IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 17, 2019, accepted December 6, 2019, date of publication December 13, 2019,
date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958960

Cascading Outage Analyses by Integrating Distribution Factor Method With AC Power Flow

YUN LIU[1] AND YUSHENG XUE[1,2], (Member, IEEE)
[1]School of Electrical Engineering, Shandong University, Jinan 250061, China
[2]State Grid Electric Power Research Institute, Nanjing 211000, China

Corresponding author: Yusheng Xue (xueyusheng@sgepri.sgcc.com.cn)

ABSTRACT Cascading failure may incur catastrophic consequences. The high order and non-simultaneousness properties are the leading causes for the massive power flow (PF) calculations of the online static security assessment (SSA). Current PF models fail to satisfy the requirements for accuracy, speed, and robustness. This paper proposes a novel method where line outage distribution factor (LODF) model and AC model are coordinated by a binary classifier. The classifier determines the feasibility of LODF model in a specific case by evaluating its potential error. Cases are preferred to be analyzed by LODF model, while the ones with large potential errors will be analyzed by AC model. Case studies of three IEEE systems and the Texas 2000-bus system are given to verify the effectiveness, flexibility, and robustness of the proposed method.

INDEX TERMS Cascading outages, classifier, feature, power flow calculation.

I. INTRODUCTION

Large-scale blackouts have happened worldwide during the past decades, resulting in economic loss, social inconvenience, and cyber physical-attacks [1]–[5]. For the sake of ensuring security and reliability of the power system operation, the principle of $N-1$ security is widely employed in industry [6], where the whole system is required to withstand the failure of any single component. However, it is still inadequate to promise the system security. One example is that the Italian 2003 blackout occurred even though the $N-1$ criterion was fulfilled [7]. In order to further improve the reliability of the bulk transmission system, system operators are suggested to add the analyses of selected high-order contingencies in their operation manuals or energy management system (EMS) [8], [9]. Besides, non-simultaneous events are considered, and the North American Electric Reliability Council (NERC) has underlined the modified requirements for the $N-1-1$ outages in transmission system planning (TPL) [10]. In light of the high order and non-simultaneousness properties of cascading outages, more accurate outage probabilities and better system adjudgments can be acquired for independent system operators (ISOs) [11].

Although the operation reliability is significantly improved with these assessments, it is computationally infeasible to analyze all cascading contingencies due to the above properties. For a system containing N elements, when k of them are sequentially out of service, the total number of contingencies will reach $N!/(N-k)!$. When we take a medium-sized system containing 500 components as an example, there will be nearly 62 billion contingencies to be analyzed when $k=4$.

The post-contingency PF offers basic information for SSA. Accurate solutions can be obtained by the traditional full AC power flow model with relatively heavy computational burdens [12]. It is computationally burdensome to analyze massive cascading contingencies with AC model.

For burden reduction, the parallel computation technique is applied and indicates superior performance, especially in the large-scale system [13]. Unfortunately, it is constrained by industrial computational resources. Remarkable speed-up is witnessed in other trials like the application of holomorphic embedding method and deep convolutional neural network (CNN) [14], [15]. However, limited by the robustness, the industrial applicability of these methods needs to be further improved.

Alternatively, approximate models are developed to accelerate the PF calculations of enormous cascading chains. Initially, DC power flow model is extensively used in contingency screening and ranking owing to its simplicity. Afterward, matrix inverse lemma has been applied [16], further reducing the PF calculations under outage scenarios. Particularly, LODF and generalized line outage distribution

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

factor (GLODF) model the single line and simultaneous multiple line outages, respectively [17], [18]. Despite the high-speed performances, the accuracies of approximate methods are inherently rough due to the nonlinearities of PF equations. Moreover, the errors of approximate models usually accumulate with the growing number of outage lines.

Apparently, neither AC model alone nor any approximate model could speed up the PF calculations with satisfying accuracies. By combining AC model and one approximate model, a hybrid method is thereafter proposed. Current researches can be categorized into topology dependent and case dependent types. The former employs different PF models for different parts of a power grid. Precise solutions can be acquired in the targeted part once different parts are ideally decoupled [19]. Addressed to the limitations of the decoupling assumption, different PF models are applied to different cases in the case dependent type. The hot-start DC model and the ACLODF-based model are two typical approaches, where an approximate model is used for the base case or the first round of cascading outages, whereas AC model for the rest [20], [21]. This method reaches notable acceleration in PF calculations, yet with poor accuracy when high order contingencies are involved. Reference [22] considers the high order property, but the accuracy gain from AC model is not fully taken as the non-simultaneousness property is ignored. Anyhow, the above methods are of insufficient robustness resulting from the empirical designation of the PF model. In order to overcome the shortcoming, the feasibility evaluation of the approximate method turns out to be a critical issue. For the sake of determining the feasibility of approximate model, the error of approximate model is experimentally explored [23]–[25]. However, it is uncertain to pre-estimate the errors of specific cases without analytical analyses. This uncertainty complicates the judgment of the feasibility of approximate model in any given case.

Addressed to the empirical designation of PF models and the neglect of high order or non-simultaneousness property of cascading outages, this paper proposes a novel hybrid method where LODF model and AC model are adaptively selected by a binary classifier. The classifier systematically determines the feasibility of LODF model by pre-evaluating its potential error without AC model verifying. According to the classification results, cases with acceptable errors will be analyzed by LODF model for rapidity, whereas cases with large potential errors will be examined by AC model to enhance accuracy. The contributions of this paper are threefold.

First, a classifier-based method is proposed to accelerate PF calculations with high precision. Instead of conventional hybrid methods with empirical judgments, the feasibility of LODF model in specific cases is systematically determined by the binary classifier designed in this paper.

Second, features of the classifier are extracted upon the analytical error of LODF model, considering both high order and non-simultaneousness properties of cascading outages. The involvement of causal inference guarantees high generalization capability.

Third, accuracies and speeds of the method are controllable and adjustable by tuning the parameter of the proposed classifier, enhancing the industrial applicability to ISOs.

The remainder of the paper is organized as follows. Section II proposes the novel hybrid method coordinating LODF model and AC model. In Section III, the classifier is designed, and the feature extraction is underlined. Section IV presents the performance evaluation. Numerical simulations are given in section V. Section VI concludes the whole paper and offers suggestions for future work.

II. METHOD FOR CASCADING OUTAGES ANALYSES

A. FRAMEWORK OF THE PROPOSED METHOD

As for the online PF calculations of cascading outages, rarely could a single PF model reach an acceptable compromise between speed and accuracy. These two demands are usually contradictory. Approximate models like LODF model significantly speed up PF calculations. However, the rough precisions impede their industrial applications. On the other hand, AC model guarantees high accuracy yet with a heavy computational burden.

Regarding the paradox, it is promising to develop a novel hybrid method, where LODF model and AC model are coordinated in this paper. To accelerate PF calculations, LODF model is the ideal option to improve rapidity, considering its outstanding speed advantage in cascading failure analyses. For the sake of guaranteed accuracy, AC model is applied to cases with large potential errors. The two models are adaptively selected according to a binary classifier, distinct from the empirical designation in previous researches.

The proposed method consists of four modules, as depicted in Fig. 1. The initialization module offers outage information, including the candidate round i of cascading and the nearest AC round r. Based on the information of the initialization module, the classification module then evaluates the potential error of LODF model, and the feasibility of LODF model is systematically determined. Finally, the calculation module analyzes the case according to the result of the classification module. For cases with small potential errors, the fast calculation module with LODF model is activated. For cases with large potential errors, the accurate calculation module with AC model works.

The focus of the proposed method lies in the binary classification. Cases are classified into feasible ones and infeasible ones, according to the potential errors of LODF model. Cases with small potential errors are labeled with feasible tags and analyzed with LODF model for rapidity. For cases with large potential errors, infeasible tags are stuck, and AC model is utilized for accuracy. Additionally, owing to the accuracy gain from the hot-start DC method, the base case of each cascading chain is analyzed by AC model. We denote by r the most recent round of outage analyzed by AC model, and r should be initialized as 0. Besides, the influence of utilizing

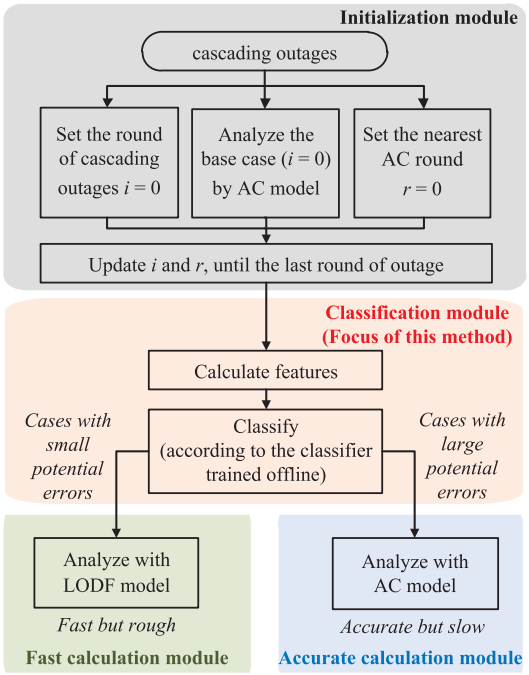


FIGURE 1. Proposed method for online PF calculations.

AC model in previous rounds of an outage is fully considered during the classifier design. With the increasing round of cascading outages, although error usually increases when LODF model is successively employed, it decreases to 0 once AC model is applied. Given such an "error elimination" ability of AC model, round r should be timely recorded and updated to calculate features. The entire designing process of the classifier will be systematically introduced in section III.

B. LODF MODEL

LODF model in the proposed method is the linear approximation of MW flow. As an incremental form of DC model in outage scenarios, it avoids the complex computations of Jacobian matrices and MVAR flows, indicating rapid performance in PF calculations for cascading outages.

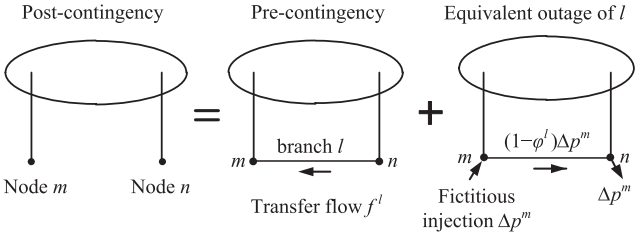


FIGURE 2. Equivalent topology for PF calculations with LODF model.

The principle of LODF model is explained in Fig. 2. Considering the non-simultaneousness property of cascading outages, we only discuss the single line outage in each round of cascading outages. Without loss of generality, we assume a power system with N buses and L transmission lines. Before the outage of branch l(between bus m and n), a transfer of f^l is observed on l.

Based on the linear assumption of PF equations, fictitious power is injected into the pre-contingency topology to model the outage. With an equal and opposite fictitious injection Δp^m at bus m and n, MW flow change on l equals to $(1-\varphi^l)\Delta p^m$ where $\varphi^l = \psi^{lm} - \psi^{ln}$. Here, we denote by $\boldsymbol{\Psi}$ the power transfer distribution factor (PTDF) matrix of $L \times N$, where ψ^{lm}, ψ^{ln} are the $(l, m)^{\text{th}}$ and $(l, n)^{\text{th}}$ element, respectively. The MW flow on l decreases to 0 when Δp^m satisfies

$$\Delta p^m = \frac{f^l}{1-\varphi^l} \tag{1}$$

We denote the active power of the outage line by $\dot{\boldsymbol{F}}$, a sparse $L \times 1$ vector with only the l^{th} element equaling f^l. The change in $\dot{\boldsymbol{F}}$ is $\Delta\dot{\boldsymbol{F}}$ and $\dot{\boldsymbol{F}} = \Delta\dot{\boldsymbol{F}}$ under any single line outage. We denote the fictitious equivalent active power injection by $\Delta\dot{\boldsymbol{P}}$, a sparse $N \times 1$ vector with the m^{th}, n^{th} element equaling Δp^m, $-\Delta p^m$, respectively. The node-branch incidence matrix $\boldsymbol{C}$ is a sparse matrix of $N \times L$ where the $(l, m)^{\text{th}}$ and the $(l, n)^{\text{th}}$ element of $\boldsymbol{C}$ are 1 and -1, respectively. Then (1) can be rewritten in matrix form

$$\Delta\dot{\boldsymbol{P}} = \boldsymbol{C} \cdot (\boldsymbol{I} - \boldsymbol{I} \circ (\boldsymbol{\psi} \cdot \boldsymbol{C}))^{-1} \Delta\dot{\boldsymbol{F}} \tag{2}$$

where $\boldsymbol{I}$ is an identity matrix of $L \times L$ and $\circ$ is the operator of Hadamard product as formulated in (A-1) of the Appendix A. We use $\boldsymbol{B}$ to denote the $N \times N$ nodal admittance matrix. As is described in DC model, the relationship between the changes of voltage angles $\Delta\boldsymbol{\theta}$ and $\Delta\dot{\boldsymbol{P}}$ is given by

$$\Delta\boldsymbol{\theta} = B^{-1} \cdot \Delta\dot{\boldsymbol{P}} \tag{3}$$

Substituting (2) into (3), we have

$$\Delta\boldsymbol{\theta} = \boldsymbol{B}^{-1} \cdot \boldsymbol{C} \cdot (\boldsymbol{I} - \boldsymbol{I} \circ (\boldsymbol{\psi} \cdot \boldsymbol{C}))^{-1} \Delta\dot{\boldsymbol{F}} \tag{4}$$

Assume an $L \times 1$ vector $\Delta\boldsymbol{F}$ describing the flow changes of all branches. Based on DC model, $\Delta\boldsymbol{F}$ is determined by

$$\Delta\boldsymbol{F} = \boldsymbol{B}_L \cdot \boldsymbol{C}^{\mathrm{T}} \cdot \Delta\boldsymbol{\theta} \tag{5}$$

where $\boldsymbol{B}_L$ represents the branch admittance matrix and ζ is the LODF matrix of $L \times L$. Substituting (4) into (5), $\Delta\boldsymbol{F}$ can be calculated by $\Delta\dot{\boldsymbol{F}}$.

$$\Delta\boldsymbol{F} = \boldsymbol{B}_L \cdot \boldsymbol{C}^{\mathrm{T}} \cdot \boldsymbol{B}^{-1} \cdot \boldsymbol{C} \cdot (\boldsymbol{I} - \boldsymbol{I} \circ (\boldsymbol{\psi} \cdot \boldsymbol{C}))^{-1} \Delta\dot{\boldsymbol{F}} = \boldsymbol{\zeta} \cdot \Delta\dot{\boldsymbol{F}} \tag{6}$$

Accordingly, ζ can be given by

$$\boldsymbol{\zeta} = \boldsymbol{B}_L \cdot \boldsymbol{C}^{\mathrm{T}} \cdot \boldsymbol{B}^{-1} \cdot \boldsymbol{C} \cdot (\boldsymbol{I} - \boldsymbol{I} \circ (\boldsymbol{\psi} \cdot \boldsymbol{C}))^{-1} \tag{7}$$

When LODF model is applied for the i^{th} round of outage, the post-contingency MW flow $\boldsymbol{F}_i$ is calculated as

$$\boldsymbol{F}_i^{\text{LODF}} = \boldsymbol{F}_{i-1} + \boldsymbol{B}_L \cdot \boldsymbol{C}^{\mathrm{T}} \cdot \Delta\boldsymbol{\theta} = \boldsymbol{F}_{i-1} + \boldsymbol{\zeta}_{i-1} \cdot \Delta\dot{\boldsymbol{F}}_i \tag{8}$$

However, the solutions of LODF model in cascading outages are inaccurate, and the errors usually accumulate with the increasing number of outage lines.

C. AC MODEL

AC model in the proposed method is a complete description of PF equations. The nonlinearity of the power system is retained to acquire accurate PF solutions.

For the power system with N buses and L branches mentioned above, the active power injection of AC power flow modeling is formulated as

$$p^m = v^m \sum_{n=1}^{N-1} v^n \left(g^{mn} \cos\theta^{mn} + b^{mn} \sin\theta^{mn}\right) \tag{9}$$

where p^m is the active power injected at m; v^m and v^n are the voltage at the node m and n, respectively; θ^{mn} is the voltage angle difference between the node m and n; g^{mn} and b^{mn} are the real and imaginary parts of the element in the bus admittance matrix, respectively. We denote by t^{mn} the transformer ratio of l and MW flow on l can be calculated as

$$f^l = v^m v^n \left(g^{mn} \cos\theta^{mn} + b^{mn} \sin\theta^{mn}\right) - t^{mn} g^{mn} \left(v^m\right)^2 \tag{10}$$

Accordingly, when the i^{th} round of outage occurs, (10) can be rewritten as

$$\boldsymbol{F}_i^{\text{AC}} = h\left(\boldsymbol{X}_i, \boldsymbol{Y}_i\right) \tag{11}$$

where $\boldsymbol{X}_i$ is the post-outage state variable including θ and v; $\boldsymbol{Y}_i$ is the network parameters including g, b, and t. Here, we ignore the ill-conditioned situation and take $\boldsymbol{X}_{i-1}$ as the starting point, and $\boldsymbol{X}_i$ can be solved by the Newton method in (9). Then, the post-contingency MW flow $\boldsymbol{F}_i$ can be acquired in (11).

Though accurate, the solutions of AC model rely on the repeated calculations of the Jacobi matrix, which is quite time-consuming.

III. CLASSIFIER DESIGN

A binary classifier is designed offline to evaluate the feasibility of the appropriate PF model for specific outage cases systematically. Generally, it consists of feature extraction (inputs), classification learning (input-output relations), and pattern generation (outputs).

A. INTEGRATION OF CAUSAL INFERENCE AND STATISTICAL PARADIGM

With the advancement in classification learning algorithms, targets can be efficiently related to features, and the classification accuracies are improved correspondingly [26]. However, supported by statistical analysis alone, the performances of these classifiers are easily affected by the over-dependence on quantity and quality of prior-knowledge database [27]. Causality underlies mechanisms and thus promotes the robustness of the statistical analysis. Therefore, causal analysis is indispensable during the classifier design. In this paper, causality inference and statistical paradigm are deeply incorporated throughout the whole offline training process, as depicted in Fig. 3.

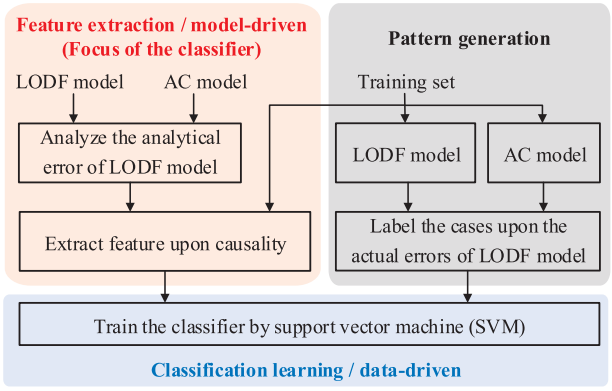


FIGURE 3. Offline training of the classifier.

With the causality involved in feature extraction, the features are designed upon the analytical error of LODF model and thus retain fundamental physical causality. In other words, the features quantify the potential error of LODF model in specific cases by logic reasoning, and the patterns are then determined. Therefore, compared with the raw empirical data, features extracted upon causality indicate higher correlations to patterns in most cases [28]. As the crucial point of classifier design, feature extraction will be comprehensively presented in part C of this section.

During the process of classification learning, the correlation between the features and the targets is efficiently analyzed. As is not the main contribution of this paper, classification learning will be briefly introduced in part D, and the support vector machine (SVM) will be applied to train the classifier.

With the robustness strength of causal analysis and the efficiency advantage of statistical analysis, the mutual benefits are highlighted by the integration of causal inference and statistical paradigm. To better clarify the purpose of the classifier design, we would introduce the pattern generation next.

B. PATTERN GENERATION

Due to the vague boundary between feasibility and infeasibility of the approximate model, the empirically-based methods are of limited robustness. The classifier is thereby designed to draw the boundary, and patterns are generated accordingly.

In this paper, the potential error of LODF model is the essential basis to quantify its feasibility. In other words, the binary patterns should be whether or not LODF model is feasible to be applied. It is feasible to apply LODF model only for cases with acceptable errors. Consequently, each case (one round of a cascading outage) is judged as a feasible one or infeasible one, as presented in Table 1. Cases with small potential errors would be tagged with feasible labels, and it is feasible to apply LODF model to these cases for rapidity. Those with large potential errors would be tagged with infeasible patterns, and AC model is utilized to avoid unacceptable rough solutions. It is worth mentioning that

the threshold of potential error is user-defined. When high-precision PF solutions are required, the threshold value is usually small so that more cases would be analyzed by AC model. On the other hand, when rapidity turns out to be the primary goal, the threshold value is usually large so that more cases would be analyzed by LODF model.

TABLE 1. Class labels of cases.

Label	Explanations
Feasible	Cases are with small potential errors, and LODF model is feasible to be applied.
Infeasible	Cases are with large potential errors, and LODF model is infeasible to be applied.

C. FEATURE EXTRACTION

Traditionally, key features are selected among raw data by statistical approaches for dimensionality reduction and performance improvement. However, arbitrary parameters may turn out to be the critical feature in a specific scenario. It is impractical to take all raw data as inputs for the classifier. Besides, the classifier is of limited robustness when little physical causality is involved in the selected feature.

In this paper, fundamental physical causality is retained in feature extraction to facilitate better data understanding and enhance the generalization capability. For the sake of learning the feasibility of LODF model, the error should be pre-estimated without being verified by AC model. Therefore, the analytical error of LODF model should be quantified under an arbitrary round of a cascading failure.

The influence of the change in voltage magnitude on active power is assumed to be ignored here. As the necessary and sufficient condition for the existence of the Taylor series, the post-contingency flow $\boldsymbol{F}_i$ is infinitely derivable in the field of real numbers. We expand $\boldsymbol{F}_i$ into Taylor series at the pre-fault equilibrium point $\boldsymbol{\theta}_{i-1}$

$$\boldsymbol{F}_i = h(\boldsymbol{\theta}_i) = \boldsymbol{F}_{i-1} + h'(\boldsymbol{\theta}_{i-1})\,\Delta\boldsymbol{\theta}_i + \frac{1}{2!}h''(\boldsymbol{\theta}_{i-1})\,\Delta\boldsymbol{\theta}_i^2 + \cdots \tag{12}$$

where $\Delta\boldsymbol{\theta}_i = \boldsymbol{\theta}_i - \boldsymbol{\theta}_{i-1}$. When LODF model has been repeatedly applied, the error consists of two parts by comparing (8) and (12). Accordingly, two types of features are extracted as follows.

1) FEATURE EXTRACTED FROM NONLINEAR ERROR

Features can be extracted from the nonlinear error $\boldsymbol{e}_{\mathrm{n}}$, which is the consequence of linear approximation of MW flow variation when the candidate outage occurs. It should be noted that the starting value $\boldsymbol{F}_{i-1}$ in (8) is assumed to be accurate to underline this type of error alone. Features extracted from $\boldsymbol{e}_{\mathrm{n}}$ include λ_1 and λ_2 deduced as follows.

In the linear approximation method, only the constant term $\boldsymbol{F}_{i-1}$ and the first-order term $h'(\boldsymbol{\theta}_{i-1})\Delta\boldsymbol{\theta}_i$ in (12) are retained. To highlight the influence of ignoring the nonlinear nature of PF equations, we assume the constant term $\boldsymbol{F}_{i-1}$ to be accurate. Therefore, $\boldsymbol{e}_{\mathrm{n}}$ is equal to the sum of infinite Taylor series truncating at second-order term.

$$\boldsymbol{e}_{\mathrm{n}} = \frac{1}{2!}h''(\boldsymbol{\theta}_{i-1})\,\Delta\boldsymbol{\theta}_i^2 + \frac{1}{3!}h'''(\boldsymbol{\theta}_{i-1})\,\Delta\boldsymbol{\theta}_i^3 + \cdots \tag{13}$$

We could extract one feature from the second-order term $h''(\boldsymbol{\theta}_{i-1})\Delta\boldsymbol{\theta}_i^2$, considering its dominant role in $\boldsymbol{e}_{\mathrm{n}}$. However, LODF model cannot provide voltage magnitude information for precise calculations of the second-order term. Disregarding the voltage magnitudes and assuming them unchanged during the contingency, the analytical calculations of the Hessian matrix are still cumbersome.

Addressed to the issues mentioned above, we apply the perturbation method, and the second-order derivative $h''(\boldsymbol{\theta}_{i-1})$ in (13) can be approximated as $h'(\boldsymbol{\theta}_i)/\Delta\boldsymbol{\theta}_i - h'(\boldsymbol{\theta}_{i-1})/\Delta\boldsymbol{\theta}_i$. The second-order term is thereby $(h'(\boldsymbol{\theta}_i) - h'(\boldsymbol{\theta}_{i-1}))\Delta\boldsymbol{\theta}_i$. Since $\boldsymbol{\theta}$ is not directly given in LODF model, we utilize the correlation between $\Delta\boldsymbol{\theta}_i$ and $\Delta\dot{\boldsymbol{F}}_i$ presented in (4), and $h''(\boldsymbol{\theta}_{i-1})\Delta\boldsymbol{\theta}_i^2$ can be estimated as $(h'(\dot{\boldsymbol{F}}_i) - h'(\dot{\boldsymbol{F}}_{i-1}))\Delta\dot{\boldsymbol{F}}_i$. The outage MW flow $\dot{\boldsymbol{F}}_i$ is defined in LODF model description (Part B, Section II) and $h'(\dot{\boldsymbol{F}}_i) = \partial\boldsymbol{F}_i/\partial\dot{\boldsymbol{F}}_i = \boldsymbol{\zeta}_i$, where $\boldsymbol{\zeta}_i$ is the LODF matrix of the i^{th} round of outage. Therefore, the second-order term is estimated as $(\boldsymbol{\zeta}_i - \boldsymbol{\zeta}_{i-1})\Delta\dot{\boldsymbol{F}}_i$. As the Hessian matrix is approximated and lots of its valuable information are lost when multiplied with the sparse vector $\Delta\dot{\boldsymbol{F}}_i$, the direct multiplication of $(\boldsymbol{\zeta}_i - \boldsymbol{\zeta}_{i-1})$ and $\Delta\dot{\boldsymbol{F}}_i$ makes little sense. For dimension reduction, the elementwise F-norm and the 1-norm are applied to $(\boldsymbol{\zeta}_i - \boldsymbol{\zeta}_{i-1})$ and $\Delta\dot{\boldsymbol{F}}_i$, respectively. Calculations of the two norms are formulated and listed in Appendix B. Feature λ_1 is defined as the product of these two norms.

$$\lambda_1 = \left\|\boldsymbol{\zeta}_i - \boldsymbol{\zeta}_{i-1}\right\|_{\mathrm{F}} \cdot \left\|\Delta\dot{\boldsymbol{F}}_i\right\|_1 \tag{14}$$

Besides (13), $\boldsymbol{e}_{\mathrm{n}}$ could also be described as follows.

$$\boldsymbol{e}_{\mathrm{n}} = h(\boldsymbol{\theta}_{i-1} + \Delta\boldsymbol{\theta}_i) - h(\boldsymbol{\theta}_{i-1}) - h'(\boldsymbol{\theta}_{i-1})\,\Delta\boldsymbol{\theta}_i \tag{15}$$

Applying with the linear approximation, we get

$$\begin{aligned}\boldsymbol{e}_{\mathrm{n}} &\approx \boldsymbol{e}_{\mathrm{n}}|_{\Delta\boldsymbol{\theta}_i=0} + \frac{\partial\boldsymbol{e}_{\mathrm{n}}}{\partial\boldsymbol{\theta}_i}\Delta\boldsymbol{\theta}_i \\ &= \left(h'(\boldsymbol{\theta}_{i-1} + \Delta\boldsymbol{\theta}_i) - h'(\boldsymbol{\theta}_{i-1})\right)\Delta\boldsymbol{\theta}_i\end{aligned} \tag{16}$$

Considering that $h'(\boldsymbol{\theta})$ is a trigonometric function of $\boldsymbol{\theta}$, term $|h'(\boldsymbol{\theta}_{i-1} + \Delta\boldsymbol{\theta}_i) - h'(\boldsymbol{\theta}_{i-1})|$ is a limited number smaller than twice the amplitude. When $\Delta\boldsymbol{\theta}$ is no longer trivial, the entire error $\boldsymbol{e}_{\mathrm{n}}$ turns out to be nonignorable. The consequence may be better understood if we compare the results between multiple line outages and single line outages. Statistically, both $\Delta\boldsymbol{\theta}$ and $\boldsymbol{e}_{\mathrm{n}}$ will increase with the growing number of outage lines. Therefore, the change in phase angle $\Delta\boldsymbol{\theta}_i$ is strongly correlated to $\boldsymbol{e}_{\mathrm{n}}$ and could be another key feature. As LODF model directly offers $\Delta\dot{\boldsymbol{F}}_i$ rather than $\Delta\boldsymbol{\theta}_i$, we may employ the correlation between $\Delta\boldsymbol{\theta}_i$ and $\Delta\dot{\boldsymbol{F}}_i$ again and use the 1-norm of $\Delta\dot{\boldsymbol{F}}_i$ to define feature λ_2 as follows.

$$\lambda_2 = \left\|\Delta\dot{\boldsymbol{F}}_i\right\|_1 \tag{17}$$

It should be noted that λ_2 can hardly quantify the whole $\boldsymbol{e}_{\mathrm{n}}$ alone. The value of $\boldsymbol{e}_{\mathrm{n}}$ is usually considerable when λ_2 is

large, but not vice versa. We may conclude that λ_2 presents better performance when it takes a larger value.

2) FEATURE EXTRACTED FROM ACCUMULATED ERROR

Another feature can be extracted from the accumulated error $\boldsymbol{e}_{\mathrm{a}}$, originating from the imprecise $\boldsymbol{F}_{i-1}$. It equals to the deviation between the AC solutions and the LODF solutions of the last round of outage. We still suppose r as the nearest AC solution outage round. The LODF solutions refer to the PF results where LODF model is continuously adapted from the $(r+1)^{\text{th}}$ round to the $(i-1)^{\text{th}}$ round. The feature extracted from $\boldsymbol{e}_{\mathrm{a}}$ refers to λ_3 designed as follows.

The inaccurate $\boldsymbol{F}_{i-1}$ in (12) leads to the nonzero $\boldsymbol{e}_{\mathrm{a}}$. Notably, this accumulated error is a zero vector if AC model analyzes the last round of outage. Due to the linear approximation property of LODF model, the accumulated error increases when this model is successively used. If AC model analyzes arbitrary round of outage (before the round i), the accumulated error is reduced or even eliminated. This situation is considered in the paper. Similarly, with round $r(r < i)$ given, we expand $\boldsymbol{F}_{i-1}$ into Taylor series at the nearest AC solution $\boldsymbol{\theta}_r$

$$\boldsymbol{F}_{i-1}=\boldsymbol{F}_r+h'(\boldsymbol{\theta}_r)(\boldsymbol{\theta}_{i-1}-\boldsymbol{\theta}_r)+\frac{1}{2!}h''(\boldsymbol{\theta}_r)(\boldsymbol{\theta}_{i-1}-\boldsymbol{\theta}_r)^2+\cdots \quad (18)$$

where $\boldsymbol{F}_r$ is the accurate PF solution calculated by AC model. The accumulated error is the truncation error of finite series presented as (19).

$$\boldsymbol{e}_{\mathrm{a}}=\frac{1}{2!}h''(\boldsymbol{\theta}_r)(\boldsymbol{\theta}_{i-1}-\boldsymbol{\theta}_r)^2+\frac{1}{3!}h'''(\boldsymbol{\theta}_r)(\boldsymbol{\theta}_{i-1}-\boldsymbol{\theta}_r)^3+\cdots \quad (19)$$

As the influence of the second-order term usually far outweighs that of the sum of other higher-order terms, we only consider the Hessian matrix $h''(\boldsymbol{\theta}_r)$. Because LODF model cannot directly acquire the voltage angle information, $h''(\boldsymbol{\theta}_r)$ is replaced by $h''(\dot{\boldsymbol{F}}_r)$ considering the strong correlation between $\boldsymbol{\theta}$ and $\dot{\boldsymbol{F}}$. Theoretically, $h''(\dot{\boldsymbol{F}}_r)$ represents the second derivative of PF when last $(i-r)$ outages coincide. Applying the perturbation method, we approximate it to the difference of the sensitivity between the $(i-1)^{\text{th}}$ round and the r^{th} round of outage. The F-norm of $h''(\dot{\boldsymbol{F}}_r)$ is taken as feature λ_3.

$$\lambda_3 = \left\|\boldsymbol{\zeta}_{i-1}-\boldsymbol{\zeta}_r\right\|_{\mathrm{F}} \quad (20)$$

In summary, the extracted features consist of λ_1, λ_2, and λ_3.

D. CLASSIFICATION LEARNING

Besides high-quality features and adequate training samples, the efficiency of the classifier also depends on smart classification learning. Machine learning algorithms can be applied to build the relationship between the outputs and inputs, such as logistic regressions, decision trees, and neural networks.

Since classification learning is not the main contribution of this paper, SVM is directly applied. Explicitly, we set C-SVM as the mathematical model and radial basis function (RBF) as its kernel function. For the sake of maximizing the classification accuracy, the gridding method is adopted to optimize the penalty parameters and the kernel function parameters. Without loss of generality, the k-fold cross-validation criterion is employed to overcome the overly optimistic estimation of training errors and improve the robustness of the classifier. Here, we take k as 5. For the classifier training offline, this paper applies the above settings to the library for support vector machines (LIBSVM) software package [29].

SVM maps the relationship between features and patterns. Therefore, accuracies and speeds of the classifier-based method can be controlled and adjusted by tuning the parameters of the classifier. Generally, when the proportion of ''feasible patterns'' increases, the accuracy deteriorates while the speed accelerates. This proportion is determined by the cost sensitivity of SVM, which is substantially related to the user-defined threshold of potential error in the user interface (UI). With the speed or accuracy requirement given by the ISOs, the desired results can be achieved.

IV. PERFORMANCE EVALUATION

Accuracy, speed, and robustness are three major concerns for evaluating the performance of the proposed method.

A. ACCURACY

Accurate results of the proposed method rely on precise classification results. Therefore, accuracy indicators should include classifier accuracy and PF accuracy of the method.

As for the accuracy of the classifier, a confusion matrix is applied. With a given classifier and an instance, four outcomes are formulated in a 2×2 matrix, as shown in Table 2. Common indicators include the true positive rate (TPR) and the false positive rate (FPR). TPR is the percentage of the TP in all actual feasible cases, and FPR is the percentage of the FP in all actual infeasible cases.

TABLE 2. Confusion matrix of classifier.

Predicted \ Actual	Feasible	Infeasible
Feasible	true positive (TP)	false positive (FP)
Infeasible	false negative (FN)	true negative (TN)

The analyses of feasible cases with AC model (the FN cases) result in an enormous waste of computations with little improvement of accuracy. Such unnecessary AC model analyses could be avoided by identifying the FN cases accurately. Otherwise, the misclassification will lead to low TPR and increase redundant computations. On the other hand, the misjudgments of actual infeasible cases (the FP cases) cause high FPR and incur potential SSA risks and inappropriate adjudgments when these infeasible results are applied.

Based on TPR and FPR, two accuracy measures are described as follows.

The receiver operating characteristic (ROC) curve describes TPRs and FPRs with various threshold settings and indicates the diagnostic ability of a binary classifier [30]. A point in the ROC space corresponds to a classifier, and one is better than another if it is northwest to the other (higher TPR and lower FPR). Notably, the point (0, 1) represents perfect classification. The diagonal line connecting the point (0, 0) and the point (1, 1) represents the strategy of random guesses. As both TPR and FPR are independent of category distribution, the ROC curve can be applied to evaluate the performance of unbalanced classification.

The area under the ROC curve (AUC) refers to the area of the curve in the normalized unit [30]. It equals to the probability that a classifier will rank a randomly choice of feasible case higher than that of infeasible one. The value is between 0 and 1. In particular, it is a random classifier when the AUC equals 0.5, whereas a perfect one when the AUC reaches 1.

Generally, in an accurate classifier, the ROC curve appears in the northwest corner of ROC space, and the value of AUC is close to 1.

As for the PF accuracy, several indicators should be applied. Single error ε is the relative error calculated as (21), where f^l and $f^{l,\mathrm{AC}}$ are the active PF solution on branch l of the proposed method and AC model, respectively, and $f^{l,\mathrm{lim}}$ is the MW limit of branch l. The value of L' equals to the sum of branches involved. Because the relative errors of low-power branches may submerge that of other branches, we rule out the branches of which the power is less than 10% of $f^{l,\mathrm{lim}}$ for better verification.

$$\varepsilon = \frac{1}{L'} \sum_{f^{l,\mathrm{AC}} > 10\% f^{l,\mathrm{im}}} \left| \frac{f^l - f^{l,\mathrm{AC}}}{f^{l,\mathrm{AC}}} \right| \times 100\% \qquad (21)$$

Based on ε, the following definitions are in Table 3. As can be expected, a high-precision method should be of small values of ε, E, σ_n, and $\varepsilon_{\max}$.

TABLE 3. Symbols of accuracy indicators.

Symbol	Quantity	Explanations
ε	Single error	Error of one case, formulated in (21)
E	Mean error	The mean value of ε of the sampled cases
σ_n	Proportion of cases with large errors	The percentage of cases with large potential errors. For example, σ_5 is the percentage of cases of which ε exceeds 5%. (in this paper, n can be 0.1, 0.8, 2, and 5.)
$\varepsilon_{\max}$	Maximum case error	The maximum value of ε of the sampled cases

B. SPEED

To achieve computational gain, we need to compare the total computation amount of the classifier-based method with that of the saved AC model calculations. If and only if the former one is less than the latter one, the application of the classifier is proved to be beneficial. Notably, the calculations of the classifier include the features and the trained SVM classifier.

However, the total calculation time of the classifier is not intuitive for the PF calculation of cascading outages. For better description purposes, the computational time of the classifier-based method and the proportion of cases with AC model analyses are provided instead. The total computational time of the method should be less than that of employing AC model alone. The speed indicators are listed in Table 4.

TABLE 4. Symbols of speed indicators.

Symbol	Quantity	Explanations
T	Computational time	The total analysis time of cases in seconds
N^{AC}	Proportion of cases with AC model analyses	The percentage of cases analyzed by AC model

A fast method is usually with a small value of T. Besides, the proportion N^{AC} is employed to replace T when samples with different numbers of cases are compared. For the same testing cases, the method is of smaller calculation amount when we take smaller N^{AC}.

C. ROBUSTNESS

For robustness concerns, case studies of the classifier and the proposed method should be evaluated in multiple simulation systems and operating conditions. Besides, the portion of training cases in the whole dataset should be limited, as an extremely large portion could hardly indicate the real performance of the classifier in industrial power systems [28]. In other words, the total number of testing cases should be much more than that of training cases.

V. CASE STUDIES

The effectiveness of the classifier and the flexibility of the proposed method are verified in the IEEE 57-bus system. For robustness verification, accuracies and speeds of the proposed method, AC model and LODF model are compared in the IEEE 30-bus, 57-bus, 118-bus, and Texas 2000-bus systems. The total number of testing cases are much more than that of training cases, for the sake of demonstrating the robustness of the proposed method. Both training and test cases are randomly sampled from the contingency set by the Monte Carlo method. All system data are provided in MATPOWER 7.0 [31].

A. EFFECTIVENESS OF THE CLASSIFIER

A case study is conducted at the standard operating condition in the IEEE 57-bus system. A total of 15,000 contingencies, ranging from one-line to six-line outages, are selected by the Monte Carlo method and forms the contingency set as described in Table 5. The PF solutions of the last round of the cascading failures are studied here when the proposed classifier is applied. To verify the effectiveness of the proposed classifier, we randomly apply AC model to an arbitrary round of cascading chains to prove that the classifier is of high generalization capability even if AC model is inappropriately applied.

TABLE 5. Descriptions of the contingency set.

Cascade	Total number of cascades	Number of samples	
		Training	Testing
$N-1$	78	18	60
$N-1-1$	2922	589	2333
$N-1-1-1$	3000	586	2414
$N-1-1-1-1$	3000	607	2393
$N-1-1-1-1-1$	3000	622	2378
$N-1-1-1-1-1-1$	3000	578	2422
Total	15000	3000	12000

TABLE 6. Classifiers with different values of V_ε.

V_ε (%)	Proportion of feasible cases (%)	Descriptions	Parameter value	
			c	g
0.5	18	Infeasible cases dominate.	3	256
1.2	50	Classification is balanced.	16	3
3.0	86	Feasible cases dominate.	3	256

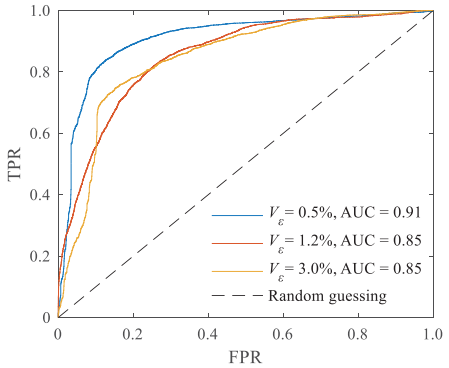


FIGURE 4. ROCs of the proposed classifier with different values of V_ε.

The proportion of feasible cases in the training set is determined by the user-defined threshold V_ε and increases with the growth of V_ε. The negative effect of an imbalanced dataset on classifier performance is considered in this part, and different values of V_ε are given as listed in Table 6. Besides, once V_ε is determined, the cost (parameter c) and gamma (parameter g) in the SVM classifier are settled by k-fold cross-validation. As indicated in Table 6, no evident regularities are found in parameter when V_ε varies. Considering that parameter tuning of SVM is not the main contribution of this paper, the classifier parameters in part B and C of this section will not be discussed in detail.

Fig. 4 illustrates the ROC curves and AUCs of the proposed classifier when V_ε varies. As the figure indicated, all of the ROC curves appear in the northwest corner of the plane, and all of their corresponding AUCs exceed 0.85 under the listed balanced or imbalanced scenarios. Admittedly, the AUC of the proposed classifier is not as outstanding as that of the pure data-driven classifier with an advanced classification learning algorithm. It should be noted that the total number of test samples is quadruple of that of training samples in our method, whereas the two usually take the opposite role in the pure data-driven classifiers. With the increasing proportion of training samples, AUC of the proposed classifier will increase to an excellent level. Besides, as our primary purpose is to acquire accurate PF solutions rather than the perfect AUC, the classification learning algorithm is not the contribution in this paper, and our goal can be reached by adjusting V_ε. Therefore, the proposed classifier remains high accuracy and ensures high generalization capability even in highly imbalanced situations.

B. FLEXIBILITY OF THE METHOD

For industrial applications, the proposed method should be flexible so that the performance can be determined according to the computational resources and the accuracy requirements of ISOs. We randomly sample 12000 $N-1-1-1-1$ cascading chains in the IEEE 57-bus systems. The standard, maximum, and minimum operating conditions are all considered in the samples. The minimum and maximum operating conditions are defined as 50% and 200% of their standard conditions in load and generation, respectively. For flexibility validation purposes, Table 7 indicates the accuracies and speeds of the proposed method with varying values of V_ε in both training and testing sets.

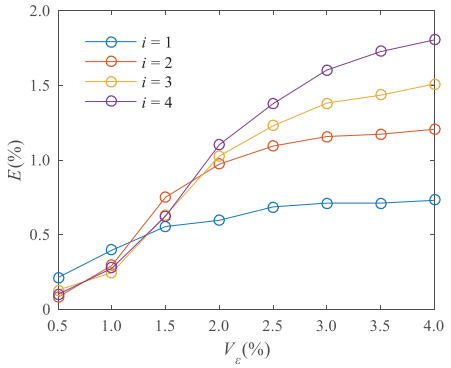
TABLE 7. Performance of the training and testing sets.

Proposed method		Training (2000 samples)			Testing (10000 samples)		
		E(%)	σ_2(%)	N^{AC}(%)	E(%)	σ_2(%)	N^{AC}(%)
Values of V_ε (%)	0.5	0.13	2	66	0.13	2	66
	1.0	0.30	4	46	0.31	4	47
	1.5	0.64	5	27	0.64	5	28
	2.0	0.92	9	15	0.92	10	16
	2.5	1.10	15	10	1.09	14	11
	3.0	1.21	18	7	1.21	18	8
	3.5	1.26	20	6	1.28	20	7
	4.0	1.31	21	5	1.34	21	6

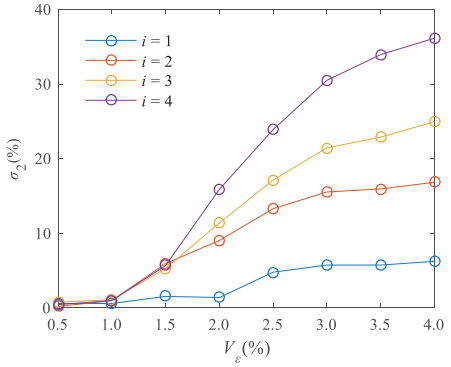
The meanings of E, σ_2, and N^{AC} are given in Table 3 and Table 4.

As for the speed description, the proportion N^{AC} is employed instead concerning the different numbers of training samples and testing samples. Though the comparison of T between training and testing sets makes no sense, the value of N^{AC} is comparable and helpful to predict the total computational time. For example, when we take $V_\varepsilon = 0.5\%$, $T = 34$s. Seeing that the number of the testing samples is five times that of the training samples, the predicted time of the testing samples should be 170s. As indicated in the simulation results, the actual time is the same as predicted. The values of E, σ_2, and N^{AC} between the training set and the testing set are very similar. Consequently, once the value of V_ε is determined, ISOs may predict the accuracy and speed of the proposed method.

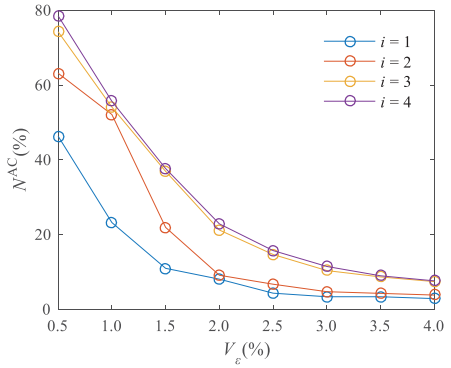
Fig. 5 further reveals the influence of V_ε on each round of cascading outages. As for the cases of the same i, when the value of V_ε increases, the values of E and σ_2 increase, whereas N^{AC} decreases. As for the cases of the same V_ε, the values of σ_2 and N^{AC} generally increase with the



(a) Influence on E.



(b) Influence on σ_2.



(c) Influence on N^{AC}.

FIGURE 5. Influence of V_ε on each round of cascading failures.

increasing values of i. Therefore, it is feasible to control the accuracy of each round i by adjusting V_ε of the corresponding round.

Three requirements on E, σ_2, and N^{AC} are listed in Table 8. Specifically, accuracies are required in Scenario A and B, whereas speeds are required in Scenario C. The values of V_ε in each round of outage is determined, according to Fig. 5. Consider the Scenario A and $i = 2$. The values of E are 0.29%, 0.75% and 1.02% when we take V_ε as 1.0%, 1.5%, and 2.0%, respectively. The value of V_ε is thereby set as 1.5% to guarantee $E \leq 1.0\%$. The relationship between V_ε and the performance of each round of outage can be learned offline so that requirements can be met by tuning the user-defined V_ε.

TABLE 8. Performance of test sets under three requirements.

Scenario A: require $E \leq 1.0\%$ in each round of outage					
Round i of outages		1	2	3	4
Setting	V_ε(%)	4.0	1.5	1.5	1.5
Simulation	E (%)	0.72	0.82	0.65	0.61
	σ_2(%)	5.5	8.5	6.2	5.9
	N^{AC}(%)	3	24	39	39
Scenario B: require $\sigma_2 \leq 1.5\%$ in each round of outage.					
Round i of outages		1	2	3	4
Setting	V_ε(%)	2.0	1.0	1.0	1.0
Simulation	E (%)	0.60	0.27	0.27	0.29
	σ_2(%)	1.5	1.0	1.0	1.5
	N^{AC}(%)	7	65	51	54
Scenario C: require $N^{\mathrm{AC}} \leq 25\%$ in each round of outage					
Round i of outages		1	2	3	4
Setting	V_ε(%)	1.0	1.5	2.0	2.0
Simulation	E (%)	0.38	0.68	0.89	1.04
	σ_2(%)	0.5	4.2	8.6	13.6
	N^{AC}(%)	24	20	20	23

The meanings of E, σ_2, and N^{AC} are given in Table 3 and Table 4.

C. ROBUSTNESS OF THE METHOD

The robustness of the proposed method is evaluated at the standard operating condition in the IEEE 30-bus, 57-bus, and 118-bus systems, respectively. In each IEEE system, all $N-1-1-1$ contingencies are sampled, while 10% of randomly selected as the training set.

TABLE 9. Accuracies and speeds of all $N-1-1-1$ chains.

System	Methods	E (%)	$\varepsilon_{\max}$ (%)	σ_2(%)	σ_5(%)	T (s)
30-bus	AC model	0	0	0	0	201
	LODF model	1.80	12.3	32.4	2.07	0.4
	Proposed method	0.58	5.66	1.28	0.05	9.7
57-bus	AC model	0	0	0	0	1023
	LODF model	2.10	38.9	29.2	6.55	3.2
	Proposed method	0.75	6.73	10.2	0.06	108
118-bus	AC model	0	0	0	0	22429
	LODF model	0.62	10.7	5.16	0.19	407
	Proposed method	0.24	4.03	0.02	0	1183

The meanings of E, $\varepsilon_{\max}$, σ_2, σ_5, and T are given in Table 3 and Table 4.

Table 9 compares the accuracies and speeds of different methods in diverse systems. One method usually indicates different performances in different systems, thereby the comparisons between different systems make no sense. Instead, the performances of methods are compared in each system only. Outstanding accuracy and speed enhancements are observed in Table 9. Compared with LODF model, accuracies are remarkably enhanced by the proposed method. Particularly, the values of σ_5 are nearly decreased by two orders of magnitude. Specifically, the values of σ_5 in 30-, 57-, 118-bus systems are dropped from 2.07%, 6.55% and 0.19% to 0.05%, 0.06% and 0%, respectively. Compared with AC model, at least one order of magnitude is reduced in computation amount with satisfying accuracies when the proposed method is applied.

The robustness of the classifier-based method is further validated in the 2000-bus Texas system, which contains 3026 branches. Since the total number of $N-1-1-1$ chains approximately reaches 33 billion, it is infeasible to list and

analyze every contingency as above. Here, 12000 $N-1-1-1-1$ cascading chains are randomly sampled by the Monte Carlo method, where 2000 chains are for training, and the remaining 10000 chains are for testing.

TABLE 10. Accuracies and speeds of random $N-1-1-1-1$ chains.

Methods	E (%)	$\varepsilon_{\max}$ (%)	$\sigma_{0.1}$(%)	$\sigma_{0.8}$(%)	T (s)
AC model	0	0	0	0	6108
LODF model	0.19	1.77	28.1	1.54	483
Proposed method	0.02	1.30	2.53	0.01	1267

The meanings of E, $\varepsilon_{\max}$, $\sigma_{0.1}$, $\sigma_{0.8}$, and T are given in Table 3 and Table 4.

The speeds and accuracies of different methods are compared in Table 10. Considering that the large branch number may lead to the relatively small values of ε, only $\sigma_{0.1}$ and $\sigma_{0.8}$ are discussed here to indicate the proportion of cases with large errors. The application of the classifier-based method acquires accurate PF solutions with satisfying speed. Outstanding improvement in precision is observed as the values of $\sigma_{0.8}$ is reduced by two orders of magnitude, and the value of E is decreased from 0.19% to 0.02%. Meanwhile, the speed of the proposed method is over five times that of AC model. To summarize, comparing with the traditional AC model, the proposed method not only notably accelerate cascading outage analyses with satisfying accuracy, but also is proved to be a method with good robustness.

VI. CONCLUSION AND FUTURE WORK

Blackouts urge the accurate and fast PF calculations for massive cascading outages. This paper proposes the classifier-based method coordinating LODF model and AC model. Rather than an empirical designation, the classifier pre-evaluates the potential error of LODF model by considering the high order and non-simultaneousness properties. The feasibility of LODF model is systematically analyzed by the proposed classifier. To guarantee the accuracy, rapid, and robust classification, causal inference and statistical paradigm are deeply integrated during the designing process of the classifier. As is indicated in multiple simulation systems, the proportion of cases with large errors could be decreased by two orders of magnitude, while the speed is improved by ten times. As accuracies and speeds of the method are predictable, the PF solutions with expected accuracy or speed can be obtained for industrial practice by tuning the parameters of the classifier offline. Afterward, ISOs are can analyze the static security or decide system adjudgments.

Admittedly, the proposed method cannot offer voltage magnitude and MVAR information when cases are applied with LODF model. It is worth trying other PF models for approximate calculations of voltage and reactive power. Apart from enhancing PF methods, other future research interests in cascading outages include fourfold. First, since protective relays are involved in 75% of major disturbances, modeling the protection devices in a more accurate manner is promising. Second, as many power system networks are already stability limited, analyzing cascading outages from a transient or mid-term view is very important. Third, identifying remedial actions or special protection schemes is necessary to determine the investments in new facilities. Fourth, improving the speed of computations is beneficial to enhance the analysis efficiency [32].

APPENDIX

A. THE DEFINITION OF HADAMARD PRODUCT

In matrix theory, the Hadamard product is an elementwise product of the two matrices of the same dimensions. If both $\boldsymbol{A} = \{a_{ij}\}$ and $\boldsymbol{B} = \{b_{ij}\}$ are $m \times n$ matrices, then their Hadamard product is given by

$$\boldsymbol{A} \circ \boldsymbol{B} = \{a_{ij} b_{ij}\} \tag{A-1}$$

B. CALCULATIONS OF THE MENTIONED NORMS

A vector norm is a function that assigns a strictly positive length or size to each vector in a vector space. A matrix norm is a vector norm in a vector space whose elements (vectors) are matrices (of given dimensions).

If $\boldsymbol{X} = \{x_i\}$ is a $n \times 1$ vector, the 1-norm of $\boldsymbol{X}$ is given by

$$|\boldsymbol{X}|_1 = \sum_{i=1}^{n} |x_i| \tag{B-1}$$

If $\boldsymbol{A} = \{a_{ij}\}$ is an $m \times n$ matrix, the Frobenius norm (F-norm) of $\boldsymbol{A}$ is given by

$$|\boldsymbol{A}| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \right)^{1/2} \tag{B-2}$$

REFERENCES

[1] O. P. Veloza and F. Santamaria, "Analysis of major blackouts from 2003 to 2015: Classification of incidents and review of main causes," *Electr. J.*, vol. 29, no. 7, pp. 42–49, 2016.
[2] Z. Ma, F. Liu, C. Shen, Z. Wang, and S. Mei, "Fast searching strategy for critical cascading paths toward blackouts," *IEEE Access*, vol. 6, pp. 36874–36886, 2018.
[3] F. Liu, J. Guo, X. Zhang, Y. Hou, and S. Mei, "Mitigating the risk of cascading blackouts: A data inference based maintenance method," *IEEE Access*, vol. 6, pp. 39197–39207, 2018.
[4] Q. Zhang, W. Fan, Z. Qiu, Z. Liu, and J. Zhang, "A new identification approach of power system vulnerable lines based on weighed H-index," *IEEE Access*, vol. 7, pp. 121421–121431, 2019.
[5] M. Tian, M. Cui, and Z. Dong, "Multilevel programming-based coordinated cyber physical attacks and countermeasures in smart grid," *IEEE Access*, vol. 7, pp. 9836–9847, 2019.
[6] EBTSO-E. (Mar. 2009). *UCTE OH—Policy 3: Operational Security*. [Online]. Available: https://www.entsoe.eu/fileadmin/user_upload/_library/publications/entsoe/Operation_Handbook/Policy_3_final.pdf
[7] A. Berizzi, "The Italian 2003 blackout," in *Proc. IEEE PES General Meeting*, Denver, CO, USA, Jun. 2004, pp. 1673–1679.
[8] NERC. (Spe. 2018). *Reliability Guidelines*. https://www.nerc.com/comm/PC_Reliability_Guidelines_DL/Reliability_Guideline_Methods_for_Establishing_IROLs.pdf
[9] WECC. (Nov. 2006). *Remedial Action Scheme Design Guide*. [Online]. Available: https://www.wecc.org/Reliability/RWG%20RAS%20Design%20Guide%20_%20Final.pdf
[10] NERC. (Spe. 2015). *Standard Application Guide TPL-001-4*. [Online]. Available: https://www.nerc.com/pa/comp/guidance/EROEndorsedImplementationGuidance/TPL-001-4_Standard_Application_Guide_endorsed.pdf
[11] L. Che, X. Liu, and Z. Li, "Screening hidden N-K line contingencies in smart grids using a multi-stage model," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1280–1289, Oct. 2017.
[12] A. J. Wood, B. F. Wollenberg, *Power Generation, Operation, and Control*. Hoboken, NJ, USA: Wiley, 1996, pp. 421–424.

[13] I. Araújo, V. Tadaiesky, and D. Cardoso, "Simultaneous parallel power flow calculations using hybrid CPU-GPU approach," *Int. J. Electr. Power Energy Syst.*, vol. 105, pp. 229–236, Aug. 2019.
[14] C. Liu, B. Wang, X. Xu, K. Sun, D. Shi, and B. C. L. Bak, "A multi-dimensional holomorphic embedding method to solve ac power flows," *IEEE Access*, vol. 5, pp. 25270–25285, 2017.
[15] Y. Du, F. F. Li, and J. Li, "Achieving 100x acceleration for N-1 contingency screening with uncertain scenarios using deep convolutional neural network," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3303–3305, May 2019.
[16] M. K. Enns, J. J. Quada, and B. Sackett, "Fast linear contingency analysis," *IEEE Trans. Power App. Syst.*, vol. 4, pp. 783–791, Apr. 1982.
[17] J. Guo, Y. Fu, Z. Li, and M. Shahidehpour, "Direct calculation of line outage distribution factors," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1633–1634, Aug. 2009.
[18] T. Guler, G. Gross, and M. Liu, "Generalized line outage distribution factors," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 879–881, May 2007.
[19] S. Kim and T. J. Overbye, "Mixed power flow analysis using AC and DC models," *IET Gener., Transmiss. Distrib.*, vol. 6, no. 10, pp. 1053–1059, May 2012.
[20] B. Stott, J. Jardim, and O. Alsac, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.
[21] P. Mitra, V. Vittal, and B. B. Keel, "A systematic approach to n-1-1 analysis for power system security assessment," *IEEE Power Energy Tech. Syst. J.*, vol. 3, no. 2, pp. 71–80, May 2016.
[22] G. Chen, Y. Dai, and Z. Xu, "A flexible framework of line power flow estimation for high-order contingency analysis," *Int. J. Electr. Power Energy Syst.*, vol. 70, pp. 1–8, Feb. 2015.
[23] M. Liu and G. Gross, "Effectiveness of the distribution factor approximations used in congestion modeling," in *Proc. 14th PSCC*, Seville, Spain, Jun. 2002, pp. 24–28.
[24] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *Proc. IEEE PES Gen. Meeting*, San Francisco, CA, USA, Jun. 2005, pp. 454–459.
[25] D. Van Hertem, J. Verboomen, and K. Purchala, "Usefulness of DC power flow for active power flow analysis with flow controlling device," in *Proc. 8th Int. Conf. AC DC Power Transmiss.*, Mar. 2006, pp. 58–62.
[26] C. C. Aggarwal, *Data Classification: Algorithms and Applications*. Boca Raton, FL, USA: CRC Press, 2014, pp. 4–14.
[27] Q. Wang, F. Li, Y. Tang, and Y. Xu, "Integrating model-driven and data-driven methods for power system frequency stability assessment and control," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4557–4568, Nov. 2019.
[28] Y. Xue and Y. Lai, "Integration of macro energy thinking and big data thinking part two applications and explorations," *Auto. Electr. Power Syst.*, vol. 40, no. 8, pp. 1–13, Mar. 2016.
[29] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:2, Apr. 2011.
[30] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
[31] R. D. Zimmerman, C. E. Murillo-Sanchez. (Jun. 2019). *Matpower User's Manual Version 7.0*. [Online]. Available: https://matpower.org/docs/MATPOWER-manual-7.0.pdf
[32] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, May 2012.

YUN LIU is currently pursuing the Ph.D. degree with the School of electrical engineering, Shandong University. Her research interest is power system static security assessment.

YUSHENG XUE received the B.E. degree from Shandong University, Jinan, China, in 1963, the M.S. degree from the State Grid Electric Power Research Institute, Nanjing, China, in 1981, and the Ph.D. degree in electrical engineering from the University of Liège, Liège, Belgium, in 1987. He is currently the Honorary President of the State Grid Electric Power Research Institute (NARI Group Corporation), and also a Professor with the Electrical Engineering Department, Shandong University. His research interest is power system automation. He is member of the Chinese Academy of Engineering.