



Article

Probability-Based Failure Evaluation for Power Measuring Equipment

Jie Liu ¹, Qiu Tang ^{1,*}, Wei Qiu ^{1,2,*}, Jun Ma ¹ and Junfeng Duan ¹

¹ College of Electrical and Information Engineering, Hunan University, Changsha 410082, China; liu_jie@hnu.edu.cn (J.L.); 19090277@hnu.edu.cn (J.M.); duanjunfeng@hnu.edu.cn (J.D.)

² Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA

* Correspondence: tangqiu@hnu.edu.cn (Q.T.); qwei4@utk.edu (W.Q.); Tel.: +86-0731-88823333 (Q.T.)

Abstract: Accurate reliability and residual life analysis is paramount during the designing of reliability requirements and rotation of power measuring equipment (PME). However, the sample dataset of failure is usually sparse and contains inevitable pollution data, which has an adverse effect on the reliability analysis. To tackle this issue, this paper first applies nonlinear regression to fuse the failure rate and environmental features of PME collected from various locations. Then, a novel binary hierarchical Bayesian probability method is proposed to model the failure trend and identify outliers, in which the outlier identification structure is embedded into hierarchical Bayesian. Integrating binary hierarchical Bayesian and the bagging method, a binary hierarchical Bayesian with bagging (BHBB) framework is further introduced to improve predictive performance in a small sample dataset by resampling. Last, the influence of typical environmental features, failure rate, and reliability are obtained by the BHBB under the real sample dataset from multiple typical locations. Experiments show that our framework has superior performance and interpretability comparing with other typical data-based approaches.



Citation: Liu, J.; Tang, Q.; Qiu, W.; Ma, J.; Duan, J. Probability-Based Failure Evaluation for Power Measuring Equipment. *Energies* **2021**, *14*, 3632. <https://doi.org/10.3390/en14123632>

Academic Editor: José Matas

Received: 24 May 2021

Accepted: 9 June 2021

Published: 18 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: power measuring equipment; binary hierarchical Bayesian with bagging; failure rate; typical environment

1. Introduction

The failure rate of power measuring equipment (PME), namely the probability of failure over a certain period of time, is an important index of reliability research [1]. Accurate evaluation and prediction of failure rate are of critical importance for instrument rotation, aiding in business decisions, and customer service [2]. Meanwhile, more and more intelligent measuring equipment are put into operation with the development of a smart grid [3]. For example, the number of electrical meters worldwide has tripled from 10.3 million in 2012 to 29.9 million in 2017 according to reports in [4]. Nevertheless, the accurate reliability prediction of PME still suffers from some unavoidable factors, especially under extreme environments. For example, the reliability data collected in the outside world usually contain noise and outliers. Different record staff or systems may record data in a biased manner due to uncontrolled sources or data, resulting in decreased accuracy [5]. Therefore, it is necessary to adopt an effective method to analyze the failure data from different perspectives.

In recent decades, numerous approaches to evaluate failure have been developed [6]. Generally, they can be classified into three categories [7]: (1) conventional methods based on system structure, (2) deterministic methods, and (3) probabilistic analysis methods. Among them, fault tree analysis (FTA) is one of the most commonly used conventional methods for failure analysis. For instance, a system-level electric filed exposure evaluation is proposed based on component-level data with a series-parallel fault tree system [8]. To deal with multiple value systems, the dynamic fault tree is developed to compensate

for the shortcomings of FTA [9]. However, FTA is suitable for failure analysis of specific instrument structures, for which it is difficult to analyze the influence of external environmental features.

To avoid issues that arise by different environment features, a number of deterministic approaches are used to aggregate the failure data with multivariate in different geographical areas. Some recent researches aim to solve failure analysis, including partial least squares regression (PLSR) and encoding methods using neural network (NN) [10]. However, the NN method requires sufficient data learning parameters, since it belongs to the data-driven methods. Additionally, the prediction methods based on data mining, such as statistics method and support vector machine (SVM), are used for reliability and uncertainty assessment [11]. They enable the integration of external factors but lack the correlation in external factors and parameter interpretability.

Different from the deterministic methods, the probabilistic analysis method can be employed to treat the information on faults sharing across time and through space. Wiener degradation (WN) and Gaussian degradation (GP) are stochastic processes. They have been widely used for data-driven degradation. However, it needs to be emphasized that they are more suitable for accelerated life analysis. In [12], an accelerated degradation testing based on time-to-failure Weibull distribution is established to predict the reliability of the smart electricity meter. Some other acceleration methods can be found in the literature [13]. However, this reliability model is difficult to predict with actual environmental operational data.

To avoid isolation of features and models, different Bayesian models have been proposed to make evaluations and predicts of failure rate. The unique advantage of Bayesian is that it combines prior information and can provide confidence intervals [14]. Yuan and Kuo [15] developed a Bayesian approach to model the hazard rate with Weibull exponential distribution. Hierarchical Bayesian (HB), a multi-information fusion method, is suitable for the estimation of the individual or multiple variables [16]. Combined with the regression model, HB can also provide accurate fitting and probability prediction ability for the failure rate model [7]. Additionally, some other probabilistic approaches, e.g., the nonparametric Bayesian methods, can also be used for failure and reliability evaluation. Different types of failure data can be fused by the probabilistic analysis method [17]. Unfortunately, the Bayesian approach strongly relies on the accuracy of raw data. That is to say, the accuracy of HB is easily affected by the outliers and shortage of sample size [18]. Otherwise, a wide confidence interval and volatility forecast results will be provided. To resolve these difficulties and limitations mentioned above, a method for identifying outliers and being able to fuse multiple data is required.

The motivation of this paper is to enhance the failure evaluation performance of PME. The contributions of the proposed probability-based method are:

1. To predict failure rate in different locations, a HB with nonlinear regression is proposed to integrate the environmental features and failure rate information of PME. The degree of influence of various environmental features can be then interpreted.
2. To reduce the effect of outliers, a binary hierarchical Bayesian (BHB) is proposed to distinguish the outliers from the raw failure rate data. Particularly, the outlier identification term is embedded in a binary structure to simplify the data preprocessing process.
3. Combined with the bagging method, a binary hierarchical Bayesian model with bagging (BHBB) is further proposed to reduce the failure prediction variance of the PME. The random resampling of BHBB avoids the weight assignment of samples without prior information compared with Bayesian bootstrap.
4. Finally, the effectiveness of the proposed method is tested and verified with actual sample data of electricity meters from three typical locations. In addition, the relationship between failure rate and typical environmental features is then quantitatively analyzed.

The remainder of this paper is organized as follows: We begin by introducing the environmental characteristics and actual failure rate data in Section 2. Then, the proposed BHBB framework for outlier identification is developed in Section 3. Section 4 further explores the way to solve the proposed BHBB framework. In Section 5, different experiments are conducted to verify the effectiveness of the method. Finally, the conclusions are presented in Section 6.

2. The Environmental Features of PME

2.1. Environmental Features Analysis

To ensure accurate metering of PME, electric energy metering equipment is used as a data instance. It is reported that the State Grid Corporation of China has to check the status of meters every few years to ensure accurate and fair measurement [19]. Moreover, in the actual environment, the normal operation of the electrical meter is affected by various factors. In addition, it is difficult to eliminate the fault of the electrical meter one by one because of its wide distribution. Hence, the failure rate of the electrical meters can be analyzed to offer suggestions for providing trade-off decisions.

Obviously, the PME are more prone to failure under typical environmental stress. Therefore, the actual data collected from Xizang (XZ), Heilongjiang (HLJ), and Xinjiang (XJ) of China are used, which have extreme altitude, cold, and dryness, respectively. The relationship between the failure rate and typical environmental features is expected to be analyzed.

Table 1 summarizes the primary environmental features of the three locations. All environmental features are calculated using the annual average [20]. It can be seen from Table 1 that four primary features are considered, including the temperature, pressure, humidity, and illumination. Meanwhile, the average annual maximum temperature (M-Tem) and the average annual minimum temperature (N-Tem) are used to represent temperature characteristics. It shows that the most dramatic temperature change is in XJ. This means that the temperature difference in the area is relatively large. Furthermore, the humidity features are particularly high in XZ during the entire year. Conversely, the pressure is the smallest in XZ. The N-Tem of XJ is the lowest when the others have no obvious characteristics compared with XZ and HLJ. It is worth mentioning that multiple typical environmental features may also be superimposed on the PME at the same time. For example, there will be very high temperatures during the day in XJ.

Table 1. Environmental features of true typical areas.

Areas	M-Tem(°C)	N-Tem(°C)	Pre. (hPa)	Hum. (%RH)	Ill. (Lux)
XZ	29	−10	605.2	61.9	103.8
XJ	47	−22	1003.5	28.3	327.64
HLJ	37	−31	925.2	55.7	253.6

2.2. Actual Failure Rate Sample of Electrical Meters

In this section, three real failure rate datasets of electrical meters from XZ, XJ, and HLJ provinces are used to illustrate the proposed method. The structure of the data acquisition system for electricity meters is shown in Figure 1. It shows that the operating state data of the electrical meters are transmitted to the concentrator through the power line. The operating state data mainly represent the electrical energy data, including current, voltage, power, and other measurement data when the meter is working. Then, the data of the concentrator is further transmitted by the base station through wireless transmission or power line. Finally, the data for number of failed meters of all locations is transferred to the server database. Moreover, if the meter fails to send the data for a certain time, its data loss state can be easily obtained and the meter may need to be replaced. As shown in Figure 1, the measuring unit mainly includes the three-phase electric energy meter and the

single-phase electric energy meter. Meanwhile, the failure rate of the data includes LCD faults, out of tolerance, battery undervoltage, clock faults, and communication faults.

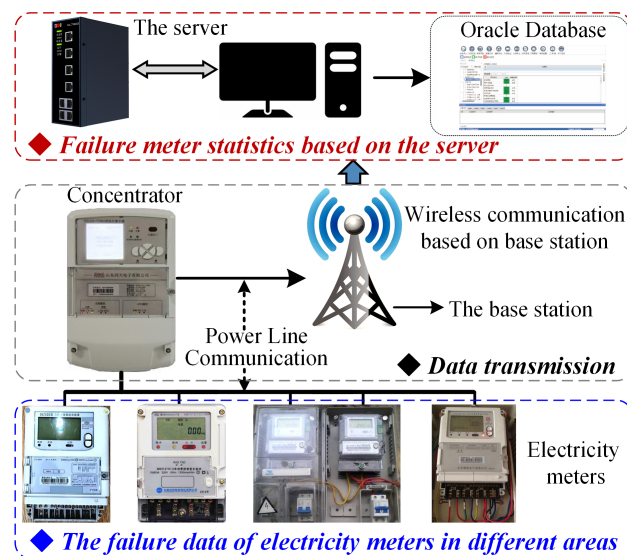


Figure 1. Operation status monitoring platform of the electrical meters.

Figure 2 shows the failure rate data in three areas. All the data are collected from the same company to avoid the difference of objects between multiple areas. The datasets record 7 years of failure numbers from 2012 to 2018. Here, A1 to A7 is the abbreviation of the dataset, where each dataset contains 7 groups of subsets that collect from different areas in the same province, namely a total of 21 groups are obtained. The failure rate is calculated through failures recorded every year divided by the remaining electrical meters.

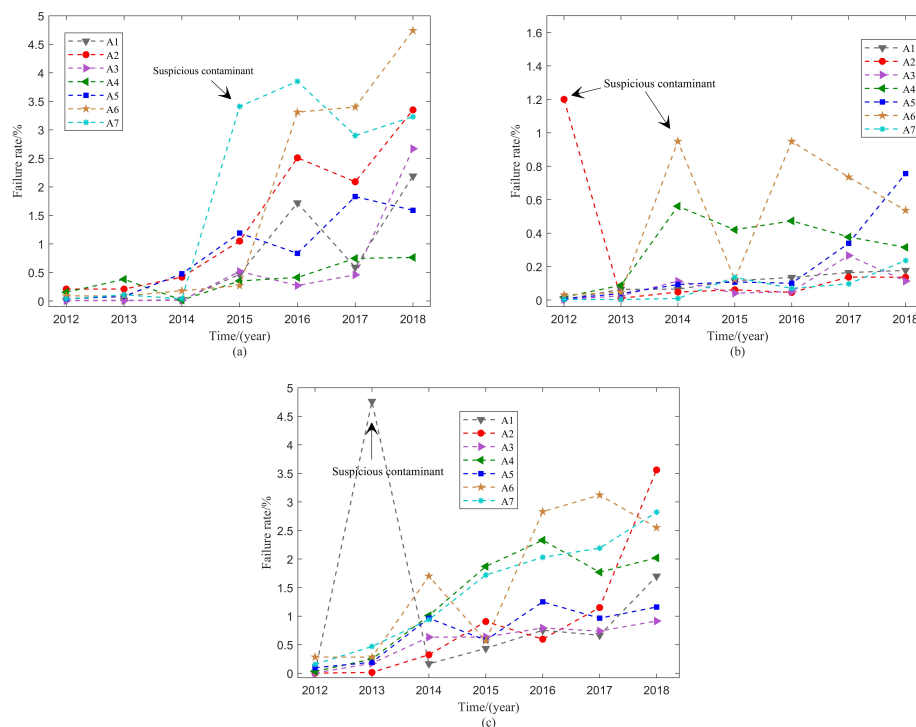


Figure 2. The raw failure rate data of electrical meters from (a) XZ, (b) XJ, and (c) HLJ.

It can be seen from Figure 2 that the failure rate keeps growing during 2012–2016 and falls slightly in 2017. Obviously, the failure tendency has a strong correlation with time. The growth rates are also different from each other because the data are collected in different areas. Furthermore, several failure rate data points are particularly large in 2012–2016 of Figure 2b and 2013 in Figure 2c. Those values are completely different from other data. Therefore, it is reasonable to treat these points that are far from intensive data as suspected outliers.

3. Binary Hierarchical Bayesian with Bagging

3.1. Motivation

The hierarchical Bayesian framework allows users to build the model with multilayered probability distributions. It is more suitable for the analysis of small samples [16].

Let $Y = \{X_{a,t,i}, y_{a,t,i}\}$ be the sample dataset where $y_{a,t,i}$ refers the i th failure rate data of PME in the area a at time t , $X_{a,t,i}$ refers the corresponding i th environmental features in area a at time t , and each $X_{a,t,i}$ is composed of different environmental features where $a = 1, 2, 3$ denotes three areas.

Generally, the failure rate can be specified to obey Weibull distribution. It is commonly used to describe electronic equipment failure [21], and the probability density function of Weibull distribution is

$$P_w(k | \alpha, \beta) = \beta \alpha (k)^{\alpha-1} \exp(-\beta k^\alpha) \quad (1)$$

where α is the shape parameter, and β represents the scale parameter of Weibull distribution. $k = y_{a,t,i}$ denotes the failure rate, and α , β , and k are greater than 0. The change of the Weibull shape parameter usually indicates the failure mechanism changes. Thus, the Weibull shape parameter α remains constant while scale parameter β changes.

To integrate environmental features with temporal information, we establish a nonlinear regression based on the scale parameter β , which can be obtained as

$$\log(\beta) = \beta_0 + \beta_{\alpha,1}t + \cdots + \beta_{\alpha,q}t^q + \beta_{\alpha,p}X_{a,t,i} + \lambda \|\beta_{\alpha,p}\|^2 \quad (2)$$

where β_0 denotes the intercept when $X_{a,t,i}$ and t equal to zero, $\beta_{\alpha,p}$ represent the p th slope in the linear equation, $q = 1, 2, \dots, n$ denotes the order of time t and $p = q + 1, \dots, q + 4$; $\lambda \|\beta_{\alpha,p}\|^2$ denotes regularization items to reduce overfitting. The log-scale $\log(\cdot)$ is used to ensure the positive interval of parameter β . The order of polynomial depends on the balance between accuracy and the simplicity of the model.

The nonlinear regression Equation (2) combines all the effects and creates an underlying structure of the model. However, it cannot identify potential outliers, thus additional outlier identification structure is in great demand.

3.2. Proposed Binary Hierarchical Bayesian

Outliers can easily cause large deviations in small sample predictions. The traditional outlier detection methods, e.g., cluster-based and density-based, will fail in a small sample dataset [22]. However, deleting data points directly can easily lead to information loss.

In this paper, a binary hierarchical Bayesian is proposed to identify outliers. The binary method assumes each failure sample data was generated by a truly process being studied or by another outlier structure and, thus, a mixture likelihood is built to specify normal values and outliers. Formally, it extends the original model to include binary likelihood which consists of the Weibull and outlier assignments.

For a particular experiment, every element in Y believed to have been generated by two models, i.e., a truly Weibull parameter model p_w parameterized by θ_w , and another outlier model p_c parameterized by θ_c . Obviously, the binary parameter $M = \{m_i\}$ can be defined as follows: $m_i = 0$ if i th data point was generated by the Weibull parameter model

and $m_i = 1$ if i th data point was generated by outlier model p_c . Furthermore, the outlier model in binary hierarchical Bayesian model can be described as

$$y_{a,t,i} \sim \begin{cases} P_w(\cdot | \theta_w), & \text{if } m_i = 0 \\ p_c(\cdot | \theta_c), & \text{if } m_i = 1 \end{cases} \quad (3)$$

where m_i can be expressed as discrete distribution

$$m_i \sim \text{Bernoulli}(\phi_i) \quad (4)$$

where ϕ_i is the probability of producing outliers for the stochastic process. Then, the likelihood of binary hierarchical Bayesian model is given as

$$L(Y, M | \theta_w, \theta_c, \phi_i) = \prod_{i=1}^N \left\{ \phi_i^{m_i} (1 - \phi_i)^{1-m_i} x[(1 - m_i)P_w(y_{a,i,t} | \theta_w) + m_i p_c(y_{a,i,t} | \theta_c)] \right\} \quad (5)$$

In practice, if the particular outlier is known, which means a known binary parameter M is obtained, the model can be inference via Markov chain Monte Carlo (MCMC) algorithm. However, sample collection staff rarely mark whether or not or which kind of data belongs in actual work, and some mistakes can then be made during the process.

Binary likelihood models data with a weight coefficient for different likelihoods. In our case of the model, the binary derives from marginalizing over the m_i , and the likelihood can then be expressed as

$$L(Y | \theta_w, \theta_c, \phi_i) = \prod_{i=1}^N \{ (1 - \phi_i) p_w(y_{a,t,i} | \theta_w) + \phi_i p_c(y_{a,t,i} | \theta_c) \} \quad (6)$$

After establishing binary hierarchical Bayesian, the inference can be performed using Bayesian updating or other schemes under this outlier likelihood. Alternatively, one can adopt no information prior to distribution Uniform(0, 1) for ϕ_i . In the proposed method, we assign probit function $\Phi(\cdot)$ to ϕ_i so that the binary structure can better fit the data due to the hierarchical structure. The probability density function of $\Phi(\cdot)$ is given by

$$\phi_i \sim \Phi(\rho_i) = \frac{1}{2} + \frac{1}{2} \text{erf}(\rho_i / \sqrt{2}) \quad (7)$$

where ρ_i is the parameter of probit function $\Phi(\cdot)$. When $\phi_i = 1$, it represents the probability p_c is generated by the outlier model. Thus, the prior distribution of p_c , namely θ_c , can be specified as Uniform(0, 1).

3.3. The Proposed BHBB Model

Failure mechanism performance is sensitive to small changes and can be easily disturbed in a small sample dataset. Compared with other integrated learning methods, bagging is more simple and efficient. Followed by the binary HB, the bagging is used to reduce the prediction variance of failure rate for PME.

In this stage, we propose the binary hierarchical Bayesian with bagging, and binary HB plays the role of the basis function of bagging. The main feature of bagging is random resampling and it can be divided into two steps: (1). The bootstrap resampling of the raw failure data is used to reduce the variance. (2). The BHBB average output of all basis functions.

For the particular subset data $\mathbf{Y}_m \in \mathbf{Y}$, where $m \in 1, \dots, n$, denotes the vector of predictions by $G(\mathbf{Y}_m)$, where $G(\mathbf{Y}_m)$ refers to the binary HB basis function of BHBB. A bootstrap sample of the data is a sample with replacement, so that $\mathbf{Y}_m = \{(X_{a,t,i})_m, (y_{a,t,i})_m\}$

with repetitions allowed. For each bootstrap sample, the model produces predictions $G(\mathbf{Y}_m) = G(\mathbf{Y}_m)_1, \dots, G(\mathbf{Y}_m)_k$, where k is the number of predictions sites. K total bootstrap samples are used, then the predicted output value of the failure rate of PME can be expressed as

$$y_{a,t,j}^* = \frac{1}{K} \sum_{m=1}^K G(\mathbf{Y}_m)_k \quad (8)$$

Thereafter, the output of the BHBB is the mean of multiple binary hierarchical Bayesian sub-models. A version of pseudocode for implementing BHBB is

1. For $m \in \{1, \dots, n\}$.
2. Draw a bootstrap sample \mathbf{Y}_m from failure rate data \mathbf{Y} randomly.
3. Set the binary hierarchical Bayesian as sub-model $G(\mathbf{Y}_m)$, then find failure rate predicted values $G(\mathbf{Y}_m)$.
4. The BHBB predictor is $(\sum^m G(\mathbf{Y}_m)/K)$. The value of K can be obtained by the tradeoff between the experimental results and the amount of calculation.

Based on the failure rate sample in Section 2, to select an appropriate parameter K , the relationship between the root mean-squared-error (RMSE) and the number of basis functions K is shown in Figure 3. It provides an example of RMSE based on the XZ. It depicts that the prediction RMSE has a huge improvement when K is less than 30. In addition, the RMSE value tends to stabilize after nearly 30 replications. It means that the results would be hard to improve with more bootstrap replicates. Hence, we set $K = 30$ in this paper. It takes about one and a half hours to complete the training of 30 sub-models in a single line for 5000 samplings.

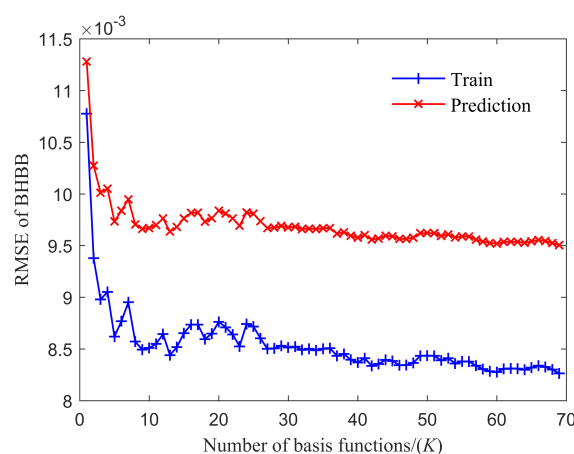


Figure 3. RMSE curve of BHBB for training and prediction in XZ.

4. Failure Evaluation of PME on BHBB

4.1. Prior Specification

After establishing the BHBB model, assigning of the prior distribution of model parameters is required. The prior distribution of model parameters is required to be assigned since it is related to the convergence of the model. Particularly, the convergence of MCMC needs to be considered when different prior distributions are specified. The selection methods of prior distribution include (1) expert prior, (2) weak information prior, and (3) the value range of the parameter [23,24]. The weak information prior, such as normal distribution, is generally recommended when an expert prior cannot be obtained. Therefore, most of the parameters can be determined by using the weak information prior if the ranges are not limited.

The value β_0 and $\beta_{a,p}$ of the Weibull regression model are not limited. To specify appropriate priors for the hierarchical model, normal distribution is selected for β_0 and $\beta_{a,p}$ as the prior probability density function

$$f(\beta_{a,p} | \mu_{a,p}, \sigma_{a,p}) = \sqrt{\frac{1}{2\pi\sigma_{a,p}^2}} \times \exp\left\{-\frac{1}{2\sigma_{a,p}^2}(x - \mu_{a,p})^2\right\} \quad (9)$$

where $\mu_{a,p}$ and $\sigma_{a,p}$ are the mean and standard deviation, respectively. λ is the regularization parameter of $\lambda\|\beta_{a,p}\|$ in $\log(\beta)$ and it should be greater than θ . Thus, it can be specified as no information distribution

$$\lambda \sim \text{Uniform}(0, 3) \quad (10)$$

In the case where mean $\mu_{a,p}$ and standard deviation $\sigma_{a,p}$ parameters are independently sampled, the distributions of the parameters are different. However, they also have an intrinsic relationship as a result of sampling from the same prior distribution $\beta_{a,p}$. It means that the inferences affect the prediction of failure rate in other areas. Generally, the hierarchical model is more powerful than a single-layer structure because they represent knowledge at different layers of abstraction. Therefore, in the second layer of BHBB, the parameters $\mu_{a,p}$ and $\sigma_{a,p}$ are estimated by normal prior and half-Cauchy prior in Equation (9) respectively, which is helpful to simplicity and computational stability

$$\begin{aligned} M_{a,p} &\sim \text{Normal}(0, \sigma_\mu^2) \\ \sigma_{a,p}, \alpha &\sim \text{HalfCauchy}(10) \end{aligned} \quad (11)$$

where σ_μ is the variance. A larger value can be set to σ_μ in the beginning and then adjust with the convergence of the model. Parameter ρ in (7) is the normal distribution, the mean and standard deviation of ρ have the same distribution as parameter $\mu_{a,p}$ in Equation (11).

4.2. Model Parameter Estimation and Prediction

To obtain the prediction results, all the observed data of PME are used to update the prior distributions and yield posterior probabilities using Bayes' theorem [25]. As for the joint posterior probability, it has the same formula of every repetition in bagging, which can be written as

$$\begin{aligned} &p(\beta_0, \beta_{a,p}, \alpha, \lambda, m_i, \theta_c | Y_m) \\ &= \frac{p(\beta_0, \beta_{a,p}, \alpha, \lambda, m_i, \theta_c, Y_m)}{p(Y_m)} \\ &\propto L(Y_m | \theta_w, \theta_c, m_i) p(\beta_{a,p} | \mu_{a,p}, \sigma_{a,p}) \\ &\quad \times p(m_i | \phi_i) p(\beta_0) p(\alpha) p(\theta_c) p(\lambda) \end{aligned} \quad (12)$$

where $L(Y_m | \theta_w, \theta_c, m_i)$ is the likelihood which has the same form as Equation (6), θ_w consists of $\beta_0, \beta_{a,p}, \alpha$ and λ , $p(Y_m)$ is the marginal density function. The posterior distribution formulations can be obtained by integrating out the rest parameters based on the marginal posterior density. To estimate parameter m_i , the marginal posterior distribution $p(m_i | Y_m)$ of m_i is evaluated as

$$\begin{aligned} &p(m_i | Y_m) \\ &= \iiint p(\beta_0, \beta_{a,p}, \alpha, \lambda, m_i, \theta_c | Y_m) d\theta_w d\theta_c \end{aligned} \quad (13)$$

The posterior distribution of BHBB parameters are calculated each repetition in bagging, K total times is used to resample from Y . The predicted mean value of failure rate can be expressed as

$$E(m_i | Y) = \frac{1}{K} \sum_{m=1}^K \left(\int p(m_i | Y_m) dm_i \right)_m \quad (14)$$

Furthermore, the known information and posterior distribution of parameters are obtained to predict the failure rate effectively in another area. If some new data $z_{x,t}$ is

observed, which contain the environmental features in the new area, the predicted failure rate can be inferred as

$$p(z_{x,t} | Y) = \frac{1}{K} \sum_{m=1}^K \left(\int f(z_{x,t} | \theta) p(\theta | Y_m) d\theta \right)_m \quad (15)$$

where $p(\theta | Y_m)$ denotes the joint posterior probability in Equation (12), θ represents parameters θ_w , m_i , and θ_c . $f(z_{x,t} | \theta)$ is the expected value under the condition of known parameters θ .

As a necessary indicator of failure rate evaluation, it is necessary to consider the reliability of PME. Generally, this prediction can be adopted to take reserve measures or set strategies and provide a guarantee for system continuous operation.

The reliability of the system indicates the ability to complete the assignment at the given time. For the BHBB, the reliability function can be inferred as

$$R(t | Y) = \frac{1}{K} \sum_{m=1}^K \left(\exp \left\{ - \int_0^t P_w(k = y_{a,t,i}) d\theta \right\} \right)_m \quad (16)$$

where the range of parameters $R(t | Y)$ is $[0, 1]$, the $P_w(k = y_{a,t,i})$ is the function of Weibull posterior probability density. Particularly, the corresponding posterior sample can be used to calculate the confidence intervals (CI) of reliability.

To verify the effectiveness of the outlier model and accuracy of the BHBB, model averaging is used to compare the relative accuracy between the BHBB and the common HB model. It should be noted that the HB does not have an outlier identification structure and bagging. Information criterion can then be computed to compare models under exactly the same observations [26].

We use the widely applicable information criterion (WAIC) to evaluate the models, which is a fully Bayesian criterion for estimating out-of-sample expectations [27]. The WAIC can be defined as the following form

$$WAIC = -2 \left(\sum_{j=1}^N \log p^*(y_{a,t,i}) - \sum_{j=1}^N V(\log P(y_{a,t,i} | \theta)) \right) \quad (17)$$

where $p^*(y_{a,t,i})$ is the Bayes predictive distribution, $V(\log P(y_{a,t,i} | \theta))$ represents the variance of log-likelihood of $y_{a,t,i}$. The first item in Equation (17) is Bayes training loss, and the second piece of WAIC is the effective number of parameters.

4.3. Proposed Failure Rate Prediction Framework

Aiming at solving the impact of environmental features and outlier in the failure rate data, a general BHBB framework is constructed and depicted in Figure 4. The framework can be divided into four steps.

1. Data preprocessing: The environmental features and failure rates are normalized based on mean and standard deviation. The failure rate data is resampled for K times for the bagging algorithm and the Y_m is obtained.
2. Model establishment: The nonlinear regression model is built to integrate time and environmental features. For a particular experiment, every element in Y_m is believed to have been generated by two models, i.e., a truly Weibull parameter model p_w parameterized by θ_w , and another outlier model p_c parameterized by θ_c . In this model, the binary parameter $M = m_i$ is defined as follows: $m_i = 0$ if i th data point was generated by Weibull parameter model and $m_i = 1$ if i th data point was generated by outlier model p_c . Then the BHBB is established to predict the failure rate of PME based on bagging and outlier identification structure.
3. Model estimation: The MCMC is used to generate samples from the posterior distribution. It takes a number of iterations for samples generated from the joint posterior

distribution. Considering the convergence of the model, automatic differentiation variational inference is used to initialize all the parameter values. The results of the BHBB will average all K times posterior parameters.

4. Model verification: The convergence analysis and WAIC are used to verify the validity of the model. If the convergence and WAIC do not meet the conditions, the model will use the trained model parameters to initialize or update priors to converge faster. Finally, the posterior distribution parameters are used to predict the failure rate and estimate the reliability.

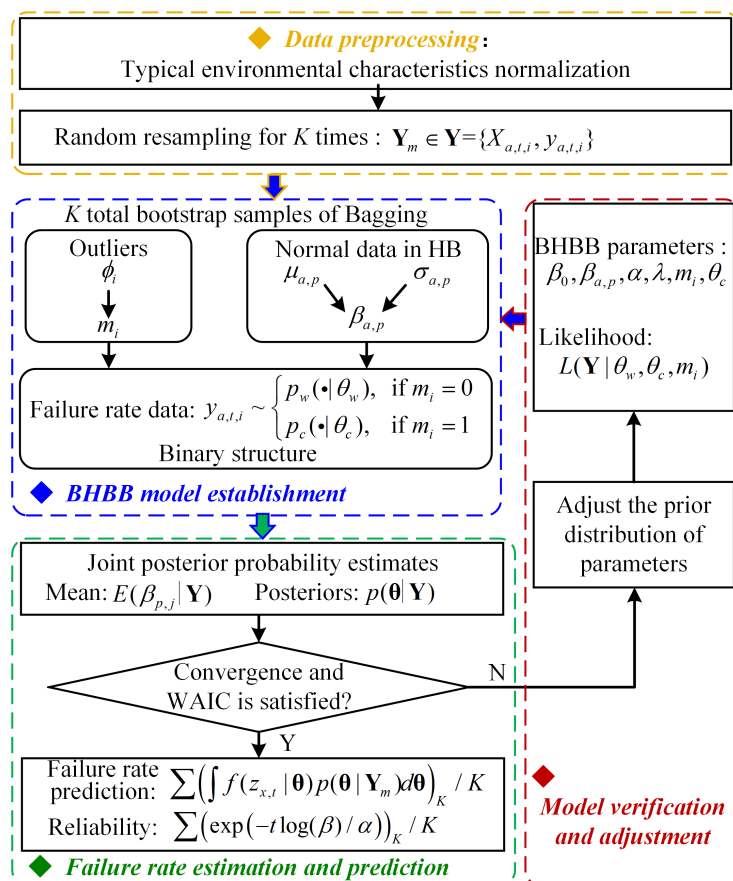


Figure 4. The BHBB framework for failure evaluation of PME.

5. Illustrative Example

To verify the effectiveness of the proposed BHBB, the Pymc3 is used, which is a kind of Python library. In the process of MCMC sampling, the number of samples is set to 5000 to fulfill all the samples and 2000 samples for burn-in. Considering the convergence of the model, automatic differentiation variational inference is used to initialize all the parameter values. Then No-U-Turn sampler and Binary-Gibbs-Metropolis are used to generate samples for continual distribution and discrete distribution, respectively [28]. To verify the effectiveness of the algorithm, leave-one-out cross-validation is used to separate data where 6 groups of data are used to training and 1 group for prediction per province.

5.1. Calculation: Model Comparison Analysis

In Equation (2), to select a suitable q value, the experiment with different parameter q is conducted to show its impact, as listed in Table 2. Here, the widely applicable information criterion (WAIC) and leave-one-out (LOO) cross-validation are used to evaluate the model, where a higher value means a better fitting effect.

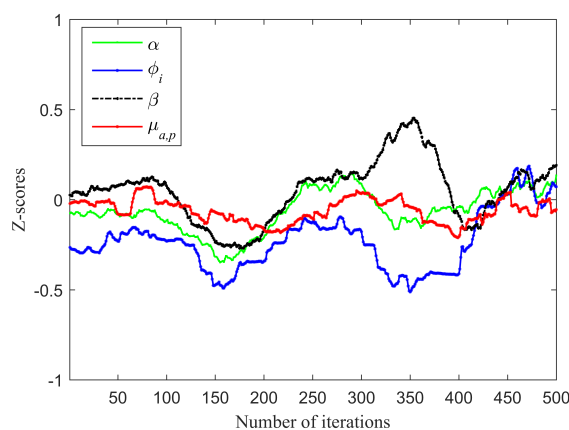
Table 2. Performance comparison of different q value.

q Value	WAIC	LOO	Training Time
1	101.2	104.7	470.5 s
2	130.5	128.8	366.7 s
3	165	146.1	799.0 s

As can be seen in Table 2, a higher q obtains better WAIC and LOO values. However, the training time is also increased when q equals 3. When $q = 1$, it has the lowest WAIC and LOO values. Here, the q is set to 2 as a compromise between the model performance and training time.

To verify the fusion performance of the Bayesian model, the BHBB is compared with regular HB, multivariate regression methods PLSR [29], WN [29], GP [30], and SVM [31]. For a fair comparison, the number of layers and basic structure of HB is the same as BHBB. The size of sample data in each repetition of bagging is set to 15 in three areas, and the number of subset data $m = 5$ for each area in BHBB. In PLSR, the number of components to keep is set to 3 and the tolerance of the iterative algorithm is set 1×10^{-6} . The nonlinear Wiener degradation structure in which the degradation parameters are integrated with environmental stress. The radial basis function (RBF) kernel is used in SVM. In addition, the penalty parameter and kernel coefficient are set to 1000 and 0.1, respectively. For GP, the RBF kernel is selected for the covariance functions. The performance of different models is evaluated through RMSE measures.

To verify the convergence of the BHBB, the z-scores proposed by Geweke are utilized to examine the MCMC chain. It compares the mean and variance at the beginning and end of the chain. The z-scores value should be located in $[-2, 2]$, which indicates the MCMC is convergence. The z-scores results of four parameters from different layers of the BHBB are shown in Figure 5. It can be seen from Figure 5 that all the z-scores of different parameters are not exceeded ± 0.5 , indicating that the BHBB has great convergence.

**Figure 5.** The BHBB framework for failure evaluation of PME.

The predicted results in three areas are presented in Figure 6. It should be noted that the gray dots in each column represent the raw failure data points. For the PLSR, GP, and SVM, it shows that the results of PLSR, GP, and SVM can follow the trend of failure data, while the WN cannot. In addition, PLSR and WN cannot meet the characteristics of the failure rate data when considering environmental features. For example, the results of PLSR and WN in Figure 6b are straight lines, and they fail to capture the features of the failure rate. Additionally, the curve of SVM completely follows the data between 2013 and 2015 in Figure 6b, thus SVM is overfitting the failure data. Moreover, the prediction of SVM and GP is lower than zero in 2012 of Figure 6a,b. However, the failure rate should always be no less than zero.

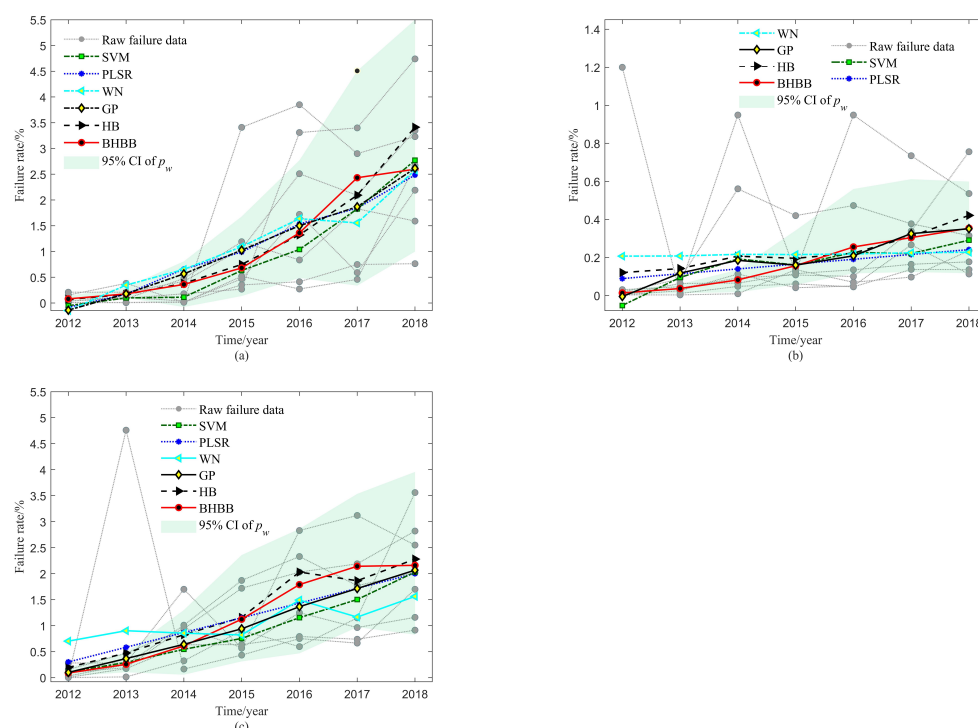


Figure 6. Failure rate prediction results of the four methods including SVM, PLSR, WN, GP, HB, and BHBB in (a) XZ, (b) XJ, and (c) HLJ.

For the regular HB, it can follow the change tendency of the failure rate data, indicating the effectiveness of the nonlinear regression. However, it shows that the curve of BHBB is smooth when integrating the features of the areas in Figure 6c. Both of the HB and BHBB can fit the data well in Figure 6a. However, it shows that the HB does not detect outliers because the curve is higher than the rest of the methods in Figure 6b.

As is clear from Figure 6b, some data points outside the 95% CI can be outliers in 2012, 2014, and 2016 of XJ. It shows that the mean curves generated by BHBB are more sensitive to outliers. However, the HB cannot identify outliers and change curve in Figure 6b,c. The 95% CI of p_w shows the normal failure data, where the outside of CI are more likely to be outliers. The outliers outside the 95% CI can be assigned a smaller weight. In this example, the BHBB recognizes that not all the failure data coming from truly Weibull parameter model p_w , the binary structure allows some outliers to fit the real world. Therefore, it is evident that the BHBB outperforms the other methods.

To highlight the effectiveness of the proposed method, the training and prediction RMSE of different methods are listed in Table 3. The RMSE is calculated from the mean of the curve. Additionally, the failure data outside the 95% CI are identified as outliers and, therefore, the outliers are replaced by the mean. It shows that the lowest predicted value of RMSE is BHBB, indicating that BHBB has a better precision performance. Furthermore, the comparison between BHBB and HB showed a significant improvement in prediction. Conversely, the RMSE of GP is lower than that of PLSR, WN, and HB. The highest RMSE is WN, which means that it has the worst fitting result. Intuitively, the results show that BHBB is more effective than traditional HB and deterministic approaches.

To verify the effectiveness of the outlier identification structure in BHBB, the mean and 95% CI of ϕ_i in Equation (4) are depicted in Figure 7. ϕ_i represents the proportion of outliers in the binary structure. It shows that the probability of outliers in XJ and HLJ is higher than XZ. The uncertain degree of outliers in XJ is larger due to more potential outliers leading to greater probability interval, which is consistent with the failure rate data according to Figure 6b. Broadly, the mean of the ϕ_i in XZ is the smallest, indicating that the

number of outliers is the least. Thus, it is worth mentioning that the BHBB has the ability to identify outliers from the failure rate data.

Table 3. Comparison of training and prediction RMSE results.

	Model	XZ	XJ	HLJ
Training	PLSR	0.861×10^{-2}	0.258×10^{-2}	0.797×10^{-2}
	SVM	0.857×10^{-2}	0.269×10^{-2}	0.869×10^{-2}
	WN	0.924×10^{-2}	0.317×10^{-2}	0.919×10^{-2}
	GP	0.862×10^{-2}	0.265×10^{-2}	0.817×10^{-2}
	HB	0.903×10^{-2}	0.259×10^{-2}	0.821×10^{-2}
	BHBB	0.855×10^{-2}	0.261×10^{-2}	0.811×10^{-2}
Prediction	PLSR	0.124×10^{-1}	0.307×10^{-2}	0.110×10^{-1}
	SVM	0.109×10^{-1}	0.296×10^{-2}	0.866×10^{-2}
	WN	0.143×10^{-1}	0.335×10^{-2}	0.132×10^{-1}
	GP	0.103×10^{-1}	0.275×10^{-2}	0.832×10^{-2}
	HB	0.127×10^{-1}	0.304×10^{-2}	0.115×10^{-1}
	BHBB	0.965×10^{-2}	0.271×10^{-2}	0.812×10^{-2}

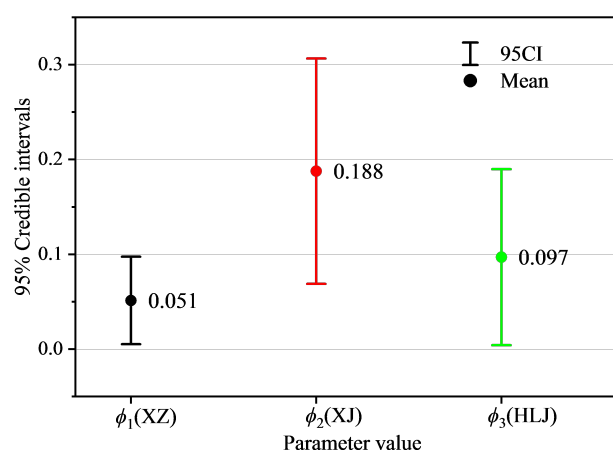


Figure 7. The confidence interval of ϕ_i for outlier identification structure in BHBB.

The value of WAIC is listed in Table 4. It worth mentioning that the pWAIC indicates the estimated effective number of parameters in models, which provides clues of how flexible each model is. A model r with larger weight w_r means a better Bayesian model. The pWAIC of BHBB is greater than HB, it shows that BHBB is flexible than HB due to the outlier model. The difference between each WAIC and the lowest WAIC is depicted by dWAIC_r. The huge dWAIC_r of HB reveals the difference between the two models. The WAIC of BHBB is lower than HB and the value w_r is higher than HB indicating the BHBB can make a better description for failure rate data.

Table 4. WAIC value of BHBB and HB under model averaging.

Model	WAIC	pWAIC	dWAIC _r	w_r
BHBB	17.11	20.84	0	1
HB	114.14	17.93	97.03	0

5.2. Model Interpretation and Prediction of Reliability

Another advantage of BHBB is that the model parameters have physical meanings. Here, the mainly posterior quantities and CI of parameters are listed in Table 5.

The β_0 denotes the intercept of nonlinear regression in Equation (2). It is assumed that all the electrical meters samples have the same intercept due to producing by the same company. The $\beta_{a,1}$ is the slope of time, $\beta_{a,3}$, and $\beta_{a,4}$ represent the slope of temperature and pressure, respectively. Different $a(a = 1, 2, 3)$ represents the slope of XZ, XJ, and HLJ in $\beta_{a,p}$, respectively.

Table 5. Main parameter values of BHBB.

Parameter $\beta_{a,p}$	Mean	SD	Confidence Interval	
			2.50%	97.50%
β_0	−0.784	1.877	−4.621	2.824
$\beta_{0,1}$	0.628	0.068	0.496	0.772
$\beta_{1,1}$	0.58	0.088	0.424	0.771
$\beta_{2,1}$	0.524	0.078	0.362	0.663
$\beta_{0,3}$	0.369	0.462	−0.54	1.21
$\beta_{1,3}$	0.312	0.441	−0.521	1.22
$\beta_{2,3}$	0.144	0.326	−0.409	0.884
$\beta_{0,4}$	−0.385	1.916	−3.879	3.32
$\beta_{1,4}$	0.845	1.944	−3.162	5.207
$\beta_{2,4}$	2.492	4.181	−4.108	11.52

Table 5 shows that all the $\beta_{a,1}$ are positive and this indicates the failure rate has a strong correlation with time. More importantly, the first feature $\beta_{a,3}$ has a higher rate of contribution than the second feature $\beta_{a,4}$, which means the higher temperature will cause more number of failures, and the electrical meters are sensitive to temperature changes while the pressure is constant.

To predict the failure rate of electrical meters in the long term, the forecasting result of reliability is presented in Figure 8. This result is calculated using the reliability index in Equation (16). The quantile denotes the upper and lower bounds of the CI.

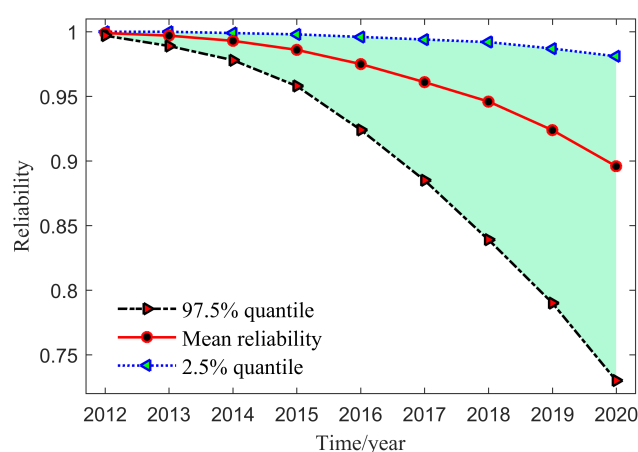


Figure 8. The prediction curve for the reliability of electrical meters. The quantile represents the boundary value of the CI.

It is observed that the reliability of electrical meter declines over time. The value of reliability is nearly 0.965 in the sixth year. It also shows the electrical meters produced by its company have high reliability during the operation. On the other hand, the slope of the reliability curve is approximate a straight line and the uncertainty of this system is

increasing, indicating the users do not need to change too much on the reserve and operate strategy in the short-term.

Overall, considering the security reliability of the PME in power system [32], a rotation strategy for practical PME reliability regulation and management can be developed so as to ensure the normal operation of PME. The PME can be replaced to prevent lack of equipment and reduce the cost of manufacturing, purchasing, and shipping.

6. Conclusions

In this paper, a binary hierarchical Bayesian with bagging method is proposed to evaluate the failure rate of PME from multiple typical locations. The inaccurate prediction results of failure rate and reliability in a small sample dataset with outliers have been addressed. This method overcomes the sole objective evaluation of failure rate, where multi-source information is embedded in BHBB. The visualization results of model parameters show that the model can identify outliers effectively for failure data of PME. Compared with the common hierarchical Bayesian and some traditional data-driven methods, the experiment manifests that the proposed method has superior performance in failure rate prediction even under a small sample dataset. The method adopts a combination of outlier identification structure and bagging, wherein the model parameters have the advantage of interpretability. Concretely, the convergence and variance experiments show that BHBB has great convergence and low RMSE. To further improve this methodology, future work will focus on the derivation of prior distribution in higher levels of hierarchical Bayesian.

Author Contributions: Conceptualization, J.L. and Q.T.; methodology, J.L. and W.Q.; software, J.M.; validation, J.L., W.Q. and J.M.; formal analysis, W.Q.; investigation, J.L.; data curation, J.D.; writing—original draft preparation, J.L.; writing—review and editing, Q.T. and W.Q.; visualization, J.M.; supervision, Q.T.; project administration, Q.T.; funding acquisition, Q.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by National Key R & D Program of China OF FUNDER grant number 2019YFF0216800, and in part by Postgraduate Scientific Research Innovation Project of Hunan Province OF FUNDER grant number CX20200426.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Conti, M.; Orcioni, S. Modeling of Failure Probability for Reliability and Component Reuse of Electric and Electronic Equipment. *Energies* **2020**, *13*, 2843. [\[CrossRef\]](#)
- Wang, H.; Liu, Z.; Xu, Y.; Wei, X.; Wang, L. Short text mining framework with specific design for operation and maintenance of power equipment. *CSEE J. Power Energy Syst.* **2020**, *1*, 1–10.
- Devlin, M.A.; Hayes, B.P. Non-Intrusive Load Monitoring and Classification of Activities of Daily Living Using Residential Smart Meter Data. *IEEE Trans. Consum. Electron.* **2019**, *65*, 339–348. [\[CrossRef\]](#)
- Alahakoon, D.; Yu, X. Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey. *IEEE Trans. Ind. Inform.* **2016**, *12*, 425–436. [\[CrossRef\]](#)
- Sun, K.; Xiao, H.; Pan, J.; Liu, Y. A Station-Hybrid HVDC System Structure and Control Strategies for Cross-Seam Power Transmission. *IEEE Trans. Power Syst.* **2021**, *36*, 379–388. [\[CrossRef\]](#)
- Liu, H.; Wang, Y.; Yang, Y.; Liao, R.; Geng, Y.; Zhou, L. A Failure Probability Calculation Method for Power Equipment Based on Multi-Characteristic Parameters. *Energies* **2017**, *10*, 704. [\[CrossRef\]](#)
- Qiu, W.; Tang, Q.; Teng, Z.; Yao, W.; Qiu, J. Failure rate prediction of electrical meters based on weighted hierarchical Bayesian. *Measurement* **2019**, *142*, 21–29. [\[CrossRef\]](#)
- Jin, L.; Peng, C.; Jiang, T. System-Level Electric Field Exposure Assessment by the Fault Tree Analysis. *IEEE Trans. Electromagn. Compat.* **2017**, *59*, 1095–1102. [\[CrossRef\]](#)
- Zhu, P.; Han, J.; Liu, L.; Lombardi, F. A Stochastic Approach for the Analysis of Dynamic Fault Trees With Spare Gates Under Probabilistic Common Cause Failures. *IEEE Trans. Reliab.* **2015**, *64*, 878–892. [\[CrossRef\]](#)

10. Lu, J.; Huang, J.; Lu, F. Sensor Fault Diagnosis for Aero Engine Based on Online Sequential Extreme Learning Machine with Memory Principle. *Energies* **2017**, *10*, 39. [\[CrossRef\]](#)
11. Lu, F.; Jiang, C.; Huang, J.; Wang, Y.; You, C. A Novel Data Hierarchical Fusion Method for Gas Turbine Engine Performance Fault Diagnosis. *Energies* **2016**, *9*, 828. [\[CrossRef\]](#)
12. Yang, Z.; Chen, Y.; Li, Y.; Zio, E.; Kang, R. Smart electricity meter reliability prediction based on accelerated degradation testing and modeling. *Int. J. Electr. Power Energy Syst.* **2014**, *56*, 209–219. [\[CrossRef\]](#)
13. Xu, D.; Wei, Q.; Elsayed, E.A.; Chen, Y.; Kang, R. Multivariate Degradation Modeling of Smart Electricity Meter with Multiple Performance Characteristics via Vine Copulas. *Qual. Reliab. Eng. Int.* **2017**, *33*, 803–821. [\[CrossRef\]](#)
14. Sundaravadivel, P.; Kesavan, K.; Kesavan, L.; Mohanty, S.P.; Kougianos, E. Smart-Log: A Deep-Learning Based Automated Nutrition Monitoring System in the IoT. *IEEE Trans. Consum. Electron.* **2018**, *64*, 390–398. [\[CrossRef\]](#)
15. Yuan, T.; Kuo, Y. Bayesian Analysis of Hazard Rate, Change Point, and Cost-Optimal Burn-In Time for Electronic Devices. *IEEE Trans. Reliab.* **2010**, *59*, 132–138. [\[CrossRef\]](#)
16. Mishra, M.; Martinsson, J.; Rantatalo, M.; Goebel, K. Bayesian hierarchical model-based prognostics for lithium-ion batteries. *Reliab. Eng. Syst. Saf.* **2018**, *172*, 25–35. [\[CrossRef\]](#)
17. Zhang, Y.; Li, M.; Dong, Z.Y.; Meng, K. Probabilistic anomaly detection approach for data-driven wind turbine condition monitoring. *CSEE J. Power Energy Syst.* **2019**, *5*, 149–158. [\[CrossRef\]](#)
18. Brown, R.E.; Frimpong, G.; Willis, H.L. Failure rate modeling using equipment inspection data. *IEEE Trans. Power Syst.* **2004**, *19*, 782–787. [\[CrossRef\]](#)
19. State Grid Corporation of China, Technical specification for single phase smart electricity meters, Q/GDW 1364-2013. 2013. Available online: <https://ishare.iask.sina.com.cn/> (accessed on 10 May 2021).
20. National Meteorological Information Center. 2021. Available online: <http://data.cma.cn/> (accessed on 20 May 2021).
21. Ciani, L.; Guidi, G. Application and analysis of methods for the evaluation of failure rate distribution parameters for avionics components. *Measurement* **2019**, *139*, 258–269. [\[CrossRef\]](#)
22. Salehi, M.; Leckie, C.; Bezdek, J.C.; Vaithianathan, T.; Zhang, X. Fast Memory Efficient Local Outlier Detection in Data Streams. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 3246–3260. [\[CrossRef\]](#)
23. Gevaert, O.; De Smet, F.; Kirk, E.; Van Calster, B. Predicting the outcome of pregnancies of unknown location: Bayesian networks with expert prior information compared to logistic regression. *Hum. Reprod.* **2006**, *21*, 1824–1831. [\[CrossRef\]](#)
24. Qiu, W.; Tang, Q.; Yao, W.; Qin, Y.; Ma, J. Probability Analysis for Failure Assessment of Electric Energy Metering Equipment Under Multiple Extreme Stresses. *IEEE Trans. Ind. Inform.* **2021**, *17*, 3762–3771. [\[CrossRef\]](#)
25. Gadrich, T.; Bashkansky, E. A Bayesian approach to evaluating uncertainty of inaccurate categorical measurements. *Measurement* **2016**, *91*, 186–193. [\[CrossRef\]](#)
26. Kontar, R.; Zhou, S.; Sankavaram, C.; Du, X.; Zhang, Y. Nonparametric-Condition-Based Remaining Useful Life Prediction Incorporating External Factors. *IEEE Trans. Reliab.* **2018**, *67*, 41–52. [\[CrossRef\]](#)
27. Watanabe, S.; Oppen, M. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *J. Mach. Learn. Res.* **2010**, *11*, 3571–3594.
28. Alexander, M.; Zagheni, E.; Barbieri, M. A Flexible Bayesian Model for Estimating Subnational Mortality. *Demography* **2017**, *54*, 2025–2041. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Hosani, E.A.; Meribout, M.; Al-Durra, A.; Al-Wahedi, K.; Teniou, S. A New Optical-Based Device for Online Black Powder Detection in Gas Pipelines. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 2238–2252. [\[CrossRef\]](#)
30. Rasmussen, C.E. Distributed Computing and Internet Technology. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 63–71.
31. Wang, S.; Ding, X.; Zhu, D.; Yu, H.; Wang, H. Measurement uncertainty evaluation in whiplash test model via neural network and support vector machine-based Monte Carlo method. *Measurement* **2018**, *119*, 229–245. [\[CrossRef\]](#)
32. Bashian, A.; Assili, M.; Anvari-Moghaddam, A. A security-based observability method for optimal PMU-sensor placement in WAMS. *Int. J. Electr. Power Energy Syst.* **2020**, *121*, 106157. [\[CrossRef\]](#)