A Sample-Enhanced Prediction Method for Insulation Life Assessment of Pumped Storage Units

Xinghe Fu [1, *], Kai Hou [2], Ling Yang [2], and Xuechun Hu [1]

[1] School of Electrical Engineering, Southeast University, Nanjing 210096, China;

[2] State Grid Electric Power Research Institute Company Limited, Nanjing 211106, China.

Abstract. Pumped storage units, serving as core equipment for electrical energy storage, peak shaving, and frequency regulation in new power systems, play a critical role in ensuring the reliability of power supply and the efficiency of energy utilization. This study reviews and compares mainstream methods for insulation life prediction, including traditional aging models, intelligent algorithm-based predictions, and dynamic prediction methods based on online monitoring data. In scenarios with limited samples, the Bootstrap method is proposed to augment sample data and enhance the reliability of parameter estimation. The augmented samples are then utilized as the data source for insulation life prediction of polyimide insulation materials, considering varying temperature and individual differences. The research demonstrates that the prediction method, which updates model parameters using posterior distributions, achieves smaller prediction errors and holds significant value for engineering applications.

Keywords: Pumped storage units; remaining life; parameter updating; sample generation.

1. Introduction

As the core equipment of pumped storage power stations, pumped storage units operate under complex and specialized working conditions. Therefore, the insulation systems of the stator windings, transformers, and switchgear in pumped storage units face severe challenges [1]. Insulation materials, as the critical components responsible for isolating current and ensuring the normal operation of electrical equipment, directly determine the operational reliability and service life of the units [2].Traditional aging model prediction methods are suitable for pumped storage units in their early service life and under relatively stable operating conditions, but their prediction accuracy is limited and insufficient for refined maintenance requirements [3]. Intelligent algorithm-based prediction methods are suitable for pumped storage units operating under complex conditions with sufficient monitoring data, enabling high-precision life predictions, but they require addressing issues such as insufficient sample data and poor model interpretability [4-5]. Dynamic prediction methods based on online monitoring data are suitable for pumped storage units with high reliability requirements and the need for real-time maintenance decision-making, offering the highest prediction accuracy and the best adaptability, but they involve higher costs and technical barriers [6-7].

An insulation life prediction method considering varying temperature and individual differences is introduced in the article. Secondly, a sample generation techniques is adopted to expand the sample size. Then, a combined approach of prior and posterior estimation is adopted to update model parameters, thereby improving the accuracy of insulation life predictions. Finally, based on accelerated aging test data for polyimide insulation materials, the insulation life prediction is completed to demonstrate the prediction accuracy of the improvement method.

2. Wiener Model Considering Varying Temperature and Individual Differences

Reference [2] proposed an insulation lifespan prediction method based on a Wiener degradation model that incorporates temperature effects. The model introduces an equivalent temperature to account for the cumulative effects of varying thermal stress and dynamically adjusts both the drift and diffusion coefficients to reflect differences among individual materials.

The Wiener degradation model under thermal stress is given as:

$$X(t) = X(0) + \exp\left(\omega_1 - \frac{\omega_2}{T}\right)t + \exp\left(\omega_3 - \frac{\omega_2}{2T}\right)B(t) \tag{1}$$

where ω1, ω2, and ω3 are constants related to the drift coefficient λ and the diffusion coefficient θ.

The temperature directly influences the drift coefficient λ and diffusion coefficient θ in the Wiener degradation model. Use an equivalent temperature T0 to reflect the effect of varying temperature, and the equivalent temperature T0 corresponding to different temperatures can be expressed as:

$$T_0 = -\omega_2 \Big/ \ln\left[\frac{1}{t_x}\int_0^{t_x} \exp\left(-\frac{\omega_2}{T(t)}\right)dt\right] \tag{2}$$

where tx denotes the current moment, and T(t) represents the varying temperature.

To more accurately characterize individual variability, the drift coefficient λ and diffusion coefficient θ in the Wiener degradation model can be assumed to be random variables, thereby enhancing the model's ability to adapt to variations across different units. When predicting remaining lifespan under constant temperature conditions, λ and θ are fixed values, and their distribution characteristics can be directly assumed. However, under varying temperature conditions, λ and θ vary with the equivalent temperature and cannot be directly treated as random variables. Since λ and θ are related to the parameters ω1, ω2, and ω3, it may treat ω1, ω2, and ω3 as random variables and assume each follows a normal distribution. The means of the three parameters are η1, η2, and η3, and the standard deviations are δ1, δ2, and δ3, respectively. These parameters are assumed to be mutually independent. Their distribution relationships can be expressed as (3). This approach more realistically reflects the differences among different devices and allows dynamic adjustment of model parameters during actual operation, thereby improving the accuracy of remaining lifespan prediction.

$$\omega_1 \sim N\left(\eta_1, \delta_1^2\right), \quad \omega_2 \sim N\left(\eta_2, \delta_2^2\right), \quad \omega_3 \sim N\left(\eta_3, \delta_3^2\right) \tag{3}$$

Put (1) and (3) together to constitute the Wiener degradation model that accounts for the effects of varying temperature and individual differences.

3. Model Parameter Estimation

It is necessary to determine the key parameters with random variable properties in the Wiener degradation model before insulation lifespan prediction. A method combining prior distribution and posterior distribution to estimate the model parameters is adopted. First, initial parameters for the model are determined by utilizing accelerated aging test data. Due to the limited amount of accelerated test data, the Bootstrap method is further applied to expand the data, thereby obtaining the prior distribution of the parameters using the expanded dataset. Subsequently, by incorporating the actual degradation data from individual operational processes, Bayesian estimation is employed to dynamically update the model parameters.

3.1 Prior Distribution

The Bootstrap method can be used to expand the data, when the amount of accelerated test data is limited, thereby obtaining the prior distribution of the model parameters. The calculation of the prior distribution can be divided into two steps: initial model parameter calculation and Bootstrap-based data expansion. The overall process is illustrated in Fig. 1.

1) Calculation of Initial Model Parameters

In accelerated degradation testing, the constant-stress accelerated testing method is primarily employed. Stress is divided into different levels, and degradation tests are conducted on samples at each stress level. Based on the performance degradation data, the initial values of the Wiener degradation model parameters ω are estimated as $\hat{\omega}$. Furthermore, the parameter estimation method provided in Reference [2] is referenced to determine the initial parameters.

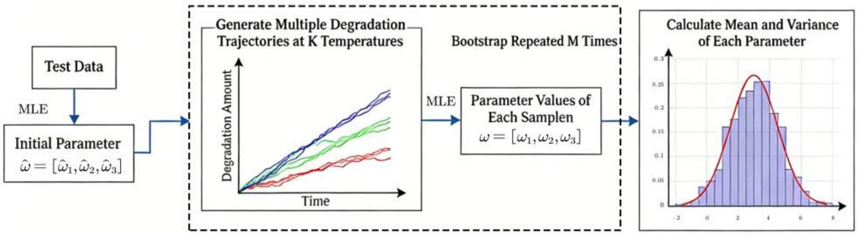


Fig. 1 Flowchart for calculating prior distribution

2) Data Expansion via Bootstrap Method

The Bootstrap method generates multiple different sample datasets through repeated resampling, thereby expanding the data volume and yielding more reliable parameter estimation results. Using the initial values of the model parameters $\hat{\omega}$ as a basis, the Bootstrap method is applied to generate multiple sets of degradation trajectories under different stress conditions. From these generated degradation trajectories, the initial distribution of the model parameters—i.e., the initial parameter η1, η2, η3, δ1, δ2, and δ3 is estimated.

The main steps of the Bootstrap method are as follows:

(a) Based on the initial parameters, the degradation trajectories Xi* at different temperature Ti is simulated.

(b) Each degradation trajectories at different temperature is composed of multiple sample data. Then, step (a) is repeated M times to obtain M sets of Bootstrap samples.

(c) Fit each parameter using a normal distribution to ultimately obtain the prior distribution of the parameters.

Assuming the parameters are mutually independent, the prior distribution of the parameters can be expressed as the joint probability density function in (4).

$$p(\omega) = \frac{1}{\sqrt{2\pi\delta_1^2}} \exp\left(-\frac{(\omega_1 - \eta_1)^2}{2\delta_1^2}\right) \cdot \frac{1}{\sqrt{2\pi\delta_2^2}} \exp\left(-\frac{(\omega_2 - \eta_2)^2}{2\delta_2^2}\right) \cdot \frac{1}{\sqrt{2\pi\delta_3^2}} \exp\left(-\frac{(\omega_3 - \eta_3)^2}{2\delta_3^2}\right) \tag{4}$$

3.2 Posterior Distribution

Considering the influence of individual differences, the practice data can be utilized to update the parameters of the degradation model. Bayesian estimation is applied to dynamically update the prior distribution of the model parameters in real-time, yielding the posterior distribution of the parameters. Due to the complexity of the prior distribution and the likelihood function of the Wiener degradation model, an analytical solution for the posterior distribution is infeasible. To address this issue, the Markov Chain Monte Carlo (MCMC) method can be introduced to numerically sample from the complex posterior distribution, thereby enabling dynamic parameter estimation and updating.

3.3 Overall Workflow

The overall framework for parameter estimation is illustrated in Fig. 2. The prior distribution of the model parameters is obtained by combining Maximum Likelihood Estimation (MLE) and the Bootstrap method with accelerated aging test data from similar products. Bayesian estimation is utilized to update the model parameters based on real-time performance degradation data at different stages, thereby adapting to the individual differences and dynamic variations of the equipment.

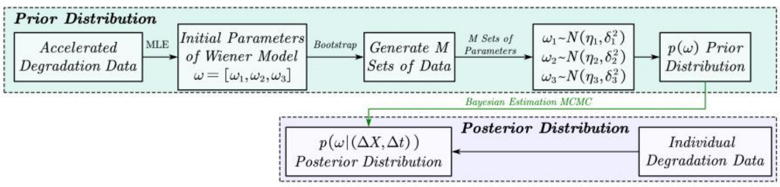


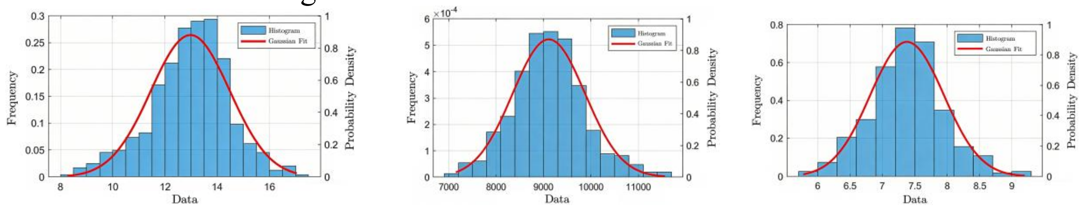
Fig. 2 Flowchart for parameter estimation

4. Case Study

4.1 Data Expansion and Parameter Estimation

Due to the difficulty in obtaining actual degradation data of insulation materials for pumped storage units, this study utilizes accelerated aging test data of 6650 polyimide insulation material, which is commonly used in these units. Based on the accelerated aging test data, the Maximum Likelihood Estimation (MLE) method is employed to estimate the parameters of the Wiener degradation model. According to the characteristics of the Wiener degradation model, the parameter estimation method yields the initial values of the model parameters $\hat{\omega} = [\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3]$. The results are $\hat{\omega}_1 = 19.80$, $\hat{\omega}_2 = 13017$, $\hat{\omega}_3 = 10.30$.

Based on the initial parameter values mentioned above, the Bootstrap method is used to generate 500 degradation trajectories at each of three different temperatures. For each of these 500 datasets, the model parameters are estimated using MLE, and the statistical characteristics of each parameter are recorded. The frequency distribution histograms and the fitted normal distribution curves for each parameter are shown in Fig. 3.



(a) Parameter ω_1 (b) Parameter ω_2 (c) Parameter ω_3

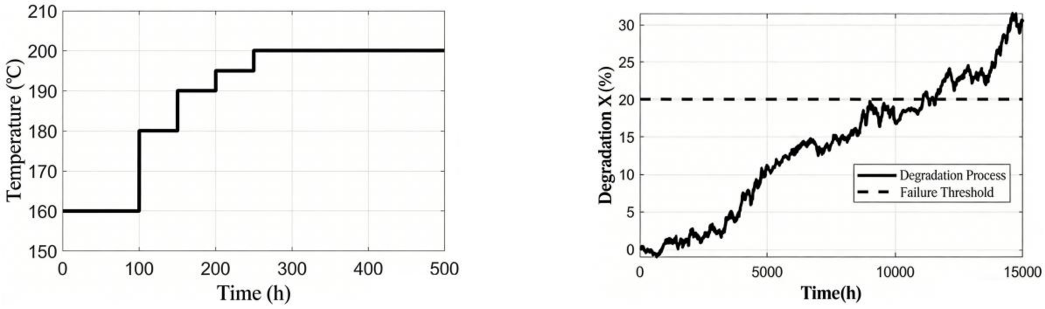
Fig. 3 Statistical characteristics of the three parameters

Finally, the mean and standard deviation of the normal distribution for each parameter is $\omega_1 : [\eta_1 = 13.1, \delta_1 = 2.47]$, $\omega_2 : [\eta_2 = 9117, \delta_2 = 794]$, $\omega_3 : [\eta_3 = 7.38, \delta_3 = 0.88]$, respectively. Since each parameter is mutually independent, the joint prior distribution function of the parameter vector ω can be obtained according to (4).

4.2 Remaining Useful Life Prediction

The prediction accuracy of two models is compared: Model 1, the Wiener model with fixed parameters, and Model 2, the dynamic Wiener model updated via Bayesian inference.

Based on the above mean and standard deviation of each parameter, degradation trajectories under varying temperature are generated through simulation experiments. Considering that temperature variations in actual operation are complex and to accelerate the degradation rate, a simplified scenario with relatively high temperatures is adopted. It is assumed that the operating temperature sequentially changes between $160^{\circ}C$, $180^{\circ}C$, and $200^{\circ}C$. A temperature cycle is set to 500 hours. Within each cycle, the operating times at the three temperatures follow a distribution ratio of 2:5:3. The temperature variation within one cycle is shown in Fig. 4. The degradation trajectory generated by the simulation is shown in Fig. 5, where the failure threshold is marked.



(a) Time-varying temperature (b) Degradation trajectory

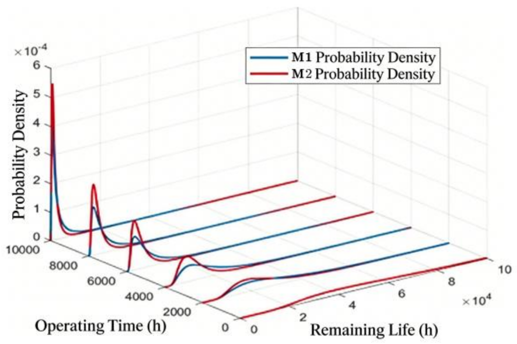
Fig. 4 Time-varying temperature and corresponding degradation trajectory

Based on this degradation trajectory, the remaining insulation life at each time can be calculated. Since the temperature varies over time, the degradation rate also changes accordingly. Therefore, the concept of equivalent temperature is introduced to more accurately characterize the degradation trend at different time points. Table 1 presents the posterior distribution parameters at different times, along with additional equivalent temperature information, to reflect the impact of temperature variations on the degradation process.

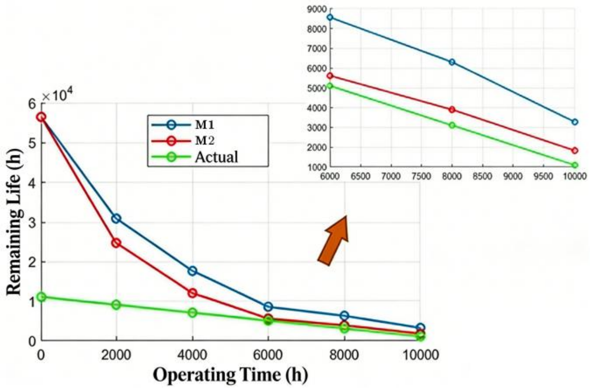
Table 1 Parameters of the model under varying temperatures

operating time（h）	equivalent temperature (℃)	ω_1		ω_2		ω_3	
		mean η_1	standard deviation δ_1	Mean η_2	standard deviation δ_2	mean η_3	standard deviation δ_3
0	160	13.10	2.47	9117	794	7.38	0.88
2000	169.78	13.32	2.45	9116	1.17	7.34	0.48
4000	177.19	13.48	2.35	9116	0.96	7.30	0.43
6000	177.83	13.52	2.19	9116	1.01	7.30	0.41
8000	178.37	13.58	2.06	9116	1.12	7.29	0.32
10000	181.57	13.68	1.88	9116	0.72	7.29	0.26

Under varying temperature conditions, the probability density curves of the remaining useful life corresponding to different operating times from Model 1 and Model 2 are shown in Fig. 5 (a). Fig. 5(a) indicates that as the operating time increases, the probability distribution gradually converges, demonstrating improved stability in the remaining useful life prediction. Fig. 5 (b) shows the comparison curves between the predicted results of each model and the actual values.



(a) Probability density curve of remaining life



(b) Comparison of different models

Fig. 5 Simulation results at time-varying temperatures

Table 2 presents the predicted values of the remaining insulation life from Model 1 and Model 2, along with comparisons with the actual remaining insulation life.

Table 2 Prediction results of remaining useful life under varying temperatures

operating time（h）	Model 1（h）	Model 2（h）	actual remaining life（h）
0	56606	56606	11100
2000	30921	24759	9100
4000	17638	12036	7100
6000	8565	5615	5100
8000	6297	3888	3100
10000	3278	1829	1100

To further quantify the prediction performance of each model under varying temperature conditions, the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are adopted as evaluation metrics. The calculated RMSE and MAPE results for the two models under the two temperature conditions are: $\mathrm{RMSE_{M0}}=21154\mathrm{h}$, $\mathrm{RMSE_{M1}}=19756\mathrm{h}$, $\mathrm{MAPE_{M0}}=61.40\%$ and $\mathrm{MAPE_{M1}}=42.33\%$. The data show that the Model 2 yields a 6.60% reduction in RMSE and a 31.06% decrease in MAPE compared with Model1 under varying temperature. This indicates that Model 2 improves prediction accuracy by introducing real-time parameter updates, and provides a more precise depiction of long-term degradation trends.

5. Conclusion

Based on the Wiener model that considers varying temperature and individual differences, this paper introduces the Bootstrap method to expand the data to the construction of a more precise model for predicting the remaining insulation life. The following conclusions is achieved: using the Bootstrap method for data expansion is an effective approach to augment the sample size. The proposed insulation life estimation method, which incorporates varying temperature and individual differences along with parameter updating, reduces RMSE by nearly 6% and MAPE by nearly 30%, demonstrating its feasibility as a method for remaining useful life prediction.

Acknowledgment

This work is supported by the State Grid Technology Project (No. 5419-202340459A-3-2-ZN).

References

[1] Gong C, Fu X H, Chen W, et al. Optimization of thermal network parameters for generator-motor based on genetic algorithm [C]//6th International Academic Exchange Conference on Science and Technology Innovation, December 6-8, 2024, Guangzhou, China. 2024: 1610-1615.

[2] Feng Z F, Fu X H, Chen W, et al. A remaining useful life prediction model of motor insulation considering individual differences at varying temperature [C]//The 19th Annual Conference of China Electrotechnical Society, September 20-22, 2024, Xi'an, China. 2024: 599-607.

[3] Shi Jinyuan. Study on life prediction of class F insulation for stator coils of large turbo-generators[J]. Journal of Chinese Society of Power Engineering, 2013, 33(7): 507-516.

[4] Ferrisi S, Cappellari P, Guido R, et al. Application of two-parameter Weibull distribution for predictive maintenance: a case study[J]. Procedia Computer Science, 2025, 253: 3160-3168.

[5] Sun J, Zhou F, Hu X, et al. Macro- and micro-spacetime feature-preference gated recurrent unit for remaining useful life prediction of electric motor in multiple working conditions[J]. Signal, Image and Video Processing, 2024, 18(11): 7953-7968.

[6] Li Y, Huang X, Ding P, et al. Wiener-based remaining useful life prediction of rolling bearings using improved Kalman filtering and adaptive modification[J]. Measurement, 2021, 182: 109706.

[7] Wang Y, Liu Q, Lu W, et al. A general varying Wiener process for degradation modeling and RUL estimation under three-source variability[J]. Reliability Engineering & System Safety, 2023, 232: 109041.