# TIME DEPENDENT PREDICTION OF MONTHLY GLOBAL SOLAR RADIATION AND SUNSHINE DURATION USING EXPONENTIALLY WEIGHTED MOVING AVERAGE IN SOUTHEASTERN OF TURKEY

by

***Bilal GUMUS[a] and Hibetullah KILIC[b*]***

[a] Department of Electrical and Electronics Engineering, University of Dicle, Diyarbakir, Turkey
[b] Diyarbakir Technical Science Vocational School, University of Dicle, Diyarbakir, Turkey

Original scientific paper
https://doi.org/10.2298/TSCI160107228K

*This paper proposes a new approach for prediction of global solar radiation and sunshine duration based on earlier years of data for the eastern region of Turkey which has a high potential of solar energy. The proposed method predicts the basic parameters using time series and an analysis method. This method is exponentially weighted moving average. This model estimates next years global solar radiation and sunshine duration and is evaluated by statistical parameters, mean absolute percentage error (MAPE) and coefficient of determination, to examine the success of the proposed technique. In our study, the result shows that this method is effective in predicting global solar radiation and sunshine duration as regards of MAPE and coefficient of determination. The calculated MAPE which are between 0-10 kWh/m² per day were assumed excellent and coefficient of determination were found significant per every year.*

Key words: *exponentially weighted moving average, global solar radiation, sunshine duration*

## Introduction

Solar energy has some basic parameters such as, global solar radiation, (GSR), and sunshine duration (SSD), which give ideas about the generated power from photovoltaic, and solar energy conversion systems. Measurement of GSR and SSD is also helpful in some other industrial and scientific areas known as wood drying, stoves, atmospheric studies, thermal load analyses on buildings, and meteorological forecasting [1, 2]. In order to predict GSR, there are several approaches including the sunshine, temperature, and cloud-based linear and non-linear regression.

Studying of these models has started in the first quarter of the 20$^{th}$ century [3]. Currently, several various alternative methodologies have been developed to explain the dynamic behavior in an Angstrom-Prescott based model [4].

Furthermore, air temperature or air temperature-cloudiness method are used to model GSR. Due to the lack of or missing SSD in some locations, these models have shown to be effective and successful to the conventional estimation which is based on SSD. To achieve better estimation in GSR, fuzzy logic can be applied and used successfully in temperature-based models [5, 6].

* Corresponding author, e-mail: hibetullah.kilic@dicle.edu.tr

Over the past two decades, the ANN approach has been utilized in solar radiation prediction or was combined with Angström-Prescott and other meteorological data in the development of models [8, 9]. It has been proven they are high-quality tools for research as they are able to be used in non-linear system behaviors, data categorization, clustering or model ordinary systems simulation. Hence, these are beneficial tools in the prediction of systems that have complicated behavior such as GSR and meteorological events [10, 11]. In addition, ANN propose a non-linear statistical method that became very popular when trying to deal with a problem in a different way such as a problem analogous to atmospheric science [7].

Another approach in predicting GSR is the time series analysis method. In meteorological science, events such as GSR, SSD, and temperatures, their behaviors are expressed in time scales [12]. Therefore, the time series analysis method are used in GSR estimation providing most knowledge of the underlying physical nature of GSR and SSD [13, 14]. Time series analysis gives some facilities that are used with conventional methods and also combined with some modern methods that are ANN and its derivation [15].

In this paper, a different approach of time series based analysis method called, exponentially weighted moving average (EWMA) is employed. This approach is an advanced method for prediction of both GSR and SSD in some regions especially in the south-eastern Anatolia region in Turkey that consists of nine cities. (Diyarbakir, Gaziantep, Sanliurfa, Batman, Adiyaman, Siirt, Mardin, Kilis, Sirnak). Due to the fact that the measurement station is removed in this region, data of GSR and SSD has not been recorded since August of 2015. In this paper, data supported by the Turkish Meteorological State Service for the cities Batman, Diyarbakir, Gaziantep, Sanliurfa, and Mardim are between 1998 and 2015. It is known that there will be no further measurement data of GSR and SSD provided for the next few years by the meteorological service. Thus, the sunshine based model will not be applicable in the prediction of GSR [17]. The aim of the EWMA time series is to achieve a new approach for prediction of GSR and SSD and also achieve acceptable prediction of GSR and SSD by using the previous measured data and eliminate the lack of some datasets or unavailable situation for the measurement station [16].

## Materials and methods

In Turkey, the south-eastern Anatolia region receives a significant amount of solar energy. The monthly average of daily GSR and the monthly average of SSD belong to the five cities of the south-eastern Anatolia region. These cities are Batman, Diyarbakir, Gaziantep, Mardin, and Sanliurfa according to the Turkish State Meteorological Service. The properties and content of datasets are given in tab. 1.

The datasets provided by Turkish State Meteorological Service are analysed by moving average approach of time series. When data is in the form of a time series, the series mean is a useful measure but does not reflect the dynamic nature of the data. Mean values computed over shorted periods, either preceding the current period or centered on the current period, are often more useful. Because such mean values will vary, or move, as the current period moves from time $t = 2$, $t = 3$, ... *etc.* they are known as moving averages

**Table 1. Data provided by Turkish State Meteorological Service**

| Location | Longitude (E) | Latitude (N) | Measured data | |
|---|---|---|---|---|
| | | | Period | Total years |
| Gaziantep | 37.22 | 37.04 | 1998-2010 | 12 |
| Sanliurfa | 38.46 | 37.07 | 1998-2010 | 12 |
| Diyarbakir | 40.13 | 37.55 | 1998-2008 | 10 |
| Batman | 41.07 | 37.52 | 1998-2006 | 8 |
| Mardin | 40.45 | 37.07 | 2012-2015 | 4 |

[17]. A simple moving average is (typically) the unweighted average of $k$ prior values. An EWMA is essentially the same as a simple moving average, but with contributions to the mean weighted by their proximity to the current time. Because there is not one, but a whole series of moving averages for any given series, the set of moving averages can themselves be plotted on graphs, analyzed as a series, and used in modelling and forecasting [18, 19]. A range of models can be constructed using moving averages, and these are known as MA models. If such models are combined with autoregressive models the resulting composite models are known as ARMA or ARIMA models [20]. However, that search will focus on EWMA that is correlated with exponentially weighted.

Since a time series can be regarded as a set of values, $y_t$, $t = 1, 2, 3,\dots, i$ the average of these values can be computed. If we assume that $i$ is quite large, and we select an integer which is much smaller than $n$, we can compute a set of block averages, or simple moving averages (of order $n$):

$$\overline{y}_{t.1} = \frac{1}{n}\sum_{t=1}^{n} y_t \tag{1}$$

$$\overline{y}_{t.2} = \frac{1}{n}\sum_{t=2}^{n+1} y_t \tag{2}$$

$$\overline{y}_{t.i-n+1} = \frac{1}{n}\sum_{t=i-n+1}^{t} y_t \tag{3}$$

where $2 \le n \le i$, each calculation of the average of the values over an interval of $n$, the data becomes:

$$\overline{y_t} = \frac{1}{n}\sum_{t=t-n+1}^{t} y_t \tag{4}$$

The aforementioned reveals that the average estimation at time, $t$, is the simple average of the $n$ value at time $t$ and this leads up to $n - 1$ time step. When weights are applied decreasing the number of $n$ that are next in time, the moving average will be then be called, exponentially smoothed [21]. Therefore, moving averages are usually provided forecasting information at a series time, $t + 1$. The S$t$+1 is considered the moving average for the period of time $t$, *e. g.*, today's forecast is based on an average of earlier values [17]. Using, eq. (4), all $n$'s are of equal weight. Said equal weights are assumed as $\mu_t$, every $n$ weight would equal $1/n$, so the sum of the weights would be 1, where $\mu_t = 1/n$, eq. (4), and will then become:

$$\overline{y_t} = \sum_{t=t-n+1}^{t} \mu_t y_t \tag{5}$$

Using EWMA, the contribution to the mean value from $n$'s that are more so removed in time is planned decreased, therefore emphasizes more local events. Basically, a smoothing parameter is $0 < \mu < 1$ where $0 < \mu < 0.5$ and will designate more weight than the prior, $y_t$.This is when $0.5 < \mu < 1$ is less weight assigned to $y_t - 1$ and more to $y_t$. In exponential smoothing it is needed to use a set of weights with sum equal to 1, to reduce in size geometrically [22]. The weights are used would be:

$$\mu(1-\mu)^k \tag{6}$$

where $k = 1, 2, 3, \ldots, \infty$. After some mathematical operations and reduction, the moving average that is weighted with eq. (5), eq. (6) becomes:

$$\overline{y}_t = \sum_{k=1}^{n} \mu \left(1-\mu\right)^{k-1} y_{t-k+1} \tag{7}$$

Then eq. (5), can be written as a repeated smoothed relation.

$$S_t = \mu y_t + \left(1-\mu\right) y_{t-1} \tag{8}$$

The algorithm of the smoothed repeated EWMA show in fig. 1 where $y(t, 1)$ and $S(t+2,1)$ is indicated during the first month of $t$ year and the first predicted month of $t+2$, respectively, whereas $N$ showed a predicted number of years. In EWMA operational algorithms, every month of the year from month one to month twelve compute separately due to smoothing operation.

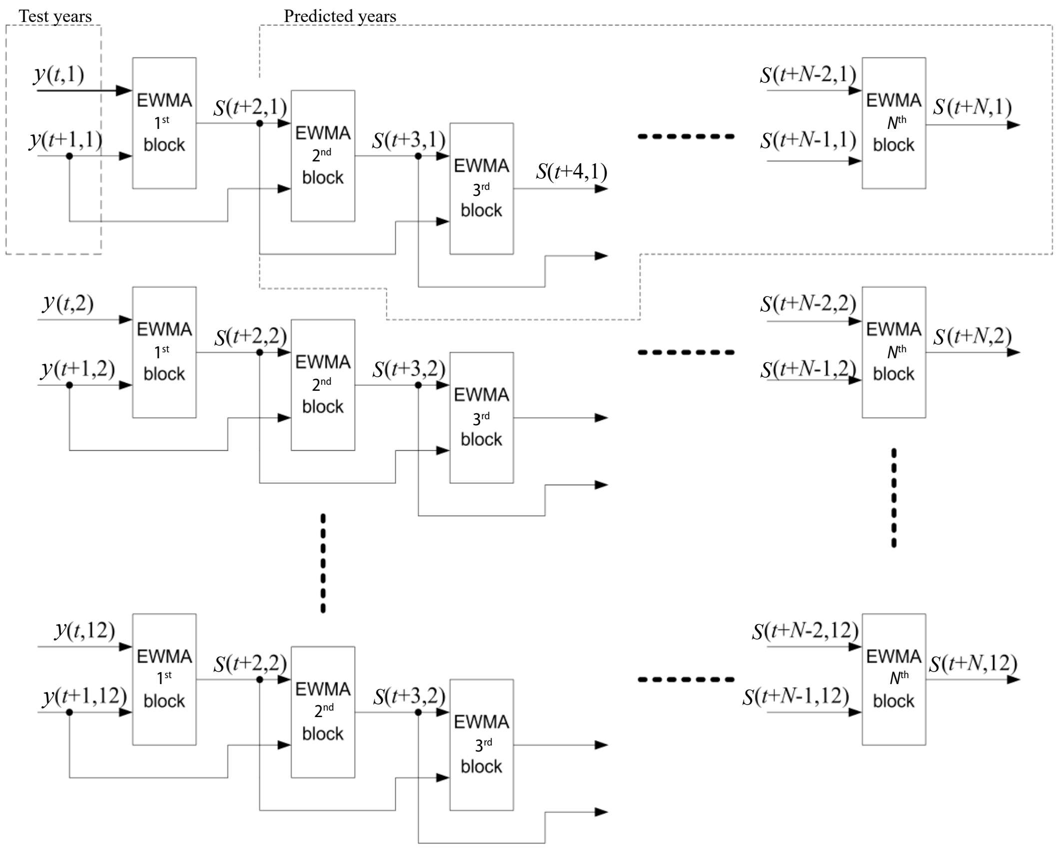


**Figure 1. Operation diagram of EWMA approach**

The data provided by Turkish State Meteorological Service are shown in tab. 2. It has 24 samples for each day and a summation of each sample which provide the total daily samples. To obtain monthly average data, total daily data is averaged for each month which constructs a dataset for one year. In addition, Microsoft EXCEL and MATLAB were used for both statistical data and other mathematical operations for analyzing and computing. The data for the first two years belonging to the city are used as a test class and the other years are then predicted. The years that are predicted from the test class data are compared with measured data in terms of coefficient of determination ($R^2$) and MAPE. The accuracy of prediction depends on the $R^2$ and MAPE [23] *i. e.*: 0 ≤ MAPE ≤ 10 indicates excellent prediction accuracy, 10 ≤ MAPE ≤ 20 indicates good prediction, 20 ≤ MAPE ≤ ≤ 50, indicates decent prediction while MEPA ≥ 50, indicates inaccurate prediction. On the other hand, $R^2$ means predicted data are fitted with measurement data. An $R^2$ closer to 1 indicates that the regression line highly fits the data, while an $R^2$ closer to 0 indicates that the predicted data does not fit the data at all. [24]. Definitions of $R^2$ and MAPE are:

**Table 2. Data used in EWMA process**

| Location | Test class years | | Prediction class years | |
|---|---|---|---|---|
| | Period | Total years | Period | Total years |
| Gaziantep | 1998-1999 | 2 | 2000-2010 | 10 |
| Sanliurfa | 1998-1999 | 2 | 2000-2010 | 10 |
| Diyarbakir | 1998-1999 | 2 | 2000-2008 | 8 |
| Batman | 1998-1999 | 2 | 2000-2006 | 6 |
| Mardin | 2012-2013 | 2 | 2014-2015 | 2 |

$$R^2 = \frac{\sum_{j=1}^{N}\left(S_j - S_{j,\text{avg}}\right)\left(y_j - y_{j,\text{avg}}\right)}{\sqrt{\left[\sum_{j=1}^{N}\left(S_j - S_{j,\text{avg}}\right)^2\right]\left[\sum_{j=1}^{N}\left(y_j - y_{j,\text{avg}}\right)^2\right]}} \tag{9}$$

$$MAPE = \frac{1}{N}\sum_{j=1}^{N}\frac{\left(S_j - y_j\right)}{y_i}\cdot 100 \tag{10}$$

where $S_j$ is the predicted value of $y_j$ and $y_j$ is the measured value. The $S_{j,\text{avg}}$ , and $y_{j,\text{avg}}$ are their average values, respectively.

## Results and discussion

EWMA is a time dependent based prediction method, which input parameters of test class years determined in tab. 2 and outputs prediction class years. The basic process of EWMA is indicated in fig. 1. The comparison of predicted years GSR (GSRp) and measured years GSR (GSRm) are shown in fig. 2. Moreover, the comparison of predicted year SSD (SSDp) and measured years SSD (SSDm) are shown in fig. 3.

Furthermore, some statistical parameters of predicted GSR and SSD values are computed to check the accuracy of predicted values such as MAPE and $R^2$ shown in tab. 3. Here it is shown that although the statistical parameters of the predicted GSRp are better than the SSDp,

**Table 3. Computed MAPE and $R^2$ values**

| Location | Predicted years | MAPE | $R^2$ | MAPE | $R^2$ |
|---|---|---|---|---|---|
| | | GSR [kWhm$^{-2}$day$^{-1}$] | GSR | SSD [hour-day] | SSD |
| Gaziantep | 2000-2010 | 7.113 | 0.970 | 9.122 | 0.950 |
| Sanliurfa | 2010-2010 | 7.569 | 0.960 | 6.427 | 0.948 |
| Diyarbakir | 2000-2008 | 6.663 | 0.968 | 7.459 | 0.935 |
| Batman | 2000-2006 | 4.780 | 0.958 | 6.689 | 0.939 |
| Mardin | 2014-2015 | 6.998 | 0.946 | 6.468 | 0.941 |

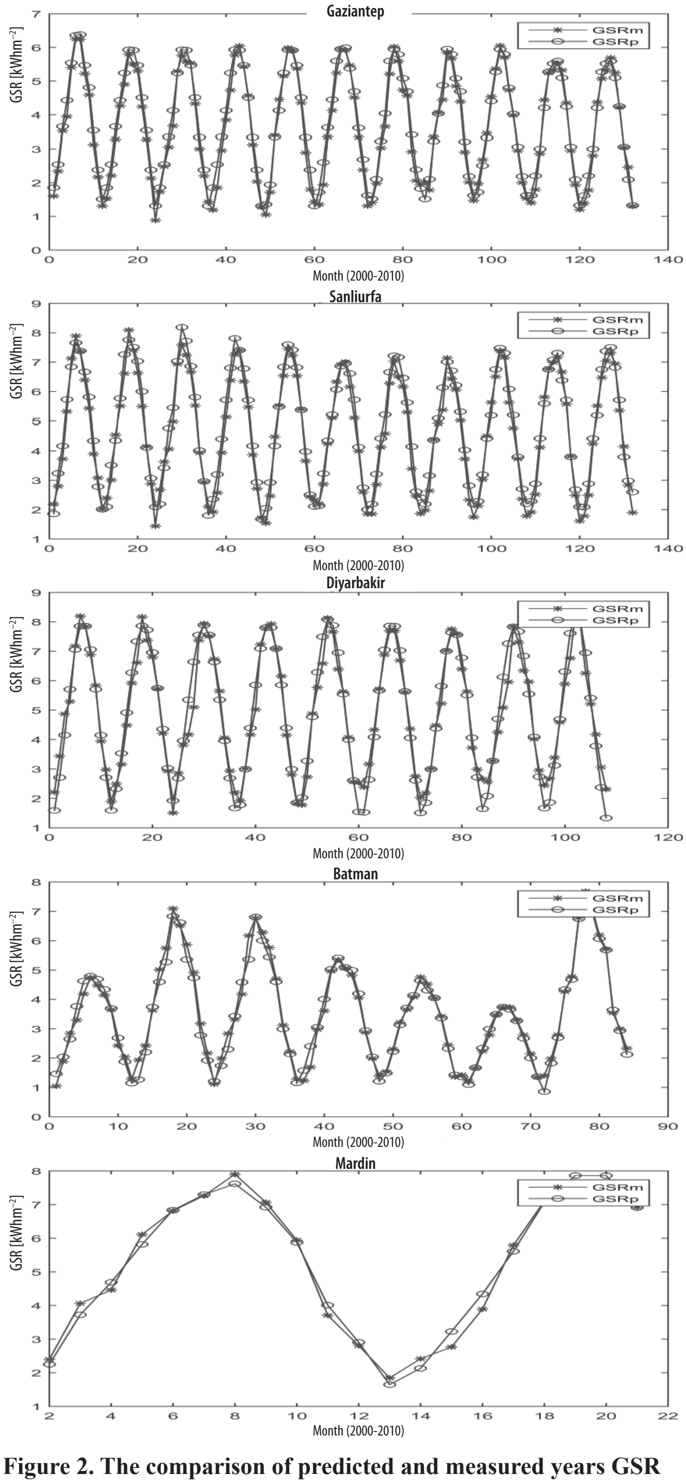


**Figure 2. The comparison of predicted and measured years GSR**

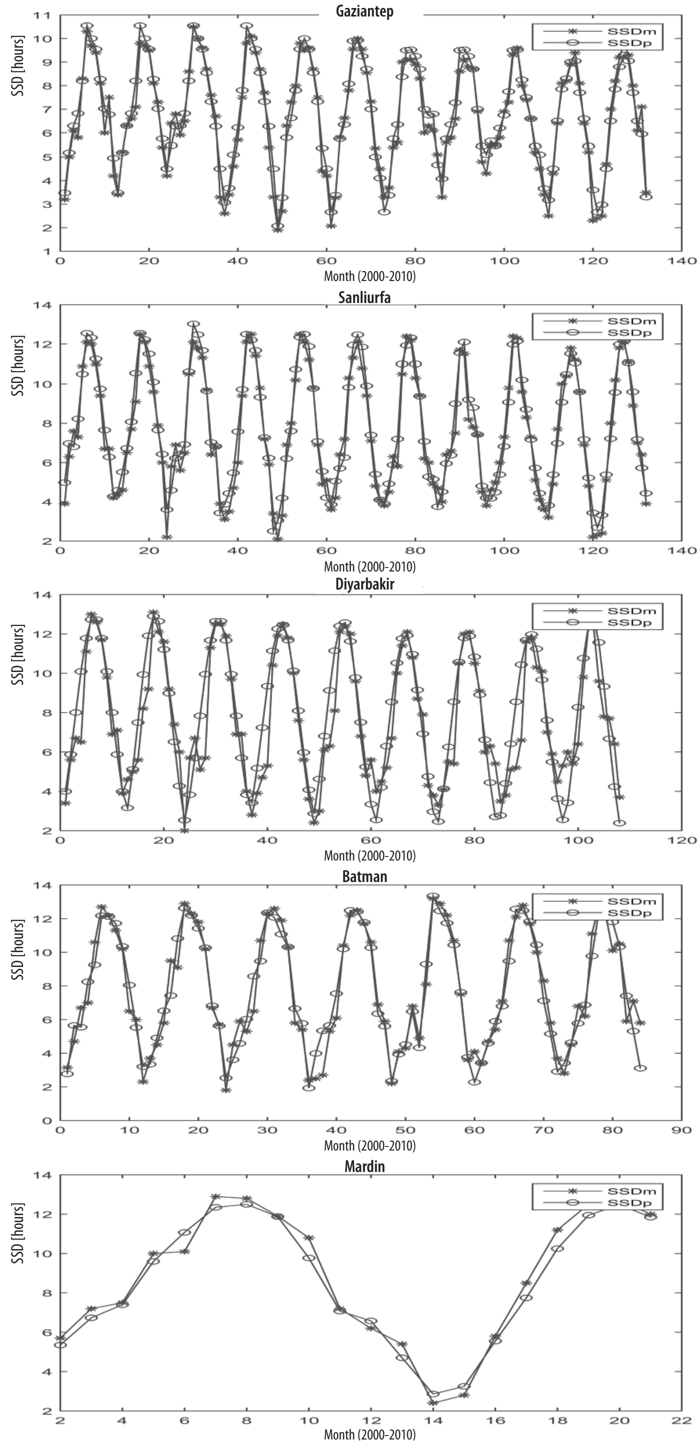


**Figure 3. The comparison of predicted and measured year's SSD**

both statistical parameters are in acceptable range in view of accuracy. The MAPE values of both GSR and SSD are smaller than 10 and means excellent prediction accuracy. Moreover, when $R^2$ values of both GSR and SSD are close to 1 that means predicted data are fitted with measurement data.

## Conclusions

In this study, time depended on EWMA series approach is used in the south-eastern region of Turkey for five biggest cities Batman, Diyarbakir, Gaziantep, Sanliurfa, and Mardin. Data of first two years from each city except Mardin are chosen for long-term prediction and Mardin's data of first two years are chosen for short-term prediction. The results show that time series based on EWMA approach is acceptable in long and short-term prediction, it is more efficient especially in cases which data of GSR and SSD is inaccessible.

Because of the fact that measurement stations are removed by Turkish State Meteorological Service in the south-eastern region of Turkey since August of 2015, data about GSR and SSD will not be gained furthermore. The method gives an idea for next years' data to obtain design criteria for solar energy systems because this region has a high potential for solar energy. Therefore, the predicted information of GSR and SSD provides noteworthy knowledge for investors in the field of solar energy generation, distribution, and transmission.

## Acknowledgment

This research was funded by a grant from the Scientific and Technological Research Council of Turkey (TUBITAK).

## References

[1] Pierro, M., *et al.*, Model Output Statistics Cascade to Improve Day Ahead Solar Irradiance Forecast, *Sol. Energy, 117* (2015), July, pp. 99-113
[2] Petrovic, I., *et al.*, Advanced PV Plant Planning based on Measured Energy Production Results – Approach and Measured Data Processing, *Adv. Electr. Comput. Eng., 14* (2014), 1, pp. 49-54
[3] Yadav, A. K., Chandel, S. S., Solar Radiation Prediction Using Artificial Neural Network Techniques: A Review, *Renew. Sustain. Energy Rev., 33* (2014), May, pp. 772-781
[4] Ravichandra, S., Rathnaraj, J. D., Analysis of Ratio of Global to Extra-Terrestrial Radiation (Clearness Index) at some Tropical Locations in India, *Thermal Science, 21* (2017), 3, pp. 1379-1387
[5] Nazaripouya, H., *et al.*, Univariate Time Series Prediction of Solar Power Using a Hybrid Wavelet-ARMA-NARX Prediction Method, *IEEE PES T&D*, Dallas, Tex., USA, 2016
[6] Shamim, M. A., *et al.*, An Improved Technique for Global Solar Radiation Estimation Using Numerical Weather Prediction, *J. Atmos. Solar-Terrestrial Phys., 129* (2015), July, pp. 13-22
[7] Adeyemi, A., *et al.*, Evaluation of Global Solar Radiation Using Multiple Weather Parameters as Predictors for South Africa Provinces, *Thermal Science, 19* (2015), Suppl. 2, pp. S495-S509
[8] Bakirci, K., Models for the Estimation of Diffuse Solar Radiation for Typical Cities in Turkey, *Energy, 82* (2015), Mar., pp. 827-838
[9] Besharat, F., *et al.*, Empirical Models for Estimating Global Solar Radiation: A Review and Case Study, *Renew. Sustain. Energy Rev., 21* (2013), May, pp. 798-821
[10] Qazi, A., *et al.,* The Artificial Neural Network for Solar Radiation Prediction and Designing Solar Systems: A Systematic Literature Review, *J. Clean. Prod., 104* (2015), Oct., pp. 1-12
[11] McCandless, T. C., *et al.*, A Model Tree Approach to Forecasting Solar Irradiance Variability, *Sol. Energy, 120* (2015), Oct., pp. 514-524
[12] Shukla, K. N., *et al.*, Comparative Study of Isotropic and Anisotropic Sky Models to Estimate Solar Radiation Incident on Tilted Surface: A Case Study for Bhopal, India, *Energy Reports, 1* (2015), Nov., pp. 96-103
[13] Kisi, O., Modeling Solar Radiation of Mediterranean Region in Turkey by Using Fuzzy Genetic Approach, *Energy, 64* (2014), Jan., pp. 429-436

[14] Guclu, Y. S., *et al.*, Solar Irradiation Estimations and Comparisons by Anfis, Angstrom-Prescott and Dependency Models, *Sol. Energy, 109* (2014), Nov., pp. 118-124
[15] Guclu, Y. S., *et al.*, HARmonic-LINear (HarLin) Model for Solar Irradiation Estimation, *Renew. Energy, 81* (2015), Sept., pp. 209-218
[16] Gerek, O. N., *et al.*, Harmonic Analysis Based Hourly Solar Radiation Forecasting Model, *IET Renew. Power Gener., 9* (2015), 3, pp. 218-227
[17] Park, J.-K., *et al.*, A New Approach to Estimate the Spatial Distribution of Solar Radiation Using Topographic Factor and Sunshine Duration in South Korea, *Energy Convers. Manag., 101* (2015), Sept., pp. 30-39
[18] De Andrade, R. C., Tiba, C., Extreme Global Solar Irradiance Due to Cloud Enhancement in Northeastern Brazil, *Renew. Energy, 86* (2015), Feb., pp. 1433-1441
[19] Teke, A., Yildirim, H. B., Estimating the Monthly Global Solar Radiation for Eastern Mediterranean Region, *Energy Convers. Manag., 87* (2014), Nov., pp. 628-635
[20] Holt, C. C., Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages, *Int. J. Forecast., 20* (2014), 1, pp. 5-10
[21] Kentli, F., Yilmaz, M., Mathematical Modelling of Two-Axis Photovoltaic System with Improved Efficiency, *Elektronika Ir Elektrotechnika, 21* (2015), 4, pp. 40-43
[22] Boland, J., Time-Series Analysis of Climatic Variables, *Solar Energy, 55* (1995), 5, pp. 377-388
[23] Brockwell, P., Davis, R. A., *Introduction to Time Series and Forecasting*, Springer-Verlag, New York, USA, 1996
[24] Bowerman, B. L., O'Connell, R. T., *Time Series Forecasting*, Duxbury Press, Boston, Mass., USA, 1979

Paper submitted: January 7, 2016
Paper revised: August 6, 2016
Paper accepted: September 20, 2016