Residual Strength Prediction of Pipeline with Single Defect Based on SVM Algorithm

Peisheng Liu, Yunping Han* and Yi Tian
The Computer and Communication Engineering College, Liaoning Petrochemical University, Fushun Liaoning, 113001

Corresponding author’s e-mail: *hanyunping@lnpu.edu.cn

Abstract. With the continuous development of society, the increasing number and size of oil-gas pipelines, how to ensure the safe operation of the pipeline has become a major problem. At the same time, under the premise of defects, it is of great significance to predict the residual strength of pipeline for the safety of pipeline operation. At present, scholars at home and abroad have conducted in-depth studies on the evaluation of residual strength of pipelines, but there are mainly two problems in the research, namely, low evaluation accuracy and inconvenient application of evaluation method. Therefore, based on SVM algorithm, a residual strength prediction model for single-defect pipelines is established from four aspects, including model construction, parameter optimization, node number determination of hidden layers and kernel function optimization. The prediction results are highly feasible and advanced compared with various algorithms and standard methods.

1. Introduction
The residual strength of the pipeline mainly refers to whether the conveying pressure of the pipeline exceeds the yield strength of the pipeline material in the process of conveying a certain medium. The evaluation study of the residual strength of the pipeline is mainly to evaluate and study the maximum operating pressure of the pipeline on the premise of understanding the pipeline defects.

As is known to all, there are a large number of defects in pipelines, and the defects must interact with each other. The current research results show that when the distance between the two kinds of defects is more than one defect of radial length, two kinds of defects will not interfere with each other, at this time, pipe can be recognized as a single defect pipe [1,2]. With the continuous improvement of pipeline anticorrosion technology, the number of defects in oil-gas pipelines decreases, and most pipelines with corrosion defects can be considered as single-defect pipelines.

In recent years, with the development of computer science, artificial intelligence technology has been widely used in the oil field, the application of artificial intelligence algorithm not only makes the work efficiency improved, but also security of oil-gas production and transportation play a major role [3]. Therefore, it can be seen that the introduction of artificial intelligence algorithm in the oil-gas field is very critical for the development of the industry. At present, some scholars in the field of pipeline residual strength has been conducted into the research of artificial intelligence algorithm [4], but the process does not use real pipeline blasting data, so the conclusion is up for debate [5]. On the other hand, the actual amount of pipeline blasting data available at present is relatively small, which also restricts the application of intelligent algorithm in this field. To solve this problem, SVM (support vector machine)
algorithm suitable for small samples is adopted in this study, and real pipeline blasting data are used for prediction research, laying a foundation for the application of artificial intelligence algorithm in the field of pipeline residual strength [6-11].

2. Model establishment

2.1. Correlation analysis of influencing factors of residual strength of pipeline

In Matlab software, the correlation analysis model is set up. In this study, considering the pipeline residual strength influencing factors mainly include eight factors, namely steel grade, pipe diameter, pipe wall thickness, defect depth, defect length, yield stress, tensile stress and defect shape, in which all the data are from the present pipeline blasting data [12-14]. The data of influencing factors of residual strength of 119 groups of pipelines and blast pressure are input into the correlation analysis model, and all influencing factors are analysed. The analysis results are shown in Table 1.

Through the analysis results in Table 1, it can be found that the eight factors and the burst pressure of the pipeline have certain relations. Especially, steel grade, pipe wall thickness, yield stress, tensile stress are positively related to the burst pressure, namely, with the increase of pipeline steel grade, pipe wall thickness, yield stress and tensile stress, residual strength of pipeline will also increase. However, there is a negative correlation between pipe diameter, defect depth, defect length with burst pressure, that is, with the increase of pipe diameter, defect depth, defect length, the residual strength of pipe will decrease.

Among all the influencing factors, the influence of defect depth and pipe wall thickness on the residual strength of pipe is relatively serious, while the influence of pipe diameter and defect length on the residual strength of pipe is relatively low. On the other hand, the correlation between steel grade and tensile stress reaches 0.991, and the correlation between steel grade and yield stress reaches 0.982, indicating that the correlation between steel grade and tensile stress and yield stress are very strong. In order to simplify the number of data of influencing factors in the prediction model, the tensile stress and yield stress of the pipe material can be replaced by the pipe steel grade.

|                      | Steel grade | Diameter | Wall thickness | Defect depth | Defect length | Yield stress | Tensile stress | Defect shape | Blast pressure |
|----------------------|-------------|----------|----------------|--------------|---------------|--------------|----------------|--------------|----------------|
| Steel grade          | 1.000       |          |                |              |               |              |                |              |                |
| Diameter             | 0.855       | 1.000    |                |              |               |              |                |              |                |
| Wall thickness       | 0.879       | 0.845    | 1.000          |              |               |              |                |              |                |
| Defect depth         | 0.632       | 0.621    | 0.698          | 1.000        |               |              |                |              |                |
| Defect length        | 0.017       | 0.089    | -0.017         | 0.129        | 1.000         |              |                |              |                |
| Yield stress         | 0.991       | 0.910    | 0.846          | 0.422        | 0.052         | 1.000        |                |              |                |
| Tensile stress       | 0.982       | 0.838    | 0.867          | 0.576        | -0.009        | 0.977        | 1.000          |              |                |
| Defect shape         | -0.220      | -0.124   | -0.289         | -0.436       | 0.101         | -0.263       | -0.303         | 1.000        |                |
| Blast pressure       | 0.318       | -0.423   | 0.701          | -0.880       | -0.333        | 0.389        | 0.397          | -0.485       | 1.000          |
### 2.2. Optimization of SVM algorithm parameters and kernel functions

#### 2.2.1. Dictionary initialization method.

In this study, for the convenience and the scientific nature, the SVM algorithm is obtained in the LIB SVM toolbox in MATLAB software. LIB SVM toolbox provides a good human–computer interaction and parameter modification and many other functions, but also can be drawn to the graphics.

In this study, the main steps of using the LIB SVM toolbox to build SVM algorithm are as follows:

1. Sort out the data of residual strength of the pipeline and relevant influencing factors.
2. Build PSO-SVM model, CS-SVM model, GA-SVM model and CV SVM model in MATLAB, so as to optimize the parameters $C$ and $\sigma$ of SVM by using PSO algorithm, CS algorithm, GA algorithm, CV algorithm.
3. After the complete of $C$ and $\sigma$ parameter optimization, the parameters of SVM algorithm are evaluated by using the optimization results of four data optimization algorithms, and a variety of kernel functions and the number of hidden layer nodes are used to set the SVM model. Then, the residual strength and influencing factor data are used to train the model to predict the unknown data, therefore, the kernel function and the number of hidden layer nodes are optimized and evaluated according to the predicted results.
4. After the kernel function and the number of nodes in the hidden layer are optimized, the date of residual strength and influencing factors of the pipeline are used to train the model, and the unknown data are predicted. According to the predicted results, the application effects of PSO-SVM model, CS-SVM model, GA-SVM model and CV SVM model in the field of residual strength of the pipeline are evaluated.
5. (5) Build the best prediction model of pipeline residual strength and conduct the prediction research of pipeline residual strength.

#### 2.2.2. SVM algorithm parameter optimization.

PSO-SVM model, CS-SVM model, GA-SVM model and CV SVM model in Matlab are constructed respectively. The data of influence factors and residual strength are input into the model. Various types of optimization algorithms are used to optimize the parameters $C$ and $\sigma$ of the SVM model. In the process of parameter optimization, in PSO algorithm, the optimization range of parameter $C$ is set as 0–180, the optimization range of $\sigma$ parameter is set as 0–150, the number of iterations is set as 200, the number of population is set as 20, and the values of $c_1$ and $c_2$ are set as 1.5 and 1.7 respectively. The parameter setting of CS algorithm is the same as that of PSO algorithm, the probability of discovery is set as 0.25, and the step size factor is set as 0.01. In GA algorithm, the number of iteration termination is set as 100, and the number of population is set as 20. The relevant settings of CV algorithm follow the default data of CV algorithm in MATLAB.

Four kinds of optimization algorithm of parameter optimization results are shown in Table 2, it can be seen that the parameters optimization results are relatively large differences. To the result of the parameter optimization and validation, the optimization results are used to set the parameters of the SVM algorithm, and the data of influence factors and the burst pressure through a pipeline data are applied to train the SVM model and then predict residual strength pipeline section, in which the error of the prediction results and conservative are analysed.

**Table 2. Parameter optimization results.**

| Parameter | PSO   | CS    | GA    | CV     |
|-----------|-------|-------|-------|--------|
| $C$       | 66.8759 | 2.3362 | 67.5820 | 3.0168 |
| $\sigma$  | 0.44126 | 3.8361 | 0.60243 | 7.58632 |
2.2.3. SVM algorithm number of hidden layer nodes and kernel function optimization. After the parameters $C$ and $\sigma$ are optimized, the number of nodes in the hidden layer is first set as 30, and then the number of nodes is superposition with 20 as the equidifference. At the same time, four types of kernel functions, namely linear kernel function, polynomial kernel function, radial basis kernel function and sigmoid kernel function, are substituted into the SVM algorithm for trial calculation. In the process of trial, the 129 groups of data of influencing factors and the actual blast pressure are randomly divided into two kinds, in which one kind is contained in 109 groups as a model of training data sets, the other contained 20 groups as a model to predict data sets. The model is trained by using the training data sets, the residual strength of pipeline in the predicted data set is predicted, and the average relative error of the predicted results is calculated. The average relative error was taken as the basic standard for the number of nodes in the hidden layer and the optimization of kernel function, and the number of nodes in the hidden layer and the optimization of kernel function are carried out.

According to the SVM algorithm, the number of hidden layer nodes and kernel function are optimized. For PSO-SVM algorithm, when the number of hidden layer nodes reaches 190, the average relative error of the prediction results basically does not change. At the same time, when the kernel function uses the radial basis function, the average relative error of the prediction results is the smallest. For CS-SVM algorithm, when the number of hidden layer nodes reaches 170, the average relative error of the prediction results basically does not change. At the same time, when the kernel function uses radial basis kernel function, the average relative error of the prediction results is the smallest. For GA-SVM algorithm, when the number of hidden layer nodes reaches 190, the average relative error of the prediction results basically does not change. At the same time, when the kernel function uses radial basis function, the average relative error of the prediction results is the smallest. For CV-SVM algorithm, when the number of hidden layer nodes reaches 150, the average relative error of the prediction results basically does not change. Meanwhile, when the kernel function uses radial basis kernel function, the average relative error of the prediction results is the smallest.

To sum up, after combining different data optimization algorithms with SVM, the optimal kernel function is radial basis function, and the optimal number of hidden layer nodes differs to some extent. The number of hidden layer nodes and kernel function optimization results of parameters $C$ and $\sigma$ are shown in Table 3. The accuracy of the intelligent algorithm in the prediction of residual strength of pipeline cannot be accurately evaluated only by the result of parameter optimization. It is necessary to use the optimal results of these data to set the SVM algorithm, train the model with data of pipeline residual. Only by calculating the prediction accuracy, the application feasibility of SVM algorithm in the field of pipeline residual strength evaluation can be analysed comprehensively.

Table 3. The optimization results of the number of hidden layer nodes and the kernel function.

| Parameter | PSO-SVM | CS-SVM | GA-SVM | CV-SVM |
|-----------|---------|--------|--------|--------|
| $C$       | 66.8759 | 2.3362 | 67.5820| 3.0168 |
| $\sigma$  | 0.44126 | 3.8361 | 0.60243| 7.58632|

| Kernel function | radial basis function |
|------------------|-----------------------|
| Number of hidden layer nodes | 190 | 170 | 190 | 150 |

2.3. Prediction of residual strength of single defective pipeline
PSO-SVM model, CS-SVM model, GA-SVM model and CV SVM model are constructed in MATLAB. Radial basis kernel function is selected for all kernel functions, and the number of hidden layer nodes is set to 190, 170, 190 and 150 respectively. The 109 samples are trained by PSO-SVM model, CS-SVM model, GA-SVM model and CV-SVM model, 20 samples are used to predict residual strength. Compare the predicted results with the actual results, calculate the relative error of each group of predicted results, calculate the average relative error of 20 groups of predicted samples, and make statistics on whether
the predicted results meet the requirements of conservatism. The comparison between the predicted results and the actual results is shown in Figure 1, the relative error is shown in Figure 2 and the average absolute error and conservatism of the predicted results are shown in Table 4.

Through the analysis of Figure 1, Figure 2 and Table 4, it can be found that among the four combination models, the prediction result of PSO-SVM model is close to the actual residual strength of pipeline, the maximum relative error of each group of data is not more than 5%, the average relative error is only 1.336%, and the maximum relative error and average relative error are smaller than the other three combination models. In terms of conservatism, the prediction results of both PSO-SVM model and GA-SVM model does not meet the requirements of conservatism. Although the conservatism is relatively poor, the conservatism of both models is better than the other two models. To sum up, it is feasible to predict and evaluate the residual strength of pipelines using PSO-SVM model, and the prediction and evaluation results are relatively accurate.

![Figure 1. SVM model prediction results.](image1.png)

![Figure 2. Relative error of prediction results and real values.](image2.png)

**Table 4. Relative error and conservativeness of SVM model prediction results.**

| Algorithm  | PSO-SVM | CS-SVM | GA-SVM | CV-SVM |
|------------|---------|--------|--------|--------|
| Mean relative error /% | 1.017   | 3.310  | 3.563  | 4.223  |
| Number of points unmatched conservativeness | 6       | 12     | 6      | 11     |
3. Conclusions
In this paper, through 129 sets of field measured data, including 8 influencing factors and bursting pressure results, the correlation analysis of the influencing factors of the residual strength of the pipeline is carried out. Four models, including PSO-SVM model, CS-SVM model, GA-SVM model and CV-SVM model, are optimized in terms of parameters, number of hidden layer nodes and kernel function. By comparing the predicted results with the actual test results, it is proved that the PSO-SVM model has the best effect in predicting pipeline residual strength when the number of hidden layer nodes is 190 and the kernel function is radial basis kernel function.

References
[1] Khalajestani, Mohsen, Khalaj, Salehi, Alireza, Bahaari, Mohammad, Reza, Shahbazi and Siamak 2015 Appl. Ocean Res. 53 15.
[2] Mahmoodian M and Li C Q 2017 J. Petrol. Sci. Eng. 151 434.
[3] Ramachandra M A, Vishnuvardhan S, Saravanan M, Saravanan M and Gandhic P 2019 Struct. Eng. Mech. 72 31.
[4] Lei N, Jiang J and Yong P 2013 J. Loss Prevent. Proc. 26 1085.
[5] Duisterwinkel E H A, Talnishnikh E, Krijnders D and Wortche H J 2018 IEEE T. Instrum. Meas. 67 655.
[6] Jaufer R, Todkar S S, Ihamouten A, Goyat Y, and X Dé Robert X 2020 18th International Conference on Ground Penetrating Radar 6 14.
[7] Nathalie H, Caradot N and Sonnenberg H ,2020 Struct. Infrastr. Eng. 27 1.
[8] Wang H and Song L 2019 Int. J. Pattern Recog. 16 1782.
[9] Abbas F, Wang L and Yan Y 2020 Measurement 4 19.
[10] Chae Y H, Kim S G, Kim H M, Kim J T and Seong P H 2020 Anna. Nucl. Energy 143 107501.
[11] Ayadi A, Ghorbel O, Bensaleh M S, Obeid A and Abid M 2017 Int. Conf. Soft. 13 1.
[12] Panda A K, Rapur J S and Tiwari R 2018 Measurement 130 44.
[13] Al-Azani K H, Elkatatny S and Abdulraheem A 2018 SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition 14 190.
[14] Al-Azani K, Elkatatny S and Ali A 2019 J. Petro. Explorat. Product. Technol. 9 2769.