*Sigma J Eng Nat Sci, Vol. 39, No. 2, pp. 159-169, June, 2021*

**Sigma Journal of Engineering and Natural Sciences – Sigma Mühendislik ve Fen Bilimleri Dergisi**
Web page info: https://sigma.yildiz.edu.tr DOI: 10.14744/sigma.2021.00005

**Research Article**

# Solar radiation forecasting by using deep neural networks in Eskişehir

**Mohammed QASEM[1], Ümmühan BAŞARAN FİLİK[2,*]**

[1]*Department of Statistics, Eskişehir Technical University, Eskişehir, Turkey*
[2]*Department of Electrical and Electronics Engineering, Eskişehir Technical University, Eskişehir, Turkey*

## ARTICLE INFO

*Article history*
Received: 09 April 2020
Accepted: 21 February 2021

**Key words:**
Daily global solar radiation forecasting, artificial neural network, deep neural network, renewable energy.

## ABSTRACT

According to the World Economic Outlook (WEO), the global demand for energy is presumably going to be increased due to growing the world's population up during the upcoming two decades. As a result of that, apprehensions about environmental effects, which appear as a result of greenhouse gases are grown and cleaner energy technologies are developed. This clearly shows that extended growth of the worldwide market share of clean energy. Solar energy is considered as one of the fundamental types of renewable energy. For this reason, the need for a predictive model that effectively observes solar energy conversion with high performance becomes urgent. In this paper, classic empirical, artificial neural network (ANN), deep neural network (DNN), and time series models are applied, and their results are compared to each other to find the most accurate model for daily global solar radiation (DGSR) estimation. In addition, four regression models have been developed and applied for DGSR estimation. The obtained results are evaluated and compared by the root mean square error (RMSE), relative root mean square error (rRMSE), mean absolute error (MAE), mean bias error (MBE), t-statistic, and coefficient of determination ($R^2$). Finally, simulation results provided that the best result is found by the DNN model.

**Cite this article as:** Qasem M, Başaran Filik Ü. Solar radiation forecasting by using deep neural networks in Eskişehir. Sigma J Eng Nat Sci 2021;39(2):159–169.

## INTRODUCTION

The increase in energy production from alternative resources as a result of the arising in traditional fuel prices makes the integration of renewable energy plants into the electric grid urgent and must be encouraged [1]. In addition, alternative energy resources have many advantages such as having the lowest effects on the environment 'environment-friendly' and sustainability [2]. Although these resources have many advantages, the output power of them is variable and changes in non-stationary time series because of these resources are interrupted. Keeping the global energy supply safe with the high integration of renewable energy sources is considered one of the serious challenges that can be faced soon [3].

*Corresponding author.
*E-mail address: ubasaran@eskisehir.edu.tr

*This paper was recommended for publication in revised form by Section Editor Necmiye Özlem Ünverdi*

Published by Yıldız Technical University Press, İstanbul, Turkey

The supply and demand of energy play an important role in the global development of human activity. Adequate supplies of clean energy are strongly connected with global stability, economic prosperity, and quality of life.

As it is known, solar energy is considered as one of the most important alternative energy resources which converts sunlight to electricity. The energy that can be generated by the solar panel is strongly dependent on many factors like solar radiation, weather temperature, and other weather variables.

The integration of solar energy into electric grids has come because of the increase in demand for energy. To make utilizing solar energy more efficient, forecasting information must be trusted. The exact prediction of DGSR variety can increase the quality of solar energy using and management [4].

The effective forecasting is considered as one of the most important factors that must be taken into consideration in solar power plant design [5]. Forecasting can be helpful for the development of different applications for power systems [6, 7]. To make DGSR forecasting more accurate, there are several methods proposed in the last decades. The results of most methods may not be acceptable because of high rates of error. The accuracy of solar radiation forecasting does not exceed 88% by using classic empirical methods [8, 10]. Therefore, we need to create new models by using novel methods in order to get results with high accuracy.

DGSR value can be affected by many weather factors, such as cloud cover, precipitation, temperature, wind speed, pressure, and humidity. Nevertheless, the precise correlation between each weather feature and the received DGSR can be hard to be find. Therefore, the originality and the main aim of this work is to create a direct relationship, by using a fully connected DNN, between various weather parameters and DGSR values, to make DGSR prediction more accurate. DNN can be used to find linear as well as non-linear relationship between features and targets without much of feature engineering or domain knowledge [11, 13].

In addition, the study [14] realized in Eskişehir proved that CPRG, CPR [6], Liu and Jordan [11], Jain model 1 [8] & Jain model 2 [9] models obtain the best results among the discussed classic empirical models. For this reason, they are chosen as the best classic empirical models to be compared with AI methods. Furthermore, two of the well-known time series models (autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average with Exogenous Regressors (SARIMAX) are implemented in this study.

The rest of the paper is structured as follows. In section 2, different DGSR methods are described. In section 3, simulations and results of each method are discussed in detail. In the last section, the conclusion and the key contributions of the paper are highlighted.

## METHODS OF SOLAR RADIATION ESTIMATION

Solar radiation estimation with a higher percentage of accuracy has the biggest role in modeling solar systems. As can be seen from Figure 1, there are various applications that use prediction results to upgrade works and future plans of the electric grid with the conforming required time-resolution of the forecasting [11]. For instance, solar radiation prediction about the upcoming second is important to operate network stability and grid voltage regulation optimally. In addition, power reserve management, dispatching, and load flowing need information about the next hours of energy forecasting. Similarly, transmission scheduling and unit commitment require data about the upcoming days of energy forecasting. And also, optimization, production, and consumption of electrical energy need predicted data about the upcoming days, weeks, or months of solar radiation. And finally, capacity and global management need energy forecasting predicted data for the upcoming years. Energy forecasting is necessary for many applications.

Because of the discontinuous and unpredictable character of renewable energy, the integration of it into the electrical grid increases the network management complexity and the balance continuity between production and consumption.

To avoid damage in the electrical loads they must be powered by constant balanced electricity all the time. However, the electrical generator and operator has some difficulties to keep the balance constant with traditional and energy production systems that can be controlled.

The reliability of the system depends on how the system can adapt to expected and unforeseen changes in production, disturbances, and consumption of electrical energy. In addition, to keep the quality of service at a good level, the energy provider should manage the system with different temporary horizons.

Grid voltage, power quality, and grid stability problems can be related to uncontrollable solar energy production [7, 8]. For these reasons, the need for forecasting models with a high accuracy becomes urgent and challenging. In view of the great importance of solar energy in energy production, the most common techniques used in the solar radiation prediction process will be addressed.

Solar radiation forecasting techniques can be classified into three main, different classes. These classes given in Figure 2 are:

- Physical structural forecasting models that depend on meteorological and geographical parameters.
- Time-series models that are based on the historical observed data of solar irradiance as input values.
- Artificial intelligence and hybrid models that are based on the historical observed data of solar radiation and considered as a combination of two or

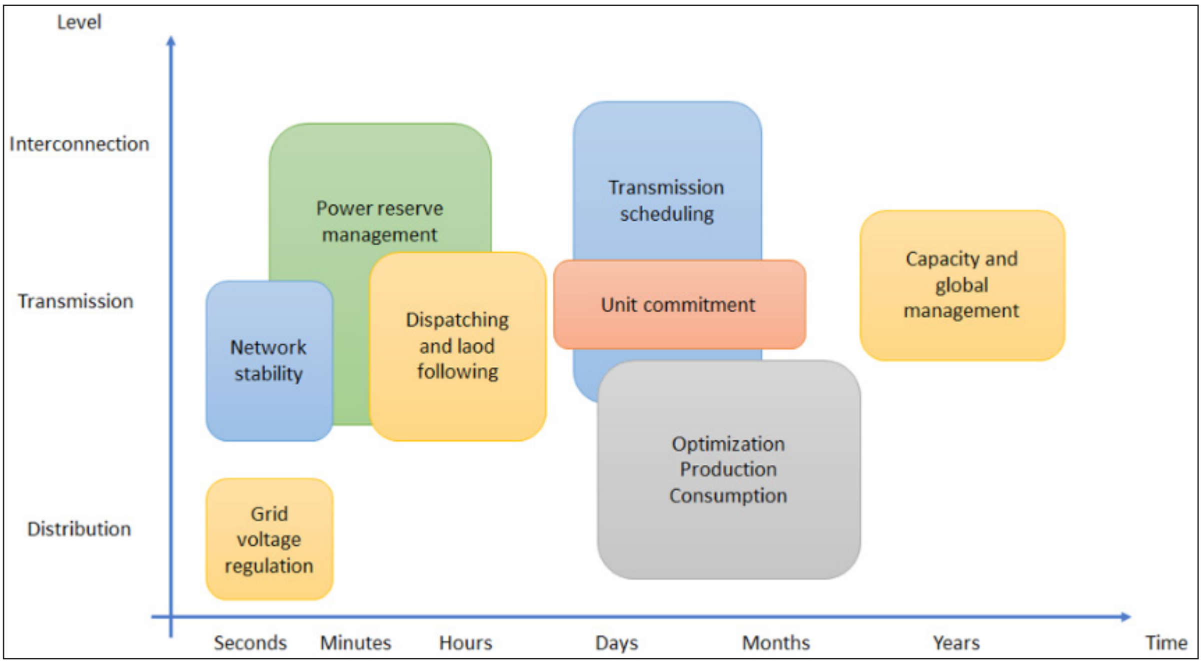


**Figure 1.** Different applications use forecasting data with different time intervals [8].

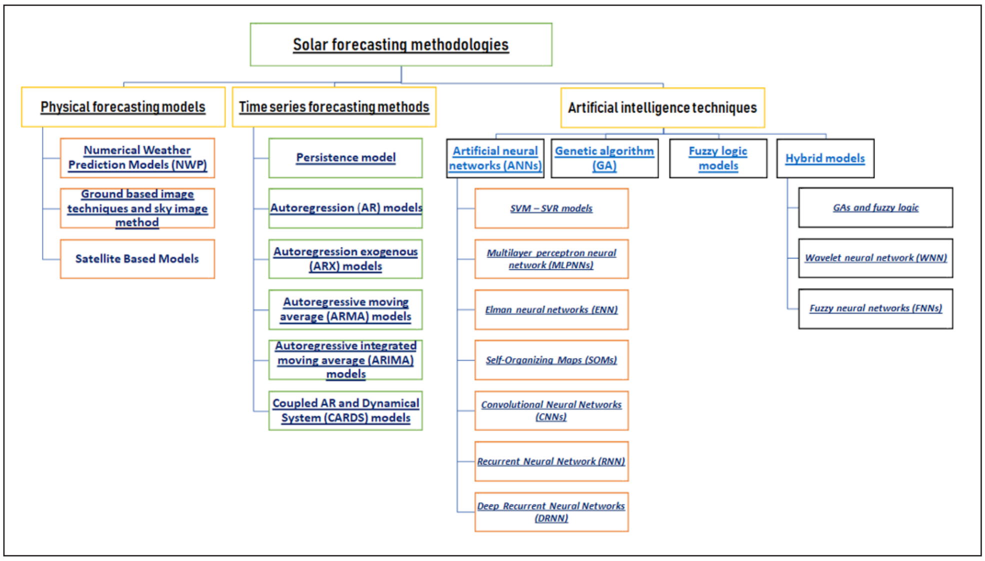


**Figure 2.** Solar irradiance prediction methods [12].

more different models of solar radiation forecasting, respectively.

### Emprical Mathematical Models

According to the study [14], which was implemented in Eskişehir, the best empirical models in that study are CPR, CPRG, Liu and Jordan, Jain model 1, and Jain model 2. For this reason, they are chosen to compare their results with DNN model results. In Table 1, the chosen empirical models which are used in this study are shown.

### ARIMA and SARIMAX Models

ARIMA model is used more than other classic empirical models because the inputs to the models are not stationary yet. The mathematical structure of ARIMA models is described as:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots \beta_1 Y_{t-p} \in_t + \varphi_1 \in_{t-1} + \varphi_1 \in_{t-2} + \dots \varphi_1 \in_{t-q} \quad (1)$$

where $Y_t$ is the predicted value, $\alpha$ is a constant, $\beta_1 Y_{t-p}$ are linear combination lags of $Y$, $\phi_1 \epsilon_{t-2}$

**Table 1.** The chosen empirical mathematical models

| Model | Mathematical structure |
|---|---|
| CPR [5] | $\frac{I}{H}=\frac{\pi\left(a+b\cos W\right)}{24}\cdot\frac{\cos W-\cos W_s}{\sin W_s-\frac{\pi W_s}{180}\cos W_s}$ |
| CPRG [6] | $\frac{I}{H}=\frac{360\left(a+b\cos W\right)r_0}{f_c}$<br>$f_c=a+0.5b\frac{\frac{\pi W_s}{180}\left(\sin W_s\cos W_s\right)}{\sin W_s-\frac{\pi W_s}{180}\cos W_s}$<br>$r_0=\frac{\pi}{24}\frac{\cos W_s\cos W_s}{\sin W_s-\frac{\pi W_s}{180}\cos W_s}$ |
| Liu and Jordan model [16] | $r_0=\frac{\pi}{24}\frac{\cos W-\cos W_s}{\sin W_s-\frac{\pi W_s}{180}\cos W_s}$ |
| Jain model 1 [8] | $\frac{I}{H}=\frac{1}{\sigma\sqrt{2\pi}}e^{-\left(\frac{(t_s-12)^2}{2\sigma^2}\right)}$<br>$\sigma = 0.192S_0 + 0.461$<br>$S_0=\frac{2}{15}W_s$ |
| Jain model 2 [9] | $\frac{I}{H}=\frac{1}{\sigma\sqrt{2\pi}}e^{-\left(\frac{(t_s-12)^2}{2\sigma^2}\right)}$<br>$\sigma = 0.2S_0 + 0.378$ |

are linear combination of lagged forecast errors [15].

The ARIMA model is considered as the reference estimators in the DGSR forecasting. The sharp transitions in radiation is captured by this model more accurately than other models [16]. SARIMA model is developed by Craggs et al. [17], accounting for about 82 and 85% of the total variation in the 10 min averaged horizontal and vertical irradiances, respectively. SARIMAX model takes exogenous variables which influences the value of time series model at time t, but it is not auto-regressed on. SARIMAX model is described as:

$$\Theta\left(L\right)^p\theta\left(L^s\right)^p\Delta^d\Delta_S^D y_t=\Phi\left(L\right)^q\varphi\left(L^s\right)^Q\Delta^d\Delta_S^D\in_t+\sum_{i=1}^{n}\beta_i x_t^i \quad (2)$$

where $\Theta(L)^p$,$\Phi(L)^p$ are non-seasonal lag polynomials, $\Delta^d\Delta_S^D y_t$ are the time series, differenced $d$ times, and seasonally differenced $D$ times, $\theta(L^s)^p$ is the non-seasonal autoregressive lag polynomial, $\varphi(L^s)^p$ is the seasonal moving average lag polynomial, $\beta_i x_t^i$ are the linear combination lags of $x$, $\epsilon_t$ is the linear combination of lagged forecast error.

**ANN**

It is considered as a processing system that connects small processing units. ANN simulates the shape and the work of the biological neural network in the human body. ANN can automatically learn to recognize patterns in data from the previous data that can be added to the network [18]. The ability to model complicated and non-linear processes between the input and output variables is considered as the best advantage of ANN that cannot be found in the other forecasting techniques.

To perform ANN for DGSR prediction, the Visual Gene Developer software is used in this study [19]. Prediction with using this software is achieved by using different ANN structures with different parameters. The model that is used for DGSR prediction consists of ANN with 3 hidden layers beside input and output layers. The shape of the ANN model created in Visual Gene Developer software is clearly shown in Figure 3.

**DNN**

it is considered as a modified ANN with many hidden layers, with more than three hidden layers. This kind of neural networks is widely used in image and speech recognition competitions, translation, and e-Marketing [18]. DNN has a fundamental one- or two-dimensional structure that can be captured and exploited by conventional layers. As a result, the forward development of IoT and increasing big data capacity, DNN models are recently taking large attention to be used in diverse research fields. DNN has the potential to give the nonlinear relationship between input features and the output targets. It takes on learning representations from data that puts an emphasis on learning successive layers of increasingly meaningful representations. The deeper it goes, the higher-level representation it is able to identify, so that it establishes a proper mapping from input features to their target [20]. The proposed network is a DNN consisting of 1 input layer with 5 nodes, 1 output layer with 1 node, and 5 hidden layers. The general structure of the proposed DNN used for forecasting is shown in Figure 4.

**Developing Regression Models for DGSR Prediction Using SPSS**

By using the regression method in SPSS [19], four different models are created. These models are developed for DGSR estimation for Eskisehir city, by using the collected data in the selected period between 01.01.2014 until 31.12.2014. The mathematical structure of the created models is expressed as shown in Eqn. (3–6):

$$SPSS_{Model_{1(SR)}}=0.191+0.233\cdot T_{max} \quad (3)$$

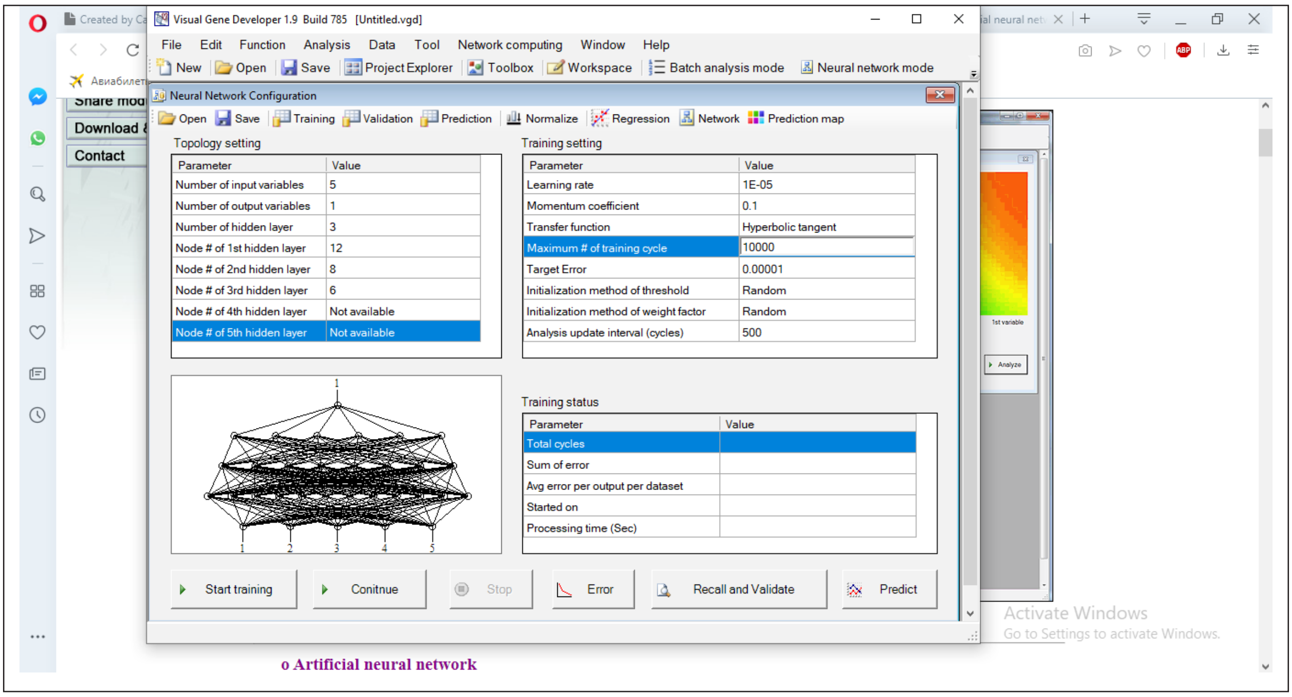


**Figure 3.** General structure of the ANN prediction model in Visual Gene Developer software.

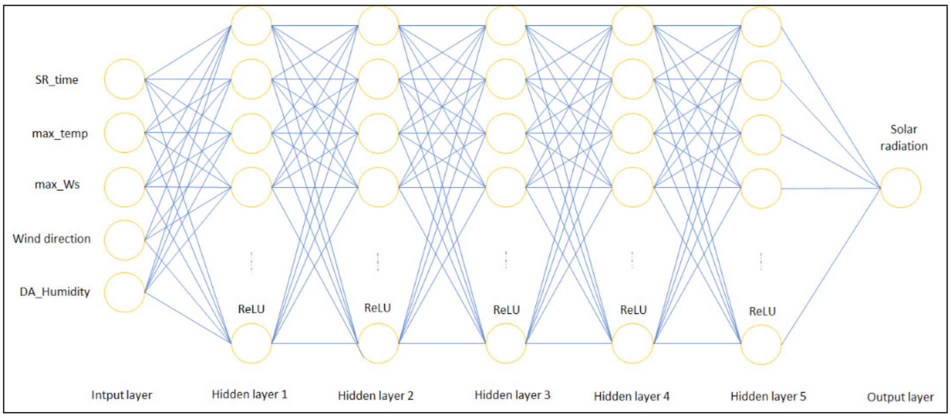


**Figure 4.** The general structure of the proposed DNN.

$$SPSS_{Model_{2(SR)}} = 0.069 + 0.209T_{max} + 0.675W_{speed} \quad (4)$$

$$SPSS_{Model\_3(SR)} = 0.343 + 0.168T_{max} + 0.604W_{speed} - 0.281.\ Humidity \quad (5)$$

$$SPSS_{Model4\_(SR)} = 0.338 + 0.089SR_t + 0.128T_{max} + 0.608W_{speed} - 0.013W_{direction} - 0.261.\ Humidty \quad (6)$$

where $SPSS_{Model4_{(SR)}}$ is the solar radiation function, $T_{max}$ is the daily maximum temperature, $W_{speed}$ is the daily wind speed, $SR_t$ is the solar radiation time. $SR_t$ it is the time that DGSR can be obtained from the sun.

## PERFORMANCE EVALUATION

The predicted DGSR values, obtained by DNN, are compared with those which are predicted by using statistical analysis methods, consequently, the work of the models used for the DGSR prediction is evaluated to obtain the best model. The most commonly used methods are as following: root mean square error (RMSE), mean absolute error (MAE), mean bias error (MBE), t-statistic, Coefficient of determination ($R^2$), and relative root mean squared error (rRMSE), which are described in detailed Eqn. (7-12) as following:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2}{n}} \quad (7)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \overline{y}_i\right| \quad (8)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right) \quad (9)$$

$$t - statistic = \sqrt{\frac{(n-1)MBE}{RMSE^2 + MBE^2}} \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \quad (11)$$

$$rRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2}{n}}}{\hat{y}_i} \quad (12)$$

where $n$ is the number of samples, $y_i$ is the true target value of sample $i$, $\overline{y}_i$ is the predicted target value of sample $i$, $\hat{y}_i$ is the mean of the selected data.

## SIMULATION AND RESULTS

In this work, DGSR data are predicted by using classic empirical models (CPR, CPRG, Liu and Jorden Lian Model 1 and Lain Model 2); time series models (ARIMA, SARIMAX), ANN model, DNN model, and 4 developed regression models for Eskisehir city. According to [14] study, the five empirical models give the best results in Eskişehir region. For this reason, they are chosen to compare their results with ANN, DNN, and developed regression models results. The classic empirical models are simulated in Microsoft Excel; however, time series models are simulated in Python. ANN model is trained by using Visual Gene Developer software. Otherwise, DNN models are trained by using the TensorFlow Keras Machine Learning Python library in the Google Colab environment. DNN models consist of one input layer contains 5 parameters, hidden layers between three and ten, and one output layer. The data are collected by the meteorology station in Eskişehir. Collected DGSR and other weather parameters (maximum temperature, maximum wind speed, wind direction, solar time, and humidity) data are recorded for four years, from 2011 until 2014. After filling the missed data feature, the vector of each sample is reshaped into a 5-dimensional vector and the target vector in a one-dimensional vector. Thus, the input layer containing 5 nodes, each node represents one feature of the weather parameters and the output layer contains only 1 node that represents DGSR. The data are divided into three sub-datasets: training, validation, and testing dataset. The percentages of the training, validation, and testing data are 60%, 20%, and 20% respectively.

DNN models are trained and tested using the collected data. The first simulations are performed with DNN. The results of 12 different DNN models are given in Table 2. The Rectified Linear Unit (ReLU) is applied to each node in the hidden layers and the number of epochs is 1000 for all of the 12 models. It can be seen that the best model with the best performance is the DNN model 4 which consists of 1 input layer, 10 hidden layers, and 1 output layer. DNN model 4 is trained with batch size equal to 32 and RMSprop equal to

**Table 2.** The results of different DNN models

| Model | Batch size | Number of hidden layers | Optimizers RMSprop | RMSE | MAE | MBE | t-statistic | $R^2$ | rRMSE |
|---|---|---|---|---|---|---|---|---|---|
| DNN Model 1 | 32 | 3 | 0.0001 | 0.104 | 0.083 | −0.0048 | 0.1532 | 0.803 | 0.2029 |
| DNN Model 2 | 32 | 6 | 0.001 | 0.106 | 0.090 | −0.0228 | 0.7341 | 0.798 | 0.2058 |
| DNN Model 3 | 64 | 10 | 0.0001 | 0.101 | 0.086 | −0.0059 | 0.1945 | 0.815 | 0.1967 |
| DNN Model 4 | 32 | 10 | 0.001 | 0.089 | 0.071 | −0.0047 | 0.1758 | 0.857 | 0.1732 |
| DNN Model 5 | 64 | 5 | 0.001 | 0.098 | 0.084 | −0.0200 | 0.6899 | 0.825 | 0.1915 |
| DNN Model 6 | 32 | 5 | 0.001 | 0.098 | 0.080 | −0.0225 | 0.7858 | 0.827 | 0.1904 |
| DNN Model 7 | 64 | 7 | 0.001 | 0.127 | 0.116 | −0.0211 | 0.5588 | 0.706 | 0.2478 |
| DNN Model 8 | 32 | 7 | 0.001 | 0.113 | 0.099 | −0.0144 | 0.4261 | 0.768 | 0.2204 |
| DNN Model 9 | 64 | 10 | 0.001 | 0.108 | 0.092 | 0.0204 | 0.6404 | 0.790 | 0.2098 |
| DNN Model 10 | 32 | 10 | 0.0001 | 0.099 | 0.080 | −0.0181 | 0.6200 | 0.823 | 0.1922 |
| DNN Model 11 | 64 | 3 | 0.0001 | 0.102 | 0.086 | −0.0106 | 0.3452 | 0.810 | 0.1997 |
| DNN Model 12 | 64 | 3 | 0.001 | 0.102 | 0.084 | −0.0036 | 0.1171 | 0.811 | 0.1990 |

0.001. The number of processed samples before updating the model is called batch data. Batch data is also known as the number of the finished processes during the training dataset. The batch size value should be an integer between one and the number of the samples in the training dataset. However, the number of epochs can be an unlimited value between one and infinity. According to the results, DNN model 4 is considered as the best choice for DGSR forecasting in order to contribute to the management of the electric power grid.

In Figure 5, it is clearly shown that DNN model 4 has the best performance among different DNN models. Its predicted data almost fits the test data.

Then, DGSR is predicted by using 5 different classic empirical models, ANN model, four new developed regression models, and two time series models. The performance of each model is shown in Table 3.

As can clearly be seen in Table 3, the DNN model 4 and SARIMAX model have the closest results to the real values. The statistical performance of the models are shown in the Table 3, DNN model 4 has the lowest value of the RMSE, when we compared it with the other models performance.

In the following graphs it is clearly shown that the DNN model 4 results have the minimum error value. DNN model 4 modeling results of the testing and predicted results are shown in Figure 6. The predicted values basically capture

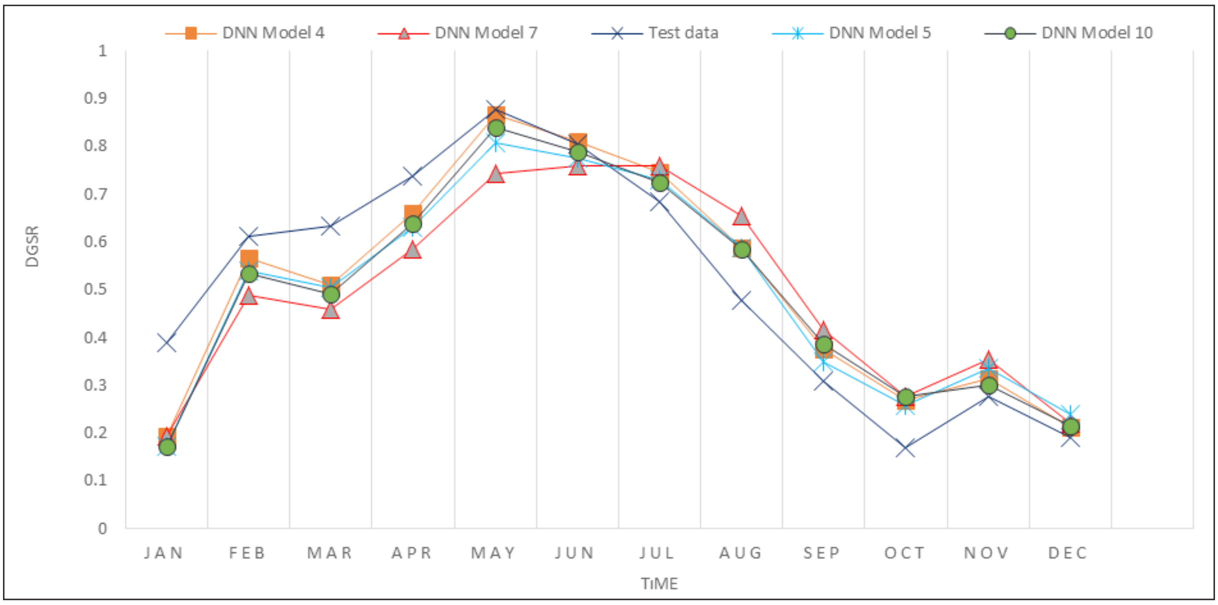


**Figure 5.** The performance of different DNN models.

**Table 3.** Statistical parameters of the daily global solar radiation of different models

| Model | RMSE | MAE | MBE | t-statistic | $R^2$ | rRMSE |
|---|---|---|---|---|---|---|
| DNN_Model 4 | 0.0888 | 0.0714 | −0.0047 | 0.1758 | 0.8566 | 0.1731 |
| ANN | 0.2789 | 0.2414 | −0.0288 | 0.3443 | -0.4153 | 0.5441 |
| $SPSS_{Model\ 1}$ | 0.2431 | 0.2219 | −0.0118 | 0.1612 | -0.0749 | 0.4742 |
| $SPSS_{Model\ 2}$ | 0.2330 | 0.2112 | −0.0103 | 0.1468 | 0.0128 | 0.4544 |
| $SPSS_{Model\ 3}$ | 0.2179 | 0.1957 | −0.0091 | 0.1386 | 0.1363 | 0.4251 |
| $SPSS_{Model\ 4}$ | 0.2185 | 0.1965 | −0.0091 | 0.1382 | 0.1312 | 0.4263 |
| CPR | 0.1591 | 0.1414 | −0.0239 | 0.5039 | 0.5398 | 0.3103 |
| CPRG | 0.1575 | 0.1387 | 0.0008 | 0.0168 | 0.5490 | 0.3072 |
| Liu | 0.1554 | 0.1377 | −0.0011 | 0.0235 | 0.5608 | 0.3031 |
| Jain model1 | 0.1592 | 0.1411 | −0.0111 | 0.2318 | 0.5392 | 0.3105 |
| Jain model2 | 0.1592 | 0.1412 | −0.0116 | 0.2423 | 0.5391 | 0.3105 |
| ARIMA | 0.1113 | 0.0866 | 0.0164 | 0.4941 | 0.7621 | 0.2353 |
| SARIMAX | 0.0940 | 0.0705 | 0.0130 | 0.4631 | 0.8327 | 0.1832 |

the true target value; however, there are some several differences around turning points, and the trends of some middle points are completely toward different directions. From Figure 6, it can be seen how DNN model 4 gives accurate results.

As it is shown in Figure 6, the predicted values can capture most of the true values. In Figure 7, it is shown The MAE for training and validation processes during training the DNN model 4, and it shows that at 844th epoch the model gives the minimum value of MAE. However, in Figure 8 it is shown the monthly average results of DGSR obtained by the different models. It can be clearly seen that DNN model 4 gives the best results among them.

## CONCLUSION

DGSR estimation plays an important role in the part of electricity production to a short and long-time interval

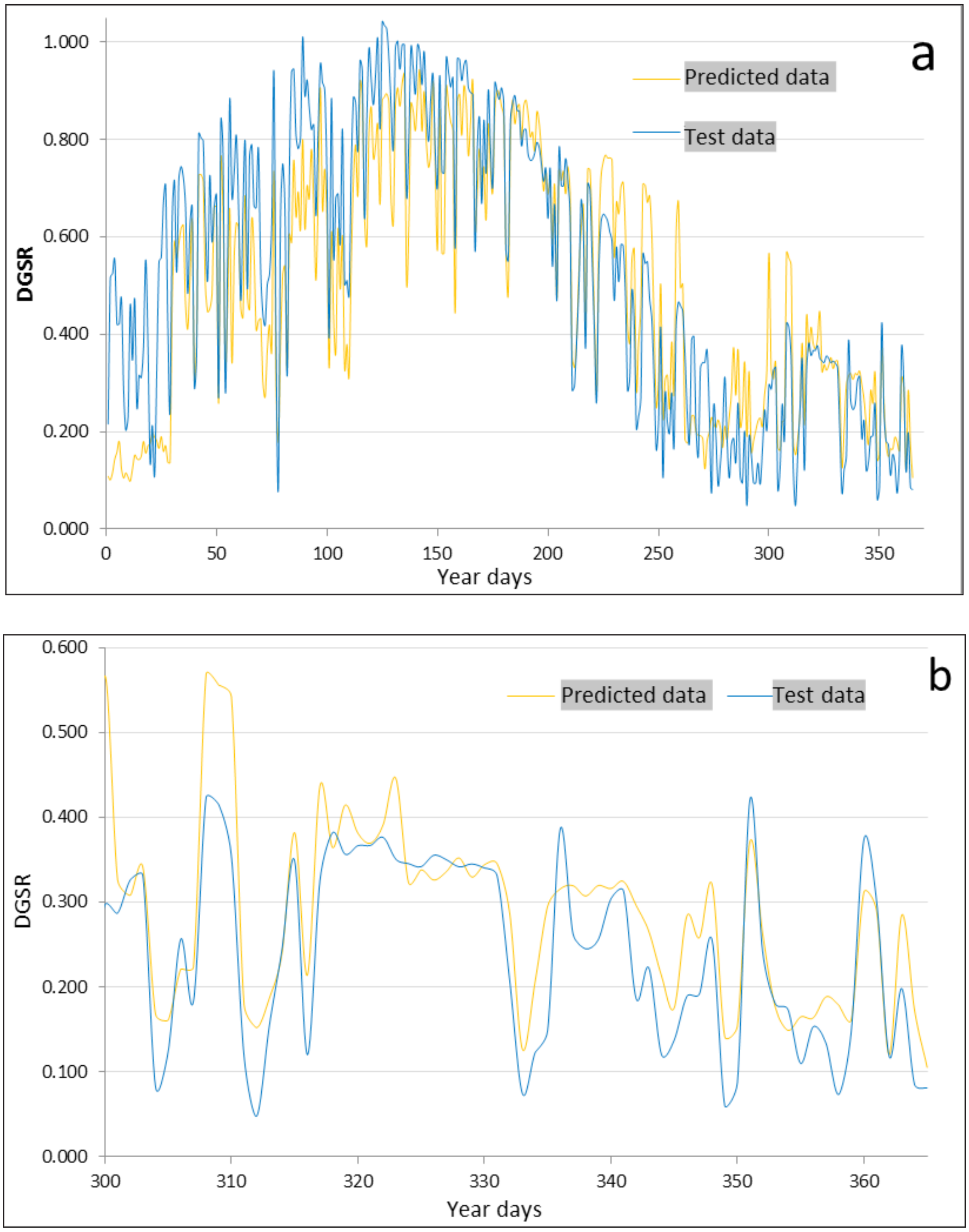


**Figure 6.** Testing result of the DNN model 4 a) results of one-year forecasting b) zoomed portion of the model results.

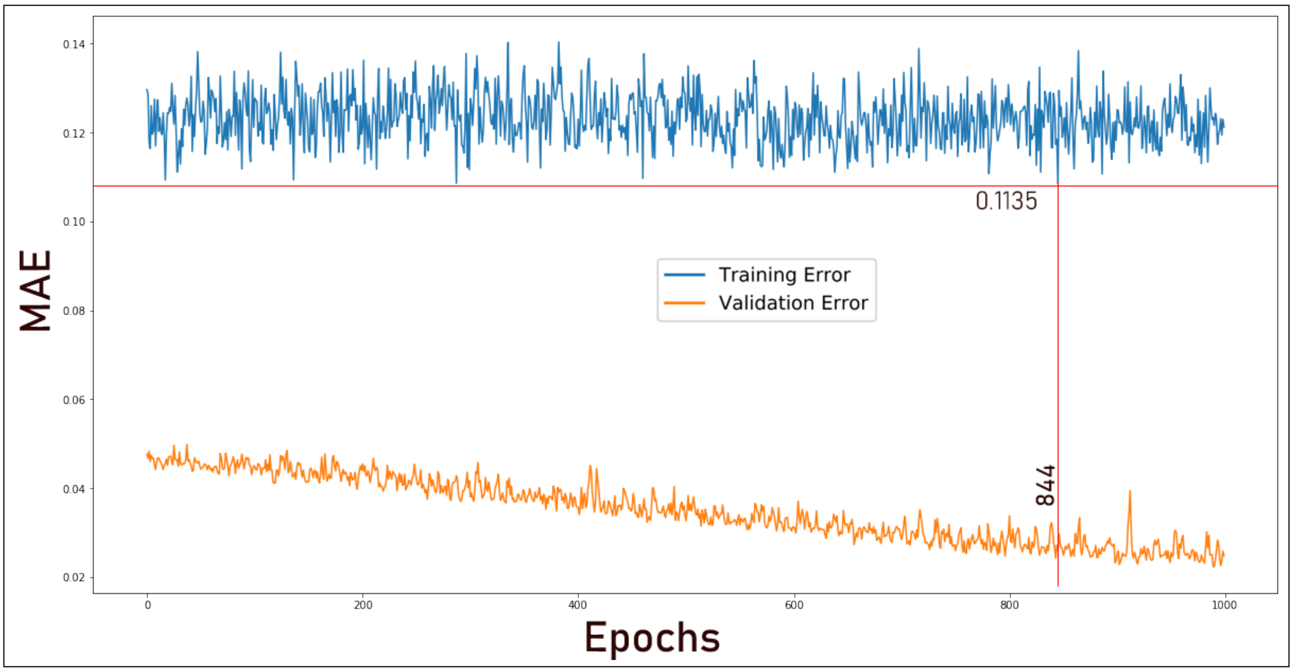


**Figure 7.** The MAE for training and validation processes of the DNN model 4.

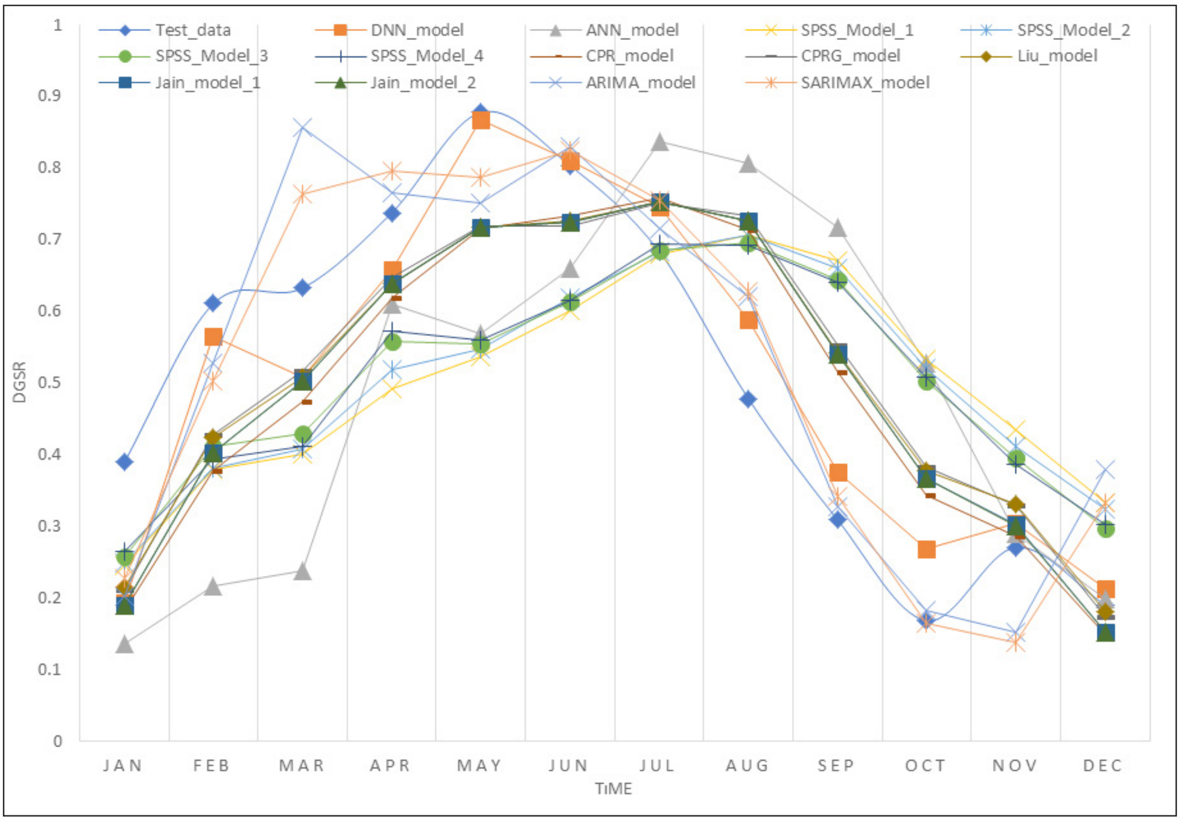


**Figure 8.** The results of monthly predicted DGSR of different used models.

in solar energy systems. Most of the published papers with many experimental results with a large number of data, support using DNN and its superiority are compared with the other methods. For this, DGSR for the given location has been predicted by using the previously proposed DNN model. The proposed DNN model contains an input layer, 10 hidden layers, and one output layer. The input layer contains 5 nodes which represent the five weather features on which the DGSR depends. The output layer has only 1 node and represents predicted DGSR values. Then, the network is trained and tested using the given data. In addition, the same data are predicted by using different models; five classic empirical models (CPR, CPRG, Lain Model 1, Lain Model 2, and Liu model), time series models (ARIMA and SARIMA), new developed regression models, and ANN model. The statistical performances of all created models are evaluated by RMSE, MABE, MBE, t-statistic, rRMSE, and $R^2$. The experiment results show that DGSR forecasting by using the DNN model gives the best result with high accuracy. That is, the results show that DNN model 4 gives the best result with a coefficient of determination equal to 0.8566; however, the second model that gives a good result is SARIMAX with $R^2$ equal 0.8327. In case that DNN model gives the best results in forecasting, they can be considered as the best alternative solution for future weather forecasting cases. Consequently, the prediction of the energy produced by renewable energy resources, including solar energy, will be more accurate. That is, an increase in prediction accuracy means a decrease in losses resulting from less accurate prediction.

Future research could focus on the improvement of the model performance and trying to implement hybrid models that combine DNN model with one or more models together. Some models that are recommended for future research are Fuzzy Logic Neural Networks (FNN) and Fuzzy Logic with Genetic Algorithms (GAs) model.

## NOMENCLATURE

**Abbreviation**

| | |
|---|---|
| AI | Artificial intelligence |
| ANN | Artificial neural network |
| ARIMA | Autoregressive integrated moving average |
| CPR | Collares-Pereira and Rabl |
| CPRG | Collares-Pereira and Rabl modified by Gueymard |
| DGSR | Daily global solar radiation |
| DNN | Deep neural network |
| FNNs | Fuzzy neural networks |
| Gas | Genetic Algorithms |
| IoT | Internet of Things |
| MAE | Mean absolute error |
| MBE | Mean bias error |
| ReLU | Rectified Linear Unit |
| RMSE | Root mean square error |
| rRMSE | Relative root mean squared error |
| SARIMAX | Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors |
| SVM | Support vector machines |
| WEO | World Economic Outlook |
| WNN | Wavelet neural network |

**Subscripts**

| | |
|---|---|
| $H$ | Daily global solar radiation on horizontal surface (W/m$^2$) |
| $I$ | Hourly global solar radiation on horizontal surface (W/m$^2$) |
| $f_s$ | Multiplying factor |
| $n$ | Number of samples |
| $r_0$ | Multiplying factor (unitless) |
| $R^2$ | Coefficient of determination |
| $SR_t$ | The solar radiation time |
| $T_{max}$ | The daily maximum temperature |
| t-statistic | t-statistic |
| $y_i$ | Measured solar radiation value |
| $\hat{y}_i$ | Mean of measured solar radiation |
| $\bar{y}_i$ | Predicted value of solar radiation |
| $W_s$ | Sunrise hour angle |
| $W_{speed}$ | The daily wind speed |
| $W$ | Solar hour angle |

**Greek symbols**

| | |
|---|---|
| $\epsilon_t$ | linear combination of lagged forecast errors |
| $\Theta(L)^p$, $\Phi(L)^p$ | non-seasonal lag polynomials |
| $\Delta^d \Delta_s^D y_t$ | the time series, differenced d times, and seasonally differenced D times |
| $\theta(L^s)^P$ | the non-seasonal autoregressive lag polynomial |
| $\varphi(L^s)^Q$ | the seasonal moving average lag polynomial |
| $\beta^i x_t^i$ | linear combination lags of $x$ |

## AUTHORSHIP CONTRIBUTIONS

Both authors contributed equally.

## DATA AVAILABILITY STATEMENT

The published publication includes all graphics and data collected or developed during the study.

## CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## ETHICS

There are no ethical issues with the publication of this manuscript.

## REFERENCES


[1] Başaran Filik, Ü., Filik, T., & Gerek, Ö. N. A hysteresis model for fixed and sun tracking solar PV power generation systems. Energies; 2018, p. 603.

[2] Alzahrani, A., Shamsi, P., Dagli, C. and Ferdowsi, M. Solar irradiance forecasting using deep neural networks. Procedia Computer Science; 2017, p. 304–313.

[3] Benali, L., Notton, G., Fouilloy, A., Voyant, C., & Dizene, R. Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components. *Renewable Energy;* 2019, p. 871–884.

[4] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. Time series analysis: forecasting and control. *John Wiley & Sons;* 2015.

[5] Chaturvedi, D. K. Solar Power Forecasting: A Review, *145*(6); 2016, p. 28–50.

[6] Collares-Pereira, M., & Rabl, A. The average distribution of solar radiationcorrelations between diffuse and hemispherical and between daily and hourly insolation values. *Solar Energy*, *22*(2), 1979, p. 155–164.

[7] Gueymard, C. Mean daily averages of beam radiation received by tilted surfaces as affected by the atmosphere. *Solar Energy*, *37*(4), p. 1986, 261–267.

[8] Http-2. date of access: July 5, 2019, from http://www.visualgenedeveloper.net/; 2019.

[9] Jain, P. C. Comparison of techniques for the estimation of daily global irradiation and a new technique for the estimation of hourly global irradiation. *Solar & Wind Technology*, *1*(2), 1984, p. 123–134.

[10] Jain, P. C. Estimation of monthly average hourly global and diffuse irradiation. *Solar & Wind Technology*, *5*(1), 1988, p. 7–14.

[11] Kalogirou, S. A. Artificial neural networks in renewable energy systems applications: a review. *Renewable and Sustainable Energy Reviews*, *5*(4), 2001, p. 373–401.

[12] Liu, B. Y., & Jordan, R. C. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy*, *4*(3), 1960, p. 1–19.

[13] Ryu, S., Noh, J., & Kim, H. Deep neural network-based demand side short term load forecasting. *Energies*, *10*(1), 3, 2017.

[14] Ayvazoğluyüksel, Ö. and Filik Başaran Ü. Estimation methods of global solar radiation, cell temperature and solar power forecasting: A review and case study in Eskişehir. Renewable and Sustainable Energy Reviews, 91, 2018, p. 639–653

[15] Atique, S., Noureen, S., Roy, V., Subburaj, V., Bayne, S., & Macfie, J. Forecasting of total daily solar energy generation using ARIMA: A case study. In 2019 IEEE 9th annual computing and communication workshop and conference (CCWC), 2019, p. 114–119.

[16] Sivhugwana, K. S., & Ranganai, E. Intelligent techniques, harmonically coupled and SARIMA models in forecasting solar radiation data: A hybridisation approach. Journal of Energy in Southern Africa, 31(3), 2020, p. 14–37.

[17] Craggs, C., Conway, E. and Pearsall N. M. Stochastic modelling of solar irradiance on horizontal and vertical planes at a northerly location. Renewable Energy 18; 1999, p. 445–463

[18] Soubdhan, T., Voyant, C., & Lauret, P. Influence of Global Solar Radiation Typical Days on Forecasting Models Error. *The Third Southern African Solar Energy Conference (SASEC2015);* 2015.

[19] Sun, Y. *Deep Neural Network Regression and Sobol Sensitivity Analysis for Daily Solar Energy Prediction Given Weather Data*. Purdue University; 2018.

[20] Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., & Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, *105*, 2017, p. 569–582.