

Wind Power Forecasting Based on Extended Latin Hypercube Sampling

Xianbing Ding¹, Minfang Peng¹, Meie Shen², Liang Zhu³, Hongwei Che³, Sheng Zhou¹, Guangming Li¹ and Rongsheng Liu¹

¹College of Electrical and Information Engineering, Hunan University, Changsha, China

²College of Computer Science and Information Engineering, Beijing Information Science and Technology University, Beijing, China

³Electric Power Corporation of Hunan Province, Changsha, China

Abstract—With the rise of distributed generation, such as wind power and photovoltaic (PV), it is necessary to consider the effect of distributed generation's output randomness. Using the method of Latin hypercube sampling (LHS) can effectively fit output scenario. Considering the sampling number of conventional LHS (CLHS) must be fixed in advance, LHS(ELHS) can be extended to predict wind power. The sample scenarios were extended exponentially on the basis of original scenarios by CLHS, taking the relative error of the output variation before and after the extension as the convergence criterion of ELHS. Numerical example results show the feasibility and accuracy of the proposed algorithm.

Keywords-distributed generation; wind power forecast; Latin hypercube sampling

I. INTRODUCTION

With the mass consumption of traditional fossil energy, the problem of environment pollution is increasingly serious, the development and use of safe, clean, flexible renewable distributed energy becomes an effective measure to reduce the pressure of energy and environmental protection[1]. When distributed power represented by wind and solar power connects to the electricity grid, on the one hand, it helps to improve the local and overall reliability of grid [2], on the other hand, distributed generation's randomness and intermittent affected by factors such as the weather have brought negative impact to economic, safe and reliable operation of the power system [3-8].

Fan power's probability density function and distribution function are known, and its influence factor of output is single. So this paper takes it for example, and considers the output randomness to make output prediction. of wind power output, it built the regression - sliding wind

This paper uses the improved Latin hypercube sampling method to consider the fan output randomness and makes wind power prediction. Latin hypercube sampling can cover the whole sample space to sample randomly, and can simulate random output of samples effectively. In literature [9] Latin hypercube sampling and scene cutting technology were adopted to establish the fan output scene model, and to evaluate the reliability of the distribution network under each scenario, eventually acquiring the system reliability. But sampling number must be fixed well in advance using the

conventional Latin hypercube sampling. Once calculation result is not ideal, we need to change the number of sampling, increasing the difficulty of calculation. Literature [10] adopted the expanded Latin hypercube sampling algorithm considering the correlation of multiple random variable for power system flow calculation, and made full use of data from traditional Latin hypercube sampling, and overcame the defects of traditional Latin hypercube sampling technology effectively at the same time.

On the basis of traditional Latin hypercube sampling method and extended Latin hypercube sampling method considering the correlation of the multiple random variable, this paper gives the improved sampling method which not only overcomes the defect of CLHS method, but also can be applied to simulate the fan random output accurately. Extended Latin hypercube sampling multiplies on the basis of the original sampling, taking the relative error of the output estimate variation before and after as a condition of extension termination. The feasibility of the improved algorithm is proved by an example.

II. BASIC METHOD

A. CLHS Basic Algorithm

Latin hypercube sampling must consider the order of the random variable sampling value, to make the relevance of independent sampling values of random variable tend to minimum [11]. For the problem of a single random variable sampling, this don't need to be considered. The basic idea of sampling is to divide the probability space of random variable into several small areas which have the same length, and sample randomly in each interval, to ensure that the sampling values of random variables can cover the entire probability space.

CLHS algorithm's steps of single random variable X are as follows:

I). At first probability space is divided evenly into N intervals. For any of them $[(i-1)/N, i/N](i=1,2,3...N)$, a random number r is generated(r obeys uniform distribution within the range of 0 to 1), which form a random number set R ;

II). Generate sampling values set P based on probability space, any element is shown as:

$$p_i = \frac{r}{N} + \frac{i-1}{N} \quad (1)$$

III). Through the inverse transformation of function, generate sampling values set U based on the variable space, any element is shown as:

$$u_i = F^{-1}(p_i) \quad (2)$$

$F^{-1}(x)$ is the inverse function of probability distribution function of a random variable x .

B. ELHS Basic Algorithm

Traditional Latin hypercube sampling need to determine sampling number beforehand, and sampling number is fixed, so this paper overcomes the lack of its technology by extended Latin hypercube sampling method. The basic idea of ELHS is to add N new sampling values on the basis of the sampling values obtained through the CLHS, and new sampling values also cover the whole sampling space. They and original sampling values which have been modified form new sampling values set to improve the accuracy of random variable output's simulation.

ELHS algorithm's steps of single random variable X are as follows:

I). Deal with the original N sampling values:

$$r_i^{(1)} = \begin{cases} 2r_i, & r_i \leq 0.5 \\ 2r_i - 1, & r_i > 0.5 \end{cases} \quad (3)$$

The modified random number set is $R^{(1)}$, the sampling values set are $P^{(1)}, U^{(1)}, i = 1, 2, 3 \dots N$;

II). Generate a new random number set $R^{(2)}$ according to CLHS algorithm, and calculate to acquire the new sampling values set $P^{(new)}, U^{(new)}$;

III). Merge sets generated in the I), II) into the new sets $R^{(2)}, P^{(2)}, U^{(2)}$, to calculate the digital features I of output (expectation or variance), stop expanding sampling when I meets the convergence condition. Among them:

$$R^{(2)} = [R^{(1)} \quad R^{(new)}] \quad (4)$$

$$P^{(2)} = [P^{(1)} \quad P^{(new)}] \quad (5)$$

$$U^{(2)} = [U^{(1)} \quad U^{(new)}] \quad (6)$$

C. Convergence Criterion

The stop of expanded Latin hypercube sampling need to meet certain convergence condition. Due to sampling points

between by CLHS and ELHS technology is not independent, which do not meet the requirements for the central limit theorem, so the coefficient of variance criterion in Monte Carlo sampling is no longer applicable. This paper proposes the relative error of the output estimate variation before and after the extension as a practical convergence criterion.

Assuming that the digital features of output random variables after the $i-1$ time sampling is $I^{(i-1)}$, and the next one is $I^{(i)}$. Setting a small positive number as a condition of threshold epsilon. If the formula 7 is right, the sampling should be stopped, otherwise continue to expand the sampling.

$$\left| \frac{I^{(i)} - I^{(i-1)}}{I^{(i)}} \right| \leq \varepsilon \quad (7)$$

III. WIND POWER PREDICTION BASED ON ELHS

A. Probability Distribution of Wind Speed

The simulation of wind speed probability distribution used most commonly two parameter's Weibull distribution. The probability distribution function can be expressed as:

$$F(v) = P(v \leq V) = 1 - \exp[-(v/c)^k] \quad (8)$$

In the equation 9 two parameters c and k are respectively the scale parameter and shape parameter of the Weibull distribution, v stands for the given wind speed. The method of maximum likelihood estimation can be used for acquiring the value of two parameters c and k .

B. Fan Output Function

The output power of the fan group depends on the speed of wind, with the following output function to fit:

$$P = \begin{cases} 0, & 0 \leq v \leq v_{in} \\ P_N(a+bv+cv^2+dv^3), & v_{in} < v \leq v_N \\ P_N, & v_N < v \leq v_{out} \\ 0, & v > v_{out} \end{cases} \quad (9)$$

In the equation 10 a, b, c, d are fitting coefficients, obtained by combining history data of wind speed and output of fan group; v is the actual wind speed, v_{in} and v_{out} stand for the speed of wind cutting in and out, v_N is the rated wind speed; P and P_N are respectively the output power and rated power of the fan.

IV. EXAMPLE ANALYSIS

Taking the actual input wind speed of fan as input random variables, output of fan as output random variables, this paper use matlab programming to achieve the extended Latin hypercube sampling algorithm. Setting the maximum of extension's time as 5, and taking the variance of fan's output as the digital features of output random variables. The convergence threshold is 0.0005.

By the method of maximum likelihood estimation the scale parameter of two parameter's Weibull distribution is 7.2814, and shape parameters is 2.0135. The speed of wind cutting in and out are 4m/s and 25m/s. The rated wind speed is 10m/s, and the rated output power of fan group is 20MW.

CLHS method, ELHS method and simple random sampling method are used to carry out scene sampling of wind speed simulating the output of the fan. The following are the comparison of the expected value and the variance of the output of the wind turbine under different sampling numbers.

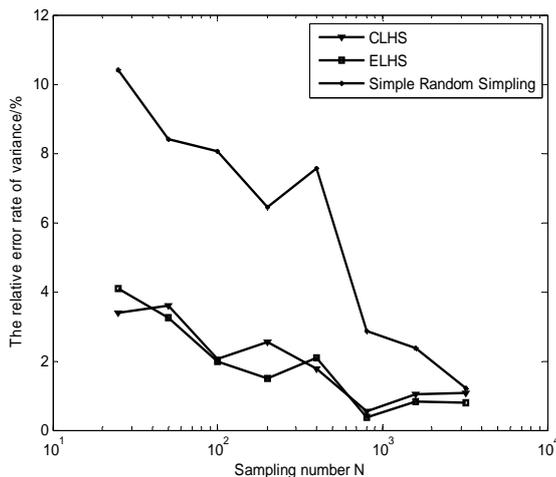


FIGURE I. COMPARISON OF THE EXPECTED VALUE FOR THREE METHODS OF FAN OUTPUT UNDER DIFFERENT SAMPLING NUMBERS

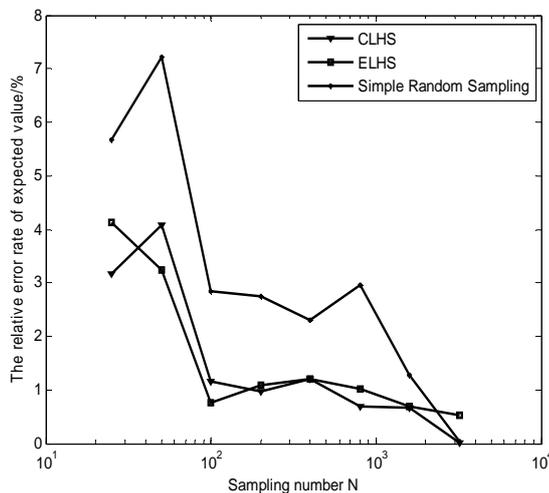


FIGURE II. COMPARISON OF THE VARIANCE FOR THREE METHODS OF FAN OUTPUT UNDER DIFFERENT SAMPLING NUMBERS.

Figure I, II show that the ELHS sampling algorithm in the aspect of convergence performance is more similar with CLHS, which also has faster convergence speed than the simple random sampling. Through the comparison of the

relative error rate of wind turbine output's variance and expected value, we can get a conclusion that ELHS algorithm in fitting scenarios of wind speed has a relatively high accuracy, so it can simulate the scenarios of wind speed effectively and predict the output energy of the wind turbine.

Sampling results show that determining times of ELHS extension and the corresponding sampling number through a given convergence threshold overcomes the CLHS algorithm's defect whose sampling number must be fixed effectively. It provides a new thought of the sampling methods, and the new algorithm combines the advantages of CLHS algorithm. At the same time, its accuracy of the simulation of output scenario is far higher compared with the method of simple random sampling. So it has a certain practical significance.

V. CONCLUSION

This paper uses extended Latin hypercube sampling algorithm to simulate the random output of wind farms. It is proved that the excellent performance of the algorithm in the convergence speed and accuracy through the comparison with the traditional Latin hypercube sampling and simple random sampling algorithm. It provides a new thought which the conditions of convergence to solve the problem for uncertainty sampling. How to use the results of fan output by extended Latin hypercube sampling effectively to analyze the impact of fan's connection to the power system is the focus of next research.

ACKNOWLEDGEMENTS

This work was supported by National Science Foundation of China under Grants 61472128 and 61173108, and Hunan Provincial Natural Science of China under Grants 14JJ2150.

REFERENCES

- [1] Kang Longyun, Guo Hongxia, Wu jie, et al. Characteristics of distributed generation system and related research issues caused by connecting it to power system[J]. Power System Technology, 2010,34(11):43-47.
- [2] Yuan Xiuguang, Huang Chun, Zhang Lei, Yu Haoming. Reliability evaluation of Distribution network considering Islanded Operation of Microgrid [J]. Power System Technology, 2015,03:690-697.
- [3] Alquthami T, Ravindra H, Faruque M O, et al. Study of photovoltaic integration impact on system stability using custom model of PV arrays integrated with PSS/E[C]//Proceeding of 2010 North American Power Symposium, Arlington, TX, USA: IEEE, 2010: 1-8.
- [4] [4] Rikos E, Tselepis E, Hoyer-Klick C et al. Stability and power quality issues in microgrids under weather disturbances study of photovoltaic integration[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2008, 1(3): 170-180.
- [5] [Paatero J V, Lund P D. Effects of large-scale photovoltaic power integration on electricity distribution networks [J]. Renewable Energy, 2007, 32(10): 216-234.
- [6] Wang Jian, Liu Tianqi, Li Xingyuan. Influences of Connecting Wind Farms and Energy Storage Devices to the Power Grid on Reliability of Power Generation and Transmission System [J]. Power System Technology, 2011,05:165-170.
- [7] He Yuqing, Peng Jianchun, Sun Qian. Fast Algorithm for Distribution System Reliability Evaluation Considering the Random Energy Output of Wind Power Generator[J]. Proceedings of the CSEE, 2010,13:16-22.
- [8] Xu Lizhong, Yang Guangya, Xu Zhao, Dong Zhaoyang, J.ΦSTERGAARD, Cao Yijia. Combined Scheduling of Electricity and Heat in a Microgrid with Volatile Wind Power[J]. Automation of Electric Power Systems, 2011,09:53-60.
- [9] Chen Can, Wu Wenchuan, Zhang Boming, Qin Jianguang, Guo Zijian. An Active Distribution System Reliability Evaluation Method Based on

- Multiple Scenarios Technique[J]. Proceedings of the CSEE, 2012,32 (34): 67-73.
- [10] Ding Ming, Wang Jingjing, Li Shenghu. Probabilistic Load Flow Evaluation with Extended Latin Hypercube Sampling [J]. Proceedings of the CSEE, 2013,33 (4):163-171.
- [11] Yu Han, Zhong Zhiyong Huang Jiebo, Zhang Jianhua. Power System Probabilistic Load Flow Calculation Method of the Latin Hypercube Sampling[J]. Automation of Electric Power Systems, 2009, 33 (21):32-35.