Received 2 June 2023, accepted 16 June 2023, date of publication 21 June 2023, date of current version 28 June 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3288341

RESEARCH ARTICLE

Resilience Assessment of Urban Distribution Network Under Heavy Rain: A Knowledge-Informed Data-Driven Approach

KE LI[1], JIE MA[1], JIANLONG GAO[1], CHANGQING XU[1], WENFENG LI[1], YUBIN MAO[1], AND SHIGONG JIANG[2]
[1]Economic Research Institute, State Grid Henan Power Supply Company, Zhengzhou 450052, China
[2]Economic Research Institute, State Grid Power Supply Company, Beijing 102211, China

Corresponding author: Jianlong Gao (jianlonggao123@sina.com)

This work was supported by the Headquarters Management Science and Technology Project of State Grid Corporation of China under Grant 5108-202218280A-2-130-XG.

ABSTRACT Heavy rains pose a great threat to the reliable and secure power supply of urban distribution networks. A knowledge-informed data-driven resilience assessment approach is proposed to evaluate urban distribution networks' abilities to resist heavy rains. Firstly, the rainstorm waterlogging process is simulated to obtain the rainstorm intensity and rainfall process, providing the input for the data-driven model. Then, input variables are grouped guided by expert knowledge, and a dynamic and static data-driven model is constructed to predict the line outages based on historical data. Finally, the Monte Carlo simulation method integrated with the data-driven model is developed to assess the resilience of urban distribution networks and the number of line outages is selected as the evaluation metric. The effectiveness of the proposed method is sufficiently validated by the historical data of an urban distribution network.

INDEX TERMS Data-driven, resilience, assessment method, heavy rains, distribution network.

I. INTRODUCTION

In recent years, extreme natural disasters which occur frequently have seriously affected people's daily life and the safe operation of power system. The 2008 southern China ice disaster hit more than 20 provinces and left hundreds of towns without power. In 2012, Hurricane Sandy caused massive power outages in the eastern United States. Except for the ice disaster and hurricane, rainstorm has also seriously affected people's lives. In July 2021, an unusually heavy rainstorm occurred in Henan Province, leading to severe waterlogging in several cities and heavy economic losses and casualties [1]. In July 2022, heavy rainfall hit Zhengzhou again, causing water flooding in many roads, posing challenge to the safety of urban power supply. In May 2023, 10 counties in Chinese Jiangxi province have been hit by heavy rainstorms, resulting in swollen rivers and flooded farmland. Large-scale urban waterlogging has affected 536,000 people and caused a number of transmission lines were damaged. Power outages caused by heavy rainstorms have exposed the weakness of the power system's ability to cope with low probability and high impact events [2]. Therefore, an effective resilience assessment of the power distribution network's ability to withstand rainstorms will be meaningful to prevent and reduce losses [3].

However, there are relatively few studies on the resilience evaluation of distribution network for rainstorm disaster at home and abroad. The existing studies mainly focus on disaster science and urban safety. At present, some studies use uncertain methods to assess the flood disaster risk, such as fuzzy mathematics method and gray system method [4], [5], [6]. Some studies also use GIS technology to build storm flood simulation model and loss assessment model [7], [8]. Most of the above studies require detailed geographic information data and hydrological information, etc. The established models mainly reflect the loss at the social and economic level, but can't reflect the impact of rainstorm and waterlogging on the urban distribution network, nor

The associate editor coordinating the review of this manuscript and approving it for publication was Baoping Cai.

the dynamic evolution process of rainstorm and waterlogging. Urban rainstorm waterlogging disaster is the result of the disaster causing factor's risk and disaster undertaker's vulnerability. Disaster causing factor risk refers to disaster intensity and occurrence probability. Vulnerability refers to the resistance ability of disaster affected body to disaster factors [9]. Problems to be solved are to determine the rainstorm waterlogging model which can effectively describe rainstorm intensity and the evaluation method which can reflect the disaster situation of distribution network.

At present, there is no unified standard for the establishment of rainstorm waterlogging model at home and abroad. The current urban flood models include hydrologic model, hydrodynamic model and simplified model [10]. Hydrological model mainly includes Storm Water Management Model (SWMM), University of Cincinnati Urban Runoff Model (UCURM) and Illinois Urban Drainage Area Simulator (ILLUDAS). Up to now, some studies use GIS technology to conduct coupling modeling with SWMM [11], [12]. No matter what kind of hydrological model can't provide the flood dynamic process at the non-node of the model, so it is not suitable to assess the resilience of the distribution network. In order to improve the accuracy and credibility of the model, some studies have proposed hydrodynamic models [13], [14]. However, most of these models require complex data, difficult modeling, and low computational efficiency. Compared with the other two models, the simplified model requires fewer data and is relatively simple, which can simulate the spatio-temporal process of runoff. Zeng et al. combined cellular automata model (WCA2D) with SWMM model to rapidly simulate urban rainstorm waterlogging [15]. References [16] and [17] put forward a submergence model applicable to large urban waterlogging area based on GIS digital topographic analysis, which effectively simulates the submergence range and submergence depth of the study area, and provides scientific and technological support for disaster warning and forecast.

In recent years, some achievements have been made in the research of damage mechanism, risk assessment and countermeasures of power system caused by extreme disasters [18]. However, the assessment and countermeasures of extreme rainstorm disasters on urban distribution network are still in the initial stage, and relevant studies are relatively few [19]. The existing research about distribution network disaster assessment mainly focuses on typhoon disaster. The evaluation methods can be divided into physical methods based on disaster mechanism and data-driven methods based on historical data. However, the disaster mechanism of extreme disasters which is difficult to describe by definite explicit physical laws is complex, so the application of physical methods is limited [20]. In recent years, with more and more data collected from Electric power companies, data-driven methods have been applied to power system fault prediction under typhoon disasters. Reference [21] applied the accelerated failure time model to predict the downtime in the case of hurricane and ice storm, and the model performance is compared with different combinations of input variables. Guikema et al. employed a data mining method to predict the number of customers without power based on meteorological data, geographic data and power grid data [22]. Random forest algorithm was applied in reference [23] to predict the number of customers' power outages and damaged poles. Compared with traditional physical mechanism methods, data-driven methods have advantages in prediction accuracy and scalability [24], [25]. However, the existing data-driven methods are not interpretable and can't be directly applied to assess the disaster bearing capacity of the distribution network.

Considering the shortage of existing research, this paper proposes a knowledge-informed data-driven method to realize urban resilience assessment under rainstorm disasters. The main innovations of this paper include the following points,

1) Propose a rainstorm waterlogging model aiming at the rainfall condition of China.

2) Establish a two-channel model to predict transmission line interruptions caused by rainstorms based on static and dynamic data.

3) Integrate the data-driven line outage prediction model with Monte Carlo simulation to implement the resilience assessment of urban distribution networks under rainstorms.

The rest of the paper is organized as follows: Section II designs the rainstorm waterlogging model. Section III illustrates the dynamic-static data-driven method. Section IV explains the Monte Carlo simulation to assess the resilience of distribution networks. Then, section V demonstrates the developed method in an actual distribution network of a city in southern China. Finally, section VI draws the conclusion.

II. RAINSTORM WATERLOGGING MODEL

The urban rainstorm waterlogging disaster with features of strong rain, short duration, linkage and mutability of loss is mainly reflected in the rainfall process and the runoff-producing with flow confluence process. In order to better simulate the rainstorm, the rainstorm waterlogging model is constructed according to three parts: the rainstorm intensity, rainfall process and runoff-producing with flow confluence.

A. RAINSTORM INTENSITY

Up to now, the expression form of the basic rainstorm intensity formula in reference [26] is generally accepted,

$$q = \frac{A(1 + C \lg P)}{(t + b)^n} \tag{1}$$

where A and C are the rainfall intensity parameter and intensity changing parameter; $P$is the rainfall return period (year); b is the rainfall duration correction parameters; n is the rainstorm attenuation parameter; $q$is the rainstorm intensity (mm/min); $t$is rainfall duration (min).

According to the historical rainfall data in different regions, the parameters of the rainstorm intensity formula are also different [27], [28]. In this paper, the parameters based on the historical rainfall data of a city in southern China are available on the public platform [29], and the final calculation formula is as follows,

$$q = \frac{21.6672(1 + 0.438 \lg P)}{(t + 11.259)^{0.750}} \tag{2}$$

B. RAINFALL PROCESS

There is no universally agreed rain pattern for the design of rainfall processes. The existing rainfall process includes uniform rain type, triangular rain type and Chicago rain type. Among them, the calculated result of uniform rain pattern is smaller than the actual precipitation and the peak flow of the triangular rain type process is obviously affected by the rainfall duration value [30]. According to the relationship between intensity duration and frequency, the Chicago rain type process can obtain the uneven rain pattern and simulate the rainstorm well in most cities [31]. Therefore, in this paper, the Chicago rain type is selected for the design of rainfall process.

Compared with bimodal rainfall, unimodal rainfall is more likely to cause heavy rain due to its rainfall concentration and distribution network faults. Therefore, this paper focuses on the unimodal rainfall. On the basis of calculating the rainfall amount, the rainfall process is designed, and the designed rainfall intensity at every moment during the rainstorm is expressed as the pre-peak rainfall process and post-peak rainfall process respectively [26],

$$\begin{aligned} Pre-peak: q(t_1) &= \frac{A}{(t_1/r + b)^n}\left[1 - \frac{nt_1}{t_1 + rb}\right] \\ Post-peak: q(t_2) &= \frac{A}{(t_2/(1-r) + b)^n} \\ &\quad \times \left[1 - \frac{nt_2}{t_2 + (1-r)b}\right] \end{aligned} \tag{3}$$

where t_1 is the pre-peak duration; t_2 is the post-peak duration (min); $q(t_1)$ represents the rainfall intensity at the pre-peak moment t_1(mm/min); $q(t_2)$ represents the rainfall intensity at the post-peak moment t_2(mm/min); the values of index A, b, and n are local experience parameters. $r \in (0, 1)$ is the rain peak coefficient of rainstorm, representing the time of peak rainfall. The rain peak coefficient of rainstorm divides the whole rainfall time series into pre-peak time series and post-peak time series, whose value is obtained by local rainfall data statistics. The following figure shows the typical rainfall process of Chicago rainfall type with 2h rainfall duration under four different return periods.

C. RUNOFF-PRODUCING WITH FLOW CONFLUENCE

After the rainstorm falls to the ground, part of the rainfall will infiltrate into the ground and not participate in the process of runoff producing with confluence. So, the remaining rainfall will become the rainfall runoff generated by the rainstorm. Rainfall runoff of the same intensity in different land types is different as a result of different land types in cities have

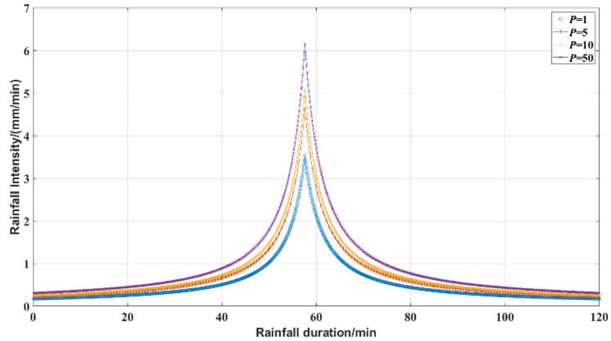


FIGURE 1. The schematic diagram of typical rainfall process of chicago rainfall pattern.

different soil infiltration rates. In this paper, the rainfall which deduct the runoff coefficient is used to simulate the flow production.

The form of waterlogging caused by heavy rain is "passive inundation" that all points with elevation lower than the given water level are included in the flood inundation area. Therefore, in this paper, the equal volume method after rasterization is adopted to calculate the water depth of flooded area [32]. The calculation formula of the equal volume method is as follows,

$$W = \sum_{\mathrm{i}=1}^{\mathrm{N}} \left[E_w - E_g(\mathrm{i})\right] \Delta\sigma \tag{4}$$

where Wrepresents the volume of water within the flooded area; N is the total number of grids; E_w is the water surface elevation; $E_g(i)$ is the ground elevation; i expresses the ith grid and $\Delta\sigma$ is the waterlogged area unit.

In this paper, the dichotomy method is used to solve the elevation of ponding water surface. Firstly, define the value range of E_{w} $(E_{\mathrm{min}}, E_{\mathrm{max}})$ to ensure the value of E_{min} is less than or equal to the minimum ground elevation of the ponding area and E_{max} is greater than or equal to the sum of the maximum ground elevation and the rainfall height in this period. Secondly, update the elevation value of ponding water surface according to the following steps.

1) Set the initial value of E_{w} as $E_{\mathrm{w}} = E_{\mathrm{a}} = (E_{\mathrm{min}} + E_{\mathrm{max}})/2$.

2) If $E_{\mathrm{w}} > E_{\mathrm{g}}(\mathrm{i})$, W adds $[E_{\mathrm{w}} - E_{\mathrm{g}}(\mathrm{i})]\Delta\sigma$. Otherwise, the value of W remains unchanged.

3) Update the total water volume value after all grid calculations are completed W.

4) Compare W with the calculated surface runoff Q. If the difference between Q and W meets the allowable error, the elevation value of ponding water surface E_{w} can be obtained. Otherwise, go to the next step.

5) If $Q > W$, update $E_{\mathrm{min}} = E_a$. Otherwise, update $E_{\mathrm{max}} = E_a$, return to step 2) for calculation.

After the water surface elevation $E_{\mathrm{w}}$is calculated, the corresponding ponding water depth value for each grid can be calculated as $D = E_{\mathrm{w}} - E_{\mathrm{g}}$.

D. SUMMARY

In this section, the selection of rainstorm intensity formula is introduced at first, then the typical rainfall process used is introduced, finally the simulation of runoff-producing with flow confluence process based on rainfall process results is elaborated in detail. In view of the lack of rainstorm waterlogging data in some areas, the rainstorm intensity calculation, rainfall process and runoff producing with flow confluence simulation can provide effective input data for the data driving link, and improve the accuracy of the evaluation model.

III. LINE OUTAGES ESTIMATION OF URBAN DISTRIBUTION NETWORK

Aiming at multivariate disaster data, this section uses feedforward neural network and gate recursive unit (GRU) to extract features of static variables and dynamic variables respectively, and then uses a multi head attention network to fuse all features to establish a mapping relationship with the total number of line interruptions [33], forming a dynamic-static dual driven distribution network disaster prediction method.

A. DYNAMIC AND STATIC VARIABLES

The static variable and dynamic variables form the disaster causing characteristic vector as the input, meanwhile the response variable forms the predictive value as the output. The relationship between them is shown in Fig.2. According to the time domain change feature of the characteristic variables, the input characteristics of the distribution network line interruption prediction model are divided into static variables and dynamic variables. Considering the difference of the time series characteristics of the two groups of variables, the deep features are extracted through the multi-layer perceptron model and long-short-term memory network respectively, and the number of disconnected lines is obtained through the multi-head attention network. This dynamic and static dual driven method improves the prediction performance and interpretability of distribution network line interruption model through the combination of expert domain knowledge and data driven method.

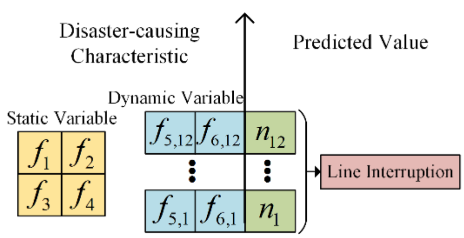


FIGURE 2. The schematic diagram of dynamic and static data.

B. RESPONSE VARIABLES

The objective of this prediction method is to help determine the line interruptions situation of urban distribution network under rainstorm disaster, to formulate relative emergency response plan and element strengthening plan for rainstorm disaster. Therefore, it is our research goal to predict the total number of disconnected lines during the whole process of rainstorm disaster. To realize this goal, the total number of power distribution network line interruptions caused by each rainstorm $\sum_{i=1}^{T} n_i$ is taken as the response variable in this paper, meanwhile, n_i represent the number of line interruptions in each period, and T is the duration of rainstorm.

On the basis of selected variables, outliers are removed from the dataset. Then we fill the missing data with linear interpolation, and then uses minimum-maximum normalization to process the eigenvalues of all samples to avoid the problem of model convergence caused by feature level deviation.

C. PREDICTION MODEL FOR LINE DISCONNECTED

In order to predict the total number of line interruptions in the distribution network under rainstorm disasters, this paper proposes a dual path deep neural network to conduct collaborative analysis on static data and dynamic data. The overall structure of the model is shown in Fig.3. Dual path prediction model mainly includes three parts: static feature extraction, dynamic feature extraction and feature fusion process. The structure and parameter optimization method of the prediction model are described below.

1) STATIC FEATURE EXTRACTION

For the static feature extraction, time invariant static variables constitute the input data. The feedforward neural network is densely connected and does not have any feedback loop, and the corresponding information propagation process is unidirectional. Therefore, feedforward neural network (FFNN) is applied to deal with static data without time correlation. In this paper, the feedforward neural network is composed of stacked linear layers with nonlinear activation functions. Considering that LeakyReLU function can effectively avoid the problem of gradient disappearance and ReLU deactivation, so it is used as a nonlinear activation function after the linear layer. Meanwhile, batch normalization layers are added after each linear layer to accelerate the training process and mitigate the vanishing gradient phenomenon.

2) DYNAMIC FEATURE EXTRACTION

For the dynamic feature extraction, the corresponding input is composed of dynamic variables in the continuous process of rainstorm disaster. A variant of the recurrent neural network (RNN), GRU can extract, store, and pass the useful information of time series data along the neural network, and has better ability in analyzing sequence data with long-term dependencies. Thereby, GRU network [34] illustrated in Fig.3 is used to extract and analyze the time-varying characteristics of dynamic data. Furthermore, in order to reduce over fitting, the dropout layer is added after each hidden layer of GRU network. A GRU cell contains two gates, namely the reset gate and the update gate. The reset gate r_t determines how much the preceding hidden state h_{t-1} contributes to the candidate hidden state h_t. The update gate z_t controls how much the preceding hidden state h_{t-1} can be directly brought to the current state h_t. The specific information processing in the

GRU cell at time step t can be described as follows,

$$\begin{aligned} r_t &= \sigma(U_r x_t + W_r h_{t-1} + b_r) \\ z_t &= \sigma(U_z x_t + W_z h_{t-1} + b_z) \\ h_t &= tanh(U_h x_t + W_h(r_t \odot h_{t-1}) + b_h) \\ h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot h_t \end{aligned} \tag{5}$$

where σ is the sigmoid activation function, U_r, W_r, U_z, W_z, U_h and W_h are weight matrices, b_r, b_z and b_h are bias vectors.

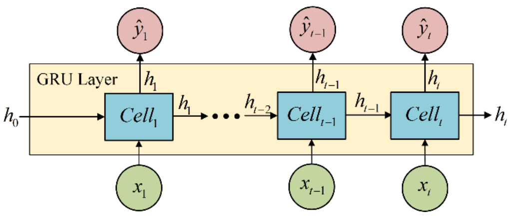


FIGURE 3. The structure of GRU networks.

3) FEATURE FUSION PROCESS

For the feature fusion process, the extracted features of static data and dynamic data are cascaded to form the input data, and the multi-head attention network [35] is used for dynamic and static feature fusion analysis. Multi-head attention network is composed of multiple representation subspaces, and the corresponding number in this paper is h in Fig.3. In each representation subspace, Q, K and V are matrices formed by static and dynamic features, and then the attention function based on scaled dot-product function is used to analyze the potential relationship between them from different perspectives.

The formula can be written as,

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}}V) \tag{6}$$

where d_k is the dimension of the matrix K.

Multi-head attention firstly maps the input to different representation subspaces and then compute outputs of attention functions in parallel. Finally, multi-head attention concatenates them and obtains the final output through a linear layer. The corresponding equations are as follows,

$$\begin{aligned} head_i &= \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \\ \text{Multihead} &= \text{Concat}(head_1, \ldots, head_h)W^O \end{aligned} \tag{7}$$

where, W_i^Q, W_i^K, W_i^V are matrices mapping to the subspace of $head_i$ and $i = 1, 2, \ldots, h$, h is the number of heads and W^O is the final linear layer.

Finally, the outputs of different representation subspaces are cascaded into vectors, and mapped into model outputs. In Fig.3, the total number of disconnected lines in the distribution network is the output of the dual path prediction model.

In order to solve the parameters of the dual path prediction model, the mean square error function is selected as the loss function to measure the difference between the predicted number of line interruptions and the actual number of line interruptions. The formula is as follows,

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_p - y_r)^2 \tag{8}$$

where $L$is the value of loss function; $N$is the total training set data; y_p is the predictive value of the number of disconnected lines; y_r is the actual value of the number of disconnected lines.

On the basis of Eq. (8), the error over time back propagation algorithm is applied to obtain the gradient value of each parameter in the model by the loss function, and the parameter value is updated with the gradient descent strategy. The gradient descent strategy chosen in this paper is Adam algorithm [36], and the parameter update formula is,

$$m_t = (\beta_1 m_{t-1} + (1 - \beta_1)\nabla_\theta L)/(1 - \beta_1) \tag{9}$$

$$v_t = (\beta_2 v_{t-1} + (1 - \beta_2)(\nabla_\theta L)^2)/(1 - \beta_2) \tag{10}$$

$$\theta_t = \theta_{t-1} - \alpha m_t/(\sqrt{v_t} + \varepsilon) \tag{11}$$

wherem_t and v_t are biased first moments and biased second moments respectively, α, β_1 and β_2 are the hyperparameters of Adam algorithm.

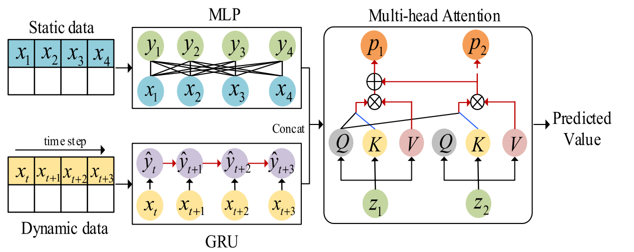


FIGURE 4. The structure of two-channel prediction model.

D. EVALUATION METRICS FOR PREDICTION MODEL

In order to assess the performance of the dual channel prediction model, the mean square error (MSE) and mean absolute error (MAE) are selected to evaluate the model prediction performance using test set samples.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_p - y_r)^2 \tag{12}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_p - y_r| \tag{13}$$

where $n$represents the total number of data in the test dataset; y_r and y_p represent the real value and predicted value of line interruptions data, respectively.

E. SUMMARY

In this section, we first introduce the dynamic data, static data and response variables used in the data driven link. Then we introduce the dual path model of distribution network line interruption prediction through dynamic and static dual data-driven. Finally, evaluation indexes are given to measure the effectiveness of the prediction model.

IV. ASSESSMENT PROCESS OF URBAN DISTRIBUTION NETWORK RESILIENCE

In this paper, the Monte Carlo simulation method is used to sample the rainstorm in different rainfall return periods. After each sampling, the rainstorm waterlogging model is calculated. Then, the dynamic static dual data-driven model is used to estimate the disaster situation of the distribution network. Finally, the evaluation result of the resilience of the urban distribution network with the number of line disconnected as the indicator is obtained. The specific process is as follows.

1) Establish a probability model and set up the probability model of rainstorm occurrence in different return periods according to the return period of rainstorm with different intensity.
2) Judge whether rainstorm in different return periods occurs in this cycle by comparing the generated random numbers with the reciprocal of the return period.
3) Substitute the rainstorm return period corresponding to the occurred rainstorm in 2) into the rainstorm intensity formula according to (2) for calculation.
4) Calculate the rainfall process according to (3) and runoff-producing with flow confluence process according to (4) in terms of the return period of the current rainstorm.
5) Take the rainstorm intensity, rainfall process calculation results and accumulated water depth calculated in the process of runoff-producing with flow confluence as the data driven input of prediction model for line disconnected to assess the disaster situation of the distribution network.
6) Repeat the steps 2) - 5) for many times to calculate the mean and variance of the number of disconnected lines in the distribution network based on the predicted results of line outages in each simulation.

The flow chart is shown in Fig.5.

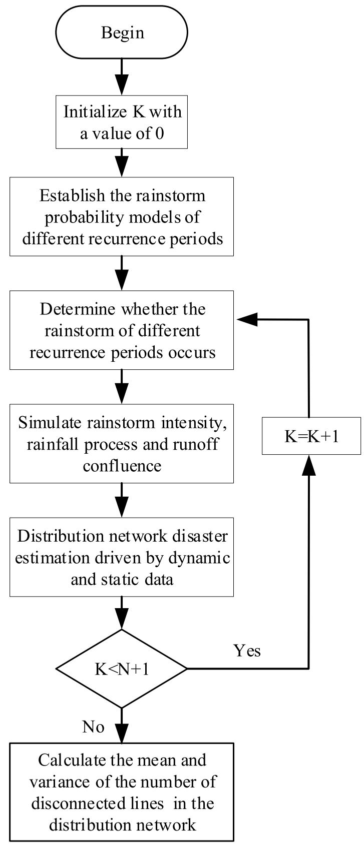


FIGURE 5. The process of resilience assessment.

TABLE 1. Static variables and dynamic variables.

Variable type	Variable name	Data source
Static variable	Terrain type	Public information platform
	Forest coverage rate	Public information platform
	Rainstorm intensity	Electric power company
	Number of customers	Electric power company
Dynamic variable	Rainfall process	Rainstorm Waterlogging Model
	Ponding water depth	Rainstorm Waterlogging Model

V. CASE STUDY

This paper selects the data of a city in southern China as historical training and test data to solve the parameters of the dynamic-static data-driven prediction model, and then generates a rainstorm of certain intensity randomly through the rainstorm waterlogging model to assess the resilience of the distribution network at the district level.

A. DATA DRIVEN MODEL

The static variables used in the case study mainly include geographic data and power grid data. In terms of geographical data, terrain type and forest coverage are selected to represent the geographical environment, which can be collected from the public information platform. Considering that terrain types are category variables, if only assign fixed values to each category, it is easy to cause the problem of partial order. Therefore, this paper adopts the unique heat coding to deal with terrain types. In terms of grid data, the number of users is an important factor affecting the interruption of distribution network lines, because the number of users represents the power consumption level of the study area, to a certain extent. At the same time, the rainstorm intensity should also be selected to express the disaster characteristics of the rainstorm. Because the static variable is time invariant, that is, the impact on the line interruption under the rainstorm disaster is stable during the rainstorm duration. Therefore, this paper selects the above four static variables, namely, scalar data f_1 to f_4, to form the static part of the disaster causing feature vector in Fig.2.

Dynamic variables mainly include meteorological data that can describe precipitation characteristics. The precipitation characteristics mainly include the rainfall per ten minutes (rainfall process) in the study area and the ponding water depth value of the studied area. The value of water depth is the value at each moment, with ten minutes as the time interval. During the duration of rainstorm disaster, the dynamic variable is time-varying, so the impact on line interruptions is time-varying and cumulative. Therefore, the dynamic data will constitute the dynamic part of the disaster feature vector. In Fig.2, $f_{5,i}$ and $f_{6,i}$ respectively represent the above dynamic variables during the rainstorm duration, and i represents the time index.

The dataset is divided into training, validation and test dataset according to the ratio of 6:2:2, which are used to fit model parameters, determine model hyperparameter and test model performance respectively. According to the method proposed in Section III of this paper, the data-driven model is trained. The effectiveness of the model is mainly evaluated by mean square error and mean absolute error.

In order to better compare the effectiveness of the proposed methods, this paper compares the index calculation results of the proposed methods with those of random forest algorithm, support vector machine method, XGBoost method and two ablation experiments. The hyperparameter configuration of the proposed method and other three comparison methods are provided in Table 2, and the specific results are shown in Fig.6.

TABLE 2. The hyperparameters of all methods.

Method	HYPERPARAMETER
Proposed methods	GRU: (3, 64), MLP: (3, 128), Attention: 8
Ablation experiment1	MLP: (3, 128)
Ablation experiment2	GRU: (3, 64)
Random forest	Tree: 100, Min_samples_leaf: 5
Support vector machine	Regularization: 20, Kernel: RBF
XGBoost	Tree: 100, Max_depth: 4

In terms of the proposed method, the num of layers of GRU and MLP network are both set to 3, and the number of each hidden neuron are determined to 64 and 128, respectively. The subspace of the attention network is set to 8 after parameter tuning. In terms of the Support vector machine, 20 is selected as the regularization parameter to punish the overfitting phenomenon, and the radial basis function is adopted as the kernel function for the intrinsic nonlinear relationship. In terms of the random forest and the XGBoost, the number of trees are both set to 100, and 5 samples are required to be split into an internal node for random forest. The maximum tree depth of XGBoost is set to 4 and the ablation experiments are conducted using the GRU and MLP networks with the same num of layers as the proposed method respectively.

The results show that the mean square error and mean absolute error of the proposed method are better than those of other methods. In terms of mean square error, the proposed method is nearly 2 times smaller than the random forest algorithm, 2.5 times smaller than the support vector machine algorithm, nearly 2 times smaller than the XGBoost algorithm, and smaller than the two ablation experiments. Meanwhile, in terms of the average absolute error, the proposed method is the smallest of the six algorithms, but

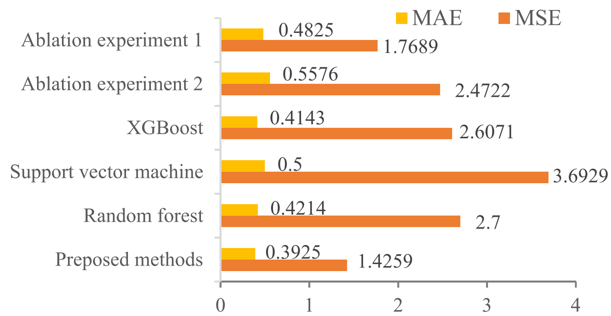


FIGURE 6. The result of the test set.

the difference between the six algorithms is not significant. Therefore, the dynamic-static dual data-driven model proposed in this paper is more feasible to assess the resilience of distribution network under rainstorm.

B. ASSESSMENT RESULTS OF DISTRIBUTION NETWORK RESILIENCE

In order to ensure the accuracy of the prediction results, the simulation time is selected as 1000 years in the Monte Carlo simulation experiment of this example. Then, in each year, judge whether there is rainstorm of this return period in this year according to the probability of rainstorm occurrence in each return period. If there is rainstorm in this year, calculate the rainstorm intensity according to Eq. (2) and (3) with the corresponding return period. After the calculation is completed, conduct runoff-producing with flow confluence simulation to calculate the ponding depth. In order to simplify the calculation, the typical rainfall duration of each rainstorm is 120 minutes, and the rainstorm return period is selected as 1 year, 5 years, 10 years, 20 years, 30 years, 50 years and 100 years. After the calculation of relevant parameters of the rainstorm waterlogging model in one year is completed, the calculation results will be input into the dynamic static dual data driven model proposed in this paper to predict the number of distribution network interruption caused by rainstorm in this return period in this year. The process of Monte Carlo is shown in Fig.7.

After calculating the number of distribution network line interruptions caused by rainstorm in each return period of this year, the number of line interruptions caused by rainstorm in this year can be summed up. The average number of disconnected lines caused by each rainstorm can be calculated after 1000 years simulation, which reflects the bearing capability of distribution network to rainstorm.

The Fig.8 shows the real-time varying process of rainfall and ponding water depth during the 120-minute rainfall duration, with a time interval of 10 minutes. It can be seen from the figure that with the increase of return period, rainfall intensity and ponding water depth become larger, and the change of ponding water depth lags behind the change of rainfall process.

The following figure shows the number of rainstorm occurrences, the value of the average rainfall intensity of rainstorm

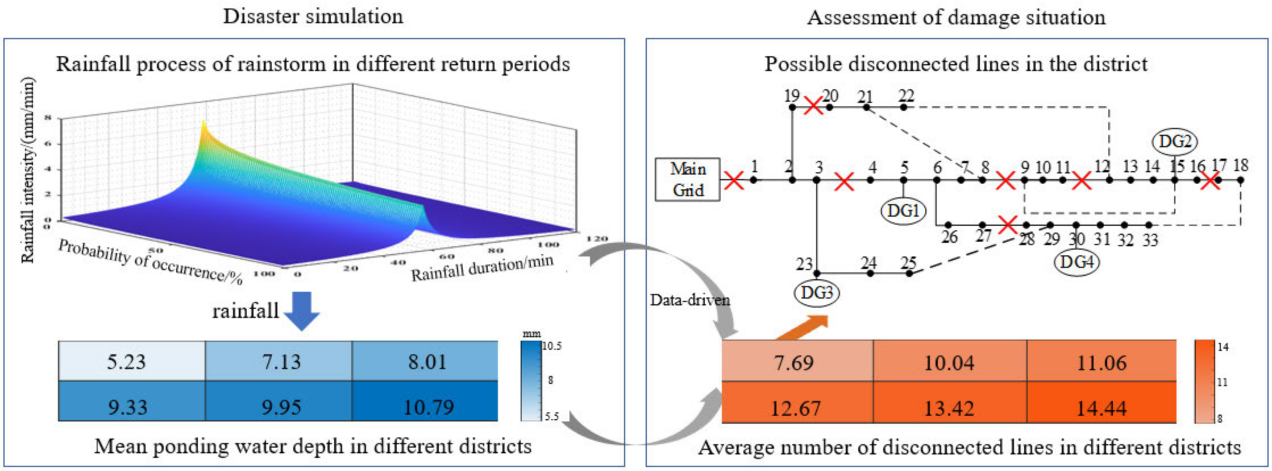


FIGURE 7. The process of monte carlo.

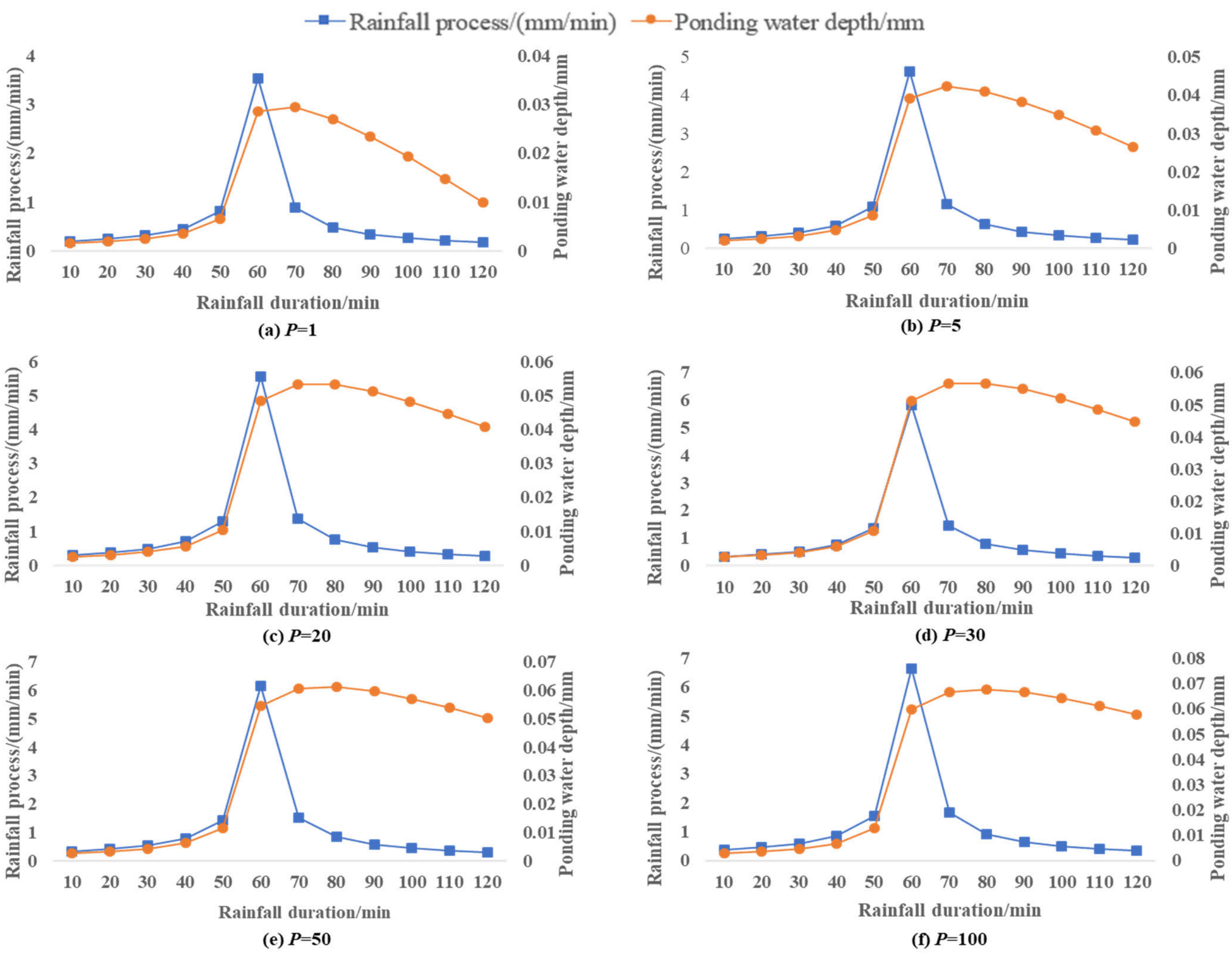


FIGURE 8. The time-varying process of simulation.

and the ponding water depth value in each return period during the simulated 1000-year time length.

It is easy to find from the figure that with the return period of rainstorm increase, the fewer times it occurs in the timescale of 1000 years, which conforms to the objective law. In addition, the static variable rainstorm intensity value as the data-driven input is also shown in the figure. The simulated average rainfall intensity of rainstorm in different

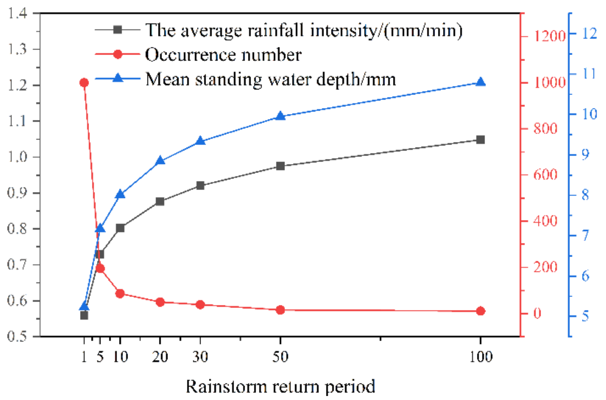


FIGURE 9. The results of simulation.

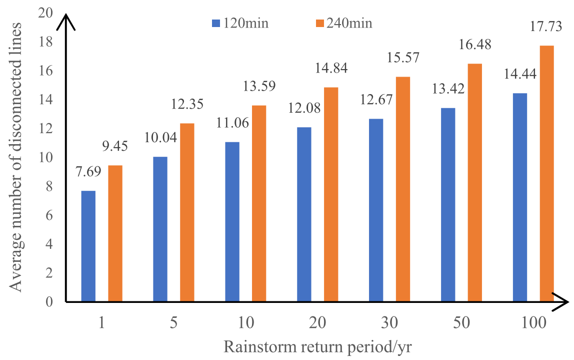


FIGURE 10. Average number of disconnected lines caused by heavy rain in different return periods.

return periods increases with the increase of return periods, which is consistent with our cognition. Meanwhile, it can be found from the figure that the simulated 120-min average ponding depth increases with the increase of the rainstorm return period. The average ponding depth caused by the 100-year return period rainstorm is about twice than that of the 1-year return period rainstorm.

In Fig.10, the average number of disconnected lines caused by simulated 1000 years, with 1-year, 5-year, 10-year, 20-year, 30-year, 50-year and 100-year return period rainstorm is given respectively.

It can be found in the figure that with the increase of the return period of rainstorm, the number of disconnected lines in the distribution network also increases. By comparing the number of line interruptions, it can be concluded that the number of disconnected lines caused by rainstorm with the 100-year return period is nearly twice as many as that with the 1-year return period. The reason is that in the rainstorm waterlogging model, the 100-year return period rainfall calculated by the rainstorm intensity calculation formula is nearly twice as strong as the 1-year return period rainfall. Therefore, it also reflects the effectiveness of the proposed method to a certain extent. At the same time, the comparison result of the rainstorm with a rainfall duration of 240 min is presented in Fig.10. It can be concluded that the average number of disconnected lines increases with the increase of rainfall duration.

The final result of the example shows that the average number of disconnected lines in the county-level unit area caused by the rainstorm with a rainfall duration of 120 min is 8.778, while that caused by the rainstorm lasting 240 min is 14.877. Compared with the historical data, this result is in line with the actual situation.

VI. CONCLUSION

In this paper, a knowledge-data-driven method for assessing the resilience of urban distribution network under rainstorm is proposed. The rainstorm waterlogging model is established by using rainstorm intensity, unimodal rainfall process and equal volume method, which depicts the dynamic time-varying process and overall static characteristics of rainstorm waterlogging, and provides input data for the data-driven model. By combining expert domain knowledge and data driven method, a dynamic-static dual data-driven model is proposed, which improves the prediction performance and interpretability of the distribution network line interruption model. Through Monte Carlo simulation, the resilience of urban distribution network under rainstorm disaster is assessed. The results of an example of a city in southern China shows that this method can effectively assess the ability of distribution network to withstand disasters.

In general, there exists following shortcomings in the work of this paper. Although the simulated scenes are typical, they are not comprehensive enough. The trained model may only fit to a single application area, that is, it is difficult to assess the resilience of urban distribution networks in other regions through the historical data of a typical area.

Therefore, according to the difference of rainstorm and precipitation data in different regions, how to comprehensively consider the impact of rainstorm with different rainfall duration on the distribution network will be the main work direction in the future. In addition, the improved method of Monte Carlo simulation algorithm should be studied to improve the calculation efficiency.

REFERENCES

[1] H. Zhang, Z. Xiao, C. Liang, and M. Xiao, "Parameters correction of rainstorm intensity formula in Zhengzhou," *J. Municipal Technol.*, vol. 40, no. 8, pp. 243–247, Aug. 2022.

[2] Z. Bie, Y. Lin, G. Li, and F. Li, "Battling the extreme: A study on the power system resilience," *Proc. IEEE*, vol. 105, no. 7, pp. 1253–1266, Jul. 2017.

[3] Y. Wang, T. Huang, X. Li, J. Tang, Z. Wu, Y. Mo, L. Xue, Y. Zhou, T. Niu, and S. Sun, "A resilience assessment framework for distribution systems under typhoon disasters," *IEEE Access*, vol. 9, pp. 155224–155233, 2021.

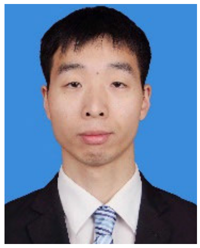
[4] P. C. Nayak, K. P. Sudheer, and K. S. Ramasastri, "Fuzzy computing based rainfall-runoff model for real time flood forecasting," *Hydrol. Processes*, vol. 19, no. 4, pp. 955–968, Mar. 2005.

[5] Q. Zou, J. Zhou, C. Zhou, L. Song, and J. Guo, "Comprehensive flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy AHP," *Stochastic Environ. Res. Risk Assessment*, vol. 27, no. 2, pp. 525–546, Feb. 2013.

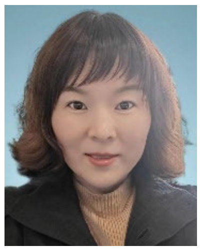
[6] M. Shao, Z. Gong, and X. Xu, "Risk assessment of rainstorm and flood disasters in China between 2004 and 2009 based on gray fixed weight cluster analysis," *Natural Hazards*, vol. 71, no. 2, pp. 1025–1052, Mar. 2014.

[7] Y. Hirabayashi, R. Mahendran, S. Koirala, L. Konoshima, D. Yamazaki, S. Watanabe, H. Kim, and S. Kanae, "Global flood risk under climate change," *Nature Climate Change*, vol. 3, no. 9, pp. 816–821, Sep. 2013.

[8] Y. Wu, P.-A. Zhong, Y. Zhang, B. Xu, B. Ma, and K. Yan, "Integrated flood risk assessment and zonation method: a case study in Huaihe River basin, China," *Natural Hazards*, vol. 78, no. 1, pp. 635–651, Aug. 2015.
[9] S. H. Erena and H. Worku, "Urban flood vulnerability assessments: The case of Dire Dawa city, Ethiopia," *Natural Hazards*, vol. 97, no. 2, pp. 495–516, Jun. 2019.
[10] W. Qi, C. Ma, H. Xu, Z. Chen, K. Zhao, and H. Han, "A review on applications of urban flood models in flood mitigation strategies," *Natural Hazards*, vol. 108, no. 1, pp. 31–62, Aug. 2021.
[11] D. Liu and J. Liu, "Intelligent planning of rain water drainage system in new urban areas considering the planning of road network," *J. Appl. Sci. Eng.*, vol. 22, no. 2, pp. 315–328, Aug. 2019.
[12] D. Wang, X. Fu, Q. Luan, J. Liu, H. Wang, and S. Zhang, "Effectiveness assessment of urban waterlogging mitigation for low impact development in semi-mountainous regions under different storm conditions," *Hydrol. Res.*, vol. 52, no. 1, pp. 284–304, Feb. 2021.
[13] Y. Tang, Y. Xue, M. Huang, Q. Wen, B. Cen, and D. Chen, "A lattice hydrodynamic model for four-way pedestrian traffic with turning capacity," *Sustainability*, vol. 15, no. 3, p. 2544, Jan. 2023.
[14] G. Ran, Q. Zhang, and Z. Chen, "Development of a three-dimensional hydrodynamic model based on the discontinuous Galerkin method," *Water*, vol. 15, no. 1, p. 156, Dec. 2022.
[15] Z. Zeng, C. Lai, Z. Wang, J. Chen, P. Chen, S. He, and S. Li, "Rapid simulation of urban rainstorm flood based on WCA2D and SWMM model," *Adv. Water Sci.*, vol. 31, no. 1, pp. 29–38, Mar. 2020.
[16] G. Zhao, Z. Xu, B. Pang, T. Tu, L. Xu, and L. Du, "An enhanced inundation method for urban flood hazard mapping at the large catchment scale," *J. Hydrol.*, vol. 571, pp. 873–882, Apr. 2019.
[17] Y. Zang, Y. Meng, X. Guan, H. Lv, and D. Yan, "Study on urban flood early warning system considering flood loss," *Int. J. Disaster Risk Reduction*, vol. 77, Jul. 2022, Art. no. 103042.
[18] Y. Du, Y. Liu, Y. Yan, and X. Jiang, "Disaster damage assessment of distribution systems with incomplete and incorrect information," *IEEE Trans. Power Del.*, vol. 38, no. 2, pp. 889–901, Apr. 2023.
[19] T. Tian, J. Zhang, F. Ye, C. Liu, T. Li, A. Tao, J. Wang, L. Yang, and Q. Wang, "Three-dimensional resilience index evaluation system for distribution network under flood disaster," *Adv. Technol. Elect. Eng. Energy*, vol. 41, no. 7, pp. 80–88, Nov. 2022.
[20] L. Zhang, R. Huang, and Z. Wang, "Optimal configuration strategy of mobile energy storage in distribution network considering balance between resilience and economy," *Autom. Electr. Power Syst.*, vol. 44, no. 21, pp. 23–31, Nov. 2020.
[21] H. Liu, R. A. Davidson, and T. V. Apanasovich, "Statistical forecasting of electric power restoration times in hurricanes and ice storms," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2270–2279, Nov. 2007.
[22] S. D. Guikema, R. Nateghi, S. M. Quiring, A. Staid, A. C. Reilly, and M. Gao, "Predicting hurricane power outages to support storm response planning," *IEEE Access*, vol. 2, pp. 1364–1373, 2014.
[23] S. Yuan, S. M. Quiring, L. Zhu, Y. Huang, and J. Wang, "Development of a typhoon power outage model in Guangdong, China," *Int. J. Electr. Power Energy Syst.*, vol. 117, May 2020, Art. no. 105711.
[24] K. Zhao, J. Hu, H. Shao, and J. Hu, "Federated multi-source domain adversarial adaptation framework for machinery fault diagnosis with data privacy," *Rel. Eng. Syst. Saf.*, vol. 236, Aug. 2023, Art. no. 109246.
[25] K. Zhao, F. Jia, and H. Shao, "A novel conditional weighting transfer wasserstein auto-encoder for rolling bearing fault diagnosis with multi-source domains," *Knowl.-Based Syst.*, vol. 262, Feb. 2023, Art. no. 110203.
[26] J. Yang, Y. Xiang, X. Xu, and J. Sun, "Design hyetograph for short-duration rainstorm in Jiangsu," *Atmosphere*, vol. 13, no. 6, p. 899, Jun. 2022.
[27] J. Liu, H. Zhou, C. Lu, and C. Gao, "A review of recent advances of urban rainfall intensity-duration-frequency relationships," *Adv. Water Sci.*, vol. 29, no. 6, pp. 898–910, Apr. 2018.
[28] P. Wang, S. Wang, J. Wang, and R. Hu, "Study on frequency distribution of short-duration rainstorm and rainstorm formula," *Water Resour. Power*, vol. 39, no. 1, pp. 22–25, Jan. 2021.
[29] *Notice of Guangzhou Water Bureau on Printing and Distributing Implementation Rules of Guangzhou Drainage Management Measures*. Accessed: Sep. 2018. [Online]. Available: https://www.gz.gov.cn/gfxwj/sbmgfxwj/gzsswj/content/post_5486145.html
[30] R. Balbastre-Soldevila, R. García-Bartual, and I. Andrés-Doménech, "A comparison of design storms for urban drainage system applications," *Water*, vol. 11, no. 4, p. 757, Apr. 2019.
[31] G. Cen, J. Shen, and R. Fan, "Research on rainfall pattern of urban design storm," *Adv. Water Sci.*, vol. 9, no. 1, pp. 41–46, Mar. 1998.
[32] Q. Huang, J. Dong, M. Li, and J. Wang, "Research on the scenario simulation method of rainstorm waterlogging hazard: A case study in the central urban area of Shanghai," *J. Geo-Inf. Sci.*, vol. 18, no. 4, pp. 506–513, Apr. 2016.
[33] L. Tang, H. Xie, Y. Wang, H. Zhu, and Z. Bie, "Predicting typhoon-induced transmission line outages with coordination of static and dynamic data," *Int. J. Electr. Power Energy Syst.*, vol. 142, Nov. 2022, Art. no. 108296.
[34] K. Cho et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Sep. 2014, pp. 1724–1734.
[35] S. Chaudhari, V. Mithal, G. Polatkan, and R. Ramanath, "An attentive survey of attention models," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 5, pp. 1–32, Dec. 2021.
[36] D. G. Triantali, K. E. Parsopoulos, and I. E. Lagaris, "Single-objective and multi-objective optimization for variance counterbalancing in stochastic learning," *Appl. Soft Comput.*, vol. 142, Jul. 2023, Art. no. 110331.



KE LI received the B.S. and M.S. degrees in electrical engineering from North China Electric Power University, in 2009 and 2011, respectively. His research interest includes power systems planning technology.



JIE MA received the B.S. degree in electrical engineering from Zhengzhou University, in 2007, and the M.S. degree in electrical engineering from North China Electric Power University, in 2010. Her research interests include power systems planning and the simulation of distribution systems.

JIANLONG GAO received the B.S. degree in electrical engineering from Shandong University, in 2017. He is currently pursuing the M.S. degree with the School of Electric Power Engineering, Xi'an Jiaotong University. He is an Intern with the Economic Research Institute, State Grid Henan Power Supply Company. His research interest includes the resilience of distribution systems.

CHANGQING XU, photograph and biography not available at the time of publication.

WENFENG LI, photograph and biography not available at the time of publication.

YUBIN MAO, photograph and biography not available at the time of publication.

SHIGONG JIANG, photograph and biography not available at the time of publication.