Caspian Journal of Applied Mathematics, Ecology and Economics
V. 2, No 1, 2014, July
ISSN 1560-4055

Features Of Regression Modeling Of Solar Radiation With Different Types Of Functions

F. Aliyev *, H. Khalilova, N. Agayev

Abstract. The paper investigates the characteristics of solar radiation process as stochastic processes. It was found that due to solar radiation multi factorial process modelling should be based on real statistical data. Time series of solar radiation was considered as a functional depending on four components:trend,more or less regular fluctuations relative to trend, seasonal fluctuations and non-systematic random effect. It is shown that non-systematic random effect can be defined as the difference between the statistics and the amount of trend and seasonality.Researches were conducted to determine the trend type on the basis of polynomial regression analysis and the Gaussian function, and seasonal variations on the basis of the Fourier series.

Key Words and Phrases: Solar radiation, time series, regression model, Gaussian function

1. Introduction

Simulation of solar radiation is necessary for design and operation of automatic control for photovoltaic systems. The ultimate goal of simulation is to calculate the dependence of the total amount of solar radiation on solar panels on geographic latitude, meteorological factors, day of year, time of day and the angle of inclination of a surface [1, 2].

Among techniques of analysis and calculation of solar radiation developed in recent years, the most fully studied one is the cloudless sky technique [3, 4].

However, a number of studies have shown that these methods give inaccurate results because of the significant influence of clouds on the amount of insolation caused by the weakening of direct radiation and an increase in diffuse fraction in most cases. Reduction of direct radiation due to increasing cloudiness is not compensated by dissipated insolation. Consequently, the total insolation, as well as the direct one, decreases with increasing cloudiness.

Simulation of the formation of clouds is a multi factorial problem. In this problem, cloudiness is usually taken as a random value with the distributive law that corresponds to real statistical data. For example, in [1] experimental observations are best described by the beta-law with the corresponding parameters of distribution. However, these results

*Corresponding author.

were based on average daily chart of the progress of solar radiation at the average cloudiness conditions. Optimal control of photovoltaic power system requires hourly values of cloudiness rate. This is not possible using the classical techniques as cloud formation is an uncontrollable time-varying process [5].

Radiation is a random variable below in Fig. 1. It is usually considered as a statistical phenomenon that develops over time according to the laws of the theory of probability. The sequence of observations is a time series, analysis of which can provide a stochastic model with a minimal number of parameters allowing calculation of the probability that some future value of insolation will lie within a certain range and at the same time adequately describing the process under study.

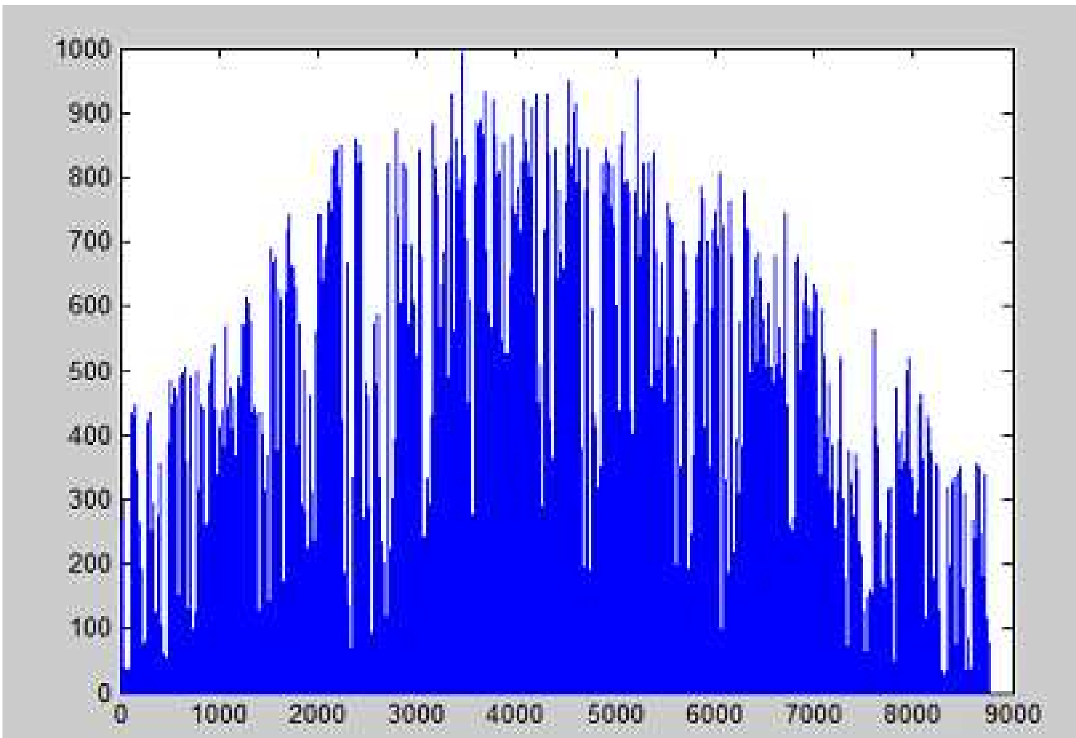


Figure 1. Hourly radiation in Baku in 2012 (X-axis shows the number of measurements, Z-axis shows total radiation in MJ/h).

Statistical approach is often used in the analysis of time series of solar radiation. Unlike other statistical objects, this time series has a characteristic feature: the observations are made sequentially in time and depend heavily on external influence. There is an unsystematic random effect of external factors in each measurement, which significantly complicates the process of simulation. If we consider several series of measurements as a multidimensional complex then we have to take into account the statistical relationship between these variables.

2. Setting objectives

It should be noted that in the general case, the time series of solar radiation is considered as a probability functional depending on four components: a trend, more or less regular fluctuations related to the trend, seasonal fluctuations and non-systematic random effect. Usually a time series is the sum of these four components. Let us consider each of them separately. The existence of a trend is explained by the presence of permanent forces operating uniformly in about the same direction of more or less regular fluctuations related to the trend which are happening due to the influence of regular perturbations that appear randomly. It is easier to present the components that are generated cyclically, such as daily and seasonal variations of solar radiation. But it should be noted that the values of these components are subject to daily and yearly changes. Nevertheless, a qualitative picture remains, which means that the seasonal effect has a trend. In general, if you can determine the trend and seasonal variations and subtract them from the data, then you have a fluctuating time series, which may represent a purely random fluctuation in one marginal case and fluent vibrational change in another.

To simulate the process of changes in solar radiation, each of these components has to be examined individually. As it was already mentioned, the first three components contain the oscillating process. Therefore it is of interest to determine the number of turning points, i.e., peak number, trough and the distribution of the distance between them.

3. Preliminary Analysis

Let's consider a finite number of values n of solar radiation q_1, q_2, ... q_n and let

$$X_i = \begin{cases} 1, & if \;\; q_i < q_{i+1} > q_{i+2} \;\; or \;\; q_i > q_{i+1} < q_{i+2}, \\ 0, & otherwise. \end{cases}$$

Then the number of turning points is equal to

$$p = \sum_{i=1}^{n-2} X_i.$$

The expectancy of any of number of turning points is equal to $\frac{2}{3}(n-2)$. Research show that the daily number of turning points for the time series of solar radiation is 2, i.e. there is one maximum and one minimum for each one-day series of measurements, which corresponds to $n \geq 5$. This means that for a study of a time series we need at least 6 values of solar radiation daily.

As it was already mentioned, the first three components can be considered as deterministic components. Various methods are used for simulation of deterministic components of insolation, including regression and harmonic analysis. Experimental or actinometrical and meteorological observations are used. In some cases of international practice, computer databases are widely used to build a model, such as US NASA SSE and Swiss

METEONORM. The comparison of the values of intensity of solar radiation in NASA SSE database with those of Russian meteorological stations revealed the possibility of using NASA SSE program in Russia. In [4], the same comparison was made on the basis of METEONORM. It was concluded in result of comparing the values of the METEONORM and NASA SSE databases of solar radiation that they are useful for designing solar power plants in the absence of reliable values of ground-based observation stations, as well as for feasibility calculations for their construction.

In current studies on the modeling of various processes, including the modeling of solar radiation, approaches to extrapolation of the trend of the time series by analytic function, building of multivariate regression or auto regression models and its extended version - method of group accounting of arguments, are common. In addition, techniques based on wavelet transform and techniques of time series prediction based on neural network technology are popular.

4. Regression models

Tabular data of solar radiation by month for the city of Baku for 2012 are used for building regression models of solar radiation. Research was conducted on Microsoft Office EXCEL. Polynomial functions of various degrees, ranging from third degree, are accepted as a regression model (Table 1).

(Table 1)

Model	R^2	Relative errors
$-4.1586t^2 + 52.24t - 10.386$	$R^2 = 0.9119$	max=50%, min=0.9% average=16.2%
$0.0411t^3 - 4.9591t^2 + 56.571t - 15.99$	$R^2 = 0.9127$	max=45, min=0.1% average=15.9%
$0.1225t^4 - 3.1446t^3 + 22.434t^2 - 30.351t + 60.465$	$R^2 = 0.9674$	max=17.5%, min=0.4% average=6.5%
$0.0075t^5 - 0.1207t^4 - 0.267t^3 + 7.4326t^2 + 2.4175t + 38.409$	$R^2 = 0.9689$	max=16.2%, min=0.2% average=6.9%
$-0.0008t^6 + 0.0388t^5 - 0.5933t^4 + 3.1937t^3 - 5.2008t^2 + 23.364t + 26.788$	$R^2 = 0.9691$	max=17%, min=0.9% average=7.4%
Fourier model (see Table 2)	$R^2 = 0.9245$	max=6.2%, min=1.3% average=2.8%
Exponential model $\text{Exp}(-[t\text{-}6)/\sigma_{max}]^2$ / $\sigma_0\sigma_{max}=21.7$ $\sigma_0=0.006$	$R^2 = 0.9134$	max=22.8%, min=0.2% average=9.73%

As shown in Table 1, you can take a polynomial of order 4 as the polynomial regression model. A similar study can be conducted for daily changes in solar radiation.

5. Simulation Using a Gaussian Function

As stated above, number of turning points is 2 for the time series of solar radiation, i.e. there is one maximum and one minimum for each one-day series of measurements and it is the same for monthly and daily average data. Therefore, we considered the possibility of modeling time series of solar radiation with the use of a Gaussian function:

$$Q\left(t\right)=\frac{1}{\sigma_0}\exp^{-\left[\frac{t-t_{cp}}{\sigma_{\max}}\right]^2}$$

We consider t_{cp} as a median (average) of the timeline, and the parameters σ_0 and σ_{max} are determined using the method of least squares.

Suppose we are given the values of the time series of solar radiation for $t_i Q\left(t_i\right))$ $i = 1, 2, \ldots n$, where n is the number of measurements. We use the following notation:

$$\text{in}\ \left(Q\left(t_i\right)\right)\left(t_i-t_{cp}\right)^2\ \text{in}\ \left(Q\left(t_i\right)\right),$$

$$S_2=\sum_{i=1}^{n}\left(t_i-t_{cp}\right)^2,$$

$$S_4=\sum_{i=1}^{n}\left(t_i-t_{cp}\right)^4,$$

$$D=\left(S_2\right)^2-nS_4.$$

Parameters σ_0 and σ_{max} are determined by the following formulas:

$$\sigma_0=\exp^{-\frac{S_2S_{1_2}-S_4S_{1_0}}{}}$$

$$\sigma_{\max}=\sqrt{\frac{D}{nS_{1_2}-S_2S_{1_0}}}$$

As shown in Table 1, the model of solar radiation built with the use of a Gaussian function gives satisfactory results, despite the high value of maximum relative error. Studies show that a maximum relative error occurs at the point of internal rotation (Fig. 2, at $t = 3$ in our case), which implies the impossibility of taking into account the point of internal rotation using a Gaussian function.

6. Simulation Using Harmonic Fourier Series

Because of the quasi-periodic nature of the time series of solar radiation, regression method of accounting for the trend and the description of periodic components using Fourier series are combined sometimes.

As is known, the Fourier series has the following form:

$$s(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos(k\omega_1 t) + b_k \sin(k\omega_1 t))$$

Here $\omega_1 = 2\pi/T$ is a circular frequency corresponding to the period T of signal repetition. Multiple frequencies $k\omega_1$ are called harmonics, numbered in accordance with the index k. Frequency $\omega_k = k\omega_1$ is called the k^{th} harmonics of signal. The coefficients of $a_k u$ are calculated by the formulas:

$$a_k = \frac{2}{T}\int_{-T/2}^{T/2} s(t)\cos(k\omega_1 t)dt.$$

The constant a_0 is calculated using the general formula for a_k. u is the average value of the signal on the period.

$$\frac{a_0}{2} = \frac{1}{T}\int_{-T/2}^{T/2} s(t)dt$$

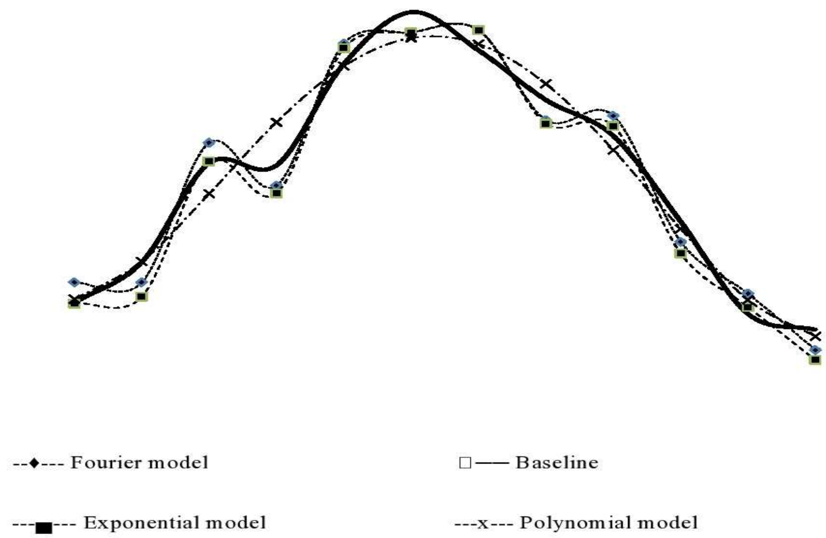
$$b_k = \frac{2}{T}\int_{-T/2}^{T/2} s(t)\sin(k\omega_1 t)dt$$

The following values of the parameter are used when building the Fourier model of solar radiation by month for 2012 in Baku:

$T = 12; k = 6; \omega_1 = 0.5236.$

Table 2. Coefficients of the Fourier model of the solar radiation by month for 2012 in Baku

k	1	2	3	4	5	6
a_k	-8.87	3.175426	1.83	1.42E-14	4.70	0
b_k	-63.77	-2.5	-7.33	7.33	0.60	-2.25
R$_m$	64.38	4.04	7.56	7.33	4.74	2.25
φ_m	-0.14	0.90	0.24	-1.94E-15	-1.44	-3.14



Solar radiation in MgJ/h

Figure 2. Models of monthly solar radiation in 2012, Baku

As shown in Table 2, some of Fourier coefficients are equal to zero, which testifies to the no true association on the basis of mathematical statistics between the Fourier coefficients and the influencing factors. Reducing the frequency (increasing the number of harmonics oscillations) increases the accuracy of the model, but it does not fully take into account all the features of process (such as various gaps, steps and peaks, Fig. 2).

7. Model Analysis

Along with above-stated drawbacks, some more difficulties arise in applying these models in practice, such as oscillation measurements, which occur due to the noise generated by the environment or the equipment. Therefore, we performed an analysis of the time series of solar radiation with the use of smoothing methods before determining the deterministic components in the model. For many years, the method of moving averages has been one of the most common methods of following the trend. Simplicity of building and interpretation largely contributed to this. Moving averages is a method that smooth's series average by its current value and its immediate neighbors in the past and in the future.

Simple moving average or arithmetic moving average is numerically equal to the arithmetic mean of the values of the initial function for a specified period and is calculated by the formula:

$$Q_t = \frac{1}{n}\sum_{i=0}^{n-1} q_{t-i} = \frac{q_t + q_{t-1} + ... + q_{t-i} + ... + q_{t-n+2} + q_{t-n+1}}{n},$$

where Q_t is the value of simple moving average at point t, n is the number of values of the original function to calculate the moving average (smoothing interval; the wider the smoothing interval, the smoother the graph of the function), q_{t-i} is the initial value of the function at point $t-i$.

It is easy to establish some properties of these moving averages:

- Sum of the weights is equal to one. This had to be this way because, as we apply the procedure of weighing to a series whose terms are equal to the same constant, the average must be equal to the same constant;

- Scales are symmetric with regard to the median value;

- Because of the symmetry, of the weight, the trend values do not depend on the direction of timing.

Findings

1. Characteristic features of the process of solar radiation are studied and investigated. It is found that due to multifactorial process of solar radiation, the modeling must be based on real statistical data.

2. The series of solar radiation functions are considered and their dependence on four components: trend, more or less regular fluctuations relative to trend, seasonal fluctuations, and non-systematic random effect is established. It is shown that the non-systematic random effect could be defined as a difference between the statistics and the amount of trend and seasonal fluctuations.

3. Some research is conducted in order to determine the type of trend based on polynomial regression analysis and a Gaussian function, and some analysis of seasonal variations based on Fourier series. A series of Fourier harmonics in E-6 could be used as a model for seasonal fluctuations.

1. shoaling process in the Eastern Aral Sea is more dynamic than in the North and Western Aral Seas;

2. effective albedo of the shoaled areas varies between $0.5 \leq A \leq 0.85$.

References

[1] M.D. Rabinovich, “Comparison of different methods of representing climatological information when calculating the performance of solar power systems” Heliotekhnika, No. 3, 1986.

[2] V.S. Simankov, A.V. Shopin, P. Y. Buchatsky “Modeling insolation in managing photo, wind energy systems” Proceedings of FORA, No. 5, 2000, pp. 67-71.

[3] M.M. Valov, B. Gorshkov, E.I. Nekrasov “On the accuracy of determining the intensity of solar radiation in calculating solar power plants” Heliotekhnika, No.6, 1982.

[4] V.S. Simankov, P. Y. Buchatsky, A.V. Chopin “Methodology of modeling physical processes in the energy complexes with non-traditional sources of energy and optimization of their parameters”. Proceedings of the Fora, No. 3, 1998, pp. 18-26.

[5] A.V. Volgin, A.V. Yurchenko, A.V. Kozlov, M.V. Kitaeva, Polzunovskii BULLETIN No. 2, 2010, pp. 150-154.

[6] Perdomo R.,Banguero E.;Gordillo G. Statistical modeling for global solar radiation forecasting in Bogota. Photovoltaic Specialists Conference (PVSC), 2010, 35th IEEE 20-25 June 2010, pp. 2374-2379.

[7] Wu Ji, CK Chan, JW Loh, FH Choo, LH Chen. Solar Radiation Prediction Using Statistical Approaches. Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on pp.1-5.

[8] Sh. Gorjian, T. Tavakkoli Hashjin, B. Ghobadian Estimation of Mean Monthly and Hourly Global Solar Radiation on Surfaces Tracking The Sun. Second Iranian Conference on Renewable Energy and Distributed Generation, 2012, pp 170-176.

[9] Samsul Ariffin Abdul Karim, Balbir Singh Mahinder Singh,Radzuan Razali and Noorhana Yahya Data Compression Technique for Modeling of Global Solar Radiation. 2011 IEEE International Conference on Control System, Computing and Engineering, 2011, pp. 348-352.

[10] A.Q. Jakhrani,A. K. Othman, A. R. H. Rigit, S. R. Samo A simple method for the estimation of global solar radiation from sunshine hours and other meteorological parameters. IEEE ICSET 2010 6-9 Dec 2010, Kandy, Sri Lanka.

[11] M. Chegaar, and A. Chibani, ”A simple method for computing global solar radiation”, Rev. Energ. Ren., Chemss, 2000, pp. 111-115.

[12] K. Bakirci, ”Correlations for estimation of daily global solar radiation with hours of bright sunshine in Turkey”, Energy, vol. 34, 2009, pp. 485-501.

[13] T. Muneer, S. Younes, and S. Munawwar, ”Discourses on solar radiation modeling”, Renewable and Sustainable Energy Reviews, vol. 11, 2007, pp. 551-602.

[14] D.H.W. Li, T.N.T Lam, and V.W.C. Chu, ”Relationship between the total solar radiation on tilted surfaces and the sunshine hours in Hong Kong”, Solar Energy, vol. 82, 12, 2008, pp. 1220-1228.

[15] J. M. Chang, J.S. Leu, M.C. Shen, and B.J. Huang, “A proposed modified efficiency for thermosyphon solar heating systems”, Solar Energy, vol. 76, 6, 2004, pp. 693-701.

[16] R.E. Childs, Donald G.S. Chuah, S.L. Lee, K.C. Tan. Analysis of solar radiation data using cubic splines. Solar Energy, Volume 32, Issue 5 , 1984 , pp. 643-653

[17] Genc, A., Kinaci, placeI., Oturanc, G., Kurnaz, A., Bilir, S. and Ozbalta, N. Statistical Analysis of Solar Radiation Data Using Cubic Spline Functions. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects. 24:12, 2002, pp. 1131-1138.

[18] Karim, S.A.A and Piah, A.R.M . Rational Generalized Ball Functions for Convex Interpolating Curves. JQMA 5:1, 2009, pp. 65-74.

[19] W.E. Alnaser, New model to estimate the solar global irradiation using astronomical and meteorological parameters, Renew Energy, 3, 1993, pp. 175–177.

[20] M. Koussa, A. Malek, M, Haddadi. Statistical comparison of monthly mean hourly and daily diffuse and global solar irradiation models and a Simulink program development for various Algerian climates, Energy Conversion and Management, 2009, 50,pp. 1227-1235.

[21] Official website of the World Meteorological Organization http://wrdc.mgo.rssi.ru.

[22] Official website of NASA Surface meteorology and Solar Energy/http://eosweb.larc.nasa.gov/sse

[23] Official website of WRDC http://wrdc.mgo.rssi.ru

[24] Official website of WRDCSOLA RGIS http://solargis.info/imaps/

Farhad F. Aliyev
International Ecoenergy Academy,5. M. Arif str., Az1073, Baku
E-mail: faliyev2010@gmail.com

Hadiyya Kh. Khalilova
Institute of Physics of ANAS, H.Cavid av.33, Az1143, Baku
E-mail: kahlilova@rambler.ru

Nadir B. Agayev
National Aviation Academy of Azerbaijan

Received 10 May 2013
Accepted 27 June 2014