



## OPEN ACCESS

## EDITED BY

Bo Li,  
Guangxi University, China

## REVIEWED BY

Shunbo Lei,  
The Chinese University of Hong Kong,  
Shenzhen, China  
Zhen Li,  
Guangxi University, China

## \*CORRESPONDENCE

Yang Fu,  
✉ hustyangfu1314@163.com

RECEIVED 15 October 2025

REVISED 11 November 2025

ACCEPTED 12 November 2025

PUBLISHED 01 December 2025

## CITATION

Fu Y, Guosong W, Yi X, Qinfeng M, Su A and Junquan C (2025) Small signal stability prediction and correction control algorithm for wind power systems based on LightGBM. *Front. Energy Res.* 13:1725371. doi: 10.3389/fenrg.2025.1725371

## COPYRIGHT

© 2025 Fu, Guosong, Yi, Qinfeng, Su and Junquan. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Small signal stability prediction and correction control algorithm for wind power systems based on LightGBM

Yang Fu\*, Wang Guosong, Xu Yi, Ma Qinfeng, An Su and Chen Junquan

Guizhou Power Grid Co., Ltd. Power Dispatch Control Center, Guiyang, Guizhou, China

With the increasing penetration of wind power in electrical power systems, its impact on small signal stability has become increasingly significant. The volatility and uncertainty introduced by wind power integration pose new challenges to power system stability analysis and control. To accurately predict and correct the small signal stability of wind power systems, this paper proposes a prediction and correction method based on the Light Gradient Boosting Machine (LightGBM) algorithm. The LightGBM model offers high computational efficiency and low memory consumption, enabling large-scale data processing and automatic handling of missing values, thereby accurately extracting grid characteristics. To verify the reliability of the proposed method, Gaussian white noise with three different signal-to-noise ratios is introduced to evaluate the model's performance and robustness. Case studies are conducted using a 3-machine 9-bus system and a 10-machine 39-bus system, in which certain conventional generators are replaced with aggregated wind farms. By applying the LightGBM model to predict the system's minimum damping ratio and constructing a correction optimization model based on damping ratio sensitivity, unstable operating states are effectively adjusted. Simulation results demonstrate that the proposed method achieves high prediction accuracy and rapid correction response, confirming its feasibility and effectiveness in small signal stability analysis and control of wind-integrated power systems.

## KEYWORDS

LightGBM, doubly fed induction generators, wind power generators, small signal stability analysis, machine learning, damping ratio sensitivity

## 1 Introduction

With the continuous progress of science and technology and the sustained development of the economy, environmental pollution and excessive consumption of resources have increasingly become the focus of social attention. In this context, new energy generation technologies mainly based on wind power are rapidly developing. Compared to traditional energy generation methods, wind power generation does not produce carbon emissions during operation, is sustainable, and power generation facilities can be built on land or deployed in the ocean. As an important component of the future power grid, wind power generation has received high attention from the country.

However, it is difficult to avoid small interference problems during the operation of the power system. Once the system becomes unstable due to small interference,

it will seriously affect its normal operation. Wind turbines have characteristics such as randomness, volatility, and intermittency. When they are integrated into a large power grid, they can have an impact on the small signal stability of the power system. Therefore, effective monitoring and adjustment of power systems containing wind power generation are crucial for maintaining the stable operation of the power system.

Common methods for analyzing small signal stability in power systems include damping torque method (Yanfeng et al., 2020; Gibbard et al., 2015) and eigenvalue analysis method. The derivation of damping torque analysis method is complex and commonly used to analyze single machine infinite bus systems. Therefore, the eigenvalue analysis method is commonly used for multi machine power systems (Jiang et al., 2021). The eigenvalue calculation method is based on the Lyapunov linearization method, which transforms nonlinear problems into linear problems and obtains the state matrix of linear differential algebraic equations based on eigenvalues to determine the stability of the system. Traditional eigenvalue calculation methods include QR (QR Algorithm) method, Arnoldi, and Jacobi Davidson. These methods require a large amount of computation and are difficult to reflect the real-time operating status of the power system.

Frequency domain analysis is a fundamental tool for small signal stability analysis of multi machine power systems. The advantage of this method is that it can handle any type of generator model; it can directly obtain frequency domain data and has good computational efficiency. However, for large-scale power systems, this method still requires a significant amount of computation and cannot meet the requirements of fast calculation (Zeng and Jun, 2020).

In recent years, the research on nonlinear theory has achieved extensive development. For instance, bifurcation theory and chaos theory have been gradually introduced into the study of small-signal stability analysis in power systems. A Hopf Bifurcation Point (HB) refers to an equilibrium point where a pair of conjugate eigenvalues crosses the imaginary axis from the left half-plane to the right half-plane in the complex plane as system parameters change. The imaginary axis in the complex plane, as described by this concept, serves as the boundary condition for the small-signal stability of a system, and thus can be applied to the research on issues related to system small-signal stability. However, with the expansion of system scale and the increase in the order of differential-algebraic equations (DAEs), the computational load of such algorithms has gradually increased. In some cases, the solution even becomes computationally intractable, which limits their application in the small-signal stability analysis of large-scale power systems. Chaos is a more complex nonlinear phenomenon than bifurcation theory. Continuous bifurcation phenomena may lead to the emergence of chaos, but its application in nonlinear dynamic systems remains relatively immature (Jia and Yixin, 2001).

The time-domain simulation method solves the time response curves of each state variable in the system after small disturbances based on numerical analysis. It then calculates key modal parameters such as the system's damping ratio and oscillation frequency through curve analysis. However, in the small-signal stability analysis of power systems, although the time-domain simulation method can determine whether the system is stable by analyzing the simulated calculation curves, it fails to provide the small-signal stability margin of the system (Ju et al., 2019).

With the expansion of power grid scale and the integration of new energy sources, the spatiotemporal relationships of power grids have become increasingly complex. Traditional small-signal stability analysis methods and manual screening approaches struggle to quickly reflect the real-time status of power grids. In recent years, machine learning has been introduced into power system research due to its advantages—such as eliminating the need to establish complex mathematical models and delivering fast and accurate prediction results (Zhou and Yang, 2016). It can address small-signal stability issues by combining the strengths of machine learning and numerical calculation methods. Specifically, damping ratio data of the power system is obtained from deterministic linear models as historical training data. Subsequently, real-time measured data is input into the trained model to generate results, thereby enabling real-time evaluation of system stability.

As machine learning is gradually integrated into power system analysis, researchers have begun to apply shallow learning algorithms to small-signal stability analysis. However, shallow learning has limited data processing capabilities and cannot handle data well in large-scale systems. When processing massive sample data in wind power-integrated power systems, it may face issues such as long prediction time and insufficient prediction accuracy (Yin et al., 2021). This paper applies the LightGBM algorithm (Light Gradient Boosting Machine) for the online evaluation of small-signal stability in power systems with wind turbines. LightGBM is an algorithm based on gradient boosting decision trees. It can handle complex nonlinear data relationships and is particularly suitable for large-scale datasets and high-dimensional feature spaces.

LightGBM features high computational efficiency and low memory consumption. It also has the capabilities of automatic feature selection and missing value handling, while supporting parallel computing and distributed training. These advantages give it a distinct edge in big data processing and significantly reduce the upfront data preprocessing work. In this paper, LightGBM will be used to establish the mapping relationship between inputs and outputs, thereby realizing the estimation of the system's minimum damping ratio. Compared with traditional shallow learning models such as BP neural networks and BLS, LightGBM has significant advantages in handling high-dimensional features and large-scale data. Its decision tree splitting strategy based on histogram, Leaf wise growth method, GOSS and EFB techniques make it superior to the compared algorithms in terms of training speed, memory usage, and prediction accuracy, especially suitable for real-time stability assessment tasks in power systems.

To quickly adjust the minimum damping ratio of a power system back to the stable range once it exceeds this range, this paper proposes a power system stability adjustment method based on the LightGBM model. First, an optimization model for damping ratio sensitivity is established using the system's operating data. This model is then used to adjust the active power of generators, thereby improving the operating state of the power system. Finally, the effectiveness of this method will be verified through tests on different power system configurations.

## 2 LightGBM algorithm

LightGBM is an efficient gradient boosting decision tree method (GBDT) (Saner et al., 2021). Its principle is similar to GBDT, which takes the negative gradient of the loss function as the residual approximation value for fitting a new decision tree to the current decision tree (Yang et al., 2017). In each iteration, the original model is kept unchanged, and a new function is added to the model to make the predicted value continuously approach the true value.

The objective function for the  $K$ -th iteration is defined in Equation 1:

$$Obj^K = \sum_i L(y_i, \hat{y}_i^{K-1} + f_K(x_i)) + \Omega(f_K) + c^{K-1} \quad (1)$$

Among them,  $y_i$  is the actual value,  $\hat{y}_i^{K-1}$  is the result of  $K-1$  learning, and  $c^{K-1}$  is the sum of the regularization terms of the first  $K-1$  trees. The objective function is to find a suitable tree  $f_K$  to minimize the value of the function.

Expand the objective function using Taylor's formula, this is achieved through a second-order Taylor expansion (Equation 2):

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \frac{1}{2}f''(x)\Delta x^2 \quad (2)$$

The second-order Taylor expansion of the loss function is applying this expansion to the loss function gives Equation 3:

$$\begin{aligned} \sum_i L(y_i, \hat{y}_i^{K-1} + f_K(x_i)) &= \sum_i [L(y_i, \hat{y}_i^{K-1}) + L'(y_i, \hat{y}_i^{K-1})f_K(x_i) \\ &\quad + \frac{1}{2}L''(y_i, \hat{y}_i^{K-1})f_K^2(x_i)] \end{aligned} \quad (3)$$

Define  $g_i$  as the first derivative of the loss function, and  $h_i$  as the second derivative of the loss function as defined in Equations 4, 5.

$$g_i = L'(y_i, \hat{y}_i^{K-1}) \quad (4)$$

$$h_i = L''(y_i, \hat{y}_i^{K-1}) \quad (5)$$

The simplified objective function can be expressed as substituting these derivatives simplifies the objective to Equation 6:

$$Obj^K = \sum_i \left[ L(y_i, \hat{y}_i^{K-1}) + g_i f_K(x_i) + \frac{1}{2} h_i f_K^2(x_i) \right] + \Omega(f_K) + c \quad (6)$$

The traditional GBDT (Gradient Boosting Decision Tree) algorithm uses a pre-sorting method to find the optimal split points when constructing decision trees. This process requires traversing all data samples and calculating the information gain for all potential split points. LightGBM adopts an improved histogram-based algorithm (Bin et al., 2023), which partitions continuous floating-point feature values into  $k$  intervals. Since the optimal split points only need to be selected from these  $k$  intervals, this significantly enhances both training speed and space utilization.

Furthermore, LightGBM adopts a leaf-wise growth strategy (instead of the traditional level-wise strategy) when building decision trees, aiming to reduce the volume of training data. It also adds a maximum depth constraint, which ensures improved efficiency while preventing overfitting. To estimate information gain accurately with a smaller dataset, LightGBM uses Gradient-based One-Side Sampling (GOSS) (Zhang et al., 2020). This method retains instances with large gradients and performs random sampling

on instances with small gradients. In terms of feature reduction, LightGBM applies Exclusive Feature Bundling (EFB) (Meng et al., 2005) to achieve dimensionality reduction—this process does not result in information loss.

## 3 Small signal stability analysis model with doubly fed fan

### 3.1 Small stability analysis model

The differential algebraic equation describing the dynamic characteristics of the power system is:

$$\begin{cases} \frac{dx}{dt} = f(x, y) \\ 0 = g(x, y) \end{cases} \quad (7)$$

$f$  is a column vector composed of the differential equation  $f_{i \in [1, n]}$ ;  $g$  is also a column vector composed of the differential equation  $g_{i \in [1, m]}$ . In the formula,  $x$  represents the dynamic characteristics of the system, including the angular velocity of the generator rotor and the power angle of the generator.  $y$  represents the operating variables of the system, such as node voltage and phase angle. At the equilibrium point, the Lyapunov linearization method is used, and the formula is expressed as the linearized system equations are shown in Equation 8:

$$\begin{cases} \frac{d\Delta x}{dt} = \tilde{A}\Delta x + \tilde{B}\Delta y \\ 0 = \tilde{C}\Delta x + \tilde{D}\Delta y \end{cases} \quad (8)$$

Among them,  $\Delta x$  represents the change in state variables,  $\Delta y$  represents the change in non-state variables,  $A$  and  $C$  are coefficients of the change in state variables, and  $B$  and  $D$  are coefficients of the change in non-state variables. Eliminating vector  $\Delta y$  from Equation 7 yields the following form:

$$\frac{d\Delta x}{dt} = A\Delta x \quad (9)$$

The state matrix  $A$  is calculated as follows Equation 10:

$$A = \tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C} \quad (10)$$

The small signal stability problem of the system is transformed into the eigenvalue problem of calculating the state matrix. In classic small signal stability calculations, the indicators used to analyze small signal stability are the eigenvalue  $\lambda$  and damping ratio  $\xi$ , as defined in Equations 11, 12 respectively:

$$\lambda = \alpha + j\beta \quad (11)$$

$$\xi = -\frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} \quad (12)$$

Establish a mathematical model of the Xi system based on the above method, and obtain the corresponding linearized model. Transform it into a state matrix  $A$  through Equation 9, and study the small signal stability problem of the system by analyzing its eigenvalues and damping ratio.

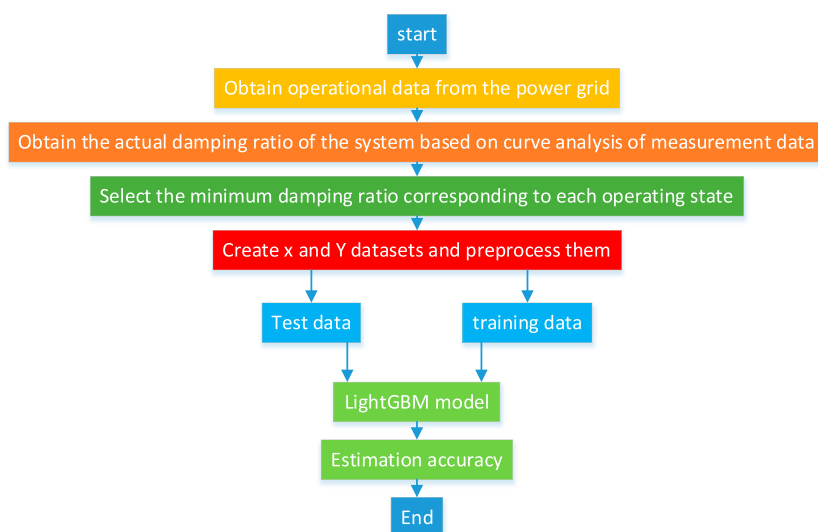


FIGURE 1  
Model training flow chart.

### 3.2 Double fed fan model

The stator side of a doubly-fed induction generator (DFIG) is directly connected to the power grid via a transformer, while its rotor side is connected to the grid through a converter. The dynamic model of the DFIG is divided into three parts: the mechanical part, the DFIG part, and the converter part. This paper assumes that each wind turbine operates under the same condition and simplifies the wind turbines into multiple parallel-connected generator units, which is equivalent to a large-scale wind farm. The parameters of the wind turbines and the converted data can be obtained from references (Huang et al., 2021; Qu et al., 2020).

## 4 Small interference stability evaluation model based on LightGBM algorithm

Traditional small-signal stability analysis requires linearizing the dynamic data in the system, calculating the state matrix, and finally obtaining the eigenvalues. For power systems that are constantly changing, traditional methods are time-consuming and cannot realize real-time monitoring of power system small-signal stability. Therefore, this paper proposes a real-time power system stability assessment method based on the LightGBM model. This method enables the monitoring speed to keep up with the changing speed of the power system and prevents the occurrence of faults. The corresponding flow chart is shown in Figure 1.

This paper uses the system's historical operating data to train the LightGBM model. The goal is to enable the model to accurately fit the relationship between the system's operating state and its stability, thereby realizing real-time analysis of power system stability and quickly identifying whether the system has small-signal stability issues. The system is adjusted to a stable operating state through a correction model.

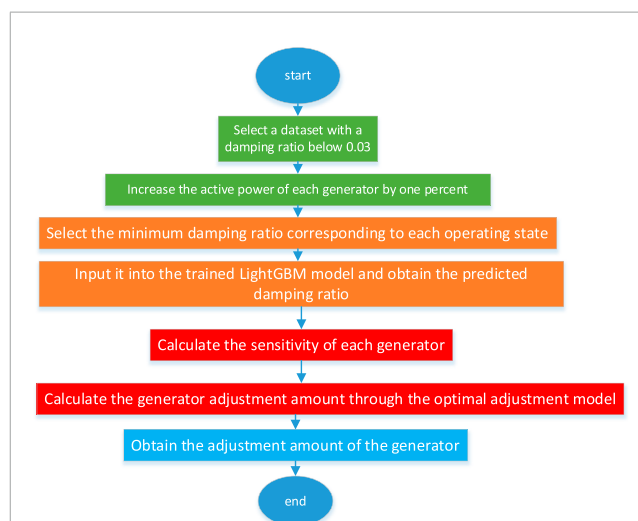


FIGURE 2  
Generator power adjustment flow chart.

### 4.1 Input and output

Currently, the Wide Area Measurement System (WAMS) has been widely applied in power systems (Tang et al., 2020) and has become a crucial monitoring and control platform for power systems (Ju et al., 2019). WAMS is a dynamic monitoring and control system for power systems based on synchrophasor technology. Its core component is the Phasor Measurement Unit (PMU), which can use GPS signals to perform high-precision synchronous measurements of voltage and current in power systems, and provide information on frequency, phase, and amplitude. Composed of synchrophasor measurement devices, phasor data concentrators, control centers, and high-speed data

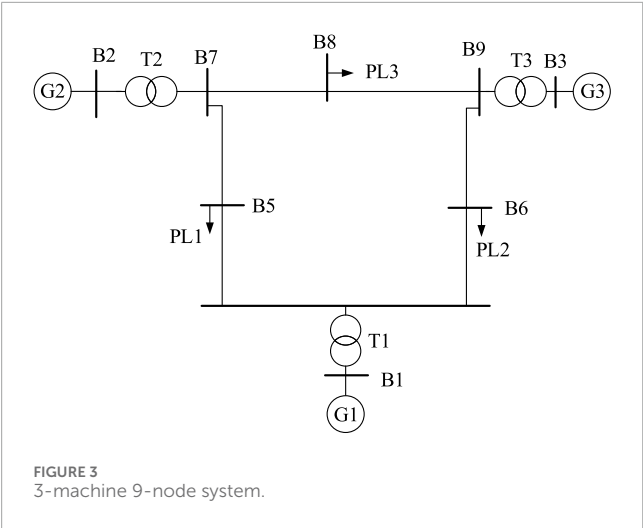


FIGURE 3  
3-machine 9-node system.

TABLE 1 Parameters of LightGBM model.

Parameter	Scope	Value retrieval
Min child weight	(0.001, 0.005)	0.003
Colsample bytree	(0.5, 1)	1
Max depth	(5, 15)	8
Reg lambda	(1e-05, 1)	0.1
Reg alpha	(1e-05, 1)	0.0001
Min child samples	(10, 200)	11
Learning rate	(0.001, 0.2)	0.1
Num leaves	(96, 256)	252
N estimators	(1,000, 5,000)	2,000
Max bin	(255, 512)	300

communication networks, WAMS enables real-time monitoring and analysis of power grid status by deploying a large number of sensors and communication equipment in power systems. In this study, data measured by WAMS from the power system is used as input. Through curve analysis of the actual measured data, the actual damping ratio is obtained. The minimum damping ratio of each operating state is selected as the output.

$\xi_{\min}$  should satisfy the following the inequality is Equation 13:

$$\xi_{\min} \geq \xi_{\text{limit}} \quad (13)$$

When the damping ratio is greater than  $\xi_{\text{limit}}$ , the system's operating state is stable, and  $\xi_{\text{limit}}$  is 0.03.

Take the power and minimum damping ratio measured from WAMS as input  $x$  and output  $y$ , respectively formulated as the input vector in Equation 14.

$$x_k = [PL \quad QL \quad PG \quad QG \quad P_{ij} \quad Q_{ij}], k = 1, 2, \dots, n \quad (14)$$

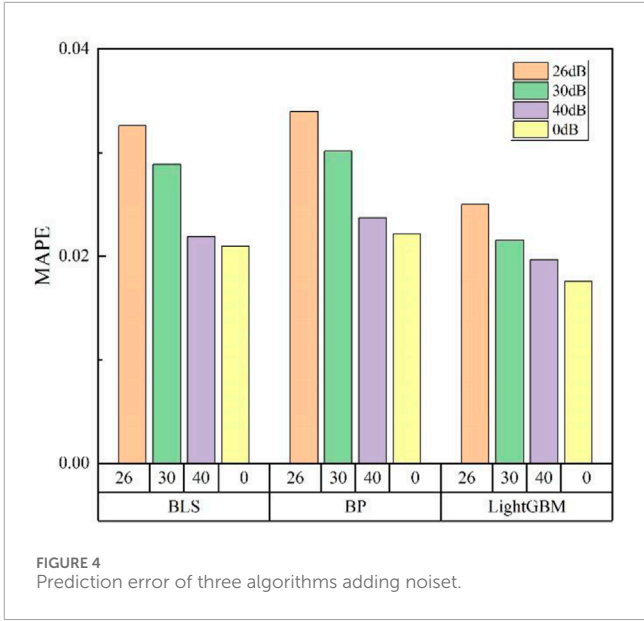


FIGURE 4  
Prediction error of three algorithms adding noiset.

TABLE 2 Predictions of the three algorithms under different noises.

Model	Noise	MAPE	RMSE
BLS	26 dB	0.03259852	0.00243456
	30 dB	0.02886344	0.00186234
	40 dB	0.02189644	0.00154578
	BLS	0.02095698	0.00104578
BP	26 dB	0.03396531	0.25698741
	30 dB	0.03015642	0.19856411
	40 dB	0.02369425	0.00159864
	0 dB	0.02215694	0.00112569
LightGBM	26 dB	0.02498634	0.00150236
	30 dB	0.02156956	0.00102355
	40 dB	0.01965823	0.00094568
	0 dB	0.01754568	0.00069045

The output is the minimum damping ratio, as shown in Equation 15:

$$y_k = [\xi_{\min}], k = 1, 2, \dots, n \quad (15)$$

PL and QL are the active and reactive power of the load. PG and QG are the active and reactive power of the generator.  $P_{ij}$  and  $Q_{ij}$  refer to the line power from node  $i$  to node  $j$ . It refers to the minimum damping ratio for each operating state.  $N$  refers to the total number of samples. In addition to the power of the generator and load, the input part of this article also considers the line power to adapt to changes in the topology of the power grid. Establishing the relationship between input variable  $X$  and output variable  $Y$



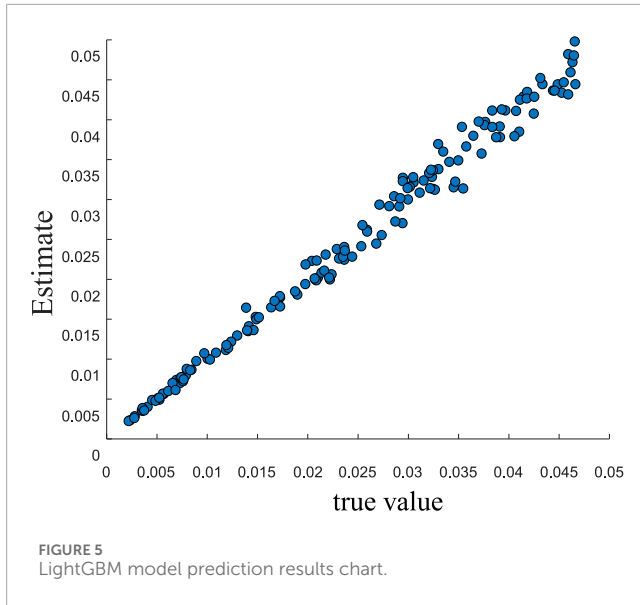


TABLE 3 Training results chart for the 3-machine 9-bus system.

Model	RMSE	MAPE	Time/s
BLS	0.00104578	0.02095698	0.068s
BP	0.00112369	0.02215694	0.091s
LightGBM	0.00069045	0.01754568	0.056s

through the LightGBM model, which are defined as the matrices in Equations 16, 17.

$$X = [x_1 x_2 \dots x_n]^T \quad (16)$$

$$Y = [y_1 y_2 \dots y_n]^T \quad (17)$$

X is the set established by (13), and Y is the set established by (14).

## 4.2 Data processing and model evaluation methods

The input data is normalized using the Z-score method (Equation 18):

$$x' = \frac{x - \mu}{\sigma} \quad (18)$$

Among them:  $x$  is the input value;  $x'$  is the value processed by normalization method;  $\mu$  and  $\sigma$  are the mean and standard deviation of  $x$ .

After normalization of the input data, the problem of “large numbers overwhelming small numbers” caused by excessive absolute differences between data can be effectively prevented. For model evaluation, the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are adopted. RMSE reflects the deviation degree and distribution of errors, while MAPE represents

the average value of relative errors. Their calculation methods are as follows, calculated using Equations 19, 20.

$$\text{MAPE} = \frac{100\%}{n} \sum_{k=1}^n \left| \frac{y_k - \hat{y}_k}{y_k} \right| \quad (19)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - \hat{y}_k)^2} \quad (20)$$

Among them,  $n$  refers to the number of datasets,  $y_k$  is the true value of the  $k$ th sample, and  $\hat{y}_k$  is the predicted value of the  $k$ th sample.

## 5 A small signal stability correction model based on LightGBM algorithm system

The damping ratio adjustment process is shown in Figure 2.

Select the damping ratio that is in an unstable operating state and correct it to a stable operating state with a damping ratio greater than 0.03 using the method mentioned above.

### 5.1 Calculation of damping ratio sensitivity

The calculation of damping ratio sensitivity can be determined by the output power  $\Delta p_m$  of the generator during normal operation and the corresponding damping ratio  $\Delta \xi_m$  to determine the sensitivity  $\Delta \xi + \Delta \xi_m$  of the generator relative to the damping ratio, as well as the output power change  $\Delta p + \Delta p_m$  corresponding to the change in damping ratio ( $\Delta p_m$  is a slight disturbance, ranging from 0.01 to 0.1 pu), as shown in Equation 21.

$$c_m = \frac{\Delta \xi_m}{\Delta p_m} \quad (21)$$

The total change in damping ratio is the sum of contributions from all generators (Equation 22):

$$\Delta \xi = C_1 \Delta p_1 + C_2 \Delta p_2 + C_3 \Delta p_3 + \dots + C_n \Delta p_n \quad (22)$$

$n$  is the total number of generators in the system;  $\Delta p_m$  is the active power adjustment corresponding to generator  $m$ ;  $\Delta \xi$  is the corresponding change in damping ratio of the system after adjusting the power generation of each unit (Wang et al., 2021).

### 5.2 Correction optimization model

In order to minimize the change in active power of each generator while achieving a damping ratio above  $\xi_{\text{limit}}$ , this section introduces an optimization model as follows this goal is formalized by the optimization model in Equation 23:

$$\min \sum_{i=1}^n \Delta P_i^2 \quad (23)$$

Subject to the following constraints (Equation 24):

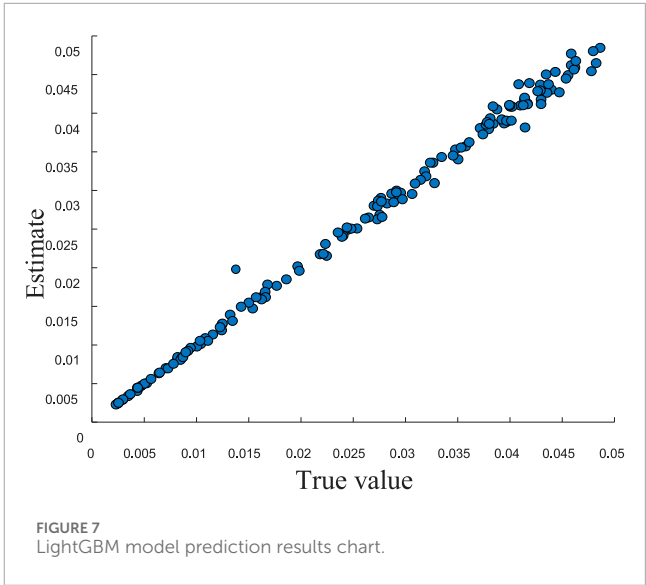
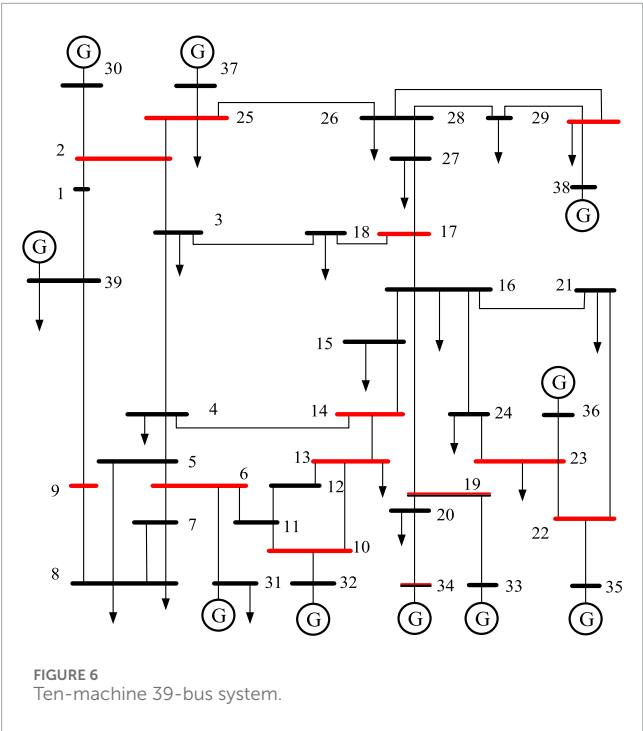
$$s.t. \begin{cases} P_i^{\min} \leq P_{i0} + \Delta P_i \leq P_i^{\max} \\ \xi_0 + \Delta \xi \geq \xi_{\text{limit}} \\ \sum_{i=1}^n C_i \Delta P_i = \Delta \xi \end{cases} \quad (24)$$

TABLE 4 Damping ratio sensitivity.

Generator	Actual damping ratio sensitivity	Predict damping ratio sensitivity
G1	-	-
G2	0.00104	0.00105
G3	0.00088	0.00085

TABLE 5 Data table of the generator active power adjustment process and results.

Generator	Initial active contribution	Active power adjustment estimated by LightGBM algorithm	Actual adjustment amount	Before adjustment $\xi$	Adjustment $\xi$	ACC
G2	179.03	20.30	20.58	0.006	0.037	0.73%
G3	80.13	11.32	11.08			



$\Delta P_i$  refers to the active power adjustment of generator  $i$ ,  $P_{i_0}$  is the initial active power of generator  $i$ ,  $P_i^{\min}$  and  $P_i^{\max}$  are the lower and upper limits of the active power adjustment of generator  $i$ , respectively.

TABLE 6 Training results chart for the 10-machine 39-bus system.

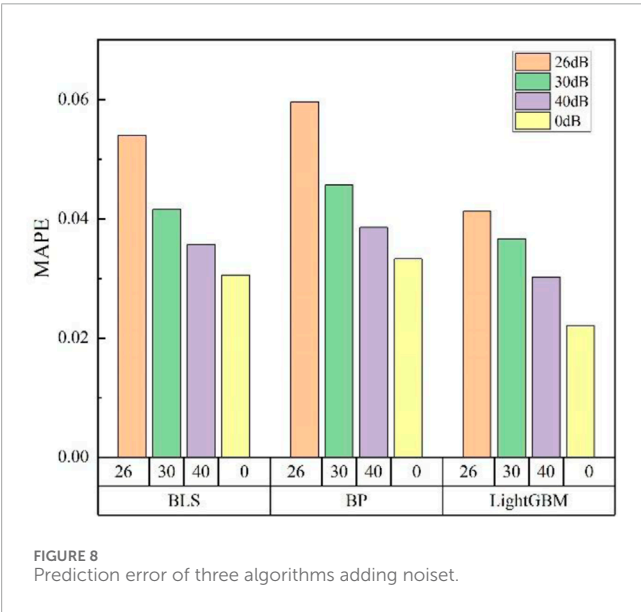
Model	RMSE	MAPE	Time/s
BLS	0.00112369	0.03056982	0.048
BP	0.00132564	0.03325691	0.110
LightGBM	0.00085491	0.02205163	0.032

## 6 Case study analysis

In this section, a 3-machine 9-bus system and a 10-machine 39-bus system are used for simulation verification. MATLAB is applied to simulate the actual operating states of the power systems. For each operating state, the active power and reactive power of generators, loads, and transmission lines, as well as the corresponding minimum damping ratio, are recorded. The collected data is used to train the

TABLE 7 Predictions of the three algorithms under different noises.

Model	Noise	MAPE	RMSE
BLS	26 dB	0.05398634	0.00290236
	30 dB	0.04156956	0.00192355
	40 dB	0.03569874	0.00167531
	0 dB	0.03056982	0.00112369
BP	26 dB	0.05963212	0.00312569
	30 dB	0.04563298	0.00213654
	40 dB	0.03856921	0.00189633
	0 dB	0.03325691	0.00132564
LightGBM	26 dB	0.04123568	0.00213695
	30 dB	0.03658974	0.00156982
	40 dB	0.03023561	0.00106531
	0 dB	0.02205163	0.00085491



LightGBM model. The dataset is divided into a training set and a test set in a ratio of 9:1.

This article uses random sampling to divide the training set and the testing set, with a ratio of 9:1. To ensure the model's generalization ability to extreme operating conditions, we introduced random fluctuations in load and generator power during the data generation process, covering various operating scenarios of the system under normal and abnormal conditions.

### 6.1 3-Machine 9-node system with wind turbines

It includes 3 generators, among which the No. 3 generator is replaced by 60 wind turbines, and these wind turbines adopt a seventh-order model. Generators G1 and G2 use a sixth-order model with their excitation systems considered. The loads employ a constant impedance model, and changes in control parameters are not taken into account. To simulate both normal and abnormal operating states that may occur in actual power system operations, the active power (P) and reactive power (Q) of the generators are randomly fluctuated by 20%, and the system load is fluctuated by approximately 20%–30%. MATLAB is used to calculate the eigenvalues for each operating state, and 10,000 sets of data are generated through random oscillations. Randomly select 1,000 data points from 10,000 as the test set, and model and train the remaining 9,000 data points. The 3-machine 9-bus system is shown in Figure 3.

Bayesian optimization was performed on the parameters of LightGBM, and the final parameters are shown in Table 1.

This article uses Bayesian optimization method to optimize the hyperparameters of LightGBM, and the selected parameters (as shown in Table 1) are determined based on their performance on the validation set. For example, num leaves controls model complexity, learning rate affects convergence speed, and max depth prevents overfitting. The combination of these parameters significantly improves training efficiency while ensuring prediction accuracy.

In practical power grids, the measurement error of data measurement units should be less than 1%. Therefore, the SNR is generally controlled above 30 dB (Nantian et al., 2021; Guanzheng et al., 2022). To verify the prediction ability of the proposed model under different noise conditions, Gaussian white noise of varying levels is added to the training data. MATLAB is used to generate Gaussian white noise with SNRs of 40 dB, 30 dB, and 26 dB, so as to evaluate the prediction accuracy of the three algorithms under interference. The prediction error graphs of three denoising algorithms are shown in Figure 4.

As shown in Figure 4, in the absence of noise, the MAPEs corresponding to the Generalized Learning System (BLS), Neural Network (BP), and LightGBM algorithms are 2.09%, 2.21%, and 1.75%, respectively. When there is a small fluctuation in the noise, there is a significant change in the relationship between the average error and the size of the noise. As shown in Table 2, under three different levels of noise, the average absolute error of the LightGBM algorithm is smaller than the prediction average error of BLS and BP.

Figure 5 shows the prediction performance of the LightGBM model.

Figure 5 shows the comparison between the predicted and actual values of the minimum damping ratio of the system using the LightGBM model. The high agreement between the predicted curve and the true value indicates that the model has good fitting ability and generalization performance.



TABLE 8 Damping ratio sensitivity.

Generator	Damping ratio sensitivity	Predict damping ratio sensitivity	Generator	Damping ratio sensitivity	Predict damping ratio sensitivity
G1	0.000208	0.000207	G2	0	−1.01e-05
G3	−9.029e-07	0.000109	G4	7.9350e-07	1.8935e-05
G5	0.002211	0.002394	G6	−0.002150	−0.001745
G7	−0.001802	−0.00229	G8	−0.002781	−0.003244
G9	0.002781	0.002584	G10	0.002875	0.003225

TABLE 9 Data table of the generator active power adjustment process and results.

Generator	Initial active contribution	Active power adjustment estimated by LightGBM algorithm	Actual adjustment amount	Before adjustment $\xi$	Adjustment $\xi$	ACC
G1	1000.04	0	0	0.008	0.041	0.89%
G2	537.80	20.35	20.35			
G3	660.98	0	−1.75			
G4	632	−0.19	9.12			
G5	467.50	0	0			
G6	623.54	61.50	61.48			
G7	593.50	−416.10	−279.02			
G8	533.71	−356.04	−366.50			
G9	828.76	−371.71	−371.70			
G10	226.48	1.24	1.24			

Table 3 shows the MAPE and training time of three algorithms. According to the data in the table, LightGBM outperforms BP and BLS in both training accuracy and speed.

ACC represents the percentage of adjustment error. The calculation formula is as shown in Equation 25:

$$ACC = \frac{|\xi_{true} - \xi_{LightGBM}|}{\xi_{true}} \quad (25)$$

$\xi_{true}$  refers to the actual damping ratio calculated after adjusting the active power of each generator, while  $\xi_{LightGBM}$  is the damping ratio obtained by incorporating the adjustment into the LightGBM model.

The damping ratio sensitivity of generators is calculated based on the data-driven model. The perturbation amount of generators is set to 10%, and the ACC is within an acceptable range. The results of the damping ratio sensitivity of each generator estimated based on the LightGBM model are shown in Table 4.

One set of data is randomly selected from the sample set with a damping ratio less than 0.03. The damping ratio adjustment is performed based on the calculated damping ratio sensitivity. Among them, Generator G1 serves as the system's balancing unit and does not require adjustment. The minimum damping ratios after adjustment via the LightGBM prediction model and the optimization model are shown in Table 5.

## 6.2 10-Machine 39-node system with wind turbines

The 10-machine 39-bus system is shown in Figure 5. It comprises 10 generators, among which the No. 4 generator connected to Bus 33 is replaced by a wind farm aggregated from 320 wind turbines.

The other generators adopt a sixth-order model, with their excitation systems taken into account. The loads use a constant impedance model, and changes in control parameters are not

considered. The wind turbines employ a seventh-order model. 10-machines with 39-nodes as shown in Figure 6.

As shown in Table 6, the RMSE, MAPE, and training time of the three algorithms were compared in a 10 machine 39 node model.

Figure 7 shows the prediction performance of the LightGBM model.

The predicted curve in Figure 7 is highly consistent with the true values, indicating that the model has good fitting ability and generalization performance in 10 machines and 39 nodes.

Gaussian white noise with signal-to-noise ratios (SNR) of 40 dB, 30 dB, and 26 dB is added to simulate the errors encountered when obtaining the minimum damping ratio in power systems. The Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are used as indicators to evaluate the prediction accuracy of the models. The prediction accuracy of the three algorithms under different noise levels is assessed, and the prediction results are shown in Table 7.

In the absence of noise interference, the Mean Absolute Percentage Errors (MAPE) of the BLS, BP, and LightGBM algorithms are 3.05%, 3.32%, and 2.21%, respectively. Their Root Mean Square Errors (RMSE) are  $9.85 \times 10^{-4}$ ,  $11.23 \times 10^{-4}$ ,  $13.25 \times 10^{-4}$ , and  $8.54 \times 10^{-4}$  (note: there may be a typo in the original RMSE count, as it lists 4 values for 3 algorithms).

After adding Gaussian white noise with signal-to-noise ratios (SNR) of 40 dB, 30 dB, and 26 dB, the MAPE and RMSE of the four models (consistent with the RMSE count) all increase. Among them, the LightGBM model shows the most stable overall variation, as illustrated in Figure 8.

Table 8 shows the true damping ratio sensitivity of each generator and the predicted damping ratio sensitivity using LightGBM algorithm.

As shown in Table 9, the minimum damping ratio adjusted by the LightGBM prediction model and optimization model is 1.0554% for ACC.

## 7 Conclusion and future work

This paper evaluates the small-signal stability of power systems with wind power integration based on the LightGBM model. It also constructs an adjustment model by combining optimization algorithms to correct the power system from weakly damped operating states to stable ones. Tests are conducted on the 3-machine 9-bus system and the 10-machine 39-bus system. The simulation results show that the proposed method has high accuracy and fast adjustment response, while also verifying its feasibility in power system stability assessment and adjustment.

LightGBM features efficient computing performance and low memory usage. It supports automatic feature selection, missing value handling, parallel computing, and distributed training—these advantages enable it to exhibit significant strengths in processing large-scale data and significantly reduce preliminary data preprocessing work. Future research will focus on the following directions: combining LightGBM with physical models to construct more interpretable hybrid learning frameworks; Expand

small disturbance stability analysis to multiple scenarios and time scales.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

YF: Writing – original draft. WG: Writing – review and editing. XY: Investigation, Writing – review and editing. MQ: Software, Writing – review and editing. AS: Data curation, Writing – review and editing. CJ: Methodology, Writing – review and editing.

## Funding

The authors declare that no financial support was received for the research and/or publication of this article.

## Acknowledgements

The authors would like to express their sincere gratitude to all the individuals and organizations that have supported this research.

## Conflict of interest

Authors YF, WG, XY, MQ, AS, and CJ were employed by Guizhou Power Grid Co., Ltd. Power Dispatch Control Center.

## Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Bin, F., Yijie, H., and Gang, H. (2023). A review of novel dispatching and scheduling optimization methods for power systems based on deep reinforcement learning. *Automation Electr. Power Syst.* 47 (17), 187–199.
- Gibbard, M. J., Pourbeik, P., and Vowles, D. J. (2015). Small-signal stability, control and dynamic performance of power systems.
- Guanzheng, L., Bin, L., and Wang, S. (2022). Dynamic frequency prediction of power system after disturbance based on feature selection and random forest. *Grid Technol.* 1–12. doi:10.13335/j.1000-3673.pst.2021.0027
- Huang, M. Z., Hu, Y. H., and Weng, Y. F. (2021). A transient stability assessment method for power systems incorporating jmm and ngboost. *Autom. Electr. Power Syst.* 45 (08), 155–165.
- Jia, H., and Yixin, Y. (2001). Chaos phenomena and related research in power systems. *Chin. J. Electr. Eng.* (07), 27–31.
- Jiang, H., Bai, Y., and Wang, S. (2021). Analysis method on the impact of wind power integration on small signal stability of power systems. *Proc. CSEE* 33 (09), 9–16. doi:10.19635/j.cnki.csu-epsa.000706
- Ju, Y., Sun, G., Chen, Q., Zhang, M., Zhu, H., and Rehman, M. U. (2019). A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting. *IEEE Access* 7, 28309–28318. doi:10.1109/access.2019.2901920
- Ke, G., Meng, Q., Finley, T., and Chen, W. (2017). Lightgbm: a highly efficient gradient boosting decision tree. *Proceeding Adv. Inf. Process. Syst.*, 3146–3154.
- Meng, J., Ming, W. M., and Zhang, L. (2005). Electromagnetic interference spectrum estimation for power converters based on modeling of igt switching transient processes estimation of electromagnetic interference spectrum. *CSEE J. Power Energy Syst.* 37 (12), 14482–14498.
- Nantian, H., Wenguang, Z., and Guowei, C. (2021). Efficient identification of LightGBM power quality disturbances constrained by IoT data transmission rate. *Chin. J. Electr. Eng.* 41 (15), 5189–5201.
- Qu, Z., Li, H., Wang, Y., Zhang, J., Abu-Siada, A., and Yao, Y. (2020). Detection of electricity theft behavior based on improved synthetic minority oversampling technique and random forest classifier. *Energies* 13 (8), 2039. doi:10.3390/en13082039
- Saner, C. B., Yaslan Genc, Y. I., and Genc, I. (2021). An ensemble model for wide area measurement-based transient stability assessment in power systems. *Electr. Eng.* 103, 2855–2869. doi:10.1007/s00202-021-01281-x
- Tang, C., Luktarhan, N., and Zhao, Y. (2020). An efficient intrusion detection method based on LightGBM and autoencoder. *Symmetry* 12 (9), 1458. doi:10.3390/sym12091458
- Wang, Y., Chen, J., Chen, X., Zeng, X., Kong, Y., Sun, S., et al. (2021). Short-term load forecasting for industrial customers based on TCN-LightGBM. *IEEE Trans. Power Syst.* 36 (3), 1984–1997. doi:10.1109/tpwrs.2020.3028133
- Yanfeng, M. A., Zheng, L., and Huo, Y. (2020). Analysis on damping torque of virtual synchronous generators connected to power systems. *Electr. Power Autom. Equip.* 40 (04), 166–171. doi:10.16081/j.epae.202003027
- Yang, Y., Yu, L., and Yuan, H. (2017). Time domain calculation and waveform analysis of small signal stability in power systems. *Adv. Technol. Electr. Eng. Energy* 36 (01), 44–51. doi:10.13334/j.0258-8013.pcsee.2001.07.007
- Yin, B., Chen, Q., Li, B., and Zuo, L. (2021). A new method for identification and classification of power quality disturbance based on modified kaiser window fast s-transform and lightgbm. *Proceed CSEE* 41 (24), 8372–8384.
- Zeng, D., and Jun, Y. (2020). Frequency small-signal stability analysis of multi-machine parallel systems with virtual synchronous generators. *Proc. CSEE* 40 (07), 2048–2061. doi:10.13334/j.0258-8013.pcsee.191217
- Zhang, R., Wu, J., and Baoqin, L. (2020). Adaptive prediction of transient stability in power systems based on transfer learning. *Grid Technol.* 44 (06), 2196–2205. doi:10.13335/j.1000-3673.pst.2019.2376
- Zhou, T., and Yang, J. (2016). Power system transient stability assessment method based on modified lightgbm. *Power Syst. Technol.* 43 (6), 1931–1940.