



Statistical Models For Sunshine Duration Related To Precipitation And Relative Humidity

Mine Tulin Zateroglu^{1*}

¹ Cukurova University, Vocational School of AOSB Technical Sciences, Department of Electrical and Energy, Adana, Turkey (ORCID: 0000-0002-1050-6174)

(International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) 2021 – 21-23 October 2021)

(DOI: 10.31590/ejosat.1022962)

ATIF/REFERENCE: Zateroglu, M.T. (2021). Statistical Models For Sunshine Duration Related to Precipitation And Relative Humidity. *European Journal of Science and Technology*, (29), 208-213.

Abstract

Energy usage leads to environmental problems especially climate change and air pollution. In urban areas, extensive consumption of fossil energy sources deteriorates urban air quality and causes health problems such as asthma and respiratory diseases. The usage of renewable energy sources mainly solar energy instead of fossil energy sources provide the healthier and cleaner environment. Some statistical approaches have been used to predict the solar radiation for many years. The formula which is named as Angström-Prescott is widely used for the estimation. This approach uses the value of relative sunshine duration which is computed by using the daily measurement of sunshine duration and the daylength. Sunshine duration has been measured at meteorological stations for a long time. But in some cases such as remote areas or non-exist stations, sunshine duration is forecasted by using statistical methods. In atmospheric environment, sunshine duration is affected from other climate parameters such as cloud cover, wind speed, relative humidity, precipitation, air temperature, pressure. This paper offers to gain the relationships for sunshine duration via precipitation and relative humidity. To construct the statistical models, cloud cover was entered the models firstly because of its association with sunshine duration. Many studies were shown that sunshine duration was highly correlated to the cloud cover. The observed daily mean values of sunshine duration and climate elements were the data of Samsun meteorological station measured by Turkish State Meteorological Service. To acquire the statistical models for sunshine duration, data was arranged as monthly mean values and then linear regression analysis method was operated. Also the graphs of time series of climate variables were visually created for a clear interpretation. According to the findings, sunshine duration can be estimated bu using statistical models over cloud cover, precipitation and relative humidity. The contributions of precipitation and relative humidity change for different time scales.

Keywords: Sunshine Duration, Cloud Cover, Precipitation, Relative Humidity, Regression Analysis.

Güneşlenme Süresinin Yağış Miktarı ve Nisbi Nem İle İlişkili İstatistiksel Modelleri

Öz

Enerji kullanımı başta iklim değişikliği ve hava kirliliği olmak üzere çevre sorunlarına yol açmaktadır. Kentsel alanlarda fosil enerji kaynaklarının yoğun tüketimi kentsel hava kalitesini bozmakta, astım ve solunum yolu hastalıkları gibi sağlık sorunlarına neden olmaktadır. Fosil enerji kaynakları yerine güneş enerjisi başta olmak üzere yenilenebilir enerji kaynaklarının kullanılması daha sağlıklı ve temiz bir çevre sağlamaktadır. Güneş radyasyonunu tahmin etmek için uzun yıllardır bazı istatistiksel yaklaşımlar kullanılmıştır. Tahmin için Angström-Prescott olarak adlandırılan formül yaygın olarak kullanılmaktadır. Bu yaklaşım, günlük güneşlenme süresi ve gün uzunluğu ölçümü kullanılarak hesaplanan bağıl güneşlenme süresi değerini kullanır. Güneşlenme süresi uzun zamandır meteoroloji istasyonlarında ölçülmektedir. Ancak istasyon bulunmayan veya uzak bölgeler gibi bazı durumlarda, istatistiksel yöntemler kullanılarak güneşlenme süresi tahmin edilir. Atmosferik ortamda güneşlenme süresi, bulutla kaplılık, rüzgar hızı, nisbi nem, yağış miktarı, hava sıcaklığı, atmosferik basınç gibi diğer iklim parametrelerinden etkilenir. Bu makale, yağış miktarı ve nisbi nem yoluyla güneşlenme süresi bağıntılarının elde edilmesini önermektedir. İstatistiksel modelleri oluşturmak için güneşlenme süresi ile olan ilişkisinden dolayı modellere ilk olarak bulutla kaplılık eklenmiştir. Birçok çalışma, güneşlenme süresinin bulutla kaplılık ile yüksek oranda ilişkili olduğunu göstermiştir. Güneşlenme süresi ve iklim elemanlarının gözlemlenen günlük ortalama değerleri, Devlet Meteoroloji İşleri tarafından ölçülen Samsun meteoroloji istasyonuna ait verilerdir. Güneşlenme süresi için istatistiksel modeller elde etmek için veriler aylık ortalama değerler olarak düzenlenmiş ve ardından doğrusal regresyon analizi yöntemi uygulanmıştır. Ayrıca iklim değişkenlerinin zaman serilerinin grafikleri de görsel olarak oluşturulmuştur. Elde edilen bulgulara göre, bulutla kaplılık, yağış miktarı ve nisbi nem

* Corresponding Author: Cukurova University, Vocational School of AOSB Technical Sciences, Department of Electrical and Energy, Adana, Turkey, ORCID: 0000-0002-1050-6174, mtzateroglu@cu.edu.tr

üzerinden istatistiksel modeller kullanılarak güneşlenme süresi tahmin edilebilmektedir. Farklı zaman ölçeklerinde yağış miktarı ve nisbi nemin katkıları değişmektedir.

Anahtar Kelimeler: Güneşlenme süresi, Bulutla Kaplılık, Yağış Miktarı, Nisbi Nem, Regresyon Analizi.

1. Introduction

Energy production and increasing energy consumption based on fossil fuels exhibit unsustainable structure for the development around the world. To avoid this problem, energy policies can be transformed to a sustainable structure which provide an increasing portion of renewable energy sources in whole energy consumption. Solar energy is basic and the most important energy source among the renewable energy sources. For a sustainable and cleaner world, it is necessary to benefit and use more solar energy.

Solar radiation for energy has been measured over the world for a long time. But these measurements are still not enough, so the prediction of solar radiation is more suitable to determine the value. Some statistical approaches have been used to predict the solar radiation. The formula which is named as Angström-Preseott is widely used for the estimation. This approach uses the value of relative sunshine duration which is computed by using the daily measurement of sunshine duration and the daylength. Sunshine duration is a key element for solar radiation and consequently for human life. Sunshine duration affects many areas such as tourism activities, planning the power plants, agriculture and etc. Also it has effects on human psychology.

Climatological variables being in the atmospheric phenomena have close interactions among themselves. So, sunshine duration which is one of the climate variable is associated with other climate variables. Especially, cloud cover which is defined as the part of the sky enclosed by clouds is highly correlated with sunshine duration according to the many studies (Randarajan et al., 1984; Palle&Butler, 2001; Essa&Etman, 2004; Weber, 1994; Angell et al., 1984; Sanchez-Lorenzo et al., 2009; Robaa, 2008; Matuszko, 2012). Clouds consist of liquid or solid ice particles in the atmosphere and obstruct the quantity of incoming direct solar irradiance which is used for measuring sunshine duration in equipments. Additionally, environmental elements in the atmosphere such air pollutants are related to sunshine duration and affect the amount of sunshine duration (Zateroglu, 2021a&2021b).

Samsun with an area of 9579 km² is located in the middle part of the Black Sea coastline, between the deltas where the Yeşilirmak and Kızılırmak rivers flow into the Black Sea. It is coordinated between the north latitudes 40° 50' – 41° 51' and east longitudes 37° 08' – 34° 25'. Samsun is surrounded by the Black Sea in the north, Ordu in the east, Sinop in the west, Tokat and Amasya in the south and Çorum in the southwest. The south of the area is mountainous, the north is coastal plains, and the middle part consists of plateaus. The province has generally a temperate climate. However, two different climate characteristics are experienced on the coastline and inland. The effects of the Black Sea climate are seen on the coastline. So, summers are hot, and winters are warm and rainy. For inland, due to the influence of the mountains, summers are cool and winters are cold, rainy and snowy. According to long term records, annual average relative humidity is 70,3(%), precipitation is 716,7 (mm), the number of rainy days is about 154, cloud cover is 5,9, sunshine duration is around 5,3(hour). The most precipitation occurs in autumn and winter whereas maximum value on November and minimum value on August. The month with the highest humidity is May and

the month with the lowest is December. Air masses coming from the sea and containing plenty of humidity cause the sky to be cloudy even in summer. The maximum value of cloud cover becomes on February with 7,6 and minimum values on July and August. For sunshine duration, the highest value is obtained on July whereas the lowest value on January. The class of climate for Thornthwaite climate classification is semi humid and 2nd mesothermal.

In this study, it is aimed to represent the relationships as statistical models to predict the sunshine duration over precipitation and relative humidity in addition to cloud cover. The long-term records of measured values for sunshine duration, cloud cover, precipitation and relative humidity were provided from Turkish State Meteorological Service for Samsun station. Daily sunshine duration data is measured by Campbell-Stokes equipment as hour per day. This sunshine recorder uses the quantity of the incoming direct solar irradiance to measure the daily sunshine duration. The measurement of the cloud cover (CLC) is provided by visual observations. Its value changes 0 (uncloudy) to 10 (completely cloudy). Precipitation (PREC) and relative humidity (RHM) is measured by pluviograph and hygograph, respectively. In lieu of the term sunshine duration, the statement relative sunshine duration is preferred in researches, therefore prior to examining the statistical studies, sunshine duration measurements were transformed to index values as relative sunshine duration.

2. Material and Method

2.1. Relative Sunshine Duration

The term of relative sunshine duration (RSD) was used in this study because of it is preferred in scientific papers instead of measured value. The value of RSD is between 0 and 1. It determines the ratio of measured sunshine duration (S) to daylength (S₀). S is measured as hour/day in meteorological station by Campbell-Stokes equipment. S₀ is calculated from formula refer to (1).

Equation (2) demonstrates the calculation of the declination angle which is defined as the angle between the equatorial plane and the incoming solar rays. φ expresses the latitude angle ($-90 \leq \varphi \leq +90$) that varies according to the location. Daylength and declination angle are computed by formulae as follows (Duffie&Beckman, 1991; Kalogirou, 2009),

$$S_0 = \left(\frac{2}{15}\right) \cos^{-1}(-\tan\delta \tan\varphi) \quad (1)$$

$$\delta = 23.45 * \sin\left(\frac{360}{365}(284 + d)\right) \quad (2)$$

where d determines the day number of the year starting from January 1.

2.2. Multiple Linear Regression

To construct the statistical models in estimating the RSD qua statistically meaningful (for the confidence interval of %95), linear regression analysis method was utilized over CLC, PREC and RHM. The regression analysis technique has been generally operated in predicting the climatological parameters. By using the linear regression, the relations between RSD and other variables may be demonstrated as mathematical models. Estimating the RSD over climate variables was implemented by using the Equation (3) which determines the relation as a mathematical model,

$$RSD = a_0 + a_1CLC + a_2PREC + a_3RHM + \varepsilon \quad (3)$$

where the statement RSD expresses the dependent variable, the expressions CLC, PREC and RHM denote the independent variables, ε determines the error phrase associated with the regression, the terms a_0 and a_1 demonstrate the regression coefficients.

There has been two measures to interpret the consistency of the mathematical models obtained via linear regression, which are called as the names standard error and multiple determination coefficient.

The statement of determination coefficient is expressed as R^2 and used as a scale to indicate the fitting of estimated model with measured values. The measure R^2 denotes the computed value as percent and varies 0 (means that the dependent variable is unexplained by the independent variables) to 1 (expresses that the dependent variable is explained by the independent variables in highest level). The value of multiple determination coefficient is calculated for the statistical model by Equation (4) as follows:

$$R^2 = 1 - \frac{\sum(Y_{e,i} - \bar{Y})^2}{\sum(Y_{o,i} - \bar{Y})^2} \quad (4)$$

where the term \bar{Y} denotes the average value of the measured values Y_i , the phrase Y_o expresses the observed value of the Y_i , the statement Y_e demonstrates the estimated value of the Y_i . The value 1 for obtained R^2 determines that the model is consistent with the dataset in high level, but the value 0 denotes the model is inconsistent with the data.

The second term to evaluate the statistical estimated models is the expression standard error. It is used to interpret the accuracy of the constructed models. The term Standard Error of Estimation (SEE) determines a scale which demonstrates the difference

between predicted and observed values. This expression indicates that the amount of the deviation the measured values from the estimated values on average and computed by the formula as shown in Equation (5),

$$SEE = \sqrt{\frac{\sum(Y_o - Y_e)^2}{n - 2}} \quad (5)$$

where the term Y_o defines the observed value, the phrase Y_e denotes the estimated value and n depicts the number of observations.

2.3. Normalization index

Prior to constructing the figures of climate variables, the studied dataset for each variable were transformed as standardized values for interpreting more clear the analysis. The all climate data were normalized over the formula shown in Equation (6),

$$x_i = \frac{x_j - x_{min}}{x_{max} - x_{min}} \quad (6)$$

where the term x_i expresses the normalized value of x_j that defines the observed data. The statement x_{min} determines the minimum value of the dataset of variables. The phrase x_{max} indicates the maximum value of the given dataset. The converted values for the interval 0 to 1 are presented by Equation (6).

3. Results and Discussion

The statistical approach was performed for climate variables determined in the study. Multiple linear regression technique was employed as statistical analysis method. This method is commonly operated in studies about estimating of meteorological parameters. Before the mathematical models were created, the climate data for variables were examined to depict if they fit the normal distribution.

The One-Sample Kolmogorov-Smirnov test which is one of the normality tests was operated on dataset to determine the distribution of data. This normality test is based on the perusal of two cumulative distribution functions (Gangam, 1998: 196). Table 1 was shown the results of Kolmogorov-Smirnov test for all climate variables used in this study. For all months, the significance level values (Asymp. Sig., 2-tailed) obtained against the calculated Z-values (Kolmogorov-Smirnov Z) were acquired as bigger than the statistical significance level 0,05 (p-value). So, the dataset was found to fit the normal distribution.

Table 1. One-Sample Kolmogorov-Smirnov Test Results

Month	Statistical Parameter	RSD	CLC	RHM	PREC
JANUARY	Kolmogorov-Smirnov Z	0,943307	0,526486	0,505865	0,806281
	Asymp. Sig. (2-tailed)	0,335773	0,944441	0,96007	0,533957
FEBRUARY	Kolmogorov-Smirnov Z	0,455231	0,768832	0,484143	1,297479
	Asymp. Sig. (2-tailed)	0,985697	0,595578	0,973191	0,06899
MARCH	Kolmogorov-Smirnov Z	0,899901	0,725153	0,455989	0,607274
	Asymp. Sig. (2-tailed)	0,392868	0,669058	0,985435	0,854508
APRIL	Kolmogorov-Smirnov Z	0,606387	0,733294	0,555444	1,054929
	Asymp. Sig. (2-tailed)	0,855715	0,655331	0,917245	0,215701
MAY	Kolmogorov-Smirnov Z	0,506201	0,461639	0,882344	0,880872
	Asymp. Sig. (2-tailed)	0,95984	0,983379	0,417564	0,419676
JUNE	Kolmogorov-Smirnov Z	0,882103	0,796396	0,575967	0,935326
	Asymp. Sig. (2-tailed)	0,41791	0,550017	0,894417	0,345837
JULY	Kolmogorov-Smirnov Z	0,6104	0,594358	0,555166	1,042735
	Asymp. Sig. (2-tailed)	0,850219	0,871664	0,917534	0,226975
AUGUST	Kolmogorov-Smirnov Z	0,40598	0,689534	0,788013	1,024812
	Asymp. Sig. (2-tailed)	0,996534	0,72858	0,563764	0,244346
SEPTEMBER	Kolmogorov-Smirnov Z	0,855049	0,486842	0,692112	0,785018
	Asymp. Sig. (2-tailed)	0,457682	0,971743	0,724323	0,568699
OCTOBER	Kolmogorov-Smirnov Z	0,527254	0,791368	0,626949	1,232445
	Asymp. Sig. (2-tailed)	0,943798	0,558248	0,826719	0,095866
NOVEMBER	Kolmogorov-Smirnov Z	0,982639	1,163308	0,563087	0,918649
	Asymp. Sig. (2-tailed)	0,289076	0,133491	0,909075	0,367501
DECEMBER	Kolmogorov-Smirnov Z	0,659392	1,01419	0,60896	1,005533
	Asymp. Sig. (2-tailed)	0,777324	0,255099	0,852201	0,264116

In Table 1, it was demonstrated the mathematical models to predict the RSD. These models were provided from stepwise regression analysis which is one of the multiple linear regression analysis methods as statistically meaningful. This method takes into account the variables that contribute the most to the model for each month. The periods were determined from January to December as JAN, FEB, MAR, APR, MAY, JUN, JUL, AUG, SEP, OCT, NOV, and DEC. Besides, the values of statistical

indexes were computed using formulas mentioned above and presented for each model. The grade of the models was high for MAR, APR, MAY, OCT and NOV, while moderate for JAN, FEB, JUN and DEC. But for JUL and AUG, the degree of produced models was obtained as weak. Probably, the other climate factors may be effective during these months or there may be errors in measurements.

Table 2. Statistical Models and Indexes for RSD

Month	Model	R	R ²	SEE	Significance
JAN	0,767-0,069*CLC-0,057*PREC-0,086*RHM	0,736	0,542	0,04756	0,000
FEB	0,817-0,077*CLC+0,001*PREC	0,815	0,664	0,04540	0,000
MAR	1,078-0,084*CLC-0,002*RHM	0,895	0,802	0,03762	0,000
APR	1,015-0,095*CLC-0,001*RHM	0,935	0,875	0,03552	0,000
MAY	1,228-0,085*CLC-0,004*RHM	0,927	0,859	0,03855	0,000
JUN	0,941-0,086*CLC	0,818	0,67	0,04984	0,000
JUL	0,808-0,055*CLC	0,452	0,204	0,07517	0,001
AUG	0,874-0,068*CLC	0,533	0,284	0,07562	0,000
SEP	0,916-0,076*CLD-0,001*RHM	0,68	0,462	0,05651	0,000
OCT	0,898-0,083*CLC-0,068*RHM-0,027*PREC	0,865	0,748	0,05111	0,000
NOV	0,884-0,082*CLC-0,024*PREC	0,872	0,761	0,04810	0,000
DEC	0,806-0,079*CLC+0,014*RHM	0,779	0,608	0,05405	0,000

Additionally, the time series for studied climate variables as pairwise were presented in Figures 1-3. The normalized index values for RSD versus CLC were constructed in Figure 1, for RSD versus PREC in Figures (2) and finally for RSD versus RHM in

Figure (3) for April session during 1960-2010. The graphs of the climate variables demonstrate the variations over the years to evaluate the temporal vision.

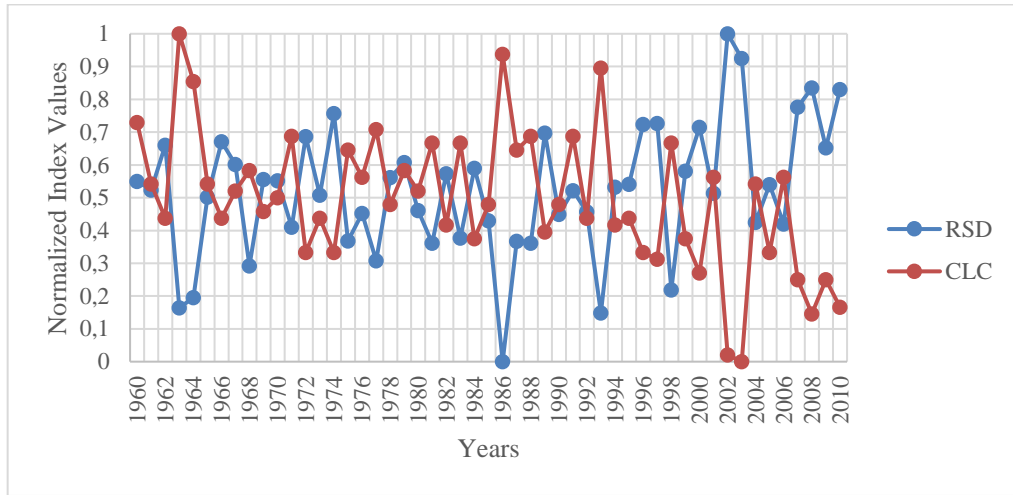


Figure 1. Variations of RSD and CLC

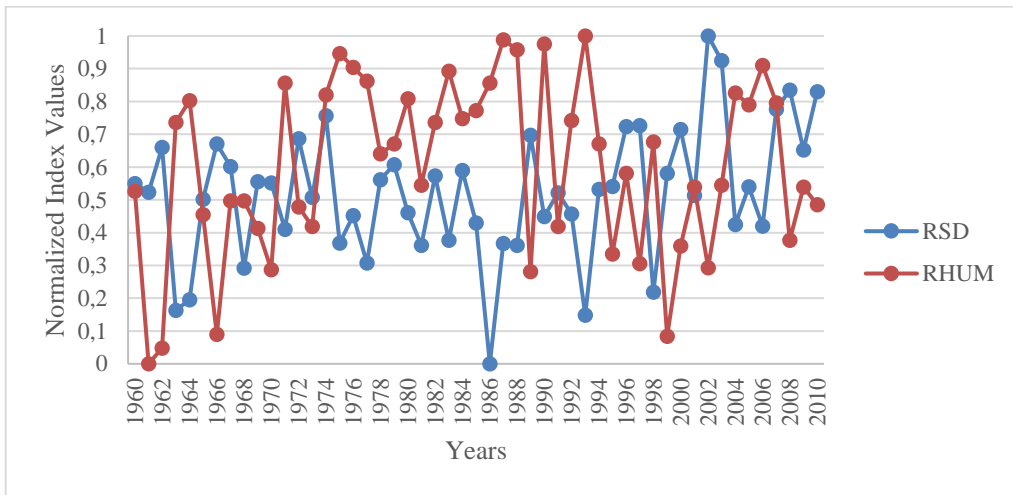


Figure 2. Variations of RSD and RHUM

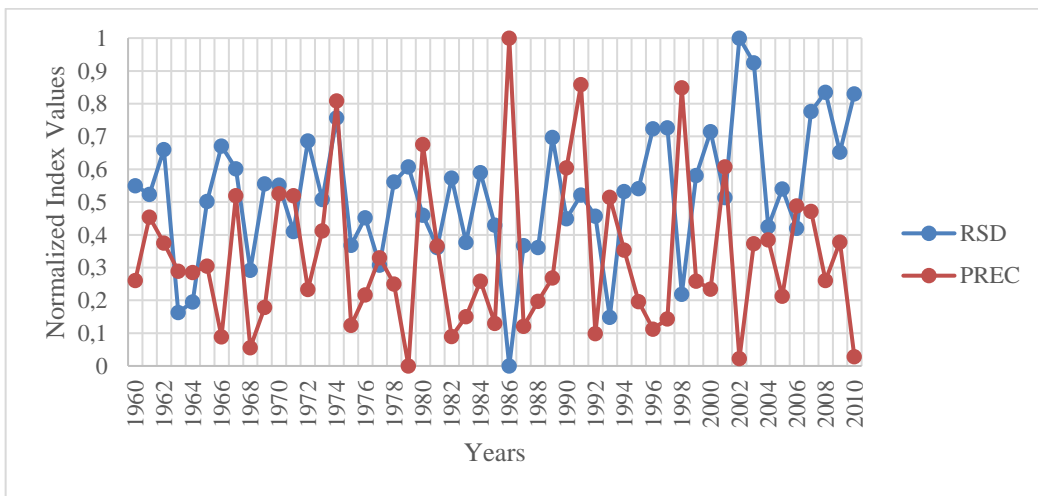


Figure 3. Variations of RSD and PREC

In all Figures 1 to 3, although there have been some exceptional circumstances, it was visualized that RSD was varied

with CLC, PREC and RHM inversely. The high values of observed RSD correspond to the low values of observed CLC,

RHM and PREC. The exceptional circumstances in the graphs may be comprised presumably due to incorrect measurement values. Further, it can be obviously interpreted that, while the sky is enveloped by the cloud cover, the quantity of sunshine duration decreases as expected.

4. Conclusions and Recommendations

This study depicted that sunshine duration was related to the climate variables i.e. cloud cover, precipitation and relative humidity. According to the findings, the constructed statistical models coming from regression analysis method were indicated that the grade of the relationships was changed for each session. Furthermore, the temporal variations of pairwise datasets were demonstrated that the variations in sunshine duration were oppositely associated with the changes of cloud cover, precipitation and relative humidity as expected in spite of some exceptional circumstances presumably because of mistaken observations. Moreover, statistical index value for determination coefficient obtained from mathematical models were acquired as high and moderate degrees but for the two months July and August, determination coefficient values were found as weak levels perhaps because of incorrect measurements or the effects of other climate factors. This situation may be discussed in further studies taking into account the other climate variables. Besides, geographical structure and atmospheric circulations for a location may affect the climate variables and so sunshine duration. The results were shown that the amount of sunshine duration gets high values with low cloud cover, precipitation and relative humidity and vice versa. Eventually, according to the findings, sunshine duration can be estimated by using statistical models over cloud cover, precipitation and relative humidity. The contributions of precipitation and relative humidity change for different time scales.

5. Acknowledge

The author would like to thank to the Turkish State Meteorological Service for providing the data for this work.

References

- Angell J.K., Korshover, J., & Cotton, G.F. (1984). Variation in United States Cloudiness and Sunshine, 1950-82. *Journal of Applied Meteorology*, 23, 752-761.
- Duffie J. A. & Beckman, W. A. (1991). Solar Engineering of Thermal Processes. *John Wiley & Sons*, New York, NY, USA, 2nd edition.
- Essa, K. S., & Etman, M. S., (2004). On the Relation Between Cloud Cover Amount and Sunshine Duration. *Meteorology and Atmospheric Physics*, 87, 235-240.
- Gamgam, H. (1998). *Parametrik Olmayan İstatistiksel Teknikler*, Ankara, Gazi Üniversitesi Yayını.
- Kalogirou, S.A. (2009). Solar Energy Engineering Processes and Systems. *Elsevier Inc*, Amsterdam, The Netherlands, 1st edition.
- Matuszko, D. (2012). Influence of cloudiness on sunshine duration. *International Journal of Climatology*, 32(10), 1527– 1536.
- Palle, E., & Butler, C.J., (2001). Sunshine Records From Ireland: Cloud Factors and Possible Links to Solar Activity and Cosmic Rays. *International Journal of Climatology*, 21, 709-729. <https://doi.org/10.102/joc.657>.
- Rangarajan, S., Swaminathan, M.S., & Mani, A. (1984). Computation of solar radiation from observations of cloud cover. *Solar Energy*, 32(4), 553–556.
- Robaa, S.M. (2008). Evaluation of Sunshine Duration From Cloud Data in Egypt. *Energy*, 33(5), 785-795.
- Sanchez-Lorenzo, A., Calbo, J., Brunetti, M., & Deser, C. (2009). Dimming/brightening over the Iberian Peninsula: trends in sunshine duration and cloud cover and their relations with atmospheric circulation. *Journal of Geophysical Research D: Atmospheres*, 114(10), Article ID D00d09.
- Weber, G.R. (1994). On the Seasonal Variation of Local Relationships Between Temperature, Temperature Range, Sunshine and Cloudiness. *Theoretical and Applied Climatology*, 50(1-2), 15-22.
- Zateroglu, M.T. (2021a). Assessment of the Effects of Air Pollution Parameters on Sunshine Duration in Six Cities in Turkey. *Fresenius Environmental Bulletin*. 30(02A), 2251-2269.
- Zateroglu, M.T. (2021b). The Role of Climate Factors on Air Pollutants (PM₁₀ and SO₂). *Fresenius Environmental Bulletin*. 30(11), 12029-12036.