

Fault Prediction of Online Power Metering Equipment Based on Hierarchical Bayesian Network

Da Cheng^{1,2}, Penghe Zhang², Fan Zhang³, Jiayu Huang⁴

¹Hunan University, Department of Electronic Science and Technology, Hunan, China.

²China electric power research institute, Beijing, China.

³Huludao Power Supply Company, Huludao, China;

⁴Beijing Normal University, Beijing, China.

Abstract: The failure rate assessment of online metering equipment is significant for power metering. For traditional methods, the performance of the model is not satisfactory especially in the case of small samples. In this paper, an online power measuring equipment fault evaluation method based on Weibull parameter hierarchical Bayesian model is proposed. Firstly, the z-score method is used to eliminate outliers in the raw failure data. Then, a generalized linear function with variable intercept is established according to the characteristics of failure data. The information of each region is merged using the characteristics of multi-layer Bayesian network uncertainty reasoning. The model parameters are updated based on the Markov chain Monte Carlo method. Thereafter, the trend of failure rate is provided with time-dependent. Finally, the proposed method is verified by the failure samples of the online measurement equipment in three typical environmental areas. The accuracy and validity of the hierarchical Bayesian model is verified by a series of experiments.

Keywords: failure rate; hierarchical Bayesian model; variable intercept; Weibull

Napovedovanje izpada na opremi merjenja moči na osnovi hierarhične Bayesianove mreže

Izveček: Ocena stopnje izpada na merilni opremi moči je zelo pomembna. Pri klasičnih metodah je učinkovitost ni zadostna, kar se pokaže predvsem pri majhnih vzorcih. V prispevku je predlagana metoda napovedi izpadov merilne opreme na osnovi Weibullovega parametra hierarhičnega Bayesianovega modela. Najprej so iz surovih podatkov odstranjeni neuporabni podatki na osnovi metode z-ocene. Informacija je nato združena z uporabo karakteristik večslojne Bayesianove mreže in nadgrajena z Markovo verigo po Monte Carlo metodi. Trendi so časovno odvisni. Rezultati so preverjeni in ovrednoteni z meritvami v treh tipičnih okoljih.

Ključne besede: stopnja odpovedi; hierarhičen Bayesianov model; Weibull

*Corresponding Author's e-mail: 13910761850@126.com

1 Introduction

Energy metering equipment, such as electric meters and collectors, have a large amount and wide distribution in the power grid. As the nerve endings of the smart grid, electricity meters play an irreplaceable role in electricity information collection and energy monitoring [1]. Meanwhile, the reliability of the metering equipment is also related to the safety of household electricity consumption. Inaccurate energy metering affects the strategic planning of power generation,

transmission and use of electricity [2]. The failure rate evaluation of the metering device over time and the fault are limited because it is difficult to collect large amounts of sample information. Therefore, it is of great significant to establish a scientific and reliability evaluation scheme for the reliability design of metering equipment.

In recent decades, a large number of methods are applied to target prediction and fault analysis. Generally,

these methods can be classified into two categories, namely deterministic method and probability method [3]. The Support Vector Regression (SVR) model is employed to identify fault and predict the remaining life of the reciprocating compressors based on the sensory data [4]. In addition, accurate bearing remaining useful life of machine's breakdown and maintenance's cost is proposed based on the artificial neural network [5]. These algorithms and methods achieve high precision prediction. However, enough samples are required to train the models before the effective prediction results is given. Moreover, the data is relatively simple, which limits its accuracy. The Principal Component Analysis (PCA) is utilized to select fault-relevant variable [6], so that a subset of variables based on training and validation data sets can be obtained to achieve better possible performance [7]. But this method has an unsatisfactory performance, especially in the absence of prior information. Consequently, the deterministic method lacks probability information for uncertain problems.

Probabilistic methods, one that can detect uncertainty in the data and provide more information, often used for failure analysis [8]. Fault Tree Analysis (FTA) is one of the most commonly used probabilistic methods. A system level electric field exposure assessment by FTA is proposed in [9]. However, the detailed system structure is needed for fault tree analysis. The Gaussian Process (GP) is another method for probability analysis [10]. The remaining useful life prediction for service units is used to improve the accuracy [11]. But the unexplained nature of this nonparametric model limits its application. Then, the Bayesian estimation is widely used in failure rate analysis [14]. For example, a Bayesian method is proposed for the hazard rate analysis of electronic devices [15]. Moreover, the fuzzy evidence theory, can be combined with Bayesian to solve the problem of insufficient samples [12][13]. Intuitively, probability method can provide more information, such as confidence interval and quantile, and is more suitable for the processing of small sample data.

To achieve failure rate analysis of online metering equipment, a Hierarchical Bayesian Network (HBN) is proposed based on a small number of electrical meters fault samples. A segmented Weibull parameter model is introduced for failure rate analysis at different times to better fit the data. Considering the different environments in different regions, part of the information fusion method is used in measurement equipment. Moreover, the real fault samples in three typical environments are used to verify the validity of HBN. The reliability of the metering device is calculated, and the parameters of the model are interpretable compared to traditional FTA and GP models.

The remaining part of this paper is organized as follows: Section II describes the outlier detection method in sample data. The proposed Weibull hierarchical Bayesian model is presented in Section III. Thereafter, the fault data of electrical meters in different regions is analyzed in Section IV. Finally, the conclusion is presented in Section V.

2 Failure rate data outlier detection

In the actual process of data collection, abnormal values may exist in the raw failure rate data due to the operational errors of researchers. The outliers not only affect the model's evaluation of the power metering equipment reliability, but also easily cause the model to be overfitting [16]. Traditional failure assessment methods are difficult to balance in small sample, failure rate data with missing data and outliers.

In order to clean up the abnormal value of the raw failure rate data and reduce the information loss, the Z-score method is used to clean the outliers in the fault rate data. The Z-score method determines the outliers by solving the relative standard distance of the data from the mean, which is suitable for the outlier detection under small samples.

Let $\mathbf{Y}=\{y_{s,t,j}\}$, where $y_{s,t,j}$ denotes the fault rate of the measuring device measured in the j^{th} time, $t=1,\dots,L$, $j=1,\dots,N$, and in the s area at the measurement time t . The standard deviation σ_t of the failure rate data for each measurement is:

$$\sigma_t = \sqrt{\frac{1}{N} \sum_{j=1}^N (y_{s,t,j} - u_{s,t,j})^2}, \quad t = 1, 2, \dots, L \quad (1)$$

where N is the total number of failure rate data when the measurement time is t , $u_{s,t,j}$ is the average of the failure rate data for each measurement and $y_{s,t,j}$ denotes the fault rate of the measuring device measured in the j^{th} time. Considering the discontinuity of the sample data and reducing the loss of valid data, a window of length three is used to analyze the outliers in failure rate data, and the window is composed of three measurements of $[t, t+1, t+2]$. That is, the standard deviation of three consecutive measurements is calculated each time, and the Z score Z_j of each data point $y_{s,t,j}$ is calculated as

$$Z_j = \frac{y_{s,t,j} - u'_{s,t,j}}{\sigma'_t} \quad (2)$$

where $u'_{s,t,j}$ are the mean failure rate of three consecutive measurements, σ'_t is the standard deviation of $u_{s,t,j} =$

$u'_{s,tj}$ in equation (1), at which point Z_j can be regarded as obeying the normal distribution $N(0, 1)$. Finally, the threshold for determining the abnormal data is set to three times the standard deviation of $N(0, 1)$. The outliers of the original failure rate data with the score $|Z_j|$ greater than the threshold is replaced by the mean $u'_{s,tj}$. Then the failure rate data $\mathbf{Y}^* = \{y^*_{s,tj}\}$ without the outlier is obtained.

3 Weibull parameter model for metering equipment failure rate

3.1 Other Recommendations

The commonly used online power metering equipment includes electrical meters, power quality detecting devices and concentrators, etc. This paper takes the electrical meter as the target to analyze. The electric energy metering equipment is a high-precision measuring instrument. The damage of the weakest link of the system will directly lead to the failure of the measuring equipment. The Weibull distribution is used as the fault description of the commonly used electronic instruments [17], so the Weibull distribution is used to fit the electric energy meter data, the probability density function of Weibull is

$$f(t | \lambda, \alpha, \chi) = \lambda \alpha (t - \chi)^{\alpha - 1} \exp[-\lambda (t - \chi)^\alpha] \quad (3)$$

$t > 0, \lambda > 0, \alpha > 0, \chi \geq 0$

where t is the fault time, λ is the scale parameter, α is the shape parameter, and χ is the location parameter. We set χ to 0 considering that the failure starts after $t=0$. Then the χ is reduced to a two-parameter Weibull distribution. When the Weibull distribution probability density function $f(t)$ is integrated, the value is always not less than 0, which is consistent with the condition that the number of failures is not less than 0.

The observed data \mathbf{Y}^* obeys the Weibull regression model, which can be expressed as

$$\mathbf{Y}^* \sim \text{Weibull}(\lambda, \alpha) \quad (4)$$

Generally, the change of the shape parameter α indicates that the failure mechanism changes. Therefore, the shape parameter α is set to obey the fixed distribution function. At this time, the regression model can be established by changing the scale parameter λ in the Weibull regression model.

When the observed data \mathbf{Y}^* is the metering device fault data, the fault number \mathbf{Y}^* will change according to the time variable x . Therefore, Weibull distribution function

will change, and a regression model can be established based on λ

$$\lambda = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k = \mathbf{x}^T \boldsymbol{\beta} \quad (5)$$

where β_k is the regression coefficient, x_k is the covariate. Considering the condition that the scale parameter λ is greater than 0, the value of the regression coefficient is limited. To select a wider distribution parameter as the prior distribution of β_k , the $\log()$ correlation function is used to limit the range

$$\log(\lambda) = \mathbf{x}^T \boldsymbol{\beta} + w_i \quad (6)$$

where w_i is the measurement error when measuring the number of failures \mathbf{Y}^* , i is the number of sub-sample categories divided by the HBN model based on the total sample information. The $\log(\lambda)$ range becomes $(-\infty, +\infty)$, and λ is inversely solved from equation (4).

$$\lambda = \exp(\mathbf{x}^T \boldsymbol{\beta} + w_i) \quad (7)$$

At this time, there is no limit to the prior distribution of β_k . When the covariate changes to Δx , the contribution to the scale factor becomes $\exp(\Delta x \boldsymbol{\beta})$.

3.2 Structure of Hierarchical Bayesian Model

In order to realize the failure rate evaluation and prediction of the measuring equipment, Bayesian networks (BN) are used to fuse data from different regions. BN takes full advantage of early prior information and sample data information to achieve a full estimate of the latest events [18]. In particular, HBN takes advantage of information between levels to provide better data fitting capabilities [19].

For BN, the posterior probability density based on prior parameters θ is

$$p(\theta | \mathbf{Y}^*) = \frac{f(\mathbf{Y}^* | \theta) p(\theta)}{\int f(\mathbf{Y}^* | \theta) p(\theta) d\theta} \quad (8)$$

where $f(\mathbf{Y}^* | \theta)$ is the likelihood function of the Bayesian model, $p(\theta)$ is the model prior distribution. The HBN model refers to the different levels of data parameters elaborated by other layer parameters, which specifies another layer of prior knowledge for a layer of parameters.

The denominator in equation (8) is independent of the parameter θ . When the Weibull distribution is used as the likelihood function of the model, the parameter posterior distribution is proportional to the numerator of (8), and it can be further expressed as

$$p(\theta | \mathbf{Y}^*) \propto f(t | \theta) p(\theta | \alpha, \boldsymbol{\beta}, w_i) \quad (9)$$

where $p(\theta|\alpha, \beta, w_i)$ represents the prior distribution of the parameters α, β, w_i .

In the Weibull distribution, the size of the parameter α determines the increase and decrease of the failure rate. When $\alpha > 1$, it indicates that the instrument is in the wear stage. And when $\alpha < 1$, which is suitable for the early failure stage of the instrument. Therefore, the parameter prior selection must meet certain conditions.

In the absence of any prior information α , according to the influence of α on the shape of Weibull distribution, the prior distribution f can be obtained by the HalfCauchy distribution.

$$p(\theta | \mathbf{Y}^*) \propto f(t | \theta) p(\theta | \alpha, \beta, w_i) \tag{10}$$

where b is the scale parameter of the distribution, and equation (10) can be abbreviated as HalfCauchy(b).

For the regression coefficient β , the setting of the log correlation function avoids limiting the range of coefficient values, and the prior distribution f of the regression coefficient β can be set to normal distribution, which can be defined as

$$f(u, \tau | t) = \sqrt{\frac{\tau}{2\pi}} \exp\left\{-\frac{\tau}{2}(t-u)^2\right\} \tag{11}$$

where u is the normal distribution mean, $1/\tau$ is the distribution variance. The measurement error w_i value is small, so the prior distribution of w_i may take a normal distribution with a small variance as well as the mean $u = 0$, and equation (11) can be abbreviated as $N(u, 1/\tau)$.

Substituting equation (10) and (11) into (7), the joint prior distribution of the parameters λ and α can be calculated as

$$p(\beta, b, w_i) = \frac{2 \exp(\mathbf{x}^T \beta + w_i)}{\pi b [1 + (t/b)^2]} \tag{12}$$

Similarly, substituting equation (12) and (3) into (9), the joint posterior distribution λ and α can be expressed as

$$p(\lambda, \alpha | \mathbf{Y}^*) \propto \frac{2 \exp(\mathbf{x}^T \beta + w_i - \lambda t^\alpha)}{\pi b [1 + (t/b)^2]} t^{\alpha-1} \tag{13}$$

Then the posterior distributions of the parameters λ and α in the hierarchical Bayesian model are

$$p(\lambda | \alpha, \mathbf{Y}^*) \propto t^{\alpha-1} \exp(w_i - \lambda t^\alpha) \times \exp(\mathbf{x}^T \sqrt{\frac{\tau}{2\pi}} \exp\left\{-\frac{\tau}{2}(t-u)^2\right\}) \tag{14}$$

$$p(\alpha | \lambda, \mathbf{Y}^*) \propto \frac{\exp(w_i - \lambda t^\alpha)}{\pi b [1 + (t/b)^2]} t^{\alpha-1} \tag{15}$$

After the parameter posterior distribution is obtained, the influence of the variation of the covariate \mathbf{x} on the failure rate of the model can be reflected by the parameter mean value and the confidence interval.

When the new covariate \mathbf{x} changes to \mathbf{z} , the fault value can be calculated based on the existing fault number \mathbf{Y}^* , and the posterior distribution of the parameter is

$$p(\mathbf{z} | \mathbf{Y}^*) = \sum \sum f(\mathbf{z} | \theta) p(\theta | \mathbf{Y}^*) d\theta \tag{16}$$

3.3 Failure rate Prediction of Metering Equipment

When the equipment fails, the rate of failure change does not always obey the fixed distribution. At this time, the regression coefficient β can be segmented according to the trend of failures number \mathbf{x} . Then β can be expressed as

$$\beta \sim \begin{cases} N(u_1, 1/\tau_1), & x < \sigma \\ N(u_2, 1/\tau_2), & x \geq \sigma \end{cases} \tag{17}$$

where σ is the time node at which the failure rate changes. Since only the trend of data can be observed, the failure rate change node cannot accurately specified. Thus, no information can be specified on the prior distribution σ

$$f(\sigma) = \frac{1}{\eta_1 - \eta_2} \tag{18}$$

where η_1 and η_2 are upper and lower limits of the density function. The regression model after segmentation can better follow the change of the failure rate of the metering equipment.

4 Fault analysis of metering equipment

4.1 Failure Data of Electrical Meters

In order to verify the effect of HBN model on the failure rate of metering equipment in different provinces, we analyze partial fault samples of Electrical Meters produced by the same company in Xizang (XZ), Xinjiang (XJ) and Heilongjiang (HLJ) province in the period of 2012-2017.

The fault data is collected from multiple locations, as shown in Fig. 1. The operating status of the electrical meter is transmitted to the concentrator via the power line carrier. Then the failure data in different areas are

transmitted to the base station through GPRS. Particularly, Fig. 1 demonstrates an example of an electricity meter with a measure fault. In this way, the number of electrical meter faults in different regions can be statistically analyzed in real time.

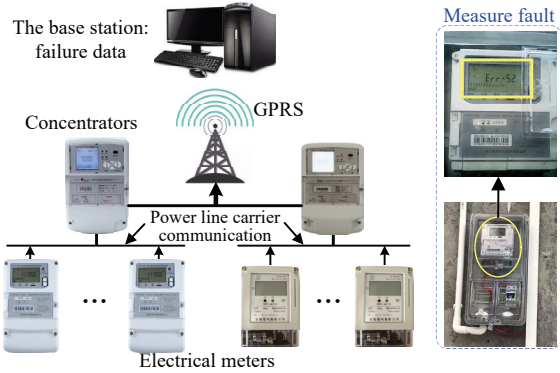


Figure 1: The fault data acquisition framework of the measuring equipment

Fig. 2-4 show the failure rates of the three areas, respectively. From the Fig. 2-4, it can be seen that there are seven groups of data in XZ and XJ, six groups in HLJ. All data are extracted independently, and each set of data contains six data points for the period 2012-2017. The failure rate is calculated by dividing the number of failures of the electrical meters by the total operation of the metering equipment each year.

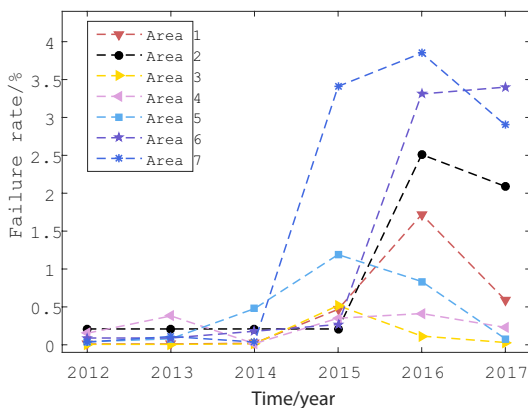


Figure 2: The electrical meters failure rate curve of XZ

As we can see from Fig. 2 and 3, the failure rate increased in the first five years and decreased in the sixth year and the failure rate of XZ is slightly higher than XJ. Fig. 2-4 show that the failure rate of metering equipment varies in different provinces. Therefore, through modeling and analysis of this example, the reliability variation relationship of metering equipment in different provinces can be found and the accurate prediction results is given. In addition, the data of 2014 in

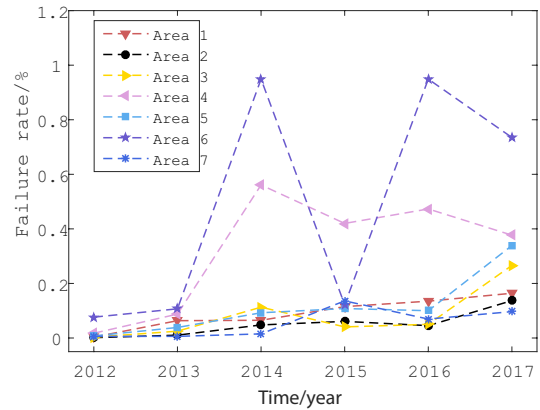


Figure 3: The electrical meters failure rate curve of XJ

Fig. 3 and one point in 2013 in Fig.4 are too large to be outliers.

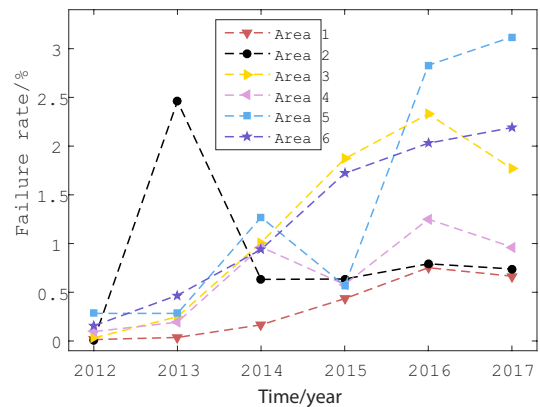


Figure 4: The electrical Meters failure rate curve of HLJ

The model analysis software uses the simulation platform based on Pycharm. Moreover, the Pymc3 based on Python is used to analyze the collected data [20].

4.2 Model structure and experimental steps

The structure of the hierarchical Bayesian model for the failure rate of online power metering equipment is as follows:

First floor:

$$Y \sim \text{Weibull}(\lambda, \alpha) \tag{19}$$

$$\lambda = \exp(\beta_i + x\beta_x + w_i) \tag{20}$$

$$\alpha \sim \text{HalfCauchy}(10) \tag{21}$$

Second floor:

$$\beta_i \sim N(0, 10^5) \tag{22}$$

$$\beta_x \sim \begin{cases} N(0,10^5), & x < \sigma \\ N(0,10^5), & x \geq \sigma \end{cases} \quad (23)$$

$$\sigma \sim \text{DiscreteUniform}(3,6) \quad (24)$$

$$w_i \sim N(0,10^{-3}) \quad (25)$$

According to the data of three provinces, the intercept i of formula (20) is set to 3 to reflect the difference of fault data of metering equipment in the three provinces. The covariant x takes one term, and the regression coefficient β_x of the three provinces is set to the same value to reflect the commonness of the products produced by the same company. According to the change of data, the failure rate change node σ can be set between 3-6 years, that is, the period of fault change fluctuation is 2014-2017. The flow chart of algorithm experiment is shown in Fig. 5.

According to flowchart Fig. 5, in order to accurately analyze the occurrence of faults in the three provinces, the algorithm flow can be summarized into the following four steps:

- Data preprocessing: The failure number of electrical meters is transformed into the failure rate. Different environmental characteristics are normalized to reduce the impact of units.
- Model establishment: Establish fault prediction model based on HBN. Then a priori distribution of model parameters is specified in conjunction with no information prior.
- Model solving: Combined simulation platform Pymc3, the MCMC posterior sampling method is used to get the target optimization parameter [21]. We use the NUTS sampling method to achieve fast convergence, where the HBN model sampling iterations are set to 10000, and the pre-firing period is set to 2000.
- Model verification: If the acceptance probability does not match the target, we than increase the number of samples or adjust the model prior distribution until it satisfies the acceptance probability.

In the process of establishing Markov chain by using MCMC sampling method to solve the Model, NUTS sampling has good effect on high-dimensional and long data, avoiding the influence of step size on sampling robustness and converges faster than Gibbs sampling method. Therefore, NUTS algorithm is used to sample the model to achieve fast convergence [22][23].

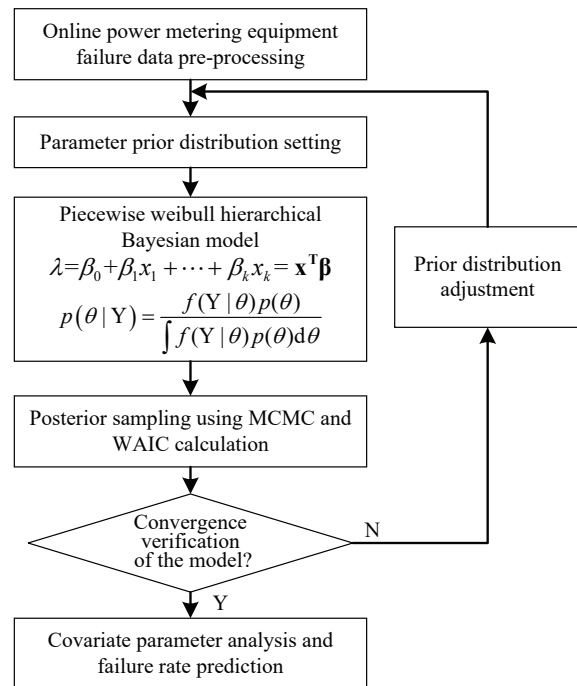


Figure 5: The experimental flowchart of HBN model algorithm

In order to verify the accuracy of the model, after solving the parameters of the prior distribution, the likelihood function is sampled to verify whether the posterior distribution of the failure rate conforms to the original failure rate data distribution. According to the posterior mean of the parameters, the prior parameters of the model are continuously adjusted to achieve the fitting of the model to the original failure rate data. Finally, the posterior distribution of linear function parameters with variable intercept is obtained, and the failure rate of metering equipment in the next year in three provinces is predicted respectively.

4.3 Experimental results analysis

The HBN model sampling iterations are set to 10,000, and the pre-firing period is set to 2,000. In order to verify the accuracy of the parameter results of the HBN model, the Maximum a Posteriori (MAP) estimate of the optimal solution is compared with the results of the HBN model. And Table I shows the results of the HBN model and the parameters calculated by MAP.

It can be seen from Table I that the intercept β difference of linear function is about 1, indicating that there is a certain difference in metering equipment in the three provinces, and this difference is related to the province. The difference of time coefficient early_β and late_β indicates that the increasing trend of failure rate has changed. In 2014, the change trend of failure rate

of metering equipment changed, and the failure rate showed a downward trend before 2015. The value of measurement error w is small, which indicates indicating that the error in the measurement process has little influence on the failure rate of metering equipment. The error between the mean value of the HBN model and that of the MAP model is small, which indicates the validity of the HBN model.

Table 1: Parameter results of HBN model

Variable	Mean	Standard deviation	2.5% quantile	97.5% quantile	MAP estimation
β_0	-3.388	0.311	-4.010	-2.803	-3.412
β_1	-4.510	0.303	-5.120	-3.939	-4.603
β_2	-2.939	0.306	-3.518	-2.293	0.932
early_βx	0.761	0.087	0.003	0.459	0.724
late_βz	0.593	0.067	0.002	0.796	0.917
α	0.938	0.066	0.001	0.816	1.075
w	2×10^{-6}	1×10^{-4}	2×10^{-6}	1.8×10^{-4}	2×10^{-4}

Fig. 6 shows the Weibull posterior distribution of the failure rate in the three provinces. The red dashed line is the mean value of the original failure rate data in the three provinces. The mean value of the Weibull posterior distribution is consistent with the location of the red line, indicating that the distribution of the failure rate data in the last six years obeys the Weibull parameter model.

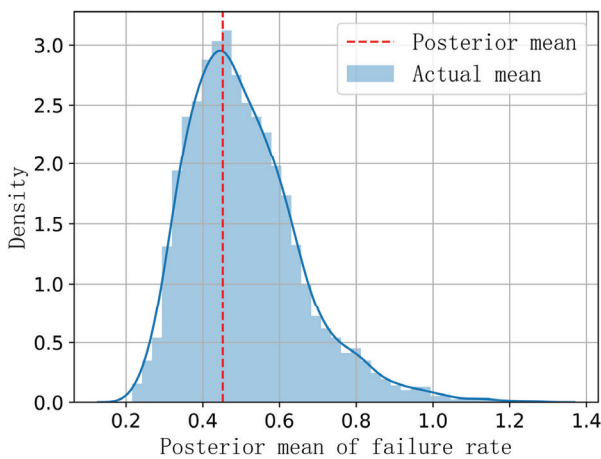


Figure 6: Posterior mean distribution of HBN model

To verify the validity of the model, we compare the proposed model with the pooling HBN, Poisson model and Cauchy model. Note that the all parameters have only one value in pooling HBN. On the other hand, the difference between proposed method and Poisson model, Cauchy model is the observed values follow different distributions. For example, the Poisson model means that the observed data Y^* obeys the Poisson re-

gression model. It's worth mentioning that all the prior distributions are the same for a fairer comparison.

Fig. 7-9 show a comparison of linear fitting curve mean, confidence interval and original failure rate data, respectively. Note that many of the gray lines are derived from predicted values within the confidence interval. The lowest and highest gray curves are confidence intervals of approximate 95% linear function. The failure data of measuring equipment are basically within 95% confidence interval after the fusion analysis of fault data using the method presented in this paper. The linear function can follow the jump trend of failure rate, which shows that the HBN model can fit the fault data of metering equipment well.

Moreover, all the other methods only have good performance on partial data. For example, the Poisson model can follow the trend of the data in XZ and HLJ. However, it have poor performance in XZ, which has a lower failure rate. At the same time, the pooling HBN cannot follow the trend of failure rate in XJ due to the model assumes the same trend for data in all regions.

In order to accurately predict the failure rate of measuring equipment, the time data is set to 7, which means the failure rate in 2018. After the model is substituted, the mean value of the failure rate in 2018 can be predicted as follows: the average failure rate of measurement equipment in XZ is 2.531, the average failure rate in XJ is 0.778 and the HLJ is 2.546. The predicted failure rate is within the confidence interval of 2017, which indicates that the model has the ability of short-term prediction.

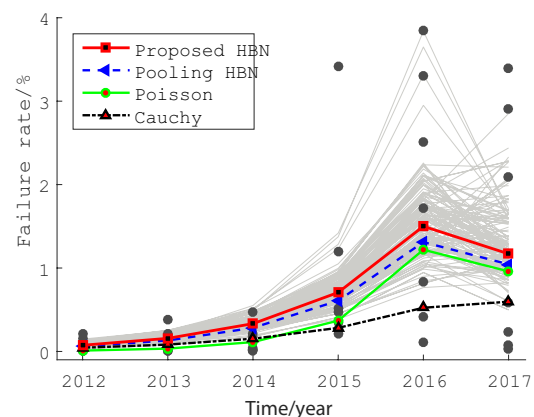


Figure 7: Comparison of fitting curve and the failure rate of the XZ Province

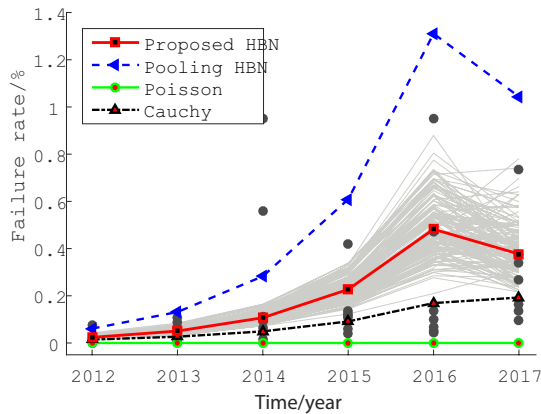


Figure 8: Comparison of fitting curve and the failure rate of XJ

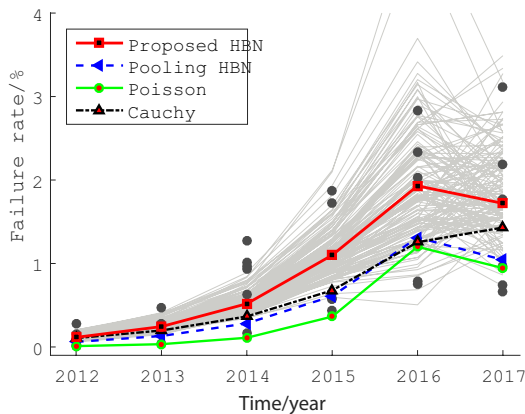


Figure 9: Comparison of fitting curve and the failure rate of HLJ

In order to evaluate the effect of partial information fusion of HBN model and realize the evaluation of different models, the Widely Available Information Criterion (WAIC) is more accurate than the traditional DIC when considering the whole posteriori distribution of the model [24]. Table II gives the WAIC calculation values of the four models under the same linear function condition and the same observation data.

Table 2: The comparison of WAIC between HBN and other models

Models	WAIC	pWAIC	weight	SE
HBN	-24.43	7.13	0.95	38.39
Pooling-HBN	7.26	3.71	0	35.39
Possion	123.3	5.03	0.05	19.32
Cauchy	138.04	14.09	0	36.5

It can be seen from Table II that the information criterion of partial information fusion HBN model is smaller, the number of effective parameters of the model is 7.13,

and the weight of model is 0.95, which is much larger than the rest models. The SE of the proposed method (HBN) is slightly larger than other models. HBN has better and the smallest WAIC value, which means that its uncertainty is consistent with its accuracy. Thus the effectiveness of partial information fusion HBN model is demonstrated. Conversely, the rest of the models have higher WAIC values, including that the models cannot effectively fit the raw data.

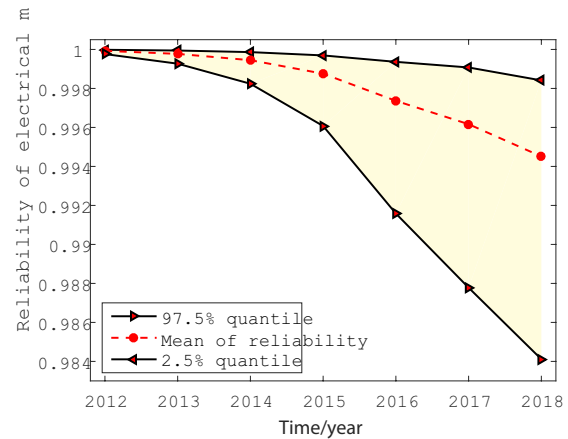


Figure 10: Reliability Curves and Confidence Intervals of power metering equipment

When the new electric energy meter equipment is put into use in the region, the failure rate of the electric energy metering equipment can be calculated according to equation (16). The reliability of the short-term prediction result of the new electric energy meter equipment is shown in Fig. 10. Fig. 10 shows that the reliability of the failure rate of the energy metering equipment decreases with time. After 6 years of operation, the reliability of the energy metering equipment is about 0.93, indicating that the type of energy metering equipment has a higher reliability under the three regional environmental conditions. In addition, the reliability curve approximates a straight line, indicating that the failure rate trend of the energy metering equipment is relatively flat in the short term. Moreover, the confidence interval of the reliability of the energy metering equipment is gradually increased, indicating that the uncertainty becomes large. Overall, the operation strategy does not need to change greatly.

5 Conclusion

By analyzing the fault data characteristics of online power metering equipment in different regions, a multi-bass Bayesian-based Weibull parameter regression model is established. Firstly, the Z-score method is used to clean the data outliers to reduce the interference of the outliers on the model. Then, according to the fault

data of the three regions, the intercept variable regression model of partial information fusion is established, and the failure rate data change node is fully considered. In the case, the influence of time factor on the failure rate of metering equipment in the three regions is obtained, and the average forecasting rate of the metering equipment failure rate in the seventh year is given. The example results verify that the method can effectively evaluate the failure rate relationship between the measuring equipment and different areas, and prove the feasibility of the scheme. Prior selection in small samples remains a challenge, and future work will focus on parameter settings for small sample models.

6 Acknowledgment

This paper is supported by Research on influence and evaluation of key components on relay protection performance.

7 References

1. H. Qing, L. Fangxing, W. Lei, et al. „Smart Meter Working Status Evaluation Method Based on Evidence Theory,” 2018 Conference on Precision Electromagnetic Measurements (CPEM 2018), Paris, 2018, pp. 1-2, 2018. <https://doi.org/10.1109/CPEM.2018.8501081>.
2. MT. Mohammadat, SM. Nafie, FB. Sharief, On electronic energy meters diagnosis exploiting components' susceptibility to failure. 2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering (ICCNEEE). pp. 418-421, 2015. <https://doi.org/10.1109/ICCNEEE.2015.7381404>.
3. Q. Xu, D. He, N. Zhang, et al. „A Short-Term Wind Power Forecasting Approach With Adjustment of Numerical Weather Prediction Input by Data Mining,” IEEE Transactions on Sustainable Energy, vol. 6, no. 4, pp. 1283-1291, 2015. <https://doi.org/10.1109/TSTE.2015.2429586>.
4. X. Li, F. Duan, I. Bennett, et al. „Canonical variate analysis, probability approach and support vector regression for fault identification and failure time prediction,” Journal of Intelligent & Fuzzy Systems, vol. 34, no. 6, pp. 3771-3783, 2018. <https://doi.org/10.3233/JIFS-169550>.
5. J. Ben Ali, B. Chebel-Morello, L. Saidi, et al. „Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network,” Mechanical Systems and Signal Processing, vol. 56-57, pp. 150-172, 2015. <https://doi.org/10.1016/j.ymssp.2014.10.014>.
6. Q. Jiang, X. Yan, B. Huang. „Performance-Driven Distributed PCA Process Monitoring Based on Fault-Relevant Variable Selection and Bayesian Inference,” IEEE Transactions on Industrial Electronics, vol. 63, no. 1, pp. 377-386, 2016. <https://doi.org/10.1109/TIE.2015.2466557>.
7. K. Ghosh, M. Ramteke, R. Srinivasan. „Optimal variable selection for effective statistical process monitoring,” Computers & Chemical Engineering, vol. 60, no. 10, pp. 260-276, 2014. <https://doi.org/10.1016/j.compchemeng.2013.09.014>.
8. M. Ram. „On system reliability approaches: a brief survey,” International Journal of System Assurance Engineering and Management, vol. 4, no. 2, pp. 101-117, 2013. <https://doi.org/urn:issn:0975-6809>.
9. L. Jin, C. Peng, T. Jiang. „System-Level Electric Field Exposure Assessment by the Fault Tree Analysis,” IEEE Transactions on Electromagnetic Compatibility, vol. 59, no. 4, pp. 1095-1102, 2017. <https://doi.org/10.1109/TEMC.2017.2647961>.
10. D. Lee, R. Baldick. „Short-Term Wind Power Ensemble Prediction Based on Gaussian Processes and Neural Networks,” IEEE Transactions on Smart Grid, vol. 5, no. 1, pp. 501-510, 2014. <https://doi.org/10.1109/TSG.2013.2280649>.
11. R. Kontar, S. Zhou, C. Sankavaram, et al. „Nonparametric-Condition-Based Remaining Useful Life Prediction Incorporating External Factors,” IEEE Transactions on Reliability, vol. 67, no. 1, pp. 41-52, 2018. <https://doi.org/10.1109/TR.2017.2717190>.
12. Yazdi M, Kabir S (2018). “Fuzzy evidence theory and Bayesian networks for process systems risk analysis”. Hum Ecol Risk Assess An Int J (0),1–30. <https://doi.org/10.1080/10807039.2018.1493679>.
13. Yazdi M, Kabir S (2017). “A fuzzy Bayesian network approach for risk analysis in process industries”. Process Saf Environ Prot, (111) ,507–19. <https://doi.org/10.1016/j.psep.2017.08.015>.
14. M. Han. „E-Bayesian estimation and hierarchical Bayesian estimation of failure rate,” Applied Mathematical Modelling, vol. 33, no. 4, pp. 1915-1922, 2009. <https://doi.org/10.1080/03610926.2010.498643>.
15. T. Yuan, Y. Kuo. „Bayesian Analysis of Hazard Rate, Change Point, and Cost-Optimal Burn-In Time for Electronic Devices,” IEEE Transactions on Reliability, vol. 59, no. 1, pp. 132-138, 2010. <https://doi.org/10.1109/TR.2010.2040776>.
16. M. Gupta, J. Gao, CC. Aggarwal, et al. „Outlier Detection for Temporal Data: A Survey,” IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 9, pp. 2250-2267, 2014. <https://doi.org/10.1109/TKDE.2013.184>.

17. B. Benedik, K. Taškova, J. Tavčar, et al. „Prediction of vacuum cleaner motor brush life: a regression approach,” IET Electric Power Applications, vol. 9, no. 9, pp. 569-577, 2015. DOI: 10.1049/iet-epa.2014.0437.
<https://doi.org/10.1049/iet-epa.2014.0437>
18. B. Cai. „Bayesian Networks in Fault Diagnosis,” IEEE Transactions on Industrial Informatics, vol. 13, no. 5, pp. 2227-2239, 2017.
<https://doi.org/10.1109/TII.2017.2695583>.
19. M. Han. „E-Bayesian estimation and hierarchical Bayesian estimation of failure rate,” Applied Mathematical Modelling, vol. 33, no. 4, pp. 1915-1922, 2009.
<https://doi.org/10.1080/03610926.2010.498643>.
20. J. Salvatier, TV. Wiecki, C. Fonnesbeck. „Probabilistic programming in Python using PyMC3,” PeerJ Computer Science, vol. 2, pp. e55, 2016.
<https://doi.org/10.7717/peerj-cs.55>.
21. Z. Hao, J. Guo, S. Zeng. „Fully Bayesian reliability assessment of multi-state systems with overlapping data,” Journal of Systems Engineering and Electronics, vol. 28, no. 1, pp. 187-198, 2017.
<https://doi.org/10.21629/JSEE.2017.01.21>.
22. S. Watanabe. „Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory,” Journal of Machine Learning Research, vol. 11, no. 18, pp. 3571-3594, 2010.
<https://doi.org/10.1002/rnc.1572>
23. S. Sun, Y. Feng, C. Dong, et al. „Efficient SRAM Failure Rate Prediction via Gibbs Sampling,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 31, no. 12, pp. 1831-1844, 2012.
<https://doi.org/10.1109/TCAD.2012.2209884>.
24. A. Vehtari, A. Gelman, J. Gabry. „Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC,” Statistics and Computing, vol. 27, no. 5, pp. 1413-1432, 2017.
<https://doi.org/10.1007/s11222-016-9696-4>.



Copyright © 2019 by the Authors.

This is an open access article distributed under the Creative Commons Attribution (CC BY) License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Arrived: 23. 02. 2019
Accepted: 19. 08. 2019