Fast Calculation of Steady-state Voltage in Power System Based on Monte Carlo and Deep Learning

Ting Zhou 1, Xiaochun Xu2, Jianyi Li3*, Xian Xu1, Peng Li3, Wen Shi2, Jiahao Wang3

1 Jiangsu Electric Power Company, Nanjing, Jiangsu, 210024, China
2 State Grid Jiangsu Electric Power Co., Ltd Huaian Branch, Huaian, Jiangsu, 223400, China
3 School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, Hebei, 071003, China
*Corresponding author’s e-mail: joeylee@ncepu.edu.cn

Abstract. It is of great significance to calculate the steady-state voltage of the power grid including the access of complex distributed power sources. To solve the problems of strong intermittent and high volatility of existing wind turbines, photovoltaics and other new energy sources, this paper generates a large number of section data in different scenarios according to the integrated probability model, and uses large data sets to train deep learning models. The steady-state voltage of each node is calculated for the distribution network under different scenarios in the IEEE 39-node system. The results of the calculation example show that the deep learning model trained by the proposed method can achieve 98% accuracy with less time of calculation. It is suitable for calculating the steady-state voltage of the large-scale complex power system.

1. Introduction

In recent years, with the increasing scale of DC transmission and new energy power generation, the proportion of electricity provided by traditional units has decreased[1]. Furthermore, with a large number of distributed new energy connected to the grid, the load share is increasing, which leads to the gradual transformation of the traditional distribution network into an active system[2]. The changes of the dual characteristics of the source and load mentioned above bring risks and uncertainties to the security characteristics of the power grid and the control of reactive power and voltage.

When distributed photovoltaic and wind generators are connected to the distribution network, it will inevitably change the power flow distribution of the network and the power flow direction of some nodes, and this change will also affect the reactive power and voltage distribution of the distribution network[3]. Aiming at the medium voltage distribution network with multiple distributed photovoltaic, the photovoltaic admittance capacity of feeder voltage not crossing the line under various distribution modes is studied in literature [4]. However, it ignores the strong volatility and uncertainty of the real photovoltaic and wind generators output models with the changes of region, time and environment. The literature [5] introduces a large number of small distributed photovoltaics and wind generators, taking into account the uncertainty of distributed power supply. The literature [6] uses Cholesky decomposition to explore the correlation between photovoltaic output and load in adjacent areas, but only considers the correlation of new energy, ignoring the limit situation that may be caused by negative correlation.
Research content in the above literature is limited to considering the conventional operating conditions of photovoltaics and wind generators, does not consider the possible extreme operating conditions of large-scale new energy loads so that the method is not suitable for the large area distribution network.

Focusing on the above problems, this paper proposes a new comprehensive probability model that combines traditional probability density and non-parametric kernel probability density which is used to extract the probability distribution characteristics of the output from the actual photovoltaic and wind generators operating data of a certain power grid in southern China. Then, using Monte Carlo simulation to randomly select samples according to the probability distribution of photovoltaics and wind generators. Perform multiple optimal power flow calculations on the IEEE 39-nodes distribution network model combined with sampled data to obtain the voltage data of each node under different scenarios. Finally, use the deep learning PYTORCH framework to establish a fully connected neural network, combine the input and output data sets in different scenarios of the optimal power flow for training, and verify the effectiveness of the method through the validation data set.

2. Power flow calculation based on Monte Carlo simulation

Monte Carlo method, also known as computer stochastic simulation method, is a numerical calculation method based on the theory and method of probability and statistics. The principle is to combine the requested problem with the probability model, carry out computer sampling or simulation, so as to obtain an approximate solution to the problem.

2.1. Comprehensive probability model of photovoltaic and wind generators

Traditional methods for simulating photovoltaic and wind generators output are commonly based on Beta distribution and Weibull distribution. The two probability models have the advantages of simplicity and practicality, but the models are not universally applicable. The non-parametric kernel (NPK) probability density is a method that does not require any prior knowledge and is completely based on the data to study the data distribution characteristics. Its probability density is based on the real measured data so that it has strong adaptability and high simulation accuracy. Considering the advantages of the two probability models, this paper proposes a comprehensive probability model that combines traditional probability density and NPK probability density. The flow chart is shown in figure 1:

![Figure 1. Flow chart of integrated probability density model](image)

(1) Based on the chi-square test ($\chi^2$ test) method, data independence can be tested whether the Beta distribution and Weibull distribution can be used to describe the probability distribution of photovoltaics and wind generators. Input the measured samples of photovoltaic output power $P_v$ and wind generators output power $P_w$ respectively. Then establish the Beta probability model of photovoltaics and the Weibull probability model of wind generators according to the following steps:
(2). Otherwise, build the non-parametric kernel density estimation model according to the following step (3).

(2) Establish a photovoltaic probability model \( f_s(p) \) based on Beta distribution and a wind generators \( f_w(v) \) probability model based on Weibull distribution. The probability model can be expressed as the following form:

\[
\begin{align*}
    f_s(p) &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} \left( \frac{p}{p_{\text{max}}} \right)^{\alpha-1} \left( 1 - \frac{p}{p_{\text{max}}} \right)^{\beta-1} \\
    f_w(v) &= \frac{k}{c} \left( \frac{v}{c} \right)^{\alpha-1} \exp \left[ -\left( \frac{v}{c} \right)^{\beta} \right]
\end{align*}
\]

Where \( p_{\text{max}} \) denotes PV maximum power, \( \Gamma(\cdot) \) is the Gamma function, \( \alpha \) and \( \beta \) are the shape parameters of the Beta distribution, \( k \) is the shape parameter of the Weibull distribution and \( c \) is the scale parameter of the Weibull distribution.

(3) Establish probability models \( f_s(p) \) and \( f_w(p) \) based on non-parametric kernel density estimation. The probability model can be expressed as the following form:

\[
\begin{align*}
    f_s(p) &= \frac{1}{n_s h} \sum_{i=1}^{n_s} K \left( \frac{p - p_s}{h} \right) \\
    f_w(p) &= \frac{1}{n_w h} \sum_{i=1}^{n_w} K \left( \frac{p - p_w}{h} \right)
\end{align*}
\]

Where \( n_s \) and \( n_w \) measured sample numbers of photovoltaic and wind generators respectively, \( h \) is the smoothing factor, which can be obtained by empirical algorithm [7], \( K(\cdot) \) is the kernel function, this article chooses the Gaussian function.

2.2. Load probability model
The load is time-varying, and the normal distribution can be used to approximate the change of the load power. The probability model of the active power and reactive power of the node \( i \) load can be expressed as following form:

\[
\begin{align*}
    f(P_i) &= \frac{1}{\sqrt{2\pi\sigma_i}} \exp \left( -\frac{(P_i - P_i^*)^2}{2\sigma_i^2} \right) \\
    f(Q_i) &= P_i Q_i / P_i^*
\end{align*}
\]

Where \( P_i \) and \( Q_i \) are active power and reactive power, \( \sigma_i \) is the standard deviation load power.

2.3. Optimal power flow calculation method based on Monte Carlo simulation
Combining the relevant probability models proposed above, Monte Carlo simulation technology is used to generate network parameters for optimal power flow calculation. The calculation steps are as follows:

(1) Use the MATPOWER flow calculation tool, input the network structure parameters, types and parameters of photovoltaic and wind generators, then set the Monte Carlo sampling scale \( n \).

(2) Generate \( n \) sets of random samples of photovoltaics, wind generators and loads based on the combination of the probability models in Sections 1.1 and 1.2.

(3) Using Newton-Raphson method to perform \( n \) times deterministic optimal power flow calculations.
(4) Calculate the voltage of each node of the distribution network by the optimal power flow and sort it into one sample data set.

3. Deep learning network

Deep Neural Network (DNN) has a simple structure which can gradually fit any nonlinear mapping through a multilayer structure, usually divided into input layer, hidden layer and output layer. The basic principles of the deep learning algorithm is to use gradient descent to update the weights of neurons, and continuously adjust the weights and bias terms of each neuron in the hidden layer through the correspondence between input and output and the back propagation of errors to ensure that the actual output and the expected output are consistent. The schematic diagram of the fully connected structure of the deep learning network is shown in figure 2:

![Figure 2. Deep learning fully connected network structure diagram](image)

3.1. Fully connected neural network frameworks

Suppose the number of training data samples is $n$, the input of the $i$ neuron in the hidden layer is $x_i^p$, then $x_i^p$ can be expressed as following form:

$$x_i^p = \sum_{j=1}^{M} \omega_{ij} o_j^p - b$$

(4)

Where $o_j^p$ is the output of neuron $j$ under the action of sample $n$, $\omega_{ij}$ is the weight of hidden layer neurons, $b$ is the bias term of the hidden layer neuron, $M$ is the number of input neurons.

3.2. Activation function and definition error

The activation function introduces a nonlinear factor to the neuron, so that the neural network can approximate any nonlinear function arbitrarily. Error is an important indicator to judge the training effect of deep learning model.

This article selects the sigmoid activation function:

$$f(x) = \frac{1}{1 + \exp\left[-(x + b_i) / b_0\right]}$$

(5)

There is a root mean square error $E$ of the output, which is defined as:

$$E = \frac{1}{2} \sum_{p=1}^{N} \sum_{k=1}^{K} (y_k^p - o_k^p)^2$$

(6)

4. Case analysis

4.1. Network structure and parameter
In this paper, the IEEE39-bus distribution network is used for simulation, there are 10 generator nodes, nodes 3, 16 and 35 are selected as wind generators access nodes, 1, 19 and 34 as photovoltaics access nodes. The data used in this article derived from the distribution network of a certain area in southern China where the climate is regular so that there are obvious gaps in power generation in different seasons, weather and moments, but there are periodic laws. Therefore, this article selects March, June, September, and December in the four quarters of 2019 as the feature samples. The parameters of each photovoltaic and wind generators access node are shown in table 1 and table 2:

| Table 1. Parameters of wind generators access node. |
|-----------------------------------------------|
| Probability | Shape | Scale | Rated capacity(MW) |
| model.     |       |       |                  |
| 3           | Weibull | 1.7   | 50               | 2     |
| 16          | Weibull | 2.5   | 30               | 3     |
| 35          | NPK    | N/A   | N/A              | 2     |

| Table 2. Parameters of photovoltaics access node. |
|-----------------------------------------------|
| Probability | Shape | Shape | Rated capacity(KW) |
| model.      |       |       |                  |
| 1           | Beta  | 1.82  | 2.52             | 200   |
| 19          | Beta  | 2.43  | 2.88             | 300   |
| 34          | NPK   | N/A   | N/A              | 200   |

According to the shape parameters and scale parameters of photovoltaics and wind generators, the output probability density models of photovoltaics and wind generators under the traditional probability distribution are drawn, as shown in figure 3:

Excluding photovoltaic and wind turbine access points, it is assumed that other distribution network node loads follow a normal distribution, the mean value is the steady-state value of the node load.

4.2. Train FCNN model

Integrate Section 3.2 according to the 150,000 sets of data generated by the probability model and the optimal power flow calculation results to generate the data set required for the deep learning model, of which 30,000 sets of data are selected as the verification data set, and the remaining sets of data are used to train model. Use the PYTORCH framework to build a fully connected neural network (FCNN). The number of layers is set to 6, and the number of neurons in each layer is 150, 300, 512, 300, 150 and 39. Set the number of training cycles to 2000, the training process is shown in figure 4:
According to figure 4, after repeated cyclic training, the loss of the model continued to decrease and stabilized, and the accuracy of the generated model gradually increased and stabilized at 98%. Compare the mean voltage $\sigma_v$ and standard deviation $\mu_v$ with that calculated under the conventional probability density model. Accuracy of some node voltage calculation is shown in Table 3:

| Mean voltage | Standard deviation | Mean voltage | Standard deviation |
|--------------|--------------------|--------------|--------------------|
| $\sigma_v$   | $\mu_v$            | $\sigma_{m}$ | $\mu_{m}$          |
| 2            | 0.923              | 0.924        | 0.00532            |
| 18           | 0.934              | 0.934        | 0.00498            |
| 39           | 0.925              | 0.925        | 0.00526            |

5. Conclusion
This paper combines the traditional probability density of wind generators and photovoltaics with non-parametric nuclear probability density, and completes the fast and accurate calculation of steady-state voltage through deep learning model training. The relevant conclusions are as follows:

1. The integrated probability density combines the features of traditional models, and has the probability features of real data, which has wide applicability.

2. After the proposed comprehensive probability density model performs the optimal power flow calculation, the node voltage has the same or higher accuracy than the traditional scheme.

3. The proposed deep learning training model is aligned with the traditional power flow calculation, which can calculate the steady-state voltage of the node more quickly with high accuracy.

Acknowledgment
This work was supported by the SGCC Science and Technology Project (5108-202018028A-0-0-00).

References
[1] Li P, Guo T, Han X, et al. The optimal decentralized coordinated control method based on the $H_\infty$ performance index for an AC/DC hybrid microgrid[J]. International Journal of Electrical Power & Energy Systems, 2021, 125. DOI: 10.1016/j.ijepes.2020.106442.

[2] Sara H, Afshin N, Kazem Z, et al. Stochastic bi-level coordination of active distribution network and renewable-based microgrid considering eco-friendly Compressed Air Energy Storage system and Intelligent Parking Lot[J]. Journal of Cleaner Production, 2021, 278. DOI: 10.1016/j.jclepro.2020.122808.

[3] Li P, Guo T, Zhou F, Yang J, Liu Y. Nonlinear coordinated control of parallel bidirectional power converters in an AC/DC hybrid microgrid[J]. International Journal of Electrical Power & Energy Systems, 2020, 122:106208.

[4] Dai J, Wei H, Bao H, et al. Random Power Flow Based on Distributed Power System Based on Non-Trajectory Transformation[J]. Electric Power Automation Equipment, 2016, 36(3):86-93.

[5] Huang W, Ge L, Hua L, et al. Day-ahead and real-time optimal scheduling for active distribution network based on probabilistic power flow[J]. Automation of Electric Power Systems, 2018,
[6] Ren Z, Yan W, Xiang B, et al. Probabilistic power flow analysis incorporating the correlations between PV power outputs and loads[J]. Transactions of China Electrotechnical Society, 2015, 30(24):181-187

[7] Zhao Y, Shen Z, Zhou N, et al. Reliability assessment of bulk power systems utilizing simulation and nonparametric Kernel density estimation[J]. Automation of Electric Power Systems, 2008, 32(6):14-19.