Application of Support Vector Regression to predict the Remaining useful life of Polymerized Styrene Butadiene Rubber of cable insulation

Bingxiu Guo, Xiaohui Wang, Yanyan Wang, Haoyun Su and Sijian Chao
School of Reliability and Systems Engineering
Beihang University
Beijing, China
xiaohuiw@buaa.edu.cn

***Abstract*—Rubber is widely used in aviation, aerospace and other important fields. Monitoring properties of rubber and predicting its remaining life is the key to ensuring timely repair and replacement, and it is related to the safety and reliability of equipment. The traditional methods of life calculation is limited by the study of environment and mechanism. The data-driven is more concise and efficient and it can characterize the coupling effect of many factors for the life of rubber. Support Vector Machine (SVM) is a data-driven method for solving small sample and nonlinear problems with good robustness. In this paper the support vector regression(SVR) algorithm was applied to the prediction of rubber life. We used a certain type Polymerized Styrene Butadiene Rubber cable insulation as an example, the temperature and the concentration of oil mist were set as the features to predict the remaining life. The model was trained by accelerated aging test data, and its remaining life was calculated according to its break elongation retention rate at the end of life. Compared with the actual test results and the pridicted results of linear regression model, the applicability of the method was discussed.**

Keywords- support vector regression(SVR); styrene-butadiene rubber; cable insulation; aging; remaining life

I. INTRODUCTION

In the fields of aviation, aerospace, etc., failure of rubber products will lead to seal failure, vibration and noise increase, resulting in failures of up to 60%. Predicting the life of rubber products in advance and monitoring the status of rubber products in time so that they can be repaired and replaced timely is an important factor to ensure flight safety and reliable task execution. Since the beginning of the last century, the aging laws and life prediction of rubber have been studied by many scholars, and many achievements have been accumulated so far. At present, the widely used life prediction method is to carry out laboratory accelerated aging test to study the change of rubber aging performance under different temperature aging conditions. According to the data accumulation of accelerated aging test, researchers have proposed linear relationship method, dynamic curve straight line method, variable line folding method, P-t-T mathematical model, statistical inference method, etc. [1]. The current study of the widely used kinetic formula (1) and the Arrhenius formula (2) are as follows:

$$y = \beta e^{-kt^{\alpha}} \tag{1}$$

$$k = Ae^{-E/RT} \tag{2}$$

The storage and service life calculation of rubber has been studied by many scholars. The life calculation of rubber is mostly accumulated by accelerated aging test in laboratory, and the dynamic formula based on failure mechanism is established. Traditional prediction methods generally require a large amount of calculation and verification, as well as kinetic theory research. Especially when it is necessary to study the effects of a variety of environmental factors on life expectancy, the dynamics formula is more tedious. With the rapid development of artificial intelligence, the research of data-driven methods has matured, and has simple and effective advantages in state monitoring and life prediction, which can solve nonlinear problems well. Some scholars have successively established BP neural network models for predicting rubber aging performance and proved its effectiveness. For example, Li et al. [2] created a Multilayer Feed-forward Neural Networks model with 3 nodes to describe the relationship between the abrasion resistance and other properties. Vijayabaskar et al. [3] established an artificial neural network to simulate the mechanical properties and volume fraction of rubber based on the experimental results.

Support Vector Machine (SVM) is a new machine learning method based on the principle of structural risk minimization and VC dimension theory of statistical learning theory. It was proposed by Vapnik et al. in the 1990s [4]. Support vector machines have good robustness when solving small sample and nonlinear problems. At present, the application of support vector machine has been applied to the field of pattern

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

recognition, probability density function estimation and regression prediction [5].

This study used the styrene-butadiene rubber cable insulation layer as an example, based on its accelerated aging test data, a support vector regression prediction model with main environmental factors and performance parameters as input and lifetime as output was established. The validated model was used to predict the life of styrene-butadiene rubber in the state of use, and the applicability of the method was discussed in comparison with the actual results and linear regression model.

II. METHOD

A. Support Vector Regression (SVR)

SVR is a support vector machine (SVM) application in the field of regression. SVR [6] maps the sample set from the original feature space to the high-dimensional feature space, and then performs regression analysis on the sample set in high-dimensional space.

Solve samples $\{(x_1, y_1), (x_2, y_2).....(x_m, y_m)\}$ to regression hyperplane: $\omega \times x_i + b = 0$ to minimize the cost of experience, where m is the number of samples.

In order to avoid over-fitting, penalty factor C (used to punish errors of samples), relax variable ξ (allowable error range of sample) and loss boundary ε (denoting no loss of points to training set with hyperplane distance less than ε) are added to the formula. The original form of the established SVR regression algorithm is:

$$\min \frac{1}{2}\|\omega\|_2^2 + C\sum_{i=1}^{m}(\xi_i^{\vee} + \xi_i^{\wedge})$$

$$-\varepsilon - \xi_i^{\vee} \le y_i - \omega \times \varphi(x_i) - b \le \varepsilon + \xi_i^{\wedge}$$

$$\xi_i^{\vee} \ge 0, \quad \xi_i^{\wedge} \ge 0 (i = 1, 2....m) \tag{3}$$

By means of Lagrangian transformation and duality, the following results can be obtained:

$$\min \frac{1}{2}\sum_{i=1, j=1}^{m}(\alpha_i^{\wedge} - \alpha_i^{\vee})(\alpha_j^{\wedge} - \alpha_j^{\vee})K(x_i, x_j) - \sum_{i=1}^{m}(\varepsilon - y_i)\alpha_i^{\wedge} + (\varepsilon + y_i)\alpha_i^{\vee}$$

$$s.t. \sum_{i=1}^{m}(\alpha_i^{\wedge} - \alpha_i^{\vee}) = 0$$

$$0 < \alpha_i^{*}, \alpha_i^{*} < C(i = 1, 2.....m) \tag{4}$$

According to the KuhnTucker theorem, the optimal solution of the extreme problem mentioned above is determined, and the corresponding regression function is obtained as follows:

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^{*})K(x, x_i) + b \tag{5}$$

$$b = y_j - \sum_{i=1}^{l}(\alpha_i^{*} - \alpha_i)K(x_i, x_j) + \varepsilon, j \in \{m | 0 < \alpha_m < C\} \tag{6}$$

$K(x_i, x_j)$ is a non-negative symmetric function called a kernel function. Table I shows four commonly used kernel functions. Generally speaking, it is better to select Gaussian Kernel (also known as Radial basis function) for nonlinear models.

TABLE I. KERNEL FUNCTION AND ITS SCOPE OF APPLICATION

Name	Equation	Scope
Linear Kernel	$K(x,z)=x \times z$	Linear SVR
Polynomial Kernel	$K(x,z) = (\gamma x \times z + r)d$	Nonlinear SVR
Gaussian Kernel	$K(x,z) = \exp(-\gamma\|x - z\|^2)$	Nonlinear SVR
Sigmoid Kernel	$tanh(x \times z+r)$	Nonlinear SVR

B. Remaining life prediction method

The SVM-based rubber residual life prediction method mainly uses the experimental data obtained to train the support vector machine model to determine the model parameters (insensitive coefficient, penalty factor, kernel function parameters, etc.). Based on the trained SVR model, the rubber actual lifetime at actual environment is predicted, and finally the remaining life is determined by the damage parameters in actual use.

Based on the above support vector regression model, a life prediction regression model is established to calculate the remaining life. The specific process is as follows:

1) Establish a data set based on the accelerated aging test data: lifetime t, environmental factors F_1 :{Temperature, humidity, irradiance, et }, Manufacturing parameters of rubber F_2 : {Composition, processing} and damage parameters P, and divide it into a training set and a test set. The data set could be pre-processed, which mainly includes eliminating outliers and missing values, and normalizing the feature set.

2) Select the applicable hyper-parameters and optimize. Select Radial basis function as kernel, there are three main hyper-parameters to be determined in support vector regression: penalty factor C, loss boundary ε , and kernel parameter coefficient. Then construct a support vector regression prediction model through training and test data sets, and use MSE and R^2_score as the evaluation indicators of the model.

$$MSE(y, \hat{y}) = \frac{1}{n}\sum_{i=1}^{n-1}(y_i - \hat{y}_i)^2 \tag{7}$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=0}^{n-1}(y_i - \hat{y}_i)^2}{\sum_{i=0}^{n-1}(y_i - \overline{y}_i)^2} \tag{8}$$

3) Input the actual use influencing factors F_1, F_2 and determine the damage parameter P_1 at the time of failure, and obtain the aging life t_1; input the actual use influencing factors F_1, F_2 and the currently measured damage parameter P_2 to obtain the aging life t_2; the remaining life of the rubber $t=t_1$-t_2.

4) The prediction of the model is verified by comparing the life measured with the actual test and the lifetime specified when the rubber leaves the factory.

5) The relationship between damage function P and life t is plotted to verify that the SVR model can explain the aging mechanism of rubber.

III. INSTANCE VERIFICATION

The cable insulation is generally a rubber product, and rubber aging is directly related to the insulation life of the cable. The aging of rubber is affected by various aspects such as environment, physics and biology. The indicators of rubber aging mainly include: breaking strength, elongation at break, permanent deformation, stress relaxation, and tensile strength. For the cable, the tensile index should be mainly investigated. The change of the fracture retention rate is more obvious than the elongation. The fracture elongation retention rate is used as the determination index of aging in this study.

In this paper, the data of styrene-butadiene rubber cable insulation layer after oven accelerated aging test was used for model verification. The test was carried out at different temperatures T and the concentration of oil mist Q. The break elongation retention rate of the test specimen P was used as the index of damage parameters. The referenced data [7] description is shown in Figure 1.

Scikit-learn is a machine learning library based on python, which can facilitate the implementation of machine learning algorithms. The five groups of temperature, the concentration of oil mist and break elongation retention rate are input, and the life is the output. The life prediction SVR model of styrene butadiene rubber(SBR) cable is established based on Scikit-learn.

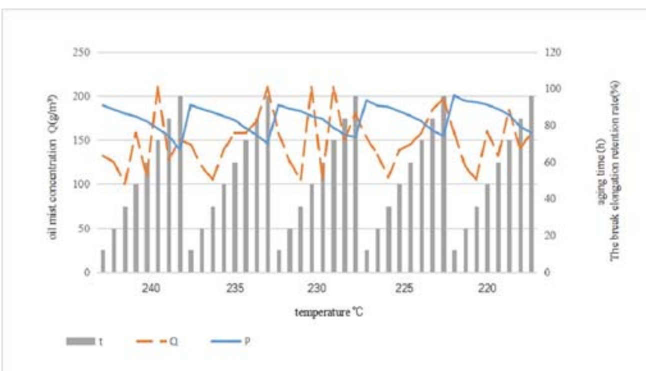

Figure 1. Experiment data description

A. Data preprocessing

Data normalization is the incorporation of differently characterized data into the same scale, which helps to accelerate convergence and improve the accuracy of the model. This paper uses the formula (9) to scale the data to the range [0:1].

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{9}$$

B. Model parameter determination

In the SVR model, the parameters could be determined:

Penalty parameter C (default=1)：C indicates emphasis on outliers, The higher the value of C, the better the training model but the risk of over-fitting will also increase. Therefore, the choice of a suitable C directly affects the quality of the model.

Epsilon (default=0.1)：It specifies the epsilon-tube within which no penalty is associated in the training loss function with points predicted within a distance epsilon from the actual value. It could be optimized by CV.

Cross-validation：A 3-fold cross-validation is used to train data to improve the generalization of the model, which means the training data set was randomly divided into eight parts and each of which taken turns as the validation data set.

Kernel function and its parameters：RBF is chosen as the kernel function of the model. The kernel parameter coefficient (gamma) needs to be optimized by GridsearchCV.

The grid tuning is performed using the function GridsearchCV. Input discrete values in the range C:[0.1,1,10,20,50], gamma:[0.1,1,10], epsilon:[0.001,0.01,0.1] and the parameter combination is selected according to the CV score.

Finally we chosed C=35, gamma=10, epsilon=0.001 as the hyperparameters of the SVR model in this study. It was calculated that MSE=7.02, R^2_score=0.9904 in this model.

C. Model verification

Due to the lack of data, use all data for training. Figure 2 shows the SVR model fitting, linear regression fitting and the aging time of the original data. It was calculated that MSE= 46.08, R^2_score=0.9390 in the linear regreesion model. It can be seen that the SVR model is more accurate than the linear regression fitting, which indicates that it can be used to predict the life of styrene-butadiene rubber cable insulation layer.

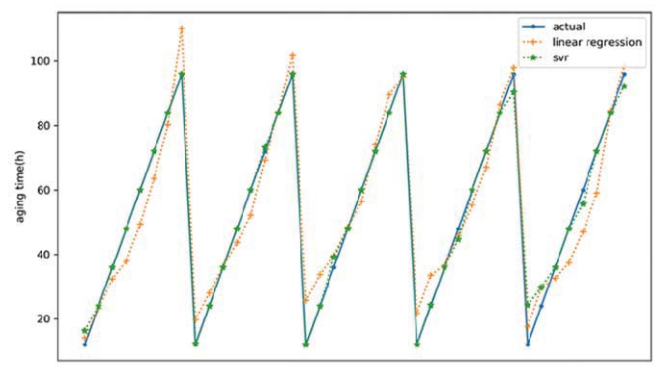

Figure 2. Fitting of SVR and linear regreesion

D. Calculation of remaining life

Referring to GB/T 11026.2 “Heat resistance of electrical insulation materials”, it is generally stipulated that the insulation life terminates when the break elongation retention rate is 30% ~50%. In this paper, the rate of 50% was taken as the index of end-of-life. For instance, we would calculate that the lifetime under the environment conditions of T=200°C and Q=100g/m^3. Enter X:{P=50%, T=200, Q=100} into the model, update the data set and predict to obtain the life of styrene-butadiene rubber. The life of the rubber is 98.4h.

Assume that the elongation at break of a styrene butadiene rubber cable is 87.5% after sampling and tensile test. It is predicted that t=26.62h. The remaining of the rubber cable is known. The life expectancy is 98.4-26.62=71.78h.

E. The relationship between break elongation retention rate P , the concentration of oil mist Q and lifetime t

To verify the interpretability of the model, enter discrete values in the range P:[65:100], Q:[100:210] to get the corresponding value t. This section shows a three-dimensional variation of P-Q-t at five different temperature values. It can be seen from the figure that the oil mist concentration Q and the aging time t are negatively correlated with the elongation at break retention P, and the performance change curve is convex. The variation law at each temperature is similar, and the general law is correct.

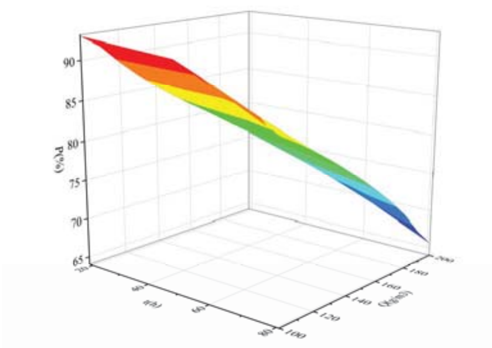

Figure 3. P-Q-t, T=240° C

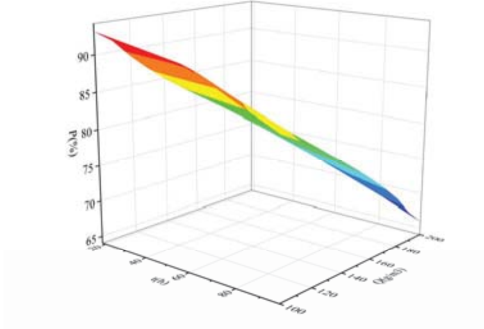

Figure 4. P-Q-t, T=235° C

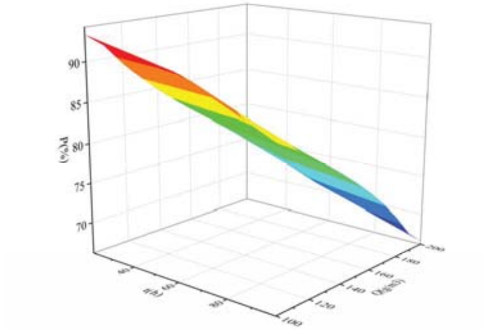

Figure 5. P-Q-t, T=230° C

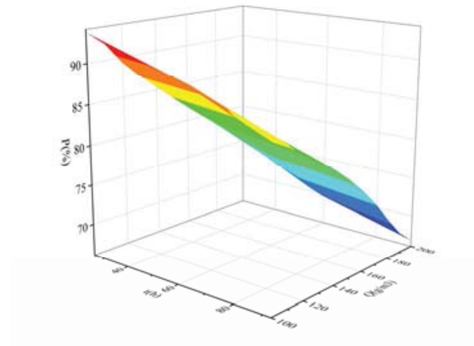

Figure 6. P-Q-t, T=225° C

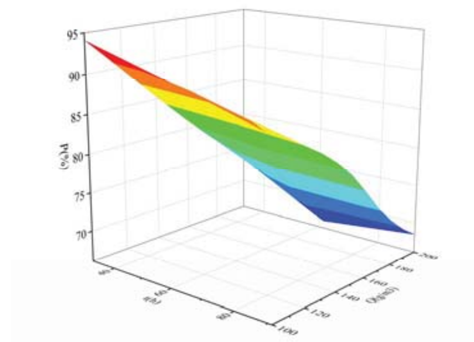

Figure 7. P-Q-t, T=220° C

IV. CONCLUSION

In this paper, SVR was applied to predict the remaining life of rubber. Based on the data of aging test, a life prediction model of styrene-butadiene rubber was established. The model is fitted with the test data and it can correctly reflect the law of rubber aging. The model could be used to calculate the cable remaining life under using environment. The model can be applied to simulate the influence of many factors on the degradation of performance parameters in real environment.

However, when the model is used to predict situations outside the known data range, the effect is not very good. Especially in this case, if the accelerated test data are used to predict the rubber life in the real use environment, the error is relatively large. This shows that the generalization of the model is not very good. There are two main reasons. First, the current training data is not enough. This model is more suitable for life prediction under complex data of real environment. Second, the accuracy of the model in the SVR model is very accurate. For large impacts, grid tuning does not maximize the optimization of hyperparameters.

For these, other optimization algorithms can be used to optimize the hyperparameters of the model, which will continue in subsequent studies.

ACKNOWLEDGMENT

Thanks for the technical guidance provided by School of Reliability and Systems Engineering in Beihang University.

REFERENCES

[1] Y. J. Li "Rubber aging and life estimation (continued) Chapter 10 Prediction of rubber storage period or performance change", Rubber Reference Materials. Xianyang, vol 39, pp. 29-71, April 2009.

[2] H. Li, D. Yang, F. Chen, Y. Zhou and Z. Xiu, "Application of Artificial Neural Networks in Predicting Abrasion, Resistance of Solution Polymerized Styrene-Butadiene Rubber Based Composites," IEEE Workshop on Electronics Computer and Application. Ottawa, pp. 581-584, June 2014.

[3] V. Vijayabaskar, R. Gupta, P. P. Chakrabarti and A. K. Bhowmick, "Prediction of properties of rubber by using artificial neural networks," Inc. J. Appl. Polym. Sci. Hoboken, vol 100, pp. 2227-2237, February 2006.

[4] C. Cortes and V. N. Vapnik, "Support-Vector Networks," Mach. Learn. Dordrecht, vol. 20, pp. 273-297, March 1995.

[5] R. N. Sujay and P. C. Deka, "Support vector machine applications in the field of hydrology: A review," Appl. Soft. Comput. Amsterdam, vol. 19, pp. 72-386, June 2014.

[6] K. W. Lau, Q. H. Wu, "Local prediction of non-linear time series using support vector regression," Pattern. Recogn. Oxford, vol 41, pp. 1539-1547, May 2008.

[7] Y. Wei, Theory and Experimental Study on Residual Life Prediction of Marine Low Voltage Cable. Dalian Maritime University, CA: University Science, 2012.