EAI Endorsed Transactions
on Energy Web

Research Article EAI.EU

Photovoltaic power generation prediction and optimization configuration model based on GPR and improved PSO algorithm

Zhennan Zhang, Zhenliang Duan, Lingwei Zhang*

Wuwei Power Supply Company, State Grid Gansu Electric Power Company, Wuwei 733000, China

Abstract

As the growing demand for energy as well as the strengthening of environmental awareness, photovoltaic power generation, as a clean and renewable energy source, has gradually attracted people's attention and attention. To facilitate the dispatching and planning of power system, this study uses historical data and meteorological data to build a photovoltaic power generation prediction and configuration optimization model on the ground of Gaussian process regression and improved particle swarm optimization algorithm. The simulation results show that the regression prediction curve of the Gaussian process regression prediction model is the closest to the real curve, and the prediction curve is stable and not easily disturbed by noise data. The Root-mean-square deviation and the average absolute proportional error of the model are small, and the disparity in the predicted value and the true value of the model is small; The integration of multi factor data has improved the accuracy of prediction data, and the regression prediction effect is good. The improved Particle swarm optimization algorithm could continuously enhance in the search for the optimal solution, and the Rate of convergence is fast. The Pareto solution can provide different solutions suitable for photovoltaic power generation optimization. Reasonable optimization configuration can effectively reduce active power line loss and voltage deviation, with the maximum reduction values reaching 132kW and 0.028, respectively. The research and design of predictive models and optimized configuration models can promote the formation of smart grids.

Keywords: Photovoltaics; Power Prediction; Multi Objective Optimization; PSO Algorithm; Gaussian Process; Optimize Configuration

Received on 30 August 2023, accepted on 28 January 2024, published on 20 February 2024

doi: 10.4108/ew.3809

*Corresponding author. Email: Lingwei_Zhang@outlook.com

1. Introduction

As the growing absence of the traditional resource, the advancement of new energy (NE) has become one of the hot topics of global concern [1]. Photovoltaic power (PPO) generation is a renewable energy generation method that utilizes photovoltaic cells to convert light energy into electricity, and is mainly composed of photovoltaic modules, inverters, racking and distribution systems. The solar photovoltaic industry occupies an important strategic position in the NE industry, and the energy structure occupies a relatively essential proportion. The extensive development of traditional resources has made environmental issues increasingly prominent; PPO generation belongs to renewable energy and is an essential way for diminishing greenhouse gas emissions in response to climate change; It possesses essential strategic meaning [2-3]. China's energy industry structure is in the transition period, and wind energy, biomass energy, solar energy and Tidal power are the key points of energy utilization in the future. PPO stations are in line with the strategy and direction of China's energy industry advancement, meeting the goals of energy conservation, emission reduction, and sustainable development, and are conducive to adjusting the existing power grid (PG) structure in China. However, the volatility, intermittency, and randomness of PPO generation are more obvious, and large-scale integration of PPO into the PG will threaten the operation of the PG. To mitigate the impact of these factors on the PPO grid, it is essential for forecasting the power generation of PPO [4]. PPO generation prediction is a short-term prediction of PPO generation power on the ground of real-time operating data of PPO generation systems, as well as solar radiation, temperature, and weather historical data monitored by meteorological stations, to reasonably arrange the operation mode of the power generation system and prepare response measures in advance [5]. PPO power generation forecasting plays an important role in the design of PPO power plants, power system operation, energy market trading and optimization of energy storage systems, which can improve the efficiency and economy of PV power generation systems and promote the sustainable development of renewable energy. Consequently, the prediction model of PPO generation needs to have high technical requirements to meet the continuous and secure operating of the PPO prediction (PP) system. To meet the demand of high technology, the research establishes the PPO generation prediction model on the ground of Gaussian process Regression (GPR), and then improves it by introducing adaptive constraints and genetic variation operation in Particle Swarm Optimization (PSO) algorithm and designs a reasonable allocation of PV power generation model. The research mainly contains two innovations, one GPR is combined with the characteristics of PV power generation system, this integrated model can predict the PV power more accurately and improve the reliability and stability of PV power generation system. Secondly, for the traditional PSO algorithm which is easy to fall into the local optimum problem, the global search capability is improved, which effectively solves the optimization problem of multi-objective, nonlinear and constraints of PV power system configuration. The contribution of the research mainly lies in enriching the theoretical study of PV power forecasting and providing new ideas and methods for the technical optimization of PSO algorithms. The model proposed in the study shows significant improvement in both the prediction accuracy of PV power and the optimization performance of system configuration.

The overall research consists of four. The first is a review of research on domestic PPO generation prediction models, which summarizes the research progress, technical novelties and shortcomings of PV power prediction models. The second proposes a PPO generation PP model and configuration optimization model on the ground of GPR and improved PSO, and describes the technical points and research ideas in the process of model design; The third conducted simulation testing experiments on the algorithm performance; The fourth summarizes the research results and describes the conclusions, significance and shortcomings of the study, and the model helps to solve the optimization problem of optimizing photovoltaic access and make sure the secure operating.

2. Related works

As an essential way of solar energy utilization, PPO generation possesses an essential influence on alleviating the Energy crisis and solving environmental problems; Lots of studies have been carried around the prediction of PPO generation. For enhancing the scheduling quality of PPO generation and maintain the response speed of PPO stations, Chen B et al. considered the influence of different meteorological factors at different time periods, designed a radiation classification coordinate method on the ground of the characteristics of radiation coordinates and power generation time series, and constructed a model training dataset. On the ground of the Long short-term memory (LSTM) recurrent neural network (RNN), the ultra-short-term prediction model of PPO generation was constructed and tested on the independent photovoltaic system of the Australian Desert Knowledge Solar Center; The results show that the robustness and accuracy of the radial classification coordinate LSTM RNN model are significantly better than other models [6]. Malik et al. used a new solar cell model to design a power generation prediction model for PPO plants; Verified by the real automatic data monitoring system, the three-parameter model is more suitable for the low solar Irradiance range, and the four-parameter model is more suitable for the medium and high solar Irradiance range. The prediction model results show that after 4 years of use, the power of PPO plants will decrease due to component degradation. The prediction model designed in the study has a higher accuracy [7]. Kim et al. designed a very short-term PPO generation prediction model on the ground of the LSTM RNN. The LSTM RNN is more sensitive to the characteristics of continuous time series data and can predict the power of ultra short-term PPO generation; The model includes two LSTM modules with different scales and uses irregular factors to affect the prediction performance. The experimental results indicate that the model can stably predict PPO generation [8]. Because the meteorological information is in a constantly changing state, the forecast of PPO generation in the peak area has become a thorny problem. Lee D et al. designed a prediction model of PPO generation power output using the LSTM network and the gated cycle unit; The model only uses meteorological information observed during the morning period as input for the data. The test outcomes of real datasets showcase that the model outperforms traditional models in predicting PPO in peak areas [9].

Aprillia et al. designed a meteorological forecasting model for PPO generation combining Convolutional neural network and Thalassia group algorithm, established different Convolutional neural network forecasting models for different weather types, and Thalassia group algorithm was utilized for adjusting the parameters of the forecasting model. Compared with the prediction model built by support vector machine and LSTM neural network, the meteorological prediction model built by the study is more suitable for the actual PPO generation model [10]. To achieve real-time scheduling of the PG, Yan et al. designed an ultra-short-term photovoltaic PP model on the ground of optimal frequency domain decomposition and deep learning; This model collects the actual photovoltaic data of rainstorm days as the test set. Compared with the Discrete wavelet transform, variable modulus decomposition and direct prediction model, the Mean absolute error of the model designed in the study is reduced by 52.97%, 64.07% and 31.21% respectively; Compared with RNN and LSTM model, mean absolute error decreased by 23.64% and 46.22% respectively, and training efficiency increased by 85.63% and 87.68% [11]. AlShafeey et al. compared the performance of multiple regression networks and artificial neural networks (ANN) in PPO generation prediction. Six different PPO generation prediction models were constructed on the ground of a combination of meteorological datasets and power historical datasets; This model predicts the 24-hour PPO generation. The performance test results show that the forecast of ANN always outperforms that of multiple regression networks. The use of mixed datasets can improve the forecast, while poorer datasets can affect the forecast [12]. Hao et al. proposed a microgrid coordination method on the ground of PPO generation PP. The PP model combines clustering algorithms and neural networks and utilizes the PPO generation output and electric vehicle charging load of the PP model for enhancing the genetic algorithm, thereby determining the equilibrium of a single leader multi follower Stackelberg game; The good performance of this coordination method has been verified through real data from PPO stations and electric vehicle charging stations [13].

Accurate prediction of photovoltaic (PV) power demand for renewable energy 5G base stations helps to reduce energy consumption. In order to cope with the power generation prediction under complex weather conditions, Yan M et al. improved the local optimal solution avoidance ability of the PV short-term prediction model based on a 5G base station with energy router by introducing an improved logistic distribution, Laplace distribution with inverse incomplete gamma function weight factor, and nonlinear mutation uptake strategy based on the sparrow search algorithm. The measured data verified the prediction accuracy of the method [14]. The size of the PV system has a large impact on the PV power generation, and the accuracy of a single model for PV power generation prediction is poor. Song H et al. considered PV power generation prediction as a multi-task prediction and proposed a multi-task recurrent neural network framework. Compared with deep neural network model, this method is superior in prediction accuracy [15]. Photovoltaic power generation to a certain extent crowds out the functional space of traditional synchronous generators and reduces the inertia of the grid itself. Facing the instability of the inertia support capacity of PV output power, Wang S et al. proposed an FM control strategy based on the dynamic characteristics of grid-side DC capacitance, and correlation analysis of different weather type data to predict PV power. Simulation is performed to analyze the system's ability to respond quickly to load changes under different control parameters. Experimental results show that the method can effectively improve the stability of the distribution network [16]. In order to reduce the impact of distributed PV power fluctuations on grid operation, Yang X et al. designed a distributed PV power plant prediction model with spatio-temporal information based on the time dependence of power series and the spatial correlation of meteorological data. The joint genetic algorithm and natural gradient-enhanced long and short-term memory neural network were used for time-varying prediction, and spatial correlation prediction was carried out by κ correlation coefficient and 2D convolutional neural network. The advanced research methodology was finally verified by utilizing the PV cluster data in Hebei Province, China [17]. Aiming at the insufficient generalization ability of the PV power prediction model, Chen G et al. proposed a time-dependent PV power prediction model combining variational mode decomposition and Bayesian regularized neural network by using mutual information analysis to select suitable meteorological sequences. The experimental results verified the accuracy of the method, and the prediction error was significantly reduced [18].

In summary, there has been some progress in the prediction model of PPO generation. However, most of the PV power prediction studies have focused on a single PV power plant without considering the impact of complex scenarios on power; at the same time, PV power is affected by several factors, including solar radiation, temperature, wind speed, etc. Existing prediction models mainly focus on the impact of a certain or a few factors, and there is still a research gap in prediction models for the combined effect of multiple factors. And research on simultaneously considering PPO generation prediction and optimizing PG configuration is still quite rare; This type of research achievement is meaningful for enhancing the stability of power generation and supply systems. Therefore, the study designed a PPO generation PP and configuration optimization model on the ground of GPR and improved PSO.

3. Design of photovoltaic power generation prediction and optimization configuration model on the ground of GPR and PSO algorithms

PPO generation is converting solar energy into electricity, which is an essential part of the energy supply process; To ensure the stability of PG supply, it is crucial for solving the issue of forecasting PPO generation. The size of solar radiation, temperature and weather change will affect the power generation. For this, the research designs the prediction model and optimal grid configuration model of PPO generation on the ground of GPR.

3.1 Design of photovoltaic power prediction model on the ground of Gaussian process regression

GPR is a machine learning (ML) method on the ground of Bayesian theory and other relevant theories; This method uses training samples to train the model, selects the optimal model, and tests it, which is the process of ML to complete the prediction, as shown in Figure 1. GPR belongs to a Stochastic process. Each Finite set is subject to Multivariate normal distribution. The distribution of Gaussian process is the joint probability distribution (PB) of all random variables. The core idea of using GPR for forecasting is to fit the isolated things into a law. In the fitting process of nonlinear laws, the kernel function (KF) plays an important role [19].

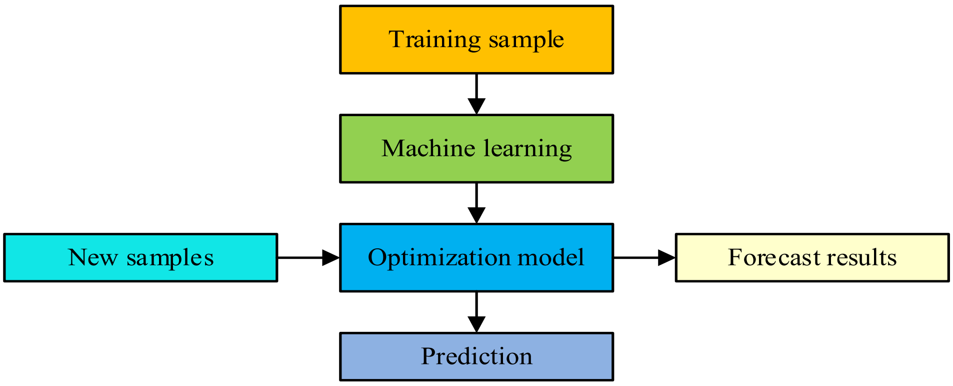

Figure 1. Machine Learning Prediction Method

The essence of GPR is the learning of KF with probability significance; First, it assumes that the sample obeys the Prior probability of Gaussian process; Then it uses Bayesian theory to obtain the Conditional probability of the prediction point and complete the calculation of Posterior probability; Finally, the final prediction model is derived from the calculation of super parameters. The relevant details of GPR prediction model is showcased in Figure 2. In Figure 2, the blue points are known, the red and green curves indicate the curves connecting the known data, where red is the one with the highest likelihood, and the yellow part indicates the range of two standard deviations of the Gaussian distribution, which indicates that there are countless possibilities for the composition of the line segments in this region. The nonparametric, flexible, and scalable advantages of GPR make it widely applicable in fields such as finance, healthcare, and weather forecasting [20].

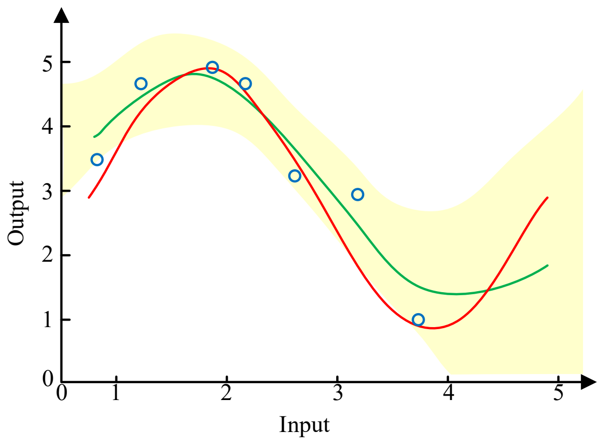

Figure 2. Schematic Diagram of Gaussian process Regression Prediction Model

If there is a linear relation in the input and output of the model, the linear regression prediction model equation of Bayesian theory is shown in equation (1). In equation (1), $f(x)$ serves as the function value; x denotes the input data; y serves as the observed value; w serves as the weight vector (WV) of the linear model; ε represents noise, $\varepsilon \sim N(0,\sigma_n^2)$. The GPR model models the noise in the input data by introducing a noise model. A Gaussian distribution is utilized to describe the noise and it is taken into account in the training process of the model. This operation allows better adaptation to data with noise.

$$y = f(x) + \varepsilon = x^T w + \varepsilon \qquad (1)$$

According to the prior distribution, the Likelihood function of the observation sample can be obtained. On the ground of Bayesian theory and Gaussian distribution (GD), the PB of the WV can be calculated by using the maximum Posterior probability. The Posterior probability PB of the final WV is showcased in Formula (2). In equation (2), $X = [x_1, x_2, \ldots, x_n]^T$ and $y = [y_1, y_2, \ldots, y_n]^T$ serve as the observed samples.

$$\begin{aligned} p(w|X,y) &\propto \exp\left(-\frac{1}{2\sigma_n^2}(y - X^T w)(y - X^T w)\right)\exp\left(-\frac{1}{2}w^T \Sigma_p^{-1} w\right) \\ &\propto \exp -\frac{1}{2}(w-\bar{w})^T\left(\frac{1}{\sigma_n^2}XX^T + \Sigma_p^{-1} w\right)(w-\bar{w}) \end{aligned} \qquad (2)$$

If $\overline{w}$ satisfies $\overline{w} = \sigma_n^{-2}\left(\sigma_n^{-2}XX^T + \sum_p^{-1}\right)^{-1} Xy$ and it lets $A = \sigma_n^{-2}XX^T + \sum_p^{-1}$, it can be inferred that the Posterior probability of the WV satisfies the GD $p(w|X,y) \sim N\left(\overline{w}, A^{-1}\right)$, also known as the Maximum a posteriori estimation estimate. The PB of the predicted value y_* is showcased in equation (3), where x_* represents the data input.

$$p(y_*|x_*,X,y) = \int p(y_*|x_*,w)\,p(w|X,y)dw = N\left(\frac{1}{\sigma_n^2}x_*^T A^{-1}Xy, x_*^T A^{-1}x_*\right) \quad (3)$$

The linear derivation of Gaussian process is mapped to high-dimensional space by function $\phi(x)$, and the model expression and prediction value after nonlinear transformation are obtained; According to the derivation of Bayesian theory, the PB expression of predicted values after mapping a nonlinear model to a high-dimensional space is shown in equation (4).

$$p(y_*|x_*,X,y) = \int p(y_*|x_*,w)\,p(w|X,y)dw = N\left(\frac{1}{\sigma_n^2}\phi(x_*)^T A^{-1}\phi(X)y\chi, \phi(x_*)^T A^{-1}\phi(x_*)\right) \quad (4)$$

After matrix transformation, it lets $K = \phi(X)^T \sum_p \phi(X)$, $k(x_*,X) = \phi(x_*)^T \sum_p \phi(X) = k_*$, $k(x_*,x_*) = \phi(x_*)^T \sum_p \phi(x_*) = k_{**}$ as well as derives from equation (4) to obtain equation (5); In equation (5), $k_*\left(K+\sigma_n^2 I\right)y$ represents the predicted value and $k_{**}\left(K+\sigma_n^2 I\right)k_*^T$ represents the variance.

$$p(y_*|x_*,X,y) = N\left(k_*\left(K+\sigma_n^2 I\right)y, k_{**} - k_*\left(K+\sigma_n^2 I\right)^{-1}k_*^T\right) \quad (5)$$

The definition formula of Gaussian process mean function and Covariance function satisfying Bayesian linear regression is shown in Formula (6), where $f(x)$ represents Bayesian linear regression function, $f(x) = \phi(x)^T w$.

$$\begin{cases} E[f(x)] = \phi(x)E(w) = 0 \\ E[f(x)f(x')] = \phi(x)^T E(ww^T)\phi(x') = \phi(x)^T \sum_p \phi(x') \end{cases} \quad (6)$$

For an input sample set without noise, the conditional posterior distribution of the predicted values is obtained as shown in equation (7). Among them, $K(x_*,X) = \phi(x_*)^T \sum_p \phi(x)$, $K(X_*,X_*) = \phi(x_*)^T \sum_p \phi(x_*)$, and $K(X,x_*) = K(x_*,X)^T$.

$$y_*|x_*,X,y \sim N\begin{pmatrix} K(x_*,X)K(X,X)^{-1}y, K(X_*,X_*) \\ -K(x_*X)K(X,X)^{-1}K(X,x_*) \end{pmatrix} \quad (7)$$

When there is noise, the definition of the regression model is $f(x) = \phi(x)^T w + \varepsilon$, and the posterior distribution probability of the output predicted value is obtained, as shown in equation (8).

$$\begin{cases} y_*|x_*,X,y \sim N\left(\overline{y}, \mathrm{cov}(y_*)\right) \\ \mathrm{cov}(y) = K(X,X) + \sigma_n^2 I \end{cases} \quad (8)$$

KF are used to measure the degree of similarity between different samples; In GPR, the main contribution of KF is to complete the transformation from nonlinear relationship to linear relationship. The Covariance function is the KF of GPR, and the KF affects the final established GPR model and the distribution of predicted values. In this research, the square index Covariance function is selected as the KF, and the definition formula is shown in Formula (9), where σ_f^2 represents the signal square difference over parameter; $M = diag(\lambda_1,\lambda_2,...,\lambda_d)$, where λ represents the Characteristic length scale; σ_n^2 represents the hyperparameter of noise variance; δ_{ij} represents the Kroneck function.

$$k_{SE} = (x_i,x_j) = \sigma_f^2 \exp\left(-\frac{(x_i-x_j)^T M^{-2}(x_i-x_j)}{2}\right) + \sigma_n^2\delta_{ij} \quad (9)$$

In the design of PPO generation prediction model, it first divides the influencing factor subspace used, establishes the Gaussian process prediction model for each subspace, and integrates to obtain the final model; The prediction process of GPR is shown in Figure 3.

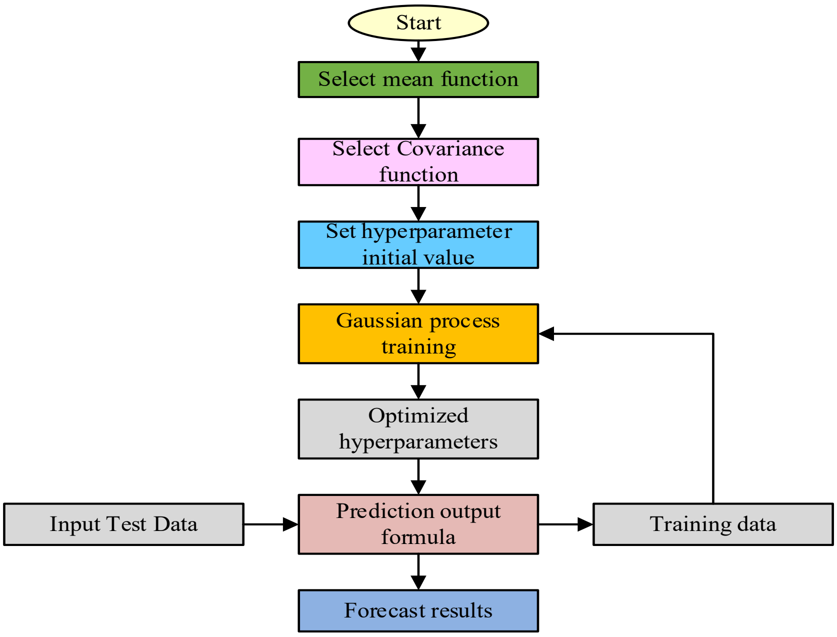

Figure 3. Prediction flow chart of Gaussian process regression

The super parameters in the KF determine the fitting effect of the nonlinear system and the prediction effect of the model; Improper selection of super parameters will lead to the invalidity of GPR, so the selection of super parameters is very important. In this study, the maximum likelihood estimation method is utilized for addressing the super parameter of GPR. The constructed Likelihood function is shown in Formula (10), and θ represents the super parameter in the KF.

$$L=\log p\left(y|\theta\right)=-\frac{1}{2}\log|K|-\frac{1}{2}y^{T}K^{-1}y-\frac{n}{2}\log 2\pi \quad (10)$$

It uses the Conjugate gradient method to solve the super parameters after initialization, as showcased in equation (11). In equation (11), $trace$ serves as the trace of the matrix.

$$\frac{\partial L}{\partial \theta_i}=\frac{1}{2}trace\left(K^{-1}\frac{\partial K}{\partial \theta_i}\right)-\frac{1}{2}y^{T}K^{-1}\frac{\partial K}{\partial \theta_i}K^{-1}y \quad (11)$$

The conversion of solar energy to electrical energy relies on PV modules to receive the sun's energy for energy conversion. Solar radiation refers to the energy of the sunlight irradiated on PV panels, and is affected by factors such as geographic location, time of day, and weather. In addition, PV cells decrease with the increase in temperature during the power generation process; humidity is related to dust, dirt and other pollutants on the surface of PV panels, and pollutants reduce the light transmission of PV panels, reducing the energy of the sunlight reaching the PV cells and lowering the efficiency of PV power generation. And the appropriate wind speed can help dissipate heat, maintain the reasonable temperature of the PV cell, and improve the power generation efficiency of the PV cell. For the GPR PPO generation prediction model, this study considers all the influencing factors. After determining the set of influencing factors, select the number of influencing factors as needed for subspace segmentation to obtain the combination and arrangement of different influencing factors; Each arrangement method is divided into subspaces of influencing factors.

3.2 Design of photovoltaic power generation system optimization configuration model on the ground of improved particle swarm optimization algorithm

When an independent PPO generation system is connected to the distribution system in a distributed generation system, the voltage distribution of the distribution system will be affected, and the circuit lines will also be damaged. The reasonable configuration of PPO generation in the PG system is crucial for ensuring the secure and reliable operation.

The total cost of PPO generation includes the initial investment cost and maintenance cost during use. The calculation process of cost and veterinary medicine is shown in equation (12); In equation (12), C_{pv}, C_{sr}, and C_{ee} respectively represent the total cost, subsidy fees, and benefits of PPO generation equipment; α_{pv} represents the purchase amount of photovoltaic equipment; r represents the discount rate; m_{pv} represents the service life of the photovoltaic equipment, and P_{pV} represents the purchased power of the photovoltaic equipment; u_{pv} represents the maintenance cost per unit power, and C_g represents the amount of electricity subsidy; C_{si} represents the grid electricity price; C_c represents the capacity coefficient.

$$\begin{cases} C_{pv} = \alpha_{pv} P_{pV} \dfrac{r(1+r)^{m_{pv}}}{r(1+r)^{m_{pv}} - 1} + u_{pv} P_{pV} \\ C_{sr} = 1289\left(C_g + C_{si}\right) P_{pV} C_c \\ C_{ee} = C_{pv} - C_{sr} \end{cases} \tag{12}$$

The current or power flows into the current through various components of the system and is distributed in various parts of the PG. The integration of PPO generation will affect the active line loss of the distribution system and change the power flow of the power system. This study considers photovoltaic as a node with known active and reactive power, and the power flow calculation process is shown in equation (13). In equation (13), f_{loss} serves as the active line loss of the power system; N serves as the total distribution system lines; G_{ij} represents the conductivity in nodes i and j; U_i, U_j represent the voltage amplitudes of nodes i and j; θ_i, θ_j represents the phase angle in nodes i and j.

$$f_{loss} = \sum_{k=1}^{N} G_{ij}\left(U_i^2 + U_j^2 - 2U_iU_j\cos\left(\theta_i - \theta_j\right)\right) \tag{13}$$

In addition, the integration of PPO generation will also possess an influence on the voltage deviation of nodes, and PPO generation will cause the voltage of nodes to increase, even exceeding the voltage limit. The functional formula for voltage deviation is shown in equation (14), where ΔU represents active line loss; U_i serves as the total number of lines; $U_{\exp}$ serves as the conductivity between nodes i and j.

$$\Delta U = \sum_{k=1}^{N} \left| U_i - U_{\exp} \right| \tag{14}$$

Incorporating the above three factors into the consideration of multi-objective optimization (MOO) in configuration, this study uses Pareto solutions to describe MOO problems [21]. The definition of dominance relationship is that if there are two solutions x_1 and x_2 that meet $f_i(x_1) f \leq_i (x_2)$, then x_1 is considered to dominate x_2. The Pareto solution is a non-dominated solution, and there is no other solution that dominates the existence of x^*.

In MOO problems, it is necessary to find as many Pareto solutions as possible [22]. To achieve this goal, PSO is introduced. PSO algorithm is a kind of swarm intelligence algorithm, which is designed to simulate the natural birds' predatory behavior. During the entire process of bird search for food sources, the location of the food source is determined through mutual information transmission and collaboration, which is for finding the optimal solution (OS) for the issue. The algorithm diagram is shown in Figure 4 [23-24].

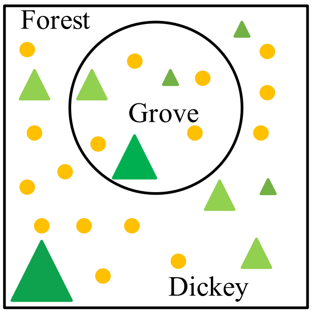

Figure 4. Schematic diagram of Particle swarm optimization algorithm

The iteratively defined formula for particle position and velocity is showcased in equation (15); In equation (15), w represents the inertia weight, c_1 and c_2 represent the acceleration factors of individuals and populations, r_1 and r_2 represent random numbers, α represents the teacher constraint factor, and m represents the number of iterations.

$$\begin{cases} v_{ik}^{m+1} = wv_{ik}^m + c_1r_1\left(p_{ik}^m - x_{ik}^m\right) + c_2r_2\left(s_k^m - x_{ik}^m\right) \\ x_{ik}^m = x_{ik}^m + \alpha v_{ik}^{m+1} \end{cases} \tag{15}$$

The update of particle position is influenced by the particles themselves and other particles, and the communication methods between particles vary, forming a variety of particle swarm topological networks, as shown in Figure 5. The particles are interconnected, and the global Rate of convergence is fast. The speed of particle updates affects the position control of particles. If the particle updates too fast, there may be a particle position vector exceeding the search limit. The study used a reset method to solve this problem. When the particle exceeds the search space, the particle's position is randomly reset.

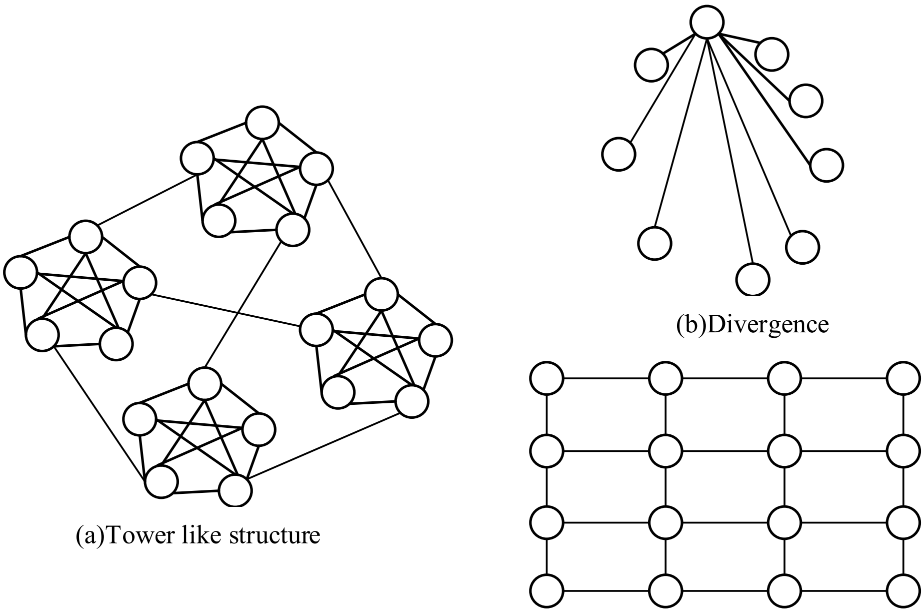

Figure 5. Different Style Particle Swarm Topology Networks

The parameter selection of PSO is very important, which determines the initial performance of the algorithm. The Particle number, the maximum particle update speed, learning factor and inertia constant affect the particle distribution uniformity. In the study, the number of particles is selected according to the actual situation, and the particle speed is limited to 15%~20% of the search space to prevent the convergence of particle search from becoming poor and local convergence. The values of learning factors and inertia constants are taken on the ground of existing research experience.

To prevent local convergence of particles, adaptive constraint factors and genetic mutation operations are first introduced into PSO [25]. The calculation of adaptive constraint factors is shown in equation (16); In equation (16), $\alpha_{\max}$ and $\alpha_{\min}$ represent the maximum and minimum values of the constraint factors, while $iter$ and $iters$ represent the current and total iterations, respectively. The adaptive operation controls the speed of the particles so that the algorithm performs a large search early in the iteration and reduces the search later in the iteration.

$$\alpha = \alpha_{\max} - \left(\alpha_{\max} - \alpha_{\min}\right) iter / iters \tag{16}$$

The purpose of mutation operation is to increase the diversity of particles, and the perturbed particle position avoidance algorithm converges to the local OS. The relevant flow of the entire improved PSO model is illustrated in Figure 6; Finally, it will combine the improved PSO algorithm on the ground of Pareto solutions with MOO to solve MOO problems, and ultimately obtain the Pareto solution set. As can be seen from Fig. 6, the iterative process needs to constantly update the position and velocity of the particles, and the multi-objective optimization problem needs to use the dominance and non-dominance relationship to determine the goodness of different solutions, and therefore needs to be the computational fitness of the particles at the same time with the dominance relationship. The multi-objective algorithm needs to select the Pareto solution set among the different particle position solutions, and thus also needs to compare the solutions with each other to get the non-inferior solution reserve set. The role of the variation operator is to prevent the particles from falling into the local Pareto front.

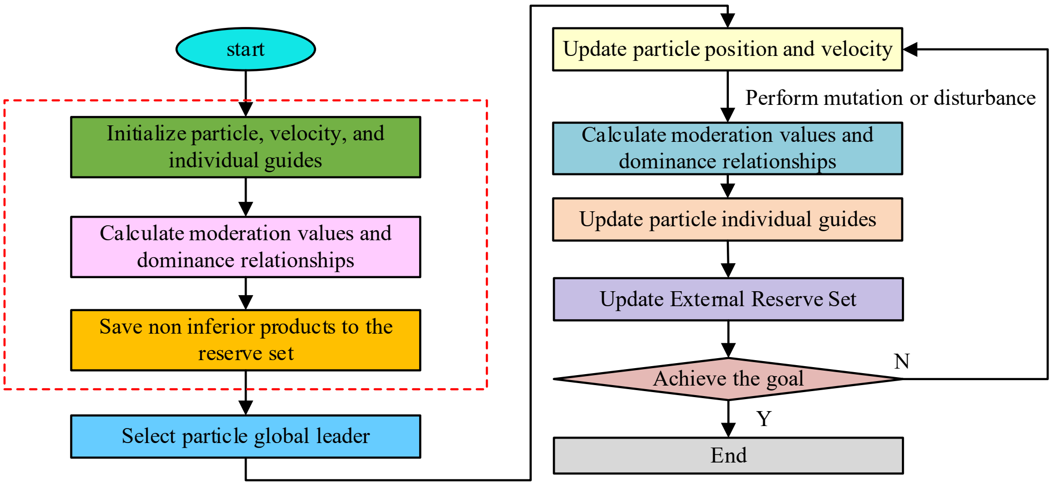

Figure 6. Algorithm flow of improved Particle swarm optimization model

4. Performance testing of photovoltaic power generation prediction and optimization configuration model on the ground of GPR and PSO algorithms

For testing the prediction model and optimized configuration model designed in the study, targeted simulation tests are designed for verifying the model, and the outcomes were analyzed and discussed.

4.1 Simulation Experiment Analysis of Photovoltaic Power Generation Power Prediction on the Ground of GPR

This study used detection data from a solar energy technology limited company and local meteorological bureau monitoring data as input data for analysis and normalized the feature vectors of the input data. First, it compared the prediction outcomes of the GPR prediction model built by the study with the real data curve and the regression prediction model on the ground of Lasso. The prediction curve is showcased in Figure 7. Figure 7 illustrated that the regression prediction curve of the GPR prediction model for the simulation data has a high coincidence with the real curve, and a high closeness to the real value. The prediction curve of Lasso regression prediction model fluctuates greatly. Although the prediction trend is close to the real data curve, the prediction data is disturbed greatly under the interference of data noise, and the prediction effect is worse than that of GPR.

It compares the GPR prediction model with Lasso based regression prediction model, ANN model and time series model ARMA, and uses Root-mean-square deviation (RMSE) and Mean Absolute Scaled Error (MASE) as evaluation indicators of experimental results; RMSE can measure the difference between predicted and true values, RMSE represents the square root of the mean of the difference between the observed value and the predicted value, the smaller the RMSE. And the MASE index is not easily affected by data changes, the MASE can more accurately predict the model's truthfulness relative to the mean value model. The industry evaluation criteria for PV power generation forecast accuracy varies by region and specific needs, Figure 7 compares with actual power generation data, while Figure 8 selects the error calculation index for evaluation based on industry standards obtained from industry reports and expert consultation. RMSE focuses more on the difference between the predicted value and the true value, while the comparison of MASE with respect to the baseline model better reflects the model's relative strengths and weaknesses. The two indicators can comprehensively evaluate the overall accuracy and relative accuracy.

It is used for measuring the prediction of the model's power generation. The experiment outcomes are illustrated in Figure 8. Figure 8 showcases that the RMSE values of the ANN model and the Lasso regression prediction model are relatively high; This shows that there is a large disparity between the predicted value and the real value. The RMSE index of the ARMA model is markedly below the two models, but the numerical polyline volatility is large, and the data value is unstable. The RMSE value of the GPR model outperforms the other three models in both numerical and stability aspects. The data integration of multiple influencing factors reduces the disparity in real data and predicted data, making predictions more accurate. For the MASE index, the GPR still shows the optimal level and is at the lowest value; This demonstrates that the regression prediction effect is good, while the values of other models are significantly higher, and the volatility is too obvious. The prediction model used in the study possesses better prediction stability and accuracy compared to other existing prediction models.

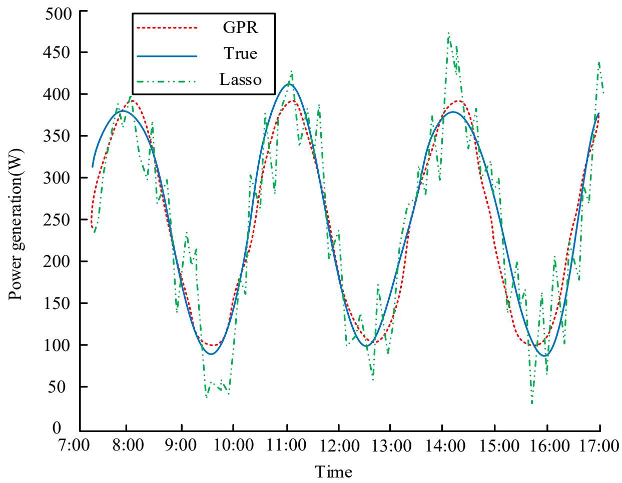

Figure 7. Comparison of Prediction Curves of Different Models

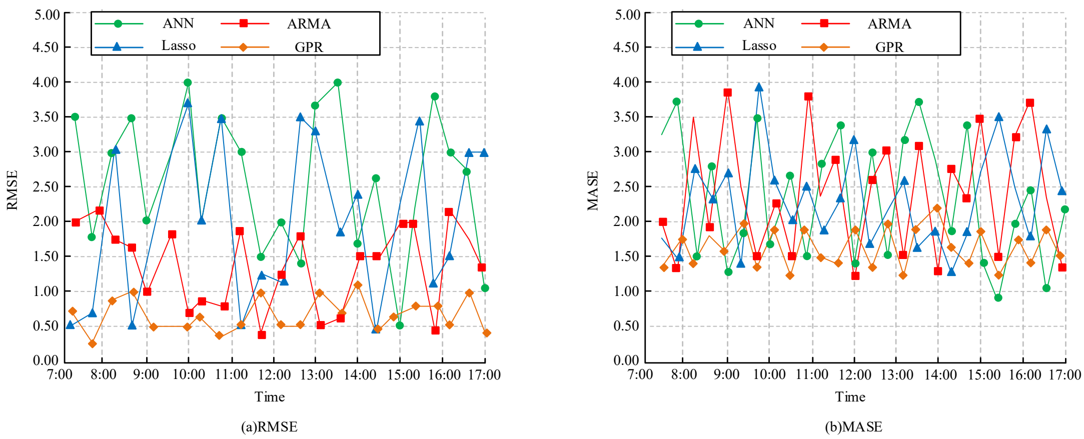

Figure 8. Comparison of Performance Indicators of Different Models

4.2 Simulation Experiment Analysis of Photovoltaic Optimization Configuration Model on the Ground of Improved PSO Algorithm

The multi-objective PSO algorithm is introduced into the MOO of distribution network connected with photovoltaic equipment for configuration. The research sets the population quantity to 150 and the maximum quantity of iterations to 100; The acceleration factor and maximum constraint factor are 1 and 2, respectively, the minimum constraint factor is set to 0.2, and the inertia weight value is 2. The optimization results of voltage deviation are shown in Figure 9, and the traditional PSO model and differential evolution model are compared respectively. Figure 9 shows that, relative to the traditional PSO model, the improved PSO model with mutation manipulation is still searching for the OS at the end of the iteration, while the traditional PSO model converges for the OS at the end of the iteration. The voltage deviation stability value is greater than 0.012, and there is no improvement of the OS. The differential evolutionary algorithm keeps searching for optimization in the iterative process. The convergence value and the improved PSO algorithm converge to the voltage deviation of about 0.011, but the rate of convergence of the OS is still lower than the improved PSO algorithm.

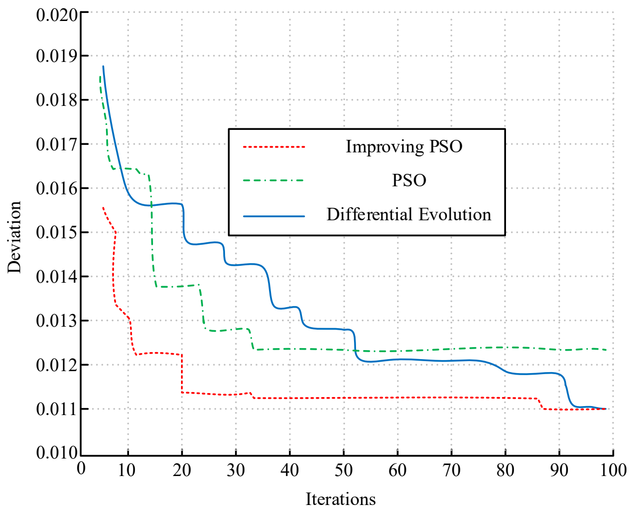

Figure 9. Voltage Deviation Curve

It sets the maximum size of the Pareto solution set to 100, and the Pareto solution set results obtained from simulation experiments are showcased in Figure 10. Figure 10 showcases that the relationship between photovoltaic cost, grid loss, and voltage deviation is uniformly dispersed in the three-dimensional space of the Pareto solution set. When the input photovoltaic cost is higher, the grid loss and voltage deviation values are smaller.

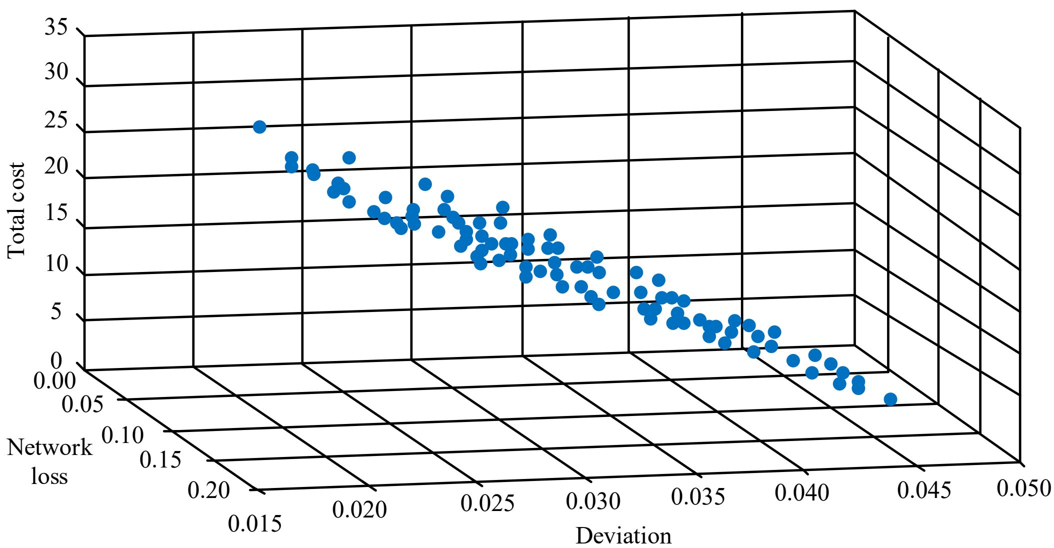

Figure 10. Pareto Solution Set 3D Space Diagram

It selects several analysable schemes among the nodes in the distribution system, and the PPO generation configuration and node parameters is showcased in Table 1. Table 1 showcases that the active line loss without PPO generation is 204kW, and the voltage deviation is 0.049; The active line loss of the other three types of connected PPO generation is less than 204kW, with a maximum reduction of 132kW; The voltage deviation has also decreased, with a maximum decrease of 0.028.

Table 1. Typical Power Grid Configuration Scheme

Programme	Scheme photovoltaic installation point/ Power (kW)			Active line loss/kW	Deviation	Total Cost/ten thousand
0	0	0	0	204	0.049	0
1	13/467	16/423	32/123	87	0.021	25.9
2	17/241	30/484	32/175	72	0.027	23.2
3	7/34	14/59	18/331	137	0.043	11.2
Starting	Starting node	End node	Branch impedance	End injection power		Node voltage
12	12	13	1.4683+j1.1551	60+j35		1
13	14	14	0.5414+j0.7106	120+j80		1
16	16	17	1.3363+j1.6984	60+j20		1
17	17	18	0.8256+j0.6827	90+j40		1
32	32	33	0.3413+j0.5312	60+j40		1

The voltage distribution diagrams of various schemes are illustrated in Figure 11. Figure 11 showcases that the minimum voltage of the scheme without PPO generation is less than 0.92, which is lower than other schemes with PPO generation connected to the distribution network. This indicates that the MOO configuration model on the ground of improved PSO designed in the research can reasonably configure photovoltaic equipment, effectively reduce network losses and improve the quality of voltage. The Pareto method can be used to obtain different solutions suitable for PPO generation optimization.

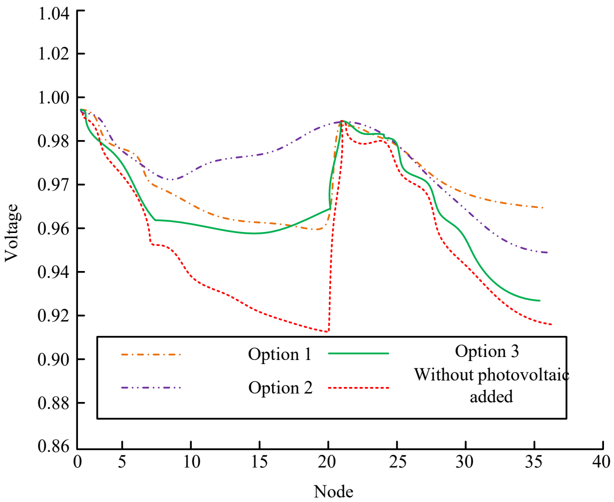

Figure 11. Voltage distribution diagram for different schemes

5. Conclusion

To solve the balance and security problems of the power system, the research builds a PPO generation prediction and configuration optimization model on the ground of GPR and improved PSO algorithm. The simulation results show that the regression prediction curve of the GPR prediction model to the simulation data is the closest to the real curve; The prediction curve of the Lasso regression prediction model exhibits significant fluctuations, resulting in poor stability of the prediction results and susceptibility to interference from noisy data. There is a significant disparity in the predicted values of the ANN model and the Lasso regression prediction model and the true values, while the RMSE index stability of the ARMA model is poor. The RMSE and MASE indicators of the GPR model have better values and stability than the other three models. The integration of multiple factor data increases the accuracy of the predicted data, and the regression prediction effect is good. Relative to the traditional PSO model, the improved PSO model performs better in improving the OS, and its rate of convergence is faster than the differential Evolutionary algorithm. In the three-dimensional space of the Pareto solution set, the larger the photovoltaic cost, the smaller the grid loss and voltage deviation values. The Pareto method can be used to obtain different solutions suitable for PPO generation optimization. Reasonably integrating PPO generation into the distribution system can effectively reduce active line loss and voltage deviation, with the maximum reduction values reaching 132kW and 0.028, respectively. The PPO generation prediction model designed through research has good performance and comprehensive prediction performance. Reasonable configuration has improved the rationality of the distribution network. But research still needs to be more in line with the century for feasibility optimization.

Funding

The research is supported by: Research and demonstration of multifunctional scenic spots integrated energy service station based on off-grid photovoltaic power distribution technology.

References

[1] Abdul R, Naila N, Muhammad I, Madad A. The current developments and future prospects of solar photovoltaic industry in an emerging economy of India. Environmental science and pollution research international,2023,30(16): 46270-46281.
[2] Guo N. The impact of energy industry structure adjustment on digital media application technology. Energy Reports, 2022, 8(S4): 1463-1471.
[3] Hui S, Al N K, Ammar K, Samaneh Sadat S, Ali Moradi A, Mahdi J, Xinghuo Y, Peter M. Multitasking recurrent neural network for photovoltaic power generation prediction. Energy Reports, 2023, 9(S3): 369-376.
[4] Bashir U T, Fuwen Y, M. S A M, Junwei L. Solar photovoltaic power forecasting for microgrid energy management system using an ensemble forecasting strategy. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2022, 44(4): 10045-10070.
[5] Xinyu Z, Yanshuang A, Xinlu W, Xifeng G, Wei D. Learning with privileged information for short-term photovoltaic power forecasting using stochastic configuration network. Information Sciences,2023,619(1): 838-848.
[6] Chen B, Lin P, Lai Y, Wu L. Very-short-term power prediction for PV power plants using a simple and effective RCC-LSTM model based on short term multivariate historical datasets. Electronics, 2020, 9(2): 289-308.
[7] Malik P, Chandel R, Chandel S S. A power prediction model and its validation for a roof top photovoltaic power plant considering module degradation. Solar Energy, 2021, 224(1): 184-194.
[8] Kim D, Kwon D, Park L, Kim J, Cho S. Multiscale LSTM-based deep learning for very-short-term photovoltaic power generation forecasting in smart city energy management. IEEE Systems Journal, 2020, 15(1): 346-354.
[9] Lee D, Kim K. PV power prediction in a peak zone using recurrent neural networks in the absence of future meteorological information. Renewable Energy, 2021, 173(1): 1098-1110.
[10] Aprillia H, Yang H T, Huang C M. Short-term photovoltaic power forecasting using a convolutional neural network–salp swarm algorithm. Energies, 2020, 13(8): 1879-1899.
[11] Yan J, Hu L, Zhen Z, Wang F, Qiu G, Li Y, zhong Yao L, Shafie-khah M, Catalão J. Frequency-domain decomposition and deep learning based solar PV power ultra-short-term forecasting model. IEEE Transactions on Industry Applications, 2021, 57(4): 3282-3295.
[12] AlShafeey M, Csáki C. Evaluating neural network and linear regression photovoltaic power forecasting models based on different input methods. Energy Reports, 2021, 7: 7601-7614.
[13] Hao Y, Dong L, Liang J, et al. Power forecasting-based coordination dispatch of PV power generation and electric vehicles charging in microgrid. Renewable Energy, 2020, 155: 1191-1210.
[14] Yan M, Guo W, Hu Y, Xu F, Chen J, Du Q, Qin T. Improved hybrid sparrow search algorithm for an extreme learning machine neural network for short‐term photovoltaic power prediction in 5G energy‐routing base stations. IET Renewable Power Generation, 2023, 17(2): 336-348.
[15] Song H, Al Khafaf N, Kamoona A, Sajjadi S S, Amani A M, Jalili M, McTaggart P. Multitasking recurrent neural network for photovoltaic power generation prediction. Energy Reports, 2023, 9(5): 369-376.
[16] Wang S, Zhu H, Zhang S. Two-stage grid-connected frequency regulation control strategy based on photovoltaic power prediction. Sustainability, 2023, 15(11): 8929-8949.
[17] Yang X, Zhao Z, Peng Y, Ma J. Research on distributed photovoltaic power prediction based on spatiotemporal information ensemble method. Journal of Renewable and Sustainable Energy, 2023, 15(3): 36102-36118.
[18] Chen G, Zhang T, Qu W, Wang W. Photovoltaic Power Prediction Based on VMD-BRNN-TSP. Mathematics, 2023, 11(4): 1033-1046.
[19] Wang Z , Zhang J , Jiang D, Liu F, Hao L. Predictive modelling for contact angle of liquid metals and oxide ceramics by comparing Gaussian process regression with other machine learning methods. CERAMICS INTERNATIONAL, 2022,48(1): 665-673.
[20] Zhao X, Zhang D, Zhang R, Xu B. A comparative study of Gaussian process regression with other three machine learning approaches in the performance prediction of centrifugal pump. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2022, 236(8): 3938-3949.
[21] Barma M, Modibbo U M. Multiobjective mathematical optimization model for municipal solid waste management with economic analysis of reuse/recycling recovered waste materials. Journal of Computational and Cognitive Engineering, 2022, 1(3): 122-137.
[22] Libotte G B, Lobato F S, Moura Neto F D, Platt G M. A Novel Reliability-Based Robust Design Multiobjective Optimization Formulation Applied in Chemical Engineering. Industrial & Engineering Chemistry Research, 2022, 61(9): 3483-3501.
[23] Liu Y, Yang Y. An extended VIKOR method based on particle swarm optimization and novel operations of probabilistic linguistic term sets for multicriteria group decision-making problem. International journal of intelligent systems, 2022, 37(8): 5381-5424.
[24] Mao L , Zhao J , Zhu Y ,Chen J .A noise-immune model identification method for lithium-ion battery using two-swarm cooperative particle swarm optimization algorithm based on adaptive dynamic sliding window. International Journal of Energy Research, 2022, 46(3): 3512-3528.
[25] Xia G, Chen J, Tang X, Zhao L, Sun B. Shift quality optimization control of power shift transmission based on particle swarm optimization–genetic algorithm. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2022, 236(5): 872-892.