1.77} \approx 0.8287$. This means that with an allowable limit value of eccentricity equal to 30 mm, the possible percentage of rejects (i.e., exceeding the limit value *R*) will be *q**u*003*d* (1–0.8287)×100≈17%. Therefore, it is necessary to ensure appropriate control of the eccentricity value during operation.

Table 4

5. Conclusions

Statistical analysis of the eccentricity according to the measurements made on pumping units of the SK 8 and SK 10 types showed that the Rayleigh law traditionally used to study the eccentricity is not correct. Since the Kolmogorov statistical criterion λ = 4.86. In this case, *P*(4.86) = 0, therefore, the hypothesis does not reflect the true state of the process.

Testing the hypothesis about the distribution of eccentricity according to the Weibull law found that the Kolmogorov statistical criterion in this case is $\lambda = 0.34$. Therefore, $P(0.34) = 0.83$, which means that with an allowable limit value of eccentricity equal to 30 mm, the possible rejection rate will be 17 %.

The above results also indicate the need for a set of studies to justify the allowable deviations of the suspension point of the rods relative to the axis of the wellhead stock.

References

- [1] Guo, B., Liu, X., Tan, X. (2017). Petroleum Production Engineering, Gulf Professional Publishing. Available at: https://www. elsevier.com/books/petroleum-production-engineering/guo-phd/978-0-12-809374-0
- [2] Kennedy, F. F., Michel, L., Frederic, L., Abbas, N., Njoke, M. (2015). Hands-on model of sucker rod pumping facility for oil well production. Journal of Petroleum and Gas Engineering, 6 (4), 45–53. doi: https://doi.org/10.5897/jpge2015.0220
- [3] Kyazimov, I. N. (2011). Razrabotka kompensatorov smescheniy osey ust'evogo shtoka i kanatnoy podveski stanka-kachalki. Materialy Mezhdunarodnoy NPK «Sovremennye problemy neftyanogo kompleksa Kazakhstana». Aktau, 297–300. Available at: http://www.rusnauka.com/27_NNM_2011/Tecnic/3_93038.doc.htm
- [4] Kyazimov, I. N. (2012). Ustranenie smescheniy osey ust'evogo shtoka i kanatnoy podveski pri nasosnoy ekspluatatsii skvazhin. Izvestiya VTUZov Azerbaydzhana, 2, 44–47.
- [5] Milovzorov, G., Ilyin, A., Shirobokov, P. (2019). Diagnostics of the condition of sucker-rod pumping units after the analysis of dynamogram cards. MATEC Web of Conferences, 298, 00137. doi: https://doi.org/10.1051/matecconf/201929800137
- [6] Gilaev, G. G. (2008). Diagnostics of deep-well pumping wells by dynamometry. RA Paratsels Publishers, 212.
- [7] Saleem, I., Aslam, M., Azam, M. (2013). The use of Statistical Methods in Mechanical Engineering. Research Journal of Applied Sciences, Engineering and Technology, 5 (7), 2327–2331. doi: https://doi.org/10.19026/rjaset.5.4660
- [8] Solonin, I. S. (1972). Matematicheskaya statistika v tekhnologii mashinostroeniya. Moscow: Mashinostroenie, 216.
- [9] Fletcher, S. J. (2017). Data Assimilation for the Geosciences. Elsevier. Available at: https://www.sciencedirect.com/ book/9780128044445/data-assimilation-for-the-geosciences
- [10] Schneeweiss, H., Komlos, J. (2008). Probabilistic rounding and Sheppard's correction. Technical Report Number 045. University of Munich, 32. Available at: https://epub.ub.uni-muenchen.de/8661/1/tr045.pdf
- [11] Orlov, A. I. (2014). Probabilistic-statistical methods in Kolmogorov's researches. Politematicheskiy setevoy elektronniy nauchniy zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. Available at: https://cyberleninka.ru/article/n/veroyatnostno-statisticheskie-metody-v-rabotah-a-n-kolmogorova
- [12] Montgomery, D. C. (2009). Introduction to Statistical Quality Control. John Wiley & Sons, Inc. Available at: https://endustri. eskisehir.edu.tr/ipoyraz/TKY302/icerik/text%20book_montgomery_6th%20edition.pdf
- [13] DeGroot, M. H. (2004). Optimal statistical decisions. John Wiley and Sons Inc., 512. Available at: https://www.wiley.com/ en-us/Optimal+Statistical+Decisions-p-9780471680291
- [14] Afifi, A., Eyzen, S. (1982). Statisticheskiy analiz. Podkhod s ispol'zovaniem EVM. Moscow: Mir, 488. Available at: https:// www.twirpx.com/file/101848/
- [15] Grigor'ev-Golubev, V. V., Vasil'eva, N. V., Krotov, E. A. (2014). Teoriya veroyatnostey i matematicheskaya statistika. Sankt-Peterburg: BKHV-Petrburg, 256. Available at: https://cdn1.ozone.ru/multimedia/1008828773.pdf
- [16] Engelhardt, M., Bain, L. J. (1973). Some Complete and Censored Sampling Results for the Weibull or Extreme-Value Distribution. Technometrics, 15 (3), 541–549. doi: https://doi.org/10.1080/00401706.1973.10489080
- [17] Bain, L. J. (1972). Inferences Based on Censored Sampling From the Weibull or Extreme-Value Distribution. Technometrics, 14 (3), 693–702. doi: https://doi.org/10.1080/00401706.1972.10488958

Received date 17.06.2021 Accepted date 09.12.2021 Published date 31.03.2022

© The Author(s) 2022 This is an open access article under the Creative Commons CC BY license

How to cite: Habibov, I., Abasova, S., Huseynova, V. (2022). Statistical analysis of the relative position of the rod hanger and the wellhead. EUREKA: Physics and Engineering, 2, 82–90. doi: https://doi.org/10.21303/2461-4262.2022.002267