

Article

A Fast Reliability Evaluation Strategy for Power Systems under High Proportional Renewable Energy—A Hybrid Data-Driven Method

Jiaxin Zhang ^{1,2}, Bo Wang ^{1,2}, Hengrui Ma ^{1,2,*}, Yunshuo Li ³, Meilin Yang ⁴, Hongxia Wang ^{1,2,5} and Fuqi Ma ^{1,2}

- ¹ Hubei Engineering and Technology Research Center for AC/DC Intelligent Distribution Network, Wuhan University, Wuhan 430072, China; 2020282070157@whu.edu.cn (J.Z.); whwdwb@whu.edu.cn (B.W.); 2018282070092@whu.edu.cn (H.W.); whumfq@whu.edu.cn (F.M.)
- ² School of Electrical and Automation, Wuhan University, Wuhan 430072, China
- ³ China Electric Power Research Institute, Haidian District, Beijing 100192, China; lys0801@163.com
- ⁴ State Grid Huaining Power Supply Company, Huaining 246121, China; shengming_yu1@163.com
- ⁵ Department of Electrical & Computer Engineering, University of Denver, Denver, CO 80208, USA
- * Correspondence: henry3764@foxmail.com

Abstract: With the increasing scale of the power system, the increasing penetration of renewable energy, and the increasing uncertainty factors, traditional reliability evaluation methods based on Monte Carlo simulation have greatly reduced computational efficiency in complex power systems and cannot meet the requirements of real-time and rapid evaluation. This article proposes a hybrid data-driven strategy to achieve a rapid assessment of power grid reliability on two levels: offline training and online evaluation. Firstly, this article derives explicit analytical expressions for reliability indicators and component parameters, avoiding the computational burden of repetitive Monte Carlo simulation. Next, a large number of samples are quickly generated by parsing expressions to train convolutional neural networks (CNNs), and the system reliability index is quickly calculated under changing operating conditions through CNNs. Finally, the effectiveness and feasibility of the proposed method are verified through an improved RTS-79 testing system. The calculation results show that the method proposed in this article can achieve an online solution of second-level reliability indicators while ensuring calculation accuracy.

Keywords: power system; hybrid data-driven strategy; explicit analytical expressions; convolutional neural network; reliability index



Citation: Zhang, J.; Wang, B.; Ma, H.; Li, Y.; Yang, M.; Wang, H.; Ma, F. A Fast Reliability Evaluation Strategy for Power Systems under High Proportional Renewable Energy—A Hybrid Data-Driven Method.

Processes **2024**, *12*, 608. <https://doi.org/10.3390/pr12030608>

Academic Editors: Chang-Hua Lin, Shiue-Der Lu and Hwa-Dong Liu

Received: 19 February 2024

Revised: 6 March 2024

Accepted: 13 March 2024

Published: 19 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Power system reliability assessment is an important part of power system planning, design, and operation management. The evaluation results have significant reference value for power system planning and operation [1,2]. However, due to the fact that the equipment reliability parameters used in traditional power system reliability assessments are mostly based on the historical fault statistical records, the data quality is not high, and it is easily affected by human factors, resulting in significant errors in reliability assessment results [3]. In addition, the complexity of operating conditions in the power system, such as large-scale renewable energy integration and diverse user demands, has increased the difficulty of power system reliability assessments [4].

At present, power system reliability evaluation algorithms to be applied in engineering can be divided into two categories: a state analysis method and Monte Carlo simulation (MCS) method [5]. The evaluation process of these two methods is basically the same, usually consisting of three parts: system state selection, system state evaluation, and indicator statistics [6,7]. Among them, system state evaluation usually requires performing optimal power flow (OPF) on system state samples to determine the minimum load shedding of the system. Although DC OPF models with high computational efficiency are often

used, state evaluation is the most time-consuming step in reliability analysis due to the involvement of tens or hundreds of thousands of samples. This significantly increases the computational burden and the quality of state samples directly affects the accuracy of evaluation results. Researchers have proposed a series of improved sampling methods, including state space pruning [8], importance sampling [9], Latin hypercube sampling [10], etc. By generating and selecting representative samples, the number of samples required to obtain specific accuracy reliability evaluation results is reduced, thereby reducing computational complexity and improving computational efficiency. Reference [11] combines Latin hypercube sampling with importance sampling and proposes an improved Latin hypercube method to improve sampling efficiency. Based on importance sampling, references [12–14] introduce the cross-entropy algorithm, which constructs the optimal probability distribution of components by changing the probability of fault states, reduces sample variance fluctuations, and accelerates convergence. In detail, reference [12] combines cross-entropy with MCS for a reliability evaluation of power generation systems; reference [13] further simplifies the evaluation process of the method proposed in [12]. After obtaining the cross-entropy probability distribution parameters, the reliability index is obtained by combining fast Fourier decomposition analysis. Reference [14] considers transmission lines and applies cross-entropy to evaluate power generation and transmission systems, accelerating the convergence of MCS. For sampling states, the power loss of the system in this state depends on the load demand that has not been met after unit scheduling. The solution of this problem is mainly based on the optimal power flow (OPF) algorithm under nonlinear constraints, and the solving process is time-consuming [15]. Essentially, if the AC OPF model is used, due to the presence of power flow constraints, the mathematical model has nonlinear and non-convex characteristics, making it difficult to ensure optimality.

On the other hand, the reliability assessment of power systems is essentially solving a series of OPF problems with minimum load shedding as the objective function. OPF establishes a mapping relationship between the system state and optimal operating indicators (such as minimum load shedding), which can be regarded as a special type of nonlinear function. Therefore, in principle, it can be fitted and approximated through data-driven methods. In order to accelerate system state evaluation, researchers have proposed methods such as neural networks [16,17], learning vector quantization networks [18], and support vector machines [19] to replace the traditional model of linear programming. However, the selection, quantity, and computational accuracy of training samples have not been uniformly solved. Reference [20] proposed a three-stage OPF solving framework and utilized a stacked extreme learning machines to build the mapping relationships of each stage. In addition, the rapidly developing deep learning technology has injected new vitality into OPF fitting. Reference [21] used a deep neural network (DNN) to fit the relationship between node injection power, active power difference of each branch before and after branch disconnection, and voltage phase, and approximately solved the DC power flow equation. Reference [22] used stacked denoising auto encoders to fit the relationships between the electrical and thermal load power of each node, as well as the gas load flow rate with voltage, pressure, pipeline flow rate, etc., achieving a data-driven multi-energy flow calculation. In terms of an OPF solution, similar to power flow analysis, reference [23] used DNN to learn the mapping relationship between system load distribution and generator power allocation, thereby solving the DC optimal power flow model. Unfortunately, methods represented by support vector machines and conventional neural networks are highly sensitive to the selection of hyperparameters such as learning rate, batch size, network structure, etc. Choosing inappropriate hyperparameters may lead to model training failure or poor performance. Furthermore, due to the complexity and nonlinearity of neural networks, their decision-making process often lacks interpretability. This makes it difficult for power system operators to understand why the model makes a certain decision, making it difficult to effectively debug and optimize it.

In all, there are still some challenges that urgently need to be addressed as follows.

1. The above methods have made some progress in reducing the calculation time of reliability assessments, but there is still a common limitation: if the reliability parameters of components or network topology change, a new reliability assessment must be conducted to obtain the changed system reliability indicators. In engineering applications that require repeated reliability evaluations, although the application of the above methods can shorten the calculation time of a single reliability evaluation, repeated evaluations still require significant time consumption.
2. These studies can effectively improve the solving speed of a large number of repetitive OPF calculations, but still face difficulties in feature extraction, overfitting, and difficulty adapting to complex network topology changes, which cannot be directly applied to reliability analysis.

To address the aforementioned issues, this paper proposes a hybrid data-driven approach to achieve a rapid reliability assessment of power systems. The innovation points of this paper can be summarized as follows.

- (1) The analytical expressions for the reliability indicators and component parameters of the power system are derived. By using this analytical expression, a large number of samples can be quickly obtained under different operating conditions.
- (2) An online and offline two-stage hybrid data-driven model has been established. Among them, a reliability index regression model based on CNN is constructed, achieving fast solving while ensuring accuracy.

The framework is organized as follows: in Section 2, the analytical expression between system reliability indicators and component reliability parameters has been derived, which can provide a large number of data samples for offline training quickly. A classification regression CNN model and its detailed design parameters are introduced in Section 3. Then, the improved IEEE RTS-79 testing system is given to validate the effectiveness and feasibility of the proposed model.

2. Analytical Expression for Reliability Indicators of Power Systems

Sequential Monte Carlo simulation is widely used in the reliability assessment of complex power systems with large-scale renewable energy integration. Currently, there are many commonly used indicators to characterize the reliability of power systems, such as loss of load probability (LOLP), loss of load frequency (LOLF), and expected energy not supplied (EENS), system average interruption frequency index (SAIFI), system average interruption duration index (SAIDI), and so on. In this paper, these three indicators (LOLP, LOLF, and EENS) are highly focused on because SAIFI and SAIDI mainly focus on the frequency and duration of power outages for users, while EENS, LOLF, and LOLP focus more on evaluating whether the power generation capacity of the system can meet the load demand from the perspective of the power system. The reliability assessment in this article focuses more on the overall power generation capacity and margin of the system, rather than the power outage situation of individual users. In addition, other indicators can also be derived from these three indicators. Common component reliability parameters include failure rate λ , repair rate μ and component availability A , etc. This section takes the LOLP, LOLF, and EENS system indicators as examples to introduce the analytical model for power system reliability evaluation. Other models can be obtained using similar methods.

Firstly, the traditional reliability index calculation formulas using the Monte Carlo method are shown in (1)–(3):

$$LOLP = \frac{\sum_{s \in \psi} t_d(s)}{T_{total}} \quad (1)$$

$$LOLF = \frac{\sum_{s \in \psi} I_f(s)}{T_{total}} \quad (2)$$

$$EENS = \frac{\sum_{s \in \psi} t_d(s) \times L_d(s) \times T}{T_{total}} \quad (3)$$

Here, T_{total} represents the total simulation duration; ψ represents the set of all failure times of the system within the total simulation duration; $t_d(s)$ represents the duration of system event s . $I_f(s)$ is a 0–1 variable, if the previous event of system event s does not lose load, $I_f(s) = 1$; otherwise $I_f(s) = 0$; $L_d(s)$ represents the amount of load shedding for system event s ; T is the given duration interval. In this paper, the value of T is set as 8760 h, i.e., one year.

It is assumed that the components involved in the reliability assessment process only have two states of operation and failure, and the faults between each component are independent of each other. For m components, the number of combined states that exist is $M = 2^m$.

Let $\varphi_j (j = 1, 2, \dots, M)$ represent the set of system events where m components are in the j -th combination state. Therefore, the set of system failure events ψ in Formulas (1)–(3) can be divided into M subsets, i.e., $\varphi_1, \varphi_2, \dots, \varphi_M$. Meanwhile, the initial Formulas (1)–(3) can be decomposed into the sum of M terms, as shown in (4)–(6) below.

$$LOLP = \frac{\sum_{s \in \varphi_1} t_d(s)}{T_{total}} + \frac{\sum_{s \in \varphi_2} t_d(s)}{T_{total}} + \dots + \frac{\sum_{s \in \varphi_M} t_d(s)}{T_{total}} \quad (4)$$

$$LOLF = \frac{\sum_{s \in \varphi_1} I_f(s)}{T_{total}} + \frac{\sum_{s \in \varphi_2} I_f(s)}{T_{total}} + \dots + \frac{\sum_{s \in \varphi_M} I_f(s)}{T_{total}} \quad (5)$$

$$EENS = \frac{\sum_{s \in \varphi_1} t_d(s) \times L_d(s) \times T}{T_{total}} + \frac{\sum_{s \in \varphi_2} t_d(s) \times L_d(s) \times T}{T_{total}} + \dots + \frac{\sum_{s \in \varphi_M} t_d(s) \times L_d(s) \times T}{T_{total}} \quad (6)$$

Here, taking the LOLP index as an example, the process of establishing a reliability analysis model is derived. Assuming event E indicates a system loss of the load event, events F_1 – F_M represent M combined states corresponding to m components. Obviously, events F_1 – F_M form a complete set of events, with any two events mutually exclusive.

Considering that the faults between each component are independent of each other, combined with the full probability formula, Formula (4) can be rewritten as (7) and (8):

$$\begin{aligned} LOLP &= \frac{\sum_{s \in \varphi_1} t_d(s)}{T_{total}} + \frac{\sum_{s \in \varphi_2} t_d(s)}{T_{total}} + \dots + \frac{\sum_{s \in \varphi_M} t_d(s)}{T_{total}} \\ &= P(E \cap F_1) + P(E \cap F_2) + \dots + P(E \cap F_M) \\ &= P(EF_1) + P(EF_2) + \dots + P(EF_M) \\ &= P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + \dots + P(E|F_M)P(F_M) \\ &= K_1P(F_1) + K_2P(F_2) + \dots + K_MP(F_M) \end{aligned} \quad (7)$$

$$K_j = P(E|F_j) = \frac{P(EF_j)}{P(F_j)} = \frac{\sum_{s \in \varphi_j} t_d(s)}{T_{total}} \times \frac{1}{P(F_j)} \quad (8)$$

Equation (8) represents the conditional probability of system load loss of event E occurring when the components are in a combined state F_j . $P(E|F_j)$ depends on the reliability level of $(n - m)$ other components under the condition that the system topology, electrical parameters, and operating parameters remain unchanged. Due to the assumption that the reliability parameters of $(n - m)$ other components are fixed and unchanging,

$K_1 - K_M$ is also a fixed and unchanging constant. Due to the independence of faults between components, the calculation formula for $P(F_j)$ is shown in (9):

$$P(F_j) = \prod_{g \in \varphi_{ju}} A_g \prod_{g \in \varphi_{jd}} (1 - A_h) \quad (9)$$

Here, φ_{ju} and φ_{jd} are respectively, the sets of components in working and faulty states in the combined state F_j .

Substituting Formula (9) into Formula (7) yields the analytical expression for the LOLP indicator, which is shown as Formula (10):

$$LOLP = \sum_{j=1}^M K_j \prod_{g \in \varphi_{ju}} A_g \prod_{g \in \varphi_{jd}} (1 - A_h) \quad (10)$$

Similarly, the analytical expressions for LOLF and EENS indicators are shown in Equations (11) and (12):

$$LOLF = \sum_{j=1}^M J_j \prod_{g \in \varphi_{ju}} A_g \prod_{g \in \varphi_{jd}} (1 - A_h) \quad (11)$$

$$EENS = \sum_{j=1}^M L_j \prod_{g \in \varphi_{ju}} A_g \prod_{g \in \varphi_{jd}} (1 - A_h) \quad (12)$$

Here, the coefficients J_j in Equation (11) can be calculated based on Equations (13)–(15):

$$J_j = J_{j,O} + \sum_{h \in \varphi_{jd}} (J_{j,h} \times \mu_h) \quad (13)$$

$$J_{j,h} = \frac{\sum_{s \in \varphi_j} t_{d,h}(s)}{T_{total}} \times \frac{1}{P(F_j)}, h \in \varphi_{jd} \quad (14)$$

$$J_{j,O} = \frac{\sum_{s \in \varphi_j} I_f(s)}{T_{total}} \times \frac{1}{P(F_j)} - \sum_{h \in \varphi_{jd}} (J_{j,h} \times \mu_h) \quad (15)$$

Here, $t_{d,h}(s)$ represents the duration of the system state s under specific conditions, and μ_h represents the repair rate of component h . The coefficients in Equation (12) can be calculated based on Equation (16):

$$L_j = \frac{\sum_{s \in \varphi_j} t_d(s) \times L_d(s) \times T}{T_{total}} \times \frac{1}{P(F_j)} \quad (16)$$

Formulas (10)–(12) provide the calculation expression for reliability indicators. When the coefficients in the analytical expression are calculated, time-consuming calculations can be avoided. By simply inputting the reliability parameters of the components into the model, relevant reliability indicators can be obtained. However, the topology, electrical parameters, and operating parameters of the power system often change, which leads to the need to repeatedly calculate the coefficients in the analytical expressions, which to some extent increases the computational burden and makes real-time online fast solving difficult. Therefore, this article adopts a hybrid data-driven strategy, dividing the rapid reliability assessment of the power system into two steps. The first step is to generate a large number of data samples (including system operating parameters and reliability indicators) using analytical expressions. Then, a convolutional neural network (CNN) is trained to quickly obtain the system reliability index under changing operating conditions through CNN.

3. Reliability Index Regression Model of Power System Based on CNN

3.1. Introduction to CNN Model

To ensure that the reliability index calculation results meet the accuracy requirements, it is usually necessary to simulate the operation of the system for millions of hours and evaluate the operation status of tens of thousands of systems. As the scale of renewable energy power increases, the number and degree of uncertainty sources in the system continue to increase, and the number of samples required to cover system operation scenarios will also significantly increase, and the amount of Monte Carlo simulation calculations will also significantly increase. Based on the reliability evaluation analytical model in the previous section, a dataset of reliability indicators under fixed operating conditions can be obtained. Changing the operating conditions of the system to obtain different data samples, providing sufficient data samples for the CNN model is covered in this section.

The calculation of reliability indicators establishes the mapping relationship between system operation status and reliability indicators, which is a key link in reliability evaluation. This section uses a CNN to fit the above mapping relationship and establish a regression model for system reliability indicators, thereby accelerating the process of reliability evaluation.

The basic structure of a CNN is shown in Figure 1, which mainly consists of convolutional layers, pooling layers, and fully connected layers, which can be represented by Equations (17)–(19).

$$h_c^l = f(X * W_c^l + b_c^l) \quad (17)$$

$$h_p^l = p(h_c^l) \quad (18)$$

$$h_{fc} = f(h_p^l * W_{fc} + b_{fc}) \quad (19)$$

where X represents the input vector of the convolutional layer; W_c^l represents the weight matrix of the l -th convolutional kernel; b_c^l represents the bias term; f represents the activation function; h_c^l represents the l -th output vector; p represents the pooling operation, such as taking the maximum value, etc.; h_p^l represents the matrix vector for the l -th pooling surface; h_{fc} represents a bias term.

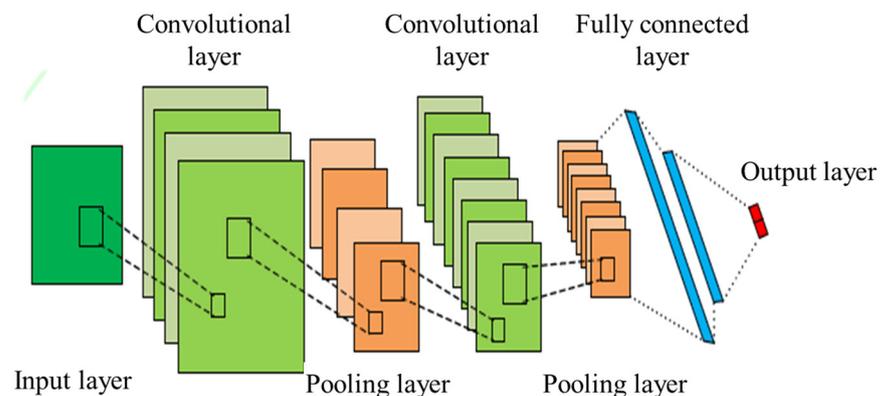


Figure 1. Basic structure of CNN model.

3.2. Input Feature Matrix Construction

When using a CNN to fit system reliability indicators, it is necessary to consider the working status of each component and its impact on system operation constraints, such as the impact of generator shutdown on its output range and line shutdown on power flow distribution. At the same time, considering the current construction of new power systems, that is, under the high proportion of renewable energy integration, the uncertainty characteristics of the power system are significant, and the fluctuation range of the system

operation status is large. Therefore, this article constructs a CNN input feature matrix as shown in Equation (20).

$$X = \begin{bmatrix} P_{D1} & P_{D2} & \cdots & P_{Dn} \\ P_{R1} & P_{R2} & \cdots & P_{Rn} \\ E_1 & E_2 & \cdots & E_n \\ B_{11} & B_{22} & \cdots & B_{nn} \end{bmatrix} \quad (20)$$

In the Formula (20), P_{Di} represents the load power of node i ; P_{Ri} represents the power of renewable energy at node i ; E_i represents the power generation capacity of conventional units at node i ; B_{ii} represents the self-impedance of node i .

In matrix X , P_{Di} and P_{Ri} reflect the time-varying characteristics of node load and renewable energy, E_i reflects the information of generator operation status, and B_{ii} can reflect topological information such as transmission line operation status. When a specific generator or transmission line is forced to shut down, the characteristic parameters E_i and B_{ii} (the first and last nodes of the i line) will change accordingly. In this paper, line losses are not taken into account because in the process of reliability assessment, in order to save calculation time, the DC power flow model is commonly used. The DC power flow model ignores line losses.

Although the number of lines, conventional units, and renewable power stations in the system varies, the input of the different dimensions mentioned above can be unified into a feature matrix that is only related to the number of system nodes n through the processing of Equation (20), for subsequent training and processing.

To obtain the mapping relationship between reliability indicators and system operating status (input feature matrix X), a CNN-based regressor model is designed, as shown in Figure 2. Among them, the kernel size of both convolutional layers is 3×3 , and the number of channels is set to 20 and 40, respectively. At the same time, to accelerate neural network training, a BN layer was introduced to perform batch normalization on the features extracted by convolution. The output dimensions of the fully connected layer are $15n$, $15n$, n , and 1, respectively. The final output dimension is 1, and the output quantity is the system reliability index. The loss function of the regression model is as follows:

$$L_r = \frac{1}{N_r} \sum_{r=1}^{N_r} (L^* - L)^2 + \frac{\alpha}{2} w_r^T w_r \quad (21)$$

Here, N_r is the number of samples used for reliability evaluation regression training, L^* is the actual reliability evaluation index value, L is the reliability index value fitted by the CNN regression network, α is the weight coefficient, and w_r is the weight matrix of the regression network. In practical applications, for specific power systems, a certain number of system operating state samples can be generated by solving the reliability analysis models (10)–(12), and the training of the regression network can be completed using known system operating states and reliability indicators as raw data. It should be noted that the BP algorithm has strong optimization capabilities. Through the backpropagation algorithm, the model can automatically adjust network parameters to reduce the value of the loss function, thereby improving the prediction accuracy of the model. In addition, the BP layer can also achieve complex nonlinear mapping relationships, enabling the model to handle various complex tasks. Therefore, when constructing deep learning models, CNN and BP layers are usually used in combination to improve the performance of the model.

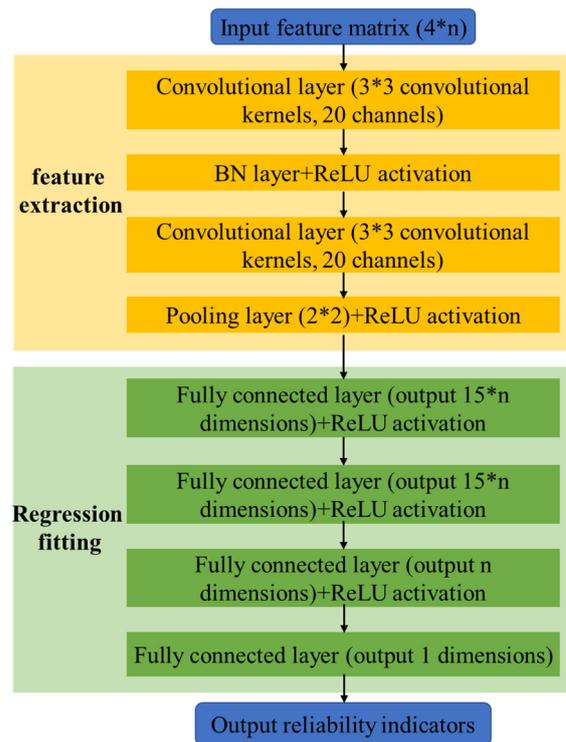


Figure 2. A CNN-based regressor model.

3.3. Improved CNN Classification Regression Model

The uncertainty characteristics of power systems containing renewable energy are significant, and the fluctuation range of system operation status is large. The CNN regression model directly trained according to Section 3.2 is difficult for achieving high-fitting accuracy on all the samples to be evaluated. Therefore, a CNN-based classification model is further introduced, which first classified the operating status samples based on whether they could be reliably powered, and then estimated the reliability indicators of the samples that could not be reliably powered, thus forming a CNN classification regression model. This processing is not only beneficial for improving the accuracy of CNN regression model training and reliability evaluation, but also avoids the evaluation and calculation of a large number of samples without load shedding, which helps to improve efficiency. At the same time, in practical applications, a more detailed classification of the system's operating status can also be made based on the binary classification in this article, further improving regression accuracy.

The input of the classification model is consistent with the regression model, that is, the input feature matrix X established by Equation (20); but its output is the confidence level of whether the system can reliably supply power. To this end, a Softmax activation function is added before the final output of the classification network. Considering that sample classification is considered simpler than reliability index regression, in terms of network structure design, the classification model introduces an additional pooling layer and reduces a fully connected layer compared to the regression model. The overall structure is shown in Figure 3. The loss function for classification is shown in Equation (22):

$$L_{cl} = -\frac{1}{N_{cl}} \sum_{cl=1}^{N_{cl}} \sum_{o=1}^2 y_{cl,o}^* \log(y_{cl,o}) + \frac{\beta}{2} w_{cl}^T w_{cl} \quad (22)$$

Here, N_{cl} is the number of samples used for classification training; $y_{cl,o}^*$ is the 0–1 label; $y_{cl,o}$ is the output of the classification model; β is the weight coefficient, and w_{cl} is the weight matrix of the classification model neural network.

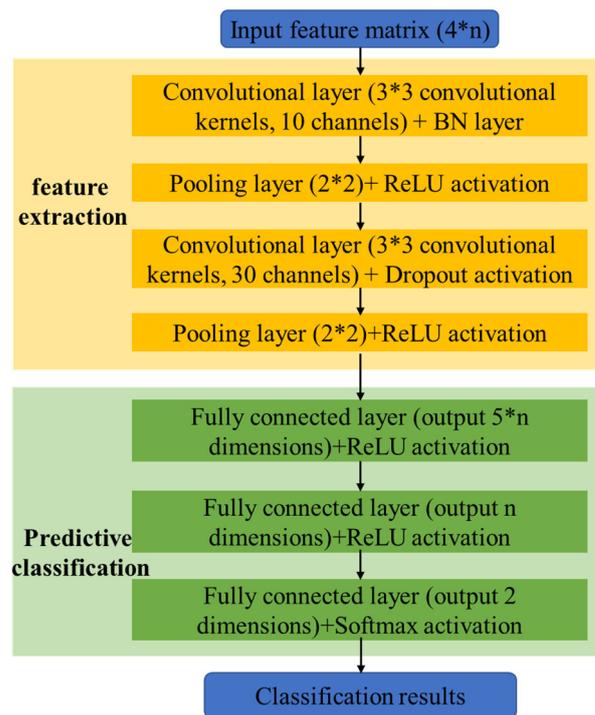


Figure 3. A CNN-based classification model.

Finally, the combination of the classification model and regression model forms a CNN classification–regression model. The detailed calculation flowchart is shown in Figure 4. During training, the classification model and regression model are two independent neural networks that can be trained separately. The classification model training set is filtered to obtain the regression model training set. After filtering, only a small portion of the samples with reliable power supply in this training set are used to improve the training efficiency and performance of the regression model. During testing, the two can be connected in series, and the system status can be classified first, followed by regression prediction of the load shedding of samples that cannot be reliably powered.

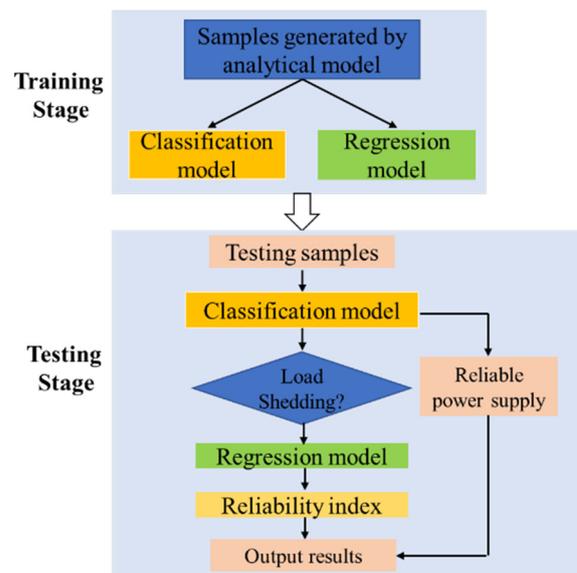


Figure 4. Detailed flowchart of the proposed method.

4. Case Study

4.1. Introduction of the Test System

This article takes the RTS-79 testing system as an example. The RTS-79 testing system is widely used in power system reliability assessments due to its advantages of standardization, scale, data accessibility, and representativeness. The IEEE-RTS testing system wiring diagram is shown in Figure 5, which includes 32 generators, 33 transmission lines, and 5 transformers. The total installed capacity of the system is 3405 MW, with a peak load of 2850 MW. The annual load curve and reliability parameters of each component are detailed in the literature [13]. At the same time, in order to facilitate the verification of the effectiveness of the method proposed in this article in high-proportion renewable energy scenarios, 70 MW wind farms are added at nodes 15, 16, 18, 19, and 21, and 50 MW photovoltaic power plants are added at nodes 13, 17, 20, 22, and 23. The output curves of wind and photovoltaic power stations are based on statistical data from a certain province in northwest China and converted according to the installed capacity [17]. Reliability assessments and CNN-related programs were written using software MATLAB 2016b. All calculations were conducted on a desktop computer with an Intel (R) Core (TM) i5-7500 CPU at 3.40GHz and 8 GB of memory.

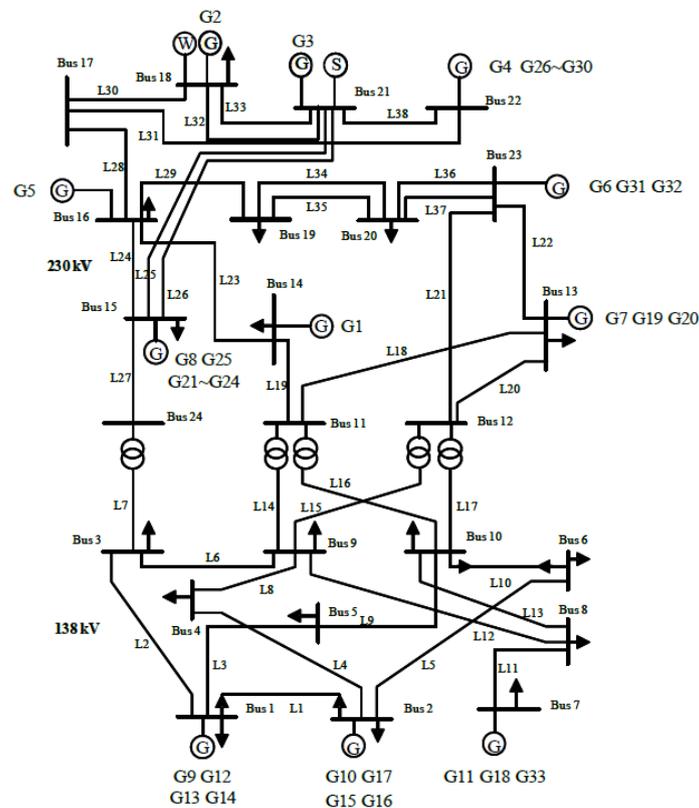


Figure 5. Topology of the RTS-79 testing system.

4.2. Effectiveness Validation of Power System Reliability Assessment of Analytical Models

We constructed multiple reliability index analysis models through different component combinations, and calculated the corresponding reliability of these cases. The calculation results were compared with traditional reliability models to demonstrate the effectiveness and feasibility of the proposed model. The set cases are as shown in Table 1.

Table 1. Detailed information of different cases.

Case	Serial Number of Fault Component
1	G13
2	G16
3	G18
4	L13
5	L18
6	L21

When calculating the model coefficients, the availability of G13, G16, and G18 is set to 0.93, the availability of L13, L18, and L21 is set to 0.9994, and the variance coefficient of convergence accuracy is set to 0.010. After the above reliability evaluation analytical models are established, the reliability indicators of IEEE-RTS can be obtained by directly substituting the availability of the interested components into the corresponding models. By comparing the evaluation results with traditional Monte Carlo simulation methods, the accuracy of the proposed analytical model can be verified. For a more intuitive evaluation of accuracy, the relative error can be calculated using Equation (23):

$$\begin{aligned}
 ER_{LOLP} &= \frac{|LOLP_{AM} - LOLP_{SMCS}|}{LOLP_{SMCS}} \times 100\% \\
 ER_{LOLF} &= \frac{|LOLF_{AM} - LOLF_{SMCS}|}{LOLF_{SMCS}} \times 100\% \\
 ER_{EENS} &= \frac{|EENS_{AM} - EENS_{SMCS}|}{EENS_{SMCS}} \times 100\%
 \end{aligned} \tag{23}$$

The detailed calculation results are shown in Table 2.

Table 2. The detailed calculation results under different cases.

Case	LOLP	LOLF	EENS	ER _{LOLP}	ER _{LOLF}	ER _{EENS}
Case 1	0.00122	2.17	1312.86	0	1.3%	0.79%
Case 2	0.00121	2.32	1314.75	0	0	0.54%
Case 3	0.00121	2.16	1302.45	0	0.26%	0.47%
Case 4	0.00122	2.25	1246.85	0.12%	0	0
Case 5	0.00121	2.18	1252.37	0.56%	0.84%	0
Case 6	0.00122	2.27	1240.68	0.07%	0.15%	0.24%

For the reliability evaluation of power systems based on Monte Carlo simulation, various computational efficiency improvement techniques have been proposed in existing research, including variance reduction techniques, system state analysis acceleration techniques, etc. However, these methods only limited the time required for a single reliability assessment, and repeated evaluations are still very time-consuming. An effective solution to these shortcomings is to establish an analytical calculation model for the variation in system/node reliability indicators with component reliability parameters. This paper uses the total probability formula and conditional probability to separate the reliability parameters to be solved from the reliability index calculation formulas, and establishes an explicit analytical model of system/node reliability index that includes component reliability parameters. In other words, analytical expressions can be seen as fitting the complex nonlinear relationship between reliability indicators and component parameters. In order to reduce computational complexity and improve computational efficiency, analytical models usually simplify the actual power system to a certain extent. These simplifications may lead to differences between the model and the actual system, thereby introducing errors. Regardless of whether the components of interest are generator sets, transmission lines, or both, the proposed analytical model obtains high accuracy reliability indicators, which can control the error within 1.5%.

On the other hand, by analyzing the model, the computational time of repetitive sequential Monte Carlo methods is also avoided. This section compares the time of two

methods in scenarios where reliability parameters vary multiple times. For convenience, this section is not intended to solve a specific optimization problem. Instead, it is assumed that during the optimization process, 40 reliability evaluations need to be conducted for different availability rates of components in cases 1–6. By multiplying the original availability rates of G13 and G16 by 40 pre-set multipliers, a corresponding 40 sets of different availability rates are obtained. For the proposed model, several Monte Carlo simulations are required to determine the coefficients of the model. After determining the model coefficients, the corresponding reliability indicators can be obtained by directly substituting the repair time of each group of components in cases 1–6 into the analytical model. Table 3 presents the proposed model and Monte Carlo method calculation time for 40 different repair times for components.

Table 3. The calculation time comparison using different methods under different cases.

Case	Calculation Time of the Proposed Method/s	Calculation Time of Monte Carlo Simulation/s
Case-1	4246.4	40,379.6
Case-2	4108.7	39,785.1
Case-3	4106.5	41,085.7
Case-4	3789.4	40,526.3
Case-5	3546.2	38,852.9
Case-6	3879.1	39,478.2

According to Table 3, in order to obtain all the reliability evaluation results, the calculation time for Monte Carlo simulation is about 30 days, while the proposed model only takes about 3 days. This is because for each set of component repair times, the former requires running a sequential Monte Carlo simulation once. However, the proposed analytical model only requires 5 temporal Monte Carlo simulations to determine the coefficients of the model. After the coefficients are determined, the reliability evaluation based on the analytical model takes very little time, and the total evaluation process time does not exceed 1 s. The results in Table 3 indicate that once the analytical model is established, the reliability indicators for the given reliability parameters can be directly obtained without the need for time-consuming Monte Carlo simulations.

4.3. Feasibility Verification of CNN Classification Regression Model

To demonstrate the effectiveness of the CNN classification regression method, the following six methods were compared on the test set formed in Section 4.2.

M1: Solve the traditional MCS model as a reference value.

M2: Train a CNN regression model using the two-dimensional matrix vector mentioned in this article as input.

M3: Train a CNN classification regression model using the two-dimensional matrix vector proposed in this article as input.

M4: The two-dimensional matrix vector proposed in this article is transformed into one-dimensional as the input for DNN, and the DNN classification regression model is trained.

M5: Use 0–1 variables to represent the status of generators and transmission lines, and use them as inputs to the DNN along with node load and renewable energy output to train the DNN classification regression model.

Among them, the learning rate of the classification model is taken as 0.0005, β is taken as 0, and the batch size is taken as 50. The SGD gradient descent algorithm is used to iterate 500 epochs. The learning rate of the regression model is taken as 0.005, a is taken as 0.2, and batch size is taken as 50. The Adam optimization algorithm is used, and the convergence criterion is that the loss function is less than 0.8 or reaches 2500 epochs. The structure of a DNN is the same as the fully connected layer in a CNN, and the activation function, optimizer, and learning rate are also consistent with the CNN model.

The accuracy index P_{acc} and F1 score (F1) are used to evaluate the performance of the classification model, and the calculation method is shown in Equations (24) and (25):

$$P_{acc} = \frac{TP + TN}{TP + TN + FP + FN} \quad (24)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (25)$$

Among them, TP is the correctly judged unreliable power supply sample, TN is the correctly judged reliable power supply sample, FP is the misjudged unreliable power supply sample, and FN is the misjudged reliable power supply sample. The performance of the regression model is evaluated using the average absolute percentage error E_{MAPE} and the weighted average absolute percentage error E_{WMAPE} , as shown in Equations (26) and (27):

$$E_{MAPE} = \frac{100\%}{N_t} \sum_{t=1}^{N_t} \left| \frac{LOLP_t^* - LOLP_t}{LOLP_t^*} \right| \quad (26)$$

$$E_{WMAPE} = \frac{100\%}{N_t} \sum_{t=1}^{N_t} \frac{LOLP_t^*}{\sum_{t=1}^{N_t} LOLP_t^*} \left| \frac{LOLP_t^* - LOLP_t}{LOLP_t^*} \right| \quad (27)$$

Here, N_t is the number of test samples. It should be noted that only the LOLP indicator is calculated here. The LOLF and EENS indicators can also be obtained based on this model.

Due to the significant uncertain characteristics of the system and the large range of operating state fluctuations, introducing necessary classification mechanisms before regression analysis can achieve more accurate regression results than just using a regressor, as shown in the comparison of M1 and M2 results in the Table 4. Comparing M2, M3, and M4, it can be found that among several common deep learning methods, CNN has the best-fitting effect, with a classification accuracy of 99.64% and a loss of load estimation deviation of only 1.42%. By comparing M3 and M5, it can be found that the input feature vectors designed in this paper contain specific system parameter information (such as generator capacity, etc.), which is more effective than simply using 0–1 variables to represent component state information, achieving higher accuracy and precision.

Table 4. Comparison of computational effectiveness using different methods.

Method	$P_{acc}/\%$	F1/%	$E_{MAPE}/\%$	$E_{WMAPE}/\%$
Method-1	100	100	0	0
Method-2	/	/	3.76	0.85
Method-3	99.64	99.19	1.42	0.16
Method-4	99.32	98.47	1.75	0.34
Method-5	98.38	96.38	6.16	0.87

At the same time, the time required for M1–M5 to evaluate the test set samples was compared. M1 took 1042.21 s to calculate the optimal load shedding model, while M2 took 13.36 s, which is only 1/78 of M0. At the same time, the time consumption of M3, M4, and M5 is 19.09 s, 16.78 s, and 14.87 s, respectively, and there is no significant difference in time consumption compared to M2. Considering the accuracy requirements, overall, the method M2 proposed in this article can maintain high solving accuracy while having good acceleration effects.

Furthermore, three different scale testing systems, including the Reliability Test System (IEEE RTS), Western Systems Coordinating Council (WSCC) Test System, and an actual provincial power grid in southwestern China, are used to verify the feasibility of the proposed model. The above three testing systems are named system 1, 2, and 3, respectively. Firstly, the actual reliability parameters are obtained using Monte Carlo simulation, and

then the reliability parameters under the same operating conditions are obtained using the data-driven strategy proposed in this paper. The accuracy is calculated by comparing the results obtained by different methods. The detailed calculation results are shown in Figure 6. It can be observed from Figure 6 that the method proposed in this paper can still achieve high computational accuracy, with a calculation accuracy of over 97.5%, while significantly improving computational efficiency, meeting the needs of practical engineering applications.

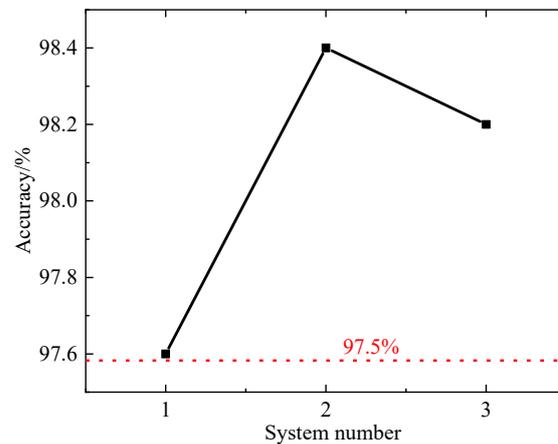


Figure 6. Calculation accuracy using different scale testing systems.

5. Conclusions

In the process of solving the inverse problem of reliability evaluation, sometimes it is necessary to use Monte Carlo simulation to perform multiple reliability evaluations on different values of the reliability parameters to be solved, which is extremely time-consuming. In response to the above issues, this chapter proposes an analytical model for reliability indicators related to component reliability parameters based on time-series MCS. Then, an input feature vector reflecting the operating status of the system was constructed, and a reliability index regression model based on a CNN was constructed. Secondly, in response to the diverse operating scenarios of the system, it is extended to a CNN classification–regression model to improve training efficiency and accuracy. The calculation results indicate the following.

- (1) The proposed reliability index analysis model can intuitively display the functional relationship between reliability indicators and the reliability parameters to be solved. The proposed analytical model can accurately and efficiently calculate the system reliability indicators that vary with the reliability parameters of the components, with an error of no more than 2% compared to the Monte Carlo simulation.
- (2) The proposed CNN classification–regression model can significantly improve computational efficiency while ensuring computational accuracy, with a computation time for only 1.3% of the traditional Monte Carlo simulation method.

Most existing reliability assessment methods for power systems focus on reliability assessment at a single time scale, such as long-term or short-term reliability, etc. However, in practical operation, the power system may face reliability issues at multiple time scales simultaneously. Therefore, studying how to comprehensively consider the reliability of multiple time scales and establish a multi-time scale reliability evaluation method for power systems is a challenging and practical research direction. On the other hand, the current power system is gradually transitioning towards the integrated energy system, and the complexity of the system is further increasing. How to establish a reliability evaluation model suitable for integrated energy systems is also a focus of future research.

Author Contributions: Conceptualization, J.Z. and B.W.; methodology, B.W. and H.M.; software, H.M. and Y.L.; validation, M.Y.; writing—original draft preparation, J.Z., B.W., H.M., Y.L., M.Y., H.W. and F.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of State Grid Corporation Limited (Contract No. 5400-202155497A-0-5-ZN).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: Author Yunshuo Li was employed by the company China Electric Power Research Institute. Author Meilin Yang was employed by the company State Grid Huaining Power Supply Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Jain, T.; Verma, K. Reliability based computational model for stochastic unit commitment of a bulk power system integrated with volatile wind power. *Reliab. Eng. Syst. Saf.* **2024**, *244*, 109949. [[CrossRef](#)]
- Firouzi, M.; Samimi, A.; Salami, A. Reliability evaluation of a composite power system in the presence of renewable generations. *Reliab. Eng. Syst. Saf.* **2022**, *222*, 108396. [[CrossRef](#)]
- Zhang, G.; Liu, H.; Zhang, J.; Yan, Y.; Zhang, L.; Wu, C.; Hua, X.; Wang, Y. Wind power prediction based on variational mode decomposition multi-frequency combinations. *J. Mod. Power Syst. Clean Energy* **2019**, *7*, 281–288. [[CrossRef](#)]
- Erdinç, F.G.; Çiçek, A.; Erdinç, O. Resiliency-Sensitive Decision Making Mechanism for a Residential Community Enhanced with Bi-Directional Operation of Fuel Cell Electric Vehicles. *Energies* **2022**, *15*, 8729. [[CrossRef](#)]
- Liu, X.; Wang, C.; Kong, X.; Zhang, Y.; Wang, W.; Lee, K.Y. Tube-Based Distributed MPC for Load Frequency Control of Power System with High Wind Power Penetration. *IEEE Trans. Power Syst.* **2024**, *39*, 3118–3129. [[CrossRef](#)]
- Dai, Y.; Noebels, M.; Preece, R.; Panteli, M.; Dobson, I. Risk Assessment and Mitigation of Cascading Failures Using Critical Line Sensitivities. *IEEE Trans. Power Syst.* **2024**, *39*, 3937–3948. [[CrossRef](#)]
- Li, G.; Huang, Y.; Bie, Z.; An, J.; Sun, S.; Qiu, Q.; Lei, Y. Review and prospect of operational reliability evaluation of integrated energy system. *Electr. Power Autom. Equip.* **2019**, *39*, 12–21.
- Hamoud, G.A. Reliability Assessment of Distribution Stations Considering Spare Transformer Sharing. *IEEE Trans. Power Syst.* **2024**, *39*, 2264–2270. [[CrossRef](#)]
- Du, J.; Guo, R.; Li, C.; Huang, D.; Hu, H.; Zhang, L. An importance-control method for power system reliability evaluation. *Autom. Electr. Power Syst.* **2015**, *39*, 69–74.
- Yin, H.; Wang, Z.; Liu, Y.; Qudaih, Y.; Tang, D.; Liu, J.; Liu, T. Operational Reliability Assessment of Distribution Network with Energy Storage Systems. *IEEE Syst. J.* **2023**, *17*, 629–639. [[CrossRef](#)]
- Zhang, W.; Che, Y.; Liu, Y. Improved Latin hypercube sampling method for reliability evaluation of power systems. *Autom. Electr. Power Syst.* **2015**, *39*, 52–57.
- Hu, D.; Sun, L.; Ding, M. Integrated Planning of an Active Distribution Network and DG Integration in Clusters Considering a Novel Formulation for Reliability Assessment. *CSEE J. Power Energy Syst.* **2023**, *9*, 561–576.
- Carvalho, L.d.M.; Gonzalez-Fernandez, R.A.; da Silva, A.M.L.; da Rosa, M.A.; Miranda, V. Simplified Cross-Entropy Based Approach for Generating Capacity Reliability Assessment. *IEEE Trans. Power Syst.* **2013**, *28*, 1609–1616. [[CrossRef](#)]
- Zhao, Y.; Tang, Y.; Li, W.; Yu, J. Composite Power System Reliability Evaluation Based on Enhanced Sequential Cross-Entropy Monte Carlo Simulation. *IEEE Trans. Power Syst.* **2019**, *34*, 3891–3901. [[CrossRef](#)]
- Green, R.C.; Wang, G.L.; Alam, M. Composite power system reliability evaluation using support vector machines on a multi-core platform. In Proceedings of the IEEE International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011; pp. 2586–2592.
- Cai, B.; Sheng, C.; Gao, C.; Liu, Y.; Shi, M.; Liu, Z.; Feng, Q.; Liu, G. Artificial Intelligence Enhanced Reliability Assessment Methodology with Small Samples. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *34*, 6578–6590. [[CrossRef](#)] [[PubMed](#)]
- Wei, W.; Cao, Y.; Hou, K. Reliability Assessment of Hybrid AC/DC Distribution Systems Using Independent Contingency States Reduction Technique. *IEEE Trans. Smart Grid* **2023**, *14*, 3439–3449. [[CrossRef](#)]
- Xu, X.; He, C.; Su, C.; Ma, J. Chenya and M. Jian, Research on System Resource Optimization of Distribution Network Producer-Consumer System Considering Resilience and Economy. *IEEE Access* **2023**, *11*, 63947–63957. [[CrossRef](#)]
- Zhu, Q.; Yan, Z.; Jia, Y.; Wang, L. Prediction of transmission operating reliability. *Autom. Electr. Power Syst.* **2010**, *34*, 18–22.
- Lei, X.; Yang, Z.; Yu, J.; Zhao, J.; Gao, Q.; Yu, H. Data-Driven Optimal Power Flow: A Physics-Informed Machine Learning Approach. *IEEE Trans. Power Syst.* **2021**, *36*, 346–354. [[CrossRef](#)]
- Yang, Y.; Yang, Z.; Yu, J.; Xiang, M.; Wang, H. Fast Analysis of N-1 Contingency Screening with Uncertainty Scenarios Based on Deep Learning. *Proc. CSEE* **2021**, *41*, 2716–2725.

22. Yu, J.; Yang, Y.; Yang, Z.; Xiang, M.; Xie, S.; Zhou, P.; Ren, P.; Zhang, Y. Fast probabilistic energy flow analysis based on deep learning. *Proc. CSEE* **2019**, *39*, 22–30.
23. Pan, X.; Zhao, T.; Chen, M.; Zhang, S. DeepOPF: A Deep Neural Network Approach for Security-Constrained DC Optimal Power Flow. *IEEE Trans. Power Syst.* **2021**, *36*, 1725–1735. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.