Neural modelling of solar radiation variability

T Ronkiewicz¹, J Aleksiejuk-Gawron¹, M Awtoniuk¹ and J Kurek²

¹ Warsaw University of Life Sciences, Institute of Mechanical Engineering, Nowoursynowska Street 164, 02-787 Warsaw, Poland
² Warsaw University of Life Sciences, Institute of Information Technology, Nowoursynowska Street 159, 02-787 Warsaw, Poland

joanna_aleksiejuk@sggw.edu.pl

Abstract. Artificial neural networks are increasingly used in engineering and technical sciences, especially to solve problems under process uncertainty. The mathematical model presented in this article describes cloud variability. The application of the model can increase the efficiency of solar systems because the response time of the solar panel to changing weather conditions is crucial. The model involves an artificial neural network that serves to determine the degree of daily cloud coverage based on three data – the month, daily solar radiation sum and total harmonic distortion factor (THD). The THD factor is determined for daily solar radiation courses using a Fast Fourier Transform. Approaching the daily variability of solar radiation as a sine wave allows employing the THD factor in an unconventional and innovative way. The modelling data have been derived from the measurements of the meteorological station of the Institute of Mechanical Engineering of the Warsaw University of Life Sciences. MATLAB Software (2019a) was used for data processing and network modelling. The model is verified using the mean square error. The performed analysis provides promising results and conclusions.

1. Introduction

Modelling of objects using artificial neural networks (ANN) is used across a number of research areas including biomedical engineering, electrical engineering, mechanical engineering, materials engineering, or nutrition and food technology [1-5]. In the solar energy area, neural networks are most often used to model the efficiency of renewable energy systems [6], as well as for the individual elements of the systems, such as solar collectors [7,8], or increasingly popular in recent times, photovoltaic panels [9,10]. Also, in the literature on the subject, references can be found to forecasting techniques related to solar radiation and its impact on the design of solar systems [11-12].

More elaborate models are employed to predict solar radiation as the main impact on solar systems efficiency. In the literature, isotropic or anisotropic models of solar radiation are described [13]; however, they are quite general. In most cases, solar radiation modelling using the ANN method is associated with forecasting the daily course of total solar radiation [14] or, less frequently, diffuse solar radiation [15-16]. To this end, a range of ANN models is available: multilayer perceptron (MLP), adaptive neuro-fuzzy inference system (ANFIS), nonlinear autoregressive recurrent exogenous neural network (NARX), and generalised regression neural networks (GRNN). In the majority of the discussed models, the results relate to a specific region or country [17-18] but may be more general and also apply to forecast monthly mean daily global solar radiation [19]. The most frequently considered model variables (input data) are atmospheric conditions, such as outside temperature, sunshine duration,
humidity, or wind speed [20]. Model input data are usually built on the results from meteorological station measurements or can be alternatively determined analytically [21]. Less often, a particular matter level (PM) can be given as a model input [22]. Hybrid models that connect neural networks with other methods are also used for daily solar radiation forecasting, e.g. hybrid method coupling artificial neural network (ANN) and simulated annealing (SA), known as ANN/SA [23], whereas in other hybrid models, a convolutional neural network is used exclusively to extract features of solar radiation [24].

In this article, the output parameter is the type of day associated with the degree of cloud coverage. As input, in addition to meteorological data, the total harmonic distortion factor (THD) is presented. Our innovative proposal is to use the THD factor, which is typically found in electrical engineering to characterise the variability of solar radiation.

2. Material and methods

The neural network was created using MATLAB (R2019a) and the Neural Network Pattern Recognition app, which is part of Deep Learning Toolbox 12.1. The block diagram of the model is shown in Figure 1. The structure of the constructed neural network is shown in Figure 2.

The neural model contains an input layer, a hidden layer, and an output layer. Each layer is represented by a given number of neurons. The number of neurons in the input and output layers is directly related to the input and output signals of the model. The number of hidden neurons was determined experimentally, by evaluating the functioning of the network. The learning was performed using scaled conjugate gradient backpropagation [25]. There was no maximum number of training epochs – the network stopped the training process when the generalisation ceased to increase. The network was set up so that 80% of the data was used to train the network, 15% for assessment and 5% for testing. The network was then re-tested on the entire set of data and the obtained values were used in quality analysis. The procedure was repeated 10 times for each model and the average results were considered. The network quality analysis was assessed using two values, cross-entropy and the percentage of incorrect qualifications: the lower the cross-entropy and the percentage of misqualification, the higher the fit of the model.

The input data consist of three coefficients, further defined as subsequent factors. The measurement data have been obtained from the solar radiation research laboratory of the Fundamental Engineering and Energy Department at the Institute of Mechanical Engineering, Warsaw University of Life Sciences. The research laboratory consists of a meteorological station located at geographical coordinates 21°03'12" E, 52°09’37” N, at an altitude of 130 m above sea level, and the measured quantity was the intensity of solar radiation. The solar radiation data is archived in daily files every 60 seconds (total of
1440 data/day). The intensity of solar radiation is measured in W/m² with the Kipp&Zonen CM11 pyranometer, the highest-class device complying with the international norm, ISO 9060.

The first quantity fed into the network input is the daily total irradiation. It was calculated from direct measurement. A good practice used in the construction of neural networks is the optimisation of its coefficients. It is much easier for the network to analyse the influence of individual inputs on the desired result when all inputs take values from a similar range, rather than when one input is ranged e.g. between 0 and 1 and the other one 1000 and 5000. Therefore, from the available data, it could be seen that the maximum total daily solar radiation sum did not exceed 8.5 kWh/m². Therefore, the first \( I_{24h} \) factor (daily solar irradiance factor) is presented by the formula:

\[
I_{24h} = \frac{\sum_{i=1}^{1440} I_{min_i} \cdot 60 \text{ min}: 1000 \text{ } [\text{kWh}] }{8.5 \text{ } [\text{kWh}] / \text{m}^2}
\]  

(1)

The second quantity fed into the network input is the total harmonic distortion factor (THD). This factor is the ratio of the effective value of higher harmonics of the signal to the value of the basic component measured for a sinusoidal signal. It is used for testing disturbance levels in different types of waves and is found in various scientific fields such as acoustics or power engineering. Defining the course of daily solar radiation as a sine wave signal allows employing this factor in the research presented in the article. Subsequent harmonics were determined using a Fast Fourier Transform (FFT). To calculate the THD factor, a formula was used that does not take into account the first and second harmonics. Subsequent harmonics represent an ever-smaller period. The first harmonic is related to the season of the year and the second to the day-night cycle [23]. These harmonics do not result from the disturbances of the radiation course and are thereby omitted. The formula for the THD factor is as follows:

\[
THD = \sqrt{\frac{h_3^2 + h_4^2 + \cdots + h_n^2}{h_0 + h_1^2 + h_2^2 + \cdots + h_n^2}}
\]  

(2)

where:

- \( h_0 \) – average harmonic
- \( h_1, h_2, \ldots, h_n \) – other subsequent harmonics

As mentioned earlier, in this research, the course of solar radiation is derived from measurements taken every minute, which finally amounts to 1440 daily entries. Therefore, although many harmonics are produced, half of them are mirrored (as the main FFT property); hence, the final result is 720 harmonics used in the above formula. Due to the design of the formula, the factor takes values from 0 to 1.

The last of the quantities in the network input is the dependence of daily sums of solar irradiation on the season factor (SF). To determine this dependence, it is necessary to look at the sums of solar irradiation for individual months. The diagram below (Figure 3) shows these sums and divides them into 6 ranges (1 – from 0 to 100, 2 – from 100 to 200, 3 – from 200 to 300, 4 – from 300 to 400, 5 – from 400 to 500, 6 – from 500 to 600 MJ/m²). For example, when examining the January days, a factor of 1, or more precisely 1/6, will be taken for optimisation of the factors. Accordingly, the factor value for June days will be 6/6.

The task of the neural network is to classify the day based on 3 input data into 5 cloud classes (as a model output called day type). Cloudiness classification is difficult and complex since there is no automatic cloudiness assessment system that analyses the daily course of solar radiation. The most commonly used method of cloudiness classification are observations and manual assessment of the degree of cloud cover based on the daily course of solar radiation. For this reason, the authors have
chosen a manual method for classifying individual days into cloudiness types (graph overview, Figures 4-8). The evaluation of days was based on the analysis of the shape of the daily distribution of solar irradiation. Cloud classes were defined as follows:

1 – a fully cloudy day
2 – a mostly cloudy day
3 – a partially cloudy day
4 – a mostly sunny day
5 – a sunny day.

Figure 3. Total solar irradiation range according to month and season.

On a fully cloudy day (1) cloudiness is completely dominant and the solar radiation intensity values do not exceed 300 W/m², as shown in Figure 4. A mostly cloudy day (2) is a day with large fluctuations in radiation intensity, which is clearly visible in Figure 5. A partially cloudy day (3) is a day when the cloud cover is temporary (Figure 6). Cloudiness is most often caused by precipitation. On a mostly sunny day (4) the values of the solar radiation intensity fluctuate slightly, as seen in Figure 7. A sunny day (5) is a day when the sun is visible all day and no cloud cover is present (Figure 8).

The determination of the degree of cloudiness is quite fluent. While specifying days type 1 and 5 is rather straightforward, day type 2, which is also quite characteristic, could be misclassified as type 3. The classification difficulty increases with respect to distinguishing day types 3 and 4. In the article, measurements and rates from 1510 days are used. An example of the conversion of input data into neural network input factors is presented in Appendix A.
Figure 4. Cloud level classification example on a scale of 1.

Figure 5. Cloud level classification example on a scale of 2.
Figure 6. Cloud level classification example on a scale of 3.

Figure 7. Cloud level classification example on a scale of 4.
3. Results and discussion

Figure 9 shows how the effectiveness of the models changes depending on the number of neurons in the hidden layer. The graph shows the percentage of misqualification (confusion) and cross-entropy for models with the number of neurons in the hidden layer from 1 to 20.

Although the figure shows a decreasing trend, it was decided that the analysis of models with a larger number of hidden neurons will not comply with the simplicity principle that every mathematical model should be as simple as possible in construction. The analysis showed that the model with 20 hidden neurons provides optimum qualification skills to recognise types of a day according to cloud coverage.

Figure 10 presents an analysis of the effectiveness of the model. The diagonal cells show the number of cases that were correctly classified and the off-diagonal cells show the misclassified cases. The dark grey cell in the bottom right shows the total percentage of correctly classified cases (in green) and the total percentage of misclassified cases (in red).
Figure 10. Analysis of the effectiveness of the chosen model.

A neural network model with 20 hidden neurons is by far the best to recognise type 1 days. This good result (87.7% of correct qualifications) could be attributed to the relative ease of observation, given the most characteristic course of radiation intensity. Type 2 days are slightly more difficult to recognise, but the number of learner examples positively affects recognition. The neural network displays the greatest problem with distinguishing days type 3 and 4. Days with such cloudiness characteristics are largely similar, which results in qualification errors. Type 5 days are highly characteristic, their radiation intensity course is almost ideally sinusoidal. However, they are least frequent in the climatic zone where measurements were taken. Nevertheless, the network correctly recognises 75.2% of such days. It should be mentioned that some miscalculations are caused by the human factor, namely the manual classification of the days to neural network teaching sequence. The neural network correctly recognises 79.2% of all days, which according to the authors is a satisfactory result.

4. Conclusions
The article discusses the use of an artificial neural network to recognise the day type according to cloud cover change. The use of ANN for this purpose seems promising, but requires several changes in the model structure.

Firstly, more tests should be carried out on the use of the THD factor for testing the solar radiation course. Secondly, the number of day types should be increased. Greater variation may have a positive effect on the performance of the network. For example, an octane scale of 0-8 could be used. This would eliminate the issue of human error (manual classification) in the evaluation of days, as the data would come from documented meteorological logs. It is also possible to go beyond the scale and enter more days.

The day types could also be correlated with the THD factor, e.g. the entry of the day type 'cloudy before noon, sunny afternoon'. It would be necessary to develop and implement an automatic cloudiness assessment that provides for the daily course of solar radiation. In addition, it would be necessary to supply a division into components of the total solar radiation (i.e. diffuse and direct) to the input parameters of the neural network. Diffuse solar radiation dominates in days when cloud cover is high.

The analysed neural network model classifies days with an efficiency of approx. 80%. This is a satisfactory result, especially with the use of the THD factor, which in this application has not yet been
thoroughly tested. Undoubtedly, the result can be improved, and the use of the THD factor to evaluate cloudiness is promising.

Appendices
Appendix A. Example of converting input data into neural network input factors.

| Date (yy-mm-dd) | Daily solar radiation sum [kWh/m²] | Day type* | Daily solar irradiance factor 12th | Total harmonic distortion factor THD | Season factor SF | Requested output of the neural network |
|----------------|------------------------------------|-----------|------------------------------------|-------------------------------------|-----------------|----------------------------------------|
| 14-05-25       | 7.10                               | 4         | .8356                               | .2805                               | .8333           | 0 0 0 1 0                               |
| 14-06-10       | 8.01                               | 5         | .9428                               | .1573                               | 1.0000          | 0 0 0 0 1                               |
| 14-07-08       | 7.70                               | 5         | .9060                               | .2240                               | 1.0000          | 0 0 0 0 1                               |
| 14-08-20       | 6.36                               | 4         | .7483                               | .2371                               | .8333           | 0 0 0 1 0                               |
| 14-09-08       | 4.67                               | 4         | .5493                               | .1901                               | .6667           | 0 0 0 1 0                               |
| 14-10-19       | 2.51                               | 3         | .2959                               | .4026                               | .3333           | 0 0 1 0 0                               |
| 14-11-23       | 1.04                               | 2         | .1221                               | .3707                               | .1667           | 0 1 0 0 0                               |
| 14-12-18       | 0.40                               | 1         | .0467                               | .4732                               | .1667           | 1 0 0 0 0                               |
| 15-01-08       | 0.75                               | 1         | .0886                               | .5215                               | .1667           | 1 0 0 0 0                               |
| 15-02-02       | 1.23                               | 2         | .1445                               | .3982                               | .3333           | 0 1 0 0 0                               |
| 15-03-16       | 3.59                               | 3         | .4225                               | .4463                               | .5000           | 0 0 1 0 0                               |
| 15-04-11       | 5.19                               | 4         | .6104                               | .3735                               | .6667           | 0 0 0 1 0                               |

* manual classification

References
[1] Szala M and Awtoniuk M 2019 Neural modelling of cavitation erosion process of 34CrNiMo6 steel *IOP Conference Series: Materials Science and Engineering* **710** 012016
[2] Salat R and Awtoniuk M 2015 Black box modeling of PIDs implemented in PLCs without structural information: a support vector regression approach *Neural Computing and Applications* **26** 723-734
[3] Salat R and Salat K 2013 The application of support vector regression for prediction of the antiallodynic effect of drug combinations in the mouse model of streptozocin-induced diabetic neuropathy *Computer Methods and Programs in Biomedicine* **111** 330–337
[4] Zagórski I, Kulisz M, Klonica M and Matuszak J 2019 Trochoidal Milling and Neural Networks Simulation of Magnesium Alloys *Materials* **12** 2070
[5] Wniczenko R, Górnicki K, Kaleta A and Janaszek-Mańkowska M 2018 Optimisation of ANN topology for predicting the rehydrated apple cubes colour change using RSM and GA *Neural Computing and Application* **30** 1795–809
[6] Kalogirou S 2001 Artificial neural networks in renewable energy systems applications: a review *Renewable and Sustainable Energy Review* **5**(4) 373-401
[7] Kalogirou S 2006 Prediction of flat-plate collector performance parameters using artificial neural networks *Solar Energy* **80**(3) 248-259
[8] Farkas I and Geczy-Vig P 2003 Neural network modelling of flat-plate solar collectors *Computers and electronics in agriculture* **40** 87-102
[9] Cortés B, Tapià Sánchez R and Flores J J 2020 Characterization of a polycrystalline photovoltaic cell using artificial neural networks *Solar Energy* **196** 157-167
[10] Huang Q and Wei S 2020 Improved quantile convolutional neural network with two-stage training for daily-ahead probabilistic forecasting of photovoltaic power *Energy Conversion and Management* **220** 113085
[11] Qazi A, Fayaz H, Wadi A, Gopal Raj R, Rahim N and Khan W A 2015 The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review Journal of Cleaner Production 104 1-12

[12] Aleksiejuk J, Chochowski A and Reshetiuk V 2018 Analog model of dynamics of a flat-plate solar collector Solar Energy 160 103-116

[13] Chwieduk B 2017 Comparison of the solar radiation models Journal Of Civil Engineering, Environment And Architecture XXXIV 223-242

[14] Pang Z, Niu F and O’Neill Z 2020 Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons Renewable Energy 156 279-289

[15] Xue X 2017 Prediction of daily diffuse solar radiation using artificial neural networks International Journal of Hydrogen Energy 42 28214-28221

[16] Božnar M Z, Grašič B, Pereira deOliveira A, Soares J and Mlakar P 2017 Spatially transferable regional model for half-hourly values of diffuse solar radiation for general sky conditions based on perceptron artificial neural networks Renewable Energy 103 794-810

[17] Zou L, Wang L, Lin A, Zhu H, Peng Y and Zhao Z 2016 Estimation of global solar radiation using an artificial neural network based on an interpolation technique in southeast China Journal of Atmospheric and Solar-Terrestrial Physics 146 110-122

[18] Çelik Ö, Teke A and Yıldırım H B 2016 The optimized artificial neural network model with Levenberg–Marquardt algorithm for global solar radiation estimation in Eastern Mediterranean Region of Turkey Journal of Cleaner Production 116 1-12

[19] Ozoegwu C G 2019 Artificial neural network forecast of monthly mean daily global solar radiation of selected locations based on time series and month number Journal of Cleaner Production 216 1-13

[20] Hussain S and AlAlili A 2017 A hybrid solar radiation modeling approach using wavelet multiresolution analysis and artificial neural networks Applied Energy 208 540-550

[21] Mghouchi Y, Chham E, Zemmouri E and El Bouardi A 2019 Assessment of different combinations of meteorological parameters for predicting daily global solar radiation using artificial neural networks Building and Environment 149 607-622

[22] Vakili M, Sabbagh-Yazdi S R, Khosrojerdi S and Kalhor K 2017 Evaluating the effect of particulate matter pollution on estimation of daily global solar radiation using artificial neural network modeling based on meteorological data Journal of Cleaner Production 141 1275-1285

[23] Mousavi S M, Mostafavi E S and Jiao P 2017 Next generation prediction model for daily solar radiation on horizontal surface using a hybrid neural network and simulated annealing method Energy Conversion and Management 153 671-682

[24] Ghimire S, Deo R C, Raj N and Mi J 2019 Deep solar radiation forecasting with convolutional neural network and long T short-term memory network algorithms Applied Energy 253 113541

[25] Fodslette Møller M 1993 A scaled conjugate gradient algorithm for fast supervised learning Neural Networks 6(4) 525-533

[26] Czekalski D, Chochowski A and Obstawski P 2012 Parameterization of daily solar irradiance variability Renewable and Sustainable Energy Reviews 16(5) 2461-2467