An Improved Fuzzy Method for Characterizing Wind Power

Yue Xiang, Shuai Hu, Junyong Liu, and Rui Wang

***Abstract*—An improved fuzzy method is proposed to derive a fuzzy number for characterizing uncertain wind power. The input measurement data are firstly converted into nested sets, and the fuzzy number is further obtained based on nested set transformation method. Numerical studies have demonstrated the effectiveness and advantages of the improved fuzzy method.**

***Index Terms*—Wind power, fuzzy method, uncertainty, nested set.**

I. Introduction

WITH the increase of wind generation, inaccurate characterizing of wind power may bring large power outage or greatly impact the stable operation and effective planning of penetrated power systems [1], [2], due to its uncertainty. It is essential for integrated energy systems to characterize the uncertainty features of wind power.

Numerous studies have been investigated for characterizing wind power, in which probabilistic and fuzzy methods are the most popular. Probability distribution is the basis for the probabilistic method. However, its application is limited by the effectiveness of generating specific probability distribution with real data and the computation complexity [3]. On the other hand, due to concise engineering linguistic expression based on fuzzy set theory [4], the fuzzy method is easy to be implemented, and mainly focuses on characterizing wind power in this paper. Briefly, the fuzzy set theory depicts the degree of an element $x \in \mathbf{R}$, which belongs to a certain set by a discrete or continuous membership function $\mu(x) \in [0,1]$, rather than an absolute 0 or 1 in the deterministic set theory. A class of special fuzzy set called fuzzy number, satisfying $\forall x_1 < x_2 < x_3 \in \mathbf{R}$, $\mu(x_2) \geq \min(\mu(x_1), \mu(x_3))$, is used to characterize uncertainty in the fuzzy modeling. Several methods are proposed for determining the fuzzy number, among which the classic ones are fuzzy factor based method (FFM) [5] and data clustering based method (DCM) [6]. However, considering the rapid real-time volatility and the proportion growth of penetrated wind power, the subjective selection of factor and partial data discarding in existing fuzzy methods ignores the cumulative information of measurement data distribution.

Thus, in order to improve the accuracy of fuzzy modeling for wind power, an improved fuzzy method is proposed. The result is expressed as a fuzzy number, which is derived based on nested set transformation method. The proposed method combines the advantages of data cumulative characteristic into fuzzy modeling, and provides a more concise and accurate method to derive the fuzzy number for characterizing wind power.

II. Methodology

Wind power, which is naturally generated according to wind speed, etc., varies within a certain range during each scheduling time period considering the limitation of cut-in and cut-out speeds, but the cumulative numbers of wind power value diversely fall in different narrowed ranges (regular set). In other words, the wind power data can be described by a group of nested intervals with different "possibility", i.e., nested set. A nested set consists of a set of regular sets and is marked by the index function, which can be formed by the data with preparation. Considering the great volume of the measurement data of aggregated wind turbines, the set of reduced time series output data set in the n^{th} time period, denoted as S_n, can be constructed at the sample interval, denoted as Δt, and formulated as:

$$S_n = \{X_1, X_2, \ldots, X_i, \ldots, X_{N_{\text{WT}}}\} \quad \forall n = 1, 2, \ldots, N_{\text{C}} \tag{1}$$

$$X_i = \{x_{i1}, x_{i2}, \ldots, x_{ik}, \ldots, x_{iN_{\text{S}}}\} \quad \forall i = 1, 2, \ldots, N_{\text{WT}} \tag{2}$$

where N_{C} is the number of time period ΔT; N_{WT} is the number of aggregated wind turbines; $N_{\text{S}} = \Delta T / \Delta t$ is the sample number; X_i is the time series data set for the i^{th} aggregated wind turbines; and x_{ik} is the k^{th} sampled data for the i^{th} aggregated wind turbines.

Whereas the variables change within a certain range in the n^{th} time period, X_i can be rearranged to form an ordered data set: $X_i' = \{x_{i1}', x_{i2}', \ldots, x_{ik}', \ldots, x_{iN_{\text{s}}}'\}$, satisfying $\forall k_1 < k_2 = 1, 2, \ldots, N_{\text{S}}$, $x_{ik_1}' \leq x_{ik_2}'$. Therefore, S_n can be converted into the set of nested set Θ_n [7], which is expressed as:

Manuscript received: March 18, 2020; accepted: September 3, 2020. Date of CrossCheck: September 3, 2020. Date of online publication: January 25, 2021.

This work was supported by National Natural Science Foundation of China (No. 51807127).

Y. Xiang and J. Liu are with College of Electrical Engineering, Sichuan University, Chengdu 610065, China (e-mail: xiang@scu.edu.cn; liujy@scu.edu.cn).

S. Hu (corresponding author) is with State Grid Ningxia Electric Power Eco-Tech Research Institute, Yinchuan 750004, China (e-mail: shuai.scu@foxmail.com).

R. Wang is with State Grid Hongguozi Power Supply Company, Shizuishan 753204, China (e-mail: elaine1126@outlook.com).

DOI: 10.35833/MPCE.2020.000169

$$\Theta_n = \{\Pi_1, \Pi_2, \ldots, \Pi_i, \ldots, \Pi_{N_{\mathrm{WT}}}\} \quad \forall n = 1, 2, \ldots, N_{\mathrm{C}} \tag{3}$$

$$\Pi_i = \{\mathcal{L}_i(\alpha) \mid \alpha \in [0, 1]\} \quad \forall i = 1, 2, \ldots, N_{\mathrm{WT}} \tag{4}$$

$$\mathcal{L}_i(\alpha): [x'_{ik}, x'_{i(N_{\mathrm{S}}+1-k)}] \quad \forall i = 1, 2, \ldots, N_{\mathrm{WT}} \tag{5}$$

$$[x'_{ik_2}, x'_{i(N_{\mathrm{S}}+1-k_2)}] \subseteq [x'_{ik_1}, x'_{i(N_{\mathrm{S}}+1-k_1)}] \quad \forall k_1 < k_2 \tag{6}$$

$$core = \mathrm{ceil}\left(\frac{N_{\mathrm{S}}}{2}\right) \tag{7}$$

$$\alpha = \frac{k}{core} \quad \forall k = 1, 2, \ldots, core \tag{8}$$

where ceil$(\cdot)$ is the rounding up function; $\mathcal{L}_i(\cdot)$ is the index function of the nested set Π_i; and Π_i is the nested set composed by $\mathcal{L}_i$ corresponding to X_i, which is graphically shown as a group of overlapped colored line segments on the data axis in Fig. 1, each line segment represents the relevant regular set $\mathcal{L}_i(\alpha)$ of the nested set Π_i.

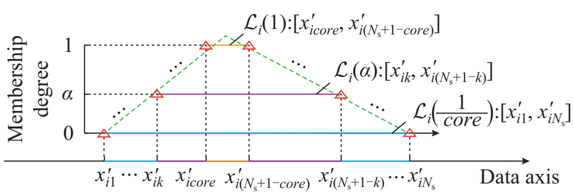

Fig. 1. Illustration for nested set and triangular shape of discrete membership function.

Considering the equivalence of fuzzy set and nested set [8], it is graphically indicated by rearranging each regular set $\mathcal{L}_i(\alpha)$ according to α in Fig. 1. When the nested set is obtained, the discrete membership function $\mu_i(x)$, i.e. the red triangles in Fig. 1, for the i^{th} aggregated wind turbines in the n^{th} time period is transformed as:

$$\mu_i(x) = \vee\left\{\alpha \,\middle|\, x \in \mathcal{L}_i(\alpha)\right\} \quad x \in X_i, \forall i = 1, 2, \ldots, N_{\mathrm{WT}} \tag{9}$$

where $\vee$ is the maximum operator.

After that, due to the formed triangular shape approximates as shown by green dotted line in Fig. 1, and its widespread utilization in electrical engineering field [4], [5], the triangular fuzzy number is utilized as the prototype for further continuous and fitting processing of $\mu_i(x)$. The triangular fuzzy number can be uniquely defined by three parameters: $\tilde{A}_{\mathrm{tria}} = (t^{\mathrm{inf}}, t^{\mathrm{ker}}, t^{\mathrm{sup}})$, where t^{ker} is the most possible value (membership degree is 1) of wind power, and t^{inf} and t^{sup} are the least possible values (membership degree is 0). Its membership function is expressed as:

$$\mu_{\tilde{A}_{\mathrm{tria}}}(x) = \begin{cases} \dfrac{x - t^{\mathrm{inf}}}{t^{\mathrm{ker}} - t^{\mathrm{inf}}} & t^{\mathrm{inf}} \le x < t^{\mathrm{ker}} \\ \dfrac{t^{\mathrm{sup}} - x}{t^{\mathrm{sup}} - t^{\mathrm{ker}}} & t^{\mathrm{ker}} \le x \le t^{\mathrm{sup}} \\ 0 & \text{others} \end{cases} \tag{10}$$

The piecewise linear regression analysis is then used to get the piecewise regression parameters for the left and right parts of (10), denoted as a_i^{I} and b_i^{I}, a_i^{II} and b_i^{II}, respectively.

Finally, the fuzzy number for characterizing uncertain power of the i^{th} aggregated wind turbines in the n^{th} time period can be derived as $\tilde{X}_i = (t_i^{\mathrm{inf}}, t_i^{\mathrm{ker}}, t_i^{\mathrm{sup}})$.

$$t_i^{\mathrm{inf}} = -\frac{b_i^{\mathrm{I}}}{a_i^{\mathrm{I}}} \tag{11}$$

$$t_i^{\mathrm{ker}} = \frac{1}{2}\left(\frac{1-b_i^{\mathrm{I}}}{a_i^{\mathrm{I}}} + \frac{1-b_i^{\mathrm{II}}}{a_i^{\mathrm{II}}}\right) \tag{12}$$

$$t_i^{\mathrm{sup}} = -\frac{b_i^{\mathrm{II}}}{a_i^{\mathrm{II}}} \tag{13}$$

Repeat the process for all the aggregated wind turbines, the set of fuzzy numbers for characterizing uncertain wind power in the n^{th} time period can be obtained as:

$$\tilde{S}_n = \{\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_i, \ldots, \tilde{X}_{N_{\mathrm{WT}}}\} \tag{14}$$

The process during other time periods is similar, and the flow chart of the proposed nested set based method (NSM) is given in Fig. 2.

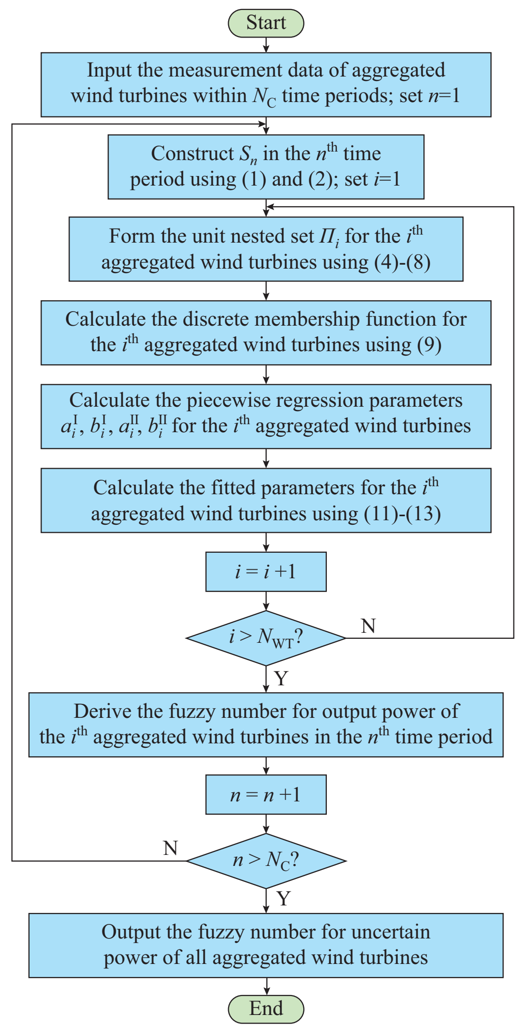

Fig. 2. Flow chart of proposed NSM.

III. Case Study

The utilized measurement output data of the aggregated wind turbines per minute within 24 hours for case study are provided in [9]. Let $\Delta T=1$ hour, $\Delta t=1$ min. The results of characterizing wind power using a fuzzy number $(t^{\inf}, t^{\ker}, t^{\sup})$ for each time period are given in the last column of Table I to show the advantages of the proposed NSM, compared with results by FFM and DCM that are given in the second and third columns of Table I.

TABLE I
FUZZY NUMBER OBTAINED WITH DIFFERENT METHODS

T (hour)	Fuzzy number (MW)		
	FFM	DCM	NSM
1	(0.43, 0.48, 0.53)	(0.44, 0.47, 0.50)	(0.43, 0.48, 0.52)
2	(0.51, 0.57, 0.63)	(0.55, 0.58, 0.62)	(0.53, 0.56, 0.63)
3	(0.49, 0.54, 0.59)	(0.52, 0.55, 0.58)	(0.50, 0.55, 0.59)
4	(0.49, 0.54, 0.59)	(0.50, 0.54, 0.57)	(0.49, 0.54, 0.59)
5	(0.50, 0.56, 0.62)	(0.53, 0.56, 0.58)	(0.52, 0.56, 0.61)
6	(0.46, 0.51, 0.56)	(0.49, 0.51, 0.54)	(0.47, 0.51, 0.55)
7	(0.57, 0.63, 0.69)	(0.60, 0.63, 0.66)	(0.58, 0.63, 0.68)
8	(0.58, 0.64, 0.70)	(0.61, 0.64, 0.67)	(0.60, 0.63, 0.69)
9	(0.49, 0.54, 0.59)	(0.51, 0.55, 0.58)	(0.50, 0.54, 0.59)
10	(0.49, 0.54, 0.59)	(0.51, 0.54, 0.56)	(0.50, 0.53, 0.58)
11	(0.47, 0.52, 0.57)	(0.49, 0.52, 0.54)	(0.47, 0.52, 0.56)
12	(0.52, 0.58, 0.64)	(0.55, 0.58, 0.61)	(0.53, 0.58, 0.62)
13	(0.56, 0.62, 0.68)	(0.59, 0.62, 0.65)	(0.58, 0.61, 0.66)
14	(0.54, 0.60, 0.66)	(0.58, 0.60, 0.63)	(0.56, 0.60, 0.65)
15	(0.51, 0.57, 0.63)	(0.54, 0.57, 0.60)	(0.53, 0.57, 0.62)
16	(0.55, 0.61, 0.67)	(0.58, 0.61, 0.66)	(0.57, 0.61, 0.67)
17	(0.54, 0.60, 0.66)	(0.57, 0.60, 0.63)	(0.56, 0.60, 0.65)
18	(0.58, 0.64, 0.70)	(0.61, 0.64, 0.68)	(0.59, 0.65, 0.69)
19	(0.54, 0.60, 0.66)	(0.57, 0.60, 0.63)	(0.55, 0.60, 0.65)
20	(0.54, 0.60, 0.66)	(0.57, 0.60, 0.64)	(0.55, 0.59, 0.66)
21	(0.54, 0.60, 0.66)	(0.57, 0.60, 0.63)	(0.56, 0.60, 0.64)
22	(0.50, 0.56, 0.62)	(0.53, 0.56, 0.59)	(0.52, 0.56, 0.61)
23	(0.50, 0.55, 0.61)	(0.52, 0.55, 0.58)	(0.50, 0.56, 0.59)
24	(0.57, 0.63, 0.69)	(0.59, 0.62, 0.65)	(0.58, 0.63, 0.67)

As shown in Table I, the values of $t^{\ker}$ at each time period with the three fuzzy methods are almost the same. However, it is quite different in terms of $t^{\inf}$ and $t^{\sup}$. Taking the 12th hour as an example, the fuzzy number for characterizing wind power by NSM is $\tilde{X}_{\mathrm{NSM}}=(0.53, 0.58, 0.62)$MW, which indicates the power output of wind power is around 0.58 MW. More quantitatively, the fluctuation range would be within (0.53, 0.62)MW and the most possible value is 0.58 MW, the membership degrees of others can be calculated by using (10). Compared with the results with FFM and DCM, which are $\tilde{X}_{\mathrm{FFM}}=(0.52, 0.58, 0.64)$MW and $\tilde{X}_{\mathrm{DCM}}=(0.55, 0.58, 0.61)$ MW, although the central value is still 0.58 MW, the characterized fluctuation ranges of wind power are different. The above indicates the proposed NSM method could keep the stability of the mean value of fuzzy numbers, but vary in the fluctuation characterizing by magnified or narrowed $t^{\inf}$ and $t^{\sup}$. Moreover, the improvement by the proposed method would be compared and explained by the following data matching index.

The coverage degree (CD), i.e., the ratio of the data range covered in the results to the whole data range, and the over-coverage degree (OCD), i.e., the ratio of the data range not covered in the results to the range of results, are designed to quantitatively evaluate the characterization performance of the results, which can be calculated by:

$$CD_i=\frac{\min(\max(X_i), t_i^{\sup})-\max(\min(X_i), t_i^{\inf})}{\max(X_i)-\min(X_i)} \tag{15}$$

$$OCD_i=1-\frac{\min(\max(X_i), t_i^{\sup})-\max(\min(X_i), t_i^{\inf})}{t_i^{\sup}-t_i^{\inf}} \tag{16}$$

The completeness of characterization performance is represented by CD, and a higher CD means a better completeness. The non-conservation of characterization performance is represented by OCD, and a lower OCD means better non-conservation. The closer the $CD\in[0,1]$ is to 1 and $OCD\in[0,1]$ is to 0, the better characterization performance would be, i.e., the best performance is attained when $CD=1$ and $OCD=0$.

As can be seen from the CD and OCD values with different methods in Fig. 3(a), FFM corresponds to high OCD values, which reaches 0.2095 in the worst case (the 21th hour) while NSM and DCM keep at 0. As shown in Table I, the results of $t^{\inf}$ and $t^{\sup}$ with FFM are always the smallest and largest ones. This phenomenon is due to the subjective selection of empirical factor when FFM is applied, so that the result is conservative. DCM corresponds to low CD values, which are 33% and 41% lower than NSM and FFM in average, respectively. As shown in the third column of Table I, the results of $t^{\inf}$ and $t^{\sup}$ of DCM are always the largest and smallest ones. The result with DCM is incomplete because of the data discarding in the clustering process.

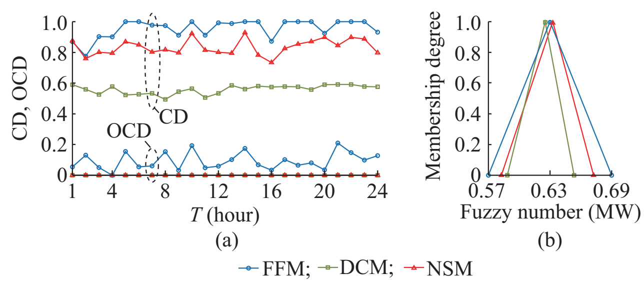

Fig. 3. CD and OCD values with different methods and fuzzy numbers at the 24th hour. (a) CD and OCD values. (b) Fuzzy number.

Delightedly, the results of the proposed NSM holds both relatively higher CD and lower OCD values compared with FFM and DCM. Taking the 24th hour as an example, fuzzy numbers are shown in Fig. 3(b), namely $\tilde{X}_{\mathrm{FFM}}$= (0.57, 0.63, 0.69)MW, $\tilde{X}_{\mathrm{DCM}}$= (0.59, 0.62, 0.65)MW, and $\tilde{X}_{\mathrm{NSM}}$= (0.58, 0.63, 0.67)MW, respectively. Although the central values are nearly the same as mentioned above, the fluctuation ranges of fuzzy numbers obtained with different methods are diverse. Specifically, FFM enlarges the variation of wind power and DCM is insufficient to characterize it because of the low data utilization, while the proposed NSM could maintain a low conservation without losing the integrity due to the cumulative feature consideration of the measurement data.

The above analysis fully indicates that the proposed improved fuzzy method could keep the similar qualitative characterization results and improve the accuracy through the obtained fuzzy number. Furthermore, the influence of sample interval Δt on fuzzy numbers using the proposed NSM is also explored. CD and OCD values with different sample intervals are shown in Fig. 4(a). It can be seen that the trend of CD slightly decreases while that of OCD increases more as Δt grows, which means the characterization performance has a slightly better completeness and worse non-conservation with a larger interval. For instance, as fuzzy numbers at the 24th hour shown in Fig. 4(b), t^{ker} at each Δt is very close, but t^{inf} gets smaller and t^{sup} gets larger as Δt increases. That is, with a shorter sample interval, the accuracy of the results can be further enhanced, while the computation burden increases as well.

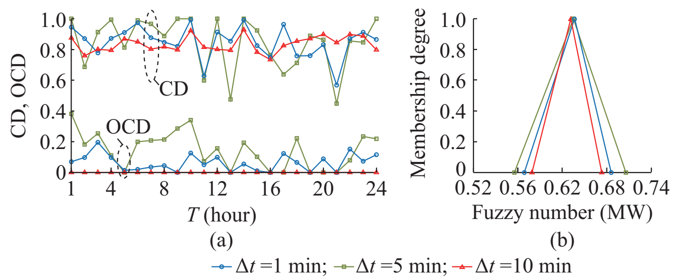

Fig. 4. CD and OCD values with different sample intervals and fuzzy numbers at the 24th hour. (a) CD and OCD values. (b) Fuzzy number.

IV. Conclusion

This letter presents an improved data-driven fuzzy method for characterizing wind power. It greatly helps improve the accuracy for estimating wind power, due to the cumulative feature consideration of measurement data distribution, compared with other fuzzy methods. The obtained fuzzy numbers could further be integrated into specific applications to strengthen the robust operation and planning of wind power integrated system.

References

[1] National Grid. (2019, Sept.). Technical report on the events of 9 August 2019. [Online]. Available: https://www.nationalgrideso.com/document/152346/downl-oad

[2] Y. Xiang, J. Liu, and Y. Liu. "Robust energy management of microgrid with uncertain renewable generation and load," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1034-1043, Mar. 2016

[3] Y. Xiang, J. Liu, F. Li *et al*., "Optimal active distribution network planning: a review," *Electric Power Components and Systems*, vol. 44, no. 10, pp. 1075-1094, May 2016.

[4] C. Gu, W. Yang, Y. Song *et al*., "Distribution network pricing for uncertain load growth using fuzzy set theory," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1932-1940, Jul. 2016.

[5] S. J. Aghili, H. Saghafi, and H. Hajian-hoseinabadi, "Uncertainty analysis using fuzzy transformation method: an application in power-flow studies," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 42-52, Jan. 2020.

[6] V. Miranda, M. A. Matos, and J. T. Saraiva, "Fuzzy load flow new algorithms incorporating uncertain generation and load representation," in *Proceedings of the 10th PSCC*, Graz, Austria, 1990, pp. 621-627.

[7] A. Piegat, "On practical problems with the explanation of the difference between possibility and probability," *Control and Cybernetics*, vol. 34, pp. 505-524, Jan. 2005.

[8] H. T. Nguyen and V. Kreinovich, "Nested intervals and sets: concepts, relations to fuzzy sets, and applications," in *Applications of Interval Computations. Applied Optimization.* Boston: Springer, 1996, pp. 245-290.

[9] Githhub. (2018, Oct.). Case data. [Online]. Available: https://github.com/eexxyy/Fuzzy/blob/master/WTdata.xlsx

Yue Xiang received the B. S. and Ph. D. degrees from Sichuan University, Chengdu, China, in 2010 and 2016, respectively. From 2013 to 2014, he was a joint Ph. D. student at the Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, USA, and also a Visiting Researcher at the Department of Electrical and Electronic Engineering, Imperial College London, London, UK, from 2019 to 2020. Now he is an Associate Professor in the College of Electrical Engineering, Sichuan University. His main research interests are power system planning and optimal operation, renewable energy integration and smart grid.

Shuai Hu received the B.S. degree in electrical engineering and automation, and the M.S. degree in power system and automation, both from Sichuan University, Chengdu, China, in 2017 and 2020, respectively. He is currently an Assistant Engineer in the Power Grid Planning and Development Research Center, State Grid Ningxia Electric Power Eco-Tech Research Institute, Yinchuan, China. His research interests include distributed renewable energy integration, distribution network planning, and optimal operation.

Junyong Liu received the Ph. D. degree from Brunel University, London, UK, in 1998. He is a Professor in the College of Electrical Engineering, Sichuan University, Chengdu, China. His main research interests are power market, power system planning, operation, stability, and computer applications.

Rui Wang received the B. S. and M. S. degrees from Sichaun University, Chengdu, China, in 2017 and 2020, respectively. She is currently an Assistant Engineer in the Department of Marketing, State Grid Hongguozi Power Supply Company, Shizuishan, China. Her research interest is power quality.