Research into various modeling methods including the artificial neural networks in studying solar radiation for different regions in the Russian Federation

Larisa Haritonova
Volgograd State Technical University, 28, Lenin Avenue, 400005, Volgograd, Russia
E-mail: haritonova410@yandex.ru

Abstract. The comparison of results obtained using the artificial neural networks (ANN) has been carried out in researching solar radiation for various regions of the Russian Federation with data modeling applying a spline interpolation preserving an interpolant shape or a polynomial (up to tenth degree) in the MATLAB environment obtained when using the MATLAB Basic Fitting - a graphical user interface (GUI). The neural network has been trained by the feed-forward backprop algorithm. The following functions were used: the Bayesian Regularization, the function of gradient descent with regard for moments, the hyperbolic tangent function. The Mean Square Error (MSE) was chosen as a loss function which was minimized. The 15 input parameters were taken. As a result, the obtained model had high values of correlation coefficients and low values of the Mean Square Error among the target and the output values. The polynomial (up to tenth degree inclusive) coefficients and the norm of the residuals were obtained which were decreasing with the increasing of the polynomial degree. It was shown that the better results are received using the artificial neural networks (ANN), somewhat worse ones employing a spline interpolant and shape-preserving interpolant. At the same time, the proposed models and methods can be used in calculation of solar radiation for various regions in the Russian Federation and other countries.

1 Introduction, problem statement and its connection with relevant scientific and practical tasks; analysis of recent achievements and publications where the solution of the problem was begun, highlighting unsolved parts of the general problem; definition of the goals and statement of problems layout age layout

One of the major directions of the research and technology advancement today is the artificial intellect [1], – in particular, the neural networks. The artificial neural network (ANN) represents a mathematical model built on the principle of organization and functioning of biological neural networks. There are attempts to use the ANN in various fields of science and technology [2-14].

The studies [2-3] try to solve the problem of environmental safety of territorial entities using Fuzzy Neural Networks. This enabled to overcome the lack of available information at the input and to assess the environmental safety. The Fuzzy Neural Networks are tolerant to the contradictions of the applied data, enable to use fuzzy parameters, the unknown or approximately known data. The article tries to predict the development changes of the Russian subjects in terms of ecological safety taking into account the environmental factors.
In the study [4] the neural network models were designed and trained to calculate the concentration of fine dust PM2.5 and PM10 in the air. As inputs of the neural networks, the weather conditions and the TSP – the total concentration of dust were used. The average errors were found with the help of models involving the relationship between the PM2.5 and PM10 concentrations and the absence of such one.

In the further study of the same authors [5] the results of research of artificial neural networks are presented to calculate the PM2.5 and PM10 concentrations of micro-particles based on the results of measuring the set of meteorological parameters and the total number of suspended particles. The article [6] presents a model of the neural network for a complex evaluation of the drinking water safety for consumers. The paper [7] shows comparison of the artificial neural networks with non-parametric regression algorithm known as the multivariate adaptive regression splines for the best approximation of the relationship between the input and output from the data-sets recorded from onboard studies of the Carbon Monoxide emissions by vehicles.

The comparison of results is of interests which were obtained when using artificial networks with data modeling applying, for instance, a spline interpolation with a shape-preserving interpolant or a polynomial of various degree including for the solar radiation analysis.

2 Presentation of the basic study material with the complete proof of the obtained scientific results

The Neural Network Toolbox (Deep Learning Toolbox) - it is a MATLAB extension package containing means of design, development and visualization for neural networks. The neural network under study was trained by the feed-forward backprop algorithm. The Bayesian Regularization, the function of gradient descent with regard for moments, the hyperbolic tangent functions were applied. The Mean Square Error had been taken as a loss function which was minimized.

The following 15 input variables were used:
- months (January through December (and yearly));
- the solar radiation volume (KWh/m²) [15] per
- vertical panel; horizontal panel; the panel rotation around the polar axis; the longitude of each city; the latitude of the cities location;
- the 50 degree panel slope in Vladivostok and Petropavlovsk - Kamchatsky; the 45 degree one in Yuzhno-Sakhalinsk and Petrozavodsk; the 40 degree slope in Moscow; the 35 degree one in the cities of Sochi and Astrakhan;
- average monthly temperature, (°C); air pressure; wind (average); number of cloudy days a month (per year); number of days with precipitation; the number of clear days in a corresponding month; the maximum and minimum temperatures in a corresponding month, (°C).

The Matrix was taken as follows 15× 364 – for the input parameters. The architecture of the neural network was in the following form: 15-15-1-1. As a result, the obtained model had high values of correlation coefficients and low values of the Mean Square Error between the target (Figure 1) and the predicted values (the output) (Figure 2).
Figure 1. The target graph.

Figure 2. Comparison of the target and the output data obtained with the artificial neural network.

The MATLAB Basic Fitting (a graphical user interface (GUI)) was also used for the following tasks in the MATLAB environment: Model data using a spline interpolant (Figure 3), a shape-preserving interpolant (Figure 4), or a polynomial (up to tenth degree) (Figure 5).
As follows from the presented figures, modeling using a spline interpolation leads to the appearance of additional “peaks” on the target graph, which were absent on the graph obtained using neural networks.

It is usually recommended to use a polynomial with the least degree that fits well. It is explained by the fact that although a polynomial with a higher degree can correspond to the initial data, however, this correspondence at the points between the initial data can behave erratically. At that, \( M \) of points when selecting a polynomial of a higher degree than \((M+1)\), the system becomes undetermined. Then the extraneous coefficients are set to 0 during the calculation.

The following equations for the polynomials up to the 10th degree were obtained:

\[
y = -23.021 \cdot z^4 + 20.703 \cdot z^5 + 83.087 \cdot z^6 - 27.389 \cdot z + 174.06
\]
Figure 5. Comparison of the target with the data modeling results using a spline interpolant, a shape-preserving interpolant, a polynomial of the 10th, 9th and 4th degree.

where $z = (x-182.5)/105.22$.

The norm of the residuals was as follows for:
- polynomials of the 4th degree: 6902.9;
- polynomials of the 9th degree: 6862.4;
- polynomials of the 10th degree: 6856.1.

As follows from the above given results, the norm of the residuals is decreasing with the increasing of the polynomial degree. However, the lesser value of the norm does not guarantee a more useful fit. Thus, the norm is not always a good test for the over-fitting the data.

When using only a part of the general matrix (only $1 \times 13$ values out of $15 \times 364$), the obtained results with a spline interpolant, a shape-preserving interpolant or polynomials (up to tenth degree) are considerably closer to the target (except for polynomials of the 4th degree) (Figure 6).
Figure 6. Comparison of the target with the data modeling results using a spline interpolant, a shape-preserving interpolant, a polynomial of the 10th, 9th and 4th degree, when using only a part of the general matrix (only $1 \times 13$ values out of $15 \times 364$).

At that, the following equations for the polynomials up to the 10th degree were obtained:

$$y = \frac{299.6 \cdot z^4 + 277.28 \cdot z^3 - 544.99 \cdot z^2 - 299.8 \cdot z + 258.15}{3.894}$$

$$y = \frac{109.3 \cdot z^9 + 136.73 \cdot z^8 - 364.59 \cdot z^7 - 412.83 \cdot z^6 + 371.24 \cdot z^5 + 408.16 \cdot z^4 - 80.064 \cdot z^3 - 260.9 \cdot z^2 - 60.927 \cdot z + 201.76}{3.894}$$

where $z = (x-7)/3.894$.

The norm of the residuals amounted for:
- polynomials of the 4th degree: 465.56;
- polynomials of the 9th degree: 20.779;
- polynomials of the 10th degree: 8.0622.

However, the use of the complete data set yields rather slightly different results (Figure 5).

3 Conclusions on the given study and perspectives of the further development of this direction

The study presents a comparison of the ANN models with the data modeling using a spline interpolant, a shape-preserving interpolant or a polynomials (up to tenth degree) in researching solar radiation for various regions of the Russian Federation. The models approximated the non-linear link of the solar radiation which was a function of 15 variables. The neural network with the architecture 15-15-1-1 trained and tested by the Bayesian Regularization, the function of gradient descent with regard for moments, the hyperbolic tangent functions among all the tested neural networks has been found as the optimum one because of the higher values of correlation coefficients and the lower MSE. Although a spline interpolant, a shape-preserving interpolant models showed fairly good results, the ANN model demonstrated a far higher accuracy. Nevertheless, the proposed approaches and methods can contribute to the improvement of calculations for solar radiation both in various regions in the Russian Federation and in other regions of the world.
In future, it seems perspective to increase the number of variables impacting the solar radiation as well as comparison and a more detailed study of other mathematical models of the process under consideration.

References
[1] Russian Long-Term S & T Foresight 2030.
[2] Glinskiy V, Serga L, Khvan M, Zaikov K 2016 Procedia CIRP. 40 615–619.
[3] Glinskiy V, Serga L, Khvan M, Zaikov K 2015 Abstracts 13th Global Conf. of Sustainable Manufacturing (Ho Chi Minh City, Binh Duong Vietnam) 64.
[4] Tunakova Y A, Novikova S V, Shagidullin A R, Kuznetsova O N 2017 Vestnik tehnologicheskogo universiteta. Publishing House: Kazan National Research Technological University 20 (17) 134-138.
[5] Tunakova Y A, Novikova S V, Shagidullin A R, Shagidullin R R 2018 Rossijskij zhurnal prikladnoj ekologii (Publishing House: Institute of Problems of Environment and Mineral Resource Management at the Academy of Sciences in the Republic of Tatarstan) 1 32-37.
[6] Tunakova Y, Novikova S, Gabdrakhmanova G, Krasnyuk I, Faizullin R 2018 Bionanoscience 8 (2) 504-510.
[7] Oduroa S D, Metiaa S, Ducb H and Haa Q P 2015 FFACE-ISARC15-3023378.
[8] Jimmy W 2016 Key, Control Engineering Russia.
[9] Haritonova L 2018 Advances in intelligent systems and computing 692 700-709.
[10] Haritonova L 2018 International Journal of Engineering & Technology 7 (2.23) (Science Publication Corporation SPC) 440-442.
[11] Haritonova L 2017 Procedia Engineering 206 1002–1008.
[12] Haritonova L 2018 E3S Web of Conferences 33 02072207 (EDP Sciences, Les Ulis, France).
[13] Haritonova L 2019 Neural Modeling of the Energy Efficiency Factor for Recuperators (Heat Exchangers) Using Impinging Jets for the Metallurgical Productions and Machine Building (Book Chapter) Lecture Notes in Mechanical Engineering. Book series Issue 9783319956299, 1269-1278 (Springer International Publishing).
[14] Solar batteries of the Russian manufacturer "SOLBAT" 2018. Information on http://www.solbat.su/fotosis/general/ date of treatment 9.04.2019