

# Mathematical simulation of the rate of carbon dioxide corrosion at the facilities of Gazprom dobycha Urengoy LLC

Aleksandr Yusupov<sup>1,\*</sup>

<sup>1</sup>Gazprom dobycha Urengoy LLC, Novy Urengoy, Russia, 629300

**Abstract.** In Gazprom dobycha Urengoy LLC, as in other oil-and-gas production enterprises, there are problems of increased equipment wear due to corrosion. A special role there plays CO<sub>2</sub> corrosion. Despite the homogeneity of the extracted fluid and even chemical composition of the working medium, the nature and intensity of corrosion damage to pipelines and equipment varies over a wide range, due to different thermobaric parameters of well operation. To determine parameters influencing the rate of corrosion most different methods of statistical analysis were used. The paper provides a methodology for compiling a mathematical model and assessing its reliability. As a result, the equation of carbon dioxide corrosion in relation to the conditions of Achimov deposits of Urengoy oil, gas and condensate field was obtained. The type of the obtained equation was chosen according to the model of the classical de Waard-Milliams carbon dioxide corrosion equation. The model proposed by the authors describes the processes of carbon dioxide corrosion more reliably than the de Waard-Milliams equation does. The disadvantage of the developed model is that it does not reliably describe the speed of corrosion in wells with corrosion rates, significantly exceeding the average values for all wells studied.

## 1 Introduction

In October 2009, the gas condensate facility was launched developing Achimov deposits new for PJSC Gazprom. Achimov deposits are characterized by abnormally high reservoir parameters: pressure is up to 60 MPa, and the temperature is up to 106°C.

After 5 years of operation the first corrosive damage to the inner surface of the wellhead equipment elements and pipelines were identified. Figure 1 shows one of the defects. The cause of these defects was carbon dioxide corrosion.



**Fig. 1.** Photo of corrosion defect of a tool flange.

The partial pressure of carbon dioxide at the wellhead exceeds 0.21 MPa. Thus, the environment has a high degree of aggressiveness in accordance with various regulatory documents (NACE SP 0106-2006, GOST R 51365-2009, STO Gazprom 9.3-011-2011) [1-3]. In gas collecting pipelines, the partial pressure of

carbon dioxide is reduced to the level of 0.1 MPa but the presence of free water and the stratified flow conditions cause the corrosion process along the lower pipe generatrix.

At the facilities of Gazprom dobycha Urengoy LLC exposed to carbon dioxide corrosion a corrosion monitoring system is implemented, which allows to measure the corrosion rate in different parts of the gas collection system [4,5]. According to the measurement results, the following pattern was noted. In areas up to the pressure regulator, characterized by high temperatures and pressures, the rate of corrosion flow is higher than in the areas after the pressure regulator, where thermobaric parameters are lower. This fact is explained by the fact that the rate of carbon dioxide corrosion depends on the temperature and partial pressure of the gas, which in turn depends on the operating pressure.

A large number of theoretical, empirical and semi-empirical models exist to describe the corresponding dependencies [6,7]. The model de Waard-Milliams is considered the most popular and - already - classic [8,9]. However, the use of this model for the conditions of Achimov deposits gives excessive corrosion rates relative to the actual corrosion rates. Based on this, the authors of the article faced the task of developing their own model describing the rate of carbon dioxide corrosion in relation to the conditions of Achimov deposits.

\* Corresponding author: [a.d.yusupov@gd-urengoy.gazprom.ru](mailto:a.d.yusupov@gd-urengoy.gazprom.ru)

## 2 Materials and Method

To carry out a multi-factor analysis assessing the impact of various parameters on the corrosion rate, the following data were selected: the results of chemical analyses of formation water samples, the results of gas condensate studies, the average values of pressure, temperature and flow rates for each well, corrosion rates obtained by gravimetric method. All data were summarized in a matrix consisting of 72 rows and 28 columns, where 72 is the number of observations; 28 - the number of factors that could have a possible impact on the corrosion rate.

The generated data array has undergone a preprocessing procedure, including:

1. Elimination of missing observations. As part of this operation, we removed from the data set observations, in which at least one of the factors was missing. For example, wells that lack data on the chemical composition of water were excluded from the data set.

2. Transformation of qualitative features into numerical (binary) ones. For example, one of the layers from which the fluid was extracted was assigned the value 0, the other layer – 1.

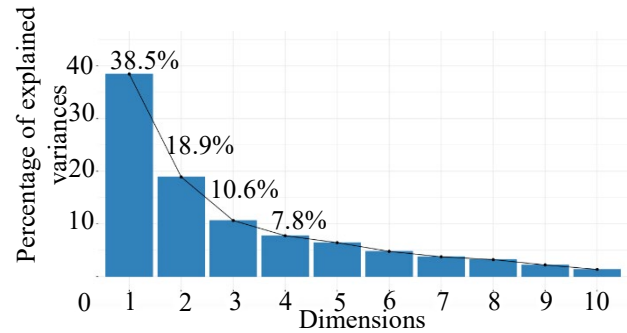
3. Application of bootstrapping method. Due to the large amount of data available for multi-factor analysis, the bootstrapping method was used to artificially increase the amount of information and improve the quality of the future model [10,11]. The essence of the method consists of forming a set of samples based on a random selection with repetitions and ultimately the entire sample assumes the normal form of distribution, which allowed further use of standard methods of mathematical statistics and data analysis.

After preliminary data processing for each analyzed factor, mean values and mean square deviation (hereinafter referred to as MSD) were found. For some factors, the MSD exceeded the average value of the factor, which indicates a large level of spread in values.

The next step was to conduct a factor analysis, the purpose of which is to find the factors that most affect the corrosion rate. For its implementation, the principal component method was chosen as the most frequently used due to ease of use and transparency. This method was used to exclude the maximum number of factors from the model, because a large number of variables in the mathematical model makes it variable, and the reliability of such a model may be in doubt.

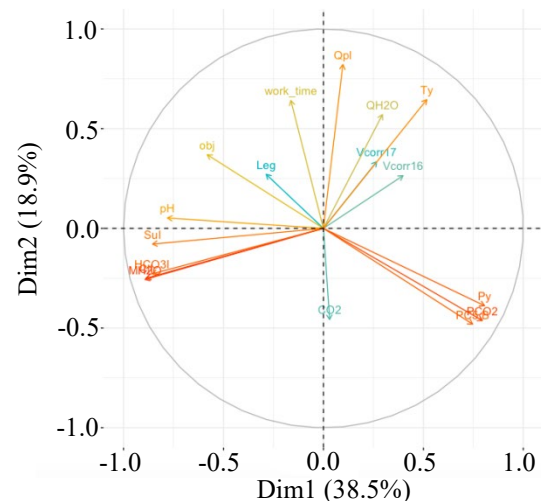
In the RStudio software product, a diagram was built (Figure 2), estimating the contribution of the total variance in each of the main components.

The higher the resulting column, the more variations include the main component being analyzed. The main components are ordered in descending order. Thus, the first two main components account for about 60% of the total number of variations.



**Fig. 2.** Main components with percentage of total variation in data.

Further, the contribution (in percent) of each of the analyzed factors to the first two main components was estimated, in other words, how strongly the analyzed factor affects the main components. After that, a correlation map was drawn in the software system (Figure 3).



**Fig. 3.** Correlation map.

The correlation map indicates positively correlated factors (grouped together), negatively correlated (located in opposite quadrants), as well as the level of variability and significance of factors: the location closer to the edge of the circle indicates greater variability, while the location closer to the center indicates lesser variability. The obtained diagram shows that the corrosion rate in 2016 (Vcorr16) and in 2017 (Vcorr17) has a relationship with the temperature (Ty) and flow rate (Qpl), also has a positive correlation with the first main component with the operating pressure (Py) and, accordingly, the partial pressure of carbon dioxide (PCO2). Some other factors that have a positive correlation with the corrosion rate (water content – QH2O, the content of hydrocarbon components C<sub>5+B</sub> – PC5B) have an overestimated mean square deviation and cannot be included in the model. Thus, the hypothesis that the maximum influence on the corrosion rate is exerted by the temperature, pressure and flow rate of wells was previously formulated. To confirm this hypothesis, we started the construction of a multidimensional regression model.

Further, also using RStudio software, a model was compiled. For this, all factors included in the original matrix were used. The result of this work was a multidimensional regression model presented in the form of a table indicating the regression coefficients, MSD for each factor, the values of the Student's t-test, p-values (significance level) for each criterion. All of the above criteria allow us to estimate the statistical significance of each factor and further exclude the least significant factors from the model.

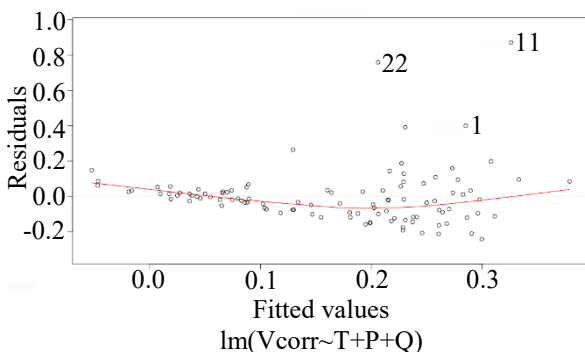
On the basis of the obtained model, it was decided to reduce the model, in particular, all factors were excluded from the model except for the temperature, pressure and flow rate of the well. After that, a model was built that includes only the above factors (Table 1).

**Table 1.** The results of calculating the parameters of the linear regression model in the first approximation.

	Estimate	Std. Error	t value	Pr(> t )
Intercept	-0.3318	0.0911	-3.642	0.0004
Temperature	0.0098	0.0029	3.358	0.0011
Pressure	0.0021	0.0029	0.728	0.4680
Flow rate	0.0021	0.0030	0.684	0.4953

The next stage of the study was to assess the reliability of the model. In this stage, we built three diagrams.

Using the Gauss – Markov theorem, an array of data was formed, on which the first diagram was built (Figure 4), which allows us to evaluate the correctness of the choice of the model type.

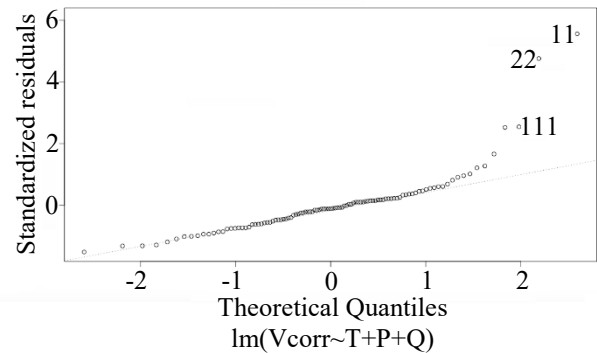


**Fig. 4.** The dependence of predicted values of rate of corrosion on the standardized residues.

The resulting dependence indicates that the model type (linear relationship) was chosen correctly. In the case of a different nature of the desired dependence, the trend line (red line) would have a geometry that is different from the linear one. Also on this diagram, it is clearly visible that there are anomalous points (No. 1, 11, 22) that do not fall on the model. The greater the spread of points from the trend line, the less adequate is the resulting model.

The following diagram (Figure 5) is a scattering diagram of observed and expected (standardized) values with a corresponding specified distribution. If the observed values fall on a straight line, then the

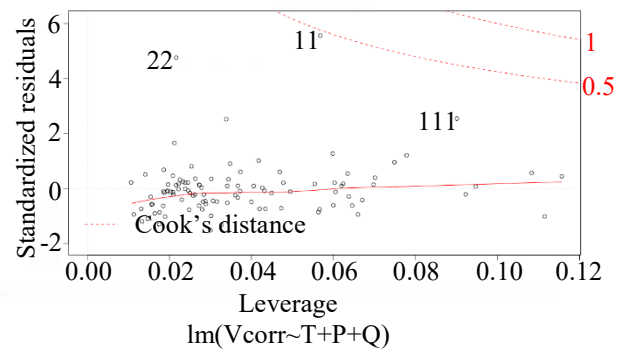
theoretical distribution is well suited with the observed data.



**Fig. 5.** Q-Q diagram of the model before detecting anomalous values.

A very close correlation with the line, except for the anomalous points (№ 11, 22, 111), indicates that the distribution of samples is normal.

The third diagram (Figure 6) shows the dependence of the degree of influence of each observation on the result of the corrosion rate on the error for each observation.

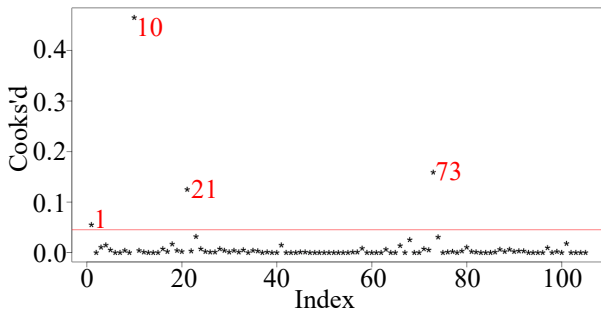


**Fig. 6.** Graphic interpretation of Cook's distance for the model.

In addition, the diagram shows a dotted line of Cook's distance (in the upper right corner). The closer to these lines the observation is, the more this observation shifts the predicted corrosion rate from the actual value. In accordance with this diagram, there are anomalous observations (№ 11, 22, 111), which coincide with the previous two diagrams.

Further, using the Cook's distance determination method [12], anomalous corrosion rate values were revealed, creating a high variability of the regression model. Cook's distance shows the difference between the calculated coefficients of the regression equation and the values that would be obtained by excluding the corresponding observation. If all Cook's distances are not the same, which indicates the adequacy of the model, we can assume that this observation shifts the estimates of the regression coefficients. Figure 7 shows the results of the algorithm in identifying anomalous values for corrosion rates. Numeric designations correspond with the index identifier of a well with an anomaly. Through

this method, four anomalous observations located above the red line were excluded from the model.



**Fig. 7.** Graphic interpretation of the Cook's distance for the model.

The next step was a similar construction of the regression model in the absence of anomalous observations (Table 2).

**Table 2.** The results of calculating the parameters of the linear regression model after eliminating anomalous observations.

	Estimate	Std. Error	t value	Pr(> t )
Intercept	-0.2011	0.0538	-3.737	0.0003
Temperature	0.0064	0.0017	3.633	0.0004
Pressure	0.0030	0.0018	1.685	0.0951
Flow rate	-0.0001	0.0018	-0.060	0.9524

Based on the evaluation of the statistical significance of each factor, it was decided to exclude the flow rate from the model due to the lack of a sufficient level of significance for the above predictor - 0.9524. After that, the model was rebuilt and its reliability was evaluated similarly to the original model. After rebuilding, the dispersion of the actual values from the predicted corrosion rates visually decreased, which is an indicator of the adequacy of the model obtained.

It is known that in the classical equation of an estimation of carbonic acid corrosion of de Waard-Milliams (1) there are the same variables as in the model obtained by authors. The only difference is that in the de Waard-Milliams equation, the pressure is represented as the partial pressure of carbon dioxide but because the CO<sub>2</sub> content is the same for all wells, respectively, the operating pressure is directly proportional to the partial pressure of CO<sub>2</sub>.

$$\lg V_{cor} = 5.8 - \frac{1710}{T} + 0.67 \lg P_{CO_2} \quad (1)$$

Therefore, the next stage of the work was the construction of a model according to the type of the de Waard-Milliams equation. Refined model parameters are presented in Table 3.

**Table 3.** Parameters of the regression model built by the type of the de Waard-Milliams equation.

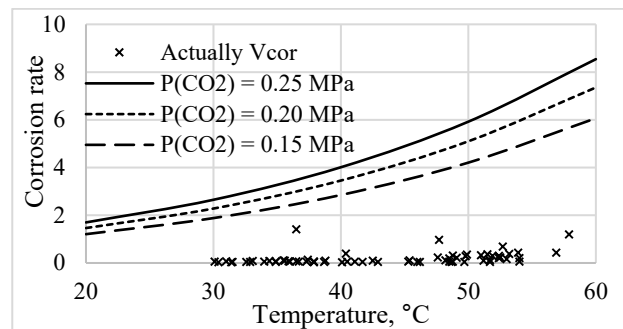
	Estimate	Std. Error	t value	Pr(> t )
Intercept	6.37	1.9966	3.191	0.0019
1/T	-2377	612.0577	3.885	0.0001
ln(P)	0.52	0.2974	1.754	0.0826

### 3 Results

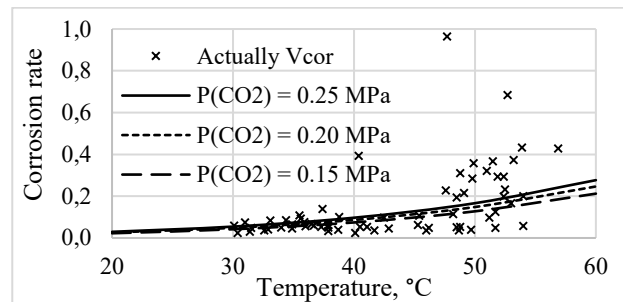
Thus the equation describing the dependence of corrosion rate on the pressure and temperature conditions for Area 2 of the Achimov deposits in the Urengoy oil, gas and condensate field can be written in the form proposed by de Waard-Milliams as follows:

$$\lg V_{cor} = 6.37 - \frac{2377}{T} + 0.52 \lg P_{CO_2} \quad (2)$$

Figures 8-9 show a graphical comparative analysis of the actually measured corrosion rates with the predicted corrosion rates for the two models.



**Fig. 8.** Dependence of the predicted corrosion rate according to the de Waard-Milliams model from the temperature.



**Fig. 9.** Dependence of the predicted corrosion rate of the developed model on temperature.

### 4 Discussion

In the model, all members are statistically significant predictors. The value of the determination coefficient (R<sup>2</sup>) for the obtained model is 0.47. Taking into account the fact that all the initial data were obtained in field conditions, the model can be considered sufficiently adequate.

The model developed by the authors describes the processes of carbon dioxide corrosion in relation to objects of Achimov deposits with a better correlation than the de Waard-Milliams model, despite the fact that the de Waard-Milliams equation has a theoretical basis on which the results of laboratory studies are imposed.

The resulting equation is of practical importance at the stage of commissioning of new wells to predict corrosion rates until it is possible to actually measure the corrosion rate.



## 5 Conclusion

1. A model describing the processes of carbon dioxide corrosion in relation to objects of Achimov deposits of the Urengoy oil, gas and condensate field was developed;

2. To identify the main factors affecting the corrosion rate, a statistical analysis was carried out using the principal component method and the construction of multidimensional regression models. It is established that the factors most correlated with the corrosion rate for the conditions of Achimov deposits are temperature and pressure;

3. An anomaly search algorithm was implemented based on the Cook's Distance determination. Using this algorithm, four observations were excluded from the total sample;

4. After exclusion of the anomalous wells an updated regression model was built by the type of de Waard-Milliams equations;

5. The proposed model describes the processes of carbon dioxide corrosion more accurately than the de Waard-Milliams equation;

6. The disadvantage of the developed model is that it does not describe the corrosion rate in wells with abnormal values of the corrosion rate.

## List of symbols

$V_{cor}$	corrosion rate, mm/y
$P_{CO_2}$	partial pressure, bar
$T$	temperature, K

## References

1. NACE Standard SP 0106 (latest revision), "Control of internal corrosion in steel pipelines and piping systems" (Houston, TX: NACE)
2. GOST R 51365-2009 Oilfield Mining wellhead. General specifications (Standartinform, Moscow, 2015)
3. System standarts of Gazprom 9.3-011-2011. Corrosion protection. Inhibitory protection against corrosion of field objects and pipelines. Primary requirements (Gazprom Expo, Moscow, 2011)
4. V. Yu. Artemenkov, A. Yu. Koryakin, D. V. Dikamov, E. O. Shishkov, A. D. Yusupov, Organization of corrosion monitoring at the sites of the second section of the Achimov deposits of the Urengoy oil and gas condensate field, GAS Industry of Russia, Special Issue No. 2, p. 74-78 (2017)
5. A. Yu. Koriakin, V. F. Kobychyev, I. V. Kolinchenko, A. D. Yusupov, Conditions of the carbon dioxide corrosion on the production facilities of achimovskie deposits, methods of monitoring and forecasting, GAS Industry of Russia, No. 12, p. 84-89 (2017)
6. S. Nestic, Key issues related to modelling of internal corrosion of oil and gas pipelines, A review, Corrosion Science **49**, p. 4308-4338 (2007)
7. R. Nyborg, CO<sub>2</sub> corrosion models for oil and gas production systems, Corrosion/2010, Paper No. 10371 (Texas, TX: NACE International, 2010)
8. C. de Waard, D.E. Milliams, Carbonic Acid corrosion of steel, Corrosion **31**, No. 5, p. 175 (1975)
9. C. de Waard, U. Lotz, Prediction of CO<sub>2</sub> corrosion of carbon steel, Corrosion/1993, Paper No. 69 (Houston, TX: NACE International, 1993)
10. S. Streukens, S. Leroi-Werelds, Bootstrapping and PLS-SEM: A step-by-step guide to get more out of your bootstrap results, European Management Journal **34** (6), p. 618-632 (2016)
11. S. Basu, Improved small sample inference procedures for epidemiological parameters under cross-sectional sampling. Journal of the Royal Statistical Society: Series D (The Statistician) **50** (3), p. 309-319 (2001)
12. Cook R. Dennis, Detection of Influential Observations in Linear Regression, Technometrics **19**, No. 1, p. 15-18 (1977)