IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 5, 2020, accepted June 8, 2020, date of publication June 15, 2020, date of current version June 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3002381

A BP Neural Network-Based Hierarchical Investment Risk Evaluation Method Considering the Uncertainty and Coupling for the Power Grid

SHUJUAN CHEN[1,2], QIN JIANG[1,2], (Student Member, IEEE), YUQING HE[3], RUANMING HUANG[4], JIAYONG LI[1,2], (Member, IEEE), CAN LI[1,2], AND JING LIAO[1,2]

[1]College of Electrical and Information Engineering, Hunan University, Changsha 410082, China
[2]Hunan Key Laboratory of Intelligent Information Analysis and Integrated Optimization for Energy Internet, Hunan University, Changsha 410082, China
[3]Hunan Key Laboratory of Energy Internet Supply-Demand and Operation, Economic and Technological Research Institute of Hunan Power Grid, State Grid Hunan Power Company Limited, Changsha 410007, China
[4]Economic and Technological Research Institute of Shanghai Power Grid, State Grid Shanghai Electric Power Company, Shanghai 200122, China

Corresponding authors: Qin Jiang (1027934842@qq.com) and Yuqing He (iheyuqing@163.com)

This work was supported by the State Grid Science and Technology Projects of China under Grant SGSHJY00GPJS1800060.

ABSTRACT Investment decision-making is affected by the uncertain and highly coupled risks in the power grid, and the inaccurate risk evaluation results in the great economic losses of power companies. In order to improve the accuracy of risk evaluation and reduce the economic losses, a hierarchical risk evaluation method considering the uncertainty and coupling of risks is proposed in this paper. At the lower level, the uncertainty and time response of risks are taken into consideration for evaluating individual risks accurately in the power grid. Through the data processing of historical risk factors based on BP neural network, the distribution regularities of risk loss and probability of occurrence are performed. According to the partition of time period, risk losses expressed by the interval and corresponding probabilities in different time periods are identified. The evaluation results at the lower level are considered as the basis for risk management and the inputs at the upper level. At the upper level, a comprehensive evaluation of multiple risks is performed for evaluating the investment scheme accurately. A discretization method is developed to transform the inputs into the probability sequences, and the sequence operation theory is applied in the comprehensive evaluation of multiple risks that considers the coupling among risks. Extensive case studies are presented to validate the effectiveness of the proposed method for risk evaluation.

INDEX TERMS Investment risk evaluation, power grid, uncertainty and coupling, BP neural network.

NOMENCLATURE

PARAMETERS AND VARIABLES

t	Time
H_t	Income of the power company in t year
S_t	Average electricity transmission and distribution price in t year
C_t	Power consumption in t year
β_t	Composite loss rate in t year
ΔS_{et}	Deviation of average electricity transmission and distribution price in t year
ΔH_t	Loss caused by average electricity transmission and distribution price risk
ΔG_1, ΔG_2	GDP growth deviation
ΔC_{G1}, ΔC_{G2}	Power consumption deviation caused by ΔG_1, ΔG_2
ΔI_1, ΔI_2	Resident income deviation
ΔC_{I1}, $\Delta C_{I2}w$	Power consumption deviation caused by ΔI_1, ΔI_2
E_g	Electricity elasticity coefficient
E_y	Elasticity coefficient of electricity income
i, j	Risk factor
R_i	Risk loss of the ith risk factor before normalization
R_i^*	Risk loss of the ith risk factor after normalization
m	Total number of risk factors

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

$min\{R_j\}$	Minimum risk loss of each risk factor
$max\{R_j\}$	Maximum risk loss of each risk factor
$min\{min\{R_j\}\}$	Minimum risk loss of all risk factors
$max\{max\{R_j\}\}$	Maximum risk loss of all risk factors
$f(x)$	Probability density function of the Normal distribution
$g(t)$	Nonlinear regression fitting function
x	Deviation of the risk factor
$\mu_{(f(x))}$	Mean of the $f(x)$
$\delta_{(f(x))}$	Variance of the $f(x)$
$P_{[t^{\cdot},t^{\cdot\cdot}]}$	Probability of occurrence of the risk in the given time period
$P_{[g(t^{\cdot}),g_{(t^{\cdot\cdot})}]}$	Probability of occurrence of risk loss in the given interval
a	Lower limit of the risk loss interval
b	Upper limit of the risk loss interval
Δi	Length of discretization
$a(k)$	Evaluation result of the single risk expressed by a sequence
k	Sequence number of the discretized sequence
$a[k]$	Value of $a(k)$ at sequence number k
N	Sequence length of $a(k)$
$[b/\Delta i]$	Largest integer not exceeding $(b/\Delta i)$
$P_{[k-1,k]}$	Probability of risk loss in the interval $[\Delta i*(k-1), \Delta i*k]$
$a_i(k_i)$	Discretized sequence of the ith risk factor
k_i	Sequence number of the ith discretized sequence
N_i	Sequence length of the ith risk after discretization
$x(K)$	Comprehensive sequence of multiple risks under the operation of the Addition-Type-Convolution
K	Sequence number of the $x(K)$
M	Length of the comprehensive sequence
E	Expected value of the comprehensive sequence

I. INTRODUCTION

As the indispensable infrastructures, power grids play a significant role in the development of the economy and society [1]. The long construction period, huge investment cost and various uncertainties result in the diverse risks in the power grid investment [2], [3]. The accuracy of risk evaluation in power grid investment directly affects the power grid investment returns [4]. Therefore, it's a significant issue to evaluate risks accurately [5].

There has been a growing number of relevant papers that study risk evaluation of projects. In [6], the decision-making tree model was proposed to analyze the influence degree and probability of occurrence of risk factors, the risks are intuitive for investors. In [7], a Monte Carlo simulation method was developed that ranks the importance of risks and analyzes the impact of risks on project schedule and cost. Through determining the weight of each risk factor, a fuzzy analytical hierarchical process was presented in [8] and [9]. In [9], the triangular fuzzy number was applied to describe the expert' judgment information of risk factors. In [10], the sensitivity analysis was applied to quantitatively analyze the probability of occurrence of risk factors. In [11] and [12], a fuzzy comprehensive evaluation model based on the expert scoring was established by distinguishing the significance of risks. In [12], a multi-layer fuzzy comprehensive evaluation method was presented that uses computer to establish models, it's more convenient to calculate the comprehensive risk value of the project.

The increasing penetration of renewable energy introduces numerous challenges to the operation and energy management of electric power systems. The application of risk-based assessment methods for short-term wind power commitment and operation planning in power systems with wind power generation has attracted high interest. In [13], a conditional probabilistic method was proposed to quantify the wind power commitment risk associated with wind power commitments at different initial conditions, it's simple for system operator to make effective wind power commitment at the acceptable risk levels. In [14], an analytical method for evaluating risks associated with frequency response inadequacy was developed, a risk index with the sum of probabilities and quantified consequences is used to assess system security. To achieve the economic dispatch schedule, the power system scheduling models with wind power forecast uncertainty information in the electricity market were developed. In [15], based on the proposed short-term operation model, an optimal tradeoff between the profit and risk for wind power penetrated system was made by value at risk and integrated risk management methods. In [16], a model to optimize the uncertainty intervals of wind power was proposed for power system scheduling problems, which achieves a better tradeoff between economics and reliability and stabilizes the energy prices. In the literature, to improve the efficiency of energy management and deal with the uncertainties of renewable energy, various approaches were employed. In [17], a PDF-based risk assessment model was proposed that considers the losses due to renewable power generation fluctuations lying outside of acceptable interval, the risk losses are considered as the optimization variable and the model provides an economically feasible energy management solution. In [18], an optimal risk-averse for heterogeneous energy storage deployment was proposed in a residential multi-energy microgrid under diverse uncertainties, which can effectively increase profits and avoid the risk.

The existing papers also study the evaluation of power system faults-induced risks and cyberattacks-induced risks. In [19], a machine learning-based anomaly detection methodology was developed to detect the cyberattacks on load forecasting and make suitable operational decisions for the electricity delivery. In [20], a flexible machine learning method using spatiotemporal patterns was developed for cyberattack detection, which performs better under different cyberattack scenarios and improves the detection accuracy in distribution systems. In [21], a risk evaluation based approach to the replacement strategy was proposed to evaluate the risk of aged High Voltage Direct Current components, which quantifies the expected system risks and risk costs due to three replacement options. In [22], failure probability model of transmission lines based on real-time operating conditions and hidden failure model based on Markov method were presented to assess operational risk on transmission system cascading failure. In [23], a model and algorithm for transmission expansion planning considering the blackout risk was proposed, the algorithm takes the power-law tail risk into consideration and helps to reduce the plans blackout risk. In [24], an event-driven emergency demand response strategy based on whale optimization algorithm was proposed to effectively mitigate system voltage instability risk.

The risk matrix method is defined by determining the level of risk impact and risk probability, which is the significant difference compared with other common risk evaluation methods [25]. In [25], the risk matrix method was proposed to assess the power company risk security, it provides the direction for the manager to improve the management and has some theoretical significance and practical value. In [26], a fuzzy risk matrix method was proposed that uses linguistic variables to express the probability and severity of the consequences, the method can be applied in different areas. A comprehensive evaluation method based on the risk matrix method was presented in [27] and [28] by constructing an expert two-dimensional matrix. In [28], the Borda ordinal method was applied to rank the importance of risks, it's effective for identifying key risks and comprehensive risk level. In [29], the risk matrix method was applied to assess the risk level in the short-term, middle-time and long-term. In [30], a risk matrix analysis framework was proposed for risk evaluation based on potential risk influence, which considers the controllability, manageability, criticality and uncertainty of risks.

These methods evaluate the risk through a definite risk value that represents the sum of losses and probabilities of all risk factors, the risk evaluation results provide basis for planning and management. Since risks have the characteristics of variability, uncertainty and coupling, the risk loss and probability of occurrence change with time, and the changes show different regularities at different time. Therefore, the single risk evaluation result expressed by a definite value is not in accordance with the reality, and the comprehensive evaluation result of multiple risks representing the sum of the single risk results cannot satisfy the accuracy required by the power grid company.

In order to evaluate risks accurately, a BP neural network based hierarchical investment risk evaluation method is proposed in this paper. Compared with previous methods, this method considers the uncertainty and coupling of risks and identifies the distribution regularities of risks that change with time. The results representing loss interval and probability of occurrence are used to evaluate the single risk, different loss intervals corresponds to different probabilities. The comprehensive probability sequences reflecting probability distribution under different loss intervals are used to evaluate multiple risks. To consider the uncertainty of risks, the distribution regularities of risks are identified and a time partition method is developed. Therefore, the single risk results can adapt to the characteristics of the risk changing with time. To consider the coupling among risks, a discretization method is developed and the evaluation results of individual risks are transformed into probability sequences. Furthermore, the sequence operation theory is applied in the comprehensive evaluation of multiple risks. The contributions of this paper are listed as follows.

1) In the proposed method, the uncertainty and time response of risks are considered at the lower level. According to the evaluation results of individual risks from the lower level, the comprehensive evaluation of multiple risks that considers the coupling among risks is performed at the upper level. The proposed method improves the accuracy of investment risk evaluation and reduces the economic losses of power companies.
2) In the single risk evaluation, a time partition method based on distribution regularities is developed to distinguish the influence of risks at different time. The risk loss changing with time is expressed by the interval that can describe the uncertainty of the risk.
3) In the evaluation of multiple risks, a comprehensive evaluation method is developed to consider the coupling among risks based on the discretization method and sequence operation theory. The evaluation results of multiple risks are expressed by probability sequences that can reflect the probability distribution of comprehensive risk under different loss intervals.
4) A central segment discretization method is developed to divide the loss interval into more small intervals, the probability of the small interval is equal to the probability of the loss in the center of the interval. After discretization, the evaluation results of individual risks are transformed into probability sequences that can reflect the probability distribution of risks under different loss intervals.

The remainder of this paper is organized as follows. In Section II, the lower level of the single risk evaluation is proposed. In Section III, the comprehensive evaluation of multiple risks at the upper level is proposed. Extensive case studies are presented and discussed in Section IV. Section V concludes.

II. SINGLE RISK EVALUATION AT THE LOWER LEVEL

The historical predicted values of risk factors are obtained by collecting the historical actual values of risk factors and establishing the time series prediction model. Then, the deviations of risk factors are determined. Through the fitting of the nonlinear regression and probability distribution, the partition of time period is finished and the distribution regularities of risks in different time periods are identified. The evaluation process of the single risk is shown in Fig. 1.

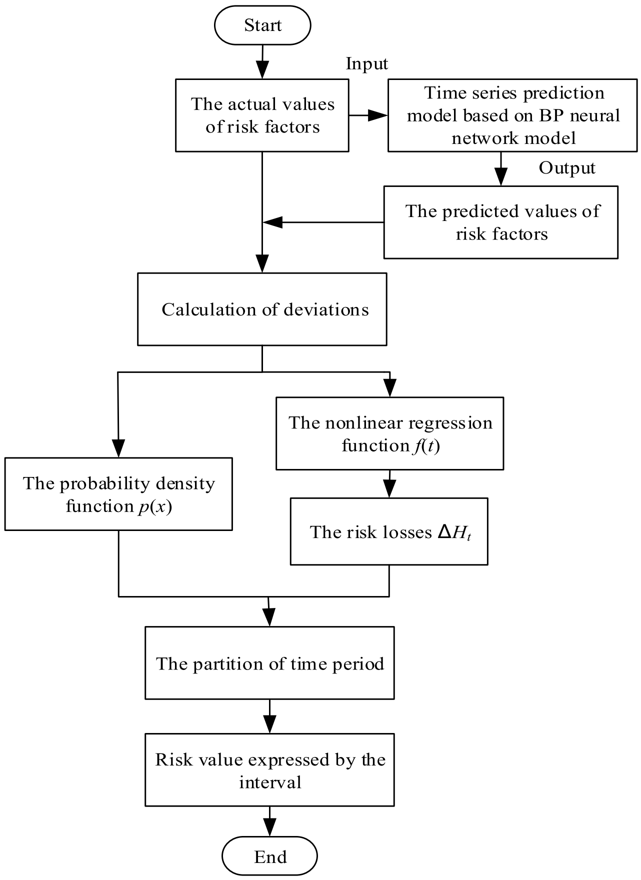


FIGURE 1. Flow chart of single risk evaluation at the lower level.

A. DATA PROCESSING OF RISK FACTORS

The time series prediction model based on BP neural network provides an effective way for time series prediction of highly nonlinear dynamic relationships [31]–[33]. The historical predicted values of risk factors are obtained by establishing corresponding time series prediction models. The historical actual values of risk factors from 1990 to 2001 are used to establish the model. The data of the adjacent three years are taken as a group and there are nine groups. The first six groups are used for model training, the others are used for model testing. In each group, the three data are used as input, and the value of the next year is used as output. After the model is established, the historical actual of risk factors from 1998 to 2017 are considered as the input data for acquiring the predicted values from 2001 to 2018. The actual values of the adjacent three years are input to the model, the predicted value of next year is output. Based on the actual values and predicted values of risk factors from 2001 to 2018, the deviations between the actual and predicted values are determined.

The deviations of risk factors are considered as the data for the probability distribution fitting and the nonlinear regression fitting. The probability distribution is fitted to obtain the probability density of the risk factor under different deviations. In order to choose the fitting function with the desirable performance, the test values of the first three fitting functions with good effect are verified, which are the Anderson-Darling test (A-D) value, the Kolmogorov-Smirnov test (K-S) value and the Chi-square goodness-of-fit test (Chi-Sq) value. A-D value is selected to verify the probability distribution fitting for the limitation of sample size, and the smaller the A-D value, the better the fitting degree [34].

B. THE MODEL OF RISK LOSS

The risk loss under the corresponding deviation can be obtained as shown in (1)

$$H_t = S_t \times C_t \times (1 - \beta_t) \tag{1}$$

When calculating the risk losses caused by deviations of risk factors, the risk loss of the single risk is only related to the deviation of the risk factor, it's equal to multiplying corresponding deviation in (1). As an example, assuming the deviation of average electricity transmission and distribution price in t year is ΔS_{et}, then the risk loss of ΔH_t caused by the risk of average electricity transmission and distribution price is obtained using (2).

$$\Delta H_t = \Delta S_{et} \times S_t \times C_t \times (1 - \beta_t) \tag{2}$$

If risk factors are not represented in (1), the corresponding risk losses are calculated by analyzing the correlation with the risk factor represented in (1). The economic development and power consumption have a positive correlation, which means the faster the economic development, the more the power consumption. The correlation between economic development and power consumption can be described by the electricity elasticity coefficient. The increase of resident income will promote the electricity demand of residents, the correlation between resident income and power consumption can be described by the elasticity coefficient of electricity income. The relationship between the deviation of the GDP growth and power consumption, the deviation of resident income and power consumption can be expressed by the formula (3) [35].

$$\begin{cases} \left[\Delta C_{G1}, \Delta C_{G_2}\right] = E_g \times [\Delta G_1, \Delta G_2], \Delta G_1 < \Delta G_2 \\ \left[\Delta C_{I1}, \Delta C_{I_2}\right] = E_y \times [\Delta I_1, \Delta I_2], \Delta I_1 < \Delta I_2 \end{cases} \tag{3}$$

Through the relationship between risk factor deviation and risk loss, the influence process and the distribution regularities of risks can be identified. After obtaining the losses caused by each risk factor, the losses are processed with normalization for the convenience of quantitative analysis. The result is the number between 0 and 1, and the formula of normalization is as follows [36].

$$R_i^* = \left(R_i - min\left\{min\left\{R_j\right\}\right\}\right) / \left(max\left\{max\left\{R_j\right\}\right\} - min\left\{min\left\{R_j\right\}\right\}\right) \tag{4}$$

The minimum and maximum risk loss of all risk factors are determined as following three steps. Firstly, the minimum and maximum deviation of each risk factor are determined. Secondly, the minimum and maximum risk loss of each risk factor are calculated according to the established risk loss model, the model represents the relationship between risk loss and risk factor deviation. Finally, the minimum and maximum risk loss of all risk factors are determined through comparing the minimum and maximum risk loss of each risk factor.

C. THE PARTITION OF TIME PERIOD

The nonlinear regression and probability distribution are fitted for describing the time response and distribution regularities of risks. The nonlinear regression fitting means that the deviation of the risk factor changes with time, plotted by time on the horizontal axis and the deviation of the risk factor on the vertical axis. The probability distribution fitting is plotted by the deviation of the risk factor on the horizontal axis and the corresponding probability density on the vertical axis.

The partition of time period is based on the distribution properties of probability density function. Such as the 3δ principle of the Normal Distribution, the probability in the interval $[\mu - 3\delta, \mu + 3\delta]$ is 99.74%, it means that the distribution can be described more accurately in consideration of the risk loss in the interval $[\mu - 3\delta, \mu + 3\delta]$. The Logistic distribution and the Normal Distribution have similar shapes, therefore, they have similar partition methods. In the effective interval, the partition of the Triangle distribution can be divided into the upper triangular and the lower triangular. The probability densities of the Uniform distribution are the same within its effective interval. Therefore, the time partition of the risk only considers the effective interval.

Assuming the probability density function of a risk obeys the Normal distribution, expressed by $f(x)$; the nonlinear regression distribution is expressed by $g(t)$. The result of time partition is shown in Fig.2.

As can be seen in Fig.2, the risk is divided into three-time periods, which are $[t1, t2]$, $[t3, t4]$, $[t5, t6]$. $t2$, $t3$, $t6$ are the intersections of the curve x equal to $g(t)$ and the straight line x equal to $x1$; $t1$, $t4$, $t5$ are the intersections of the curve x equal to $g(t)$ and the straight line x equal to $x2$, the formula is as follows.

$$\begin{cases} (t2, t3, t6) = (\{x = x1\} \cap \{x = g(t)\})|_t \\ (t1, t4, t5) = (\{x = x2\} \cap \{x = g(t)\})|_t \\ x1 = \mu_{(f(x))} - 3\delta_{(f(x))} \\ x2 = \mu_{(f(x))} + 3\delta_{(f(x))} \end{cases} \tag{5}$$

Through the partition of time period, the evaluation results of individual risks representing the loss interval and probability of occurrence are determined. In general, in the same time period, different risks obey different distribution regularities, and the partition results are different. In different time periods, the partition results of the same risk can also be different. Therefore, to evaluate multiple risks, it's necessary to calculate the probability of risk loss in the given time period and the given interval.

In the effective time period, the area enclosed by probability density function and loss interval corresponds to the probability of risk loss in the given interval. Risk loss at other time or other intervals can be ignored for the probability of occurrence is too small. The probability of risk loss within

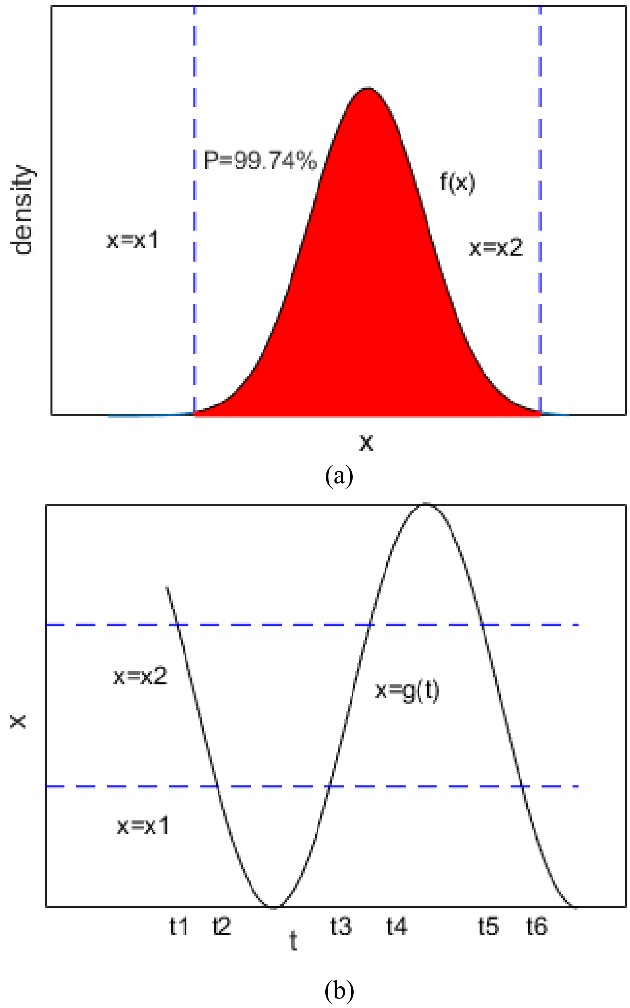


FIGURE 2. **Process of time partition. (a) Basis of time partition. (b) Result of time partition.**

the effective interval and the effective time period can be calculated as follows.

$$\begin{cases} \boldsymbol{P}_{\left[t', t''\right]} = \boldsymbol{P}_{\left[g(t'), g\left(t''\right)\right]} = \int_{t'}^{t''} f\,(g(t))dt \\ \left[g\left(t'\right), g\left(t''\right)\right] \subseteq [x1, x2] \\ \left[t', t''\right] \subseteq \{[t1, t2]\,, [t3, t4]\,, [t5, t6]\} \end{cases} \tag{6}$$

III. COMPREHENSIVE EVALUATION OF MULTIPLE RISKS AT THE UPPER LEVEL

After the partition of time period, the risk loss expressed by the interval and corresponding probability of occurrence in the multi-time periods are obtained at the lower level. Considering the coupling among risks, the sequence operation theory is applied in the comprehensive evaluation of multiple risks. The flow chart of the comprehensive evaluation is shown in Fig. 3.

A. THE PROCESS OF DISCRETIZATION

The Addition-Type-Convolution of the sequence operation theory is applied in the comprehensive evaluation of multiple risks, which means the summary of multiple

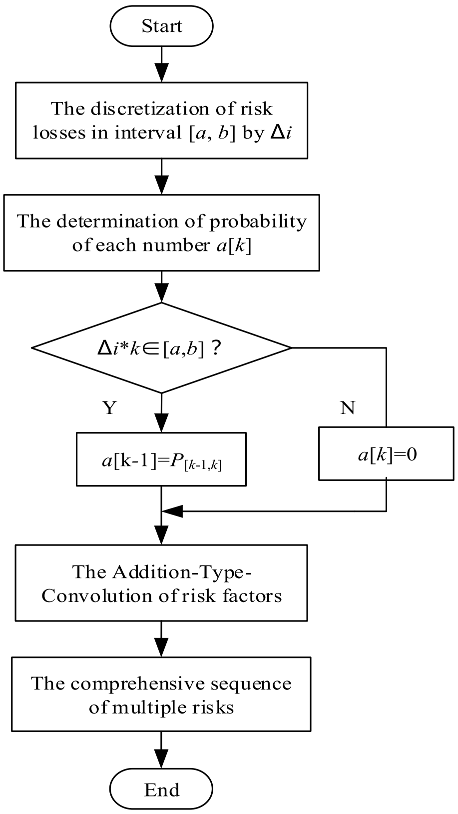


FIGURE 3. Flow chart of comprehensive evaluation of multiple risks at the upper level.

random variables [37]. The sequence operation theory can deal with problems of probability sequences. Therefore, the evaluation results representing the loss interval and probability are discretized to form corresponding probability sequences.

Before the Addition-Type-Convolution, the loss interval is discretized by the length of Δi. If the discretized value is not included in the interval of the risk loss, the probability is zero, it means that $a[k]$ is equal to 0. Otherwise, the probability is discretized by the central segment. The probability of the risk loss in the interval $[\Delta i * (k-1), \Delta i * k]$ is equal to the probability of the risk loss in the center of the interval, and the probability is the value at the sequence number $(k-1)$ when the risk loss is expressed by a sequence.

$$\begin{cases} a(k) = a_{(1\times(N-1))} = \{a[k]\} \\ k = \{0, 1, \cdots, N-1\} \end{cases} \tag{7}$$

$$N = [b/\Delta i] \tag{8}$$

$$a[k-1] = \begin{cases} 0, & \Delta i * k \notin [a,b] \\ P_{[k-1,k]}, & \Delta i * k \in [a,b] \end{cases} \tag{9}$$

B. THE COMPREHENSIVE EVALUATION

After the discretization, the evaluation result of the single risk is expressed by a sequence. In consideration of the coupling among risks, the Addition-Type-Convolution is applied in the comprehensive evaluation of multiple risks. The evaluation result of multiple risks is expressed by a comprehensive probability sequence, and the formula is as follows.

TABLE 1. Index system of economic development risk.

First-order Risk	Secondary Risks
Economic development risk B	Average electricity transmission and distribution price B1
	Resident income B2
	GDP growth B3
	Electrical load B4

$$\begin{aligned} x(K) &= a_1(k_1) \oplus a_2(k_1) \oplus a_i(k_i) \oplus \cdots \oplus a_m(k_m) \\ &= \sum_{k_1+k_2+\cdots+k_i=K} \left[\prod_{i=1}^{m} a_i(k_i) \right] \end{aligned} \tag{10}$$

$$M = \sum_{i=1}^{m} N_i,\ K = \{0, 1, \cdots, M\} \tag{11}$$

$$E = \Delta i \sum_{0}^{M-1} (K-1) * x[K] \tag{12}$$

The comprehensive sequence $x(K)$ can describe the comprehensive probability distribution of multiple risks under different losses, which is more intuitive for investors to make investment decisions. In order to make a comparison with the risk matrix method that assesses risks by a single and determined level, the expected value of the sequence is calculated by (12). The expected value of the sequence represents the average risk loss. Furthermore, the expected value of the single risk expressed by a sequence can also be calculated based on (11) [38], and it can reflect the average loss of the single risk.

Different risks have different lengths of discretized sequences. The discretized sequences of different risks can be expressed by the form similar to a matrix.

$$\begin{bmatrix} a_{1(1\times(N_1-1))} \\ \vdots \\ a_{m(1\times(N_m-1))} \end{bmatrix} = \begin{bmatrix} a_1[0] & \cdots & a_1[N_1-1] \\ & \vdots & \\ a_m[0] & \cdots & a_m[N_m-1] \end{bmatrix} \tag{13}$$

IV. CASE STUDY

There are many uncertain factors that lead to risks in the power grid investment. In order to ensure the comprehensibility and objectivity of evaluation indexes, eight uncertain factors that may exist in the investment process of the power grid are identified through the investigation. The fuzzy theory is applied to identify five key uncertain factors. On the basis of the key uncertain factors, the model of GA-BP is applied to identify four risk factors for the feasibility of the selected risk factors. Finally, the evaluation index system of economic development risk is established in Table 1.

The GDP growth risk represents uncertain losses and the probability of occurrence caused by the variation of the GDP growth. The variation of the GDP growth is evaluated by the deviation between the actual value and the predicted value. The bigger the deviation, the bigger the risk caused by the GDP. The deviation of the GDP growth that changes with

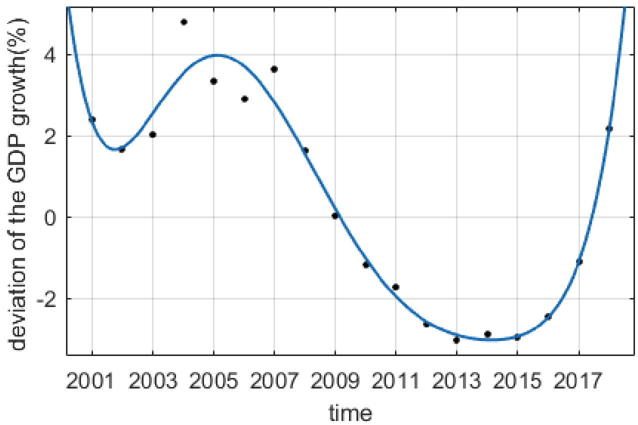


FIGURE 4. Nonlinear regression fitting of the GDP growth risk.

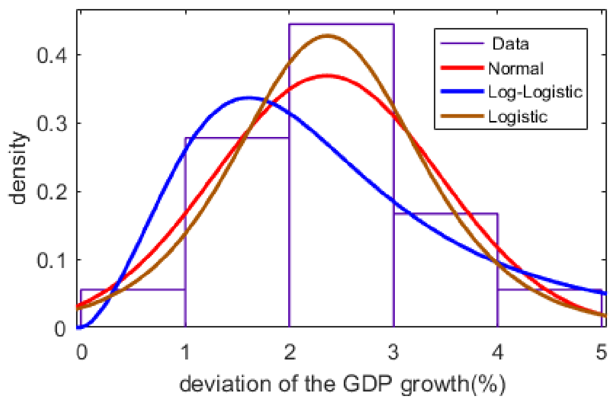


FIGURE 5. Probability distribution fitting of the GDP growth risk.

TABLE 2. Test values of the GDP growth risk.

Distribution Function	A-D Test	K-S Test	Chi-Sq Test
Log Logistic distribution	0.1257	0.0656	1.7343
Logistic distribution	0.1048	0.0638	0.5961
Lognormal distribution	0.1325	0.0721	1.7342

time is identified, and its fitting function is shown in Fig 4. The horizontal axis shows time and 2001 is the first year, the vertical axis shows the deviation between the actual and predicted value of the GDP growth. According to the fitting result, the distribution of the GDP growth risk that changes with time is fitted by an approximate polynomial.

The probability distribution of the GDP growth risk is fitted, and the fitting result is shown in Fig 5. The horizontal axis shows the deviation of the GDP growth, the vertical axis shows the corresponding probability density. In order to choose the fitting function with the desirable performance, the test values of the first three fitting functions with the good effect are verified. The test values of the GDP growth are shown in Table 2.

As can be seen in Table 2, A-D test value of the Logistic distribution is the smallest. Therefore, the Logistic distribution is the fitting function with the desirable performance,

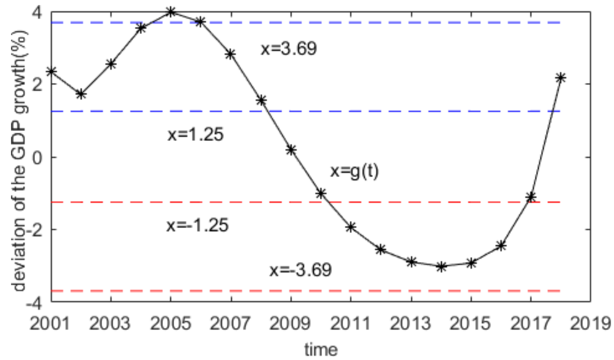


FIGURE 6. The partition result of the GDP growth risk.

and the deviation of GDP growth tends to the Logistic distribution.

The partition result of the GDP growth risk in time period is shown in Fig. 6. The horizontal axis shows time and 2001 is the first year, the vertical axis shows the deviation of the GDP growth. The four straight lines represent the deviation of the GDP growth in the interval $[\mu - 3\delta, \mu + 3\delta]$ and $[-\mu - 3\delta, -\mu + 3\delta]$, μ and δ are the mean and variance of the Logistic distribution, which are equal to 2.47 and 0.407. The curve represents the fitting result of the nonlinear regression distribution. The deviation of the GDP growth that changes with time is fitted by an approximate polynomial, and the expression is as follows.

$$g(t) = 4.61e^{-5}t^6 - 0.0025t^5 + 0.05t^4 - 0.47t^3 + 1.82t^2 - 2.03t + 2.34 \quad (14)$$

It can be seen from Fig. 6 that the GDP growth risk is divided into four-time periods. In the first time period, the risk of GDP growth is approximate to the invented triangular distribution. In the second time period, the GDP growth risk is approximate to the linear distribution. In the third time period, the GDP growth risk is approximate to the invented triangular distribution. In the fourth time period, the GDP growth risk is approximate to the linear distribution. At other time, the probability of the GDP growth risk is too small so that it can be ignored.

The GDP growth risk from 2015 to 2018 is chosen to calculate the comprehensive evaluation result of multiple risks. In this time period, the GDP growth risk can be divided into two-time periods. In the first time period, the lower limit of time is 2015, the corresponding deviation of the GDP growth is -2.76%; the upper limit of time is the intersection of the curve and the straight line x equal to -1.25, the corresponding deviation of the GDP growth is -1.25%. In the second time period, the lower limit of time is the intersection of the curve and the straight line x equal to 1.25, the corresponding deviation of the GDP growth is 1.25%; the upper limit of time is 2018, the corresponding deviation of the GDP growth is 2.17%. In the first time period, the deviation of the GDP growth belongs to the interval $[-2.76, -1.25]$, and the corresponding probability of risk loss in this interval is 0.48. In the second time period, the deviation of the GDP growth

TABLE 3. Risk loss and probability in the first time period.

Risk Factor	B1	B2	B3	B4
Risk loss R	[0.22,0.61]	[0.18,0.52]	[0.21,0.67]	[0.29, 0.48]
Probability P	0.44	0.49	0.48	0.56

TABLE 4. Risk loss and probability in the second time period.

Risk Factor	B1	B2	B3	B4
Risk loss R	[0.33,0.82]	[0.14,0.54]	[0.21,0.53]	[0.24,0.68]
Probability P	0.84	0.69	0.24	0.41

belongs to the interval [1.25, 2.17], and the corresponding probability of risk loss in this interval is 0.24.

Through the process of normalization, the risk loss expressed by the interval and corresponding probability of each risk factor in different time periods are listed in Table 3 and Table 4.

After obtaining the evaluation results of individual risks at the lower level, the risk losses expressed by the interval are discretized by the length of 0.1. The principle of discretization is based on (8) and (9). The discretized sequences in the first time period are listed in (15).

$$\begin{bmatrix} a_{B1(1\times7)} \\ a_{B2(1\times6)} \\ a_{B3(1\times7)} \\ a_{B4(1\times6)} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.11 & 0.04 & 0.08 & 0.15 & 0.06 \\ 0 & 0.08 & 0.13 & 0.07 & 0.16 & 0.05 & \\ 0 & 0 & 0.04 & 0.06 & 0.17 & 0.09 & 0.12 \\ 0 & 0 & 0.13 & 0.14 & 0.17 & 0.12 & \end{bmatrix} \tag{15}$$

In the second time period, the discretized sequence of each risk is listed as follows.

$$\begin{bmatrix} a_{B1(1\times9)} \\ a_{B2(1\times6)} \\ a_{B3(1\times6)} \\ a_{B4(1\times7)} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0.24 & 0.15 & 0.08 & 0.16 & 0.11 & 0.1 \\ 0.05 & 0.23 & 0.18 & 0.09 & 0.07 & 0.07 & & & \\ 0 & 0 & 0.03 & 0.07 & 0.08 & 0.06 & & & \\ 0 & 0 & 0.04 & 0.13 & 0.08 & 0.07 & 0.08 & & \end{bmatrix} \tag{16}$$

After the discretization, the sequence operation theory is applied in the comprehensive evaluation of multiple risks. The Addition-Type-Convolution is applied to obtain the comprehensive probability sequence of multiple risks, the comprehensive sequence of risk factors in the first time period is listed as follows.

$S_1 =$ [0, 0, 0, 0, 0, 0, 7.8960e-04, 0.0036, 0.0114, 0.0272, 0.0511, 0.0844, 0.1167, 0.1433, 0.1537, 0.1422, 0.1157, 0.0790, 0.0444, 0.0199, 0.0059, 7.4543e-04].

The probability distribution of multiple risks in the first time period is shown in Fig. 7, and the expected value of the comprehensive sequence is 1.436.

The comprehensive sequence of risk factors in the second time period is listed as follows.

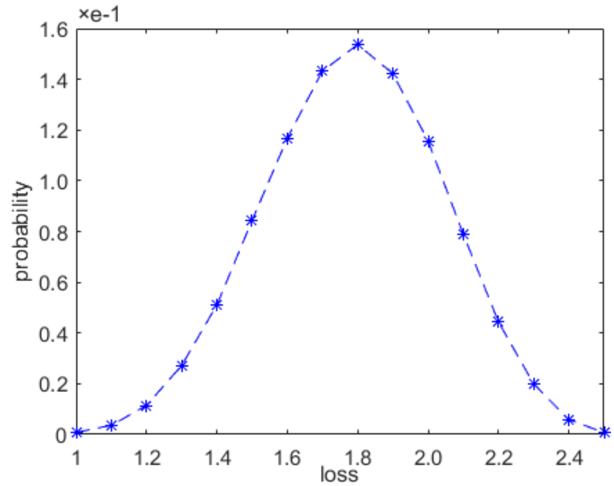


FIGURE 7. The probability distribution in the first time period.

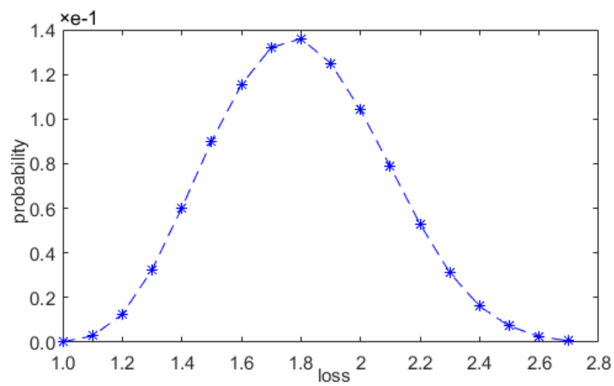


FIGURE 8. The probability distribution in the second time period.

TABLE 5. Assessment basis of the risk level.

Mark	Influence Degree	Probability	Assessment Level
Severity	≥80%	> 90%	5
High	≥50%	> 60%	4
Medium	≥20%	> 40%	3
Low	≥5%	> 10%	2
Very low	< 5%	≤10%	1

$S_2 =$ [0, 0, 0, 0, 0, 0, 2.5880e-04, 0.0028, 0.0125, 0.0324, 0.0602, 0.0901, 0.1154, 0.1320, 0.1360, 0.1250, 0.1040, 0.0787, 0.0529, 0.0311, 0.0161, 0.0072, 0.0026, 6.0386e-04].

The probability distribution of multiple risks in the second time period is shown in Fig. 8, and the expected value of the comprehensive sequence is 1.448.

In the single risk evaluation, the risk matrix method and the hierarchical risk evaluation method proposed in this paper quantify risks from two sides, which are the influence degree and probability of occurrence. The risk matrix method evaluates the risk level based on the expert scoring method, the assessment basis is listed in Table 5 [39]. The hierarchical risk evaluation method evaluates the risk by identifying its distribution regularities.

TABLE 6. The assessment results of the risk level.

	Level\Risk factor	B1	B2	B3	B4
The first time period	loss degree	[2,4]	[2,3]	[2,4]	3
	Probability	3	3	3	3
The second time period	Loss degree	[3,4]	[2,3]	[2,3]	[2,3]
	Probability	4	4	2	3
Risk matrix method	Risk loss	3	2	4	3
	Probability	4	4	3	3
Actual value	Loss degree	3	3	2	3
	Probability	3	2	4	2

TABLE 7. Risk evaluation results of individual risks

Risk Factor	B1	B2	B3	B4
Risk matrix method	12	8	12	9
Actual value	9	6	8	6
weight	0.2	0.15	0.4	0.25
Expected value in first time period	7.97	5.32	10.12	8.19
Expected value in second time period	10.42	6.88	9.05	8.00

The hierarchical risk evaluation method considers the uncertainty of the risk, the risk loss is expressed by the interval. In order to compare the two methods, the risk losses expressed by the intervals in Table 3 and Table 4 are transformed into corresponding risk levels based on Table 5. The evaluation results of the two methods are listed in Table 6. Compared with the risk matrix method, the risk loss obtained by the hierarchical risk evaluation can have multiple levels, instead of a definite level. The risk levels expressed by the interval are the integers no less than the lower limit, and no more than the upper limit of the interval.

In different time periods, the levels of risk loss can be different. In the same time period, the risk loss can have multiple levels, and the actual level of risk loss belongs to one of these levels. According to Table 6, the evaluation result of the single risk based on the risk matrix method is calculated. The result is the product of the loss degree and the probability of occurrence. The risk evaluation result obtained by the hierarchical risk evaluation method is the expression of probability sequence. To compare the two methods, the expected value of the probability sequence is calculated, and the evaluation results of the single risk are listed in Table 7.

It can be concluded from Table 7 that compared with the risk matrix method, the hierarchical risk evaluation method has better accuracy in the evaluation of the single risk. In different time periods, the evaluation results of the single risk are different. According to the method proposed in this paper, the risk loss of the single risk is expressed by the interval, and the evaluation result can reflect the probability distribution of the risk under different loss intervals.

In the comprehensive evaluation of multiple risks, the risk matrix method determines the weight of individual risks based on the analytic hierarchy process, as shown in Table 7. The hierarchical risk evaluation method applies the sequence operation theory to evaluate multiple risks.

TABLE 8. Comprehensive evaluation results of multiple risks.

Risk value	Comprehensive value	Risk level
Risk matrix method	10.65	3
Actual value	7.4	2
The first time period	8.17	2
The second time period	8.24	2

The comprehensive evaluation results of multiple risks obtained by the two methods and corresponding risk levels are listed in Table 8.

The evaluation result of multiple risks obtained by the risk matrix method is a comprehensive level, which is the weighted average value of the single risk evaluation result and corresponding weigh. The result obtained by the hierarchical risk evaluation method is expressed by a probability sequence. To compare the two methods, the expected value of the comprehensive probability sequence is calculated.

It can be seen from Table 8 that compared with the risk matrix method, the hierarchical risk evaluation method has higher accuracy in the comprehensive evaluation of multiple risks. The obtained result of multiple risks is not a definite level, but a probability sequence which can reflect the probability distribution of the economic development risk under different losses. It can be concluded that the method proposed in this paper considers the uncertainty and coupling of risks, its accuracy is verified higher.

V. CONCLUSIONS

In this paper, a hierarchical risk evaluation method is proposed for power grid investment that considers the uncertainty and coupling of risks. This method includes the evaluation of the single risk and multiple risks. At the lower level, the single risk is evaluated by considering the uncertainty of the risk and recognizing its distribution regularities. Through the partition of time period, the evaluation results of individual risks expressed by the interval can describe the uncertainty of risks better. At the upper level, the comprehensive evaluation method of multiple risks that considers the coupling among risks is developed. The sequence operation theory is applied in the comprehensive evaluation of multiple risks to express the result by an intuitive probability sequence, instead of a definite value. Finally, the method proposed in this paper is compared with the risk matrix method in the case study. The results show that the method proposed in this paper has high accuracy in the evaluation of the single risk and multiple risks.

REFERENCES

[1] A. K. David and F. Wen, "Transmission planning and investment under competitive electricity market environment," in *Proc. Power Eng. Soc. Summer Meeting. Conf.*, Jul. 2001, pp. 1725–1730.
[2] M. Wang, Z. Tan, and R. Zhang, "Risk Evaluation Model of the Power Grid Investment Based on Increment Principle," *Trans. China Electrotech. Soc.*, vol. 21, no. 9, pp. 18–24, 2006.
[3] Z. L. Hu, Y. J. Zhang, C. B. Li, J. Li, Y. J. Cao, and D. S. Luo, "Utilization efficiency of electrical equipment within life cycle assessment: Indexes, analysis and a case," *Energy*, vol. 88, pp. 885–896, Aug. 2015.

[4] X. Wang, "Analysis of investment risks of power grid enterprises under incremental distribution business," *Sci. Technol. Innov.*, vol. 15, pp. 143–144, 2019.

[5] J. F. Wang, H. Chen, Y. Zhang, and H. Huang, "Basic studies on risk management of power grid operation," *Southern Power Syst. Technol.*, vol. 9, no. 2, pp. 1–8, 2015.

[6] K. D. Prasanta, "Decision support system for risk management," *Manage. Decis.*, vol. 39, no. 8, pp. 634–649, 2001.

[7] T. W. Gao, "Risk Quantification and analysis for the research-typed projects based on MCS method," *Sci. Technol. Rev.*, vol. 30, no. 13, pp. 31–35, Apr. 2012.

[8] J. R. Ribas, M. E. Arce, F. A. Sohler, and A. S. García, "Data and calculation approach of the fuzzy AHP risk assessment of a large hydroelectric project," *Data Brief*, vol. 25, Aug. 2019, Art. no. 104294.

[9] S. L. Zhu, P. Zhou, Y. Han, and H. C. Hai, "Research on risk analysis based on fuzzy AHP method," *Comput. Integr. Manuf. Syst.*, vol. 10, no. 8, pp. 980–984, Aug. 2004.

[10] M. Ketabdari, F. Giustozzi, and M. Crispino, "Sensitivity analysis of influencing factors in probabilistic risk assessment for airports," *Saf. Sci.*, vol. 107, pp. 173–187, Aug. 2018.

[11] L. Hanfang, B. Li, T. Wang, and X. Wang, "Application of the fuzzy comprehension decision model to analyze the generation company's inner risk," *Mod. Electr. Power*, vol. 23, no. 3, pp. 74–78, Jun. 2006.

[12] J. H. Zhou, "Research on quantitative evaluation of engineering project risk," *Project Manage.*, vol. 5, pp. 31–35, May 2010.

[13] R. Karki, S. Thapa, and R. B. Billinton, "A simplified risk-based method for short-term wind power commitment," *IEEE Trans. Sustain. Energy*, vol. 3, no. 3, pp. 498–505, Jul. 2012.

[14] M. Negnevitsky, D. H. Nguyen, and M. Piekutowski, "Risk assessment for power system operation planning with high wind power penetration," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1359–1368, May 2015.

[15] X. Li and C. Jiang, "Short-term operation model and risk management for wind power penetrated system in electricity market," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 932–939, May 2011.

[16] Y. Wang, Z. Zhou, A. Botterud, and K. Zhang, "Optimal wind power uncertainty intervals for electricity market operation," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 199–210, Jan. 2018.

[17] Z. Liang, H. Chen, X. Wang, S. Chen, and C. Zhang, "Risk-based uncertainty set optimization method for energy management of hybrid AC/DC microgrids with uncertain renewable generation," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1526–1542, Mar. 2020.

[18] Z. Li, Y. Xu, X. Feng, and Q. Wu, "Optimal stochastic deployment of heterogeneous energy storage in a residential multi-energy microgrid with demand-side management," *IEEE Trans. Ind. Informat.*, early access, Feb. 3, 2020, doi: 10.1109/TII.2020.2971227.

[19] M. Cui, J. Wang, and M. Yue, "Machine learning-based anomaly detection for load forecasting under cyberattacks," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5724–5734, Sep. 2019.

[20] M. Cui, J. Wang, and B. Chen, "Flexible machine learning-based cyberattack detection using spatiotemporal patterns for distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1805–1808, Mar. 2020.

[21] W. Li, P. Choudhury, D. Gillespie, and J. Jue, "A risk evaluation based approach to replacement strategy of aged HVDC components and its application at BCTC," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1834–1840, Jul. 2007.

[22] X. Wu, J. H. Zhang, L. W. Wu, and Z. Huang, "Method of operational risk assessment on transmission system cascading failure," *Proc. CSEE*, vol. 32, no. 34, pp. 74–82, Dec. 2012.

[23] Y. J. Cao, L. H. Cao, C. B. Li, X. R. Li, and L. Yu, "A model and Algorithm for transmission expansion planning considering the blackout risk," *Proc. CSEE*, vol. 34, no. 1, pp. 138–145, Jan. 2014.

[24] M. Amroune, T. Bouktir, and I. Musirin, "Power system voltage instability risk mitigation via emergency demand response-based whale optimization algorithm," *Protection Control Mod. Power Syst.*, vol. 4, no. 1, pp. 1–14, Dec. 2019.

[25] W. Xu, C. Tao, W. Yujie, G. Qinrui, and T. Yangxin, "Risk assessment of power information risk security based on risk matrix," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Oct. 2018, pp. 1494–1498.

[26] Y. L. Yao, Y. J. Peng, X. X. Li, and A. R. Zhang, "Research on safety risk management of civil construction projects based on risk matrix method," in *Proc. IOP Conf., Mater. Sci. Eng.*, 2018, pp. 1–5.

[27] J. Skorupski, "The simulation-fuzzy method of assessing the risk of air traffic accidents using the fuzzy risk matrix," *Saf. Sci.*, vol. 88, pp. 76–87, Oct. 2016.

[28] Y. G. He and L. F. Zhao, "Strategic risk assessment of power generation enterprises based on risk matrix," *Electr. Power Sci. Eng.*, vol. 31, no. 7, pp. 25–30, Jul. 2015.

[29] Q. Sun, Y. Zhang, D. Han, Z. Yan, and J. Zhao, "Multi-elements and multi-dimensions risk evaluation of smart grid," in *Proc. IEEE PES Innov. Smart Grid Technol.*, May 2012, pp. 1–6.

[30] Y. Duan, J. Zhao, J. Chen, and G. Bai, "A risk matrix analysis method based on potential risk influence: A case study on cryogenic liquid hydrogen filling system," *Process Saf. Environ. Protection*, vol. 102, pp. 277–287, Jul. 2016.

[31] H. M. Yang, Z. S. Pan, and W. Bai, "Review of time series prediction methods," *Comput. Sci.*, vol. 46, no. 1, pp. 21–29, Jan. 2019.

[32] F. Guo, B. Wang, and M. Liu, "Time series prediction research based on BP neural network," *Value Eng.*, vol. 12, pp. 128–129, 2010.

[33] H. Hamdi, C. B. Regaya, and A. Zaafouri, "A sliding-neural network control of induction-motor-pump supplied by photovoltaic generator," *Protection Control Modern Power Syst.*, vol. 5, no. 1, pp. 306–332, Dec. 2020.

[34] W. H. Yang, "Model of risk evaluation and aversion method of urban power network planning," Ph.D. dissertation, Dept. Bus. Admin. Eng., North China Electr. Power Univ., Beijing, China, 2010.

[35] G. Zhao, "A study on enterprise risk management of provincial grid company," Ph.D. dissertation, Dept. Technol. Econ. Manage. Eng., Wuhan Univ., Wuhan, China, 2012.

[36] X. T. Liu, "Study on data normalization in BP neural network," *Mech. Eng. Automat.*, vol. 3, pp. 122–124, Jun. 2016.

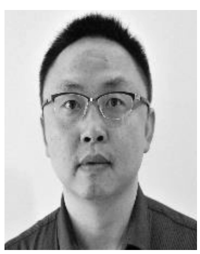
[37] C. Q. Kang, Q. Xia, N. D. Xiang, and L. C. Bai, "Sequence operation theory and its applications," *Autom. Electr. Power Syst.*, vol. 26, no. 17, pp. 6–11, Sep. 2002.

[38] C. Q. Kang, G. F. Yang, and Q. Xia, "Analysis of the uncertainty of electric power demand," *Autom. Electr. Power Syst.*, vol. 29, no. 17, p. 14G19, Sep. 2005.

[39] L. Anthony TonyCox, "What's wrong with risk matrices?" *Risk Anal.*, vol. 28, no. 2, pp. 497–512, Apr. 2008.

SHUJUAN CHEN was born in Shaoyang, Hunan, China, in 1996. She received the B.E. degree in electrical engineering from the Changsha University of Science and Technology, Changsha, China, in 2018. She is currently pursuing the M.S. degree with the College of Electrical and Information Engineering, Hunan University, Changsha. Her major research interests include data mining and optimization.

QIN JIANG (Student Member, IEEE) was born in Hengyang, Hunan, China, in 1995. She received the B.E. degree in electrical engineering from the Changsha University of Science and Technology, Changsha, China, in 2018. She is currently pursuing the M.S. degree with the College of Electrical and Information Engineering, Hunan University, Changsha. Her major research interests include data mining and optimization.



YUQING HE was born in Hunan, China, in 1982. He received the B.S., M.S., and Ph.D. degrees from Hunan University, Hunan, China, in 2005, 2008, and 2011, respectively, all in electrical engineering. He has been with Hunan University, since 2001. He is currently with the State Grid Hunan Electric Power Corporation Economical and Technical Research Institute. His interests include power system economics, power system reliability, and operation.

RUANMING HUANG was born in 1980. He received the M.S. degree in electrical engineering from Southeast University, Nanjing, China, in 2007. He is currently with the State Grid Shanghai Economic Research Institute. His main research directions are grid planning and power system analysis.

JIAYONG LI (Member, IEEE) received the B.Eng. degree from Zhejiang University, Hangzhou, China, in 2014, and the Ph.D. degree from The Hong Kong Polytechnic University, Hong Kong, in 2018. He is currently an Assistant Professor with the College of Electrical and Information Engineering, Hunan University, Changsha, China. He was a Postdoctoral Research Fellow with The Hong Kong Polytechnic University and a Visiting Scholar with Argonne National Laboratory, Argonne, IL, USA. His research interests include power economics, energy management, distribution system planning and operation, renewable energy integration, and demand-side energy management.

CAN LI was born in Kaifeng, Henan, China, in 1996. She received the B.E. degree in electrical engineering from Zhengzhou University, Zhengzhou, China, in 2018. She is currently pursuing the M.S. degree with the College of Electrical and Information Engineering, Hunan University, Changsha, China. Her major research interests include data mining and optimization.

JING LIAO was born in Loudi, Hunan, China, in 1997. She received the B.E. degree in building electrical and intelligence from Xiangtan University, Xiangtan, China, in 2018. She is currently pursuing the M.S. degree with the College of Electrical and Information Engineering, Hunan University, Changsha, Her main research interests are power system planning and optimization, and new energy consumption.