

Article

A Data-Driven Approach for Online Inter-Area Oscillatory Stability Assessment of Power Systems Based on Random Bits Forest Considering Feature Redundancy

Songkai Liu ^{1,2}, Dan Mao ^{1,2}, Tianliang Xue ^{1,2,*}, Fei Tang ³, Xin Li ^{1,2}, Lihuang Liu ^{1,2}, Ruoyuan Shi ^{1,2} , Siyang Liao ³ and Menglin Zhang ⁴

¹ College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China; skliu0120@outlook.com (S.L.); danmao0707@outlook.com (D.M.); ctgxl0903@outlook.com (X.L.); rickliu96@outlook.com (L.L.); cf6221452@outlook.com (R.S.)

² Hubei Provincial Collaborative Innovation Center for New Energy Microgrid, China Three Gorges University, Yichang 443002, China

³ School of Electrical Engineering, Wuhan University, Wuhan 430072, China; ft202134@outlook.com (F.T.); liaosiyang@whu.edu.cn (S.L.)

⁴ School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China; zhangml1207@foxmail.com

* Correspondence: tlx202134@outlook.com

Abstract: To utilize the rapidly refreshed operating data of power systems fully and effectively, an integrated scheme for inter-area oscillatory stability assessment (OSA) is proposed in this paper using a compositive feature selection unit and random bits forest (RBF) algorithm. This scheme consists of offline, update, and online stages, and it can provide fast and accurate estimation of the oscillatory stability margin (OSM) by using the real-time system operating data. In this scheme, a compositive feature selection unit is specially designed to realize efficient feature selection, which can significantly reduce the data dimensionality, effectively alleviate feature redundancy, and provide accurate correlation information to system operators. Then, the feature set consisting of the selected pivotal features is used for the RBF training to build the mapping relationships between the OSM and the system operating variables. Moreover, to enhance the robustness of the scheme in the face of variable operating conditions, an update stage is developed. The effectiveness of the integrated scheme is verified on the IEEE 39-bus system and a larger 1648-bus system. Tests of estimation accuracy, data processing speed, and the impact of missing data and noise data on this scheme are implemented. Comparisons with other methods reveal the superiority of the integrated scheme. In addition, the robustness of the scheme to variations in system topology, distribution among generators and loads, and peak and minimum load is studied.

Keywords: inter-area oscillatory stability; feature redundancy; random bits forest; data-driven



Citation: Liu, S.; Mao, D.; Xue, T.; Tang, F.; Li, X.; Liu, L.; Shi, R.; Liao, S.; Zhang, M. A Data-Driven Approach for Online Inter-Area Oscillatory Stability Assessment of Power Systems Based on Random Bits Forest Considering Feature Redundancy. *Energies* **2021**, *14*, 1641. <https://doi.org/10.3390/en14061641>

Received: 12 January 2021

Accepted: 4 March 2021

Published: 15 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the increasing penetration of clean energy and the wide-area interconnection of modern power systems, the secure operation of such systems is confronted with serious threats [1,2]. In particular, the inter-area oscillatory stability problem is a crucial issue that should be emphasized [3–5]. Power system operation is usually affected by several sources of uncertainty, such as the adjustment of dispatched power generation and variations of load distribution, which will result in changes to the equilibrium points of power system operation [6]. When a system operating point (OP) suffers from a small disturbance caused by such factors, inter-area oscillatory conditions will emerge. If system damping ratio (DR) is insufficient to dampen the oscillation, the oscillatory amplitude will grow, leading to the instability of the system operation or even system splitting [7]. The inter-area oscillatory stability depends on the system's ability to maintain stable operation when faced with the

small naturally and continuously occurring variations [8]. Considering that oscillations accompanied by negative damping and affecting large numbers of machines will lead to economic losses to systems, realizing accurate and real-time inter-area oscillatory stability assessment (OSA) for the system secure operation is necessary.

Traditionally, time-domain (TD) simulation methods [9–11] are employed as primary tools for system stability analysis, which typically rely on modal analysis of the dynamic equations of the system. The DR is commonly used as a criterion for determining whether the system can remain stable [12]. TD simulation methods are straightforward and accurate in theory, and the reliability and accuracy of such methods can be guaranteed when the system model and measurements are of adequate quality. Nevertheless, as modern power systems become more complex and larger in scale, the application of such conventional methods is becoming greatly constrained by their excessive computational burdens and massive time consumption [13]. When faced with changeable operating conditions of a power system, online OSA is a more challenging task.

To realize online OSA for power systems while making full use of rapidly refreshed phasor measurement unit (PMU) measurements, it is recommended to apply data-driven methods. Data-driven methods can provide OSA results immediately once the data from a certain snapshot of power system operation are obtained. In the related literature, several data-driven methods have been utilized to date, including convolutional neural network (CNN) [14], support vector machine (SVM) [15], extreme learning machine (ELM) [16], and decision tree (DT) [13]. Although such methods exhibit good assessment accuracy, they also have notable limitations. The structure of a CNN is usually complex, and massive parameters need to be repeatedly tuned; consequently, expensive computations are difficult to avoid. The SVM and ELM are difficult for application to large systems. The DT usually performs poorly in the case of missing data. Moreover, for data-driven methods, the number of spatial dimensions of the input will grow as the system scale increases, which may seriously affect the computational efficiency [17]. Therefore, to overcome the curse of dimensionality and choose the pivotal variables related to the output index, effective feature selection techniques are also critical for the application of data-driven methods.

An integrated scheme using a compositive feature selection unit and random bits forest (RBF) algorithm is proposed to achieve online OSA in this paper. This scheme is divided into offline, update, and online stages, through which can achieve fast and accurate estimation of the oscillatory stability margin (OSM) using the real-time operating data from a power system. The compositive feature selection unit is specially designed for efficient feature selection, enabling a significant decrease in the data dimensionality, the selection of pivotal features and a reduction in feature redundancy. The pivotal feature set is then utilized to train RBF to build accurate mapping relationships between the OSM and the system operating variables. Finally, online OSM estimation for the system can be rapidly obtained using the real-time PMU measurements of selected features received from a wide area measurement system (WAMS).

The prominent contributions of the proposed integrated scheme can be summed up as follows.

1. In this paper, a data-driven scheme that can improve the OSM assessment accuracy, provide rapid data processing speed and reduce the computing time of the real-time OSA is proposed.
2. A compositive feature selection unit is specially proposed to facilitate OSA. Not only can the pivotal features be selected to enhance the computational efficiency of the scheme, but the problem of feature redundancy is also effectively mitigated.
3. To improve the robustness of the integrated scheme for the unseen network topologies, an update stage is proposed in the scheme considering the impacts of variations in system topology, distribution among generators and loads, and peak and minimum load.
4. This paper analyzes the advantages of the proposed compositive feature selection unit by the comparisons with several other feature selection techniques. Tests of

the robustness of the scheme to missing data are reported and discussed. Moreover, comparisons with other methods illustrate the applicability and superiority of the integrated scheme.

2. Problem Formulation and Supporting Methods

2.1. Problem Formulation of Oscillatory Stability Margin (OSM)

During the operation of a power system, when the system suffers from a small disturbance, it may be difficult for the damping torque to make the system reach a new steady-state operating condition that is same or close to the system condition prior to the disturbance. Such a deficiency of the damping torque may lead to system oscillation or even more serious damage to systems.

By means of modal analysis, the inter-area oscillatory stability of a power system can usually be assessed based on the results of analyzing the nonlinear differential algebraic equations of the system, as shown in Equation (1).

$$\begin{cases} \dot{x} = f(x, y, z) \\ 0 = g(x, y, z) \end{cases} \quad (1)$$

where x represents a state vector, y represents an algebraic vector, and z represents a control vector. Different algebraic equations can be formulated based on differential mathematical models of the dynamic and static components of power systems [18]. The mathematical models for dynamic components are characterized by differential equations, and the mathematical models for static components are represented by algebraic equations.

The nonlinear equations in (1) are linearized around the specific OP, and the Equation (2) is obtained.

$$\begin{cases} \Delta \dot{x} = A \Delta x + B \Delta u \\ \Delta y = C \Delta x + D \Delta u \end{cases} \quad (2)$$

From the perspective of modal analysis theory, each pair of complex conjugate eigenvalues of matrix A represents a system oscillation mode. Accordingly, matrix A needs to be decomposed in the form shown in (3) to analyze the oscillation modes.

$$A = \Phi \Lambda \Psi \quad (3)$$

where Φ and Ψ denote the left eigenvector matrix and right eigenvector matrix, respectively, and Λ denotes the diagonal eigenvalue matrix. Generally, the conjugate eigenvalues of the i th system oscillation mode can be represented in (4).

$$\lambda_i = \sigma_i \pm j\omega_i \quad (4)$$

where σ_i is the damping, and ω_i is the frequency. Finally, the DR of the i th mode is expressed as (5).

$$DR_i = \frac{-\sigma_i}{\sqrt{\sigma_i^2 + \omega_i^2}} \quad (5)$$

Figure 1 exhibits the schematic relation between DR and the oscillatory stability in a system. DR can afford a clear boundary between stable and unstable states, a smooth movement trajectory, and an explicit distance from an unstable OP. Therefore, this paper regards DR as the OSM indicator. Notably, when the DR of a particular mode becomes progressively more insufficient, the corresponding OSM gradually decreases.

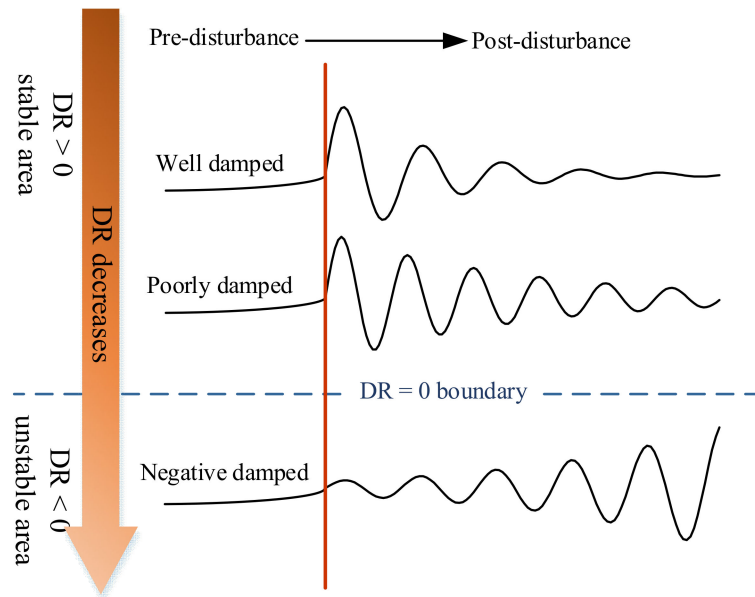


Figure 1. Use of damping ratio (DR) as oscillatory stability margin (OSM).

2.2. Introduction of Supporting Methods

(1) Population Maximal Information Coefficient (MIC_e)

As the power systems continues to develop, the system size and complexity are becoming progressively larger; consequently, it is inevitable that the amount of operating data collected for such a system will continue to increase rapidly. For data-driven OSA, it is important to accomplish the dimensionality reduction in the system operating data. In the machine learning, detecting the correlations between input variables and indicators is a feature selection method that is commonly employed to choose the pivotal variables that are most strongly related to the indicators of interest in order to overcome the curse of dimensionality.

In this paper, the population maximal information coefficient (MIC_e) can be used as a new measure of dependence to detect the degree of correlation between two variables and assign a corresponding score efficiently and equitably [19]. Compared with the previous heuristic approach of the maximal information coefficient (MIC) presented in [20], MIC_e exhibits lower computational complexity, faster processing speed, and better bias and variance properties. MIC_e has been found to be efficient for application to large data sets in biology. Thus, MIC_e is introduced into the field of OSA in this paper. By using MIC_e to detect the correlations between the OSM and the system operating variables, irrelevant features can be eliminated, and then the pivotal feature set can be obtained with the features strongly related to the OSM.

Given a grid G and a point (x, y) , the function $row_G(y)$ is defined to return the row of G containing y and $col_G(x)$ is analogously defined. For a pair of jointly distributed random variables (X, Y) , $(X, Y)|_G$ is adopted to denote $(col_G(X), row_G(Y))$, and $I((X, Y)|_G)$ is used to denote the discrete mutual information (MI) between $row_G(y)$ and $col_G(x)$. For two natural numbers k and l , the set of all $k \times l$ grids can be denoted by $G(k, l)$. Once $G(k, l)$ is determined, $col_G(x)$ will be partitioned into k bins, and $row_G(y)$ will be partitioned into l bins. Moreover, all rows and columns of $(X, Y)|_G$ have the same probability mass in the grid partition. Let

$$I^*((X, Y), k, [l]) = \max_{G \in G(k, [l])} I((X, Y)|_G) \quad (6)$$

where $G(k, [l])$ denotes the set of $k \times l$ grids whose y -axis partition is an equipartition of size l . $I^*((X, Y), [k], l)$ is similarly defined.

Define $I^{[*]}((X, Y), k, l)$ to equal $I^*((X, Y), k, [l])$ if $k < l$ and $I^*((X, Y), [k], l)$ otherwise. Then, a population characteristic matrix $[M](X, Y)$ can be obtained and given by (7).

$$[M](X, Y)_{k,l} = \frac{I^{[*]}((X, Y), k, l)}{\log \min\{k, l\}} \quad (7)$$

where $k > 1$ and $l > 1$.

Finally, the MIC_e between two variables (X, Y) is defined by (8).

$$MIC_{e,B}(X, Y) = \max_{kl \leq B(n)} [M](X, Y)_{k,l} \quad (8)$$

where $B(n) = n^\alpha$, and n is the length of X or Y . $\alpha = 0.6$ is usually recommended because this value has been found to work well in practice [19]. MIC_e has some characteristics as follows.

1. The value of MIC_e falls between 0 and 1.
2. A stronger correlation tends to be assigned a higher score.
3. A correlation between statistically independent variables tends to be assigned a score of 0.

(2) Random Bits Forest (RBF)

For some traditional data analysis approaches, the computational efficiency, memory consumption and applicability to large data sets are routine problems. As a solution, the RBF algorithm is an advanced machine learning method that combines boosting, neural networks, and random forest (RF) [21]. The superiority of RBF in regression prediction has been demonstrated in tests on datasets from the University of California, Irvine (UCI) Machine Learning Repository. In this paper, RBF is applied into the power system field to realize the fast and accurate estimation for OSM by using the system operating data. A schematic of the RBF is shown in Figure 2, and the internal flow can be summed up as the following three steps.

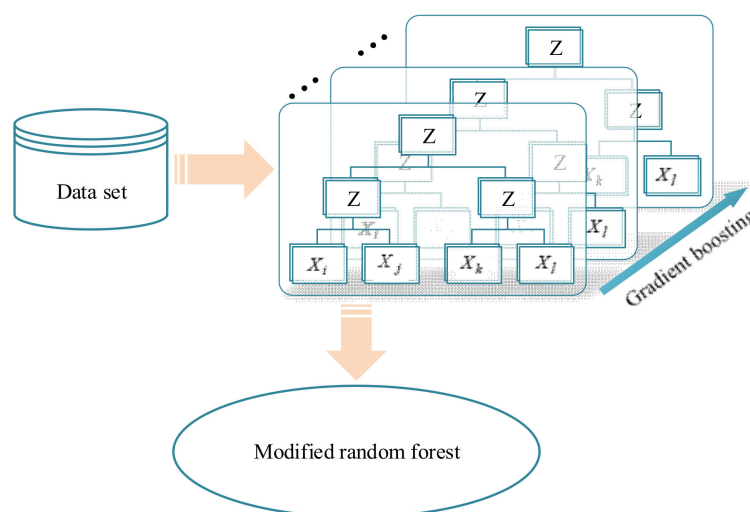


Figure 2. Summarized process of Random Bits Forest (RBF).

First, the input features are standardized by subtracting the mean and dividing by the standard deviation. Second, the standardized features are transformed by means of gradient boosting and random bits (RB). RBs are massive 3-layer sparse neural networks with random weights for each network. The construction of an RB depends on parameters: twist 1 (the number of features connected to each hidden node) and twist 2 (the number of hidden nodes). The standardized features connected to the hidden nodes are randomly assigned. According to a standard normal distribution, the interlayer weights are extracted.

The hidden and top nodes in the network are threshold units, where the threshold for each node is determined by calculating the linear summation of its input for the i th sample Z_i and choosing a random Z_i from among the samples as the threshold. By using a gradient boosting scheme, massive RBs are generated. Finally, the obtained RBs are sent into an RF modified for processing speed. In the modified RF, each tree is grown with a bootstrapped sample and bits, and the best bit is chosen for each split. Furthermore, an acceleration in data processing can be realized by special coding and Streaming SIMD Extensions (SSE).

Based on the above construction procedure, RBF shows better performance in terms of accuracy and robustness than other popular methods [21]. By using RBF as a high-performance predictor, the following integrated scheme for estimating OSM in online OSA is proposed.

3. Proposed Integrated Scheme for Estimating Oscillatory Stability Margin (OSM)

As shown in Figure 3, the proposed integrated scheme consists of offline, update, and online stages. The first stage mainly includes the database construction and the offline training of RBFs. The second stage is proposed to handle the changeable operating condition. The last stage is online OSA based on the real-time operating data of the system. In particular, this paper designs a compositive feature selection unit, and the application of the compositive feature selection unit is incorporated into each stage of the integrated scheme. A more detailed introduction to the integrated scheme is as follows.

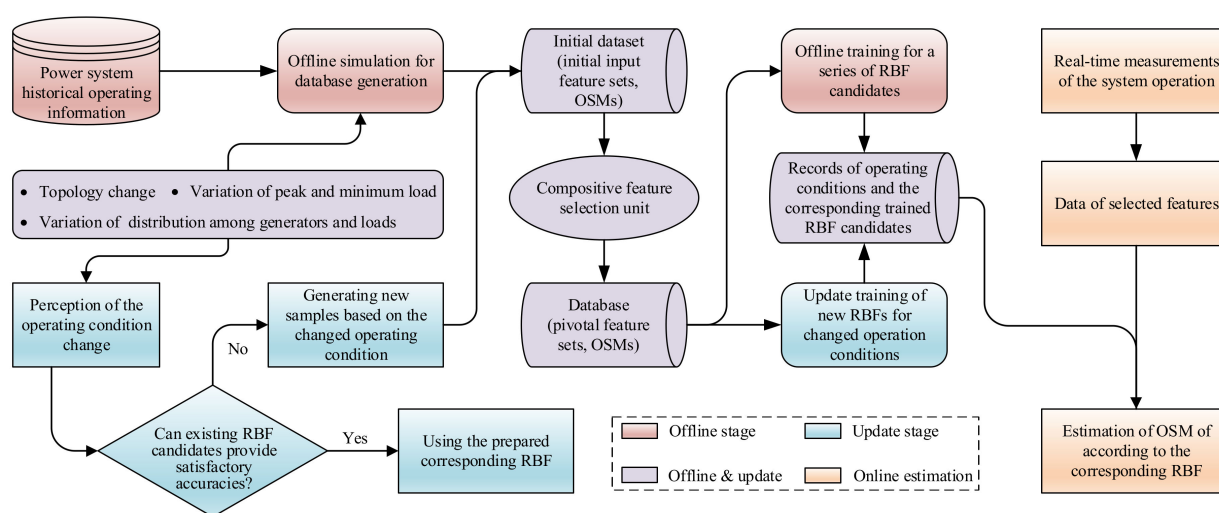


Figure 3. Flow chart of the proposed integrated scheme.

3.1. Process Flow of the Integrated Scheme

(1) Offline Stage

The successful application of a data-driven scheme is inseparable from an efficient database that can supply abundant empirical data for the training of the scheme. The database is composed of the operating data of massive OPs and the corresponding OSMs of the system. RBF can be trained using the operating data as input and the corresponding OSM as output, then accurate mapping relationships between the OSM and the system operating variables can be built. In this scheme, the operating data consist of steady-state operating variables, such as bus voltage magnitude, branch power flows, and bus voltage phase angle, etc.

During the practical operation of the systems, the oscillatory stability is closely related to the variation trend and the composition of system generators and loads [13]. Accordingly, different OPs can be obtained based on the variation of the system generators/loads. With the help of the commercial software PSS/E, the characteristic matrix A can be conveniently obtained, and then the OSMs corresponding to different OPs can be calculated by modal

analysis. In this study, the generation of the database is conducted according to the flow chart shown in Figure 4, following the procedure described below.

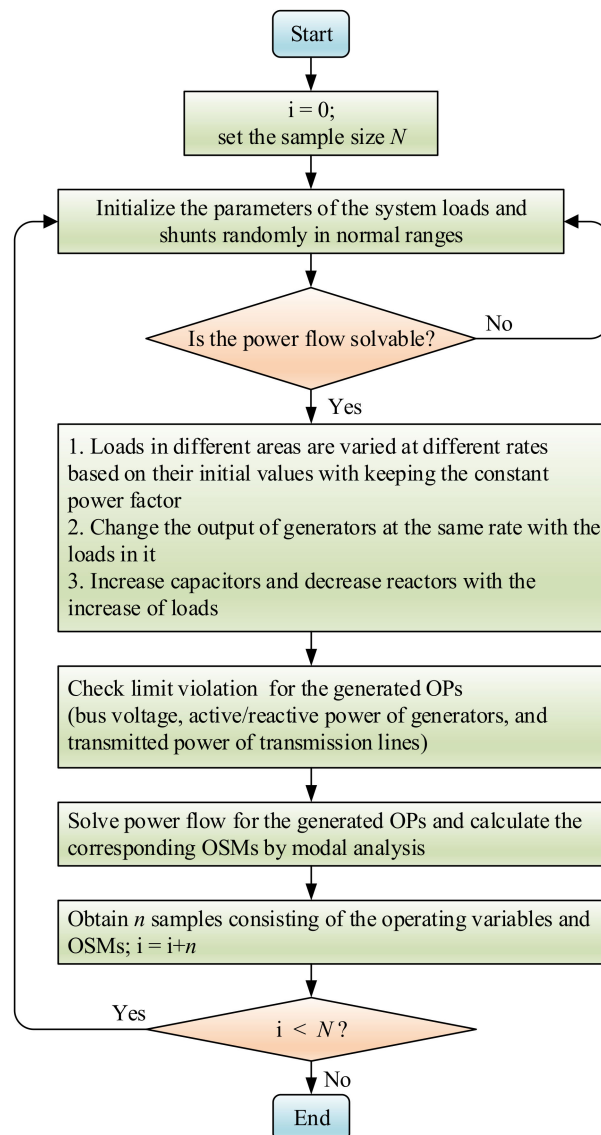


Figure 4. Flow chart of the database generation.

1. Randomly initialize the parameters of the system loads and shunts in their normal ranges by introducing reasonable perturbations in the corresponding parameters.
2. Iteratively change the system load level. Loads in different areas are varied with different rates based on their initial values while keeping a constant power factor. Concurrently, the balance of the load variations mainly relies on the generators in the same area.
3. Increase capacitors and decrease reactors with the increase in loads to simulate the practical operating condition of the systems.
4. Consider various factors influencing the operation of the system during database creation, including variations in system topology, distribution among generators and loads, and peak and minimum load. Contingencies, scheduled maintenance, and economic dispatch can lead to topology change. Optimal power flow considerations may produce the variation of distribution among generators and loads. The peak and minimum load values tend to be different in different seasons, especially between winter and summer. In practice, the system operating condition hardly stays the

same because of such influence factors, and large condition variations may result in an unacceptable decrease in the assessment accuracy of data-driven methods [22]. To accommodate new operating conditions, the retraining using new samples corresponding to the new conditions is usually considered necessary [23]. Nevertheless, retraining is more or less time-consuming and may not meet the requirements for seamless estimation of OSM. Usually, a credible list of possible system operating conditions can be acquired from historical operating information collected and stored by utility companies. Thus, a recommended solution is to prepare an abundant database that includes multiple sample sets corresponding to potential system operating conditions on the basis of the credible list, and then use the prepared sample sets to train a series of RBF candidates beforehand in the offline stage.

In general, the more possible operating conditions and the corresponding trained RBFs are contained in the database, the lower the probability of encountering an unseen condition and the greater the likelihood of realizing seamless OSM estimation.

(2) Update Stage

The update stage is essential to promote the robustness to the complex operating conditions and the generalization ability of the integrated scheme. As shown in Figure 3, the perception to operating condition variations is utilized to actuate the update of the integrated scheme.

In online application of the integrated scheme, when a changed operating condition is encountered due to the abovementioned influence factors, the following strategy will be executed.

If the changed operating condition has previously been recorded in the database and the corresponding RBF candidate has been trained in the offline stage, the prepared candidate will be immediately selected out to replace the original one.

If a match cannot be found, the estimation errors of the available trained RBFs are checked using the new operating condition. If the errors of some candidates are acceptable, then RBF with the highest accuracy among these candidates will be used to conduct OSA for the changed operating condition.

If none of the available RBFs can provide an acceptable accuracy for the changed operating condition, then retraining is required, and a new RBF can be constructed. For this purpose, the new RBF should be trained using the sample set of the changed operating condition. Finally, the changed operating condition and the corresponding RBF will be recorded and added into the database.

With the ongoing application of the integrated scheme, progressively fewer unseen operating conditions will be encountered. In this way, not only can the estimation accuracy be guaranteed, but seamless online OSA can also be achieved.

(3) Online Stage

As shown in Figure 3, online OSA is conducted using the real-time system operating data. With the development of PMU and WAMS, the collection of system operating data has become more convenient and rapid. While the real-time PMU measurements are obtained for a new OP, the data of the input features will be immediately delivered to the corresponding RBF, and then, the online estimation of OSM for this OP can be provided to the system operators.

Furthermore, at the same time that the system operators acquire the OSM value, a threshold can be established to distinguish whether the assessed OP is stable or unstable. By checking the corresponding threshold, any unstable OP will be detected immediately, and the possible event will be sent to the system operators. Simultaneously, the corresponding preventive control strategies will be executed.

3.2. Compositive Feature Selection Unit

As introduced above, an abundant database containing massive samples can be created. However, two issues remain: the features considered in the database may include many variables that are weakly related to the OSM; and some features that are all strongly

related to the OSM may be highly redundant. Regarding the first issue, with an increase in power system operation scale, the feature dimensionality of the database and the number of weakly correlated variables will rapidly increase. Using the raw database with such weakly correlated variables to train RBFs is not conducive to improving the estimation accuracy and will seriously affect the computational efficiency [24]. Regarding the second issue, the selected features are usually faced with high redundancy, meaning that some of them are strongly correlated, and the training of RBFs with such strongly correlated features may obtain redundant relationships between the OSM and the system operating variables. It can lead to a waste of computational resources.

To overcome the issues discussed above, a composite feature selection unit is designed for use in the integrated scheme to achieve efficient feature selection, decrease the feature dimensionality, and alleviate feature redundancy. The composite feature selection unit consists of three steps, elaborated as follows.

Step 1: The flow chart of the first step is shown in Figure 5. This step aims to split the initial input feature set into multiple feature subsets while ensuring that the pairwise correlations between features from different subsets are relatively weak, whereas the pairwise correlations between features in the same subset are relatively stronger.

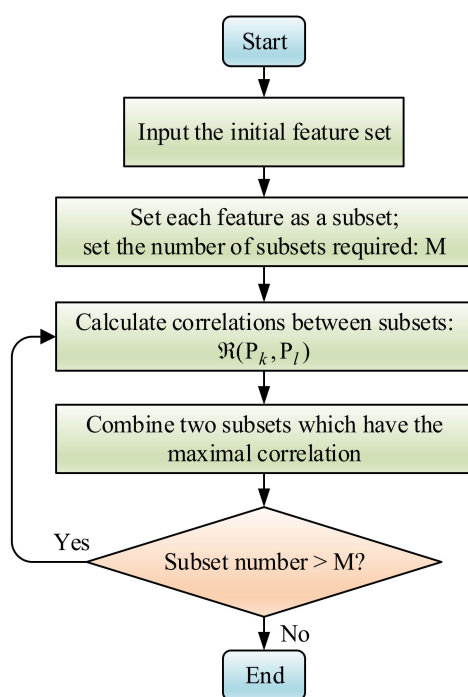


Figure 5. First step in the composite feature selection unit.

Let $\Re(X_i, X_j)$ represent the absolute value of the Pearson correlation coefficient (PCC) between variable X_i and variable X_j , where X_i and X_j are two initial input features. The different feature subsets after partitioning are denoted by P_k and P_l ($k, l = 1, 2, \dots, M$), where M is the user-defined number of subsets. As shown in Figure 5, $\Re(P_k, P_l)$ is used to measure the pairwise correlation between different feature subsets, where

$$\Re(P_k, P_l) = \min\{\Re(X_i, X_j)\} \quad (9)$$

$$X_i \in P_k, X_j \in P_l$$

In accordance with Figure 5, M subsets are acquired and used as the input to the next step.

Step 2: The flow chart of the second step is shown in Figure 6. This step aims to remove the redundant features in each subset, which are significantly related to other features in the same subset. Through this step, the feature redundancy in each subset

can be reduced, and thus, the total number of features and the database dimensionality are decreased.

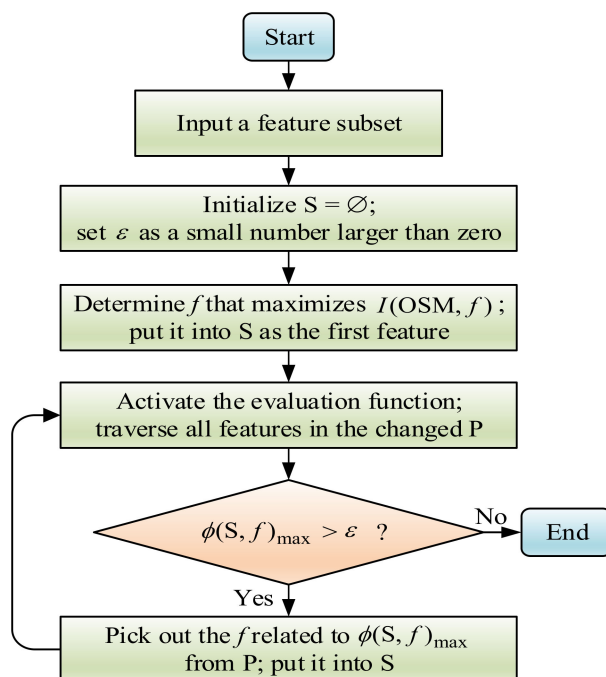


Figure 6. Second step in the compositive feature selection unit.

As shown in Figure 6, an evaluation function ϕ is defined for the features in a given subset. The calculation of ϕ is shown in (10).

$$\phi(S, f) = I(\text{OSM}, f) - \varepsilon \sum_{s \in S} I(s, f) \quad (10)$$

where S is a newly generated feature set consisting of features selected from P , f denotes a feature in subset P from the previous step, $I(\text{OSM}, f)$ represents the MI between the feature and the OSM, ε is a feature selected from P for inclusion in S , and ε is a user-defined parameter for adjusting the number of features that are finally selected. Based on empirical experience, a value of ε in the range of $[0.5, 1]$ is recommended [25].

It should be noted that the processing of each feature subset is independent in this step. In this way, after each subset is processed, an intermediate feature set can be created by combining the processed subsets. From this intermediate feature set, the features with low redundancy among them are significantly related to the OSM. Finally, the intermediate feature set is delivered to the final step for further feature selection based on correlation detection.

Step 3: This step aims to choose the pivotal features from the intermediate feature set on the basis of MIC_e . The different correlations between the features and the OSM can be detected, and then each feature is assigned a score with the corresponding MIC_e value. Based on the score ranking of the features, the highly ranked features are finally selected to establish the pivotal feature set.

Generally, the application of this unit is regarded as a data preprocessing before RBF training. In the offline stage, this unit is used to perform efficient feature selection on the created database to obtain the pivotal feature set. The input for training RBF consists of the obtained pivotal feature set and the corresponding OSM. In the update stage, new RBFs may need to be trained to accommodate unseen operating conditions. In this case, the compositive feature selection unit will be used in a similar way. In the online stage, the real-time data of the features selected based on this unit will be sent to the corresponding RBF to obtain the online OSA results.

4. Application to the IEEE 39-Bus System

The performance of the integrated scheme is tested on the IEEE 39-bus system. As shown in Figure 7, the system includes 19 load buses, 34 transmission lines, and 10 generator buses. The tests are conducted on an Intel Core i7 3.40-GHz CPU with 8 GB of RAM.

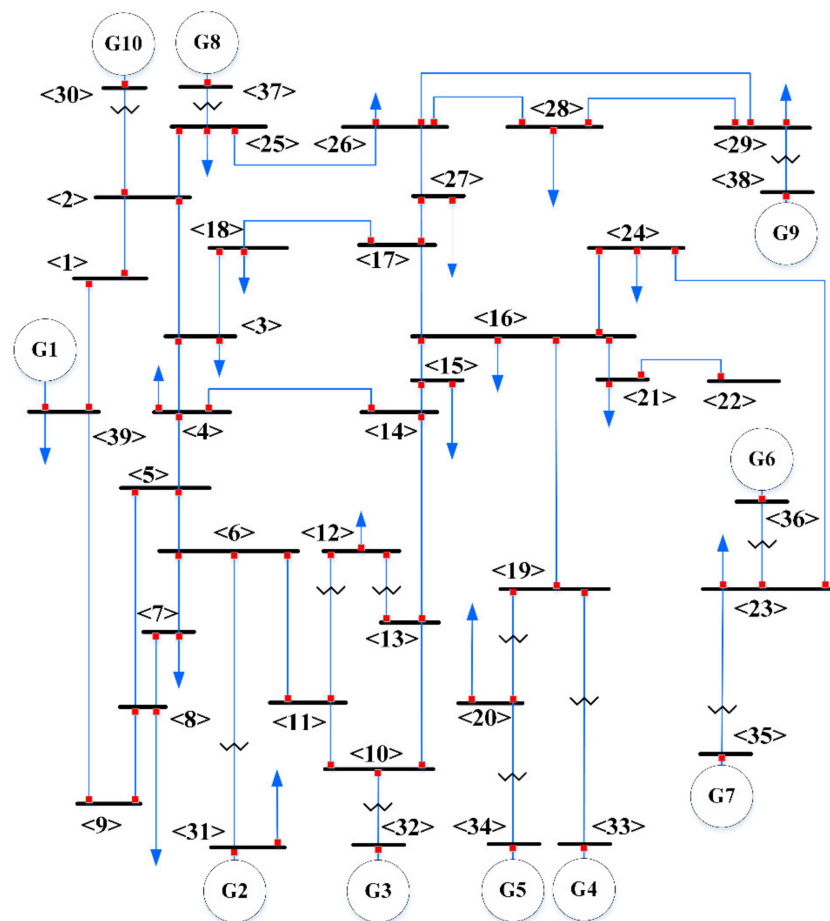


Figure 7. Diagram of the IEEE 39-bus system.

To capture more operation behaviors of the system, the generators and loads parameters are initialized by randomly varying their original distributions within a range of 80–120%. Then, the loads are changed between 70 and 130% of their initial values to simulate load level variations. Moreover, several arbitrary N-k scenarios are considered in database creation for these tests. In the tests, the active power and reactive power of generators, and the transmitted power of transmission lines for all generated OPs conform to their limits. Meanwhile, the bus voltage magnitudes are limited to 0.9–1.1 p.u.

The above simulation work is accomplished with the software PSS/E, Python and MATLAB. PSS/E is adopted to achieve power flow analysis and acquire the characteristic matrix for different OPs, and Python and MATLAB are used to implement automatic dynamic simulations. Subsequently, a total of 2816 cases are generated based on the IEEE 39-bus system, with 465 initial features included in each case. In the tests for this system, 80% of the generated cases are randomly selected as training set, and the remaining cases are used to test the estimation accuracy. To obtain reliable test results, 5-fold cross-validation method is applied in this work.

4.1. Feature Selection Process

In the tests, the following system steady-state operating variables are considered: power levels of loads and generators, reactive power levels of shunts, branch power losses, bus voltage amplitudes and phase angles, and so on. By using a composite feature

selection unit, the variables with corresponding MIC_e scores in the top 5% are selected to construct the pivotal feature set which is used as the input to the RBF.

In general, the raw data can provide the most abundant information about the connotative relationships between the OSM and the system operating variables. However, the fitting of the scheme to massive irrelevant variables will inevitably lead to a decrease in the OSM estimation accuracy. Meanwhile, the high data dimensionality of the raw data will usually result in a long training time. Therefore, with the help of the proposed composite feature selection unit, the pivotal features are selected to eliminate the negative influence of irrelevant and redundant variables, and then, the valuable computing resources will be economized.

In the literature, Fisher discrimination and Relief are typically used as feature selection techniques in the field of the stability assessment of power systems [16,26]. Fisher discrimination and Relief are usually applied to classification problems. Moreover, the high computational burden limits the applicability of Fisher discrimination in large power systems. To expand the application scope of Relief to regression prediction, ReliefF is proposed. Nevertheless, for ReliefF, the inability to effectively remove redundant features remains an inevitable disadvantage.

Compared with such techniques, the composite feature selection unit proposed in this paper is more readily applicable for the regression estimation of OSM. The proposed unit can effectively eliminate redundant features to overcome the feature redundancy problem. Moreover, the unit is a transparent and interpretable tool that ranks features based on the correlation scores assigned by MIC_e , with a shorter computing time and a lower computational burden. The corresponding scores intuitively represent the correlations between the OSM and the system operating variables. System operators can flexibly determine the number of selected high-ranking features by adjusting the score threshold according to different requirements. Furthermore, system operators can gain an understanding of the relations between the parameter changes of certain pivotal features and the OSM variations through visualization in an appropriate coordinate system. Based on such visualizations, the trends of variation of OSM for current system can be grasped roughly by observing the increases or decreases in feature parameters, rather than employing all selected features to estimate the system OSM in each snapshot.

4.2. Oscillatory Stability Assessment (OSA) Performance Test

(1) Evaluation Indices

To measure the OSA accuracy of the proposed integrated scheme, two statistical indices are used in this paper: the residuals squared error (R^2) and the root mean squared error (RMSE) [27,28].

R^2 can be calculated by the equation shown in (11).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

where n represents the number of cases used for calculation, y_i represents the actual OSM_{*i*}, \hat{y}_i is the corresponding OSM estimation value obtained via the integrated scheme, and \bar{y} is the mean of \hat{y}_i .

Generally, an R^2 value closer to 1 indicates a better regression estimation accuracy. In this paper, $R^2 > 0.90$ is used as an acceptable accuracy standard [13,22]. If system operators have stricter accuracy requirements, higher thresholds can be adopted, such as $R^2 > 0.95$.

RMSE especially applies to the condition that specific differences between actual and estimation values are desired. The RMSE can be calculated by the equation shown in (12).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

Ordinarily, the smaller the value of RMSE is, the better regression estimation accuracy. The value of RMSE strongly relies on the initial magnitude of the OSM.

(2) Estimation Test

For practical applications, the proposed integrated scheme should perform well when tested on unseen cases; otherwise, the scheme would be unacceptable due to a lack of generalization ability. In Table 1, the OSM estimation accuracy results of the integrated scheme on the IEEE 39-bus system are summarized. The R^2 value tends to be close to 0.99, and the RMSE value is lower than 0.02. It is observed that the proposed integrated scheme exhibits a desirable OSM estimation accuracy on the IEEE 39-bus system.

Table 1. Accuracy of the proposed integrated scheme.

Test System	R^2	RMSE
IEEE 39-bus	0.9831	0.00187
1648-bus	0.9643	0.00179

RFMSE—Root Mean Squared Error.

5. Performance Test in a 1648-Bus System

To better prove the superiority of the proposed integrated scheme, a 1648-bus system supported by PSS/E is adopted as a larger test system which includes 313 generator buses, 1220 load buses, 2294 transmission lines, and 182 shunts [29].

The same methodology used to create the cases for the IEEE 39-bus system is also utilized for this system, and the hardware configuration and tools used for testing are also the same. A total of 9138 cases with 25,958 features are generated for this test. Using the compositive feature selection unit, the variables with corresponding MIC_e scores in the top 1% are selected to construct the pivotal feature set used for training and testing. As shown in Table 1, the corresponding test results of the 1648-bus system indicate that the proposed integrated scheme also exhibits encouraging OSA accuracy on the 1648-bus system.

To further validate the performance of the proposed integrated scheme in practical applications, additional tests are carried out as follows, including tests for data processing speed, impact of missing data, comparisons with other methods, and tests for variations in system topology, distribution among generators and loads, and peak and minimum load.

5.1. Data Processing Speed

To meet the demands of real-time OSA, the fast data analysis capability of the proposed integrated scheme is crucial. In practical power system operations, the PMU measurements are usually refreshed at least 30 times per second [23]. Therefore, if system operators desire to take advantage of the rapidly refreshed PMU data and realize OSA in each snapshot, in theory, data processing time of each snapshot must be less than $1/30 \approx 0.033$ s.

As shown in Table 2, the results of processing speed tests on the IEEE 39-bus and 1648-bus systems are summarized. It is obvious that after offline training, the scheme can be used to assess more than 400 cases per second for these two systems. Thus, the test results indicate that the data processing speed of the integrated scheme can achieve real-time OSA of power systems.

Table 2. Data processing speed test results.

Test System	Offline Training	New Case Estimation
IEEE 39-bus	89.18 s (2252 case)	1.07 s (564 cases)
1648-bus	857.12 s (7310 cases)	3.84 s (1828 cases)

5.2. Impact of Missing Data

In general, the long-distance transmission of measurements in the power system may lead to missing data, which is an important issue that cannot be ignored. Robustness to

missing data is a crucial characteristic of the proposed integrated scheme for practical applications. In this test, various missing data situations are simulated to explore the impact of missing data on the OSA performance of the integrated solution, where σ represents the number of missing features as a percentage of the initial test set. In this test, four situations are considered: $\sigma = 5\%$, $\sigma = 10\%$, $\sigma = 15\%$, and $\sigma = 20\%$. Offline training is performed using the complete pivotal feature set for each situation.

The corresponding test results of the proposed integrated scheme are shown in Figure 8, the OSA accuracy gradually declines with the increase in the percentage of missing features for both the IEEE 39-bus and 1648-bus systems; however, the scheme still hold a satisfying accuracy even with the maximum value of σ . Thus, the proposed integrated scheme reflects a good robustness to the situation of missing data.

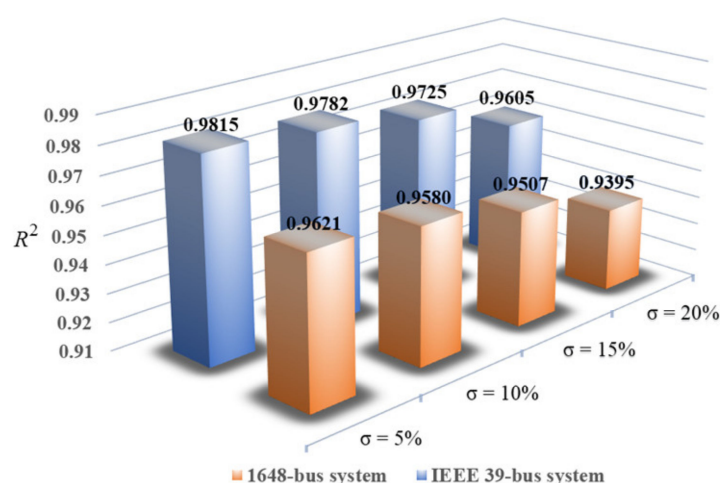


Figure 8. Accuracies under different missing data situations.

5.3. Impact of Noise Data

In practice, the data sets of PMUs in the WAMS may contains noises when PMU data sent to the monitoring center. To verify the impact of noise data on OSA performance of the proposed scheme, the following two scenarios about noise data are tested in this paper. In the test, random noise is added to the data sets, and the total vector error offered by PMUs should be controlled below 1% [16].

1. Noise is added only to the test set.
2. Noise is added to both the training set and test set.

The test results for the considered scenarios are shown in Table 3. Compared with Table 1, although the noise data indeed decreases the OSA accuracy of the proposed scheme, the scheme can still provide acceptable OSA accuracy, and the OSA accuracy of Scenario 1 is less than that of Scenario 2.

Table 3. Accuracies of the proposed integrated scheme for two scenarios of noise data.

Scenario	IEEE 39-Bus System		1648-Bus System	
	R^2	RMSE	R^2	RMSE
Scenario 1	0.9613	0.00221	0.9534	0.00253
Scenario 2	0.9724	0.00195	0.9608	0.00186

5.4. Comparisons with Other Methods

To clarify the superiority of the proposed integrated scheme, comparisons with the CNN, SVM, ELM, and DT methods are performed. Figure 9 provides the corresponding results based on replicated tests for the IEEE 39-bus and 1648-bus systems. It is clearly

observed that the integrated scheme exhibits relatively good accuracy in this test. Compared with other methods, the integrated scheme also has some other advantages, which are discussed as follows.

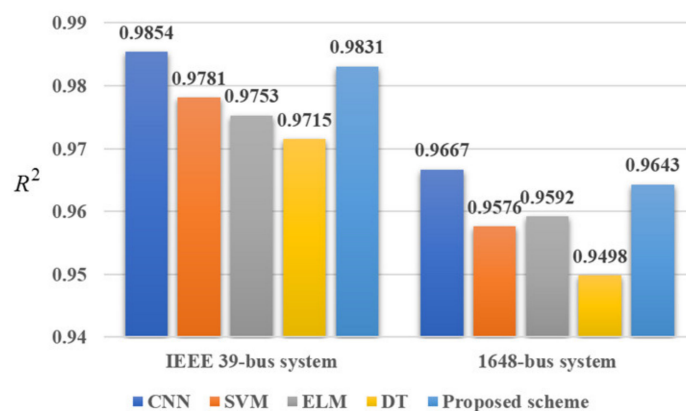


Figure 9. Comparison results between the scheme and other methods.

(1) Comparison with the CNN method: The CNN method exhibits a higher accuracy than the integrated scheme. Nevertheless, since the CNN essentially belongs to the category of deep learning, an enormous computational burden and a high model complexity may be inevitable, and the high-performance hardware demand for the CNN is usually indispensable. During the training of a CNN, massive network parameters are repeatedly tuned, resulting in a long training time. When CNN is applied to a larger system, the number of system operating variables will rapidly increase, which may cause a deeper CNN to be needed. In this case, the complexity of network parameters of the CNN also increases dramatically.

In contrast, although the proposed integrated scheme has a relatively lower accuracy than the CNN method, this scheme has far fewer parameters that need to be repeatedly and manually adjusted. In this scheme, the parameters of the compositive feature selection unit and the RBF obtained via experimental rules are robust. Meanwhile, the use of gradient boosting in the RBF can also promote the computing speed [21]. Thus, the computational burden can be effectively reduced by using this scheme. Furthermore, the features selected by the compositive feature selection unit can be intuitively understood since the scores assigned by this unit directly reflect the correlations between the features and the corresponding OSM. Moreover, by visualizing the correlations, system operators can intuitively grasp the relations between the parameter changes of the operating variables and the variation trend of the corresponding OSM. Therefore, by observing the changes in certain pivotal features, the approximate resulting increase or decrease in the corresponding OSM can be anticipated.

(2) Comparison with the SVM and ELM methods: Figure 9 shows that the accuracies of the SVM and ELM methods are lower than that of the proposed integrated scheme. In the practical application of the SVM, solving for the support vector of a large matrix will consume massive quantities of machine memory and computing time; consequently, SVM is difficult to apply to a large-scale power system [30]. Meanwhile, the ELM has difficulty dealing with complex works because of its natural structure [21].

In this test, the integrated scheme performs better than the SVM and ELM methods. Through special coding and SSE, the application of the RBF to large-scale data sets can be effectively supported, and it performs better when applied to a larger number of samples. The multiple independent boosting chains in the RBF also enable computing speed acceleration.

(3) Comparison with the DT method: As seen in Figure 9, the accuracy of DT method is also lower than that of the proposed integrated scheme. For the application of DT, the missing data is difficult to be handled. Since the DT is essentially a sequential processing

method, incorrect decisions in earlier phases may result in a negative influence on the accuracy of the end result. Furthermore, the DT is easily susceptible to overfitting on large data sets, limiting their performance [22].

In the RBF construction of the proposed integrated scheme, the processed features are sent to a modified RF which is regarded as the final estimator. In other words, the final OSA estimation is achieved by the modified RF. Therefore, each of which is trained and generates predictions independently due to the multiple parallel trees of RF. Different feature groups are assigned to different trees. Therefore, the missing features in a certain group will not obviously influence the performance of trees trained on other feature groups that do not contain the missing features. The multiple independent boosting chains in RBF can also mitigate the local optimum problem and lead to better estimation results.

5.5. Tests for Variations in System Topology, Distribution among Generators and Loads, and Peak and Minimum Load

Robustness to the influence factors of power system operations is an important capacity for the proposed integrated scheme. To examine the impacts of variations in system topology, distribution among generators and loads, and peak and minimum load on the performance of the proposed integrated scheme, corresponding tests are performed. Various unseen system topologies of the 1648-bus system are used to generate test cases, with the results of a part of them are shown in Table 4. Meanwhile, the test results of the variation of distributions among generators and loads are tested for the 1648-bus system and shown in Table 5. Table 6 shows the test results obtained by considering different peak and minimum loads. In these tests, two scenarios are considered for comparison, as follows:

Table 4. Partial test results for variation of topology change.

Out of Service	Type	RMSE ₁	RMSE ₂
Line 683–711	N-1	0.00197	0.00176
G 202	N-1	0.00211	0.00182
G 1515 and Line 1218–1220	N-2	0.00232	0.00185
G 398 and Shunt 1306	N-2	0.00219	0.00178
G 749, Line 315–317 and Line 635–804	N-3	0.00257	0.00180
G 303, Line 535–546 and Shunt 561	N-3	0.00247	0.00174

Table 5. Partial test results for variation of distribution among generators and loads.

Variation Range of the Distribution Relative to the Original Distribution	RMSE ₁	RMSE ₂
70–110%	0.00227	0.00179
75–115%	0.00243	0.00175
85–125%	0.00249	0.00181
90–130%	0.00236	0.00185

Table 6. Partial test results for variation of peak and minimum load.

Minimum Load to Peak Load	RMSE ₁	RMSE ₂
60–120%	0.00208	0.00183
65–125%	0.00221	0.00182
75–135%	0.00214	0.00173
80–140%	0.00231	0.00186

Scenario 1: The integrated scheme is not updated for these operating condition variations. In other words, the used scheme is trained only on the data from the original system topology, with random variations of the original distribution within 80–120%, and variations of the loads within 70–130% of their initial values. By contrast, the test cases are

generated from new system operating conditions. $RMSE_1$ denotes the OSA test accuracy in this scenario.

Scenario 2: The integrated scheme is updated. For each changed operating condition, the composite feature selection unit is newly utilized to choose the pivotal features, and a new RBF is trained accordingly and used for OSA. $RMSE_2$ denotes the OSA test accuracy in this scenario.

The results shown in Tables 3–5 indicate that the integrated scheme can maintain acceptable accuracies for unseen conditions in Scenario 1, although the accuracies are somewhat decreased (in the experiments of this study, $R^2 > 0.90$ is approximately equivalent to $RMSE < 0.0030$ for the cases of the 1648-bus system). Nevertheless, the performance of the scheme in Scenario 2 is better than that in Scenario 1. According to the results for $RMSE_2$, if the system operators require a higher standard of prediction accuracy when faced with variable operating conditions, then updating the integrated scheme and training new RBF candidates is recommended to guarantee good OSA accuracies.

The test results indicate that the integrated scheme has good robustness to operating condition variation. Meanwhile, the update of the scheme is more or less time-consuming. With this in mind, utilizing high-performance and distributed computing platforms is a good solution that can significantly reduce the training time for the update. Moreover, preparing a series of trained RBF candidates and continuing to store new RBFs corresponding to new operating conditions can reduce the probability of encountering unseen conditions over time, so as to alleviate the impact of frequent updates on real-time OSA for power systems.

5.6. Additional Test

To better verify the performance of the proposed approach, it is necessary to compare the proposed approach with some classical methods such as Prony, autoregressive moving average exogenous (ARMAX), and vector fitting (VF) algorithm. In this part, Prony is chosen as the representative of classic methods for the comparative test. Table 7 shows the test results of data processing speed and OSM estimation accuracy of the proposed approach and Prony on the 1648-bus system. According to Table 7, although the proposed approach in this paper requires offline training and has a relatively slower data processing speed than Prony, the calculation time for each case in the online stage can also be maintained in few ms, which is sufficient to satisfy the processing speed requirement of real-time OSA. On the other hand, the results in the Table 7 also indicate that the proposed approach exhibits a better OSM estimation accuracy than Prony.

Table 7. Data processing speed and Oscillatory Stability Margin (OSM) estimation accuracy of the proposed integrated scheme and Prony.

Method	Data Processing Speed	RMSE
Proposed Integrated Scheme	0.0021 s/case	0.00179
Prony	0.0019 s/Case	0.00214

Meanwhile, when the system scale gradually increases, the calculation time of Prony for a case will increase proportionally, and therefore the performance of Prony is limited by the capacity of the computing platform [31]. However, the proposed OSA approach in this paper uses the trained candidate RBFs to perform OSA based on the real-time operating data of the system. Based on the feature selection unit and the mapping relationships constructed in the offline stage, the data processing speed of the approach will not be significantly affected by the increase in input features brought by the system scale expansion.

In addition, Prony is sensitive to noise data and has poor anti-interference ability [32]. ARMAX model is a prediction model for functional time series which is highly dependent on time series data [33,34]. If the complete data in a whole time series of the system is

defective, the ARMAX model will be invalid. The VF algorithm is essentially an iterative process of solving linear least squares problems. When some fitting problems with noise data are encountered, VF algorithm may not be able to achieve convergence [35,36]. In contrast, the test results for noise data of part 5.3 of Section 5 show that the proposed approach in this paper has a good anti-interference ability to noise data. Moreover, the proposed approach does not rely on time series data to establish a database training OSA prediction model, and the OSA results can be quickly given when the operating data of any system state are obtained. Therefore, compared with Prony, ARMAX, and VF algorithm, the method proposed in this paper is a better choice for system operators.

6. Conclusions

This paper proposes an integrated scheme that uses a compositive feature selection unit and the RBF algorithm to realize real-time OSA for increasingly complex power systems based on the rapidly refreshed system measurements. The three stages are designed in the scheme to achieve offline training, scheme update and online OSA for power systems. The designed compositive feature selection unit can not only realize efficient feature selection, but also effectively reduce feature redundancy. This scheme considers influence factors of power system operations including variations in system topology, distribution among generators and loads, and peak and minimum load, and then the update stage is developed to enhance the robustness of the scheme in the face of operating condition variations.

Based on the results of tests on the IEEE 39-bus system and a 1648-bus system, the encouraging OSA accuracy, fast computing speed and robustness to missing data of the integrated scheme are verified. Compared with other data-driven methods (CNN, SVM, ELM, and DT methods), the scheme exhibits relatively high accuracy, and the superiority of this scheme is more clearly reflected through an expanded contrastive analysis. In conclusion, the proposed integrated scheme in power systems has theoretical and practical significance for the online OSA.

Author Contributions: Conceptualization, S.L. (Songkai Liu), D.M., and F.T.; Data curation, D.M., T.X., and F.T.; Formal analysis, X.L. and L.L.; Project administration, S.L. (Songkai Liu) and T.X.; Software, T.X. and M.Z.; Supervision, S.L. (Songkai Liu) and S.L. (Siyang Liao); Visualization, R.S. and X.L.; Writing—original draft, R.S. and L.L.; Writing—review and editing, D.M. and L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Open Foundation of Yichang Key Laboratory of Intelligent Operation and Security Defense of Power System under Grant 2020DLXY06, in part by the University Applied Fundamental Research Project of Yichang City under Grant A19-402-a15, in part by the Open Foundation of Hubei Provincial Key Laboratory for Operation and Control of Cascaded Hydropower Station under Grant 2019KJX11, in part by the Natural Science Foundation of Hubei Province under Grant 2019CFB331, and in part by the National Natural Science Foundation of China under Grant 51607104 and 51807109.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable. No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liao, S.; Xu, J.; Sun, Y.; Bao, Y.; Tang, B. Control of energy-intensive load for power smoothing in wind power plants. *IEEE Trans. Power Syst.* **2018**, *33*, 6142–6154. [\[CrossRef\]](#)
2. Tang, F.; Wang, B.; Liao, Q.; Pisani, C. Research on forced oscillations disturbance source locating through an energy approach. *Int. Trans. Electr. Energy Syst.* **2016**, *26*, 192–207. [\[CrossRef\]](#)
3. Chung, C.Y.; Wang, L.; Howell, F.; Kundur, P. Generation rescheduling methods to improve power transfer capability constrained by small-signal stability. *IEEE Trans. Power Syst.* **2004**, *19*, 524–530. [\[CrossRef\]](#)

4. Ma, J.; Wang, S.; Gao, X.; Wang, Z. Contribution factors based optimal selection of wide-area signals for power system damping control. *Elect. Power Compon. Syst.* **2014**, *42*, 935–944. [\[CrossRef\]](#)
5. Teeuwssen, S.P.; Erlich, I.; El-Sharkawi, M.A.; Bachmann, U. Genetic algorithm and decision tree-based oscillatory stability assessment. *IEEE Trans. Power Syst.* **2006**, *21*, 746–753. [\[CrossRef\]](#)
6. Krismanto, A.U.; Mithulanathan, N.; Kamwa, I. Oscillatory stability assessment of microgrid in autonomous operation with uncertainties. *IET Renew. Power Gener.* **2017**, *12*, 494–504. [\[CrossRef\]](#)
7. Wang, T.; Bi, T.; Wang, H.; Liu, J. Decision tree based online stability assessment scheme for power systems with renewable generations. *CSEE J. Power Energy Syst.* **2015**, *1*, 53–61. [\[CrossRef\]](#)
8. Preece, R.; Milanovic, J.V. Risk-based small-disturbance security assessment of power systems. *IEEE Trans. Power Deliv.* **2015**, *30*, 590–598. [\[CrossRef\]](#)
9. Ma, J.; Dong, Z.Y.; Zhang, P. Comparison of BR and QR eigenvalue algorithms for power system small signal stability analysis. *IEEE Trans. Power Syst.* **2006**, *21*, 1848–1855. [\[CrossRef\]](#)
10. Ma, J.; Wang, S.; Wang, Z.; Thorp, J.S. Power system small-signal stability region calculation method based on the guardian map theory. *IET Gener. Transm. Distrib.* **2014**, *8*, 1479–1488. [\[CrossRef\]](#)
11. Rueda, J.L.; Colome, D.G.; Erlich, I. Assessment and enhancement of small signal stability considering uncertainties. *IEEE Trans. Power Syst.* **2009**, *24*, 198–207. [\[CrossRef\]](#)
12. Kundur, P.; Wong, G.J.; Wang, D.Y.; Lauby, M. A comprehensive computer program package for small signal stability analysis of power systems. *IEEE Trans. Power Syst.* **1990**, *5*, 1076–1083. [\[CrossRef\]](#)
13. Zheng, C.; Malbasa, V.; Kezunovic, M. Regression tree for stability margin prediction using synchrophasor measurements. *IEEE Trans. Power Syst.* **2013**, *28*, 1978–1987. [\[CrossRef\]](#)
14. Tan, B.; Yang, J.; Pan, X.; Li, J.; Xie, P.; Zeng, C. Representational learning approach for power system transient stability assessment based on convolutional neural network. *J. Eng.* **2017**, *2017*, 1847–1850. [\[CrossRef\]](#)
15. Mohammadi, H.; Khademi, G.; Dehghani, M.; Simon, D. Voltage stability assessment using multi-objective biogeography-based subset selection. *Int. J. Electr. Power Energy Syst.* **2018**, *103*, 525–536. [\[CrossRef\]](#)
16. Xu, Y.; Dong, Z.Y.; Zhao, J.H.; Zhang, P.; Wong, K.P. A reliable intelligent system for real-time dynamic security assessment of power systems. *IEEE Trans. Power Syst.* **2012**, *27*, 1253–1263. [\[CrossRef\]](#)
17. Sun, M.; Konstantelos, I.; Strbac, G. A deep learning-based feature extraction framework for system security assessment. *IEEE Trans. Smart Grid* **2019**, *10*, 5007–5020. [\[CrossRef\]](#)
18. Li, X.; Zheng, Z.; Ma, Z.; Guo, P.; Shao, K.; Quan, S. Real-time approach for oscillatory stability assessment in large-scale power systems based on MRMR classifier. *IET Gener. Transm. Distrib.* **2019**, *13*, 4431–4442. [\[CrossRef\]](#)
19. Reshef, Y.A.; Reshef, D.N.; Finucane, H.K.; Sabeti, P.C.; Mitzenmacher, M. Measuring dependence powerfully and equitably. *J. Mach. Learn. Res.* **2016**, *17*, 1–63.
20. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* **2011**, *334*, 1518–1524. [\[CrossRef\]](#)
21. Wang, Y.; Li, Y.; Wen, K.; Shugart, Y.Y.; Xiong, M.; Jin, L. Random bits forest: A strong classifier/regressor for big data. *Sci. Rep.* **2016**, *6*. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Fan, Y.; Liu, S.; Qin, L.; Li, H.; Qiu, H. A novel online estimation scheme for static voltage stability margin based on relationships exploration in a large data set. *IEEE Trans. Power Syst.* **2015**, *30*, 1380–1383. [\[CrossRef\]](#)
23. Liu, S.; Liu, L.; Fan, Y.; Zhang, L.; Huang, Y.; Zhang, T.; Cheng, J.; Wang, L.; Zhang, M.; Shi, R.; et al. An integrated scheme for online dynamic security assessment based on partial mutual information and iterated random forest. *IEEE Trans. Smart Grid* **2020**, *11*, 3606–3619. [\[CrossRef\]](#)
24. Wang, B.; Fang, B.; Wang, Y.J.; Liu, H.S.; Liu, Y.L. Power system transient stability assessment based on big data and the core vector machine. *IEEE Trans. Smart Grid* **2016**, *7*, 2561–2570. [\[CrossRef\]](#)
25. Battiti, R. Using mutual information for selecting features in supervised neural net learning. *IEEE Trans. Neural Netw.* **1994**, *5*, 537–550. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Jensen, C.A.; El-Sharkawi, M.A.; Marks, R.J. Power system security assessment using neural networks: Feature selection using fisher discrimination. *IEEE Trans. Power Syst.* **2001**, *16*, 757–763. [\[CrossRef\]](#)
27. Liu, S.; Liu, L.; Yang, N.; Mao, D.; Zhang, L.; Cheng, J.; Xue, T.; Liu, L.; Yan, G.; Qiu, L.; et al. A data-driven approach for online dynamic security assessment with spatial-temporal dynamic visualization using random bits forest. *Int. J. Electr. Power* **2021**, *124*, 106316. [\[CrossRef\]](#)
28. Liu, S.; Shi, R.; Huang, Y.; Li, X.; Li, Z.; Wang, L.; Mao, D.; Liu, L.; Liu, S.; Zhang, M.; et al. A data-driven and data-based framework for online voltage stability assessment using partial mutual information and iterated random forest. *Energies* **2021**, *14*, 715. [\[CrossRef\]](#)
29. *PSS/E Users Manual*; Shaw Power Technologies: Baton Rouge, LA, USA, 2004.
30. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Saraf, P.; Venayagamoorthy, G.K.; Luitel, B. Online oscillation monitoring of synchronous generators using Parallel-Prony analysis. In Proceedings of the 2014 IEEE Power & Energy Society Innovative Smart Grid Technologies, Washington, DC, USA, 19–22 February 2014; pp. 19–22.

-
32. Yalcin, N.A.; Vatansever, F. A new hybrid method for signal estimation based on Haar transform and Prony analysis. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 6501409. [[CrossRef](#)]
 33. González, J.P.; Roque, A.M.S.M.S.; Pérez, E.A. Forecasting functional time series with a new hilbertian ARMAX model: Application to electricity price forecasting. *IEEE Trans. Power Syst.* **2018**, *33*, 545–556. [[CrossRef](#)]
 34. Liu, H.; Zhu, L.; Pan, Z.; Bai, F.; Liu, Y.; Liu, Y.; Patel, M.; Farantatos, E.; Bhatt, N. ARMAX-based transfer function model identification using wide-area measurement for adaptive and coordinated damping control. *IEEE Trans. Smart Grid.* **2017**, *8*, 1105–1115. [[CrossRef](#)]
 35. Shi, G. On the nonconvergence of the vector fitting algorithm. *IEEE Trans. Circuits Syst. II* **2016**, *63*, 718–722. [[CrossRef](#)]
 36. Górnaiak, P.; Bandurski, W. A new universal approach to time-domain modeling and simulation of UWB channel containing convex obstacles using vector fitting algorithm. *IEEE Trans. Antenn. Propag.* **2014**, *62*, 6394–6405. [[CrossRef](#)]