Short-term forecast methods of electricity generation by solar power plants and its classification.

D A Tyunkov, A S Gritsay, V I Potapov, R N Khamitov, A V Blohin, L K Kondratukova

Omsk State Technical University, 11 Mira ave., Omsk, 644050, Russia

Abstract. The article provides a classification of existing forecast models the generation of electricity by solar power plants and discusses various options for forecast methods for each of the selected models. As a result of the study, it was concluded that the most promising forecast methods are hybrid statistical-adaptive methods.

1. Introduction
One of the most promising and rapidly developing areas of renewable energy is solar generation of electrical energy. Worldwide, the pace of development of solar power plants since 2016 exceeds the pace of development of coal-fired power plants [1].

Currently, in the world of solar energy, short-term forecast of solar energy production does not have a proven and fully tested technology and is often associated with large errors, sometimes reaching 60–65%. At the same time, the forecast for a longer period gives a more accurate result, the same is true for the region – a more pointed forecast gives a less accurate result.

In fact, the forecast of the amount of electrical energy produced by a solar power plant is the forecast of the amount of solar radiation produced by solar panels. What is influenced by many factors, the main of which are meteorological and climatic conditions, the position of the sun in the sky, the duration of daylight hours, precipitation, cloudiness, wind strength, etc.

All currently existing forecast methods the generation of electric energy by solar power plants can be divided into four main groups presented in Figure 1 [1, 2]:

1. Physical models describe the physical relationship between weather conditions and solar radiation, obtained using numerical weather prediction, and the generation of electrical energy at the station.
2. Statistical models describe the relationship between the solar radiation flux density obtained using numerical weather prediction and the generation of electricity in a solar power plant by statistical analysis of time series of retrospective data without taking physical factors into account.
3. Adaptive models use artificial intelligence systems to determine the relationship between predicted weather conditions and power plant output.
4. Hybrid models in most cases are a combination of physical, statistical and adaptive models.
2. Numerical weather prediction

Numerical weather prediction models in recent years are often used to perform short-term meteorological variable predictions in many different applications. Typically, these models are classified according to the space-time characteristics of their weather forecasts. Each model of numerical weather prediction predicts atmospheric variables with a degree of quality depending on the geographical expansion and temporal resolution of their forecasts [3]. Thus, models with a high spatial scale receive predictions about shorter temporal confidence, while longer horizons of forecasts are achieved using models of numerical weather prediction of smaller spatial scale.

3. Physical models

The input variables for the physical forecast models are numerical weather prediction, local meteorological measurements, relief data and the type of the earth's surface, as well as historical data on the output power of a solar power plant. In addition, there is the possibility of using satellite systems to track the direction and speed of clouds, which provides the ability to predict solar radiation in real time.

The practical use of various physical models in forecast the generation of electric energy by solar power plants has been shown in [4–7]. In [4], the effect of the aerosol optical depth on the quality of solar radiation forecast using the MM5 or Eta numerical weather prediction models is considered. As a result, it was shown that when the value of the aerosol optical depth is less than 0.1, the forecast error is acceptable and is about 3–4%, and when the value of the aerosol optical depth is greater than 0.1, the average deviation is about 100 W/m².

In [5], one of the physical forecast models is presented, based on short-term Oktas-scale variations and temperature changes to determine the average hourly output power of photovoltaic systems in small solar power plants. This model has an acceptable accuracy in forecast the generation of solar energy in cloudy weather, while in solar it produces several times worse results.

4. Statistical models

Numerical weather prediction data, information on solar radiation and retrospective data on the generation of electrical energy by a solar power plant, are used as input data for statistical forecast models. Statistical models are most widely used for medium- and long-term forecast.

4.1. Regression methods

One of the widely used statistical methods for predicting the generation of electrical energy by solar power plants are regression methods [8]. They allow you to consider many factors affecting the forecast, for example, meteorological.
When using regression methods, it is possible to evaluate causal relationships and dependencies in the data. In addition, the advantage of these methods is that they forecast the value of the dependent variable by the values of independent, but it is necessary that the values of the signs be uncollectible. The linear regression equation has the following form

\[ Y = a + b_1x_1 + b_2x_2 + \ldots + b_nx_n + \varepsilon \]  

(1)

where \( Y \) — resulting sign; \( x_1, x_2, \ldots, x_n \) — factor signs; \( b_1, b_2, \ldots, b_n \) — regression coefficients; \( a \) — free term of the equation; \( \varepsilon \) — model “error”.

The advantage of regression methods is the simplicity of their implementation, and the main disadvantage is the unpredictability of parameters that affect the actual values. If sudden changes occur in the available historical data, there will be a violation of the forecast accuracy. The use of regression methods is well applicable for finding patterns in the data and for determining the significant factors of the model, but they do not provide high accuracy in constructing short-term models for predicting the generation of electricity.

4.2. Time series-based methods

One of the most commonly used time series-based methods is the Box-Jenkins model — ARIMA (autoregressive integrated moving average) [9]. This model is applied to non-stationary time series, reducible to stationary, by taking the difference of some order from the original values of the time series.

For a non-stationary time series, the ARIMA model has the following form

\[ \Delta^d X_t = c + \sum_{i=1}^p a_i \Delta^d X_{t-(i-d)} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t \]  

(2)

where \( \varepsilon_t \) — stationary time series; \( c, a_i, b_j \) — model parameters; \( \Delta^d \) — time difference operator of order \( d \) (sequential taking \( d \) times the differences of the first order — first from the time series, then from the received differences of the first order, then from the second order, etc.)

Also, this model is interpreted as \( ARMA(p + d, q) \) — model with \( d \) single roots. When \( d = 0 \), possible to use ARMA models only.

Methods based on the theory of time series are widely used in the construction of short-term forecasts of electricity generation, since enterprises have a large amount of historical data on the generation of electricity by solar power plants, and the methods in this group are aimed at processing large data arrays, and allow us to find patterns in them, as well as to use these patterns when building forecasting models. The main disadvantage of these methods is the failure to ensure the required forecast accuracy. But when using time series-based methods in a complex, for example with adaptive models, the required accuracy is achievable, but the methods used in this case will relate to hybrid models.

In the works [10–15] the practical application of statistical models to the prediction of solar energy is shown.

In [10], a two-stage method for forecasting electricity by a solar power plant is described. At the first stage, a statistical normalization of solar energy is performed using the clear sky model, the clear sky model was proposed in [16]. At the second stage, forecasts of normalized solar energy are calculated using time series-based models, in particular, an autoregressive model and an autoregressive with exogenous input model, which is supplied with a numerical weather forecast. As a result, it was shown that the use of an autoregressive with exogenous input model gives a 12% better result than using a simple autoregressive model in forecasting a very short horizon and 23% in predicting the next day.

In [11], a model for solar energy forecast was proposed that considers the stochasticity of clouds using various parameters that consider the attenuation of power. Based on the statistical behavior of the parameters, a simple process of switching between three classes was proposed: "sunny", "cloudy", "partly cloudy". The forecast is built by identifying the current mode and assuming it is saved in this mode.

In [12], the authors proposed a short-term forecast of solar energy using the classical concept of "seasonality", highlighting the average value or trend and rapid fluctuations around it. A feature of this work is a very short time horizon of 1, 15 and 60 minutes.
5. Adaptive models
Adaptive models have a “learning” process based on retrospective data analysis. These models, unlike statistical ones, can implicitly describe complex non-linear relationships between weather conditions and power generated by a solar power plant. The main factor affecting the accuracy of the forecast is the sample and structure of the source data used to build the model [17].

5.1. Methods using neural network
Methods using artificial neural networks have recently become widespread not only in solving problems of short-term forecast of solar energy generation, but also in medium- and long-term forecast. Artificial neural networks consist of a set of neurons of input, hidden and output layers interacting with each other. Neurons have an activation function that depends on the weights of the connections between neurons and displacement.
To obtain the forecast training of the neural network is required. In the process of learning, the values of displacements and weights are selected for each neuron so that the output signal of the neural network is as close as possible to the actual value. There are many neural network training methods [18] which have found application in predicting the generation of electricity by solar power plants. The advantage of neural networks are fast learning algorithms and the ability to work in the presence of noise input signals. Compliance with the requirements for constructing the structure of a neural network considering the redundancy of neurons which depends on the number and sample of informative features formed for training an artificial neural network ensures high reliability of such a network [19].

5.2. Support Vector Machine
Support vector machine belongs to the family of linear classifiers and is used in regression analysis and classification problems. The main idea of the method is the translation of vectors into a space of higher dimension and the search for a separating hyperplane with a maximum gap in it [20]. From the hyperplane separating classes on both sides, two parallel hyperplanes are constructed. A hyperplane, in which the distance between two parallel planes is maximal and will be separating. The main advantages of the support vector machine are the ability to obtain the correct solution of the problem in the presence of incomplete and distorted data, and the possibility of considering many additional factors affecting the quality of forecasting. The disadvantage is the need for training and increased requirements for software and hardware resources. The use of adaptive models to predict the generation of solar energy has been shown in [21–29].

In [21] demonstrate the use of decision trees for very short-term forecasting of the generation of electricity by a solar power plant. The work shows that the accuracy of the prediction using gradient boosting was 75–65%. The forecast is built 1 hour ahead.
In [22], the use of an artificial neural network, an adaptive neuro-fuzzy inference system and a generalized neural network is considered. The input parameters for the proposed model are the level of solar radiation, ambient temperature, wind speed and module temperature. As a result the work shows that the use of a generalized neural network gives the best result.
In [23], the authors use a radial basis functions network for forecasting solar energy. The input data are data on power measurements and meteorological forecasts of solar radiation, relative air humidity and temperature at the location. A feature of this work is a preliminary classification of the type of weather (sunny, cloudy, rainy) and the use of a different network structure for each class.
In [24] proposed the use of a deep belief network to forecast the generation of solar energy. As an input for the network data are used on the level of solar radiation, ambient temperature, relative humidity and historical data on the generation of solar energy five days before the forecast. The work also showed that wind speed has little effect on the generation of solar energy. As a result of the work it was shown that a model based on a deep belief network shows a better result than a model based on a neural network with back propagation of error.

6. Hybrid models
Hybrid methods combine various combinations of methods from other known groups. So, for example, often there are combinations of methods of physical, statistical and adaptive models. The use of physical models is not always justified due to the difficulty of accurately considering certain factors which requires the introduction of a statistical approach to determine them. Statistical approaches have higher accuracy when calculating averaged values of solar radiation over a long period (day, month, year), while over a shorter period (minutes, hours) when physical processes, such as cloudiness, cannot be averaged for a given interval they have much lower accuracy. For optimal accuracy of hybrid models the statistical model is adapted to constantly changing conditions which are described by physical models. Most of the combined approaches to forecast the generation of solar energy, as a rule, can be used either to calculate clear days or cloud days. Another example of hybrid models can be methods combining various methods from statistical and adaptive models. Such hybrid methods have prospects because they allow to consider the specifics of the physical process and use the capabilities of adaptive methods. The methods of this group are developing, and experts find various combinations of methods that provide the necessary accuracy. He uses of hybrid models for predicting the generation of electricity by solar power plants was shown in [30–36]. In [30], a hybrid physical-statistical model was used to predict the generation of electricity. Two parameters, the transparency coefficient and the diffusion coefficient, were determined using statistical models. All other parameters were determined using physical models. As a result of the application of the method the error was 22.3%.

In [31], a two-step model is used that combines an autoregressive integrated moving average, a least square support vector machine, an artificial neural network, and an adaptive neuro-fuzzy inference system with a genetic algorithm in the second stage. The method error as a result of the study was about 5.21%.

7. Conclusion

The paper presented the main classification of models for forecasting the generation of electric energy by solar power plants and described some of the basic methods of forecasting. In addition, various applications of forecasting methods in the works of other authors were considered. Average forecast error for various physical models is about 21–26%, statistical models – 20–24%, adaptive – 15–19%, hybrid physical-statistical models – 19–24%, statistical-adaptive – 5–10%. Based on this, we can conclude that the use of hybrid statistical-adaptive models gives the best forecasting results and is the most promising.

8. Reference

[1] Kochneva E 2018 MATEC Web of Conferences 208 04005
[2] Snegirev D A, Valiev R T, Eroshenko S A and Khalyasmaa, A I 2017 International Conference on ENERGY and ENVIRONMENT (CIEM) 349–353
[3] Al-Yahyai S, Charabi Y and Gastli A 2010 Renewable and Sustainable Energy Reviews 14(9) 3192–3198
[4] Zamora R J, Dutton E G, Trainer M, McKeen S A, Wilczak J M and Hou Y-T 2005 Monthly Weather Review 133(4) 783–792
[5] Gandoman F H, Abdel Aleem S H E, Omar N, Ahmadi A and Alenezi F Q 2018 Renewable Energy 123 793–805
[6] Wittmann M, Breitkreuz H, Schroedter-Homscheidt M and Eck M 2008 IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1(1) 18–27
[7] Lorenz E, Hurka J, Heinemann D and Beyer H G 2009 IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2(1) 2–10
[8] Abuella M and Chowdhury B 2015 Solar Power Probabilistic Forecasting by Using Multiple Linear Regression Analysis SoutheastCon 1–5
[9] Kaur H and Ahuja S 2017 Proceedings of Sixth International Conference on Soft Computing for Problem Solving 347–358
[10] Bacher P, Madsen H and Nielsen H A 2009 Solar Energy 83(10) 1772–1783

[11] Ramakrishna R; Scaglione A and Vittal V A 2016 Stochastic Model for Short-Term Probabilistic Forecast of Solar Photo-Voltaic Power Proceedings of the Asilomar Conference on Signals, Systems and Computers (Pacific Grove, CA, USA) 1–13

[12] Mora-López L, Martínez-Marchena I, Piliouigne M and Sidrach-de-Cardona M 2011 Advances in Intelligent Data Analysis X 294–305

[13] Mazorra-Aguiar L, and Díaz F 2018 Wind Field and Solar Radiation Characterization and Forecasting 171–200

[14] Sheng H, Xiao J, Cheng Y, Ni Q and Wang S 2018 IEEE Transactions on Industrial Electronics 65(1) 300–308

[15] Xiao C, Zhi-bao C, Yuyu D, Hai Z and Jie D 2012 AASRI Procedia 2 174–179

[16] Chowdhury B and Rahman S 1987 IEEE Photovoltaic Specialists Conference, 171–176

[17] Romanov A, Romanov M, and Kharchenko A 2017 Proceedings of Russian Young Researchers in Electrical and Electronic Engineering (EICON) 976–981

[18] Haykin S O 2008 Neural Networks and Learning Machines (3rd Edition) (London: Pearson) p. 936

[19] Potapov V I 2010 Omsk Scientific Bulletin 3(93) 202–203

[20] Statnikov A, Aliferis C F and Hardin D P 2011 A Gentle Introduction to Support Vector Machines in Biomedicine: Theory and methods (Singapore: World Scientific) p. 200

[21] Eroshenko S, Khalyasmaa A and Snegirev D 2018 E3S Web of Conferences 51 02004

[22] Singh V P, Vijay V, Siddhartha Bhatt M and Chaturvedi D K 2013 Proceedings of 13th International Conference on Environment and Electrical Engineering (EEEIC) 1–5

[23] Chen C, Duan S, Cai T and Liu B 2011 Solar Energy 85(11) 2856–2870

[24] Li L-L, Cheng P, Lin H-C and Dong H 2017 Advances in Mechanical Engineering 9(9) 1–13

[25] Earfan Ali M and Khondaker M E A 2018 Journal of Climatology & Weather Forecasting 6(3) 1000238

[26] Wang J, Ran R and Zhou Y 2017 Applied Sciences 7(4) 423

[27] Fentis A, Bahatti L, Mestari M and Chourri B 2017 Proceedings of 15th IEEE International New Circuits and Systems Conference (NEWCAS) 405–408

[28] Behera M K, Majumder I and Nayak N 2018 Engineering Science and Technology, an International Journal 21(3) 428–438

[29] Potapov V I, Gritsay A S, Tyunkov D A and Sinitzins G E 2016 Bulletin of the Tomsk Polytechnic University, Geo Assets Engineering 327(8) 44–51

[30] Snegirev D A, Eroshenko S A, Valiev R T and Khalyasmaa A I 2017 IEEE II International Conference on Control in Technical Systems (CTS) 228–231

[31] Wu Y-K, Chen C-R and Abdul Rahman H 2014 International Journal of Photoenergy 2014 1–9

[32] Filipe J M, Bessa R J, Suamiall J, Tome R and Sousa J N 2015 Proceedings of 18th International Conference on Intelligent System Application to Power Systems (ISAP) 1–6

[33] Abedinia O, Amjad Y and Ghadimi N 2017 Computational Intelligence 34(1) 241–260

[34] Khan I, Zhu H, Yao J, Khan D and Iqbal T 2017 International Journal of Photoenergy 1–9

[35] Khamitov R N, Gritsay A S, Tyunkov D A and Sinitzins G E 2017 IOP Conference Series: Materials Science and Engineering 189 012005

[36] Potapov V, Khamitov R, Makarov V, Gritsay A, Florensot A and Tyunkov D 2018 Proceedings of Dynamics of Systems, Mechanisms and Machines (Dynamics) 1–5