Machine learning techniques for short-term solar power stations operational mode planning

Stanislav Eroshenko[1], Alexandra Khalyasmaa[1,*] and Denis Snegirev[1]

[1]Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, 620002 Mira str. 19, Russia

Abstract. The paper presents the operational model of very-short term solar power stations (SPS) generation forecasting developed by the authors, based on weather information and built into the existing software product as a separate module for SPS operational forecasting. It was revealed that one of the optimal mathematical methods for SPS generation operational forecasting is gradient boosting on decision trees. The paper describes the basic principles of operational forecasting based on the boosting of decision trees, the main advantages and disadvantages of implementing this algorithm. Moreover, this paper presents an example of this algorithm implementation being analyzed using the example of data analysis and forecasting the generation of the existing SPS.

1 Introduction

At present, machine-learning algorithms that use decision trees are very common and universal for most applications. They allow predicting the real response for each object, that is, solving the regression problem. Composing several derived trees, combining the responses of each of them, it is possible to get a stable and much more qualitative solution than many other algorithms can provide.

Forecasting the generation of a solar power station (SPS), both short-term (day ahead) and operational (one hour ahead), consists of five main stages:

- Determining the solar radiation flux density (SRFD) at the boundary of the atmosphere;
- Determining the SRFD incident on the horizontal surface of the earth;
- Determining the SRFD incident on the inclined panel surface;
- Determining the generation of a photoelectric DC converter;
- Determining the power output of AC inverters.

The most difficult part of the SPS generation forecasting process is the second stage, which is the determination of SRFD incident on the horizontal surface of the earth, since this value depends on a variety of unstable and difficult for forecasting factors, the greatest influence of which is exerted by cloud cover.

In case of short-term forecasting, the dependence of SRFD incident on the horizontal surface of the earth on cloud cover can be restored by solving the regression problem based on least-squares method (or other similar method) and choosing the optimal function [1,2]. The use of such traditional methods to solve the regression problem for operational forecasting of SRFD incident on the horizontal surface of the earth is not possible, since in the case of variable clouds it is not possible to restore the relationship between SRFD and the current measurement data, i.e. variables that allow describing this dependence.

2 SPS operational forecasting model

2.1 Problem formulation

In this research the task of SRFD operational forecasting using retrospective data was set, based on which the output power of the SPS was calculated [3]. There is a learning sample S :

$$S = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{1a} & y_1 \\ x_{12} & x_{22} & \cdots & x_{2a} & y_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{l1} & x_{l2} & \cdots & x_{la} & y_l \end{bmatrix}, \qquad (1)$$

where y_j is the parameter, which is the measured value of SPFD within the task; x_{ij} is the attribute matching the parameter y_i (within the task, attributes can be both calculated values, solar declination angle δ , SRFD on the boundary of atmosphere $\overline{G}_0$, and retrospective parameters, such as transparency factor $\overline{k}_T$ as well as weather data); l is the number of observations in the sample; a is the number of attributes.

The task consisted in composing the decision trees, which with great accuracy will determine the value of the new parameters y according to the relevant attributes x_{ij} [4,5]; in other words, the task consisted in development of a *model* (function) f that, having received x on input,

[*]Corresponding author: lkhalyasmaa@mail.ru

would predict the value of the response y . In this research, a gradient-boosting algorithm was used with decision trees.

2.2 Gradient boosting

Gradient boosting is essentially a gradient descent in the space of all possible algorithms [6]; each step of this descent is done by the basic algorithm according to the following sequence of actions $b_n(x)$ according to the following sequence of actions.

- Initializing of the first basic algorithm b_0

$$b_0(x) = 0. \tag{2}$$

- For $n = 1, \ldots, N$ the following steps are repeated:
  1. The shift vector s , which shows how to correct the predictions of the composition already constructed to reduce the error in the learning sample, is calculated:

$$s_n = \left(-2(a_{n-1}(x_1) - \mathrm{y}_1), \ldots, -2(a_{n-1}(x_l) - y_l)\right) \tag{3}$$

  2. A basic algorithm b_n is constructed by approximating its responses to the learning sample to the shift data s_i :

$$b_N(x) = \arg\min_{b} \frac{1}{l}\sum_{i=1}^{l}\left(b(x_i) - s_i\right)^2 = \sum_{j=1}^{J}\left[x \in R_{Nj}\right]b_{Nj} \tag{4}$$

  3. After the algorithm is found, it is added to the composition:

$$a_n(x) = a_{n-1}(x) + \eta\sum_{j=1}^{J}\left[x \in R_{Nj}\right]b_{Nj} \tag{5}$$

  4. Steps 1-3 are repeated until a stop criterion, for example, a specified number of iterations, is satisfied.

The settings shown in Table 1 were used to work with the gradient-boosting algorithm. The *colsample_bytree* and *colsample_bylevel* settings were used to avoid overfit. To solve the problem, authors developed an approach based on gradient boosting on decision trees. The task was proposed to be solved in three ways: teaching an algorithm without history, teaching only with history and teaching with history and weather data.

Computational experiments were performed to evaluate the possibility of using the gradient-boosting algorithm on solution trees for the task of SPS generation operational forecasting within the research work. Experiments varied according to the following criteria:

- Teaching without history

The following calculated values were used as attributes:

1. Number of a day in a year n ;
2. Solar declination angle δ , degrees;
3. Solar time t_s , hours;
4. SRFD at the boundary of the atmosphere G_0 , W/m^2.

- Teaching with history

As attributes, the calculated values and the data obtained by measurements were used:

1. Transparency factor $k_T = \frac{\overline{G}_m}{\overline{G}_0}$, p.u.;
2. Number of a day in a year n ;
3. Solar declination angle δ , degrees;
4. Solar time t_s , hours;
5. SRFD at the boundary of the atmosphere G_0 , W/m2.

- Teaching with history and weather data

As attributes, the calculated values, the data obtained by measurements, as well as actual weather data were used:

1. Transparency factor $k_T = \frac{\overline{G}_m}{\overline{G}_0}$, p.u.;
2. Average temperature per hour temp , K;
3. Maximum temperature per hour temp_max , K;
4. Minimum temperature per hour temp_min , K;
5. Average pressure per hour pressure , hPa;
6. Average humidity per hour humidity , %;
7. Average wind speed per hour wind_speed , m/s;
8. Average wind direction per hour wind_deg , degrees;
9. Precipitation in the last 3 hours rain_3h , mm;
10. Average cloud cover per hour clouds_all , %;
11. Weather identifier (presence or absence of cloud cover), weather_id ;
12. Number of a day in a year n ;
13. Solar declination angle δ , degrees;
14. Solar time t_s , hours;
15. SRFD at the boundary of the atmosphere G_0 , W/m2.

Table 1. Gradient boosting configuration

Configuration options	Types of input data		
	Teaching without history	Teaching with history	Teaching with history and weather data
Step length (*learning rate*) η	0.005	0.01	0.01
Trees maximum depth (*max_depth*)	4	4	6
Number of trees (*n_estimators*)	3000	3000	3000
Number of iterations (*n_boost_round*) N	3000	3000	3000
Share of variables used at each iteration (*colsample_bytree*)	0.8	0.8	0.8
Share of variables used at each level of the tree (*colsample_bylevel*)	0.7	0.7	0.7
Threshold value $\overline{G}_{thres}$	10	10	10

2.3 Developed model evaluation

A cross-validation was used to evaluate the analytical model and its behavior on independent data. The available data were divided into 5 parts for the model evaluation. The algorithm for the learning subsample is configured for each partition; then its mean error at the objects of the control subsample was estimated. The cross-validation estimation was the error average for all partitions on the control subsamples [6].

To analyze the prediction error, the mean absolute percentage error MAPE and the mean-square error (MSE) were used [7,8]. The mean absolute percentage error (MAPE) was calculated by the formula:

$$MAPE = \frac{1}{l}\sum_{i=1}^{l}\frac{\left|\overline{G}_m^i - \overline{G}_f^i\right|}{\overline{G}_m^i}\cdot 100\%, \qquad (6)$$

where *MAPE* is the mean absolute percentage error, [%]; l is the number of objects in the sample; $\overline{G}^m$ is the average measured value of the SRFD near the ground, [W/m^2]; $\overline{G}^f$ is the average forecast value of the SRFD near the ground, [W/m^2]. *MAPE* does not allow adequately estimating the error in the range of small values $\overline{G}_m$ so a threshold value $\overline{G}_m > 100$ is used to calculate *MAPE*.

Mean-square error (MSE) is calculated by the formula:

$$MSE = \frac{1}{l}\sum_{i=1}^{N}\left(\overline{G}_m^i - \overline{G}_f^i\right)^2, \qquad (7)$$

where *MSE* is the mean-square error, [W^2/m^4]; N is the number of time intervals t in the day considered (between sunrise and sunset); $\overline{G}^m$ is the average measured value of the SRFD near the ground, [W/m^2]; $\overline{G}^f$ is the average forecast value of the SRFD near the ground, [W/m^2].

3 Calculation example of SPS operational forecasting

The results of the operational forecasting correction for the SRFD incident on the horizontal surface of the earth $\overline{G}$ for all three teaching modes in case of the 1-hour forecast horizon are shown in Fig. 1 - 2.

Parameters of the forecast error for considered days are presented in the Table 2.

It is seen from the Figures 1 and 2 that in case of teaching without history (without using current measurement data) the algorithm allows determining the relationship between SRFD at the boundary of the atmosphere $\overline{G}_0$ and SRFD incident on the horizontal surface of the earth $\overline{G}$ without considering cloud cover. This teaching mode physically takes into account only the change in the solar altitude angle during the day, as well as factors that are generally constant (solar absorption by ozone, dust, air molecules, water).

In case of teaching with history, when the nearest available data of current measurements were used for the prediction, the algorithm allows to approximate the forecast value $\overline{G}_f$ to the measured values $\overline{G}_m$. For the case of slightly variable clouds on 28.05.2017, the algorithm allows to correct the forecast even without reliance on weather data. For the case of sharply variable clouds on 27.05.2017, an accurate correction was not possible; the error for 8 of 16 hours was 50% or more.

The operational forecast correction for such days necessitates the use of reliable weather data. A decrease in the prediction error is observed for both days considered in case of teaching with history and weather data, even though low quality weather data were used to teach the algorithm.

Table 3 shows the parameters of the forecast error for the entire sample considered. This data is from 14.01.2017 to 28.05.2017. It can be seen from the table that the proposed algorithm has a lower average error for the 1-hour forecast horizon within the entire considered time interval (4.5 months) than for a single day. This fact points at the significant accuracy of operational forecasting and the success of the proposed algorithm application for most of the considered days.

From Table 3 it is also clear that the use of low quality weather data provided only a slight error reduction compared with teaching only with history. Obviously, using more accurate weather data will result in making more accurate operational forecasts.

The main factor affecting the error of operational forecast is the remoteness of the forecast horizon. Figures 3 and 4 show the dependence of the mean-square error *MSE* and the mean absolute percentage error *MAPE* on the magnitude of the forecast horizon. It can be seen from the diagrams that the accuracy increases as the forecast horizon approaches, since the current measurements are more informative (better describe the predicted value).

From the analysis of the results obtained, it can be concluded that it is possible to apply the gradient-boosting algorithm on solution trees in the context of the task of SPS generation operational forecasting.

Table 2. Forecast error parameters ($H = 1\ hour$)

Day	Type of input data					
	Without history		With history		With history and weather data	
	MAPE, %	*MSE*, (W/m^2)^2	*MAPE*, %	*MSE*, (W/m^2)^2	*MAPE*, %	*MSE*, (W/m^2)^2
27.05.17	91,01	47514,50	59,03	24835,58	49,89	24461,12
28.05.17	26,56	10056,30	22,96	9224,44	18,46	9720,42

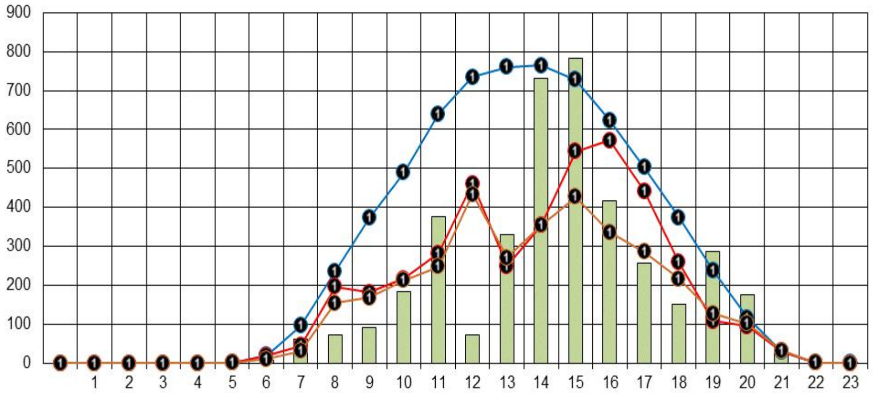

Figure 1. 1-hour ahead operational forecast for 27.05.2017 (blue line – without history, red line – with history, orange line – history and weather)

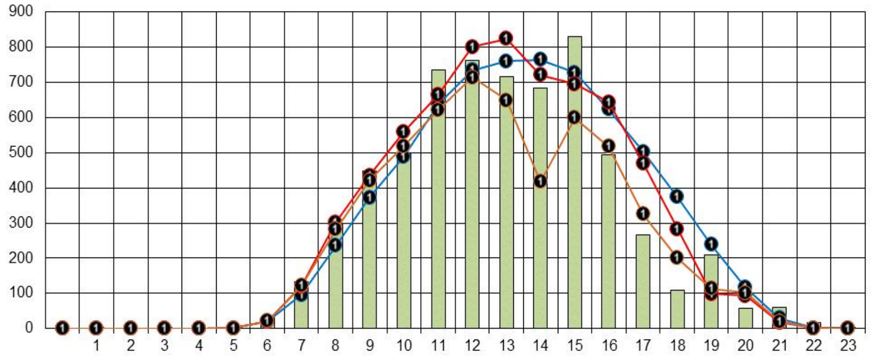

Figure 2. 1-hour ahead operational forecast for 28.05.2017 (blue line – without history, red line – with history, orange line – history and weather)

Table 3. Forecast error parameters for the entire sample considered

Forecast horizon H	Types of input data					
	Without history		With history		With history and weather data	
	MAPE, %	*MSE*, (W/m^2)^2	*MAPE*, %	*MSE*, (W/m^2)^2	*MAPE*, %	*MSE*, (W/m^2)^2
1	45.84	15950.74	21.06	3651.02	19.88	3363.92
2	45.84	15950.74	28.23	6092.28	24.94	4904.88
3	45.84	15950.74	32.77	8188.28	28.09	5908.13
4	45.84	15950.74	36.77	10476.08	30.63	6627.15
5	45.84	15950.74	40.25	12644.29	32.46	7474.96
6	45.84	15950.74	43.11	14272.88	35.01	8896.09
7	45.84	15950.74	44.82	15489.63	38.07	10589.30
8	45.84	15950.74	45.80	16235.96	39.60	11424.11
9	45.84	15950.74	46.48	16630.79	40.82	11943.67
10	45.84	15950.74	46.34	16556.05	41.48	12665.08
11	45.84	15950.74	46.78	16852.75	41.19	12596.49
12	45.84	15950.74	46.46	16733.06	42.63	13686.00
13	45.84	15950.74	46.64	17000.71	43.40	14260.52
14	45.84	15950.74	46.36	16646.83	44.39	15398.75
15	45.84	15950.74	45.73	16042.63	44.15	15073.16
16	45.84	15950.74	45.61	15718.92	44.47	15030.38
17	45.84	15950.74	45.22	15446.75	45.45	15741.26
18	45.84	15950.74	45.75	15794.79	45.19	15567.77
19	45.84	15950.74	45.52	15842.28	46.40	16543.24
20	45.84	15950.74	45.41	15923.70	47.10	17016.06
21	45.84	15950.74	45.50	16186.40	46.43	16885.56
22	45.84	15950.74	45.81	16655.97	46.38	17081.73
23	45.84	15950.74	46.20	17348.61	45.74	17143.63
24	45.84	15950.74	46.47	17483.31	45.53	17396.71

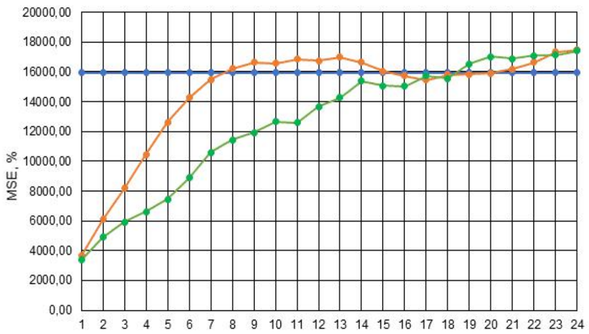

Figure 3. MSE dependence on the forecast horizon value (blue line – without history, orange line – with history, green line – history and weather)

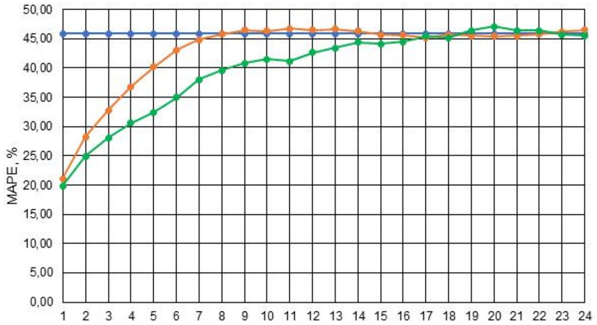

Figure 4. MAPE dependence on the forecast horizon value (blue line – without history, orange line – with history, green line – history and weather)

4 Conclusion

Within the presented research, the analysis of gradient boosting application on decision trees in the task of SPS generation operational forecasting is performed.

To improve the accuracy of generation forecasting based on the data of a local meteorological station, the authors developed a new mathematical model of operational forecasting for software implementation as part of the existing software package. Mathematical model of operational generation forecasting is implemented based on decision trees.

Within this research, test calculations of the operational generation forecast for one of the existing SPS in Russia in the form of a Python program in the Jupyter software environment were performed. In the course of the research, the accuracy of the SPS generation forecast for the operational forecast model was 75-65%, which allowed increasing the forecast accuracy by 20% in comparison with similar approaches to the SPS generation forecasting.

References

1. V. Margaret, J. Jose. Exponential Smoothing Models for Prediction of Solar Irradiance. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 2, February 2015.
2. D.A. Snegirev, S.A. Eroshenko, R.T. Valiev, A.I. Khalyasmaa. Algorithmic realization of short-term solar power plant output forecasting. Proceedings of 2017 IEEE 2nd International Conference on Control in Technical Systems, CTS 2017. Pp. 228-231.
3. C. Severiano, F. G. Guimarães and M. W. Cohen, "Very short-term solar forecasting using multi-agent system based on Extreme Learning Machines and data clustering," 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, 2016, pp. 1-8.
4. C.A. Severiano, P. C. L. Silva, H. J. Sadaei and F. G. Guimarães, "Very short-term solar forecasting using fuzzy time series," 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Naples, 2017, pp. 1-6.
5. A.D. Orjuela-Cañón, J. Hernández and C. R. Rivero, "Very short term forecasting in global solar irradiance using linear and nonlinear models," IEEE Workshop on Power Electronics and Power Quality Applications (PEPQA), Bogota, 2017, pp. 1 -5.
6. S. Gupta, N. A. Shrivastava, A. Khosravi and B. K. Panigrahi, "Wind ramp event prediction with parallelized gradient boosted regression trees," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 5296-5301.
7. W. Glassley, J. Kleissl. Current state of the art in solar forecasting. California Renewable Energy Collaborative Final Report. 2013.
8. V. Prema. K. Uma Rao. Development of statistical time series models for solar power prediction // Renewable energy. 2015.