APPLICATION OF A RANK FUZZY REGRESSION MODEL TO PREDICT THE TECHNICAL CONDITION OF WELL PIPES

Ibrahim Habibov, Oleg Dyshin, Gulnara Feyziyeva, Irada Ahmadova, Zohra Garayeva

•

Azerbaijan State Oil and Industry University
h.ibo@mail.ru
oleg.dyshin@mail.ru
gulnara.feyziyeva5@gmail.com
ahmadovairada@gmail.com
zabiyevaadnsu@gmail.com

Abstract

Oil and gas pipes used in well operations are undergo to aggressive environments. In this case, corrosion wear of the thickness of their walls occurs, which leads to various difficulties. In order to assess the technical condition of well pipes, geophysical methods are used, one of which is the electromagnetic inspection method.

The paper proposes a method for predicting the maximum loss of pipe thickness based on the results of electromagnetic inspection by the step values of the depth of immersion of the lower part of the pipe into the well.

Based on the use of fuzzy regression with fuzzy input/fuzzy output, a method for assessing the level of impact of the main formation parameters on the technical condition of well pipes is proposed.

Keywords: rank transformation, electromagnetic inspection, fuzzy emissions, moving forecasting, membership function

I. Introduction

Despite a fairly wide range of methods for assessing the technical condition of oil and gas pipes, the most widely used method is electromagnetic inspection. (EMI) [1-4].

Experimental studies were carried out for an offshore field, using an electromagnetic inspection of the EMI-43 type, at a depth of 2000-2400 m. The parameters of the pipes under study were $D_o \times d_i$=127.0×108.6 mm, $D_o \times d_i$ =339.7× 313.6 mm and $D_o \times d_i$ = 473.1 × 446.1 mm (D_o and d_i are the outer and inner string diameter, respectively).

II. The purpose of the work

Is to development of a method for predicting the loss of thickness of the outer pipe of a technical string not accessible to the depth interval for measurements based on previous inspection measurements.

III. Results and discussions

We will demonstrate the application of the RT method to fuzzy regression with a fuzzy input/output the case of measurements using an electromagnetic inspection EMI, which allows us to determine the loss of thickness of the outer pipe of a technical string in a well.

Tables 1 and 2 correspondingly show the values of the maximum loss of pipe thickness and the dependence of the formation parameters (density ϱ and viscosity ν) on the depth of immersion of the pipe into the well in the period 2019-2022 obtained by EMI methods.

Table 1: *Dependence of the maximum loss of pipe thickness on the depth of immersion into the well*

Top part of the pipe, (m)	Bottom of the pipe, (m)	Pipe length, (m)	Nominal thickness of pipe, (mm)	Actual minimum thickness of pipe (mm)	Depth of maximum loss of pipe thickness, (m)	Maximum loss of pipe thickness, (%)	Classification of losses
1	2	3	4	5	6	7	8
2284,10	2292,40	8,29	13,06	12,40	2285,10	5,1	B
2293,70	2302,00	8,29	13,06	12,51	2300.00	4,2	A
2303,30	2311,60	8,34	13,06	12,14	2308,30	7,1	B
2312,90	2321,20	8,29	13,06	12,26	2321,20	6,1	B
2322,40	2330,80	8,34	13,06	12,18	2323,50	6,8	B
2332,10	2340,60	8,50	13,06	12,41	2335,70	5,0	A
2341,90	2350,00	8,11	13,06	12,57	2347,40	3,7	A
2351,30	2359,20	7,98	13,06	8,54	2359,20	34,6	E
2361,20	2369,10	7,86	13,06	8.34	2366,80	36,1	E
2370,40	2381,00	10,61	13,06	8,54	2380,70	34,6	E

Table 2. *Dependence of the formation parameters (density ρ, kg/m³ and viscosity ν, s·poise) on the depth of immersion of the pipe into the well in the period 2019-2022*

Years							
2019		2020		2021		2022	
ρ	ν	ρ	ν	ρ	ν	ρ	ν
899	6,25	898	5,73	890	5,55	890	5,15
890	6,20	897	5,68	890	5,50	890	5,05
889	6,17	895	5,66	890	5,48	888	5,00
889	6,14	892	5,60	886	5,46	886	4,98
822	6,10	886	5,51	880	5,40	882	4,90
891	6,05	885	5,51	876	5,38	880	4,87
890	6,03	880	5,51	874	5,33	878	4,82
890	6,01	880	5,48	876	5,28	870	4,80
889	5,95	878	5,45	866	5,25	863	4,76
868	5,93	875	5,40	863	5,06	860	4,72
885	5,90	871	5,33	880	5,15	850	4,69
886	5,87	870	5,30	870	5,10	845	4,67
882	5,81	866	5,27	871	5,07	840	4,65
880	5,78	862	5,25	870	5,02	834	4,61
874	5,73	860	5,22	866	5,00	830	4,58
871	5,73	857	5,20	865	4,95	822	4,55
870	5,71	855	5,20	863	4,90	818	4,52
869	5,73	851	5,18	860	4,86	815	4,48
868	5.70	848	5,16	858	4,81	810	4,40
865	5,68	844	5,16	856	4.76	810	4,35

IV. Numerical implementation of the predicting method and discussion of the results

1. According to the table 2 in a year 2021 for variable X_k T°C (T°C is a temperature) build a sequence of points (x_i, y_i) $(i = 0, 1, \dots, 5)$, where $x_i = h_i$ – the values of variable depth h: $x_i = 2000 + i \cdot \Delta x$ $(\Delta x = 100)$, $i = 0, 1, \dots, 5$; $y_i = y(x_i)$ – the values of variable $y = X$. By sample $V = (x_i, y_i)$ $(i = 0, 1, \dots, 5)$ and by method of MNL build a multinomial regression model

$$y = f(x), \quad f(x) = \sum_{j=0}^{4} \beta_j x^j = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 \tag{1}$$

Entering the variables of $\tilde{X}_0 = 1$, $\tilde{X}_1 = x$, $\tilde{X}_2 = x^2$, $\tilde{X}_3 = x^3$, $\tilde{X}_4 = x^4$, we obtain relative variables of $\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_4$ of linear regression model

$$y = \sum_{j=0}^{4} \beta_j \tilde{X}^j \tag{2}$$

or in matrix notation of $Y = \tilde{X} \cdot \beta$, where $\tilde{X} = (\tilde{X}_{ij})$ – (6×5) – matrix of variable values $\tilde{X}_j$ in the i-th observation; the $Y = (y_0, y_1, \dots, y_5)$ is the vector of вектор observations of y variable. According to MNL, coefficient of regression $\beta = (\beta_0, \beta_1, \dots, \beta_4)^T$ (T – is the matrix transpose sign) calculated by the formula

$$\beta = \left(\tilde{X}^T \tilde{X}\right)^{-1} \cdot Y \tag{3}$$

β coefficients can be calculated using the LIN program in EXCEL with the determination of the error in calculating the predicted value of the output variable

$$\varepsilon_{\hat{y}} = RSS = \sum_{i=0}^{3} (\hat{y}_i - y_i)^2, \tag{4}$$

where $\hat{y}_i$ – predicted value of a variable by regression y_i

$$\hat{y}_i = \sum_{j=0}^{4} \hat{\beta}_j \cdot \tilde{X}_{ij}, \tag{5}$$

$\hat{\beta}_j$ – MNL- estimated β coefficient calculated by the formula (41).

If the LIN program is not available for estimating of β, then it calculated by formula (3) using a matrix inversion program to calculate the matrix of $\left(\tilde{X}^T \tilde{X}\right)^{-1}$. In our case if $n = 6$ at number $m = 4$ independent variables of $\tilde{X}_1, \dots, \tilde{X}_4$.

2. Total interval of $[x_0, x_5]$ divide the values of the variable x into intervals of $[x_{i-1}, x_i]$ $(i = 1, \dots, 5)$ with length of 100 m. In each of these intervals select an interpolation node of $x_i^* = x_i - \frac{10}{3}$ and construct the interpolation polynomial of Lagrange for $x_i^* \neq x_i$ $(i = 1, \dots, n; n = 5)$:

$$g_n(x) = \sum_{i=1}^{n} f(x_i^*) \prod_{j \neq i} \frac{x - x_j^*}{x_i^* - x_j^*} \tag{6}$$

Since $f^{(n)}(x) \equiv 0$ (due to the fact that the polynomial degree $l = 4$, for all $x \in [x_0, x_5]$, satisfying the condition $x \neq x_i^*$ $(i = 1, \dots, n)$), it is true that $f(x) \equiv g_n(x)$ and in particular for all

$$x_i = 2000 + i\Delta\tilde{x} \ (\Delta\tilde{x} = 10; i = 0,1, \dots, 50). \tag{7}$$

Therefore, for each variable X from table 3, by formula $f(x) = g_n(x)$ calculated values of $f(x)$ for all $x = x_i$, defined by the formula (7).

3. In the example under consideration, can assume (with a sufficiently small error) that $h_i = 2380$ – the maximum depth of immersion into the well of the bottom part of the pipe, which was accessible to inspection measurements in the depth range of [2000; 2500]. Further we will use normalized depth values of $\tilde{h}_i = h_i 2000$. Then $\tilde{h}_{i_0} = 1{,}19$.

4. Further we will assume that $i = 1, 2, \dots$ - length interval numbers $\Delta\tilde{x} = 0{,}005$ (10/2000) at the depth interval of $\tilde{h}_i$, equal to [1; 1,25] and i_0 - number of subsequent available length interval of Δi. For this example $i_0 = 38$.

Starting from the 1st line of table 1, these intervals can be numbered with a series of numbers $I = \{\varepsilon | i = 1, 2, \dots, i_0\}$, every fixed value i correspond to the x_i interval with length of $\Delta x = 0{,}005$.

There is a maximum loss of the pipe thickness (in fractions of a unit to three decimal places) along this interval.

5. By sample $V = \{x_i, y_0\}(i = 1, \dots i_0)$ built the best polynominal regressive model of $y = F(x)$ degree $l \leq 10$ by using of MNL. Let's denote $f(x) = F(x)$ ($F(x)$ – it is derivative function of $F(x)$). Let's apply the 4th order Runge-Kutta method with a step of $\Delta\tilde{h} = 0{,}005$, assuming i_0 is an even number (otherwise V is considered starting from i=2):

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{8}$$

$$k_1 = \Delta\tilde{h} \cdot f(x_i), \qquad k_2 = \Delta\tilde{h} \cdot f\left(x_i + \frac{1}{2}k_1\right),$$

$$k_3 = \Delta\tilde{h} \cdot f\left(x_i + \frac{1}{2}k_2\right), \qquad k_4 = \Delta\tilde{h} \cdot f(x_i + k_3).$$

By formula (8), knowing that $y_i = F(x_i)$ if $n = n_0$, it is possible to calculate the forecasting value of $y_{i_0+1} = F(x_{i_0+1})$. The forecast error is estimated according to the following formula

$$\rho_1 = \frac{1}{2880}\left|f^{(4)}(x_{i_0})\right|\left(\left(\Delta\tilde{h}\right)^5 + O(\Delta h)^6\right) \tag{9}$$

where $f^{(4)}(x_{i_0})$ – is a 4th derivative function of f(x) at a point x_{i_0}.

If the degree of the polynom $F(x)$ $l \leq 1$, then $f^{(4)}(x) - F^{(5)}(x) \equiv O$ and $\rho_1 = O\left(\left(\Delta\tilde{h}\right)^6\right)$, where $O((\Delta h)^6)$ – infinitesimal order of $\left(\Delta\tilde{h}\right)^6$.

Shifting the sample V forward by step Δh˜ (taking the obtaining predicted value y_{i_0+1} as the actual value), we will get predicted value of $y_{i_0+2} = F(x_{i_0+2})$ with predicted error

$$\rho_2 = \frac{1}{2880} \cdot \rho_1 \cdot \left(\Delta\tilde{h}\right)^5 + O\left(\Delta\tilde{h}\right)^6,$$

Calculated by fomula (8) with replacement on the right side of x_{i_0} by x_{i_0+1}. If $l \leq 4$ we also obtain that $\rho_2 = O\left(\Delta\tilde{h}\right)^6$. In this way, forecasts can be calculated by $y = F(x_{max})$, where $x_{max} = \tilde{h}_{i_{max}}$, $i_{max} = 50$ (limiting value in the table 3) in this case the forecast error $x_{n_0+\nu}$ will equal to $\rho_{n_0+\nu} = O\left(\Delta\tilde{h}\right)^{5(\nu_0-1)}$.

6. By sample of last 20 values y: $y_{i_0-19}, \dots, y_{i_0}$ fuzzify the variable Y. Denote that $y^{(1)} = y_{i_0-19}$, $y^{(2)} = y_{i_0-18}$, $y^{(N)} = y_{i_0}$ $(N = 20)$ and calculate for indexes $i = i_0 - 9$, $i_0 - 8$, ..., i_0 of value

$$\hat{a}^0 = \frac{1}{2}\left(\max_{1\leq i\leq n} y^{(i)} + \min_{1\leq i\leq n} y^{(i)}\right),$$
$$\hat{b}^0 = \frac{\left(\max_{1\leq i\leq n} y^{(i)} + \min_{1\leq i\leq n} y^{(i)}\right)}{2\varepsilon_\alpha}, \quad \varepsilon_\alpha = 2. \tag{10}$$

Then with probability $1 - \alpha$, $\alpha = \varepsilon^{-4}$ will satisfy the following ratio

$$\mu_Y(y) = \sigma(Y = y) = \exp\left[-\left(\frac{y - \hat{a}}{\hat{b}}\right)^2\right]. \tag{11}$$

İn order, $L_Y(y) = R_Y(y) = \exp\left[-\left(\frac{y-\hat{a}^0}{\hat{b}^0}\right)^2\right]$.

If $\hat{a}^0 < y^{(n)}$, then subsequent $V_0 = \{y^{(1)}, y^{(2)}, \dots, y^{(n)}\}$ shifts to right by 1 and replaced by subsequent of $V_1 = \{y^{(2)}, y^{(3)}, \dots, y^{(n)}, y^{(n+1)}\}$,
where $y^{(n+1)}$ at $n \geq n_0$ then has a forecast value calculated by formula (8).

By formula (41) is calculated new values of a^1, b^1 values of a and b etc. until, at a certain step of $\nu_{k,n}^0$ where the condition is satisfied of $\hat{a}^{\nu_0} \approx y^{(n)}$ with a certain error ε_0 (for example, $\varepsilon_0 = 10^{-2}$). Then, with the approximation error ε_0, can accept that the mode of the fuzzy number of $y^{(n)}$ is equal to $\hat{a}^{\nu_0}$.

Truth, for fixed depth of $\tilde{h}_{i_0}$ the value of $X_{n,k}$ of each fuzzy parameter of X_k at a depth of $\tilde{h} = \tilde{h}_i$ find number of $\hat{a}_{i,k}^{\nu_{i,k}^0}$ and $\hat{b}_{i,k}^{\nu_{i,k}^0}$ (denote their as $\hat{a}_{i,k}^{\nu_{i,k}^0}$ and $\hat{b}_{i,k}^{\nu_{i,k}^0}$), where $X_{n,k}$ represented by LR-form $X_{n,k} = \left(\hat{a}_{i,k}^{\nu_{i,k}^0}, \hat{b}_{i,k}^{\nu_{i,k}^0}, \hat{b}_{i,k}^{\nu_{i,k}^0}\right)$ with the membership function $\mu_{X_{n,k}}(y) = \exp\left[-\left(\frac{y-\hat{a}_{i,k}^{\nu_{i,k}^0}}{\hat{b}_{i,k}^{\nu_{i,k}^0}}\right)^2\right]$, if $\left|y - \hat{a}_{i,k}^{\nu_{i,k}^0}\right| < \hat{b}_{i,k}^{\nu_{i,k}^0}$ and $\mu_{X_{n,k}}(y) = 0$, if $\left|y - \hat{a}_{i,k}^{\nu_{i,k}^0}\right| > \hat{b}_{i,k}^{\nu_{i,k}^0}$.

7. Via $\hat{X}_k(\tilde{h})$ denote the output fuzzy variable in linear fuzzy regression with close-cut input $\tilde{h}$:

$$\hat{X}_k(\tilde{h}) = \beta_0^{(k)} + \beta_1^{(k)}\tilde{h}, \tag{12}$$

Which the coefficients $\beta_0^{(k)}$ and $\beta_1^{(k)}$ calculated by sample of $V = \{x_i, y_i\}$ $(i = 1, \dots, i_0)$, where $x_i = \tilde{h}_i$ – the depth interval length $\Delta\tilde{h} = 0.005$ with number of n, i.e. $x_i = 1 + i\Delta h$ and $\hat{Y}_i$ – the value of variable $\hat{Y} = \hat{X}_k$ at depth interval x_k, obtained by calculation according to the equation (12).

Denote $x_k \equiv 0$, $x_1 = \tilde{h}$ $(x_{k,0} \equiv 1, x_{k,1} = \tilde{h}_i)$, via X (10×2) – matrix of $X = (x_{i,j})$ $(i = i_0 - 8, \dots, i_0 + 1;\ j = 0,1)$. $Y = \left(Y_{i_0-8}, \dots, Y_{i_0+1}\right)^T$, $Y_i = X_k(\tilde{h}_i)$, $i = i_0 - 8, \dots, i_0 + 1$; $\beta = \left(\beta_0^{(k)}, \beta_1^{(k)}\right)^T$ (T – is matrix transpose sign). Then the value of $\hat{\beta}$ determined by the formula

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{13}$$

or calculated by the program of или LIN in EXCEL for simple linear regression with single input variable of x_1.

8. Will differ the fuzzy number of Y_i in the record, obtained by fuzzification using the formula (11), fron fuzzy number of $\hat{Y}_i$, obtained by calculation with regression (12).

Due to the symmetry of y the left span of $l_{Y_i}(\alpha)$, $\alpha \in [0,1]$ of the formula (49), will be equal to the right span $r_{Y_i}(\alpha)$, and $l_{Y_i}(1) = r_{Y_i}(1) = \hat{b}_i^{\nu_i^0}$.

9. Let's put $\alpha^* = 0{,}5$.

By the sample of $\left\{R(\tilde{h}_i), R\left(l_{Y_i}(\alpha^*)\right) : i = i_0 - 8, \dots, i_0 + 1\right\}$ will construct following regressive model

$$R\left(l_{Y_i}(\alpha^*)\right) = \beta_0(\alpha^*) + \beta_1(\alpha^*) R(\tilde{h}_i), \tag{14}$$

where $R(\tilde{h}_i)$ – rang of number $\tilde{h}_i$ in the subsequent of $\{\tilde{h}_i\}$ $(i = i_0 - 8, \dots, i_0 + 1)$, first number in ascending number order $\tilde{h}_i$, $R = 1$ rank is assigned, to the second $R = 2$ and so on; in the case of two identical numbers, equal for example r, in an unordered row, the first one is given rank of $R = {[(r-1)+r]}/{2}$, but the second rang represented by the following $R = {[(r+1)+r]}/{2}$.

For simple regression of (14), assum that $Y_i = l_{Y_i}(\alpha^*)$ and $R(X_i) = R(\tilde{h}_i)$ rank transformation (14) reduces to the following equation

$$R(Y_i) = \frac{(n+1)}{2} + \beta\left[R(X_i) - \frac{(n+1)}{2}\right], \tag{15}$$

where n – is the size of selected subsequent of $\{R(X_i), R(Y_i)\}$ (in this case $n = (i_0 + 1) - (i_0 - 8) + 1 = 10$).

Assume that $y_i = R(Y_i) - \frac{(n+1)}{2}$, $x_i = R(X_i) - \frac{(n+1)}{2}$, in this case we will get following regression equation $y_i = \beta \cdot x_i$, in which the MNL- is the value of $\hat{\beta}$ cofficient of β determined from the equation of $\frac{dZ}{d\beta} = 0$, $Z = \sum_{i=1}^n (y_i - \beta x_i)^2$, the solution of which is determined by the formula:

$$\hat{\beta} = {\sum_{i=1}^n y_i x_i}\Big/{\sum_{i=1}^n x_i^2} \tag{16}$$

Then from the equation of (15) find the predicted value

$$\hat{R}\left(l_{Y_i}(\alpha^*)\right) = \frac{(n+1)}{2} + \left[R(X_i) - \frac{(n+1)}{2}\right] \frac{\sum_{i'=1}^n \left[R(Y_{i'}) - \frac{(n+1)}{2}\right] \cdot \left[R(X_{i'}) - \frac{(n+1)}{2}\right]}{\sum_{i'=1}^n \left[R(X_{i'}) - \frac{(n+1)}{2}\right]^2}. \tag{17}$$

10. Let's calculate the values $\bar{l}_{\hat{Y}_i}(\alpha^*)$ by the following rules

$$\bar{l}_{\hat{Y}_i}(\alpha^*) = \begin{cases} l_{Y_{(1)}}(\alpha^*), \ if \ \hat{R}\left(l_{Y_i}(\alpha^*)\right) < R\left(l_{Y_{(1)}}(\alpha^*)\right), \\ l_{Y_{(n)}}(\alpha^*), \ if \ \hat{R}\left(l_{Y_i}(\alpha^*)\right) > R\left(l_{Y_{(n)}}(\alpha^*)\right), \\ l_{Y_{(j)}}(\alpha^*), \ if \ \hat{R}\left(l_{Y_i}(\alpha^*)\right) < R\left(l_{Y_{(j)}}(\alpha^*)\right). \end{cases} \quad (18)$$

If $R\left(l_{Y_{(j)}}(\alpha^*)\right) < \hat{R}\left(l_{Y_i}(\alpha^*)\right) < R\left(l_{Y_{(j+1)}}(\alpha^*)\right)$, then

$$\bar{l}_{\hat{Y}_i}(\alpha^*) = l_{Y_{(j)}}(\alpha^*) + \left(l_{Y_{(j+1)}}(\alpha^*) - l_{Y_{(i)}}(\alpha^*)\right) \frac{R\left(l_{Y_{(i)}}(\alpha^*)\right) - R\left(l_{Y_{(j)}}(\alpha^*)\right)}{R\left(l_{Y_{(j+1)}}(\alpha^*)\right) - R\left(l_{Y_{(j)}}(\alpha^*)\right)},$$

where $\bar{l}_{\hat{Y}_i}(\alpha^*)$ - is a value of the left span $l_{\hat{Y}_i}(\alpha^*)$ of the fuzzy number $\hat{Y}_i(\alpha) = Y_i(X_i)(\alpha)$, constructed by the rang regression (14); $l_{Y_{(j)}}(\alpha^*)$ - j^{th} is the ascending value of l of subsequent $\left\{l_{Y_{(i)}}(\alpha^*)\right\}$, $i = i_0 - 8, \dots, i_0 + 1$. Since the left-side value of the α-level set of a fuzzy number must be no more than its mode, then the estimation of the left-side range of the fuzzy number Yi (α^*) will be written as

$$\hat{l}_{\hat{Y}_i}(\alpha^*) = \min\{\bar{l}_{Y_i}(\alpha^*), \hat{y}_i\}, \quad (19)$$

where $\hat{y}_i$ – a priori estimation of the mode of a fuzzy number $\hat{Y}_i$, for which the assessment can be taken as $\hat{a}_{Y_i}^{(\nu_0)}$ mode of моды fuzzy number Y_i, obtained due to the fuzzification according to the formula (11). This estimate is subsequently corrected using a parametric estimator of span.

Similarly, according to the sample of $\{R(\tilde{h}_i), R(r_{Y_i}(\alpha^*): i = i_0 - 8, \dots, i_0 + 1\}$ with replacement in the formulas (17)-(19) of l by r will get value of right span of the fuzzy number $Y_i(\alpha^*)$ as following form

$$\hat{r}_{\hat{Y}_i}(\alpha^*) = \max\{\bar{r}_{Y_i}(\alpha^*), \hat{y}_i\}. \quad (20)$$

Under the numbers of $l_{Y_i}(\alpha)$ and $r_{Y_i}(\alpha)$ understood the projection to the y axes of intersection points with line $\mu = \alpha$ correspondingly of left branch $L(x)$ and rigth branch $R(x)$ membership functions of $\mu_{Y_i}(y)$ of fuzzy number Y_i.

In the formulas (17) and (19) mode of a_{Y_i} of the fuzzy number $Y_i(\alpha)$ satisfies, respectively, the equalities

$$\begin{aligned} a_{Y_i} &= l_{Y_i}(\alpha) + l \cdot L^{-1}(\alpha), \\ a_{Y_i} &= r_{Y_i}(\alpha) - r \cdot L^{-1}(\alpha), \end{aligned} \quad (21)$$

where $L^{-1}(\alpha) = l_{Y_i}(\alpha)$.

In the case, when the membership functions of of fuzzy number Y_i is represented by the Gaussian $\mu_{Y_i}(y) = \exp\left[-\left(\frac{y - a_{Y_i}}{b_{Y_i}}\right)^2\right]$ that get

$$l = r = b_{Y_i}. \quad (22)$$

11. Parametric assessment of spans and modes of the fuzzy number $Y_i(X_i)(\alpha)$ with $\alpha \neq \alpha^*$ at a close-cut input $X_i = \tilde{h}_i$ is constructed as follows.

Based on the obtained estimations of $\bar{r}_{\hat{Y}_i}(\alpha^*)$ and $\bar{l}_{\hat{Y}_i}(\alpha^*)$ are built following estimations

$$\begin{aligned} \hat{r}_{\hat{Y}_i}(\alpha) &= \begin{cases} \max\left\{\max_{\{\alpha \le s < \alpha^*\}}\{\bar{r}_{Y_i}(s), \hat{y}_i\}\right\}, \ if \ \alpha < \alpha^*, \\ \max\left\{\min_{\{\alpha^* < s \le \alpha\}}\{\bar{r}_{Y_i}(s), \hat{y}_i\}\right\}, \ if \ \alpha^* < \alpha, \end{cases} \\ \hat{l}_{\hat{Y}_i}(\alpha) &= \begin{cases} \min\left\{\max_{\{\alpha^* \le s < \alpha\}}\{\bar{l}_{Y_i}(s), \hat{y}_i\}\right\}, \ if \ \alpha^* < \alpha, \\ \min\left\{\min_{\{\alpha < s \le \alpha^*\}}\{\bar{l}_{Y_i}(s), \hat{y}_i\}\right\}, \ if \ \alpha < \alpha^*, \end{cases} \end{aligned} \quad (23)$$

Since $\hat{l}_{\hat{Y}_i}(\alpha)$, $\hat{r}_{\hat{Y}_i}(\alpha)$ increase by decreasing of α, then the $\min_\alpha \hat{l}_{\hat{Y}_i}(\alpha) = \hat{l}_{\hat{Y}_i}(0)$ and the $\max_\alpha \hat{l}_{Y_i}(\alpha) = \hat{l}_{\hat{Y}_i}(1)$; $\min_\alpha \hat{r}_{Y_i}(\alpha) = \hat{r}_{\hat{Y}_i}(1) = \hat{y}_i$ and $\max_\alpha \hat{r}_{\hat{Y}_i}(\alpha) = \hat{r}_{\hat{Y}_i}(0)$, in this case $L_{Y_i}\left(\hat{l}_{Y_i}(1)\right) = R_{Y_i}\left(\hat{r}_{Y_i}(1)\right) = 1$ and $\hat{y}_i$ – is the value of the fuzzy number $Y_i(X_i)(\alpha)$ where $\alpha = 0$, to be confirmed.

For each fixed i, intend sample data of $\left\{\hat{l}_{\hat{Y}_i}(\alpha_k), \alpha_k : k = 0, 1, \dots, k_0\right\}$ and $\left\{\hat{r}_{\hat{Y}_i}(\alpha_k), \alpha_k : k = 0, 1, \dots, k_0\right\}$ with increasing subsequent of $\{\alpha_\nu\}$ the value parameters α (for instance, $\alpha_k = k/10\,, k = 0, 1, \dots, 10$) and by the values of $\hat{l}_{\hat{Y}_i}(\alpha_\nu)$ and $\hat{r}_{\hat{Y}_i}(\alpha_\nu)$, determined by the formula (23), the measure of the fuzzy number $\hat{Y}_i = Y_i(X_i)$ is adjusted and the membership function $\mu_{\hat{Y}_i}(y)$ is approximated. However, if $\alpha = \alpha_\nu$ $L_{\hat{Y}_i}\left(\hat{l}_{\hat{Y}_i}(\alpha_k)\right) = \alpha_\nu$ $(k = 0, 1, \dots, k_0)$ ($L_{\hat{Y}_i}(y)$ – that is the left branch of function $\mu_{\hat{Y}_i}(y)$), then by the regression

$$L_{\hat{Y}_i}(y) = \beta_0^{(l)} + \beta_1^{(l)} \cdot y + \beta_2^{(l)} \cdot y^2 \tag{24}$$

Based on the sample of $\left\{\hat{l}_{\hat{Y}_i}(\alpha_k), \alpha_k : k = 0, 1, \dots, k_0\right\}$ it is possible to obtain of predicted value $\hat{L}_{\hat{Y}_i}(y)$ for all of $y \in supp\hat{Y}_i = \left[-\hat{b}_{Y_i}^{(\nu_0)}, \hat{b}_{Y_i}^{(\nu_0)}\right]$. Similarly, taking into account the equalities of $R_{Y_i}\left(\hat{r}_{\hat{Y}_i}(\alpha_k)\right) = \alpha_k, k = 0, 1, \dots, k_0$, the right branch is being restored $R_{Y_i}(y)$ by the following regression

$$R_{\hat{Y}_i}(y) = \beta_0^{(r)} + \beta_1^{(r)} \cdot y + \beta_2^{(r)} \cdot y^2. \tag{25}$$

Further, denoting through $\hat{y}_i^{(l)} = \hat{l}_{\hat{Y}_i}(1)$ and $\hat{y}_i^{(r)} = \hat{r}_{\hat{Y}_i}(1)$ obvious estimates of the mode of a fuzzy number, for evaluation of $\hat{y}_i$ the mode of the fuzzy number $Y_i(X_i)$ can be accept as following

$$\hat{y}_i = \frac{\hat{y}_i^{(l)} + \hat{y}_i^{(r)}}{2}. \tag{26}$$

The spans $l_{\hat{Y}_i}$ and $r_{\hat{Y}_i}$ of the fuzzy number $\hat{Y}_i = Y_i(X_i)$ estimate as following

$$\hat{l}_{\hat{Y}_i} = \hat{y}_i - L_{\hat{Y}_i}^{-1}(0), \quad \hat{r}_{\hat{Y}_i} = R_{\hat{Y}_i}^{-1}(0) - \hat{y}_i, \tag{27}$$

where $L_{\hat{Y}_i}^{-1}(0)$ and $R_{\hat{Y}_i}^{-1}(0)$ are solutions, respectively, of the equations $L_{\hat{Y}_i}(y) = 0$ and $R_{\hat{Y}_i}(y) = 0$. However, the fuzzy number $\hat{Y}_i$ represented by the LR-form as following

$$\hat{Y}_i = \left(\hat{y}_i, \hat{y}_i - L_{\hat{Y}_i}^{-1}(0), R_{\hat{Y}_i}^{-1}(0) - \hat{y}_i\right)_{LR} \tag{28}$$

Tuth, for the fuzzy variables $\hat{Y}_i = X_{ik}(\tilde{h}_i), i = i_0 - 8, \dots, i_0 + 1$ is obtained as the form (28) represented by the LR-form.

12. Suppose that the Y_i – is the variable, characterized by the maximum losses of the pipe thickness at the of $\tilde{h}_i$, $i = i_0 - 8, \dots, i_0 + 1$, the value of which $i = i_0 - 8, \dots, i_0$ determined in the 7th column of the table 1 and expressed in fractions of a unit (so if $i = i_0$ $Y_i = 0{,}346$, i.e. 34,6%), and the value Y_{i_0+1} calculated by the formula of predicting (8).

Let's make fuzzification of number Y_i $(i = i_0 - 8, \dots, i_0 + 1)$ by the formula (10) with appropriate modes $\hat{a}_{Y_i}^{(\nu_0)}$ and spans $\hat{b}_{Y_i}^{(\nu_0)}$.

To describe the main characteristics of a fuzzy number of $Y_i(X_i)$, $X = (X_1, X_2, X_3)$ – are vector of parametrs of the formation with the values from the table 3, integrated using the formula (6) whole of 10-meter depth scale by step $\Delta\tilde{h} = 0{,}005$ along all depth interval [1, 1.25].

Consider the last 11 values of the vector X at a depth $\tilde{h}_i$ with indexes $I = \{i | i = i_0 - 8, \dots, i_0 + 1\}$. By sample $\left\{l_{X_1}(\alpha), l_{X_2}(\alpha), l_{X_3}(\alpha), l_{Y_i}(\alpha)\right\}$, $i \in I$, let's construct a rank regression instead of (52)

$$R\left(l_{Y_i}(\alpha)\right) = \beta_0^{(l)}(\alpha) + \sum_{p=1}^{3} \beta_p^{(l)} R\left(l_{X_{ip}}(\alpha)\right) + \sum_{p'=1}^{3} \sum_{p=1}^{3} \beta_{pp'}^{(l)}(\alpha)\, R\left(l_{X_{ip}}(\alpha)\right) \cdot R\left(l_{X_{ip'}}(\alpha)\right) \tag{29}$$

however by selected data $\left\{r_{X_1}(\alpha), r_{X_2}(\alpha), r_{X_3}(\alpha), r_{Y_i}(\alpha)\right\}$, $i \in I$ – a rank regression.

$$R\left(r_{Y_i}(\alpha)\right) = \beta_0^{(r)}(\alpha) + \sum_{p=1}^{3} \beta_p^{(r)} R\left(r_{X_{ip}}(\alpha)\right) + \sum_{p'=1}^{3} \sum_{p=1}^{3} \beta_{pp'}^{(r)}(\alpha)\, R\left(r_{X_{ip}}(\alpha)\right) \cdot R\left(r_{X_{ip'}}(\alpha)\right) \tag{30}$$

From the equations (29) and (30) based on the MNL-estimations of their coefficient the predicted values $\hat{R}\left(l_{Y_i}(\alpha)\right)$ and $\hat{R}\left(r_{Y_i}(\alpha)\right)$ are found, of which using the formulas (18)-(23) the values of $\hat{r}_{\hat{Y}_i}(\alpha)$ and $\hat{l}_{\hat{Y}_i}(\alpha)$ are calculated, and using formulas (24)-(28) the left and right branches of the membership function and the LR-form of the fuzzy number $\hat{Y}_i = Y_i(X_i)$ are calculated.

13. Discrepancy between fuzzy numbers of Y_i и $\hat{Y}_i$ is calculated by the formula (11):

$$d\left(Y_i, \hat{Y}_i\right) = \frac{\int_{-\infty}^{\infty}\left|\mu_{Y_i}(x) - \mu_{\hat{Y}_i}(x)\right|dx}{\int_{-\infty}^{\infty}\mu_{Y_i}(x)dx} + h_d\left(Y_i(0), \hat{Y}_i(0)\right), \tag{31}$$

where $Y_i(0) = suppY_i$, $\hat{Y}_i(0) = supp$ – carriers of fuzzy numbers Y_i and $\hat{Y}_i$, represented by intervals

$$\begin{aligned} Y_i(0) &= \left[\hat{a}_{Y_i}^{(\nu_0)} - L_{Y_i}^{-1}(0),\ R_{Y_i}^{-1}(0) - \hat{a}_{Y_i}^{(\nu_0)}\right] \\ \hat{Y}_i(0) &= \left[\hat{y}_i - \hat{l}_{\hat{Y}_i}(0),\ \ \hat{r}_{\hat{Y}_i}(0) - \hat{y}_i\right] \end{aligned} \tag{32}$$

Let's denote for intervals of $Y_1(0)$ иand $Y_2(0)$ respectively via $[a_1, b_1]$ and $[a_2, b_2]$. The first term on the right side (31) characterizes the relative error of approximation of the membership function of the fuzzy number Yi by the membership function of the fuzzy number $\hat{Y}_i$. In order for the second term to characterize the relative error in approximating the interval $Y_i(0)$ to the interval of $\hat{Y}_i(0)$ will wrote as a following form

$$\begin{aligned} h_d\left(Y_i(0), \hat{Y}_i(0)\right) = \Big\{\frac{1}{2}\left[\left(\hat{y}_i - \hat{l}_{\hat{Y}_i}(0)\right) + \left(\hat{r}_{\hat{Y}_i}(0) - \hat{y}_i\right)\right] - \\ -\frac{1}{2}\left(\hat{a}_{Y_i}^{(\nu_0)} - L_{Y_i}^{-1}(0)\right) + \left(R_{Y_i}^{-1}(0) - \hat{a}_{Y_i}^{(\nu_0)}\right)\Big\} \cdot \frac{1}{R_{Y_i}^{-1}(0) + L_{Y_i}^{-1}(0) - 2\hat{a}_{Y_i}^{(\nu_0)}}, \end{aligned} \tag{33}$$

where is the first multiplier in (33) is the distance between the centers of the intervals $\hat{Y}_i(0)$ and $Y_i(0)$, and the second multiplier – is an inverse number to the length of the interval Yi (0) Simpson's compound rule. Assume that the points of $a = t_0 < t_1 < t_2 < \cdots < t_{s-1} \approx b$ split the cut $[a, b]$ to the *s*-subcutes (elementary cuts). Having put the $h = (b - a)/2s$, $t_i = a + 2th$ will obtain the following

$$I = \int_a^b f(x)dx = \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} f(x)dx.$$

Simpson's compound formula is written as following

$$\begin{aligned} I = \frac{h}{8}[f(a) + 4f(a+h) + 2f(a+2h) + 4f(a+3h) + 2f(a+4h) + \\ +4f(a+5h) + f(a+6h)] - (b-a)\cdot\frac{h^4}{180}\cdot f^{(4)}(a+5h), \quad h = \frac{b-a}{4}, \end{aligned} \tag{34}$$

with remaining members

$$R_{2s+1} = -(b-a)\frac{h^4}{180}\cdot f^{(4)}\left(a + \frac{5}{4}(b-a)\right). \tag{35}$$

With increasing number of 5 elementary cuts $R_{2s+1} \to 0$.

The integral in the denominator of the fraction on the right-side of (69) is written as

$$\int_{-\infty}^{\infty}\mu_{Y_i}(x)dx = \int_{a_1}^{b_1} f(x)dx\,, \quad f(x) = \exp\left[-\left(\frac{x - \hat{a}_{Y_i}^{(\nu_0)}}{\hat{b}_{Y_i}^{(\nu_0)}}\right)^2\right]. \tag{36}$$

In this case, the 4th derivative is written as following

$$f^{(4)}(x) = \left[-2\left(\frac{x-a}{b}\right)\cdot\frac{1}{b}\right]^4 \exp\left[-\left(\frac{x - \hat{a}_{Y_i}^{(\nu_0)}}{\hat{b}_{Y_i}^{(\nu_0)}}\right)^2\right].$$

By choosing s, any degree of smallness of the remainder term R_{2s+1} is achieved in (35).

To calculate the integral in the numerator on the right side of (68), calculate $a = \min\{a_1, a_2\}$, $b = \max\{b_1, b_2\}$ and calculate of

$$\int_{-\infty}^{\infty}\left|\mu_{Y_i}(x)-\mu_{\hat{Y}_i}(x)\right|dx=\int_{a}^{b}\left|\mu_{Y_i}(x)-\mu_{\hat{Y}_i}(x)\right|dx, \tag{37}$$

where $\mu_{Y_i}(x)$ is the $f(x)$ function from (36), but $\mu_{\hat{Y}_i}(x)$ has a left $L_{\hat{Y}_i}(x)$ and right $R_{\hat{Y}_i}(x)$ branches, modelling by the regression of (24) and (25), respectively. Therefore, the derivatives $\mu_{\hat{Y}_i}(x)\equiv 0$ and remaining members of (35) for the function $f(x)=\left|\mu_{Y_i}(x)-\mu_{\hat{Y}_i}(x)\right|$ is completely determined by the 4th derivative of the function of

$$\mu_{Y_i}(x)=\exp\left[-\left(\frac{x-\hat{a}_{Y_i}^{(\nu_0)}}{\hat{b}_{Y_i}^{(\nu_0)}}\right)^2\right].$$

Tuth, the relative approximation accuracy (RAA) of the fuzzy number Y_i with a fuzzy number $\hat{Y}_i$ expressed as the following

$$\mathrm{RAA}\left(Y_i=\hat{Y}_i\right)=1-d(Y_i,\hat{Y}_i), \tag{38}$$

moreover, equality (38) is true with accuracy (38) and confidence probability

$$P_{\partial}\left(Y_i=\hat{Y}_i\right)=\left(\overset{0}{\alpha}\right)^{\nu_0},\quad \overset{0}{\alpha}=e^{-4}, \tag{39}$$

where ν_0 – iteration step, on which the assessment $\hat{a}_{Y_i}^{(\nu_0)}$is achieved for the mode of the number Y_i fuzzified by formula (11).

14. According by the formula (30) the relative approximation of $\hat{Y}_i$ to Y_i is calculated, in the case, when the Y is the parametr of the formation X_k and X is the depth of the well and $Y_i=X_{ik}$ and $X_i=h_i$ (let's denote their as $d\left(X_{ik},\hat{X}_{ik}\right)$) for each parametrs of the formation X_k $(k=1,2,3)$ and in the case when Y is the maxumum losses of the pipe thickness and the $X=(X_1,X_2,X_3)$ is the vector parametrs of the formation and $Y_i=Y_i(X_i)$, $X_i=(X_{i1},X_{i2},X_{i3})$ (let's denote their as $d\left(Y_i,\hat{Y}_i\right)$. Taking into account the error ρ_1 depth predictions h_i valuves of y_{i_0+1} macsimum losses of the thickness, determined by the formula (9), total (general) error of predictions $\hat{Y}_{i_0+1}$ are calculated as following

$$\varepsilon_1=\rho+\left(\prod_{k=1}^{3}d\left(X_{i_0+1,k},\hat{X}_{i_0+1,k}\right)\right)\cdot d\left(Y_{i_0+1},\hat{Y}_{i_0+1}\right), \tag{40}$$

where $\hat{X}_{i_0+1,k}$ –predicted parameter value X_k at a depth h_i by regression (12) and Y_{i_0+1} – predicted value of the maxumum losses of thickness, determined by the rank regressions (24) and (25). To error of (40) corresponds to relative approximation accuracy $\hat{Y}_{i_0+1}$ to the true (unobserved) value Y_{i_0+1} (i.e. the accuracy of predict $\hat{Y}_{i_0+1}$), defined by equality

$$\mathrm{RAA}\left(\hat{Y}_{i_0+1},Y_{i_0+1}\right)=1-\varepsilon_1, \tag{41}$$

which is valid with confidence probability

$$P_{\partial,1}=\left[\prod_{k=1}^{3}P_{\partial,k}\left(X_{i_0+1,k}=\hat{X}_{i_0+1,k}\right)\right]\cdot P_{\partial}\left(Y_{i_0+1},\hat{Y}_{i_0+1}\right), \tag{42}$$

where $P_{\partial,k}\left(X_{i_0+1,k}=\hat{X}_{i_0+1,k}\right)=(e^{-4})^{\nu_0,i_0+1,k}$ $(\nu_0,i_0+1,k$ – is iteration step, on which achieved by the way of fuzzification mode estimation of a fuzzy number $\hat{X}_{i_0+1,k})$ and $P_{\partial}\left(Y_{i_0+1},\hat{Y}_{i_0+1}\right)=(e^{-4})^{\nu_0,i_0+1}$ $(\nu_0,i_0+1$ – is iteration step, on which achieved by the way of fuzzification mode estimation of a fuzzy number $\hat{Y}_{i_0+1})$. Printing the forecast y_{i_0+1}.

According to the rolling forecasting scheme, shifting a set of indices $I=\{i|i=i_0-8,\dots,i_0+1\}$ forward one step $\Delta i=0{,}005$, we will get a set of indices $I_1=\{i|i=i_0-9,\dots,i_0+2\}$.

Then the forecasting error $\hat{Y}_{i_0+2}$ will be equal

$$\varepsilon_2 = \varepsilon_1 \cdot \rho_2 \cdot \left(\prod_{k=1}^{3} d\left(X_{i_0+2,k}, \hat{X}_{i_0+2,k}\right) \right) \cdot d\left(Y_{i_0+2}, \hat{Y}_{i_0+2}\right), \quad (43)$$

where ρ_2 – is forecasting To error of value y_{i_0+2} maxumum losses of thickness, determined by the formula (19) and following

$$\mathrm{RAA}\left(\hat{Y}_{i_0+2}, Y_{i_0+2}\right) = 1 - \varepsilon_2$$

with confidence probability

$$P_{\partial,2} = \left(\prod_{k=1}^{3} P_{\partial,k}\left(X_{i_0+2,k}, \hat{X}_{i_0+2,k}\right) \right) \cdot P_{\partial}\left(Y_{i_0+2}, \hat{Y}_{i_0+2}\right), \quad (44)$$

At step m of this algorithm, a forecast of values y_{i_0+m} will be obtained with the error

$$\varepsilon_m = \varepsilon_{m-1} \cdot \rho_m \cdot \left(\prod_{k=1}^{3} d\left(X_{i_0+m,k}, \hat{X}_{i_0+m,k}\right) \right) \cdot d\left(Y_{i_0+m}, \hat{Y}_{i_0+m}\right), \quad (45)$$

If $\varepsilon_m < \varepsilon_0$ (ε_0 – is the permissible forecast error), then $m = m + 1$ and go to point 6.). Otherwise STOP. We print.

Using this algorithm, the following predictions were obtained:

$$i = i_0 + 1 = 39 \text{ (2390 м)},\ \hat{y}_{39} = 0{,}357;$$
$$i = i_0 + 1 = 40 \text{ (2400 м),)},\ \hat{y}_{40} = 0{,}362.$$

Thus, based on the above studies, the following main conclusions can be drawn.

V. Conclusion

1. Inspection measurements carried out on the basis of studying the electromagnetic field created by a sounde located inside casing or pumping and compression pipe (tubing) require significant costs, especially at large well depths and an increased degree of aggressiveness of their environment. In this regard, the problem of predicting the results of the inspection at greater depth of immersion of the pipe into the well becomes fundamentally important.

2. For this purpose, we have proposed a new methodology for predicting the maximum loss of pipe thickness by step of 10 m after any certain depth of immersion of the bottom part of the pipe into the well, for which the maximum loss of thickness (in percentage) is considered given.

3. Forecasting is carried out by sliding in the direction of increasing losses with a forward shift of depth by 10 m. The error of the resulting forecast is estimated using a numerical procedure for calculating the measure of difference between fuzzy numbers.

References

[1] Koskov, V.N. Comprehensive assessment of the condition and operation of oil wells using field geophysical methods: textbook. allowance /V.N. Koskov, B.V. Koskov, I.R. Yushkov. Perm: Perm State Technical University Publishing House, 2010, – 226.

[2] Huseyinov, G.G. Accelerated technology for eliminating accidents in wells and the mechanisms used in this case. Oilfield business, 2012, 8, 36-38.

[3] Technical guidance on the application of geophysical surveys and operations in oil and gas wells. SOCAR guidance document. Baku, 2019, 320.

[4] Feyziyeva, G.E Results of using geophysical methods in assessing the operating condition of oil and gas wells. Equipment, Technologies, Materials, 2022, 38-42.