IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 15, 2020, accepted September 18, 2020, date of publication September 23, 2020,
date of current version October 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3026171

A Data Mining Approach for Transformer Failure Rate Modeling Based on Daily Oil Chromatographic Data

WEI HUANG[1], XUAN LI[1], BO HU[1], (Member, IEEE), JIAHAO YAN[1], (Member, IEEE), LVBING PENG[1], YUE SUN[1], XIN CHENG[1], JINFENG DING[1], KAIGUI XIE[1], (Senior Member, IEEE), QINGLONG LIAO[2], (Member, IEEE), AND LINGYUN WAN[2], (Member, IEEE)

[1]State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400044, China
[2]State Grid Chongqing Electric Power Research Institute, Chongqing 400000, China

Corresponding author: Kaigui Xie (kaiguixie@vip.163.com)

This work was supported by the Science and Technology Project of State Grid Corporation of China under Grant 5100-201999332A-0-0-00.

ABSTRACT Evaluating the real-time failure rate of transformers can effectively guide the planning of maintenance and reduce their failure risk. This paper proposed a novel transformer failure rate model that considers the impact of maintenance based on daily oil chromatographic monitoring data mining. Firstly, to ensure the quality of the modeling data, an improved k-nearest neighbor (KNN) algorithm based on genetic algorithm (GA) is proposed to repair the missing monitoring data. The repaired data is then mapped to the equivalent state duration (ESD) by the M-BPNN proposed, which is used to modify the multistate Markov process of transformers so as to quantify the impact of maintenance on failure rate. Considering the changing characteristics of the dissolved gases' content in the short period, the ESD is further merged in sequential periods to obtain the merged equivalent state duration (MESD). Finally, an analytical function of the transformer failure rate with respect to the MESD is obtained. Case studies on a typical substation demonstrate that the proposed approach has the ability to characterize the impact of maintenance and the actual failure rate, thereby improving the accuracy of the substation reliability assessment.

INDEX TERMS Oil-immersed transformer, data mining, oil chromatographic data, refined failure rate model, substation reliability.

NOMENCLATURE

Most of the symbols and notations used throughout this paper are defined below for quick reference.

A. ABBREVIATIONS

DGA	Dissolved gas analysis
SD	State duration
ESD	Equivalent state duration
MESD	Merged equivalent state duration

B. INDICES

n	Index of days
k	Index of fault types
l	Index of dissolved gases
u, v	Index of parent samples
s	Index of child samples
y	Index of the missing dissolved gas
w	Index of nearest DGA data samples
t	Index of operation time
f	Index of data samples from a specific transformer

C. PARAMETERS

$\boldsymbol{G}$	The target matrix $\boldsymbol{G}$ of storing the DGA data
$\boldsymbol{G}_k$	The DGA data matrix of fault type k
N_k	The number of days in fault type k
$\boldsymbol{g}_s$	The child samples before mutation
$\boldsymbol{g}'_s$	The child samples after mutation
$\boldsymbol{g}_{k,\min}$	The minimum values of each dissolved gas content in G_k
$\boldsymbol{g}_{k,\max}$	The maximum values of each dissolved gas content in G_k

The associate editor coordinating the review of this manuscript and approving it for publication was Ying Xu.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

$\boldsymbol{g}_{train,w}$	The w^{th} sample in $\boldsymbol{G}_{train}$
$\boldsymbol{g}_{test}(l)$	The l^{th} dissolved gas in $\boldsymbol{g}_{test}$
$\hat{\boldsymbol{g}}_{test}(y)$	The repairing value of the missing data
$\zeta_{i,j}$	The transition rate from state i to state j
t_i	The SD corresponding to operation time t
$t_{i_n}^{''}$	The ESD corresponding to operation time t
$t_i^{''}$	The MESD corresponding to operation time t
$\boldsymbol{C}$	The external random failure rate
$\lambda(t)$	The failure rate at time t
$Dist_{f,f+1}$	The distance between the DGA data samples g_f and g_{f+1}.
t_b'	The ESD that corresponds to the b^{th} data sample in period p.
$t_p^{''}$	The MESD that corresponds to period p

I. INTRODUCTION

As the core electrical equipment in a substation, the transformer plays a significant role in the substation reliability [1]. The failure rate is a significant parameter that measures the ability of a transformer to maintain its normal operations. Evaluating the failure rate of transformers accurately can lay a solid foundation for the reliability assessment of substations.

So far, the average failure rate model has been widely used in the substation reliability assessment. However, this model assumes the constant failure rates and cannot reflect the short-term fluctuation of the reliability level of an individual transformer [2], [3]. In reality, the failure rate of a transformer changes in real time according to its own operating condition [4], [5]. Therefore, it is necessary to establish a novel time-varying transformer failure rate model [6], [7].

Transformers faults can be divided into internal latent fault and external random fault [8]. Owing to the differences in the development modes, these two kinds of fault should be considered separately when modeling the failure rate model [5]. This paper focuses on the modeling of internal latent failure rate of transformers.

At present, the main approach for establishing the internal latent failure rate model is to analyze the internal physical and chemical processes of the transformers [1]. The degree of the internal latent failure is characterized by a distribution function, such as the Arrhenius-Weibull distribution. Accordingly, the trends of the failure rate can be evaluated. Recently, modeling the failure rate of transformers based on oil chromatographic data has become a popular method [9].

Oil chromatographic data, also known as dissolved gas analysis data (DGA data), include the content of the dissolved gases in a transformer's oil [9]. These dissolved gases are generated from the operational and fault events of the transformer [10]. Therefore, the content of the dissolved gases has a tight connection to the fault type and fault severity of the transformers [11]. By now, DGA data has been widely applied in the transformer faults diagnostics and remaining useful life (RUL) estimation aided by some powerful machine learning algorithms such as dynamic bayesian networks and temporal fault trees [13]–[17]. Reference [1] captures the qualitative causal relationship between dissolved gases and transformer faults, then a probabilistic diagnosis framework is proposed. Reference [18] established an improved transformer RUL prediction approach combining the data-driven thermal forecasting models and model-based lifetime experimental models. These studies excavate the value of DGA data promote its application. To the best of the authors' knowledge, DGA data also accurately characterize the development degree of internal latent faults and can be used to establish the models of internal latent failure rate of transformer [19]–[25]. Reference [19]–[21] choose the DGA data as the features and solve the multistate Markov process to obtain the time-varying failure rate model of transformers [5]. However, the transformer failure rates obtained by those models are still equal as long as the state and the state duration are the same although operating at different operation time. That is these models cannot describe the impact of maintenance on the transformer failure rate [22]. The reason for this defect is that those models ignored the actual operating conditions of the day when calculating the real-time failure rate of the transformer. To overcome this problem, some improved models introduced more covariates so as to describe the operating conditions and quantify the impact of maintenance when modeling the failure rates [22]. Reference [23] proposed a transformer failure rate model based on Proportional hazard model. In this model, the information of real-time hot spot temperatures (HST) is introduced to reflect the real-time operating conditions. Reference [24] further used some covariates such as the moisture content of insulating paper to modify the Markov process. Reference [25] selected the service life of the equipment as new features to establish the failure rate model.

Unfortunately, these external covariates such as service life, HST are hard to accurately describe the internal operating conditions of the transformer on daily basis. Besides, the acquisition of these covariates relies on complex accelerated aging experiments and it is difficult to obtain the actual operating data. An alternative method is to mine the daily DGA data, which is the most direct reflection of the daily operating conditions of the transformer. However, the sampling resolution of the DGA data in the above models is usually measured in months [26], making it difficult to capture the short-term changes in the DGA data.

In addition, the quality of DGA data is another important factor that affects the accuracy of the transformer failure diagnosis and failure rate modeling [27]. Nonetheless, a repairing model for missing DGA data is absent in the existing literature.

For the system reliability assessment considering component failures, a whole range of different approaches to the problem have been proposed. Reference [28] established an uncertainty-aware dynamic reliability analysis framework without the need of exact failure data of its components. Reference [29] presented a framework for prognostics-updated dynamic dependability assessment based on the online data analysis. Reference [30] used the FTA for performing reliability analysis of an management system. In this

model, the system reliability was also evaluated based on the given failure rates of the components. Reference [31] established a conceptual framework to incorporate complex basic events in hierarchically performed hazard origin and propagation studies (HiP-HOPS), which can guarantee the modelling capability on complex failures and the efficiency of Model-based safety analysis (MBSA) effectively. Reference [16] gave the detailed classification for safety models associated with machine learning (ML) and further proposed a novel approach based on ML and real-time operational data to reduce the difficulty in modeling and evaluation these complex safety-critical systems. In [32], the state analysis method is introduced for generic power system reliability assessment. Considering the requirements of the case studies, state analysis method is adopted in this paper for the substation reliability evaluation.

This paper established a novel modeling framework for transformer failure rate based on daily DGA data mining. The framework solves two main problems of repairing the missing DGA data and quantifying the impact of maintenance on failure rates. The short-term change characteristics of dissolved gases' content are also analyzed in this paper.

In summary, the main contributions of this paper are as follows:

1. A data repairing method derived from the improved KNN is proposed; thus, the quality of the modeling data is well guaranteed.

2 In order to eliminate the influence of unbalanced sample space on the repairing accuracy, the Genetic algorithm (GA) is introduced to construct a balanced sample space.

3. The parameter of ESD, which can reflect the real-time operational conditions when calculating the failure rate, is proposed to modify the Markov failure rate model. And then, a Multi-back-propagation neural network (M-BPNN) model is constructed to map the DGA data to ESD.

4. The MESD is further obtained by merging the ESD based on the analysis of the variation in the content of the dissolved gases. Thus, an analytical function of the failure rate with respect to the MESD is obtained based on a multistate Markov process.

The rest of this paper is organized as follows. The model for repairing the missing DGA data based on the GA and improved KNN algorithm is proposed in Section II. Section III established the transformer failure rate model proposed and its validation method. Case studies are given in Section IV. Finally, Section V concludes the paper.

II. REPAIRING MODEL FOR MISSING DGA DATA

The quality of DGA data is an important factor that affects the accuracy of the transformer failure rate model. The DGA data could be missing in the process of collection, transmission and storage [33]. It is necessary to repair the missing data. Most existing data repairing methods are based on time series analysis [34]. These methods have performed well when the missing data occurs under the normal state of transformers. However, the missing data often appears when the transformer works in abnormal/fault states [35]. Therefore, this paper proposes an improved KNN algorithm to repair the missing data. However, the number of DGA data samples in the abnormal or fault state is much smaller than that in the normal state due to the low occurrence probability of abnormal and fault states. It could result in an unbalanced data sample space and errors for the KNN algorithm. To avoid this problem, a balanced sample space of DGA data is first constructed through sample expansion based on a GA.

A. EXPANSION OF DGA DATA BASED ON GA

The genetic algorithm imitates the process of biological evolution and includes operations such as selection, crossover and mutation [36]. This paper proposes a sample expansion method that is inspired by the population generation method of the GA.

A target matrix $\boldsymbol{G}$ is created for storing the daily DGA data, including the content of each type of gas in N days. The n^{th} row of the N^*L matrix $\boldsymbol{G}$ represents the DGA data sample of n^{th} day. It can be expressed as:

$$\boldsymbol{g}_n = (g_{n1}, g_{n2}, \ldots\ldots g_{nL}) \tag{1}$$

where $1 \leq n \leq N$, and L is the total number of types of dissolved gases. The dissolved gases include H_2, C_2H_2, CH_4, C_2H_6, C_2, CO, CO_2 and total hydrocarbons.

The fault types for these N days could be determined through their content ratios of dissolved gases [37], [38]. The 12 fault types and their corresponding ranges for the ratios of dissolved gases are given in Table 1.

Accordingly, each DGA data sample in $\boldsymbol{G}$ can be allocated to one of these fault types. The number of days when the transformer in fault type k ($1 \leq k \leq 12$) is N_k. It has the following relationship with N:

$$N = \sum_{i=1}^{12} N_k \tag{2}$$

The DGA data belonging to type k constitutes an N_k^*L matrix $\boldsymbol{G}_k$:

$$\boldsymbol{G}_k = \begin{bmatrix} G_{1,1} & G_{1,2} & G_{1,L} \\ G_{2,1} & G_{2,2} & G_{2,L} \\ G_{Nk,1} & G_{Nk,2} & G_{Nk,L} \end{bmatrix} \tag{3}$$

Choose $\boldsymbol{G}_k$ as the initial parent population and randomly match the data samples contained in this population into pairs. A total of $N_k/2$ pairs of parent samples can be obtained.

Take a pair of parent samples, for example, $\boldsymbol{g}_{k,u}$ and $\boldsymbol{g}_{k,v}$, and perform the crossover operation on this pair, which is expressed as follows:

$$\begin{cases} \boldsymbol{g}_{s1} = 1.5\boldsymbol{g}_{k,u} - 0.5\boldsymbol{g}_{k,v} \\ \boldsymbol{g}_{s2} = -0.5\boldsymbol{g}_{k,u} + 1.5\boldsymbol{g}_{k,v} \\ \boldsymbol{g}_{s3} = (\boldsymbol{g}_{k,u} + \boldsymbol{g}_{k,v})/2 \end{cases} \tag{4}$$

where $\boldsymbol{g}_{s1}$, $\boldsymbol{g}_{s2}$ and $\boldsymbol{g}_{s3}$ are newly generated DGA data samples, named here child samples. Formula (4) presents a typical linear crossover in GA proposed by Wright [36].

TABLE 1. Rogers ratio method [35].

Fault types	$\frac{CH_4}{H_2}$	$\frac{C_2H_6}{CH_4}$	$\frac{C_2H_4}{C_2H_6}$	$\frac{C_2H_2}{C_2H_4}$
Normal	0.1 to 1	< 1	< 1	< 0.5
Partial discharge	< 0.1	< 1	< 1	< 0.5
Overheating <150 °C	> 1	< 1	< 1	< 0.5
Overheating 150 °C to 200 °C	> 1	< 1	< 1	< 0.5
Overheating 200 °C to 300 °C	0.1 to 1	> 1	< 1	< 0.5
Conductor overheating	0.1 to 1	< 1	1 to 3	< 0.5
Winding circulating currents	1 to 3	< 1	1 to 3	< 0.5
Core and Tank circulating	1 to 3	< 1	> 3	< 0.5
Flashover without power follow through	0.1 to 1	< 1	< 1	0.5 to 3
Arc with power follow through	0.1 to 1	< 1	> 1	> 0.5
Continuous sparking to floating potential	0.1 to 1	< 1	> 3	> 3
Partial discharge with tracking	< 0.1	< 1	< 1	> 0.5

In formula (4), the coefficient of $\boldsymbol{g}_{k,u}$ and $\boldsymbol{g}_{k,v}$ are set to 1.5 and -0.5 when generating $\boldsymbol{g}_{s1}$, to prevent $\boldsymbol{g}_{s1}$ from being too similar to $\boldsymbol{g}_{k,u}$ or $\boldsymbol{g}_{k,v}$. The same reason holds for $\boldsymbol{g}_{s2}$ and $\boldsymbol{g}_{s3}$. By keeping all the parent samples and child samples, the number of DGA data samples corresponding to fault type k will be $2.5N_k$.

To ensure the heterogeneity between the child samples and parent samples, an uncertain increment is added to the original values of each child sample. The step length variation method [39] is used to process those child samples obtained from (5).

$$\begin{aligned} \boldsymbol{g}'_s &= \boldsymbol{g}_s \pm 0.5\boldsymbol{g}_\Delta \cdot \Delta \\ \Delta &= \sum_{q=1}^{m} \frac{a(q)}{2^q} \end{aligned} \tag{5}$$

where $a(q)$ is equal to 1 with a probability of $1/m$. Otherwise, $a(q) = 0$. Here, m is set to 20 normally, and $\boldsymbol{g}_\Delta$ is a variation vector with the value range $[\boldsymbol{g}_{k,\min}, \boldsymbol{g}_{k,\max}]$.

Formula (5) uses an uncertain probability Δ multiplied by an uncertain step length $\boldsymbol{g}_\Delta$ to describe the uncertain increment for $\mathbf{g}_s$. Note that the step length $\boldsymbol{g}_\Delta$ is bounded by the value range of each dissolved gas in $\boldsymbol{G}_k$.

After the process of mutation for child samples, the validity of each child sample generated should be examined; that is, only the samples with values within [0.95 $\boldsymbol{g}_{k,\min}$, 1.05 $\boldsymbol{g}_{k,\max}$] can be added to $\boldsymbol{G}_k$. Otherwise, the samples should be removed from $\boldsymbol{G}_k$.

Perform the above operations iteratively until the sample size meets the requirement. At this point, the algorithm ends, and $\boldsymbol{G}_k$ is updated. The sample expanding to consider other fault types can be treated with the same method. Finally, the data matrixes for all fault types constitute the updated $\boldsymbol{G}$.

B. REPAIRING METHOD BASED ON IMPROVED KNN

The repairing method proposed in this paper is developed from the observation that the DGA data under the same abnormal/fault state have similar patterns and the missing data can be estimated according to these DGA data [34]. This paper uses the KNN algorithm [40] to find adjacent samples of the DGA data samples. The missing data is then repaired according to these adjacent samples.

Due to the missing data, the dimensionalities of the test samples and the training samples are not equal. Therefore, it is hard to choose the nearest samples by the Euclidean distance used in the traditional KNN. In this paper, the Pearson similarity [41] is selected to be the measure of the distance because it is not affected by the dimensionality mismatches between the samples.

To avoid the error in the distance calculation caused by the dimensionality mismatch and elementwise numerical difference, the DGA data samples should first be centralized. For the n^{th} DGA data sample in $\boldsymbol{G}$, $\boldsymbol{g}_n$, its centralization can be realized according to the following formula (6):

$$\boldsymbol{g}_n(l) = \boldsymbol{g}_n(l) - \mu_n \tag{6}$$

where μ_n is the mean value of all elements in $\boldsymbol{g}_n$.

For the testing sample $\boldsymbol{g}_{test}$ to be repaired, if the y^{th} element of $\boldsymbol{g}_{test}$, $\boldsymbol{g}_{test}(y)$, is missing, then the centralization formula should be described as follows:

$$\boldsymbol{g}_{test}(l) = \begin{cases} 0, & l = y \\ \boldsymbol{g}_{test}(l) - \mu'_{test}, & l \neq y \end{cases} \tag{7}$$

where μ'_{test} is the mean of all elements in $\boldsymbol{g}_{test}$ after setting the y^{th} element $\boldsymbol{g}_{test}$ (y) to 0.

Similarly, the other data samples in $\boldsymbol{G}$ can be processed by formulas (6) and (7). The advantage of (6) and (7) is that each centralized element becomes a value regardless of its original unit and magnitude.

The distance D_n between the centralized test sample, $\boldsymbol{g}_{test}$, and the n^{th} sample in $\boldsymbol{G}$, $\boldsymbol{g}_n$, can be calculated based on the Pearson similarity:

$$D_n = \frac{\mathrm{cov}(\boldsymbol{g}_{test}, \boldsymbol{g}_n)}{\sigma_{\boldsymbol{g}_{test}} * \sigma_{\boldsymbol{g}_n}} \tag{8}$$

where cov represents the covariance between $\boldsymbol{g}_{test}$ and $\boldsymbol{g}_n$. σ means the standard deviation of the vector. A smaller D_n implies a closer distance between $\boldsymbol{g}_{test}$ and $\boldsymbol{g}_n$ and more similarity between them.

According to formulas (8), the distance between $\boldsymbol{g}_{test}$ and each sample in the updated $\boldsymbol{G}$ can be calculated. Select r nearest training samples and record these samples as $\boldsymbol{G}_{train}$. The weighted average value in $\boldsymbol{G}_{train}$ is calculated as the

repaired value of the missing data.

$$\hat{\boldsymbol{g}}_{test}(y) = \left(\sum_{w=1}^{r} \eta_{train,w} \cdot \boldsymbol{g}_{train,w}(y) \right) / \boldsymbol{r} \tag{9}$$

where $\eta_{\text{train},w}$ is the weight of each data sample in $\boldsymbol{G}_{train}$, and $1 \leq w \leq r$. Details of the weight calculation method can be found in [42]. A smaller distance in (8) results in a larger weight in the estimation of the missing data, which indicates its greater similarity to the test sample $\boldsymbol{g}_{test}$.

III. REFINED FAILURE RATE MODEL OF THE TRANSFORMER

A. MULTISTATE MARKOV PROCESS OF A TRANSFORMER

According to IEEE Std C57.104 [35], the operational states of transformers can be divided by their DGA data into four states: normal state, abnormal state, warning state and failure state.

As a typical repairable component, the operating state of the transformer generally transfers between the above four states. Specifically, the transformer may change from a normal state to an abnormal/warning state with an increasing failure rate during the operation. It can also return to the normal state from the abnormal/warning state through maintenance. The multistate Markov process of the transformer is illustrated in Fig 1.

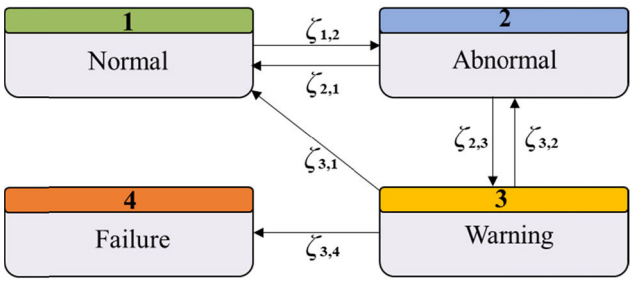


FIGURE 1. Multistate Markov model of the transformer.

In Fig 1, $\zeta_{i,j}$ represents the transition rate from state i to state j, which can be calculated as follows:

$$\zeta_{i,j} = \frac{count_{i,j}}{T_i} \tag{10}$$

where $count_{i,j}$ is the times that the transformer transfer from state i to the j. T_i is the total duration that the tranformer stays in state i.

Formula (10) can be used to calculate the transition rate between states with a direct transition path. Based on the relationships in Fig 1, the transition rates between all states form a Markov time-varying state transition matrix $\mathbf{A}$, as follows:

$$\mathbf{A} = \begin{bmatrix} -\zeta_{1,2} & \zeta_{2,1} & \zeta_{3,1} & 0 \\ \zeta_{1,2} & -(\zeta_{2,1}+\zeta_{2,3}) & \zeta_{3,2} & 0 \\ 0 & \zeta_{2,3} & -(\zeta_{3,1}+\zeta_{3,2}+\zeta_{3,4}) & 0 \\ 0 & 0 & \zeta_{3,4} & 0 \end{bmatrix} \tag{11}$$

where $\mathbf{A}_{i,j} = 0$ indicates that there is no direct transition path between state i and state j.

Assume that the probability of the transformer operating in various states at time t is $\mathbf{P}(t) = [\mathrm{P}_1(t), \mathrm{P}_2(t), \mathrm{P}_3(t), \mathrm{P}_4(t)]$. Since the Markov process established in this paper is continuous in operation time, this section uses the Kolmogorov differential equation [43] to calculate the $\mathbf{P}(t)$. The Kolmogorov differential equation that corresponds to equation (11) can be expressed as follows:

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{A}\mathbf{P}(t) \tag{12}$$

where the left side of equation (12) is the derivative of $\mathrm{P}(t)$ with respect to the operation time t.

The probability of the transformer operating in various states $\mathbf{P}(t)$ can be obtained by solving (13):

$$\mathbf{P}(t) = e^{\mathbf{A}t}\mathbf{P}(0) \tag{13}$$

where

$$e^{\mathbf{A}t} = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} + \alpha_2\mathbf{A}^2 + \alpha_3\mathbf{A}^3 \tag{14}$$

When operating in the failure state (i.e. state 4), the transformer needs to be out of service and protected immediately. Therefore, the failure state is considered as the ending state of the internal latent fault of the transformer. Consequently, the internal latent failure rate at time t can be considered as the probability that the transformer transfers from the operating state to the failure state at time t, that is $\mathbf{P}_4(t)$.

It should be noted that P(0) is the initial conditions of the differential equation (12). It denotes the initial probability that the transformer operates in each state at time t. The value of the i^{th} element is set to 1 if the transformer is in state i at time t. Otherwise, it is set to 0. For example, $\mathrm{P}(0) = [1\ 0\ 0\ 0]$ is the initial conditions for solving the analytical expressions of the time-varying internal latent failure rate when the transformer is operating in normal state at time t. Similarly, $\mathrm{P}(0) = [0\ 1\ 0\ 0]$ and $\mathrm{P}(0) = [0\ 0\ 1\ 0]$ are the initial conditions for obtaining the failure rate expressions when the transformer operates in abnormal state and warning state at time t respectively. Additionally, α_0, α_1, α_2, and α_3 are exponential functions determined by the eigenvalues of matrix A. These functions can be solved by the Hamilton-Hailey method [24], [46].

Define the SD t_i as the duration that the transformer has stayed in state i at time t, and $1 \leq t_i \leq T_i$. From equations (10)-(14), the analytical expressions of the failure rate $\lambda_i(t)$ for the latent failure of the transformer can be expressed by equation. As shown in formula (15), as shown at the bottom of the next page, the time-varying internal latent failure rate of transformer at time t under different operating states has different analytical expressions.

When calculating the latent failure rate of the transformer at operation time t, the variable used should be the SD that corresponds to t, t_i, rather than t. A constant C is chosen to represent the external random failure rate of the transformer [27]. The total failure rate at time t can be obtained

as follows:

$$\lambda(t) = \lambda_i(t_i) + C \tag{16}$$

The above failure rate model is called the Markov failure rate model (Markov model). An illustrative example is given in Fig 2, which shows the change in the failure rate with the operation time, according to the Markov failure rate model.

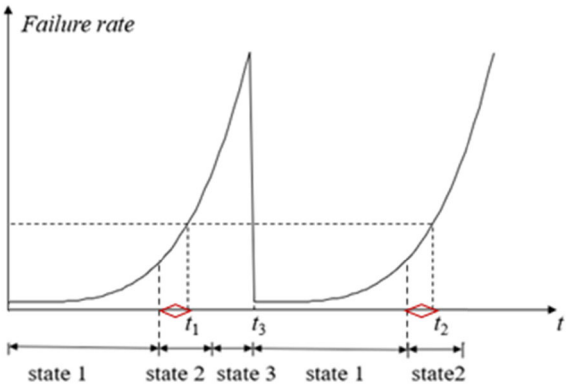


FIGURE 2. Failure rate based on Markov failure rate model.

Suppose that the maintenance is performed on the transformer at time t_3 in a warning state, which results in a significant decrease in the failure rate at time t_3. The SDs in state 2 at time t_1 and t_2 are equal, and the failure rates at time t_1 and t_2 calculated by formula (16) and are also the same. The reason for equal SDs at times t_1 and t_2 is that the Markov model always assumes that the transformer can return to the initial state after the maintenance is performed at time t_3 [22].

Obviously, the values of SD t_i used in the Markov model are determined only by the current state and the corresponding state duration at time t. As along as the transformer is under the same state and has the same state duration, the values of SD remain the same. However, the maintenance might not return the transformer to its original condition [22]. The SD at time t_2 after maintenance should be no longer the same as that at time t_1. Therefore, the Markov failure rate model tends to have an optimistic view of the impact of maintenance by ignoring the actual operating condition after maintenance.

The next section introduces a modification of SD combined with the actual operating condition after the maintenance to effectively describe the impact of the maintenance.

B. REFINED MODEL OF THE TRANSFORMER FAILURE RATE

1) MODELING OF ESD

To quantify the impact of the maintenance, the SD t_i in formula (16) is modified to its corresponding ESD t_i'. Compared with SD, ESD is able to simultaneously consider the current state, the corresponding state duration and the operating condition at time t. The DGA data should be accounted for in the calculation of ESD because it can efficiently describe the operating condition of the transformers.

Fig 3 illustrates the modification of SD to ESD. The DGA data that corresponds to time t is used to obtain t_i'. The quantitative relationship with non-linear correlation between the ESD and the DGA data is modeled by BPNN in this paper. Note that the other neural network models like Radial Basis Function neural network (RBF) can also be used for modeling [44], [45].

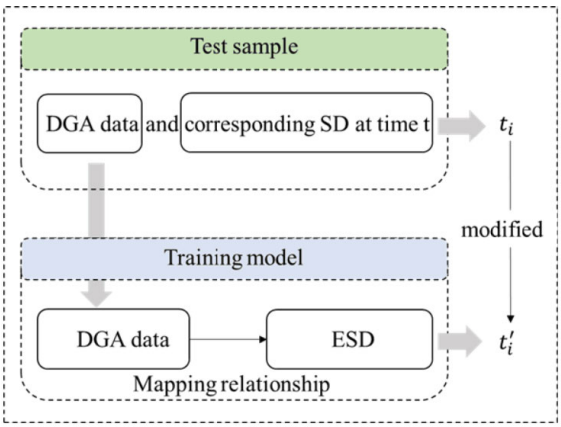


FIGURE 3. The process of modifying SD to ESD.

BPNN is a multilayer feedforward network that is trained by an error back-propagation algorithm. This network is able to describe nonlinear relationships and has a good learning ability [46]. In addition, BPNN has been proved that it has a reasonable explainability [47].The network structure of the BPNN is shown in Fig 4:

As seen from Fig 4, the BPNN includes an input layer, hidden layer and output layer. Formula (17) describes the feature delivered from the input layer to the hidden layer:

$$\boldsymbol{h}_1 = \sigma_1(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1) \tag{17}$$

where $\boldsymbol{h}_1$ is the feature vector for the first hidden layer. X is the input vector. $\boldsymbol{W}_1$ is the weight coefficient matrix of

$$\begin{aligned}
\lambda_1(t) &= \alpha_3\zeta_{1,2}\zeta_{2,3}\zeta_{3,4} && \text{if P(0)} = [1\ 0\ 0\ 0]\\
\lambda_2(t) &= \zeta_{2,3}\zeta_{3,4}[\alpha_2 - \alpha_3(\zeta_{2,3} + \zeta_{3,4} + \zeta_{3,2} + \zeta_{3,1} + \zeta_{2,1})] && \text{if P(0)} = [0\ 1\ 0\ 0]\\
\lambda_3(t) &= \alpha_1\zeta_{3,4} - \alpha_2\zeta_{3,4}(\zeta_{3,4} + \zeta_{3,2} + \zeta_{3,1}) + \alpha_3\zeta_{3,4}[\zeta_{2,3}\zeta_{3,2} + (\zeta_{3,4} + \zeta_{3,2} + \zeta_{3,1})^2] && \text{if P(0)} = [0\ 0\ 1\ 0]
\end{aligned} \tag{15}$$

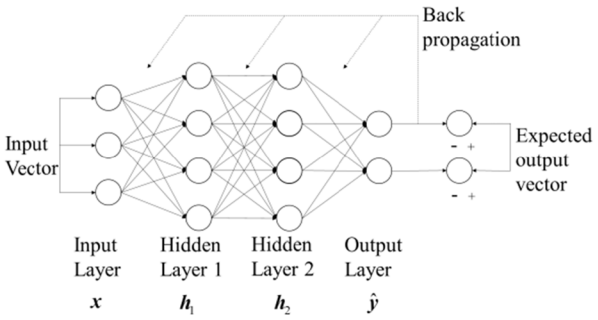


FIGURE 4. The structure of the BPNN.

the hidden layer. $\boldsymbol{B}_1$ is the bias vector. Σ_1 is the activation function.

A sigmoid function is used as the activation function, which is shown in(18). The relationships between the hidden layers can be expressed as (19) for the multilayer neural networks.

$$\sigma_1(z) = \frac{1}{1+e^{-z}} \tag{18}$$

$$\boldsymbol{h}_e = \sigma_e(\boldsymbol{W}_e\boldsymbol{h}_{e-1} + \boldsymbol{b}_e) \tag{19}$$

where e is the index of the hidden layers. $\boldsymbol{W}_e$ is the weight coefficient matrix of the e^{th} hidden layer. $\boldsymbol{B}_e$ is the bias vector of the e^{th} hidden layer.

Assuming that there are X hidden layers in total, the predicted output $\hat{\boldsymbol{y}}$ can be described as:

$$\hat{\boldsymbol{y}} = \sigma_e(\boldsymbol{W}_y\boldsymbol{h}_X + \boldsymbol{b}_y) \tag{20}$$

Since the calculation of the output layer is exactly the same as the hidden layer, the output layer can be regarded as the $X + 1^{\text{th}}$ hidden layer. The error $\boldsymbol{\delta}_y$ between the predicted output and the actual output can be calculated as

$$\boldsymbol{\delta}_y = \hat{\boldsymbol{y}} - \boldsymbol{y} \tag{21}$$

Combining (20) and (21), the back derivation of the error of the e^{th} hidden layer can be expressed as:

$$\boldsymbol{\delta}_e = \boldsymbol{W}_e^{\mathrm{T}}\boldsymbol{\delta}_{e+1}. \times (\boldsymbol{h}_e \cdot (\boldsymbol{1} - \boldsymbol{h}_e)) \tag{22}$$

where the weight coefficient matrix of the e^{th} hidden layer can be updated as:

$$\boldsymbol{W}_e = \boldsymbol{W}_e + \boldsymbol{\delta}_{e+1} \cdot \boldsymbol{h}_e^T \tag{23}$$

The above is an iteration in the training process of the BPNN. The learning parameters of the BPNN can be tuned after several iterations to obtain a more accurate BPNN model.

The ESD is equal to its SD at any time before the maintenance of the transformer. Nonetheless, this circumstance is not the case after maintenance. To correctly represent the relationships between the ESD and DGA data, only the DGA data and its corresponding ESD (equal to SD) before maintenance should be selected as the training samples of the BPNN.

The size of a single BPNN model could be too large if it is designed to estimate the ESD under all possible operating states. It reduces the computational efficiency of the model. To handle this problem, the Multi-BP neural network (M-BPNN) is trained. It contains four BPNNs, one for each operating state. The structure of the M-BPNN is shown in Fig 5.

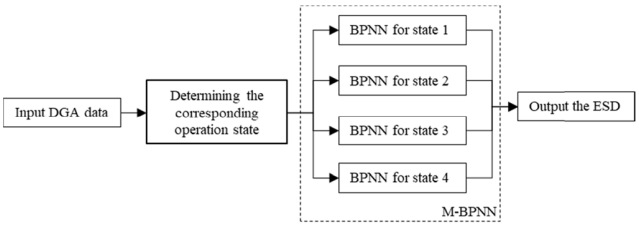


FIGURE 5. The structure of the M-BPNN.

The steps of estimating ESD by M-BPNN are as follows:

Training:

STEP1: Select the DGA data $\boldsymbol{G}_{init}$ in each state before maintenance as the input. The corresponding SD $\boldsymbol{T}_{init}$ are used as the output. These data together form the training samples of M-BPNN.

STEP2: Tune the learning parameters. The training samples are inputted into M-BPNN for learning, in such a way that the mapping relationship between the input and output can be constructed.

Testing:

STEP3: Choose the DGA data $\boldsymbol{G}_{aft}$ after maintenance and determine the operating state of the transformer. Next, input the data into the different BPNN models according to the corresponding states. Finally, the ESD $\boldsymbol{T}'_{aft}$ corresponding to the DGA data $\boldsymbol{G}_{aft}$ can be estimated from the BPNN model.

2) MODELING OF THE MESD

Minor changes in the dissolved gases within a short period of time are commonly seen in the DGA data. Such changes do not necessarily indicate changes in the failure rates. If the changes in the dissolved gases between any two days are less than a given threshold in this period, then the ESD and failure rate of the transformer can be considered to be the same. Therefore, the ESD in this period could be merged into the same value. To achieve this goal, this paper proposes a clustering algorithm based on the distance matrix between the time series of the DGA data. By this method, the time series of the DGA data can be partitioned into different sequential periods.

The first step of the clustering algorithm proposed is to create a distance matrix, which can be expressed as:

$$\mathbf{Dist} = \begin{bmatrix} Dist_{1,1} & Dist_{1,2} & Dist_{1,F} \\ Dist_{2,1} & Dist_{2,2} & Dist_{2,F} \\ Dist_{F,1} & Dist_{F,2} & Dist_{F,F} \end{bmatrix} \tag{24}$$

where F is the number of DGA data samples of a particular transformer. Each element in matrix **Dist** adopts the standard

Euclidean distance:

$$Dist_{f,f+1} = \sum_{i=1}^{L} \sqrt{\frac{\left[\boldsymbol{g}_f(i) - \boldsymbol{g}_{f+1}(i)\right]^2}{s_i}} \tag{25}$$

where $1 \leq f < F$, and s_i is the standard deviation of each element. The advantage of using the standard Euclidean distance is that this formula can avoid the influence of dimensionality differences in the distance calculation.

Next, the procedure of the clustering algorithm proposed in this section is shown in Table 2.

TABLE 2. The algorithm flow for obtaining MESD.

```
Initialize f=1, z =1, input the threshold σ
While z ≤F
    χ=z
For m=f to z, n=f to z
    If any Dist (m,n) < σ
        z= z+1
    ELSE
        cluster samples from f to z-1 to a period
        f = z
    End If
End For
```

Assuming that time t is clustered into period p, the MESD t_i'' corresponding to time t can be obtained by averaging the ESD of each day in period p:

$$t_p'' = \frac{\sum_{b=1}^{q} t_b'}{q} \tag{26}$$

where q is the number of data samples in period p. Here, t_b' is the ESD corresponding to the b^{th} data sample in this period.

The total failure rate at time t can be modified from (15) and (16) by replacing the parameter ESD t_i' with MESD t_i'' (t_p''):

$$\lambda(t) = \lambda_i(t_i'') + C \tag{27}$$

Combining (26) and (27), the analytical failure rate model proposed in this paper is established.

C. VERIFICATION METHOD OF THE FAILURE RATE MODEL

Although the failure rate of an individual transformer is affected by factors such as the region, operation time, and statistical methods, the transformers' failure rates have similar changing trends over time [48], [49]. The trend in the individual transformer failure rates should be consistent with the statistical data of the failure rate that is observed from multiple transformers. This paper proposes a verification method of the transformer failure rate model by comparing the similarity in the changing trend of the failure rate over time.

Suppose that $\lambda_{pro} = \{\lambda_{pro,1}, \lambda_{pro,2}, \lambda_{pro,3}, \ldots, \lambda_{pro,\tau}\}$ is the time series of the time-varying failure rates of an individual transformer obtained from the failure rate model. Additionally, $\lambda_{sta} = \{\lambda_{sta,1}, \lambda_{sta,2}, \lambda_{sta,3}, \ldots, \lambda_{sta,\upsilon}\}$ is the time the series of failure rates from the statistical data of multiple transformers. The similarity between the overall failure rate and the individual transformer failure rate can be characterized by the parameter RMSD as follows:

$$RMSD = DTW(\tau, \upsilon) \tag{28}$$

where τ *and* υ *are the* number of points in the sequence λ_{pro} and λ_{sta} *respectively.* $DTW(\tau, \upsilon)$ represents the distance between λ_{pro} and λ_{sta} with dynamic time regularization [50]:

$$\begin{aligned} DTW(o, z) = {} & DTW(o, z) \\ & + \min\{DTW(o-1, z), DTW(o, z-1) \\ & + DTW(o-1, z-1)\} \end{aligned} \tag{29}$$

where $2 \leq o \leq \tau$, $2 \leq z \leq \upsilon$. $DTW(o, z)$ represents the Euclidean distance between the o^{th} point of the sequence λ_{pro} and the z^{th} point of the sequence λ_{sta}. A smaller RMSD value indicates greater similarity between the trends of the two failure rate time series. As a result, the change in the failure rate of an individual transformer calculated by the failure rate model is more consistent with that obtained from the statistical data; i.e., the model is more accurate.

To better explain how different methods mentioned above join together to established the novel transformer failure rate model above, the flowchart of the transformer failure rate modeling process proposed is given in Fig 6. Note that the proposed validation method in Section III-C does not need to be considered in the flowchart description below.

As shown in Fig 6 above, the failure rate modeling proposed can be divided into three main parts including repairing the missing DGA data, modifying the model parameters and calculating the transformer failure rate. The detailed explanation of part 1, part 2 and part 3 can be found in the section II and section III.

IV. CASE STUDY

This paper uses the DGA data of multiple 110 kV oil-immersed power transformers from an electrical company in southwest China in 2017 to verify the proposed failure rate model of a transformer. The model is applied to two transformers of a typical substation, as shown in Fig 7. The two 220/110 kV transformers, T1 and T2, use the same DGA data. The external failure rate of transformer C is set to 0.005. The constant failure rate λ_c of the two transformers is set to 0.2. The reliability data and load data of this substation can be found in [51]. The failure rates from the statistical data of multiple transformers can be found in [52].

In this section, study 1 verifies the effectiveness of the repairing method for the missing DGA data. In Study 2, the failure rate of the transformer without maintenance was calculated. Study 3 calculates the failure rate of the transformer while considering maintenance first. The results are then applied in the reliability assessment of the substation by the state analysis method [53]. Study 4 verifies the accuracy

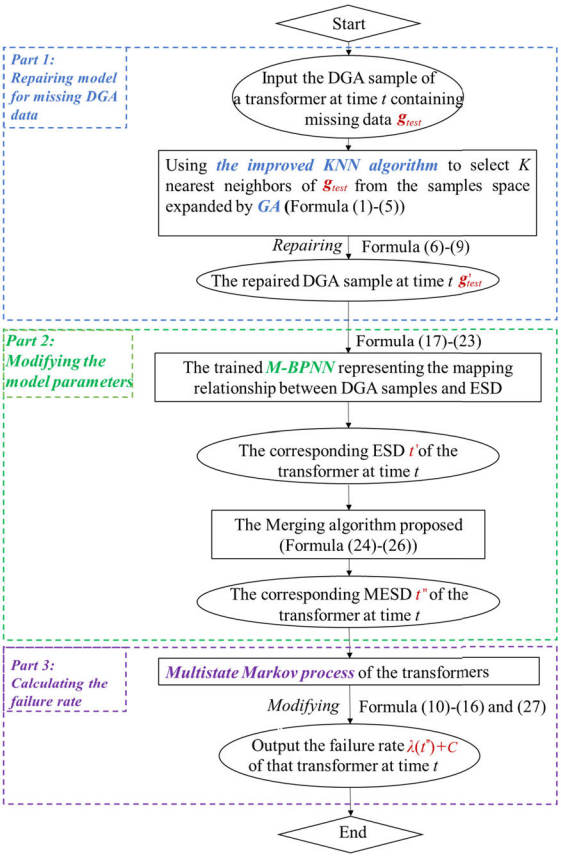


FIGURE 6. The flowchart of transformer failure rate modelling.

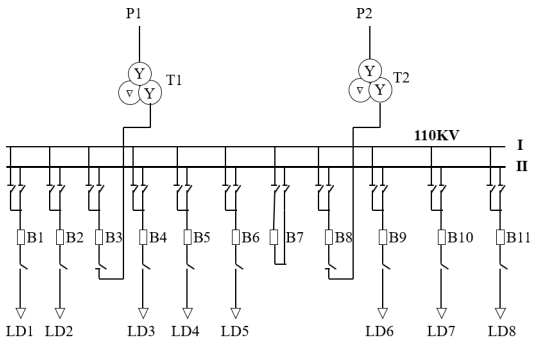


FIGURE 7. A typical substation model.

of the proposed model. All simulations are implemented on a personal computer with 3.0 GHz Intel Core i5-9500 and 8 G RAM.

A. STUDY 1. REPAIRING OF MISSING DGA DATA

To verify the effect of the proposed data repairing method, this case created multiple sets of DGA data. The created data sets represent different fault types of the transformers through random deletion. The RV (Repaired Value) obtained by the KNN (non-time series) and Arima (time series) methods are both given in this case.

The samples with missing data are shown in Table 3. The unit of each gas content is μL/L. '$*$' represents the missing DGA data in each sample. Table 4 gives the corresponding RV, RE (Relative Error) and FTJ (Fault Types Judged) obtained by the three methods.

It can be seen from Table 4 that the three methods all have low relative errors ($<10\%$) when the data missing occurs in a normal state (sample 1 and sample 2). At the same time, all three methods make the correct judgment of the fault types based the repaired DGA data. On closer examination, the repaired value based on the proposed method keeps the relative error less than 1%, which means that it is better than the other two methods when the data missing occurs in a normal state.

When the missing data occurs in abnormal or fault states (sample 3 and sample 4), the proposed method still shows good performance. The relative error is still less than 15%, which implies its accuracy for the fault type judgment. However, the relative errors obtained by the KNN and Arima methods are larger than 45%, which results in mistakes in the fault type judgment. The results confirm that the repairing method proposed is suitable for repairing the missing data that occurs in abnormal or fault states compared with the traditional KNN method and Arima method.

Table 5 gives the maximum, minimum and the average values of RE obtained from the three methods.

It can be seen from Table 5 that the proposed repairing method always has a lower RE regardless of the operating states of the transformers. The average RE is reduced by 23.69% and 17.07% compared with KNN and Arima, respectively. Therefore, the method can provide accurate repaired data for the failure rate model.

B. STUDY 2. SHORT-TERM FAILURE RATES WITHOUT MAINTENANCE

In this case, the daily DGA data of a transformer from January 22, 2010 to February 15, 2010 is used for the failure rate modeling. The transformer had never undergone maintenance before May 24, 2010. Fig 8 shows the SD, ESD and MESD that correspond to the daily DGA data during this period.

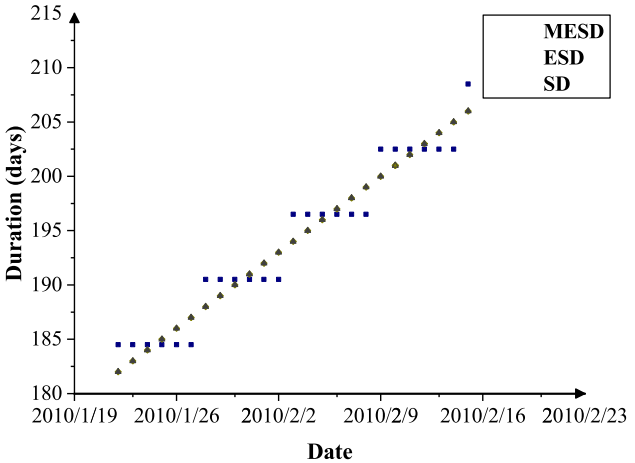


FIGURE 8. The values of SD, ESD and MESD.

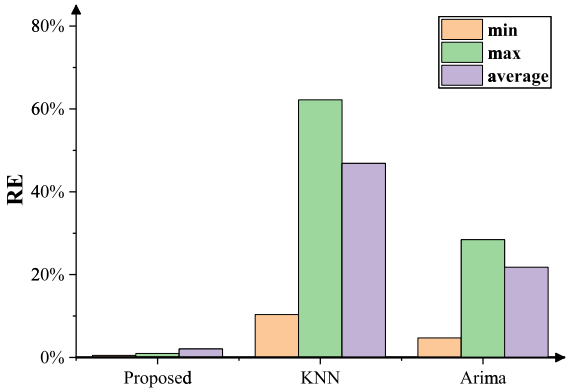
TABLE 3. The samples with data missing.

Sample	Monitoring Time	H_2	C_2H_2	CH_4	C_2H_6	C_2H_4	CO	CO_2	Total Hydrocarbons
1	2015-11-08	*	0	2.740	0.430	0.420	533.48	1929.79	3.60
2	2015-11-04	*	0	0.740	0.520	0.390	1.93	279.27	178.66
3	2014-01-05	14.310	0.490	3.690	0.790	*	307.11	595.06	4.68
4	2011-04-01	32.490	0.300	3.230	0.560	*	286.43	726.63	4.58

TABLE 4. The samples with data missing.

Sam-ple	True Value	True Fault Type	Proposed			KNN			Arima		
			RV	RE	FTJ	RV	RE	FTJ	RV	RE	FTJ
1	19.530	Type 1	19.632	0.52%	Type 1	20.450	4.71%	Type 1	19.323	1.06%	Type 1
2	8.380	Type 1	8.445	0.78%	Type 1	8.300	0.95%	Type 1	8.839	6.19%	Type 1
3	0.960	Type 10	1.030	7.29%	Type 10	1.400	45.83%	Type 6	0.510	46.88%	Type 9
4	1.350	Type 5	1.490	10.37%	Type 5	0.510	62.22%	Type 1	0.903	33.11%	Type 9

TABLE 5. The comparison of RE.



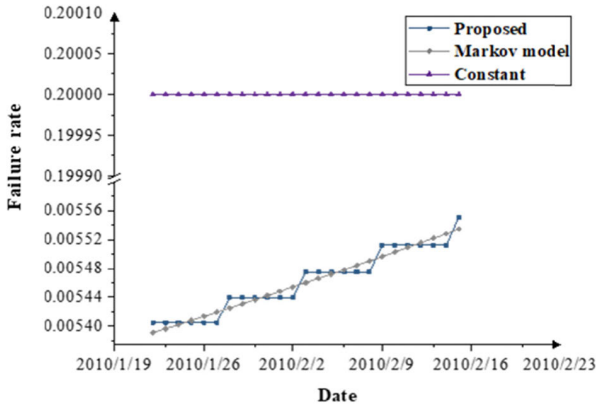


FIGURE 9. Failure rates from 2010-1-22 to 2020-2-15.

Fig 8 shows that the SD of the transformer is equal to its ESD for each day from 2010-1-22 to 2020-2-15, because no maintenance is performed on the transformer before 2020-5-24. The reasons are detailed in section III. The 25 days are partitioned into 5 periods, and the MESD for all periods take the same value. In addition, the EMSD curve in the stepped increase in Fig 8 suggests that the severity of the internal latent faults also increases with time.

Substituting the MESD into the failure rate analysis formula (30), the daily failure rate from 2010-1-22 to 2010-2-15 can be obtained, as shown in Fig 9. For comparison, the failure rates obtained by the Markov failure rate model and the constant failure rate model are also shown in this figure.

According to Fig 9, the failure rate given by the proposed model is maintained at a constant value in a short time period, even though the DGA data are changing every day. For example, the failure rates from 2010-1-22 to 2010-1-27 maintain the same value. The reason is that the model proposed considers short-term changes in the dissolved gases to be a normal phenomenon. Overall, the failure rate curve still shows a trend of stepped increase, whereas that obtained from the Markov model is continuously increasing. Moreover, the failure rate curve obtained by the proposed model fluctuates around the Markov failure rate curve.

In fact, the transformer had just been put into service from 2010-1-22 to 2020-2-15, and the failure rate will not reach 0.2 in this period. The failure rate only increases from 0.00541 to 0.00555, and the internal latent failure rate only increases from 0.00041 to 0.00055. The reason is that the transformer is in a normal state during 2010-1-22 to 2010-2-15. The main cause of transformer failures should be attributed to external random failures rather than internal latent failures.

C. STUDY 3. SHORT-TERM FAILURE RATES WITH MAINTENANCE

To characterize the impact of maintenance on the transformer failure rate, study 3 selects the DGA data from 2013-11-3 to 2013-12-1 as the modeling data for transformer T1 and T2. The maintenance of the transformer is scheduled on 2013-11-12.

Table 6 shows part of the data from 2013-11-3 to 2013-12-1. It can be seen from Table 6 that the ESD will

TABLE 6. Data from 2013-11-8 to 2013-11-15.

Date	SD	ESD	MESD	State
2013-11-08	140	140	135.999	3
2013-11-09	141	141	135.999	3
2013-11-10	142	142	142.5	3
2013-11-11	143	143	142.5	3
2013-11-12	1	168.8864	171.890	1
2013-11-13	2	169.8877	171.890	1
2013-11-14	3	170.889	171.890	1
2013-11-15	4	171.890	171.890	1

not be equal to the SD for each day after the maintenance. This table also shows that the MESD in 2013-11-11 is 142.5 days in state 3, and the MESD in 2013-11-12 is 171.890 days in state 1, after maintenance. However, the Markov model considers that the transformer is able to return to state 1 of the first day after maintenance.

The failure rates from 2013-11-3 to 2013-12-1 based on the proposed model are displayed in Fig 10. They are compared with the results obtained by the Markov failure rate model.

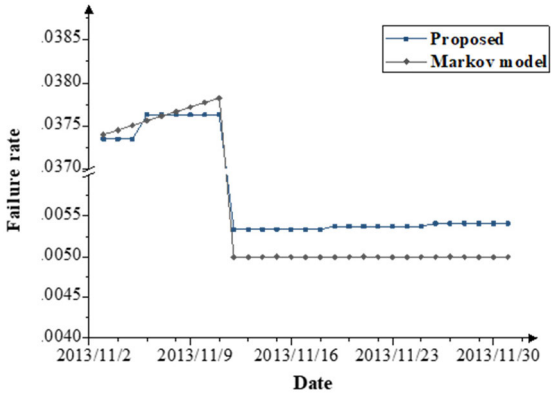


FIGURE 10. Failure rate curve of the transformer after maintenance.

This case focuses on the analysis of the impact of maintenance on the transformer failure rate. The increment of the failure rate caused by maintenance is the difference between the failure rate obtained by the proposed method and the Markov model. The increments recorded in the period from 2013-11-12 to 2013-12-1 are shown in Table 7. Note that the unit for the values in the table below is 10^{-2}.

After maintenance, the failure rate obtained by the Markov failure rate model is significantly smaller than that obtained by the proposed model. The reason is that the actual effect of the maintenance is overestimated in the Markov model. As can be seen in Table 7, the failure rates obtained from Markov model are keep in 0.005 from 2013-11-12 to 2013-12-1 because the internal latent failure rates calculated are ignored as they are very close to 0. Although maintenance indeed can reduce the failure rate, the transformer cannot be repaired to the initial operating state. The results of the reliability assessment for the substation can be calculated based on the two failure rate models. This paper selects Loss of load probability (LOLP) and Expected energy not supplied (EENS) as the indices for the reliability assessment [54].

TABLE 7. Increments in the failure rate.

Date	Proposed model	Markov model	Increment
2013-11-12	0.534	0.500	0.034
2013-11-13	0.534	0.500	0.034
2013-11-14	0.534	0.500	0.034
2013-11-15	0.534	0.500	0.034
2013-11-16	0.534	0.500	0.034
2013-11-17	0.534	0.500	0.034
2013-11-18	0.534	0.500	0.034
2013-11-19	0.537	0.500	0.037
2013-11-20	0.537	0.500	0.037
2013-11-21	0.537	0.500	0.037
2013-11-22	0.537	0.500	0.037
2013-11-23	0.537	0.500	0.037
2013-11-24	0.537	0.500	0.037
2013-11-25	0.537	0.500	0.037
2013-11-26	0.541	0.500	0.041
2013-11-27	0.541	0.500	0.041
2013-11-28	0.541	0.500	0.041
2013-11-29	0.541	0.500	0.041
2013-11-30	0.541	0.500	0.041
2013-12-01	0.541	0.500	0.041

The LOLP and EENS are shown in Fig 11. Before the maintenance of the transformer, the results of the assessments from the two models are close to each other. However, the LOLP and EENS calculated based on the proposed model are significantly higher than the results by the Markov failure rate model after maintenance. The results show that the maintenance operation has a large influence on the substation reliability assessment. If the actual impact of the maintenance is not considered in the reliability assessment, the assessment results will be overoptimistic.

D. STUDY 4. VERIFICATION OF THE FAILURE RATE MODEL

In this case, the *RMSD* between the failure rate curve obtained by the proposed model and the statistical failure rate curve is used to measure the accuracy of the proposed model. Details on the *RMSD* can be found in section III.

The daily DGA data used in the proposed failure rate mode is collected from the transformers whose operation time range is from 7.4 to 11.3 years. In total, 7 sampling points are selected, and the intervals between the sampling points are uneven. The average failure rates at the sampling points are then calculated separately.

The failure rate curves obtained from the proposed model and the statistical data are shown in Fig 12. For comparison, Fig 12 also gives the curve obtained from the Markov model. The failure rate curve in blue is the overall failure rate curve based on the statistical data. The curve in orange is obtained from the proposed model. The gray curve is the transformer failure rate curve based on the Markov failure rate model.

To verify the effectiveness of the different failure rate models, the RMSD between the orange curve and the blue curve, and that between the gray curve and the blue curve, are calculated. The results are shown in Table 8.

Table 8 shows that the *RMSD* of the proposed model is smaller than that of the Markov failure rate model and is

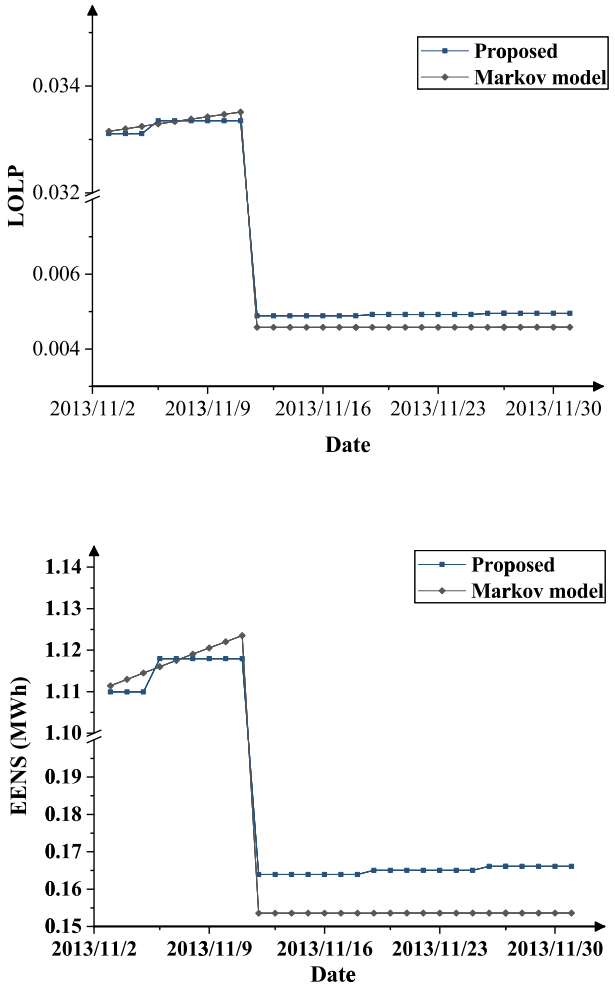


FIGURE 11. Reliability assessment of the substation.

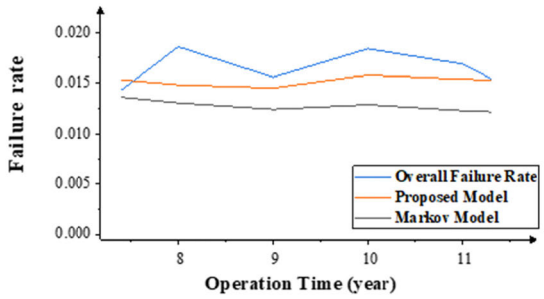


FIGURE 12. Comparison between three failure rate curves.

TABLE 8. The RMSD of the two models.

	Proposed Model	Markov Model
RMSD	9.4	23.7

decreased by 60.34%. The trends in the failure rate curve proposed are more consistent with the overall failure rate curve. This finding shows that the effectiveness of the model is indeed improved compared with the Markov failure rate model. The results also show that maintenance has a large influence on the failure rate, and its impact cannot be ignored.

V. CONCLUSION

This paper proposes a new approach for transformer time-varying failure rate modeling by mining the DGA data on a short-time scale. The model obtained by the approach builds a connection between the DGA data and the time-varying failure rate of the transformer, which quantifies the impact of maintenance on the failure rate.

Case studies show that DGA data can effectively describe the operating conditions of transformers and the impact of maintenance on the transformer failure rate. After maintenance, the failure rate will decrease, but it is still higher than that obtained from the Markov failure rate model. On the other hand, the LOLP and EENS calculated based on the proposed model are higher than that obtained from the Markov failure rate model. The results show that the trend of the failure rate curve obtained from the model proposed is more in line with the statistical failure rate curve. Therefore, the proposed model can better characterize the actual transformer failure rate considering the impact of maintenance. Furthermore, the proposed repairing method for missing data can effectively improve the accuracy of the modeling data.

With the installation and operation of a large amount of transformer monitoring equipment, massive quantities monitoring data will be used in the modeling of transformer failure rates. Determining how to effectively identify the erroneous monitoring data and improve the accuracy of the transformer failure rate model will be an important research direction in the future.

ACKNOWLEDGMENT

The authors would like to thank Prof. Baoping Cai from the China University of Petroleum (East China), China, for the useful comments and guidance of this research.

REFERENCES

[1] K. Xie, K. Cao, and D. C. Yu, "Reliability evaluation of electrical distribution networks containing multiple overhead feeders on a same tower," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2518–2525, Nov. 2011.

[2] J. H. Jurgensen, L. Nordström, and P. Hilber, "Estimation of individual failure rates for power system components based on risk functions," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1599–1607, Aug. 2019.

[3] S. L. Lima, O. R. Saavedra, and V. Miranda, "A two-level framework to fault diagnosis and decision making for power transformers," *IEEE Trans. Power Del.*, vol. 30, no. 1, pp. 497–504, Feb. 2015.

[4] Q. Li "Research on prediction model of insulation failure rate of power transformer considering real-time aging state," in *Proc. IEEE 3rd Conf. Energy Internet Energy Syst. Integr. (EI2)*, Nov. 2019, pp. 800–805.

[5] Y. Liang, K. Li, L. Niu, and J. Zhao, "EWDGA and Markov process based failure rate estimation of transformer internal latent fault," in *Proc. IEEE PES Innov. Smart Grid Technol.*, May 2012, pp. 1–5.

[6] W. Li, J. Zhou, K. Xie, and X. Xiong, "Power system risk assessment using a hybrid method of fuzzy set and Monte Carlo simulation," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 336–343, May 2008.

[7] J. Qiu, H. Wang, D. Lin, B. He, W. Zhao, and W. Xu, "Nonpara-metric regression-based failure rate model for electric power equipment using lifecycle data," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 955–964, Jan. 2015.

[8] A. A. A. Etumi and F. J. Anayi, "The application of correlation technique in detecting internal and external faults in three-phase transformer and saturation of current transformer," *IEEE Trans. Power Del.*, vol. 31, no. 5, pp. 2131–2139, Oct. 2016.

[9] Y. Zhang, X. Li, H. Zheng, H. Yao, J. Liu, C. Zhang, H. Peng, and J. Jiao, "A fault diagnosis model of power transformers based on dissolved gas analysis features selection and improved krill herd algorithm optimized support vector machine," *IEEE Access*, vol. 7, pp. 102803–102811, 2019.

[10] J. I. Aizpurua, V. M. Catterson, B. G. Stewart, S. D. J. McArthur, B. Lambert, B. Ampofo, G. Pereira, and J. G. Cross, "Selecting appropriate machine learning classifiers for DGA diagnosis," in *Proc. IEEE Conf. Electr. Insul. Dielectr. Phenomenon (CEIDP)*, Oct. 2017, pp. 153–156.

[11] J. I. Aizpurua, B. G. Stewart, S. D. J. McArthur, B. Lambert, and J. G. Cross, "Towards a comprehensive DGA health index," in *Proc. IEEE 2nd Int. Conf. Dielectrics (ICD)*, Budapest, Hungary, Jul. 2018, pp. 1–4.

[12] M. K. Pradhan and T. S. Ramu, "On the estimation of elapsed life of oil-immersed power transformers," *IEEE Trans. Power Del.*, vol. 20, no. 3, pp. 1962–1969, Jul. 2005.

[13] B. Cai, Y. Liu, and M. Xie, "A dynamic-Bayesian-network-based fault diagnosis methodology considering transient and intermittent faults," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 1, pp. 276–285, Jan. 2017.

[14] S. Kabir, T. K. Geok, M. Kumar, M. Yazdi, and F. Hossain, "A method for temporal fault tree analysis using intuitionistic fuzzy set and expert elicitation," *IEEE Access*, vol. 8, pp. 980–996, 2020.

[15] B. Cai, X. Shao, Y. Liu, X. Kong, H. Wang, H. Xu, and W. Ge, "Remaining useful life estimation of structure systems under the influence of multiple causes: Subsea pipelines as a case study," *IEEE Trans. Ind. Electron.*, vol. 67, no. 7, pp. 5737–5747, Jul. 2020.

[16] Y. Gheraibia, S. Kabir, K. Aslansefat, I. Sorokos, and Y. Papadopoulos, "Safety+AI: A novel approach to update safety models using artificial intelligence," *IEEE Access*, vol. 7, pp. 135855–135869, 2019.

[17] S. Kabir, K. Aslansefat, I. Sorokos, Y. Papadopoulos, and S. Konur, "A hybrid modular approach for dynamic fault tree analysis," *IEEE Access*, vol. 8, pp. 97175–97188, 2020.

[18] J. I. Aizpurua, S. D. J. McArthur, B. G. Stewart, B. Lambert, J. G. Cross, and V. M. Catterson, "Adaptive power transformer lifetime predictions through machine learning and uncertainty modeling in nuclear power plants," *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4726–4737, Jun. 2019.

[19] *Mineral Oil-Impregnated Electrical Equipment in Service: Guide to the Interpretation of Dissolved and Free Gases Analysis*, Standard IEC 60599, AMD, 2007, pp. 1–8.

[20] F. Jakob, P. Noble, and J. J. Dukarm, "A thermodynamic approach to evaluation of the severity of transformer faults," *IEEE Trans. Power Del.*, vol. 27, no. 2, pp. 554–559, Apr. 2012.

[21] I. B. M. Taha, D.-E.-A. Mansour, S. S. M. Ghoneim, and N. I. Elkalashy, "Conditional probability-based interpretation of dissolved gas analysis for transformer incipient faults," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 4, pp. 943–951, Mar. 2017.

[22] Y. Wang, M. Shahidehpour, and C. Guo, "Applications of survival functions to continuous semi-Markov processes for measuring reliability of power transformers," *J. Modern Power Syst. Clean Energy*, vol. 5, no. 6, pp. 959–969, Nov. 2017.

[23] S. Lin, X. Sun, and D. Feng, "A failure rate model for traction transformer based on PHM considering multiple factors," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chengdu)*, Oct. 2016, pp. 1–6.

[24] S. Han *et al.*, "Dynamic failure rate model for transformer considering insulation aging and oil chromato-graphic monitoring data," *Power Syst. Technol.*, vol. 42, no. 10, pp. 3275–3281, 2018.

[25] W. Zhao, H. Wang, and J. Qiu, "Analysis of dynamic reliability of transformer based on monitoring data of oil chromatography," *Automat. Electr. Power Syst.*, vol. 38, no. 22, pp. 38–42 and 49, 2014.

[26] J. Pathak, Y. Jiang, and V. Honavar, "Condition data aggregation with application to failure rate calculation of power transformers," in *Proc. Hawaii Int. Conf. Syst. Sci.*, 2006, p. 241.

[27] J. Jia, H. O. U. Huijuan, and D. U. Xiuming, "Optimum maintenance policy for power delivery equipment based on Markov decision process," *High Voltage Eng.*, vol. 43, no. 7, pp. 2323–2330, 2017.

[28] S. Kabir, M. Yazdi, J. I. Aizpurua, and Y. Papadopoulos, "Uncertainty-aware dynamic reliability analysis framework for complex systems," *IEEE Access*, vol. 6, pp. 29499–29515, 2018.

[29] J. I. Aizpurua, V. M. Catterson, Y. Papadopoulos, F. Chiacchio, and G. Manno, "Improved dynamic dependability assessment through integration with prognostics," *IEEE Trans. Rel.*, vol. 66, no. 3, pp. 893–913, Sep. 2017.

[30] S. Kabir, T. Azad, M. Walker, and Y. Gheraibia, "Reliability analysis of automated pond oxygen management system," in *Proc. 18th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dhaka, Bangladesh, Dec. 2015, pp. 144–149.

[31] S. Kabir, K. Aslansefat, I. Sorokos, Y. Papadopoulos, and Y. Gheraibia, "A conceptual framework to in-corporate complex basic events in HiP-HOPS," in *Proc. Int. Symp. Model-Based Saf. Assessment*. Cham, Switzerland: Springer, Oct. 2019, pp. 109–124.

[32] W. Li, *Risk Assessment of Power Systems: Models Methods and Applications*. Hoboken, NJ, USA: Wiley, 2014.

[33] Y. Liu, L. Wenpeng, and X. Yan, "Data cleaning method for distribution transformer," *Power Syst. Technol.*, vol. 41, no. 3, pp. 1008–1014, 2017.

[34] J. Chen, W. Liao, J. Xin, Q. Zhou, C. Kang, S. Ma, and D. Song, "On-line fault diagnosis method for power transformer based on missing data repair," *Power Syst. Protection Control*, vol. 47, no. 15, pp. 86–92, 2019.

[35] *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, IEEE Standard C57.91-1995, Institute of Electrical and Electronics Engineers, IEEE Standards Board, USA, 1996.

[36] L. M. O. Queiroz and C. Lyra, "Adaptive hybrid genetic algorithm for technical loss reduction in distribution networks under variable demands," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 445–453, Feb. 2009.

[37] R. Rogers, "IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis," *IEEE Trans. Electr. Insul.*, vol. EI-13, no. 5, pp. 349–354, Oct. 1978.

[38] J. Faiz and M. Soleimani, "Dissolved gas analysis evaluation in electric power transformers using conventional methods a review," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 2, pp. 1239–1248, Apr. 2017.

[39] J. Du, Z. Liu, and Z. Yi, "A KNN algorithm for unbalanced data set classification," *Sci. Technol. Eng.*, vol. 11, no. 12, pp. 2680–2685, 2011.

[40] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient kNN classification with different numbers of nearest neighbors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1774–1785, May 2018.

[41] H. Xu and Y. Deng, "Dependent evidence combination based on shearman coefficient and pearson coefficient," *IEEE Access*, vol. 6, pp. 11634–11640, 2018.

[42] Y. Sai, R. Jinxia, and L. Zhongxia, "Learning of neural networks based on weighted mean squares error function," in *Proc. 2nd Int. Symp. Comput. Intell. Design*, 2009, pp. 241–244.

[43] X. Luo and S. S.-T. Yau, "Hermite spectral method to 1-D forward kolmogorov equation and its application to nonlinear filtering problems," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2495–2507, Oct. 2013.

[44] D. Zhai, C. Xi, J. Dong, and Q. Zhang, "Adaptive fuzzy fault-tolerant tracking control of uncertain nonlinear time-varying delay systems," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 5, pp. 1840–1849, May 2020.

[45] C. Xi and J. Dong, "Adaptive neural network-based control of uncertain nonlinear systems with time-varying full-state constraints and input constraint," *Neurocomputing*, vol. 357, pp. 108–115, Sep. 2019.

[46] T. Kaczorek, "Generalization of Cayley-hamilton theorem for n-D polynomial matrices," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 671–674, May 2005.

[47] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM Comput. Surv.*, vol. 51, no. 5, pp. 93.1–93.42, Jan. 2019.

[48] E. N. A. Franz and L. Bertling, *State of the Art-Life Time Modeling and Management of Transformers*. Stockholm, Sweden: Royal Institute of Technology, 2007.

[49] D. Zhou, Z. Wang, P. Jarman, and C. Li, "Data requisites for transformer statistical lifetime modelling—Part II: Combination of random and aging-related failures," *IEEE Trans. Power Del.*, vol. 29, no. 1, pp. 154–160, Jul. 2014.

[50] D. P. Borgers, V. S. Dolk, and W. P. M. H. Heemels, "Riccati-based design of event-triggered controllers for linear systems with delays," *IEEE Trans. Autom. Control*, vol. 63, no. 1, pp. 174–188, Jan. 2018.

[51] J. Zhong, W. Li, R. Billinton, and J. Yu, "Incorporating a condition monitoring based aging failure model of a circuit breaker in substation reliability assessment," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3407–3415, Nov. 2015.

[52] F. Zhao and H. Su, "Short-term load forecasting using Kalman filter and elman neural network," in *Proc. 2nd IEEE Conf. Ind. Electron. Appl.*, May 2007, pp. 1043–1047.

[53] H. Yang, L. Wang, Y. Zhang, H.-M. Tai, Y. Ma, and M. Zhou, "Reliability evaluation of power system considering time of use electricity pricing," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 1991–2002, May 2019.

[54] Y. Zhao, J. Kuang, K. Xie, W. Li, and J. Yu, "Dimension reduction based non-parametric disaggregation for dependence modeling in composite system reliability evaluation," *IEEE Trans. Power Syst.*, early access, Jul. 8, 2020, doi: 10.1109/TPWRS.2020.3007692.

WEI HUANG received the B.S. degree from the China University of Petroleum, Qingdao, in 2019. He is currently pursuing the M.S. degree with Chongqing University, China. His research interests include power system reliability and planning.

XUAN LI received the B.S. degree from Chongqing University, China, in 2019, where he is currently pursuing the M.S. degree. His research interests include power system protection and control.

BO HU (Member, IEEE) received the Ph.D. degree in electrical engineering from Chongqing University, China, in 2010. He is currently working as a Full Professor with the School of Electrical Engineering, Chongqing University. His research interests include power system reliability and parallel computing techniques in power systems.

JIAHAO YAN (Member, IEEE) received the B.S. degree from the Hefei University of Technology, China, in 2016. He is currently pursuing the Ph.D. degree with Chongqing University, China. His research interests include power system risk assessment and optimization.

LVBING PENG received the B.S. degree from the Hefei University of Technology, China, in 2015. He is currently pursuing the Ph.D. degree with Chongqing University, China. His research interests include power system reliability and planning.

YUE SUN received the B.S. degree from Guizhou University, in 2019. He is currently pursuing the M.S. degree with Chongqing University, China. His research interest includes power system reliability.

XIN CHENG received the B.S. degree from the North University of China, in 2019. She is currently pursuing the M.S. degree with Chongqing University, China. Her research interest includes power system reliability.

JINFENG DING received the B.S. degree from Chongqing University, China, in 2019, where he is currently pursuing the M.S. degree. His research interest includes power system reliability.

KAIGUI XIE (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Chongqing University, China, in 2001. He is currently working as a Full Professor with the School of Electrical Engineering, Chongqing University. His main research interests include power system reliability, planning, and analysis.

QINGLONG LIAO (Member, IEEE) received the Ph.D. degree in electrical engineering from Chongqing University, Chongqing, China, in 2017. He is currently working with the State Grid Chongqing Electric Power Research Institute.

LINGYUN WAN (Member, IEEE) is currently working with the State Grid Chongqing Electric Power Research Institute.