Comprehensive Evaluation of Electric Power Prediction Models Based on D-S Evidence Theory Combined with Multiple Accuracy Indicators

Qiong Cui, Jizhong Zhu, Jie Shu, Lei Huang, and Zetao Ma

***Abstract*—A comprehensive evaluation method of electric power prediction models using multiple accuracy indicators is proposed. To obtain the preferred models, this paper selects a number of accuracy indicators that can reflect the accuracy of single-point prediction and the correlation of predicted data, and carries out a comprehensive evaluation. First, according to Dempster-Shafer (D-S) evidence theory, a new accuracy indicator based on the relative error (RE) is proposed to solve the problem that RE is inconsistent with other indicators in the quantity of evaluation values and cannot be adopted at the same time. Next, a new dimensionless method is proposed, which combines the efficiency coefficient method with the extreme value method to unify the accuracy indicator into a dimensionless positive indicator, to avoid the conflict between pieces of evidence caused by the minimum value of zero. On this basis, the evidence fusion is used to obtain the comprehensive evaluation value of each model. Then, the principle and the process of consistency checking of the proposed method using the entropy method and the linear combination formula are described. Finally, the effectiveness and the superiority of the proposed method are validated by an illustrative instance.**

***Index Terms*—Dempster-Shafer (D-S) evidence theory, multiple accuracy indicators, electric power prediction model, comprehensive evaluation.**

I. Introduction

ACCURATE electric power prediction is the basis and premise of power planning and design, and is also an important guarantee for the safe and economic operation of power grids. Electric power prediction mainly include wind power forecasting, photovoltaic power forecasting and load forecasting. At present, there are many methods for electric power prediction. For example, some methods are based on statistical models: grey linear regression model (GLRM), autoregressive moving average model, Markov model, etc. Also, some methods are based on artificial intelligence: wavelet neural network (WNN), least squares support vector machine (LSSVM), fuzzy prediction method, etc. [1]-[4]. To improve the accuracy of electric power prediction and obtain the preferred models, it is necessary to evaluate the prediction accuracy of the models first.

At present, many references such as [5], [6], have evaluated the prediction accuracy of the models, which normalize the accuracy indicators of the evaluation prediction model, multiply the coordination factors of different indicators, and finally integrate them into a comprehensive indicator to evaluate the prediction accuracy of the model. Reference [7] evaluates the effect of each prediction model with three accuracy indicators: mean absolute error (MAE), mean absolute percentage error (MAPE), and sum of squared error (SSE). Reference [8] adopts a single indicator, i. e., relative error (RE), to evaluate and combine different models. Reference [9] measures MAPE in model parameter identification, and then combines and evaluates the models based on the variable weight coefficients of the optimization algorithm. Most of the commonly used screening methods of prediction models are carried out through a single accuracy indicator, i. e., the single indicator value of each model is calculated, and the model whose indicator value exceeds the threshold value is removed to complete the screening process. In fact, any single indicator cannot fully reflect the effectiveness of the prediction models [10], [11]. Different indicators can reveal the generation mechanism and characteristics of the prediction accuracy of the models from different perspectives. If the models are screened by a single indicator, some useful data and information may be lost. Therefore, this paper proposes a comprehensive evaluation method of electric power prediction models based on multiple accuracy indicators to illuminate the prediction effect of the models from multiple aspects.

Dempster-Shafer (D-S) evidence theory was first proposed by Dempster and further promoted and developed by Shafer. D-S evidence theory deals with uncertain problems [12]. It has the advantage of integrating multiple types of random and fuzzy information and can optimize the decision-making

Manuscript received: July 23, 2020; revised: October 27, 2020; accepted: January 1, 2021. Date of CrossCheck: January 1, 2021. Date of online publication: May 17, 2021.

This work was supported by National Key R&D Program of China (No. 2016YFB0901405), Guangdong Provincial Science and Technology Planning Project of China (No. 2020A0505100004, No. 2018A050506069), and Guangdong Provincial Special Fund Project for Marine Economic Development of China (No. GDNRC[2020]020).

Q. Cui and J. Zhu (corresponding author) are with the School of Electric Power, South China University of Technology, Guangzhou, China, and Q. Cui is also with the Key Laboratory of Renewable Energy, Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, Guangzhou, China (e-mail: ciqg0716@163.com; zhujz@scut.edu.cn).

J. Shu, L. Huang, and Z. Ma are with the Key Laboratory of Renewable Energy, Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, Guangzhou, China (e-mail: shujie@ms. giec. ac. cn; huanglei@ms. giec. ac. cn; mazt@ms.giec.ac.cn).

DOI: 10.35833/MPCE.2020.000470

scheme [13]-[16]. This paper uses D-S evidence theory for reference, combines independent evaluation results of the prediction models by multiple accuracy indicators, and proposes a new method for comprehensive evaluation of the models based on D-S evidence theory combined with multiple indicators.

II. Characteristics of Accuracy Indicators

At present, the commonly used accuracy indicators are mainly divided into two categories: single-point prediction accuracy indicators and predicted data correlation indicators. The former includes RE, MAE, MAPE, root mean square error (RMSE), and standard deviation of error (SDE). The latter includes relative coefficient (RC) and forecasting effective measure (FEM) [17].

Among the above indicators, RE is based on the individual sample, and the number of its evaluation values is that of samples. Other indicators are based on the total sample and each has a single evaluation value. The indicators of MAE and MAPE are error measurement tools based on the idea of point-by-point summation and then averaging, which reflect the overall average performance of the prediction models. RMSE measures the dispersion of the deviation between the predicted power and the actual power. SDE reflects the prediction accuracy and the discretization level of the error, which is independent of the average value of the prediction error. RC reflects the correlation level between the predicted value and the actual value, and can be used to analyze the errors caused by uncertain factors. The value of FEM reflects the effectiveness of the prediction method, and it describes the mean value of the prediction accuracy and the mean variance of the dispersion degree. FEM can reflect the fitting situation between the predicted sequence and the actual sequence.

When a single accuracy indicator is used to evaluate the electric power prediction models, it is difficult to fully reflect the prediction effect of the models. The single-point prediction accuracy indicator only reflects the single-point error of the prediction models, which can indicate the prediction accuracy of the models to a certain extent, but cannot accurately illuminate the trend of the prediction errors. While the single-point error values of two models may be the same, the curves of the prediction errors may be very different. Therefore, it is necessary to evaluate the models with the predicted data correlation indicators. Considering the average value of the prediction errors, the level of discretization, the correlation between the predicted values and the actual values, and other factors, the indicators that can reflect the different prediction conditions of the models from different aspects are selected to independently or comprehensively evaluate the prediction models. When different indicators are used to independently evaluate the prediction models, the evaluation results may change or even conflict with each other, which causes difficulties in exploring the change rules of the prediction accuracy of the models. Meanwhile, it brings difficulties to the optimal selection or combination of the prediction models. Therefore, it is a problem to be solved how to use the independent evaluation results of different indicators to comprehensively evaluate the model.

III. D-S Evidence Theory

The evidence theory can integrate information from multiple evidence sources and provide correct analysis and decision-making.

A. Relevant Definitions of Evidence Theory

For set C, if set A satisfies: $A=\{B|B\subseteq C\}$, then the power set of A is C. For a discriminant problem, it corresponds to a subset of C. Set C represents all possible results that can be recognized, so set C is called a frame of discernment, recorded as Ω.

If $m(C)\in[0,1]$, $m(\varnothing)=0$ and $\sum_{C\subset\Omega} m(C)=1$, then m is the basic probability assignment in Ω. The probability assignment reflects the size of reliability. $m(C)$ is the basic probability number for C, where $\forall C\subset\Omega$. If $m(C)>0$, then set C of all elements that satisfy this condition is called the focal elements of m. If $Bel(C)=\sum_{A\subset C} m(B)$ and $\forall C\subset\Omega$, then Bel is called the belief function in Ω. The belief function reflects the degree of reliability, and it follows that $Bel(\varnothing)=0$ and $Bel(\Omega)=1$.

B. Belief Functions Fused by Using Dempster's Rule

Dempster's rule is an important part of the evidence theory, which reflects the joint effect of pieces of evidence. We use it to calculate a new belief function, marked as Bel_{new}, which is generated by the combined action of n pieces of evidence. Bel_{new} is called the direct sum of the pre-synthesis belief functions, expressed by $Bel_{\text{new}}=Bel_1\oplus Bel_2\oplus\ldots\oplus Bel_n$. Assume that Bel_1 and Bel_2 are the belief functions based on two independent pieces of evidence in the same Ω. m_1 and m_2 are the corresponding basic probability assignments, and the number of focal elements is N_{ele}. Suppose the focal elements corresponding to m_1 are $E_{1,1}, E_{1,2}, \ldots, E_{1,N_{\text{ele}}}$, and the focal elements corresponding to m_2 are $E_{2,1}, E_{2,2}, \ldots, E_{2,N_{\text{ele}}}$, which satisfy $\sum_{E_{1,p}\cap E_{2,q}=\varnothing} m_1(E_{1,p})m_2(E_{2,q})<1$, where $p,q=1, 2, \ldots, N_{\text{ele}}$. Then the basic probability assignment of the composition can be expressed as:

$$m_{1\to 2}(E)=\begin{cases} 0 & E=\varnothing \\ \dfrac{\sum\limits_{E_{1,p}\cap E_{2,q}=E} m_1(E_{1,p})m_2(E_{2,q})}{1-\sum\limits_{E_{1,p}\cap E_{2,q}=\varnothing} m_1(E_{1,p})m_2(E_{2,q})} & E\neq\varnothing \end{cases} \tag{1}$$

where E is the focus element, and it satisfies $E_{1,p}\cap E_{2,q}=E$.

Equation (1) is the expression of the combination of Bel_1 and Bel_2, i.e., $Bel_1\oplus Bel_2$. We can use the same algorithm to continue the synthesis with the next belief function until all the belief functions are fused.

IV. Comprehensive Evaluation Method of Electric Power Prediction Models

To comprehensively evaluate the electric power prediction models, this paper first considers the average value of the prediction errors, the level of discretization, and the correlation between the predicted values and the actual values, etc. Then, seven indicators that can reflect the prediction situation of the models are selected. They are RE, MAE, MAPE, RMSE, SDE, and two other indicators that can reflect the correlation between the predicted values and the actual values, i.e., FEM and RC. All or a few of them can be selected as required, and the proposed method is still applicable. The comprehensive evaluation process of the electric power prediction models is shown in Fig. 1.

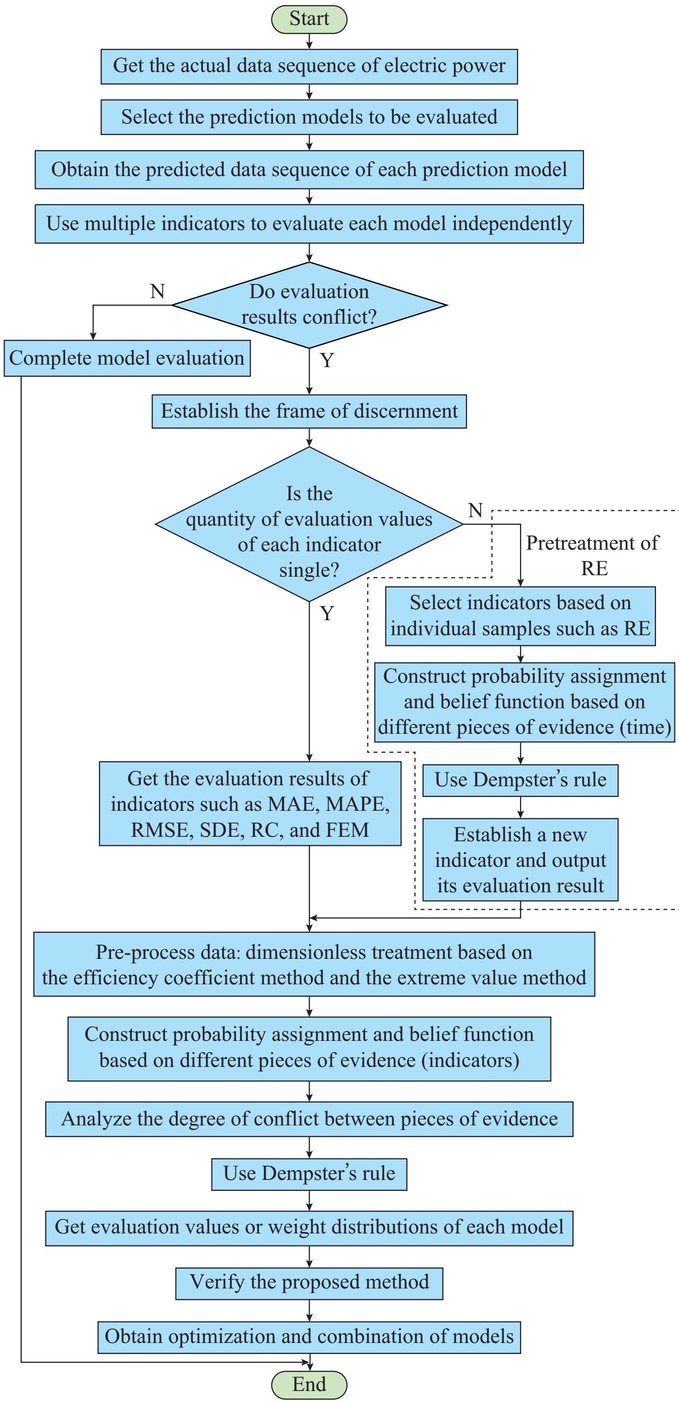

Fig. 1. Comprehensive evaluation process of electric power prediction models based on multiple accuracy indicators.

Based on the actual data of electric power, the above indicators are used to independently evaluate the electric power prediction models. If there are no conflicts between the evaluation results, the model evaluation is completed. If they are conflicted, D-S evidence theory is used to fuse multiple indicators for the comprehensive evaluation of the models, and the steps are as follows.

A. Establishment of Frame of Discernment

N_{ind} and N_{mod} represent the sample numbers of indicators and electric power prediction models, respectively. The sets of evaluation results of N_{mod} models by N_{ind} precision indicators are $Y_1, Y_2, \ldots, Y_{N_{\text{mod}}}$, where $Y_j=\{y_{1,j}, y_{2,j}, \ldots, y_{N_{\text{ind}},j}\}$ and $j=1, 2, \ldots, N_{\text{mod}}$. Then for set $\{Y_1, Y_2, \ldots, Y_{N_{\text{mod}}}\}$, the number of its power sets is $2^{N_{\text{mod}}}$, where the nonempty subsets are $\{Y_1\}, \{Y_2\}, \ldots, \{Y_{N_{\text{mod}}}\}$. The frame of discernment can be expressed by $\Omega=\{\{Y_1\}, \{Y_2\}, \ldots, \{Y_{N_{\text{mod}}}\}\}$. The evaluation of each indicator of the model is taken as evidence information, and the quantity of the evaluation values of each model may vary due to different indicators. In the same sample number of time intervals, MAE, MAPE, RMSE, SDE, FEM, and RC all have a single evaluation value for each model. Note that RE is an exception. If the set of selected indicators contains RE, RE should be pre-processed.

B. Pre-processing of RE

To address the problem that RE and other indicators cannot be compared and adopted at the same time due to the inconsistent quantity of evaluation values for the same model with the same sample number of time intervals, RE needs to be pre-processed, but its values cannot be simply averaged. The reasons are analyzed through Fig. 2 which shows the values of RE for different models.

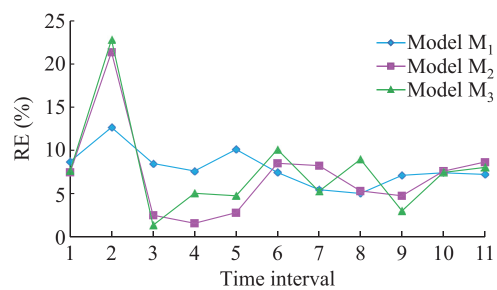

Fig. 2. Values of RE for different models.

As shown in Fig. 2, the average values of RE for the three prediction models of M_1, M_2 and M_3 are very close, which are 7.9%, 7.1% and 7.6%, respectively. However, the model with a small average value may have the largest value of RE. For example, the maximum value of RE for M_3 is 22.8%, the maximum value of RE for M_2 is 21.3%, and the average value of the RE for M_1 is the maximum. To dig out more useful information about RE and enable it to comprehensively evaluate the models by combining other indicators, the values of RE at each time interval are fused by the evidence theory, i.e., a new indicator is constructed. The process is as follows.

1) Establish the frame of discernment of RE.

The frame of discernment of RE can be expressed by $\Omega'=\{\{R_1\}, \{R_2\}, \ldots, \{R_{N_{\text{mod}}}\}\}$, where R_j is the power set of Ω', $R_j=\{R_{1,j}, R_{2,j}, \ldots, R_{N_{\text{tim}},j}\}$, $R_{t,j}$ ($t=1, 2, \ldots, N_{\text{tim}}$, $j=1, 2, \ldots, N_{\text{mod}}$) is the value of RE of model j at time interval t, and N_{tim} is the sample number of time intervals.

2) Construct the basic probability assignment and the belief function in Ω'.

$m'_t(R_j)$ represents the basic probability number about R_j, and it is obtained by the normalization of $R_{t,j}$, where m'_t represents the basic probability assignment at time interval t in Ω', $m'_t(R_j)\in[0, 1]$, and $\sum_{R_j\subset\Omega'} m'_t(R_j)=1$. According to the definition of the belief function, Bel'_t corresponds to m'_t one to one at time interval t, where Bel'_t represents the belief function at time interval t in Ω', $Bel'(\varnothing)=0$, and $Bel'(\Omega')=1$.

3) Generate a new accuracy indicator.

The belief functions are fused two by two based on Dempster's rule. Taking Bel'_1 and Bel'_2 as examples, the equation can be expressed by $Bel'_{1\to 2}=Bel'_1\oplus Bel'_2$, where Bel'_1 and Bel'_2 are two belief functions of independent pieces of evidence in Ω'. m'_1 and m'_2 are the basic probability assignments corresponding to Bel'_1 and Bel'_2, respectively. For $\forall R_j\subset\Omega'$, there is $m'_t(R_j)>0$, so R_j is the set of focal elements of m' in Ω'. The number of focal elements is N_{mod}. $m'_{1\to 2}(R)$ is the basic probability assignment corresponding to $Bel'_{1\to 2}$, where R is the focal element. According to Dempster's rule, the two functions of $Bel'_{1\to 2}$ and the next belief function continue to be combined until the fusion of N_{tim} belief functions is completed. The resultant belief function is $Bel'_{1\to N_{\text{tim}}}$, where $Bel'_{1\to N_{\text{tim}}}=Bel'_1\oplus Bel'_2\oplus\ldots\oplus Bel'_{N_{\text{tim}}}$, and the corresponding basic probability assignment $m'_{1\to N_{\text{tim}}}(R)$ is the value of the new indicator.

C. Dimensionless Treatment of Accuracy Indicators

In the comprehensive evaluation, there may be differences in positive and negative type inconsistencies and dimensionality inconsistencies among various indicators. Therefore, to eliminate the influence of these differences, dimensionless processing is required for the indicators. The common treatment methods include summation standardization treatment, standard deviation treatment, extreme value method, and efficacy coefficient method [18]. For RE, MAE, MAPE, RMSE or SDE, the smaller the value is, the better the model evaluation will be, i.e., these are negative indicators. For FEM or RC, the larger the value is, the better the model evaluation will be, i.e., these are positive indicators. Because the summation standardization method and the standard deviation method cannot unify the indicators into the same type of positive indicators or negative indicators, they are not applicable to D-S evidence theory. Although some indicators that have become dimensionless through the extreme value method can be unified as the same type of positive and negative indicators, some of the focal element values are zero because the minimum value is zero, leading to a greater conflict between pieces of evidence. Therefore, Dempster's rule is no longer applicable. Although the efficiency coefficient method can avoid the situation in which the minimum value is zero by adjusting the coefficient value, it cannot unify the positive and negative indicators. Therefore, this paper proposes a method combining the efficiency coefficient method and the extreme value method to conduct dimensionless treatment of indicators. This method can unify positive and negative indicators into the same type, and avoid the situation in which the minimum value being zero, so that it is applicable to the evidence theory. The calculation process can be expressed by the following steps.

1) Construct the data matrix of indicator values used to evaluate the prediction models. The data matrix can be expressed as:

$$\boldsymbol{Y}=(y_{i,j})_{N_{\text{ind}}\times N_{\text{mod}}} \tag{2}$$

where $y_{i,j}$ ($i=1, 2, \ldots, N_{\text{ind}}$, $j=1, 2, \ldots, N_{\text{mod}}$) is the value of indicator i of electric power prediction model j.

2) For the indicators of RE, MAE, MAPE, RMSE, SDE, and other negative indicators, the calculation equation can be expressed as:

$$x_{i,j}=\frac{M_{\max,i}-y_{i,j}}{M_{\max,i}-M_{\min,i}}\alpha+(1-\alpha) \tag{3}$$

where $M_{\max,i}$ and $M_{\min,i}$ are the maximum and minimum values of the models evaluated by indicator i, respectively; and α is the efficiency coefficient.

For the positive indicators such as FEM and RC, the calculation equation can be expressed as:

$$x_{i,j}=\frac{y_{i,j}-M_{\min,i}}{M_{\max,i}-M_{\min,i}}\alpha+(1-\alpha) \tag{4}$$

where $x_{i,j}$ is the value of $y_{i,j}$ after dimensionless treatment by combining the efficiency coefficient method with the extreme value method; and $0<\alpha<1$.

3) As α grows, the influence of $(M_{\max,i}-y_{i,j})/(M_{\max,i}-M_{\min,i})$ in (3) grows, while the influence of translation amplitude $1-\alpha$ diminishes. When $\alpha=1$, the corresponding method is the extreme value method.

D. Comprehensive Evaluation of Electric Power Prediction Models

$x_{i,j}$ ($i=1, 2, \ldots, N_{\text{ind}}$, $j=1, 2, \ldots, N_{\text{mod}}$) represents the basic probability number in Ω. The basic probability numbers corresponding to m_i are $x_{i,1}$, $x_{i,2}$, $\ldots$, $x_{i,N_{\text{mod}}}$. According to the definition of the belief function, Bel_i corresponds to m_i one to one. Using Dempster's rule, the belief functions are fused to obtain the evaluation values of each model. According to the value of z_j ($j=1, 2, \ldots, N_{\text{mod}}$), where z_j is the comprehensive evaluation value of model j by the proposed method, the models are sorted, or they are used as the model weight coefficients for the optimal combination of models to improve the prediction accuracy. Taking the combination of N_{mod} models as an example, the weight of model j can be expressed as:

$$w_{M_j}=\frac{z_j}{z_1+z_2+\ldots+z_{N_{\text{mod}}}} \tag{5}$$

E. Validation Method

1) Analysis on Degree of Conflict Between Pieces of Evidence

The degree of conflict between pieces of evidence affects whether the fusion results obtained by Dempster's rule deviate from or even contradict the actual situation. The classical conflict coefficient in the evidence theory reflects the degree of non-mutual inclusion between focal elements, which is denoted by K_{e}. According to (1), K_{e} is expressed as:

$$K_{\text{e}}=\sum_{E_{1,p}\cap E_{2,q}=\varnothing} m_1(E_{1,p})m_2(E_{2,q}) \tag{6}$$

It is not enough to measure the conflict between pieces of evidence considering only non-inclusiveness, which cannot reflect the actual degree of conflict, but we must also consider the differences between pieces of evidence. To reflect the differences between pieces of evidence, we take the distance proposed in [19] for measuring the similarity between sets of basic belief assignments [19], [20].

If $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ are two pieces of independent evidence in Ω and the pieces of evidence are represented as vectors in space, the distance proposed in [19] between $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ can be expressed as:

$$D(\boldsymbol{m}_1, \boldsymbol{m}_2)=\sqrt{\frac{\left\|\boldsymbol{m}_1\right\|^2+\left\|\boldsymbol{m}_2\right\|^2-2\left\langle\boldsymbol{m}_1, \boldsymbol{m}_2\right\rangle}{2}} \tag{7}$$

where $\left\|\boldsymbol{m}_1\right\|^2=\left\langle\boldsymbol{m}_1, \boldsymbol{m}_1\right\rangle$; $\left\|\boldsymbol{m}_2\right\|^2=\left\langle\boldsymbol{m}_2, \boldsymbol{m}_2\right\rangle$; and $\left\langle\boldsymbol{m}_1, \boldsymbol{m}_2\right\rangle$ is the inner product of two vectors, which can be defined as:

$$\left\langle\boldsymbol{m}_1, \boldsymbol{m}_2\right\rangle=\sum_{p=1}^{N_{\text{ele}}}\sum_{q=1}^{N_{\text{ele}}}\boldsymbol{m}_1(E_{1,p})\boldsymbol{m}_2(E_{2,q})\frac{\left|E_{1,p}\cap E_{2,q}\right|}{\left|E_{1,p}\cup E_{2,q}\right|} \tag{8}$$

where $E_{1,p}, E_{2,q}\in\Omega$.

For K_{e} and D, only if both of them are large, can we determine that there is a great conflict between pieces of evidence. If one of them is relatively small, the conflict between the two pieces of evidence can be considered minor [21].

2) Verification Based on Entropy Method and Linear Combination Formula

This paper proposes a method to evaluate prediction models based on multiple indicators, which cannot be verified by a single indicator. The entropy method [22]-[24] is also a measure of uncertainty, which is an objective weighting method to determine the weight according to the difference of the ordering degree of the information contained in each indicator. For a certain indicator, the greater the variation of its value is, the smaller the information entropy is, then the larger the weight coefficient is, and vice versa. Therefore, we can use the information entropy of each indicator to calculate the weights of indicators, and use the weight values to rank the validity of all indicators to verify the validity of the indicator constructed in this paper. Then, the dimensionless indicator values are used to calculate the comprehensive evaluation value of each model by using the linear combination formula [25] to sort the models, and verify the proposed comprehensive evaluation method. However, when the amount of test sample data of indicators increases, the model evaluation based on the entropy method and the linear combination formula needs to be overturned, and the results need to be recalculated, which is a tedious process, and is the disadvantage of the entropy method.

The calculation process of the weight and model evaluation values by using the entropy method is as follows.

1) The matrix for evaluation values of the indicators of the models after dimensionless treatment based on the efficiency coefficient method and the extreme value method can be expressed as:

$$\boldsymbol{X}=(x_{i,j})_{N_{\text{ind}}\times N_{\text{mod}}} \tag{9}$$

2) The information entropy value of the indicators and the avail value, which can reflect the difference degree of the indicators, are calculated. According to the definition of entropy, they can be expressed as:

$$e_i=-\frac{1}{\ln N_{\text{mod}}}\sum_{j=1}^{N_{\text{mod}}}\frac{x_{i,j}}{\sum_{j=1}^{N_{\text{mod}}}x_{i,j}}\ln\frac{x_{i,j}}{\sum_{j=1}^{N_{\text{mod}}}x_{i,j}} \tag{10}$$

$$d_i=1-e_i \tag{11}$$

where e_i and d_i are the information entropy value and the avail value of indicator i, respectively.

The greater the difference of $x_{i,j}$ is, the smaller the value of e_i is; and indicator i plays a more important role in the evaluation of the model; i.e., the larger the value of d_i is, the more important the evaluation of the indicator is.

3) The weight of indicator i can be expressed as:

$$w_i=\frac{d_i}{\sum_{i=1}^{N_{\text{ind}}}d_i} \tag{12}$$

4) For the sake of simplicity and without loss of generality, the comprehensive evaluation value of model j is calculated by using the linear combination method, which is expressed as:

$$v_j=x_{i,j}w_i \tag{13}$$

where v_j is the comprehensive evaluation value of model j.

5) The models are sorted or optimally combined according to the value of v_j to verify the proposed method. The weights of the optimal combination of the models are calculated by (5).

V. Instance Analysis

The following three prediction models are adopted: WNN (Mod_1), GLRM (Mod_2), and LSSVM (Mod_3). Taking the historical load data of a certain region in 6 months as the training sample data and the load data at 9 a.m. of the following 11 days as the test sample data, the predicted and actual values are compared, and the above three models are evaluated with seven indicators. The set of $R_{t,j}$ ($t=1, 2, \ldots, 11; j=1, 2, 3$) represents the values of RE for the three models, and the set of the values of RE for each model can be expressed by $R_j=\{R_{1,j}, R_{2,j}, \ldots, R_{11,j}\}$. The specific values are shown in Table I.

TABLE I
Values of RE for Prediction Models

Time interval t	$R_{t,j}$ (%)		
	Mod_1 ($j=1$)	Mod_2 ($j=2$)	Mod_3 ($j=3$)
1	8.696667	7.4548928	7.6395670
2	12.652160	21.3784764	22.8073358
3	8.470515	2.4901021	1.3106897
4	7.601239	1.5578847	5.0190567
5	10.110100	2.7698618	4.7236615
6	7.446321	8.4928080	10.0979779
7	5.441113	8.2333768	5.2564103
8	5.010369	5.2903360	8.9485691
9	7.099335	4.7285990	2.9736108
10	7.391576	7.5834028	7.4457882
11	7.217011	8.6366479	8.0195956

The set of $y'_{i,j}$ ($i=1, 2, \ldots, 6; j=1, 2, 3$) represents the values of the indicators of MAE, MAPE, RMSE, SDE, FEM, and RC for the three models, and the set of the values of the six indicators for each model can be expressed by $\{Y'_j\}=\{y'_{1,j}, y'_{2,j}, \ldots, y'_{6,j}\}$. The specific values are shown in Table II.

TABLE II
EVALUATION VALUES OF OTHER PRECISION INDICATORS EXCEPT FOR RE

Prediction model	Set	$y'_{i,j}$					
		MAE (kW) (i=1)	MAPE (%) (i=2)	RMSE (kW) (i=3)	SDE (kW) (i=4)	FEM (%) (i=5)	RC (%) (i=6)
Mod_1 ($j=1$)	Y'_1	8.6966670	12.6521600	8.4705150	7.6012390	10.1101000	7.4463210
Mod_2 ($j=2$)	Y'_2	7.4548928	21.3784764	2.4901021	1.5578847	2.7698618	8.4928080
Mod_3 ($j=3$)	Y'_3	7.6395670	22.8073358	1.3106897	5.0190567	4.7236615	10.0979779

It can be concluded from Table I and Table II that RE is different from other indicators in the number of evaluation values of the samples, and different indicators are inconsistent in the evaluation of the models, as shown in Fig. 3.

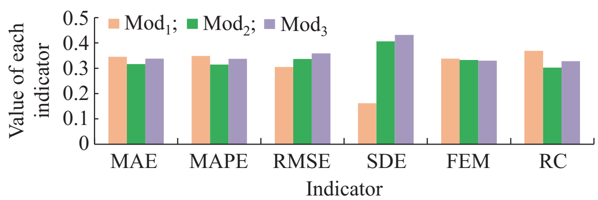

Fig. 3. Inconsistency of different indicators for prediction models (all indicators are dimensionless).

Figure 3 shows that the values of both MAE and MAPE for Mod_1 are greater than those of the other two models, indicating that the prediction accuracy of Mod_1 is worse than those of Mod_2 and Mod_3. The evaluation values of SDE, FEM, and RC indicate that the prediction accuracy of Mod_1 is better than those of the other two models. It is concluded that the independent evaluation values of the indicators for the models are in conflict, so the comprehensive evaluation with multiple indicators for each model is needed.

D-S evidence theory is used to pre-process RE. First, the values of RE are dimensionless. Then, the frame of discernment is established, which can be expressed by $\Omega'=\{\{R'_1\}, \{R'_2\}, \{R'_3\}\}$. Finally, the basic probability assignment is constructed, expressed by m'_t $(t=1, 2, \ldots, 11)$, as shown in Table III.

TABLE III
BASIC PROBABILITY ASSIGNMENT FOR RE

m'_t	Normalized value of RE		
	Mod_1	Mod_2	Mod_3
m'_1	0.365542470	0.313347608	0.32110992
m'_2	0.222600433	0.376130198	0.40126937
m'_3	0.690269997	0.202920700	0.10680930
m'_4	0.536122312	0.109879032	0.35399866
m'_5	0.574319352	0.157346086	0.26833456
m'_6	0.285988791	0.326180945	0.38783026
m'_7	0.287419652	0.434917354	0.27766299
m'_8	0.260288728	0.274833012	0.46487826
m'_9	0.479634730	0.319466589	0.20089868
m'_{10}	0.329675439	0.338231190	0.33209337
m'_{11}	0.302305273	0.361770870	0.33592386

Bel'_t $(t=1, 2, \ldots, 11)$ is gradually fused by using Dempster's rule, which corresponds to m'_t in Table III, and the fusion is completed after 10 steps. The results obtained after the fusion of each step are shown in Table IV.

TABLE IV
FUSION RESULTS OF BASIC PROBABILITY ASSIGNMENT BASED ON RE

$m'_{1\to t}$	Fusion result		
	Mod_1	Mod_2	Mod_3
$m'_{1\to 2}$	0.248018	0.359239039	0.392743
$m'_{1\to 3}$	0.598505	0.254844750	0.146651
$m'_{1\to 4}$	0.800602	0.069867610	0.129530
$m'_{1\to 5}$	0.909503	0.021745321	0.068751
$m'_{1\to 6}$	0.885128	0.024136665	0.090735
$m'_{1\to 7}$	0.876967	0.036186329	0.086847
$m'_{1\to 8}$	0.819377	0.035699228	0.144924
$m'_{1\to 9}$	0.906534	0.026307143	0.067159
$m'_{1\to 10}$	0.905469	0.026958183	0.067572
$m'_{1\to 11}$	0.894011	0.031852780	0.074137

After the pre-processing of RE with D-S evidence theory, a new indicator is obtained, represented as F_{RE}, and its evaluation values for the three models are 0.894011, 0.03185278, 0.074137, respectively. F_{RE} is consistent with other indicators in the number of values for the models. The values of the indicators of F_{RE}, MAE, MAPE RMSE, SDE, FEM and RC for the three models can be expressed by $y_{i,j}$ ($i=1, 2, \ldots, 7; j=1, 2, 3$), and the set of the values of the seven indicators for each model can be expressed by $\{Y_j\}=\{y_{1,j}, y_{2,j}, \ldots, y_{7,j}\}$.

According to (3) and (4), the dimensionless treatment for the indicators of RE, MAE, MAPE, RMSE, SDE, FEM, and RC is conducted.

First, the value of α should be determined, which affects whether the evidence theory is available or not. The trends of K_e and D with α are shown in Fig. 4. The blue and green paths show the changes of K_e and D with α, respectively.

As shown in Fig. 4, when the value of α gradually increases from 0 to 1, both K_e and D increase, and the changing ranges of K_e and D are [0.6667, 0.7070] and [0.0198, 0.5658], respectively, i.e., the rate of change of K_e is less than that of D. When $\alpha=0.5$, the values of K_e and D are 0.6736 and 0.2552, respectively. The smaller K_e and D are, the less conflict there is between pieces of evidence. Therefore, to improve the accuracy of the fusion result obtained by Dempster's rule, the value range of α is set as $0<\alpha<0.5$. If the value of α is adjusted within this range, the rate of change of D is slightly larger, while the rate of change of K_e is not more than 2%.

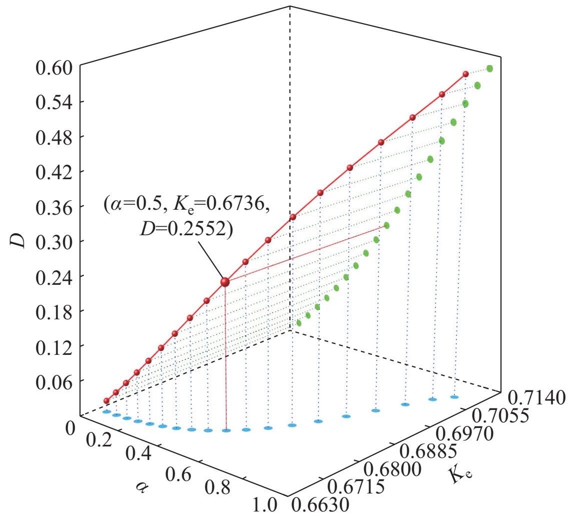

Fig. 4. Trends of K_e and D with α.

For simplicity, we take $\alpha=0.4$ as an example to show the process of dimensionless treatment and fusion of accuracy indicators in detail. The results are shown in Table V.

TABLE V
EVALUATION VALUES OF INDICATORS FOR MODELS AFTER DIMENSIONLESS TREATMENT

Indicator	$x_{i,j}$		
	$Mod_1\,(j=1)$	$Mod_2\,(j=2)$	$Mod_3\,(j=3)$
F_{RE} $(i=1)$	0.6	1.00000000	0.980382250
MAE $(i=2)$	0.6	1.00000000	0.703687436
MAPE $(i=3)$	0.6	1.00000000	0.735917313
RMSE $(i=4)$	1.0	0.76392189	0.600000000
SDE $(i=5)$	1.0	0.63771174	0.600000000
FEM $(i=6)$	1.0	0.72802030	0.600000000
RC $(i=7)$	1.0	0.60000000	0.753042587

The fusions of $x_{i,j}$ $(i=1, 2, \ldots, 7; j=1, 2, 3)$ are conducted by using Dempster's rule. First, each indicator in Table V is normalized to construct m_i. Then, Bel_i corresponding to m_i is gradually fused, and the fusion is completed after six steps. The results obtained after each step of fusion are shown in Table VI.

The comprehensive evaluation values of the three models based on multiple indicators are $\{z_j\}=\{0.422387, 0.416127, 0.161486\}$. Therefore, the ordering of evaluation values of the models is $Mod_1>Mod_2>Mod_3$.

When the selected values of α are different, the corresponding comprehensive evaluation values of the three models are shown in Fig. 5.

It can be seen from Fig. 5 that, according to the comprehensive evaluation values, the models are ranked as $Mod_1>Mod_2>Mod_3$. The comprehensive evaluation values of the first two models are relatively close, and the difference between the former two and the last one is relatively large.

TABLE VI
FUSION RESULTS BASED ON ALL INDICATORS

$m_{1\to i}$	Fusion result		
	Mod_1	Mod_2	Mod_3
$m_{1\to 2}$	0.175619807	0.48783280	0.336547394
$m_{1\to 3}$	0.125312076	0.58014850	0.294539423
$m_{1\to 4}$	0.168153599	0.59470470	0.237141697
$m_{1\to 5}$	0.243810836	0.54988594	0.206303221
$m_{1\to 6}$	0.317494726	0.52131428	0.161190994
$m_{1\to 7}$	0.422387488	0.41612653	0.161485986

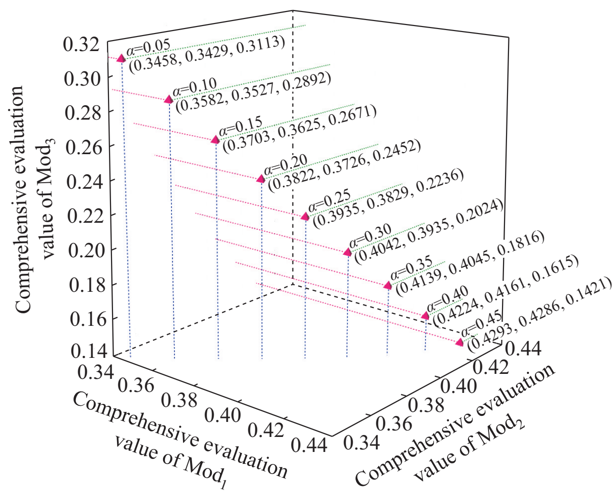

Fig. 5. Comprehensive evaluation values of three models corresponding to different values of α.

The entropy method is used for verification, and the calculation results are shown in Table VII.

TABLE VII
WEIGHTS OF EACH INDICATOR CALCULATED BY ENTROPY METHOD

Indicator	Information entropy value	Avail value	Weight
F_{RE}	0.977819000	0.022181000	0.148028000
MAE	0.978418000	0.021582000	0.144031000
MAPE	0.979617000	0.020383000	0.136032000
RMSE	0.980328000	0.019672000	0.131282000
SDE	0.974527000	0.025473000	0.169999000
FEM	0.979361885	0.020638115	0.137732414
RC	0.980087000	0.019913000	0.132896000

Note: $\alpha=0.4$.

The weights of the indicators of F_{RE}, MAE, MAPE, RMSE, SDE, FEM, and RC calculated by the entropy method are 0.148028000, 0.144031000, 0.136032000, 0.131282000, 0.169999000, 0.137732414, 0.132896000, respectively. The weight of F_{RE} is second only to SDE, which verifies the rationality and effectiveness of the proposed new indicator. According to (13), the weights of the three models are calculated as $\{v_j\}=\{0.828764, 0.816800, 0.710069\}$, so the ordering of weights of the models is $\text{Mod}_1>\text{Mod}_2>\text{Mod}_3$.

When the selected values of α are different, the corresponding weights of the three models calculated by the entropy method are shown in Fig. 6.

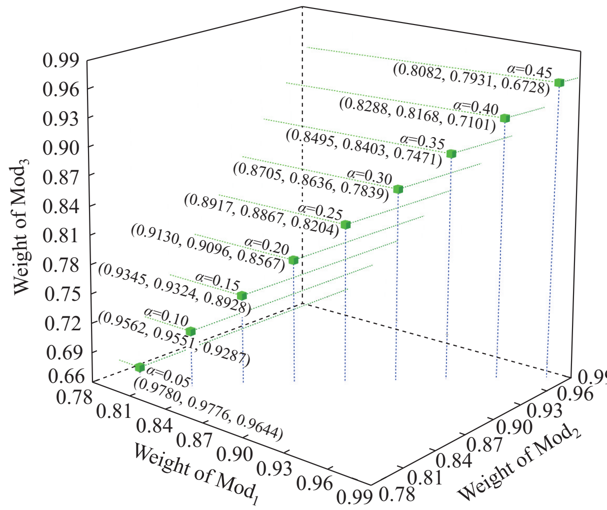

Fig. 6. Weights of three models corresponding to different values of α.

It can be seen from Fig. 6 that, according to the weight values, the models are ranked as $\text{Mod}_1>\text{Mod}_2>\text{Mod}_3$, and the weight values of the first two models are relatively close.

Furthermore, to make the verification clearer, the two models at the top of the ranking are combined. According to (5), the weights of Mod_1 and Mod_2 with different methods are calculated, which are shown in Table VIII.

TABLE VIII
COMPARISON OF WEIGHTS OF TWO MODELS WITH DIFFERENT METHODS

α	Weight (the proposed method)		Weight (the entropy method)	
	Mod_1	Mod_2	Mod_1	Mod_2
0.05	0.502075	0.497925097	0.50010919	0.499891
0.10	0.503867	0.496132977	0.50029262	0.499707
0.15	0.505313	0.494686955	0.50056034	0.499440
0.20	0.506334	0.493665750	0.50092390	0.499076
0.25	0.506832	0.493168223	0.50139658	0.498603
0.30	0.506680	0.493319947	0.50199370	0.498006
0.35	0.505718	0.494282465	0.50273302	0.497267
0.40	0.503733	0.496266633	0.50363519	0.496365
0.45	0.500448	0.499552216	0.50472435	0.495276

As shown in Table VIII, the weights of Mod_1 and Mod_2 calculated by the proposed method differ very little from those calculated by the entropy method. The percentage differences between the weights of the two models calculated by the proposed method and those calculated by the entropy method are both less than 1.1%.

The results calculated by the two methods are consistent whether one model or the combination of two models is preferred, i.e., the correctness and the effectiveness of the proposed method are verified. It should be noted that the ordering of these models is based on the specific instance. Different test sample data in various instances may lead to different orderings of the models, i.e., the ordering for prediction accuracy of power prediction models is related to the test sample data. The proposed method can be used to select the optimal models based on the test sample data which is the predicted value of different models and the actual value in the same instance. When the comprehensive evaluation values of the models differ greatly, the first model is selected for the prediction. When the top j models differ little, the weight coefficients are calculated according to (5), and the combined model is used for prediction.

VI. CONCLUSION

Firstly, this paper uses the idea of the evidence theory for reference and establishes a new accuracy indicator based on RE. Secondly, the weights of indicators are calculated by using the information entropy, and the effectiveness of the new indicator is verified. Then, a new dimensionless processing method that combines the efficiency coefficient method and the extreme value method is proposed. The proposed method not only unifies each indicator into a dimensionless positive indicator, but also avoids the conflict between pieces of evidence caused by the minimum value of zero. On this basis, the values of indicators are fused by using Dempster's rule to obtain the comprehensive evaluation value of each model, and the proposed method is verified by using the entropy method and the linear combination formula. Finally, an example is given to verify the effectiveness of the proposed method. When the quantity of evidence information for indicators increases, the proposed method can directly fuse the new evidence based on the original fusion results, while the result of the model screening method based on the entropy method needs to be recalculated. This flexibility shows that the proposed method has more practical significance.

REFERENCES

[1] T. Liu, "The study on optimization criteria and fusion strategy of wind power forecasting models," M.S. thesis, College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan, China, 2017.

[2] Y. Feng, N. Cui, L. Zhao *et al*., "Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China," *Journal of Hydrology*, vol. 536, pp. 376-383, May 2016.

[3] L. Kumar, S. K. Sripada, A. Sureka *et al*., "Effective fault prediction model developed using least square support vector machine (LSSVM)," *Journal of Systems and Software*, vol. 137, pp. 686-712, Mar. 2018.

[4] S. Atuahene, Y. Bao, Y. Y. Ziggah *et al*., "Short-term electric power forecasting using dual-stage hierarchical wavelet- particle swarm optimization- adaptive neuro-fuzzy inference system PSO-ANFIS approach based on climate change," *Energies*, vol. 11, no. 10, pp. 1-19, Oct. 2018.

[5] M. Xu, Y. Qiao, and Z. Lu, "A comprehensive error evaluation method for short-term wind power prediction," *Automation of Electric Power Systems*, vol. 35, no. 12, pp. 117-125, Jun. 2011.

[6] X. Li, "Research on combined forecasting methods for wind power and its application," M.S. thesis, College of Renewable Energy, North China Electric Power University, Beijing, China, 2016.

[7] L. Dai, S. Huang, K. Huang *et al*., "Combination forecasting model

based on neural networks for wind speed in wind farm," *Proceedings of the CSU-EPSA*, vol. 23, no. 4, pp. 27-31, Aug. 2011.

[8] J. Liu, "Study on short term forecasting method of photovoltaic power generation based on combination model," M.S. thesis, College of Water Conservancy and Hydroelectric Power, Xi'an University of Technology, Xi'an, China, 2017.

[9] Z. Wang, "Research on combined model for short-term load forecasting based on crisscross optimization algorithm," M.S. thesis, College of Automation, Guangdong University of Technology, Guangzhou, China, 2016.

[10] S. Ren, X. Yang, Y. Zhang *et al*., "A real time optimization strategy for microgrid integrated with schedulable ability and uncertainties," *Proceedings of the CSEE*, vol. 37, no. 23, pp. 6866-6877, Dec. 2017.

[11] Y. Xue, C. Yu, J. Zhao *et al*., "A review on short-term and ultra-short-term wind power prediction," *Automation of Electric Power Systems*, vol. 39, no. 6, pp. 141-151, Mar. 2015.

[12] Z. Luo and Y. Deng, "A vector and geometry interpretation of basic probability assignment in Dempster-Shafer theory," *International Journal of Intelligent Systems*, vol. 35, no. 6, pp. 944-962, Jun. 2020.

[13] Q. Jia, L. Shi, N. Wang *et al*., "A fusion method for ground fault line detection in compensated power networks based on evidence theory and information entropy," *Transactions of China Electrotechnical Society*, vol. 27, no. 6, pp. 191-197, Jun. 2012.

[14] Y. Wu, S. Lou, S. Lu *et al*., "The medium and long-term load forecasting based on improved D-S evidential theory," *Transactions of China Electrotechnical Society*, vol. 27, no. 8, pp. 157-162, Aug. 2012.

[15] Q. Sun and W. Yang, "Fusion of gray model load forecasting based on Dempster synthesize rule of evidential theory," *Electric Power Science and Engineering*, vol. 22, no. 4, pp. 4-8, Dec. 2006.

[16] O. Ozdikis, H. Ogurtuzun, and P. Karagoz, "Evidential estimation of event locations in microblogs using the Dempster-Shafer theory," *Information Processing & Management*, vol. 52, no. 6, pp. 1227-1246, Nov. 2016.

[17] F. Wang, "Step-by-step selection method of combined forecasting single item model based on validity," *Statistics and Decision Making*, vol. 408, no. 12, pp. 69-71, Jun. 2014.

[18] X. Zhu and G. Wei, "Discussion on the standard of dimensionless method in entropy method," *Statistics and Decision Making*, vol. 1, no. 2, pp. 12-15, Feb. 2015.

[19] A. L. Jousselme, D. Grenier, and É. Bossé, "A new distance between two bodies of evidence," *Information Fusion*, vol. 2, no. 2, pp. 91-101, Jun. 2001.

[20] F. Xiao, "CED: a distance for complex mass functions," *IEEE Transactions on Neural Networks and Learning Systems*, Apr. 2020. doi: 10.1109/TNNLS.2020.2984918

[21] W. Jiang, J. Peng, and Y. Deng, "New representation method of evidential conflict," *Systems Engineering and Electronics*, vol. 32, no. 3, pp. 562-565, Mar. 2010.

[22] D. Li, B. Xu, and C. Chang, "Entropy-based comprehensive evaluation for the operating conditions of large-scale wind turbine," *Journal of Engineering for Thermal Energy and Power*, vol. 31, no. 4, pp. 1-5, Apr. 2016.

[23] Y. Liu, F. Guo, Z. Wang *et al*., "Research on the spectral characteristics of series arc fault based on information entropy," *Transactions of China Electrotechnical Society*, vol. 30, no. 12, pp. 488-495, Jun. 2015.

[24] J. Wang, W. Pang, L. Wang *et al*., "Synthetic evaluation of steady-state power quality based on combination weighting and principal component projection method," *CSEE Journal of Power and Energy Systems*, vol. 3, no. 2, pp. 160-166, Jun. 2017.

[25] R. Guo, Q. Zheng, J. Liu *et al*., "Ecosystem health assessment of Chongqing City based on entropy and GM (1, 1) model," *China Environmental Science*, vol. 32, no. 6, pp. 1148-1152, Jun. 2012.

Qiong Cui received the B.S. degree in electrical engineering and automation from Chongqing University, Chongqing, China, in 2005, and the M.S. degree in electrical engineering from South China University of Technology, Guangzhou, China, in 2015. She is currently pursuing the Ph.D. degree in South China University of Technology. Currently, she is a Research Assistant in Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences. Her research interests include optimal operation of distributed energy system and load forecasting.

Jizhong Zhu received the B.S., M.S., and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 1985, 1987, and 1990, respectively. His work experience includes Chongqing University, Brunel University, London, UK, National University of Singapore, Singapore, Howard University, Washington, USA, Alstom Grid, Redmond, USA, and Electric Power Research Institute, China Southern Power Grid, Guangzhou, China. He is a Fellow of IEEE and a Professor at School of Electric Power, South China University of Technology, Guangzhou, China. His research interests include power system operation and control as well as renewable energy application.

Jie Shu received the B.S. degree in electric machines from Shenyang University of Technology, Shenyang, China, in 1995, and the M.S. degree in computer application from South China University of Technology, Guangzhou, China, in 2004. He received the Ph.D. degree in operation research and control theory from Sun Yat-sen University, Guangzhou, China, in 2009. He is currently a Professor, Research Fellow, Chief of the Renewable Energy Generation and Microgrid Technology Laboratory in Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, Guangzhou, China. His research interests include power electronics control for renewable energy generation, microgrid technology for distributed renewable energy generation, solar power generation, and building integrated photovoltaic.

Lei Huang received the B.S. and M.S. degrees in electrical engineering from Harbin Institute of Technology, Harbin, China, in 2009 and 2011, respectively. She received the Ph.D. degree in thermal engineering from University of Chinese Academy of Sciences, Beijing, China, in 2020. She is currently a Research Assistant in Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, Guangzhou, China. Her research is on optimal planning of distributed energy system.

Zetao Ma received the B.S., M.S. and Ph.D. degrees in electrical engineering from Harbin Institute of Technology, Harbin, China, in 2011, 2013, and 2020, respectively. He was a Visiting Student with Chalmers University of Technology, Gothenburg, Sweden, in 2017, and a Post Doctor with University of Lille, Lille, France, from 2018 to 2019. He is currently a Research Assistant with the Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, Guangzhou, China. His research interests include design and energy management of electric vehicles, optimization of distributed energy systems and hardware-in-the-loop simulation.