Model and Data Driven Machine Learning Approach for Analyzing the Vulnerability to Cascading Outages With Random Initial States in Power Systems

Hongji Zhang, *Student Member, IEEE*, Tao Ding, *Senior Member, IEEE*, Junjian Qi, *Senior Member, IEEE*, Wei Wei, *Senior Member, IEEE*, João P. S. Catalão, *Fellow, IEEE*, and Mohammad Shahidehpour, *Life Fellow, IEEE*

***Abstract*— In this paper, a hybrid machine learning model is applied to evaluate the relationship between random initial states and the power system's vulnerability to cascading outages. A cascading outage simulator (CS), which uses off-line AC power flows, is proposed for generating training data. The initial states are randomly selected and the CS model is deployed for each initial state, where power system generation and loads are adjusted dynamically and power flows are redistributed to quantify the vulnerability metric. Furthermore, the proposed hybrid machine learning model deploys a combined Support Vector Machine (SVM) classification and Gradient Boosting Regression (GBR) to improve the learning precision. The classification model is trained by SVM, which divides the data into two categories with and without load shedding. Then, GBR is adopted only for the data with load shedding to determine the relationship between input power outage states and the vulnerability metric. The proposed vulnerability analysis approach is applied to several test systems and the results are analyzed.**

***Note to Practitioners*—The power system vulnerability can be quantified by cascading outage simulations. However, there are two challenges: i) there are a huge number of possible initial states and we cannot enumerate all these initial states for the cascading outage simulation. Neither can we precisely quantify the bus vulnerability. ii) The cascading outage simulation may be time-consuming for large-scale power systems, which is challenging for the online application. To address the above challenges, we expect to design a machine learning technique to predict the power system vulnerability, which can train the model in an offline way and then use it for the online application. Firstly, since there is not enough operation data from practical power systems, we develop a cascading outage simulator, using off-line AC power flows, for generating synthetic training data. Secondly, we observe that the training precision by directly applying the regression model may be very poor because the output of the machine learning model may take on an uneven distribution concerning input parameters. Thus, we propose a hybrid machine learning model with a combined classification and regression method, where the classification model is employed to remove the data without the load shedding, and the regression model then determines the relationship between input power outage states and the vulnerability metric. The proposed model and method have been tested on several systems including a practical large-scale Polish power system to show the effectiveness.**

***Index Terms*— Machine learning, cascading outages, vulnerability analysis, gradient boosting regression.**

Manuscript received 30 January 2022; revised 24 April 2022 and 29 July 2022; accepted 23 August 2022. This article was recommended for publication by Associate Editor M. Franceschelli and Editor Q. Zhao upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 51977166, in part by the Natural Science Foundation of Shaanxi Province under Grant 2022JC-19, and in part by the Science and Technological Project of Northwest Branch of State Grid Corporation of China under Grant SGNW0000DKQT2100172. *(Corresponding author: Tao Ding.)*

Hongji Zhang and Tao Ding are with the School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (e-mail: tding15@mail.xjtu.edu.cn).

Junjian Qi is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA.

Wei Wei is with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China.

João P. S. Catalão is with the Faculty of Engineering, University of Porto (FEUP), 4099-002 Porto, Portugal, and also with INESC TEC, 4200-465 Porto, Portugal.

Mohammad Shahidehpour is with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASE.2022.3204273.

Digital Object Identifier 10.1109/TASE.2022.3204273

NOMENCLATURE

A. Indices and Sets

t	Index for time intervals
i, j	Index for buses
g	Index for generators
l	Index for power loads
d	Index for islands
n,p	Index for training data
m	Index for test data
k	Index for iterations in GBR
$\mathbb{B}$	Set of the buses
$\mathbb{L}$	Set of the transmission
$\mathbb{G}$	Set of the generators
$\mathbb{D}$	Set of the islands

B. Parameter and Constants

r	The ramp rate of generators
Δt	The time interval for generation output ramping
U_0^{ref}	The voltage magnitude at the reference bus.
θ_0^{ref}	The voltage angle at the reference bus.
w_i	The weight of the load shedding for load bus i

$n_{l,i}$	The allowable power factor at the load bus i
$U_i^{min}(t)$	The lower bound of voltage magnitude at bus i
$U_i^{max}(t)$	The upper bound of voltage magnitude at bus i
$F_{ij}^{max}(t)$	The transferred power limit on the line ij.
$P_{L,i}^0$	The initial active power load at bus i.
N_b	Number of buses
N_{train}	Number of training data
N_s	Number of training data with load shedding
N_{test}	Number of test data
ω, b	The parameters of the hyperplane in SVM
γ	The nonnegative parameter in the kernel function in SVM
C	The parameter that controls the tightness of margin in SVM
M	The number of regression trees in GBR
μ	The learning rate for GBR

C. Functions and Variables

$P_{g,i}(t)$	The active power generation at bus i and time t
$P_{l,i}(t)$	The active power load at bus i and time t
$Q_{g,i}(t)$	The reactive power generation at bus i and time t
$Q_{l,i}(t)$	The reactive power load at bus i and time t
$U_i(t)$	The voltage magnitude of bus i at time t
$\theta_i(t)$	The voltage angle of bus i at time t
$P_{g,0}(t)$	The generation output at the reference bus and time t
$P_{g,0}^*(t)$	The active power flow solution at the reference bus
T_d^{sp}	The duration of generator ramping
T_{min}^{sp}	The minimum time of T_d^{sp}
$P_{l,i}^s(t)$	The active load power at bus i after the load shedding
$Q_{l,i}^s(t)$	The reactive load power at bus i after the load shedding
$F_{ij}(t)$	The power flow on the transmission line ij
$P_{l,i}^\infty$	The final active power load at bus i.
x_n	The n-th set of input data for generator output and load level
z_n	The n-th classification label referring to whether the load shedding happens
y_n	The n-th regression label for LOSS
K	The kernel function for SVM
$\varphi(x)$	The feature space for SVM
$f_k(x_n)$	The model of k-th regression tree in GBR
L	The loss function in GBR
$r_{k,n}$	The negative gradient for the k-th tree in GBR
$h_k(x_n, a_k)$	The parameterized function for the k-th tree in GBR
β_k	The steepest descent direction of k-th regression tree in GBR

I. Introduction

RESILIENCE and vulnerability analysis for power systems is critical for system operation and dispatch. There are many methods for vulnerability assessment of the power system, such as the Monte Carlo simulation method, transient analysis method, and cascading outage simulation (CS) methods [1], [2]. Particularly, the CS method is considered an effective tool for tracking fault propagation and analyzing the power system vulnerability to such incidents. Recent years have witnessed several blackouts in electric power systems worldwide, with substantial socioeconomic impacts. Many of these blackouts were caused by local faults which then propagated by triggering cascading outages in multiple geographical regions.

The mainstream of CS includes applications of the complex network theory to power flow analysis. The complex network theory treats a power system as a model with a large number of components, considers interactions among corresponding components, and analyzes its critical characteristics in regard to cascading outages. In addition, the power flow analysis method divides the continuous fault process into several stages and calculates the power flow recursively to find the overloaded lines. However, most of these power flow-based CSs only focus on the power flow redistribution, while lacking comprehensive consideration for generation adjustment and load shedding strategies. It results in a large error for the final vulnerability evaluation results compared with practical cases.

Moreover, when some extreme events come, the operating states of the power system (i.e., generator output and load levels) may change before the cascading outage occurs. In different operating states, the locations of the vulnerable parts for power systems are different [3]. It is very time-consuming and impractical to conduct CS for all possible initial boundary conditions in online applications. Therefore, we need to quickly evaluate the vulnerability of system components in a stochastic initial state before cascading outages. In this way, the operators can give priority to protecting and strengthening these vulnerable critical components according to the prediction results. To solve this problem, machine learning methods can be employed to analyze the impacts of uncertainties on CS.

At present, machine learning methods have been successfully applied to power system studies, which mainly include deep learning, reinforcement learning, neural networks, etc. These methods were used to predict load curves [4], [5], distributed generation forecasting [6], [7], fault diagnosis [8], [9] energy management [10], [11]. Many of these practical applications of machine learning methods can be modeled as classification and regression problems. For these machine learning algorithms, we can train the model in an offline way and then use it for the online application.

However, few studies characterized the relationship between random initial states and the vulnerability under CS. Besides, there are two challenges: i) there are a huge number of possible initial states and we cannot enumerate all these initial states for the CS simulation. Neither can we precisely quantify the bus vulnerability. Thus, we expect to use the machine learning technique to predict the vulnerability of the power system according to the input characteristics (i.e., uncertain generator output and load level); ii) The CS simulation may be time-consuming for large-scale power systems, which is challenging for the online application. In contrast, the machine learning method can be employed to solve these challenges.

Among these state-of-the-art machine learning methods, Support Vector Machine (SVM) was developed in many studies to provide a reliable solution with a large number of features for binary classification problems. Compared with other advanced classification algorithms, SVM algorithm usually has a shorter training time and higher efficiency due to its simple structure under the same prediction accuracy requirements [12], [13]. Besides, with the development of some advanced ensemble learning algorithms, Gradient Boosting Regression (GBR) algorithm was widely used in various data analyses of power systems and has been proved to be more effective than other traditional machine learning algorithms [14].

Moreover, we have observed that the training precision by directly applying the regression model is very poor due to the following two reasons: 1) The number of features (i.e., bus generation and load states) is generally high. Therefore, the accuracy may suffer due to the over-fitting problem when considering a limited number of available training data. 2) Many initial states will not lead to load shedding. Therefore, the output of the machine learning model may take on an uneven distribution concerning input parameters. To address the above challenges, we propose a hybrid machine learning model with a combined classification and regression method. SVM algorithm is effectively employed for the classification with and without load shedding categories under multidimensional features. Then GBR regression prediction is carried out for the data with load shedding. The main contributions of this paper are summarized as follows:

(i) A CS is proposed based on off-line AC power flows for the vulnerability analysis of each bus, which divides the outage propagation process into multiple cascading outage events (CEs). Meanwhile, the initial states are randomly selected and the CS model is deployed for each initial state to generate CS samples. Using CS, power generation and loads are adjusted and power flows are redistributed dynamically to mitigate any power imbalances and quantify the vulnerability metric.

(ii) A hybrid machine learning model with the combined classification and regression methods is proposed for bus vulnerability analyses considering random initial states. Here, the classification method is chosen by SVM, and the regression method is chosen by GBR. A sequential strategy is proposed in which the training data are classified into with and without load shedding categories. Then, a regression model is applied for training with load shedding data to predict the vulnerability metric in random initial states.

The rest of the paper can be summarized as follows. In Section II, we review the related work. Section III provides the details of the vulnerability analysis of cascading outages where the simulation strategy for cascading outages is proposed. Section IV designs a combined classification and regression machine learning model for analyzing the impact of stochastic inputs on vulnerability. Simulation results based on CS and machine learning strategy are presented in Section V. Section V has drawn the conclusions of this paper.

II. Related Work

Based on the complex network theory, [15] proposed a cascading outage model that concerned electrical load characteristics. The result showed that the proposed model could analyze network vulnerabilities more effectively. In [16], the Galton-Watson branching process method was introduced in CS to estimate the cascading outage process and the corresponding blackout size. A Lagrange-Good inversion-based multi-type branching process method was proposed in [17] to quantify the blackout propagation and analyze the interdependencies among different infrastructure systems.

In addition, power flow-based methods are also widely used for CS analysis. Ref. [18] formulates an optimal power flow problem for the cases of single and multiple generator failures is addressed as an example, which could accurately capture the outage propagation. DC power flow was used in [19] for CS due to its efficiency and simplicity. Besides, [20] quantified the interdependence of power and communication networks by an interactive cascading model, indicating that a greater interdependence would lead to a lower probability of power outages. In [21], a Markovian-tree-based multi-timescale model was set up for simulating cascading outages, and a forward-backward searching strategy was employed to speed up the simulation. In [22], a multi-time scale dynamic simulation model was established with a sensitivity-based dispatch strategy. Ref. [23] developed a novel distribution system restoration model in response to multiple outages caused by extreme weather.

Other scholars studied CS via optimization and statistical methods. A nonlinear convex optimization model solved by saddle point dynamics was established in [24] for the prediction of the cascading outage path, which could change cascading outages by adjusting the injected power. Ref. [25] proposed a dynamic programming model, focusing on identifying key network branches and initial disturbances that caused cascading outages. It also developed a key risk identification algorithm based on the maximum value principle. Then, risk constraints were introduced to describe the impact of cascading outages in the economic dispatch model, and a risk management optimization model was proposed in [26] to balance economy and risk. Besides, a statistical method was used to identify critical network devices based on large data for cascading outages. Based on this idea, a state-outage network model with empirical probabilities was formulated in [27] to reduce the blackout risk resulting from cascading outages.

Furthermore, CS has immediate significance in components vulnerability analysis. Ref. [28] used the kernel fuzzy C-means method to predict the outage chain, which located key components and interactions among cascading outages. To identify key components and quantify corresponding impacts, [29] established a probabilistic model to identify the propagation patterns of cascading outages. Extending the above single-layer network, a multi-layer interactive graph was proposed in [30] to predict outage propagations and search mitigation measures, which could more effectively identify the critical components that affect propagations among layers.

Some efforts focused on contingency analysis by machine learning methods. A fuzzy inference data fusion technique that was not affected by fault types, fault resistance, and data asynchronous was proposed in [31] to improve fault location accuracy. Ref. [32] proposed a prediction method based on machine learning to determine the impact of the hurricanes, where components states were divided into damage and operation for obtaining the decision boundary. A single classifier was trained in [33] by using the transient energy function as input to the machine learning algorithm. In [34], a novel reliability evaluation method for integrated power-gas systems was proposed, where the Random Forest feature selection method and Extreme Gradient Boosting regression algorithm were employed. Ref. [35] suggested an online detection model for cloud systems based on SVM, where a systematic parameter-search method called SVM-Grid is established to optimize parameters in the SVM algorithm.

However, due to the complexity of data structure, we often need to improve the traditional machine learning algorithms to solve the actual data analysis problem. In [36], a hybrid unsupervised and supervised machine learning method-based classification system is suggested, which can be efficiently applied to the inspection of related scanning electron microscope images of the electrospun nanofiber. Ref. [37] utilized a multilayer perceptron and the suggested Dendritic Neuron Model (DNM) for multiple application scenarios involving classification, approximation, and prediction. The results showed that the suggested DNM is effective and promising in addressing these problems. Ref. [38] combined SVM and Gaussian process regression (GPR) to evaluate the vulnerability of the equipment excited by transient electromagnetic disturbances. Ref. [39] has introduced a hybrid machine learning method and proved its effectiveness from the perspective of data mining. In [40], the event-based load shedding (ELS) problem was hierarchically modeled as a multi-output classification subproblem for identifying the best shedding location and a regression subproblem to predict the minimum shedding amount. Ref. [41] applied the hierarchical classification and regression approach to combine several machine learning methods together to achieve better predicting performance. To deal with the overfitting problem, [42] combined the bound optimization approach with variation Bayesian inference to derive a novel L1 norm-based Extreme learning machine. Moreover, a novel classification method was suggested in [43] based on neighbor searching and kernel fuzzy c-means approach, which can reduce the impact of parameters uncertainties with dataset classification for the cyber-physical system. In [44], a network attack detection method integrated a flow calculation and deep learning was developed to process high-speed network data and detect complex network attacks. To overcome data imbalance problems in the machine learning model, [45] proposes a weighted undersampling scheme for SVM based on space geometry distance.

III. Vulnerability Analyses of Cascading Outages

A. Cascading Outage Simulation

In cascading outage analysis, any alterations in types and locations of initial outages could lead to different effects. We consider that the outage of a bus or substation will trip all the connecting lines. Accordingly, the power imbalance will occur during cascading outages. Since the cascading outages are propagated very fast, the optimal dispatch of generators cannot be deployed effectively in real time. Hence, the automatic generation control (AGC) along with possible load shedding will be deployed to redistribute line flows.

Specifically, once a power imbalance occurs, the generator and load will take the following actions, where $P_{G,i}(t)$ is the power generation at bus i and time t and $P_{L,i}(t)$ is the load at bus j and time t:

1) Generation Output Ramping: Generators will adjust their power outputs according to their respective ramp rates in the following three cases:

Case 1 (Power Generation Increment): If the power generation is less than the load, i.e., $\sum_{i=1}^{N_b} P_{g,i}(t) < \sum_{i=1}^{N_b} P_{l,j}(t)$, all generators will ramp up their outputs with a ramp rate r at each time interval Δt.

Case 2 (Power Generation Decrement): If the power generation is larger than the load, i.e., $\sum_{i=1}^{N_b} P_{g,i}(t) \geq \sum_{i=1}^{N_b} P_{l,j}(t)$, all generators will ramp down their outputs with a ramp rate r at each time interval Δt.

Case 3 (Fixing the Minimum and Maximum Limits): Once the generation output of a generator reaches its minimum or maximum value, it will remain generation unchanged.

2) Generator Tripping and Load Shedding: In Case 1, load shedding will occur if the power imbalance exceeds the available generation capacity increment. In this case, the curtailment will start with less critical loads and continue until the power balance is restored. In Case 2, if the power imbalance exceeds the available generation capacity decrement, the system will trip some of the generators sequentially. In this case, the lowest priced generator will be tripped first and the process will be continued until the load balance is restored. At any point, if the generation is less than load the condition will be shifted to Case 1.

3) Constrained Power Flow Model: The line flows in the grid will be redistributed after any generation or load adjustments. A constrained AC power flow model is set up for the cascading outage analysis, which is described as

$$\begin{cases} U_0(t) = U_0^{ref}, \qquad \theta_0(t) = \theta_0^{ref} \\ P_{g,i}(t) - P_{l,i}(t) = U_i(t) \sum_{j=1}^{N_b} U_j(t) \big(G_{ij} \cos\theta_{ij}(t) \\ \quad + B_{ij} \sin\theta_{ij}(t)\big) \\ Q_{g,i}(t) - Q_{l,i}(t) = U_i(t) \sum_{j=1}^{N_b} U_j(t) \big(G_{ij} \sin\theta_{ij}(t) \\ \quad - B_{ij} \cos\theta_{ij}(t)\big) \quad i \in \mathbb{B}/0,\ ij \in \mathbb{L} \end{cases} \tag{1a}$$

$$P_{g,0}(t-\Delta t) - r\Delta t \leq P_{g,0}(t) \leq P_{g,0}(t-\Delta t) + r\Delta t \tag{1b}$$

The solution of the power flow equations (1a) will check the constraint (1b) for power balance. If condition (1b) is not satisfied, we will set $P_{g,0}(t)$ as in (1c) and go to 2) for generator ramping and potentially tripping and load shedding.

$$P_{g,0}(t) = \begin{cases} P_{g,0}(t-\Delta t) + r\Delta t \\ \quad P_{g,0}^*(t) - P_{g,0}(t-\Delta t) > r\Delta t \\ P_{g,0}^*(t) \\ \quad -r\Delta t \leq P_{g,0}^*(t) - P_{g,0}(t-\Delta t) \leq r\Delta t \\ P_{g,0}(t-\Delta t) - r\Delta t \\ \quad P_{g,0}^*(t) - P_{g,0}(t-\Delta t) < -r\Delta t \end{cases} \tag{1c}$$

If condition (1b) is satisfied, the power balance will be guaranteed and we can analyze the redistributed power flow on each transmission line. Once the updated power flow on any transmission line exceeds the corresponding transmission capacity, the relays will trip overloaded lines.

4) Optimal Power Flow Model: If the power balance is still not satisfied after a given number of time intervals, i.e., T^{sp}, the load shedding will happen by the optimal power flow. Since the generator ramping can be interrupted by a new cascading outage event (CE) for island d, the duration of generator ramping T_d^{sp} between two CEs can be expressed as follows:

$$T_d^{sp} = \min_{ij\in\mathbb{L}} \left\{T_{ij,d}\right\}, \quad d \in \mathbb{D} \tag{2}$$

$T_{ij,d}$ means that there is no overload line tripping online ij in island d within $T_{ij,d}$ time. Since the load demand at different buses may have different importance, the least important one will be shed first until the system power balance is satisfied, which gives the following optimal power flow model,

$$\min \sum_{i\in\mathbb{L}} \omega_i \left(P_{l,i}(t) - P_{l,i}^{s}(t)\right) \tag{3a}$$

$$\begin{cases} P_{g,i}(t) - P_{l,i}^{s}(t) \\ \quad = U_i(t)\sum_{j=1}^{N_b} U_j(t)\left(G_{ij}\cos\theta_{ij}(t) + B_{ij}\sin\theta_{ij}(t)\right) \\ Q_{g,i}(t) - Q_{l,i}^{s}(t) \\ \quad = U_i(t)\sum_{j=1}^{N_b} U_j(t)\left(G_{ij}\sin\theta_{ij}(t) - B_{ij}\cos\theta_{ij}(t)\right) \\ \quad\quad i\in\mathbb{B},\ ij\in\mathbb{L} \end{cases} \tag{3b}$$

$$P_{g,i}(t-\Delta t) - r\Delta t \le P_{g,i}(t) \le P_{g,i}(t-\Delta t) + r\Delta t, \quad i\in\mathbb{G} \tag{3c}$$

$$0 \le P_{l,i}^{s}(t) \le P_{l,i}(t), \quad \left|Q_{l,i}^{s}(t)\right| \le \eta_{l,i}P_{l,i}^{s}(t), \quad i\in\mathbb{L} \tag{3d}$$

$$U_i^{\min}(t) \le U_i(t) \le U_i^{\max}(t), \quad i\in\mathbb{B} \tag{3e}$$

$$\left|F_{ij}(t)\right| \le F_{ij}^{\max}(t), \quad ij\in\mathbb{L} \tag{3f}$$

5) Vulnerability Metric: For each given initial outage at buses or substations, we simulate the cascading outage process according to the above analysis. The load shedding ratio is calculated for quantifying the system vulnerability [16], given by

$$LOSS = \sum_{i\in\mathbb{B}} \left(P_{l,i}^{0} - P_{l,i}^{\infty}\right) \Big/ \sum_{i\in\mathbb{B}} P_{l,i}^{0} \tag{4}$$

B. Islanding Issue

In a CS, overload line tripping may lead to several islands. Cascading outages can be propagated on each island independently, and realize a steady state in the end. However, when cascading outages propagate on each island, the response times T_d^{sp} for each island are different. To take a synchronization of the cascading simulation for each island in the same CE, the minimum time T_{min}^{sp} of T_d^{sp} is set as the simulation time step to promote further propagation.

TABLE I
ISLANDING PROCEDURE

input	Network topology Ω
output	N and $(\Omega_1,\ldots,\Omega_N)$
Step 1	Initialize N=1 and Ψ=Ω
Step 2	Visit an initial bus i_0 in Ψ and set Φ={}
Step 3	Find all the buses connected to the bus i_0 but haven't been visited before, denoted as $i_{11}, i_{12},\ldots, i_{1n}$. Then, $\Phi=\Phi \cup \{i_{11}, i_{12},\ldots, i_{1n}\}$;
Step 4	for each bus from i_{11} to i_{1n}
Step 5	Find all the buses connected to the bus i_{11} to i_{1n} but haven't been visited before, denoted as i_{21}, $i_{22},\ldots, i_{2n}$. $\Phi=\Phi\cup\{i_{21}, i_{22},\ldots, i_{2n}\}$;
Step 6	if $(i_{21}, i_{22},\ldots, i_{2n})$ is null
Step 7	stop;
Step 8	Else
Step 9	recursion $(i_{21}, i_{22},\ldots, i_{2n})$
Step 10	End
Step 11	End
Step 12	Specify Ω_N=Φ
Step 13	if $\bigcup_{i=1}^{N}\Omega_i = \Omega$
Step 14	return
Step 15	Else
Step 16	N=N+1 and $\Psi = \Omega - \bigcup_{i=1}^{N}\Omega_i$
Step 17	goto Step 2
Step 18	End

Specifically, we use the following steps to determine the islands in CS. At first, we determine the number of islands according to the flowchart in Table I, where either Depth-First-Search or Breadth-First-Search method is employed in Steps 3-9 for traversing the system topology by using recursion [35]. Furthermore, islands are divided into three types: with only generators, with only loads, and with both generators and loads. For islands with only loads or generators, all loads are curtailed and all generators are tripped out, respectively. For islands with both generators and loads, we set a reference bus at each island to balance generation with the load. In this case, the generator with the largest capacity is often chosen as the reference bus in the simulation. Then, CS is conducted in which parallel computing techniques are used to accelerate the simulation

C. Flowchart of CS

The CS flowchart of one bus outage is shown in Fig. 1 and its detailed steps are summarized as follows:

Step 1: Set the initial outage and change the system topology according to the system information and parameters.

Step 2: For each island with imbalanced power, deploy generation ramping and tripping and load shedding as necessary.

Step 3: Solve the constrained power flow model to check the bus power balance. If the power is imbalanced, go to Step 2; otherwise, go to the next step.

Step 4: If the power balance is still not satisfied after the time interval T_d^{sp} the load shedding will happen by the optimal power flow (3).

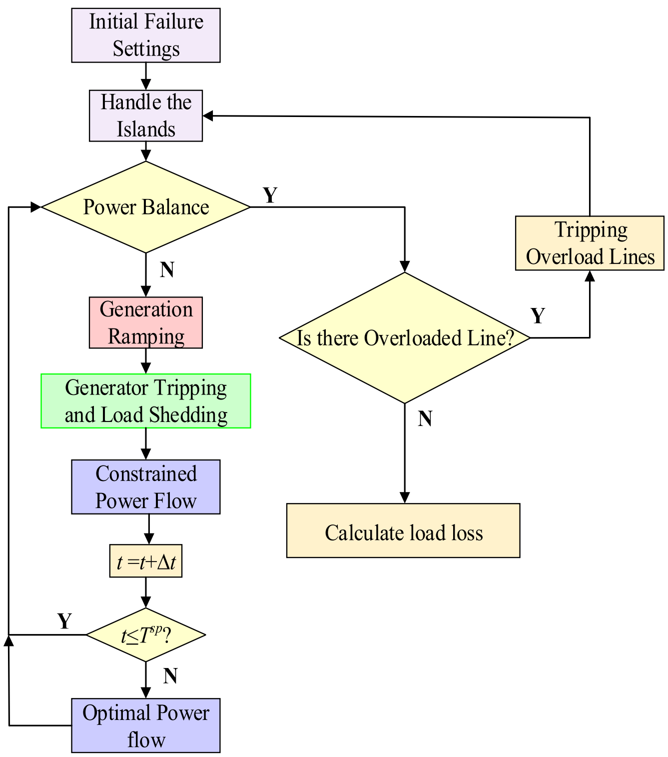

Fig. 1. Flowchart of CS.

Step 5: For the balanced power flow solution, check the overloaded transmission lines. If there is no overloaded line, calculate the vulnerability metric; otherwise, trip all the overloaded lines and go to Step 2.

It should be noted that each bus needs its individual model and the above CS flowchart only obtains one bus outage. Therefore, we can repeatedly conduct the above flowchart with each bus serving as the initial outage, and then we can obtain the vulnerability metric of all buses.

IV. Vulnerability Analysis of Cascading Outage in Random Initial State Based on Machine Learning

A. Feature Selection and Data Acquisition

The random initial power system states will require a stochastic vulnerability analysis using pertinent scenarios. In this paper, we consider the random initial states for generator output and load levels for the stochastic vulnerability analysis. The power system parameters (e.g., resistances and reactances of transmission lines) are considered fixed since they have a minute influence on load losses in power system cascading outages. Moreover, machine learning techniques will be properly devised to analyze the impact of stochastic inputs on vulnerability analysis. Since an N-k contingency has a low probability, the corresponding historical data are deemed limited. To obtain a sufficient number of training data, we employ the simulation method depicted in Fig. 1 and apply the following steps: 1) Choose a scenario with an arbitrary generation output and load level within a certain range near the rated power. 2) Perform the CS according to the flowchart in Fig. 1 and calculate the load shedding for the given initial state.

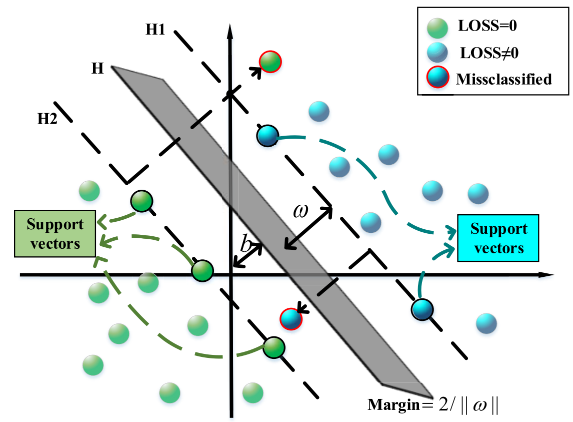

Fig. 2. Schematic diagram of SVM.

Then, a hybrid machine learning model with the combined classification and regression is designed. In this model, the classification model by SVM is employed to judge whether the cascading outage will result in any load shedding. Then, the regression model using the GBR method is employed to describe the relationship between input features and the output indicator only for cases where load shedding is prescribed.

B. Support Vector Machine Classification

The SVM algorithm aims to classify the training data into two sets of with and without load shedding. The approach follows the principle of structural risk minimization for minimizing the deviation between actual and ideal outputs. SVM applies a few support vectors to represent the data set [46], and the kernel function of the original space is used to deal with the nonlinear classification problem by mapping the input variables into a higher-dimensional space.

The N_{train} training data $\{(x_n, y_n, z_n)|n = 1, 2, \ldots, N_{\text{train}}\}$ are chosen for vulnerability analysis. Since the computational error of LOSS is inevitable, a positive threshold is set to determine the value of the classification label z_n, such that $z_n = -1$, if $y_n \leq 10^{-6}$; otherwise $z_n = 1$. The SVM algorithm will separate these data by a hyperplane $\omega x + b = 0$. As shown in Fig. 2, H_1 and H_2 are two planes parallel to a hyperplane H, where the distance between them is called the classification interval. They cross a small number of data points closest to H, which are the support vectors. The hyperplane H separates the two categories and the SVM classifier aims to find the optimal hyperplane that maximizes the classification interval $2/||\omega||^2$, which is the same as minimizing $||\omega||^2$, stated as

$$\min \frac{1}{2}||\omega||^2 \tag{5a}$$

$$s.t.\ z_n\left(\omega^T x_n + b\right) \geq 1, \quad n = 1, \ldots, N_{\text{train}} \tag{5b}$$

To solve the nonlinear problem, it is necessary to introduce the kernel function to map the data into a higher-dimensional space, so that the data can be linearly separable. Here, the kernel function $K(x_p, x_n)$ is chosen as the Radial Basis

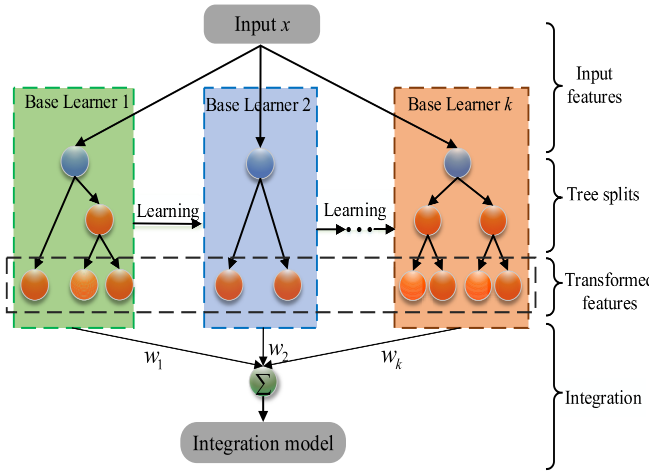

Fig. 3. Schematic diagram of GBR.

Function (RBF) by

$$K\left(x_p, x_n\right) = \exp\left(-\gamma \, ||x_p - x_n||^2\right) = \varphi\left(x_p\right)\varphi\left(x_n\right) \quad (6)$$

Furthermore, the kernel SVM is formulated as

$$\min \frac{1}{2}||\omega||^2 + C\sum_{n=1}^{N_{train}} \xi_n \quad (7a)$$

$$z_n\left(\omega^T \varphi\left(x_n\right) + b\right) \geq 1 - \xi_n,$$
$$\xi_n \geq 0, \ n = 1, \ldots, N_{train} \quad (7b)$$

A larger C will lead to a tighter margin.

The performance of a binary SVM classification model is usually assessed by the four values: true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Then, the accuracy of the SVM classification model is formulated as

$$Accuracy_{SVM} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

C. Gradient Boosting Regression

Once the training data $\{(x_n, y_n, z_n)|n = 1, 2, \ldots, N_{\text{train}}\}$ is classified by SVM, the regression model is set up only for the data with load shedding to find a latent function that maps x_n to y_n. Assume that there are N_s data with load shedding, and giving $\{(x_n, y_n, z_n) \cup z_n = 1|n = 1, 2, \ldots, N_s\}$. However, the learning ability of a single decision tree may be weak for regression, leading to poor generalization performance and over-fitting problems. Hence, GBR uses several least squares regression trees as base learners, fits the residual value of the previous regression tree through the iterative method, and finally adds the results of all regression trees to get the final predicted value [47], which contains input features, three splits, transformed features and integration of several trees. As shown in Fig. 3, sequential improvement will be deployed to enhance the regression trees, where each tree is to reduce the residual and improve the performance by using information from the previous trees.

A loss function $L_n(y_n, f(x_n)) = \|y_n - f(x_n)\|_2^2$ is defined for the joint distribution of x_n and y_n, where the number of regression trees is given. Initialize the model by a constant value $f_0(x)$ as

$$f_0(x) = \arg\min_{f(x)} \sum_{n=1}^{N_s} L_n(y_n, f(x_n)) \quad (9)$$

Apply M iterations to reduce the residual value along the gradient descent direction of the loss function. For the k-th tree, the negative gradient is defined as

$$r_{k,n} = -[\frac{\partial L(y_n, f(x_n))}{\partial f(x_n)}]_{f(x_n)=f_{k-1}(x_n)}, \quad n = 1, \ldots, N_s \quad (10)$$

Form a new training data set $\{(x_n, r_{k,n})|n = 1, 2, \ldots, N_s\}$ and choose the parameterized function $h_k(x_n, a_k)$ to fit the base regression tree as

$$a_k = \arg\min_{\beta_k a_k} \sum_{n-1}^{N_S} \left[r_{k,n} - \beta_k h_k(x_n, a_k)\right] \quad (11)$$

Replace the negative gradient $r_{k,n}$ by $h_k(x_n, a_k)$ and apply a greedy-stagewise approach to approximate the k-th data space by the (k-1)-th data space, leading to

$$f_k\left(x\right) = f_{k-1}\left(x\right) + \mu\beta_k h_k(x), \quad 0 < \mu < 1 \quad (12)$$

β_k is the steepest descent direction that is obtained by a line search method as

$$\beta_k = \arg\min_{\beta_k} \sum_{n=1}^{N_s} L(y_n, f_{k-1}(x_n) + \beta_k h_k(x_n, a_k)) \quad (13)$$

It is observed that the number of trees M and the learning rate μ may affect the solution performance. Generally, the training error is reduced with the increasing number of trees. However, the generalization ability of the model will be weakened if there is an excessive number of trees, and the prediction performance of the model will be reduced due to the over-fitting phenomenon. In order to address this problem, GBR introduces the shrinkage to gradually approach the best result with small steps to avoid the over-fitting problem. In general, a smaller learning rate can improve the generalization ability of the model.

In order to get a prediction model with high accuracy, we employ the ten-fold cross-validation method. It means that the data are randomly divided into ten parts, where nine are randomly used as the training set, and one is applied as the test set. Since the prediction models by using different selected training sets are different, we need to scramble the samples and reselect the training and testing set several times. Finally, we choose the model with the highest accuracy as the final prediction model.

Meanwhile, in order to determine the main parameters affecting the SVM algorithm (i.e., γ and C) and GBR algorithm (i.e., M and μ), this paper adopts the grid search method. Firstly, determine all possible combinations of important parameters affecting the two algorithms within a certain range. Then, train all possible parameter combinations by using the corresponding classification and regression algorithm. Finally,

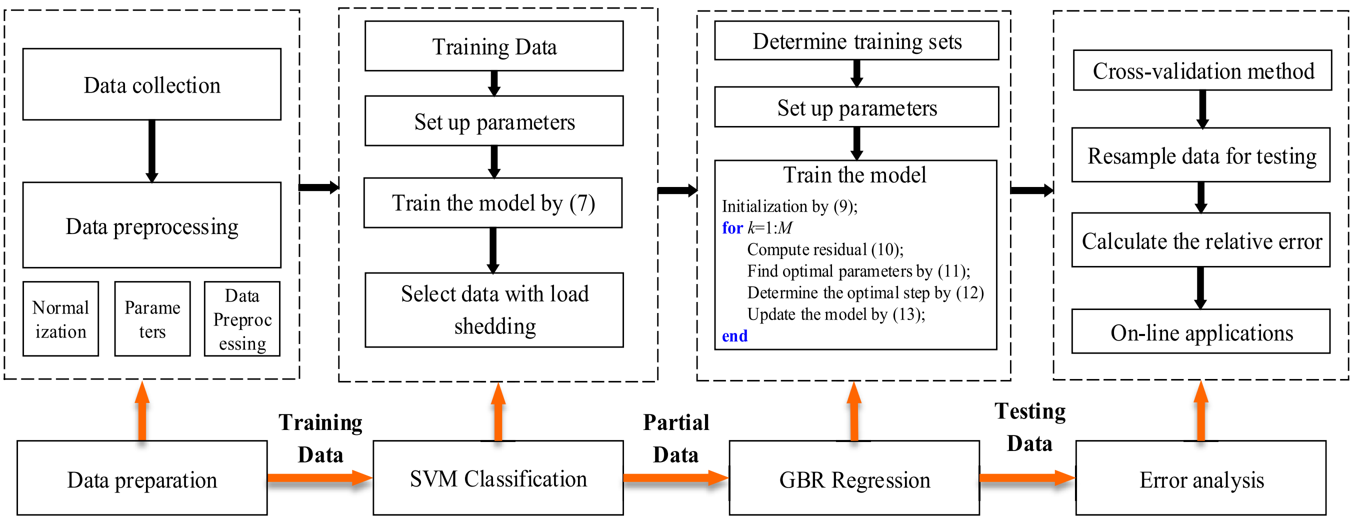

Fig. 4. Flowchart of the proposed hybrid machine learning technique.

the parameter combination with the highest prediction accuracy is taken as the optimal parameters for the two algorithms.

The R-square coefficient is employed as the metric for the accuracy of the GBR model, which is formulated as

$$Accuracy_{GBR} = 1 - \frac{\sum_{m=1}^{N_{test}} (y_m - y_{j,pre})^2}{\sum_{m=1}^{N_{test}} (y_m - y_{mean})^2},$$
$$m = 1, 2, \ldots, N_{test} \tag{14}$$

Finally, since the proposed method is a sequential computation for the classification and regression, the final prediction accuracy of the proposed method is the product of the SVM classification accuracy and the GBR regression accuracy, which can be determined as follows:

$$Accuracy_{prediction} = Accuracy_{SVM} \times Accuracy_{GBR} \tag{15}$$

D. Flowchart of the Proposed Combined Classification and Regression Machine Learning Model

A hybrid machine learning model with the combined SVM classification and GBR regression method is proposed for the power system vulnerability analysis on cascading outages in the presence of random initial states. The detailed flowchart is shown in Fig. 4 with the following four steps:

Step 1: Perform the data collection by CS for randomly selected initial states. Apply data preprocessing, e.g., normalization, to avoid numerical instability resulting from uneven data quantities.

Step 2: Select proper parameters for the SVM classifier and solve the optimization model (7) to determine the optimal hyperplane for dividing the training data into two categories with and without load shedding.

Step 3: Conduct GBR regression by selecting the proper regression model parameters for M iterations using (9)-(13). Based on the classification results, only the training data with load shedding are collected for further regression.

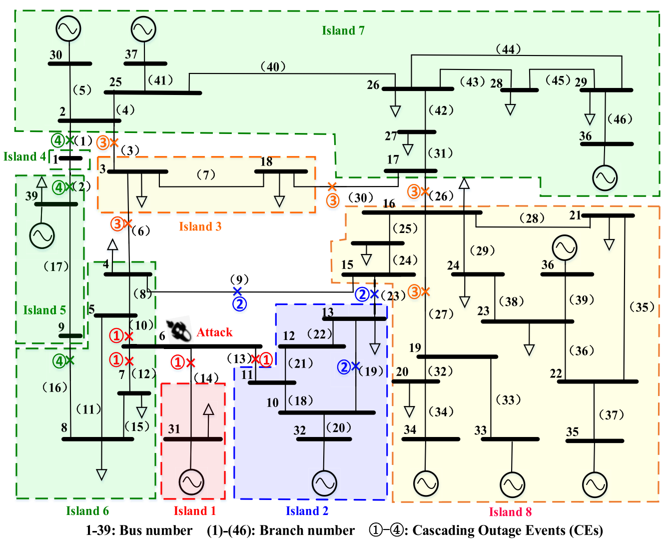

Fig. 5. CS for the initial outage at bus 6.

Step 4: Employ the cross-validation method to resample the data set, evaluate the proposed machine learning model, and quantify the relative error between predicted regression and actual values.

V. Case Studies

A. CS for the IEEE 39-Bus System

To investigate CS, the standard IEEE 39-bus system is chosen with parameters available in MATPOWER [48]. The following assumptions are considered in the simulation: 1) The generator ramp rate r at each Δt is set as 5% of the corresponding capacity. 2) The load with a larger capacity is more important.

The CS results are shown in Fig. 5 and Table II, which contain four cascading outage events (CEs). The initial outage was at bus 6 where four lines connected to bus 6 (i.e., lines 10, 12, 13, 14) were tripped subsequently, splitting the system into two islands with a LOSS of 0.0015%. One island contains only one bus, i.e., 31, where the power balance is straightforward.

TABLE II
ISLANDS AND LOSS OF CS

CEs	Overloaded Lines	# of islands	Buses on each island	LOSS (%)
1	10, 12, 13, 14	2	31	0.0015
			Others	
2	9, 19, 23	3	31	0.0015
			10,11,12,13,32	
			Others	
3	3, 6, 26, 27, 30	4	31	0.0767
			10,11,12,13,32	
			3,18	
			Others	
4	1, 2, 16	8	31	0.1607
			10,11,12,13,32	
			3,18	
			1	
			9,39	
			4,5,7,8	
			2,17,25,26,28,29,30,36,37	
			Others	

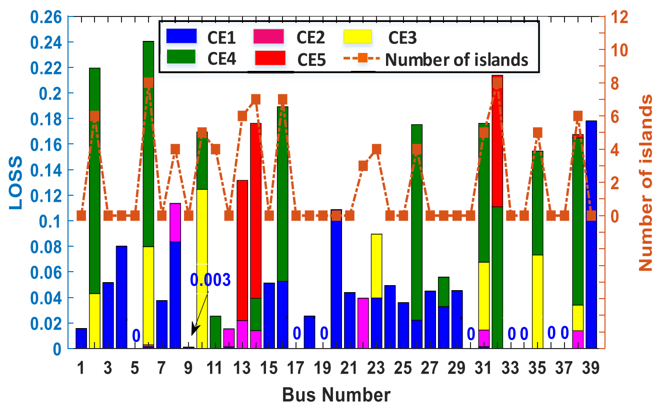

Fig. 6. LOSS for different initial outages.

The other large island has 37 buses, where we learn that active flows on lines 9, 19 and 23 exceed rated values once we recalculate the power flow. Accordingly, the 37-bus island is further split into two islands, i.e., 5-bus and 32-bus islands, leading to a 0.015% LOSS.

The cascading outages will continue to propagate in the 32-bus island, where lines 3, 6, 26, 27, and 30 are disconnected and this 32-bus island is decomposed into 2-bus and 30-bus islands with a LOSS of 0.0676%. Finally, by tripping lines 1, 2, and 16, the 30-bus island is split into 5 islands: single-bus, 2-bus, 4-bus, and 9-bus. This CE contributes to a 0.1607% LOSS. The cascading outage will lead to 8 islands and a total LOSS of 0.2404%. The main LOSS is derived from the third and the fourth CEs, while the first and the second CEs only contribute to 0.03% when the load at bus 9 is curtailed.

The LOSS for different initial outages is shown in Fig. 6, where there are five CEs at most for buses 13, 14, and 32. Meanwhile, the initial outages at either load buses (e.g., 17 and 26) or generator buses (e.g., 31 and 32) will lead to cascading outages. Since a higher LOSS will lead to a higher vulnerability, the initial outage at bus 6 in the test system will result in the largest LOSS and the corresponding highest vulnerability. In contrast, initial outages at buses 5, 17, 19, 30, 33, 35, 36, and 37 do not lead to any load shedding, so the vulnerability on these buses is very low. Besides, when the initial outage occurs at buses 3, 4, 20, and 39, there is only one CE but the LOSS is still large. This is because the corresponding loads are large which will be curtailed after the initial outage. For example, the load at bus 39 accounts for 17.65%. Besides, the bus a larger number of CEs have a higher vulnerability in general (e.g., buses 2, 6, 14, and 32). According to the cumulative results of LOSS, the vulnerability of each bus can be obtained corresponding to vulnerability in this initial state, respectively. Then, the operators can take some specific reinforcement measures for those vulnerable critical buses of the power system.

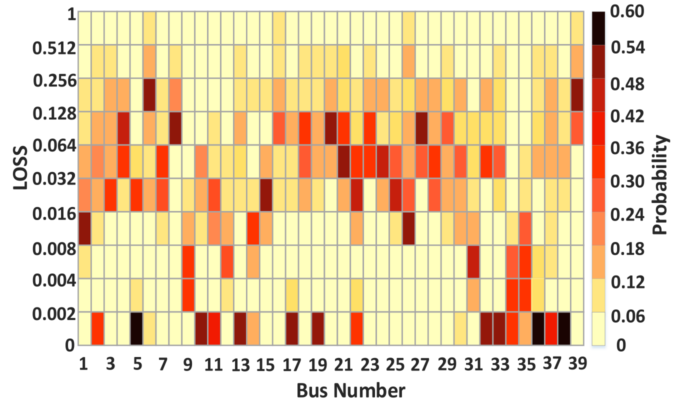

Fig. 7. Probability of LOSS for different initial states.

B. Machine Learning for Vulnerability Analysis

The vulnerability analysis is performed by the proposed hybrid machine learning method for different initial states. Here, the 39-bus power system is utilized again and the data sets are constructed by the following steps:

1) Data sets with 10,000 initial states are generated where generator outputs and load levels are randomly changed between 0.8 and 1.2 of the power ratings.
2) Use the proposed CS model for each data set to calculate the vulnerability metric. Obtain the input features of each data set x_n (i.e., uncertain power and gas load levels), classification labels z_n (i.e., whether the load shedding happens), and regression labels y_n (i.e., the vulnerability metric LOSS).
3) Normalize the 10,000 data sets. Split them into a training set and a testing set by the ten-fold cross-validation method for the SVM classification.
4) Determine the optimal SVM parameters by the grid search method as $C = 0.8$ and $\gamma = 1/30$.

The LOSS distribution for different initial states is depicted in Fig. 7, where certain buses (e.g., 5, 36, 38) have very small LOSS and some other buses (e.g., 6, 39) have large LOSS. Compared with that in Fig. 6 with a given initial state, the LOSS may vary significantly under different initial states. Therefore, initial states will have a major impact on LOSS.

Take buses 6 and 16 with the classification results shown in Fig. 8, where only 10 feature vectors include 5 large generators

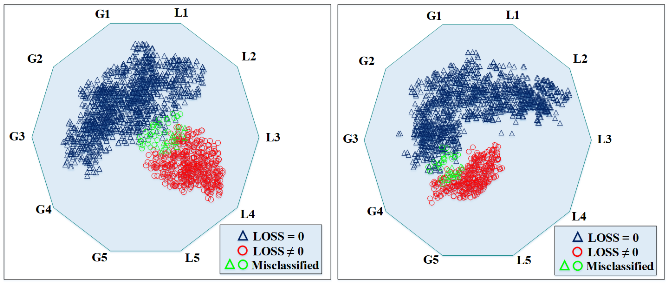

Fig. 8. Classification of data sets in buses 6 and 16.

and 5 large loads due to the large feature dimensions. The vertices of polygons represent the input features, and the points with different colors represent different kinds of data. It can be observed that it is feasible to classify the data and select the data with LOSS≠0 for the subsequent GBR regression training because the data have obvious clusters. The output vectors are in two categories of zero and non-zero LOSS. According to the classification results, we adopt only the data sets with load shedding for the GBR regression method, i.e., LOSS≠0. For bus 6, there are 7,656 data sets with LOSS≠0, which are still split into a training set and a testing set. The parameters are set as learning rate $\mu = 0.01$ with the number of learners $M = 9$.

The precision of the proposed method is shown in Fig. 9. The blue curve shows the accuracy of the first-step SVM classification. The green curve represents the accuracy of the second-step GBR regression by using the data with LOSS≠0 after the SVM classification. The red curve represents the accuracy of the hybrid machine learning method. The accuracy of the proposed machine learning method (i.e. the red curve) is the product of the SVM classification accuracy and the GBR regression accuracy. Therefore, it will be lower than that of individual ones.

It can be found that the minimum and average accuracies of the SVM classification method are 94.35% and 96.91%, respectively. The minimum and average accuracies of the GBR regression method are 97.65% and 99.12 %, respectively. Generally, the hybrid method has good performance and the overall minimum and average accuracies are about 93.12% and 96.05%. In addition, there are several data sets with relatively small LOSS values, which are intuitively on the boundary of two categories. In this scenario, the SVM misclassification may also happen.

All buses are traversed to train their corresponding prediction model. For a certain initial state in the power system, the proposed method is employed to rapidly predict the vulnerability metrics of each bus under cascading outages. Then, the buses with high vulnerability metrics are regarded as the vulnerable critical components in this power system.

C. Comparisons of Large-Scale Systems

The direct GBR regression with the same parameters is employed for comparison. The precision and computation time are shown in Fig. 10. It is obvious that the proposed method will increase the computation time since the additional SVM classification is conducted. Generally, the computation time

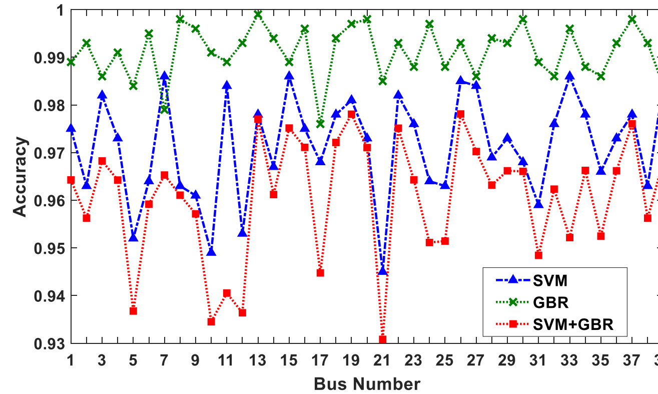

Fig. 9. Classification and regression errors of the proposed method.

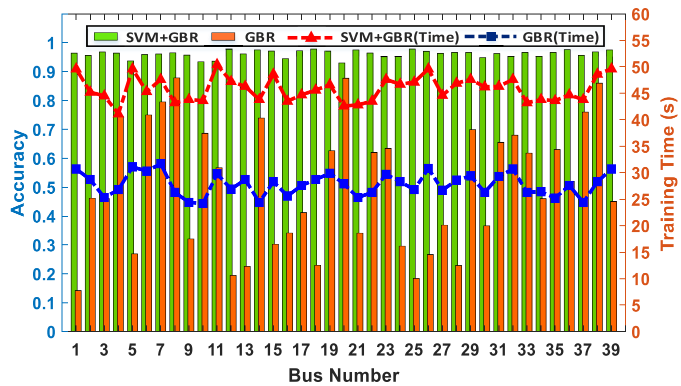

Fig. 10. Computation time and precision for the 39-bus system.

TABLE III
COMPARISONS FOR VULNERABILITY ANALYSIS TIME WITH DIFFERENT METHODS

Test systems	Method 1	Method 2	Method 3	
	Time (s)	Time (s)	Training time (s)	Predict Time (s)
14-bus system	31.3	43.8	46.7	0.739
39-bus system	39.7	76.9	47.6	0.593
57-bus system	69.3	90.6	48.1	0.763
118-bus system	98.4	123.5	51.2	0.839
300-bus system	128.0	213.5	54.8	0.651
2383-bus system	315.2	379.3	59.6	0.778

of the proposed method is less than half of that of the GBR method. In addition, the precision of the proposed method is improved significantly. In particular, the GBR regression method will lead to a very poor precision, which is as low as nearly 20% at some buses. In contrast, the precision of the proposed method exceeds 90%. This is because different initial states will lead to an uneven distribution of LOSS. It will affect the learning performance. The proposed hybrid machine learning method utilizes the classification to select non-zero data sets for regression with an even data distribution of output variables.

Furthermore, the application of the proposed method is simulated for verification on several large-scale test systems including 14-bus, 39-bus, 57-bus, 118-bus, 300-bus, and 2383-bus systems [48]. Each test system adopts the same simulation settings as the 39-bus system. The results are shown

TABLE IV
COMPARISON OF THE PROPOSED METHOD WITH OTHER MACHINE LEARNING METHODS

Methods	14-bus system		39-bus system		57-bus system	
	Accuracy	Training Time (s)	Accuracy	Training Time (s)	Accuracy	Training Time (s)
SVM+GBR	0.988	46.7	0.967	47.6	0.960	48.1
Logistic Regression+GBR	0.397	39.4	0.387	41.9	0.342	42.9
Random Forest+GBR	0.985	73.4	0.961	79.7	0.957	83.4
SVM+ Linear Regression	0.295	34.1	0.243	35.5	0.279	36.7
SVM+ Polynomial Regression	0.437	38.7	0.335	41.7	0.353	43.3
SVM+SVR	0.677	37.5	0.687	42.4	0.619	43.2
Methods	118-bus system		300-bus system		2383-bus system	
	Accuracy	Training Time (s)	Accuracy	Training Time (s)	Accuracy	Training Time (s)
SVM+GBR	0.960	51.2	0.952	54.8	0.949	61.2
Logistic Regression+GBR	0.305	49.7	0.311	48.9	0.287	51.3
Random Forest+GBR	0.956	92.3	0.951	99.7	0.943	114.9
SVM+ Linear Regression	0238	39.2	0.223	42.7	0.205	52.1
SVM+ Polynomial Regression	0.368	45.6	0.335	58.9	0.298	57.5
SVM+SVR	0.628	46.9	0.578	49.7	0.433	58.3

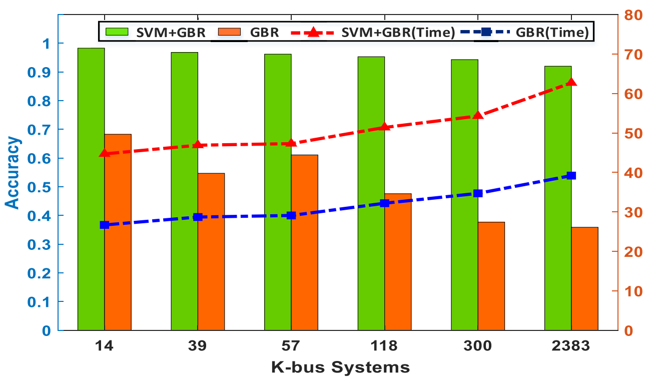

Fig. 11. Computation time and precision on large-scale test systems.

in Fig. 11. Here, the computation time increases slightly for larger systems, while the precision is always higher than 90%. This indicates that the machine learning method is not sensitive to the system size and can also be applied to large-scale systems. In contrast, the GBR regression method will always demonstrate a low accuracy (i.e., smaller than 70%), which decreases for larger systems. For the 2383-bus system, the precision is only 30%.

Finally, to verify the advantages of the proposed method on the vulnerability analysis under random initial states, the CS model based on DC power flow mentioned in [19] (Method 1), the CS model proposed in Section III (Method 2), and the hybrid machine learning model proposed in Section IV (Method 3) are compared in terms of vulnerability analysis time. The simulation settings are as follows: 1) Each method calculates the average vulnerability analysis time for each bus under five random sets of random initial states; 2) Each test power system adopts the same simulation setting and calculation methods of vulnerability metric in Method 1 and Method 2. The results are shown in Table III.

It can be found that when the initial states of a power system change, the average time of re-conducting CS for the traditional model-driven Method 1 and Method 2 is relatively long. In addition, with the scale of the power system increasing, the time for Method 1 and 2 Method also increases significantly (e.g. more than 6 minutes in the 2383-bus power system). In contrast, although the hybrid machine learning-based Method 3 also needs about 1 minute to train the prediction model, the training process can be carried out offline. The online prediction time of Method 3 is no more than 1s, and it is hardly affected by the scale of power systems. Therefore, the proposed hybrid machine learning method can effectively predict the vulnerability to cascading outages of the power systems in random initial states.

D. Comparisons With Other Machine Learning Methods

In order to verify the effectiveness of the proposed machine learning method, some other classical machine learning algorithms including two classification algorithms (Logistic Regression and Random Forest Classification) and three regression algorithms (Linear Regression, Polynomial Regression, and Support Vector Machine Regression) are employed to generate six hybrid machine learning methods for the comparison with the proposed method. Comparisons are investigated on the six test systems. The average accuracy and training time of all buses are shown in Table IV.

It can be found the SVM + GBR leads to the highest accuracy among all the hybrid machine learning methods. In contrast, the SVM + Logical Regression has the lowest accuracy. The reason is that the proposed model takes on strong nonlinear characteristics and the Logical Regression cannot well handle this nonlinear problem. Moreover, the Random forest + GBR has similar accuracy as the proposed SVM + GBR method. However, it is more complex and computationally expensive than the SVM + GBR method. As a result, this paper selects SVM + GBR as the hybrid machine learning method to investigate the relationship between the initial states and the vulnerability metric.

VI. CONCLUSION

This paper proposes a hybrid machine learning method for the power system vulnerability analysis. Initial states are randomly sampled and a cascading outage simulation for each initial state is proposed to quantify the vulnerability metric and generate the training data. Furthermore, a hybrid machine learning model with the combined classification and regression is set up to characterize the relationship between the initial states and the vulnerability metric. Simulation results suggest that the proposed cascading outage simulation can effectively deal with islands and reflect the power system vulnerability. Moreover, the proposed hybrid machine learning model, which applies a combined classification and regression method, can achieve higher precision than a single regression model.

References

[1] Y. Tan, Y. Cao, Y. Li, K. Y. Lee, L. Jiang, and S. Li, "Optimal day-ahead operation considering power quality for active distribution networks," *IEEE Trans. Automat. Sci. Eng.*, vol. 14, no. 2, pp. 425–436, Apr. 2017.

[2] L. Li, Q.-S. Jia, H. Wang, R. Yuan, and X. Guan, "A systematic method for network topology reconfiguration with limited link additions," *J. Netw. Comput. Appl.*, vol. 35, no. 6, pp. 1979–1989, Nov. 2012.

[3] T. Ding, Y. Lin, G. Li, and Z. Bie, "A new model for resilient distribution systems by microgrids formation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 4145–4147, Sep. 2017.

[4] P. Liu, T. Ding, Z. Zou, and Y. Yang, "Integrated demand response for a load serving entity in multi-energy market considering network constraints," *Appl. Energy*, vol. 250, pp. 512–529, Sep. 2019.

[5] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019.

[6] H. Zang, L. Cheng, T. Ding, K. W. Cheung, Z. Wei, and G. Sun, "Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning," *Int. J. Electr. Power Energy Syst.*, vol. 118, Jun. 2020, Art. no. 105790.

[7] H. Zang *et al.*, "Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network," *IET Gener. Transm. Distrib.*, vol. 12, no. 20, pp. 4557–4567, Nov. 2018.

[8] H. Li, G. Hu, J. Li, and M. Zhou, "Intelligent fault diagnosis for large-scale rotating machines using binarized deep neural networks and random forests," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 2, pp. 1109–1119, Apr. 2022, doi: 10.1109/TASE.2020.3048056.

[9] C. Yang, B. Gunay, Z. Shi, and W. Shen, "Machine learning-based prognostics for central heating and cooling plant equipment health monitoring," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 1, pp. 346–355, Jan. 2021.

[10] Z. Zhang *et al.*, "A review of technologies and applications on versatile energy storage systems," *Renew. Sustain. Energy Rev.*, vol. 148, Sep. 2021, Art. no. 111263.

[11] K. Ojand and H. Dagdougui, "Q-learning-based model predictive control for energy management in residential aggregator," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 1, pp. 70–81, Jan. 2022, doi: 10.1109/TASE.2021.3091334.

[12] W. M. Lin, C. H. Wu, C. H. Lin, and F. S. Cheng, "Detection and classification of multiple power-quality disturbances with wavelet multiclass SVM," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2575–2582, Oct. 2008.

[13] L. S. Moulin, A. P. A. Da Silva, M. A. El-Sharkawi, and R. J. Marks, II, "Support vector machines for transient stability analysis of large-scale power systems," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 818–825, May 2004.

[14] D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Gradient boosting feature selection with machine learning classifiers for intrusion detection on power grids," *IEEE Trans. Netw. Service Manag.*, vol. 18, no. 1, pp. 1104–1116, Mar. 2021.

[15] F. Wenli, L. Zhigang, H. Ping, and M. Shengwei, "Cascading failure model in power grids using the complex network theory," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 15, pp. 3940–3949, Nov. 2016.

[16] J. Qi, I. Dobson, and S. Mei, "Towards estimating the statistics of simulated cascades of outages with branching processes," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3410–3419, Aug. 2013.

[17] J. Qi, W. Ju, and K. Sun, "Estimating the propagation of interdependent cascading outages with multi-type branching processes," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1212–1223, Mar. 2017.

[18] Q. Kang, M. Zhou, J. An, and Q. Wu, "Swarm intelligence approaches to optimal power flow problem with distributed generator failures in power networks," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 2, pp. 343–353, Apr. 2013.

[19] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading failure analysis with DC power flow model and transient stability analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 285–297, Jan. 2015.

[20] Y. Cai, Y. Li, Y. Cao, W. Li, and X. Zeng, "Modeling and impact analysis of interdependent characteristics on cascading failures in smart grids," *Int. J. Elect. Power Energy Syst.*, vol. 89, pp. 106–114, Jul. 2017.

[21] R. Yao *et al.*, "Risk assessment of multi-timescale cascading outages based on Markovian tree search," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2887–2900, Jul. 2017.

[22] R. Yao, S. Huang, K. Sun, F. Liu, X. Zhang, and S. Mei, "A multi-timescale quasi-dynamic model for simulation of cascading outages," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 3189–3201, Jul. 2016.

[23] T. Ding, Z. Wang, W. Jia, B. Chen, C. Chen, and M. Shahidehpour, "Multiperiod distribution system restoration with routing repair crews, mobile electric vehicles, and soft-open-point networked microgrids," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4795–4808, Nov. 2020.

[24] C. Zhai, H. Zhang, G. Xiao, and T.-C. Pan, "A model predictive approach to protect power systems against cascading blackouts," *Int. J. Electr. Power Energy Syst.*, vol. 113, pp. 310–321, Dec. 2019.

[25] H. Zhang, C. Zhai, G. Xiao, and T. Pan, "Identifying critical risks of cascading failures in power systems," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 12, pp. 2438–2445, Jun. 2019.

[26] R. Yao, K. Sun, F. Liu, and S. Mei, "Management of cascading outage risk based on risk gradient and Markovian tree search," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4050–4060, Jul. 2018.

[27] L. Li, H. Wu, Y. Song, D. Song, and Y. Liu, "Quantify the impact of line capacity temporary expansion on blackout risk by the state-failure–network method," *IEEE Access*, vol. 7, pp. 183049–183060, 2019.

[28] Y. Pan, F. Mei, C. Zhou, T. Shi, and J. Zheng, "Analysis on integrated energy system cascading failures considering interaction of coupled heating and power networks," *IEEE Access*, vol. 7, pp. 89752–89765, 2019.

[29] J. Qi, K. Sun, and S. Mei, "An interaction model for simulation and mitigation of cascading failures," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 804–819, Mar. 2015.

[30] W. Ju, K. Sun, and J. Qi, "Multi-layer interaction graph for analysis and mitigation of cascading outages," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 2, pp. 239–249, Jun. 2017.

[31] Z. Jiao and R. Wu, "A new method to improve fault location accuracy in transmission line based on fuzzy multi-sensor data fusion," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4211–4220, Jul. 2019.

[32] R. Eskandarpour, A. Khodaei, and J. Lin, "Event-driven security-constrained unit commitment with component outage estimation based on machine learning method," in *Proc. North Amer. Power Symp. (NAPS)*, Sep. 2016, pp. 1–6.

[33] J. Geeganage, U. D. Annakkage, T. Weekes, and B. A. Archer, "Application of energy-based power system features for dynamic security assessment," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1957–1965, Jul. 2015.

[34] S. Li, T. Ding, C. Mu, C. Huang, and M. Shahidehpour, "A machine learning-based reliability evaluation model for integrated power-gas systems," *IEEE Trans. Power Syst.*, vol. 37, no. 4, pp. 2527–2537, Jul. 2022, doi: 10.1109/TPWRS.2021.3125531.

[35] P. Zhang, S. Shu, and M. Zhou, "An online fault detection model and strategies based on SVM-grid in clouds," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 2, pp. 445–456, Mar. 2018.

[36] C. Ieracitano, A. Paviglianiti, M. Campolo, A. Hussain, E. Pasero, and F. C. Morabito, "A novel automatic classification system based on hybrid unsupervised and supervised machine learning for electrospun nanofibers," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 1, pp. 64–76, Jan. 2021.

[37] S. Gao, M. Zhou, Y. Wang, J. Cheng, H. Yachi, and J. Wang, "Dendritic neuron model with effective learning algorithms for classification, approximation, and prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 601–614, Feb. 2019.

[38] Y.-H. Chen *et al.*, "Vulnerability assessment of equipment excited by disturbances based on support vector machine and Gaussian process regression," *IEEE Trans. Electromagn. Compat.*, vol. 63, no. 1, pp. 103–110, Feb. 2021.

[39] S.-W. Ke and C.-W. Yeh, "Hierarchical classification and regression with feature selection," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manag. (IEEM)*, Macao, China, Dec. 2019, pp. 1150–1154.

[40] Q. Li, Y. Xu, and C. Ren, "A hierarchical data-driven method for event-based load shedding against fault-induced delayed voltage recovery in power systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 699–709, Jan. 2021, doi: 10.1109/TII.2020.2993807.

[41] S. Li, T. Ding, W. Jia, C. Huang, J. P. S. Catalao, and F. Li, "A machine learning-based vulnerability analysis for cascading failures of integrated power-gas systems," *IEEE Trans. Power Syst.*, vol. 37, no. 3, pp. 2259–2270, May 2022.

[42] X. Shi, Q. Kang, J. An, and M. Zhou, "Novel L1 regularized extreme learning machine for soft-sensing of an industrial process," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1009–1017, Feb. 2022.

[43] L. Liu, A. Yang, W. Zhou, X. Zhang, M. Fei, and X. Tu, "Robust dataset classification approach based on neighbor searching and kernel fuzzy c-means," *IEEE/CAA J. Autom. Sinica*, vol. 2, no. 3, pp. 235–247, Jul. 2015.

[44] H. Zhang, Y. Li, Z. Lv, A. K. Sangaiah, and T. Huang, "A real-time and ubiquitous network attack detection based on deep belief network and support vector machine," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 3, pp. 790–799, May 2020.

[45] Q. Kang, L. Shi, M. Zhou, X. Wang, Q. Wu, and Z. Wei, "A distance-based weighted undersampling scheme for support vector machines and its application to imbalanced classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 4152–4165, Sep. 2018.

[46] F. Deng, S. Guo, R. Zhou, and J. Chen, "Sensor multifault diagnosis with improved support vector machines," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 1053–1063, Apr. 2017.

[47] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.

[48] R. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2010.

Hongji Zhang (Student Member, IEEE) received the B.S. degree from the School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 2022, where he is currently pursuing the M.S. degree. His major research interests include energy system and power system resilience.

Tao Ding (Senior Member, IEEE) received the B.S.E.E. and M.S.E.E. degrees from Southeast University, Nanjing, China, in 2009 and 2012, respectively, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2015. From 2013 to 2014, he was a Visiting Scholar with the Department of Electrical Engineering and Computer Science, The University of Tennessee, Knoxville, TN, USA. From 2019 to 2020, he was a Visiting Scholar with the Robert W. Galvin Center for Electricity Innovation, Illinois Institute of Technology. He is currently a Professor with the State Key Laboratory of Electrical Insulation and Power Equipment, School of Electrical Engineering, Xi'an Jiaotong University. His current research interests include electricity markets, power system economics and optimization methods, and power system planning and reliability evaluation. He is also an Editor of IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE POWER ENGINEERING LETTERS, *IET Generation, Transmission & Distribution*, and *CSEE JPES*.

Junjian Qi (Senior Member, IEEE) received the B.E. degree in electrical engineering from Shandong University, Jinan, China, in 2008, and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2013. He was a Visiting Scholar with Iowa State University, Ames, IA, USA, in 2012; a Research Associate with the Department of EECS, The University of Tennessee, Knoxville, TN, USA, from 2013 to 2015; a Post-Doctoral Appointee with the Energy Systems Division, Argonne National Laboratory, Lemont, IL, USA, from 2015 to 2017; and an Assistant Professor with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL, USA, from 2017 to 2020. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, USA. His research interests include cascading blackouts, microgrid control, cyber-physical systems, and synchrophasors. He was a recipient of the NSF CAREER Award in 2020, and the Best Paper Awards from IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE PES General Meeting, and IEEE ISGT Asia. He is also an Associate Editor of *IET Generation, Transmission & Distribution* and IEEE ACCESS.

Wei Wei (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2008 and 2013, respectively. From 2013 to 2015, he was a Post-Doctoral Research Associate with Tsinghua University, where he is currently an Associate Professor. He was a Visiting Scholar with Cornell University, Ithaca, NY, USA, in 2014; and a Visiting Scholar with Harvard University, Cambridge, MA, USA, in 2015. His research interests include applied optimization and energy system economics.

João P. S. Catalão (Fellow, IEEE) received the M.Sc. degree from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2003, and the Ph.D. and Habilitation for Full Professor (Agregação) degrees from the University of Beira Interior (UBI), Covilha, Portugal, in 2007 and 2013, respectively. He is currently a Professor with the Faculty of Engineering, University of Porto (FEUP), Porto, Portugal; and a Research Coordinator with INESC TEC. He was the Primary Coordinator of the EU-funded FP7 project SiNGULAR (Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration), a 5.2-million-euro project involving 11 industry partners. His research interests include power system operations and planning, power system economics and electricity markets, distributed renewable generation, demand response, smart grid, and multienergy carriers.

Mohammad Shahidehpour (Life Fellow, IEEE) received the Honorary Doctorate degree from the Polytechnic University of Bucharest, Bucharest, Romania, in 2009. He is currently a University Distinguished Professor, a Bodine Chair Professor, and the Director of the Robert W. Galvin Center for Electricity Innovation, Illinois Institute of Technology, Chicago, IL, USA. He is a member of the U.S. National Academy of Engineering, and a fellow of the American Association for the Advancement of Science and the National Academy of Inventors.