Prediction of Corrosion Rate of Submarine Oil and Gas Pipelines Based on IA-SVM Model

WANG Chen¹, MA Gang¹, LI Junfei², DAI Zheng¹, LIU Jinyuan³

¹School of Petroleum Engineering, Xi'an Shiyou University, Xi'an, ShaanXi, 710065, China;
²China Petroleum and Natural Gas Pipeline Bureau;
³School of Petroleum and Natural Gas Engineering, Southwest Petroleum University

Corresponding author: Li Junfei, male, engineer, born in 1983, he graduated from the China University of Petroleum (East China) mechanical engineering and automation, is mainly engaged in research work of the professional direction of oil and gas pipeline construction. Address: No. 87, Guangyang Road, Langfang City, Hebei Province, China, 065000. Tel: 15903162921; Email: lijunfei@cppndc.com.

About the author: Ma Gang, male, born in 1995, is a master of Xi'an Petroleum University. Now he is mainly engaged in the research of pipeline intelligent algorithm application and pipeline safety evaluation. Address: Xi'an Petroleum University, Electronic Second Road, Yanta District, Xi'an, Shaanxi Province, China, 710065. Tel: 13571995451, Email: mag.10@163.com.

*Corresponding author’s e-mail: 837786304@qq.com

Abstract. In view of the corrosion problem of submarine oil and gas pipelines, this paper proposes a predictive model based on support vector machine (SVM), according to the related factors affecting the corrosion of submarine oil and gas pipelines, and uses the immune algorithm based on the problem of relatively low corrosion rate and influencing factors. (IA) prefers the penalty parameters C and ε of the SVM. The IA-SVM model is combined to form the IA-SVM model, and the number of hidden nodes and the kernel function of the SVM are optimized based on the absolute error. Finally, the model is verified according to the actual corrosion rate of the submarine pipeline in a certain sea area of China, and with the PSO-SVM, the GA-SVM and LS-SVM models are used to compare the prediction errors to verify the feasibility and advancement of the IA-SVM model. The research shows that the preferred results of IA for SVM penalty parameters C and ε are 43,6213 and 0.0483, the preferred result of SVM hidden layer nodes is 260, and the kernel function preferred result is Sigmoidal function. At this time, the predicted mean absolute error and root mean square error of the combined model are 1.45% and 0.0159165, respectively, the error of the model is smaller than other prediction models. The research results show that the prediction error of the corrosion rate of submarine oil and gas pipelines based on IA-SVM model is relatively small, and the data training time is short, which can be used to predict the corrosion rate of submarine oil and gas pipelines.

1. Introduction

It is a feasible way to transport offshore oil and gas resources to port cities through oil and gas pipelines.
However, due to the complex marine environment, it will cause serious corrosion to oil and gas pipelines [1-2]. For onshore pipelines, the corrosion status of the pipeline can be understood through various internal detection methods, but in the marine environment, the implementation of internal detection technology is relatively difficult, and predicting the corrosion rate of pipeline by algorithm is a better means.

At present, scholars at home and abroad have conducted a lot of research on the prediction of corrosion rate of submarine oil and gas pipelines. Wang Rui [3] established the Frechet distribution model for predicting the corrosion rate of submarine oil and gas pipelines, the model only predicts the corrosion depth and pipeline life of oil and gas pipelines, but not indicate the prediction error. Luo Zhengshan et al [4] established the GM-RBF seabed oil pipeline corrosion rate prediction model. In the model, the GM model is used to predict the corrosion rate, and the predicted value is input into the RBF model to output the prediction residual. The residual is used as the compensation value of the prediction result of the new GM model, and the corrosion rate prediction result is obtained. The prediction error of the model is 6.37%, and the prediction error is relatively large. It can be seen that in the current research results, the average minimum absolute error of the prediction rate of the corrosion rate of the submarine pipeline is 3.79%, and the prediction accuracy needs to be improved.

In this study, the objective factors that may cause pipeline corrosion in the marine environment are fully considered. The data of pipeline corrosion and influencing factors are collected. The influencing factors are used as input to the support vector machine (SVM) algorithm, and the immune algorithm (IA) is used, too. The penalty parameters and optimization of the SVM algorithm are preferred for meeting accurate prediction of the corrosion rate of submarine pipelines.

Analysis of Factors Affecting Corrosion of Submarine Oil and Gas Pipelines:

1. Temperature. When the temperature of seawater is relatively high, the anode-anode reaction speed of the primary battery composed of pipeline and marine environment will be accelerated, and the corrosion rate of pipeline will be accelerated [5].

2. Dissolved oxygen content. The effect of dissolved oxygen on the corrosion rate of pipelines is mainly reflected in two aspects. When the dissolved oxygen content in seawater increases, it will cause cathodic depolarization reaction of the primary battery composed of pipeline and marine environment, and the pipeline corrosion rate will increase [6].

3. Salinity. In general, when the salinity in the marine environment increases, the conductivity of seawater will also increase, which will promote the electrochemical corrosion of the pipeline [7].

4. The pH value. When the pH in the marine environment rises, a layer of calcium deposits easily forms on the surface of the pipe, which causes the pipe to be corroded [8].

5. Oxidation reduction potential. The redox potential will directly affect the redox of seawater, which will affect the redox reaction rate of the primary battery composed of pipeline and marine environment, and finally affect the corrosion rate of the pipeline [9].

6. Seawater flow rate. The increase of seawater flow rate will accelerate the diffusion rate of oxygen atoms, which will accelerate the corrosion rate of pipelines [10]. However, in the deep sea, the submarine pipeline is at the bottom of the ocean, and the seawater is basically at rest, so the influence of this factor on the corrosion rate of the pipeline can be ignored.

In summary, for submarine pipelines, temperature, dissolved oxygen content, salinity, pH value, and redox potential are the main factors causing corrosion of pipelines, and the influence of seawater velocity is negligible.

2. Model Establishment

2.1 Basic Principles of Immune Algorithm (IA)

The immune algorithm (IA) is a new algorithm developed based on the basic functions of the immune system in biology. This algorithm combines the practical application with the basic mechanism of the immune system, resulting in an artificial intelligence program. The IA algorithm has the functions of fast search, self-memory storage, etc., and also has the function of antigen differentiation. Under the strong
interference of the external system, the algorithm can still maintain its own balance well. Therefore, the IA algorithm has strong stability [11]. It is based on the stability of the IA algorithm. In this study, this algorithm is used to select the penalty factor of the SVM algorithm.

2.2 Basic Principles of Support Vector Machine (SVM)
Support Vector Machine (SVM) algorithm is one of the most common intelligent predictive programs. It is widely used in data analysis and regression analysis. It is developed based on the statistical principle of mathematics and adopts the principle of risk minimization. An algorithm with strong generalization ability can adapt to the problem of small sample and high dimension in the process of application[12]. The basic prediction function formula of the SVM algorithm is as follows:

\[ f(x) = \sum_{i \in S_V} (a_i - a_i^*) K(x_i, x_j) + b \]  \hspace{1cm} (1)

In the above formula, \( a_i \) and \( a_i^* \) represent the Lagrangian multiplier; \( b \) represent the offset; \( S_V \) is the data set representing the prediction process; and the \( K(x_i, x_j) \) is the kernel function representing the SVM prediction. \( a_i \) and \( a_i^* \) can be obtained by the following formula:

\[
\begin{align*}
\max \sum_{i=1}^{n} y_i (a_i - a_i^*) - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (a_i - a_i^*)(a_j - a_j^*)K(x_i, x_j) \\
\text{s.t.} \sum_{i=1}^{n} (a_i - a_i^*) = 0, 0 \leq a_i^* \leq C
\end{align*}
\]  \hspace{1cm} (2)

In the above formula, \( C \) is a penalty parameter is indicated, \( \epsilon \) indicating an insensitive loss parameter of the prediction process. At present, there are many methods for the preferred use of \( C \) and \( \epsilon \), the most common methods are particle swarm optimization (PSO), genetic algorithm (GA), least squares algorithm (LS), etc. Each method is only applicable in different fields, and there is no advantage. Inferior points.

2.3 IA-SVM Model Construction
In this study, the IA algorithm is used to optimize the parameters in the SVM algorithm, and the number of kernel functions and hidden layer nodes is selected in a comparatively preferred manner, in order to improve the generalization ability of the SVM algorithm and improve the SVM. The prediction accuracy of the algorithm is as shown in Fig.1.
3. Instance Verification

3.1 Data Sources

The research object is three oil pipelines in a certain sea area of China. These three oil pipelines are all made of X60 steel. The outer diameter of the pipeline is 340.8mm, the wall thickness is 10.6mm, and the pipeline operating pressure fluctuates within 10±0.3MPa. A total of 26 sets of data were obtained for the environment and corrosion rate changes of the three pipelines during the year. The 26 groups of data are divided into two categories, one category contains 20 sets of data, as the model training data set, and the other A category contains 6 sets of data as a model prediction and validation data set. The prediction results are verified using the mean absolute error and the root mean square error.

3.2 Parameter Preference

There are four models of IA, PSO, GA and CV, which are established in the MATLAB programming software. The parameters of the SVM algorithm are optimized by using these four models. The preferred results are shown in Table 1 below. Using the four model parameters $C$ and $\varepsilon$ preferred results, the number of hidden layer nodes in the SVM prediction algorithm is initially set to 20, and superimposed in units of 20. The model is trained using the three most common kernel functions (Sine function, Sigmoidal function, Radial basis function), The average absolute error of the verification sample is used as the preferred criterion for the number of hidden layer nodes and the type of kernel function, and selection of the number of nodes and the type of kernel function. Under the premise of using the optimal sum of the four models IA, PSO, GA, and CV, the results of the preferred number of hidden layer nodes and the type of kernel function are shown in Table 2 below.

Table 1. Parameters C and Preferred result

| Model                          | C      | $\varepsilon$ |
|--------------------------------|--------|---------------|
| Immune algorithm (IA)          | 43.6213| 0.0483        |
| Particle swarm optimization (PSO)| 30.8542| 0.8423        |
| Genetic algorithm (GA)         | 52.934 | 1.3894        |
| Least squares algorithm (LS)   | 28.4523| 9.0723        |
Table 2. Summary of the number of hidden layer nodes and kernel function types

| Model                        | Best hidden layer nodes | Optimal kernel function   |
|------------------------------|-------------------------|----------------------------|
| Immune algorithm (IA)        | 260                     | Sigmoidal Function         |
| Particle swarm optimization (PSO) | 220                  | Sigmoidal Function         |
| Genetic algorithm (GA)       | 240                     | Sigmoidal Function         |
| Least squares algorithm (LS) | 300                     | Sigmoidal Function         |

3.3 Result
The prediction results are shown in Figure 2. The average absolute error and root mean square error of the four models of training actual and predicted results are shown in Table 3 below. As can be seen from the following chart, the prediction results of the LS-SVM model and the GA-SVM model are larger than the actual corrosion rate. At the same time, the data training time of the two combined models is relatively high. Long, indicating that the two combined models are not suitable for the prediction of corrosion rate of submarine pipelines. The prediction results of the IA-SVM model and the PSO-SVM model are less error than the actual corrosion rate, and the maximum error is less than 10%. However, compared with the two combined models, the IA-SVM model is more applicable and the average is absolute. The error is only 1.45%, the root mean square error is only 0.01591645, and the data training time is the shortest, which proves that the IA-SVM model can be used to predict the corrosion rate of submarine oil and gas pipelines.

![Figure 2. Forecast result change chart](image-url)

Table 3. Model training actual, average absolute error and root mean square error table

| Model    | Average absolute error | Root mean square error | Training time (s) |
|----------|------------------------|------------------------|-------------------|
| IA-SVM   | 1.45%                  | 0.01591645             | 5.36              |
| PSO-SVM  | 3.24%                  | 0.04213273             | 6.89              |
| GA-SVM   | 7.89%                  | 0.11425264             | 11.73             |
| LS-SVM   | 9.38%                  | 0.12031279             | 14.52             |

4. Conclusion
(1) It is relatively difficult to detect the environment and corrosion rate of the submarine pipeline, which makes the related data obtained relatively small, which makes the use of intelligent algorithms for prediction in this field relatively difficult, and the SVM algorithm not only has strong generalization ability and small sample adaptability, therefore, the field can use SVM intelligent algorithm for prediction, but in the use of SVM intelligent algorithm, the choice of parameter $C$ and $\varepsilon$ hidden layer nodes and the type of kernel function will have a greater impact on the forecast results.

(2) Simultaneously use the four models of IA, PSO, GA and CV to optimize the parameters $C$ and $\varepsilon$ of SVM algorithm. Under the premise of determining the sum of parameters $C$ and $\varepsilon$, the number of hidden layer nodes and the type of kernel function are based on the minimum average absolute error.
Finally, the SVM algorithm is used for prediction. The prediction results show that the average absolute error and root mean square error of the IA-SVM combination model are relatively small. At the same time, the data training time is relatively short, which proves that the algorithm model is suitable for submarine pipelines corrosion rate prediction.

Acknowledgment
Fund Project: Funded by Shaanxi Province Science and Technology Coordination Innovation Project (221516001)

REFERENCES:
[1] LI Li, LI Jiafeng, WANG Hong, et al. Corrosion Causes of a Subsea Pipeline[J]. Corrosion & Protection, 2017, 38(3): 240-242.
[2] JIANG Xinde, LI Yantao, DU Fanglin. Progress in Study of Corrosion and Protection of Seabed Pipelines[J]. Materials Protection, 2010, 43(4): 65-67.
[3] WANG Rui. Research on Corrosion Failure Prediction of Offshore Oil and Gas Pipelines [D]. Xi'an University of Architecture and Technology, 2017.
[4] LUO Zhengshan, YUAN Hongwei. GM-RBF Model based error compensation for prediction of Submarine Pipeline corrosion[J]. China Safety Science Journal, 2018, 28(03): 96-101.
[5] TIAN Yu, LAN Xu. Influences of Temperature and Velocity on Corrosion Rate of X70 Subsea Pipelines [J]. Total Corrosion Control, 2018(3):62-68.
[6] LI Yanwei. Study on Corrosion Behavior of Seawater Pipelines in Submarine Pipelines [D]. Qingdao University of Science and Technology, 2012.
[7] ZHAO Xiaoyu. The Research of External Corrosion Failure Modes and improvement Method of Single Layer Insulation Pipeline [D]. Dalian University of Technology, 2015.
[8] LUO Wenqiang. Research of Grey System on Oil-Gas Gathering and Transferring Pressure Vessel in Marine Engineering Platform [D]. Lanzhou University of Technology, 2016.
[9] XU Yingbo, HE Sujuan, YAN Huayun, et al. Application of Flow Loop on Flow Rate Study in Oilfield Pipeline[J]. Total Corrosion Control, 2013(11): 66-69.
[10] LIU Bo. Study on Simulative Crevice Corrosion of Common Metal on Submarine Pipelines [D]. Civil Aviation University of China, 2016.
[11] Cutello V, Nicosia G, Pavone M, et al. An Immune Algorithm for Protein Structure Prediction on Lattice Models[J]. IEEE Transactions on Evolutionary Computation, 2007, 11(1):101-117.
[12] Bagheri A, Zandieh M, Mahdavi I, et al. An artificial immune algorithm for the flexible job-shop scheduling problem[J]. Future Generation Computer Systems, 2010, 26(4):533-541.