Prediction of water-light output in water-light complementary systems

Yan, Tian Xiao [1,a], Fan, Hui Ying[1,b], Guo Su[1,c*]

[1] College of Energy and Electrical Engineering, Hohai University, Nanjing, Jiangsu,211100, China

Abstract: With the increasing demand of electricity consumption for social development, China has vigorously promoted the development of renewable energy such as wind and solar energy in recent years. Since photovoltaic power generation has the characteristics of randomness, intermittency and volatility, which are not conducive to dispatching and grid-connection, and hydroelectric power generation has the characteristics of fast start-up and good peak adjustment, etc. Therefore, through the study of multiple linear regression model neural network prediction model Markov chain and other methods, the hydraulic resource output prediction model and photovoltaic resource output prediction model are built. The combination of the two can be used to predict the output of water and light resources more accurately and promote the development of water and light complementary industry in China. At present, there are few researches on the coupling of hydroelectric output and photovoltaic output at home and abroad. In this paper, the prediction of water-light output in water-light complementary systems by other scholars is reviewed.

1 Preface

The number of rivers in our country is large, and the difference between the terrain is large, and the water energy resources are rich. According to the results of our country's water resources review, the total amount of hydropower resources in China ranks first in the world, and the solar energy resources in China are also extremely rich.

China is a big energy consumer. With the large consumption of non-renewable energy such as coal, oil and natural gas, we are in urgent need of a new energy utilization method to replace the traditional energy forms. Solar power generation is safe, pollution-free and has the characteristics of short construction period [1]. However, the photovoltaic power generation is affected by seasonal changes, day and night, weather and other factors, resulting in great variation of photovoltaic output within the day, if direct grid connection will aggravate the difficulty of peak adjustment of the power grid. At the same time, hydropower plays a significant role in the peak regulation of power grid, which has the advantages of flexible start-up and fast regulation. Combined with the advantages of China's water power resources and light energy resources, we can complement each other's advantages, and then make greater use of these resources.

2 Research status

2.1 Foreign research status

Overseas research on grid-connected photovoltaic power station mainly focuses on reforming the power grid and improving the efficiency of power generation. Jure Margetal [2] proposed that solar energy should be used as the main energy source and pumped storage energy should be used to solve the problems of fluctuation and intermittency of photovoltaic power generation Solar. For system [3], developed by Enfor Company in Denmark, combines the short-term weather forecast information, historical output power data, geographical location and other information to predict the short-term output power of photovoltaic power generation by using an adaptive statistical model. Atsushi Yona, Et al. [4] took the pressure, temperature, relative humidity and wind speed of the previous 18 hours as the input of the neural network, and used the feedforward neural network radial basis neural network and the recursive neural network respectively to predict the solar radiation amount.

2.2 Domestic research status

In domestic pv power generation system power output prediction technology research and development of this block is in the stage. Chen Chang song [5]and others through the analysis of pv array output power and the

Yan, Tian Xiao, [a] 2395588586@qq.com
Fan, Hui Ying, [b]1366717902@qq.com
*Corresponding author: Guo, Su,[c]:guosu81@126.com

historical meteorological data and history data, using BP neural network methods of RBF neural network, the complex weather first preliminary identification. Then divided into the sunny day, cloudy day and rainy day three categories, on this basis, set up photovoltaic power generation system, based on short-term weather forecast information, output prediction model. Ding Ming [6] put forward a kind of directly based on the Markov chain prediction method of photovoltaic power station output power. At present, there are relatively few theoretical system studies and coordinated operation mathematical model studies on water-light complementarity in China, which makes the development of photovoltaic grid-connected power generation relatively slow. Therefore, the research on the theory and development of water-light complementarity plays a huge role in promoting the development of photovoltaic power generation in China.

3 Analysis of influence factors of output

3.1 Analysis of main influencing factors of photovoltaic output

The main factors affecting photovoltaic output can be roughly divided into external factors and internal factors [7,8]. External factors include meteorological factors, geographical factors and daily weather types. Meteorological factors are: the amount of solar irradiance, temperature, humidity, wind speed, sunshine time, etc. Geography: the geographical position, pressure, etc. Day weather types are divided into: sunny, cloudy, rain, fog, etc. Internal factors are affected by the quality of photovoltaic modules, photovoltaic inverter quality, bus box, transformer station quality and so on [9]. The relationship between external factors and photovoltaic output is analyzed by correlation coefficient method, and then the influence degree of each factor on photovoltaic output can be learned.

Correlation coefficient can be used to represent the correlation between the two vectors. Correlation analysis is shown in Table 1. According to the table, the daily average sunshine minimum relative humidity and average relative humidity have the most significant influence on photovoltaic output, while the daily minimum temperature and daily average temperature are more significant.

Tab.1 Photoelectric output with each factor correlation coefficient table [10].

Month	The average sunshine	Minimum relative humidity	Mean relative humidity
Jan.	0.83	-0.76	-0.63
Feb.	0.81	-0.60	-0.47
Mar.	0.63	-0.60	-0.28
Apr.	0.87	0.01	-0.49
May.	0.88	-0.53	-0.39
Jun.	0.94	-0.90	-0.70
Jul.	0.88	-0.77	-0.69
The maxim um	**0.94**	**-0.90**	**-0.70**

Month	Daily mean temper ature	Daily maximum temperatur e	Daily minimum temperatu re
Jan.	0.01	0.49	-0.12
Feb.	0.03	-0.09	-0.47
Mar.	0.12	-0.10	-0.10
Apr.	0.32	0.01	-0.21
May.	-0.19	-0.47	-0.62
Jun.	0.56	-0.33	-0.25
Jul.	0.43	0.03	-0.34
The maxim um	**0.56**	**-0.49**	**-0.47**

data presentation

1.The variation trend of sunshine duration is consistent with that of photovoltaic output.

2.The trend of daily maximum temperature and daily average temperature is basically similar to that of photovoltaic output, but the correlation coefficients are different in different months.

3.Relative humidity is basically opposite to the trend of photovoltaic output.

4.The variability of cloud cover has strong randomness, which makes it difficult to predict photovoltaic output.

3.2 Analysis of influencing factors of hydroelectric output

Hydropower output is affected by runoff inflow, reservoir upper and lower water levels, generator unit efficiency, reservoir evaporation and other factors, among which, runoff inflow has a greater impact. The climate has the greatest influence on runoff, and the annual runoff can be divided into dry year, normal year and wet year. A year can be divided into: dry season, normal season, flood season, with different geographical environment in the basin, the beginning and end of each period has a certain difference. Climate change has an impact on the meteorology of each river basin and will change the runoff in the future. The change of runoff will directly affect the generation output process of the reservoir. Therefore, the prediction of future runoff variation has a profound impact on the prediction of hydropower output.

4 PV output prediction model

4.1 Method for predicting output power of photovoltaic power station

The forecasting methods of photovoltaic power generation can be divided into two categories: indirect forecasting method and direct forecasting method.

(1) Indirect forecasting/data-driven model forecasting.

Indirect prediction method is based on the rational modeling of the historical data of solar radiation intensity. Then, according to the physical principles of voltaic cells

and the definition of photovoltaic conversion efficiency, the expression between the output power of the photovoltaic power generation system and solar radiation intensity, photoelectric conversion rate, MPPT conversion efficiency and other factors is established. Finally, the above model can be used to calculate the output power of photovoltaic power generation [11]. Although this method is effective, it relies too much on the complex solar radiation intensity prediction model and detailed and accurate public weather forecast information. At the same time, for different types of photovoltaic power generation unit systems, there are also differences in conversion efficiency, installation Angle and other parameters, which will increase the difficulty in determining the output power model parameters [12].

(2) Direct prediction method

This method is based on the historical data of the output power of photovoltaic power generation system. Firstly, the output power historical data are reasonably modeled (commonly used artificial neural network, Markov chain, support vector machine, gray scale model, etc.). Then, the short-term output power value of the photovoltaic power generation system is directly predicted using the information of weather forecast [13]. This method avoids the complicated modeling process, saves the investment of solar radiation intensity measurement device, and simplifies the prediction process. However, it requires the accumulation of historical data over a period of time as a guarantee, and has high requirements for the parameter selection of model input [14].

4.2 Medium - and long-term photovoltaic output forecast

In order to predict the medium and long term photovoltaic output, the method of multiple linear loop analysis can be used to forecast [15]. The monthly average temperature, monthly minimum temperature, monthly average relative humidity and monthly average sunshine hours were used as the multiple linear regression PV output prediction model.

The multiple linear regression model is as follows：

$$y = a_0 + a_1 x_1 + \ldots + a_m x_m + \varepsilon \quad \varepsilon \sim N \ （0，\ \sigma^2 ）$$
$$(4\text{-}1)$$

In the formula, $a_0, a_1, \ldots, a_m$ is the regression coefficient, ε is the random error.

In the photovoltaic output prediction, historical data are substituted into Equation (4-1). The least square method is used to obtain the estimator of parameter $a_0, a_1, \ldots, a_m$, and then the regression equation is substituted into (4-1).

$$\hat{y} = \hat{a}_0 + \hat{a}_1 x_1 + \ldots + \hat{a}_m x_m \qquad (4\text{-}2)$$

If the correlation degree between the predicted object and the predictor in the regression model is statistically tested, the R test method and the F test method are usually adopted.

4.3 Short-term PV output forecast

4.3.1 Processing of input data for short term forecasting models

Since the field measurement time is short and the amount of solar radiation is random, in order to make more accurate prediction, we can first process the data to a certain extent [16].

（1）Processing of historical photovoltaic output data.

The historical data are divided according to a certain period of time (1h, 1d), and the data are classified respectively according to season and weather.

（2）Meteorological data processing

The representative sunshine duration data were selected from the long series of meteorological data and the engineering representative year was established to fully reflect the variation rule of long-term sunshine duration. The temperature data and humidity data were processed to obtain the average data with a duration of 1 day for the convenience of the model.

（3）Selection of similar days

According to the weather information provided by the meteorological department, two vectors $x_0(1), x_0(2)$ are used to represent two different influencing factors of the forecast day. The sunshine duration and the minimum relative humidity that have the greatest influence on photovoltaic output are usually selected. At some point in history, the vector of these two factors was x_j $=[x_j(1), x_j(2)]$. Find the correlation coefficient between x_0 and x_j. Starting from the adjacent historical days, the similarity between them and the predicted days is calculated and compared one by one. The historical day with the greatest similarity is regarded as the similar day of the date to be measured.

4.3.2 Establishment of short - term PV prediction model

（1）Establishment of prediction model of Markov chain [17]：

1. Make statistics of similar daily average photovoltaic output in different seasons and weather conditions.

2. The output is divided into four intervals, and the power falling within one interval is regarded as a state.

3. The state transfer times of photovoltaic power output were counted in the study period, and the one-step transfer matrix of each period was obtained.

4. The initial state of the next unit time is the output value predicted at the end of the previous unit time, and its probability is denoted as 1.

5. Calculate the one-step transfer probability and judge the average output range of the predicted daily photovoltaic.

（2）The establishment of neural network prediction model：

1.The determination of input and output quantity

As the photovoltaic power generation curve has certain similarity in the shape of the same weather type [18].

Therefore, the relevant power output data of the previous day or two with similar climate types are used as input variables. Considering that there are many influencing factors on photovoltaic power generation, solar irradiation, temperature and relative humidity with the greatest influence are finally recommended as meteorological influencing factors.

2. Selection of hidden layer and number of nodes of hidden layer.

Elman neural network usually adopts the network structure of single hidden layer[19]. When the input factors are too many, the structure of multiple hidden layers can be adopted. However, the number of hidden layers should not be too many, otherwise it may increase the complexity of the network, lead to the phenomenon of "overfitting", the training time will become longer, and the prediction accuracy will decline. Therefore, it is recommended to use the appropriate number of hidden layers.

3. Optimization of neural network by grid search method.

（3）Ultra-short - term PV output prediction model based on stepwise regression.

1. Classify historical data according to season and weather.

2. The rolling prediction method is adopted to make predictions hour by hour. The input data is the data of the factors affecting the photovoltaic output of the day before the forecast and the photovoltaic output value of the first few hours of the hour.

3. The stepwise regression equation is obtained through the application of the Stepwise statement in the software.

4.4 Evaluation index of photovoltaic output prediction model

In order to evaluate the forecast results, the following three evaluation indicators are recommended [20].

(1) Absolute Error (AE)

$$AE = \hat{x}(k) - x(k) \qquad （4\text{-}3）$$

In the formula, $x(k)$ is the measured value $\hat{x}(k)$ is the predicted value.

(2) Normalized Root Mean Square Error (NRMSE)

$$NRMSE = \frac{1}{N\sigma^2}\sum_{k=1}^{N}[x(k) - \hat{x}(k)]^2 \qquad （4\text{-}4）$$

In the formula, N is the number of prediction tests, and σ^2 is the variance of the predicted measured sequence

(3) Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}\left|\frac{x(k) - \hat{x}(k)}{x(k)}\right| \qquad （4\text{-}5）$$

5 Hydropower output prediction model

The hydropower output prediction benefits from the runoff prediction of the reservoir, the control water level at the beginning and the end of the reservoir, and the comprehensive consideration of the constraint conditions. In this way, the change of reservoir water level can be determined and the output of water resources can be predicted.

5.1 Runoff prediction model

With the development of computer technology, fuzzy mathematics, chaos theory, wavelet analysis and support vector machine (SVM) have been widely used in the medium and long term runoff prediction.

5.1.1 SWAT Hydrological Model

SWAT watershed scale hydrological model has a relatively complete physical mechanism and has continuity. The historical measured data were used to calibrate the SWAT model, and the rainfall and temperature data obtained from the SDSM statistical downscaling model were input into the SWAT model, so that the runoff process of each outlet section of the basin could be obtained, and then the runoff process into the reservoir could be calculated. The model is computationally efficient and can be used to simulate the distributed watershed over a long period of time.

5.1.2 Hydrological model of the Xin 'an River

The hydrological model of the Xin 'an River divides the basin into several blocks, calculates the runoff production and confluence of each block respectively, and obtains the discharge process of the outlet section after summing up. The parameters were calibrated in the order of evapotranspiration, runoff generation, water source separation and confluence. By inputting the calibrated data into the Xin 'an River model, the runoff process of each outlet section of the basin can be obtained, and then the runoff process into the reservoir can be calculated. The model is suitable for wet and subhumid areas, but the accuracy is sometimes not high enough.

5.1.3 Artificial neural network model

The multi-year historical monthly runoff data simulated by the SAR[21] model (seasonal autoregressive model) were used to calculate the partial correlation coefficient of each month according to the description[22] of the selection method of forecasting factors in the literature, and the historical data of appropriate months were selected as the forecasting factors. Through continuous attempts to select the appropriate model structure and appropriate training proportion, the model with the best forecast data effect is selected.

5.1.4 Consistency check

The consistency between future runoff and historical runoff can be tested by square Newman ratio method and cumulative deviation method.

The square Newman ratio is defined as:

$$N = \frac{\sum_{i=1}^{n-1}(Y_i - Y_{i+1})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \qquad (5\text{-}1)$$

Where $\bar{Y}$ represents the average of the series Y_i. If the sample is discontinuous, N is usually lower than expected.

The cumulative deviation method needs to construct statistics:

$S^* = 0 \quad S_k^* = \sum_{i=1}^{k}(Y_i - \bar{Y}) \quad$ k=0,1,2,…,n(5-2)

Divide S^* by the sample standard deviation to get the sum of the adjusted parts of the readjustment:

$S^{**} = \frac{S^*}{D_Y}$ \qquad\qquad k=0,1,2,…,n \quad (5-3)

$D_Y^2 = \frac{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}{n}$ \qquad k=0,1,2,…,n \quad (5-4)

Review consistency parameter R:

$R = max_{0 \le k \le n}S^{**} - min_{0 \le k \le n}S^{**}$

The high value of R reflects the change in the level, and the consistency is usually judged by the size of $R/\sqrt{n}$.

5.2 Hydropower output forecast

The use of water in the reservoir needs to consider the downstream irrigation water, the downstream ecological water and the generating capacity of the power station. Therefore, it is constrained by many constraints, such as the water balance of the reservoir, the water level limit of the reservoir, the discharge of the reservoir, the output limit of the hydropower station, and the non-negative variable. Combined with the runoff prediction model above, the solution space is worked out, and then the water scheduling is optimized, and finally the forecast data of water resource output is obtained.

6 Conclusions

In this paper, the main influencing factors of hydroelectric output and photovoltaic output are analyzed, and the runoff prediction modeling and photovoltaic output prediction modeling are preliminary summarized. The main conclusions of this paper are summarized as follows:

1. Data are usually classified according to seasons and weather in data processing.

2. The factors that have the greatest influence on photovoltaic power generation are roughly as follows: sunshine hours, minimum relative humidity, daily average temperature, daily maximum temperature and daily minimum temperature.

3. Influencing factors of hydroelectric power generation include: runoff inflow, upper and lower water levels of reservoirs, generator unit efficiency, and reservoir evaporation, etc. The biggest factor affecting runoff is the amount of runoff into the reservoir, and the biggest factor affecting runoff is climate change.

4. For medium and long term predictions, multiple linear regression model is recommended. The model is simple and the results are relatively accurate.

5. For short-term prediction, Markov chain model, BP neural network model, Elman neural network model and stepwise regression model are recommended. All four models have their own advantages.

6. Three different evaluation model indicators are recommended: absolute error (AE), standard root mean square error (NMSE), and mean absolute percentage error (MAPE).

7. Runoff prediction models include SWAT model, artificial neural network model, etc. Square Newman ratio method and cumulative deviation method can be used to test the consistency between future runoff and historical runoff.

Due to climate change, the consistency between future runoff and historical runoff has changed, and the characteristics have been sharpened. The most direct manifestation is that the runoff in flood season is increased and the runoff in dry season is reduced. As a result, the consistency between the future runoff and the historical runoff obtained after the direct use of SWAT model is reduced, which will make the prediction of future runoff more difficult. Due to the instability of irradiation, the error between the predicted value and the real value is relatively large in the forecast of photovoltaic power generation in cloudy and rainy days. These problems remain to be further studied.

Acknowledgments

References

1. Duo-sheng ,F, Miao, Z, Hui, Z, et al. Solar Power Generation Technology and Application [M]. Bei Jing: Posts and Telecommunications Press, 2009

2. Jure Margeta. Zvonimir Glasnovic .Theoretical settings of photovoltaic-hydro energy system forsustainable energy production[J].Solar Energy,2012.

3. Bacher P,Madsen H,Nielsen H A.Online short-term solar power forecasting[J].Solar Energy,2009,83 (10):1772-1783.

4. Yona A,Senjyu T,Saber A Y,et al.Application of neural network to 24-hour-ahead generating powerforecasting for PV system[C].IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century，2008.

5. Chang song, C, Shan xu, D, Jin jun, Y. Design of photovoltaic array power generation prediction model based on neural network [J]. Transactions of China Electrotechnical Society, 2009,24 (9) : 153-158

6. Ming, D, Ning zhou, X. Short-term forecasting method of output power of photovoltaic power generation system based on Markov chain Power Network Technology,2011,35(1):152-157

7. Li bo, W, Zheng ming, Z, Jian zheng, L, et al. Research on Stability of Maximum Power Point Tracking Algorithm for Single-stage Grid-connected PV Inverter System [J]. Proceedings of the CSEE,2006,26(6):73-77.

8. Zheng ming, Z, Yi, L, Fanbo, H, et al. Review of large capacity grid-connected photovoltaic power station technology [J]. Automation of Electric Power

Systems,2011,35(12):101-107.

9. 10. Yuan, A, Research on Theory and Method of Water and Light Complementary Coordinated Operation [D]. Xi 'an University of Technology, 2016:53-79

11. Lei, W, Research on Short-term Prediction Technology of Output Power of Photovoltaic Power Generation System [D]. Hefei University of Technology,2012

12. Ming, D, Lei, W, Rui, B. A Short-term Prediction Model for Output Power of Photovoltaic Power Generation System Based on Improved BP Neural Network [J]. Power System Protection and Control,2012,40(11):93-99.

13. Lei, W, Research on Short-term Prediction Technology of Output Power of Photovoltaic Power Generation System [D]. Hefei University of Technology,2012

14. Ming, D, Lei, W, Rui, B. A Short-term Prediction Model for Output Power of Photovoltaic Power Generation System Based on Improved BP Neural Network [J]. Power System Protection and Control,2012,40(11):93-99.

15. Fei, W, Zeng qiang, M, Yin xun, Y, et al. Prediction method of photovoltaic power generation based on neural network and associated data [J]. Acta Energiae Solari Sinica,2012,33(7): 1171-1177

16. 17. Yuan, A, Research on Theory and Method of Water and Light Complementary Coordinated Operation [D]. Xi 'an University of Technology, 2016:53-79

18. Mi, Y. Research on key technologies of grid-connection and parallel connection of photovoltaic power generation [D]. Huazhong University of Science and Technology,2009.

19. Guo quan, Q, Yan jun, X, Hong yi, Y. Optimal calculation of solar radiation model for clear days [J]. Acta Energica Sinica,2001,22(4):456-460.

20. Qin man,T, Research on maximum power tracking control method in photovoltaic power generation system [D]. Tianjin University,2008.

21. Cheng, C, Reservoir Probabilistic Constraint Scheduling Model and Method Based on Linear Programming [D]. Huazhong University of Science and Technology, 2019:15-69

22. Tran H D, Muttil N, Perera B J C.Selection of significant input variables for time series forecasting[J].Environmental Modeling & Software,2015,64:156-163