Bulletin of Environment, Pharmacology and Life Sciences
Bull. Env. Pharmacol. Life Sci., Vol 2 (6) May 2013: 54- 63

Online ISSN 2277-1808
Journal's URL:http://www.bepls.com

ORIGINAL ARTICLE

Predicting Global Solar Radiation using Genetic Algorithm (GA)

Farzad Fathian
Department of Mechanical Engineering, Ashkezar Branch, Islamic Azad University, Yazd, Iran
*Corresponding author. Tel.: +98 912 2301158; fax: +98 352 3625953

ABSTRACT
Many studies are performed by researchers to predict global solar radiation (GSR) but the Genetic Algorithm (GA) has never been used for such a study. In this study, GA technique is applied to estimate monthly average daily global solar radiation (GSR) on horizontal surface of central arid deserts of Iran. To achieve this, 10 sunshine-based models were developed and the empirical coefficients for all models were separately determined using GA technique. Eventually, the obtained results of GA technique were compared with the obtained results of statistical regression techniques (SRTs). The results indicate that obtained empirical coefficients for all 10 models based on GA have more accuracy than SRTs.
Keywords: *Genetic Algorithm (GA); Statistical Regression techniques (SRTs); Global Solar Radiation (GSR); Sunshine Hours; Modeling.*

Received 01.04.2013 Accepted 02.05.2013

INTRODUCTION

Global solar radiation (GSR) measurements are the most important parameters for the solar energy applications [1]. For low-priced and effective development and utilization of solar energy, a comprehensive knowledge about the accessibility and variability of solar radiation intensity in time and special domain is of great importance [2]. Many studies are performed by researchers to predict solar radiation using available meteorological and geographical parameters [1- 13].
The important key to develop a proper solar radiation model at any location is the availability of meteorological parameters. The most commonly parameter to predict GSR is sunshine duration mainly due to the fact that sunshine duration can be easily and reliably measured [1].
All models which are presented in the literature have used different statistical regression techniques (SRTs) to find empirical coefficients of models or used other techniques in order to predict solar radiation. Recently, Behrang et al. have applied Particle Swarm Optimization technique to more accurately determine empirical coefficients in solar radiation conventional models for 17 cities in Iran. They also showed that the obtained models for given latitude can be generalized to estimate solar radiation in cities at similar latitude [13].
In this study, GA technique, as a well-known intelligent optimization technique, is applied to estimate monthly average daily GSR on horizontal surface for central arid deserts of Iran.

GENETIC ALGORITHM

GAs encode candidate solutions as binary strings. Each string (chromosome) is built by chaining a number of sub-strings, each sub-string representing one of the candidate solution's features. Biological genes are in this case equivalent to the substrings encoding the parameters, while each binary digit can be related to the nucleotides composing the DNA. In most of the cases, one individual is fully described by a single bit-string, thus leading to the identification of the genotype with one single chromosome. Several other encoding procedures have been explored leading to a debate on the most appropriate choice. Holland [14] showed that binary coding allows the maximum number of schemata to be processed per individual. On the other hand, the mapping to binary coding introduces Hamming cliffs onto the search surface. Moreover, non-binary representations may be more natural for some problem domains and may reduce the computational burden of the search. The canonical binary-coded GA as described here is now rarely used for continuous function

optimization as it has been shown that solutions are too easily disrupted (the Hamming cliff issue). Therefore researchers tend to use less disruptive coding such as Gray coding [15].
Similarly to the other Evolutionary Algorithms (EAs), canonical GAs use generational replacement. Popular alternatives are elitism and steady-state replacement [16]. In the first case, the best solution(s) are directly copied into the new population while in the second case only a fraction of the population is replaced at each generation. Both variants aim to improve the preservation of good genetic material at the expense of a reduced search space exploration. A comparison between the behavior of generational and steady-state replacement is given in [17]. Individuals are selected for reproduction with a probability depending on their fitness. Canonical GAs allocate the mating probability of each individual proportionally to its fitness (proportional selection) and draw the parents set (mating pool) through the roulette wheel selection procedure [18]. Other popular selection schemes are fitness ranking [19] and tournament selection [20]. For a comparison of selection procedure, the reader is referred to [20].
Crossover is the main search operator in GAs, creating offsprings by randomly mixing sections of the parental genome. The number of sections exchanged varies widely with the GA implementation. The most common crossover procedures are one-point crossover, two-point crossover and uniform crossover [16]. In canonical GAs, a crossover probability is set for each couple. Couples not selected for recombination will generate two offsprings identical to the parents.
A small fraction of the offsprings are randomly selected to undergo genetic mutation. The mutation operator randomly picks a location from a bit-string and flips its contents. The importance of this operator in GAs is however secondary, and to the main aim of mutation is the preservation of the genetic diversity of the population.
GAs require the tuning of some parameters such as the mutation rate, crossover rate and replacement rate in the case of steady-state replacement. This task is often not trivial as the chosen values may strongly influence the search process [21]. Moreover, the optimal value for the GA parameters may vary according to the evolution of the search process. For all these reasons, several adaptive schemes have been investigated. A survey of adaptation in GAs is given in [19- 21] proposed an off-line tuning approach giving an optimal mutation rate schedule. Problem-specific operators are sometimes employed in addition to the canonical ones. The introduction of such operators results an increase in the search power of the algorithm but a loss of general applicability. Population size, selection function, reproduction, crossover, mutation and generation are considered as important factors in GA.

PROBLEM DEFINITION

In this study, GA technique is applied to estimate monthly average daily GSR on horizontal surface for central arid deserts of Iran based on the measured data, provided by Iranian Meteorological Office (IRIMO). For this purpose, 10 sunshine–based models are developed. Table 1 presents information for the considered cities in this study. As it can be seen in this table, the valid months included in the study are large enough to obtain the significant results. The valid months are described in Appendix A. The 10 models developed in this study are listed in Table 2.

Table1.Information for considered cities.

Location	Long. (°E)	Lat. (°N)	Elev. (m)	Period of installation data series	Validation data series		
					Period	Number of Valid months [a]	Valid months / Total months (%)
Birjand	59.20	32.87	1491.0	1982-2001	2004-2005	19	79%
Esfahan	51.67	32.62	1550.4	1985-2001	2002-2005	45	94%
Kerman	56.97	30.25	1753.8	1984-2001	2002-2005	24	50%
Khoor-Biabanak	55.08	33.78	845.0	1988-2001	2002-2005	47	98%
Tabass	56.92	33.60	711.0	1986-2000	2001-2003	24	67%
Tehran	51.32	35.68	1190.8	1974-1987	1998-1999	23	96%

[a] Valid months has been described in Appendix A.

Table 2. The models developed in this study.

Model no.	Equation form	Source
1	$H/H_0 = a^{(S_0/S)}$	El-Metwally (1990)
2	$H/H_0 = a + b(S/S_0)$	Angstrom (1924); Prescott (1940)
3	$H/H_0 = a + b(S/S_0) + c(S/S_0)^2$	Akinoglu and Ecevit (1990)
4	$H/H_0 = a + b\exp(S/S_0)$	Almorox and Hontoria (1967)
5	$H/H_0 = a + b(S/S_0) + c\exp(S/S_0)$	Bakirci (2009)
6	$H/H_0 = a + b(S/S_0)^c$	Behrang et al (2011)
7	$H/H_0 = a + bCos(c(S/S_0)) + dSin(c(S/S_0)) + eCos(2c(S/S_0)) + fSin(2c(S/S_0))$	Behrang et al (2011)
8	$H/H_0 = aSin(b(S/S_0) + c) + dSin(e(S/S_0) + f) + gSin(h(S/S_0) + i)$	Behrang et al (2011)
9	$H/H_0 = a + bSin(c(S/S_0) + d) + eSin(f(S/S_0) + g) + hSin(i(S/S_0) + j)$	Behrang et al (2011)
10	$H/H_0 = a + bCos(c(S/S_0)) + dSin(c(S/S_0)) + eCos(2c(S/S_0)) + fSin(2c(S/S_0)) + gCos(3c(S/S_0)) + hSin(3c(S/S_0))$	Behrang et al (2011)

RESULTS AND DISCUSSION

Determining empirical coefficients using GA

In this section, optimal empirical coefficients corresponding to the actual data are found using a GA algorithm implemented in MATLAB 2007 (Math Works, Natick, MA). The following steps are carried out to find optimal values of empirical coefficients of each model:

STEP (a): Collected data is divided into two parts:

1) Installation data series.
2) Validation data series.

Table 1 shows the details about the periods of installation and validation data series. Figure 1 shows the measured values of fraction of possible monthly average daily global radiation (H/H_0) and fraction of possible monthly average daily sunshine duration (S/S_0), in both installation and validation periods, for Esfahan and Khoor-Biabanak.

STEP (b): Installation data series are used in GA algorithm to find candidates of the best empirical coefficients. The criteria to select optimal coefficients is the minimum fitness function defined by

$$F(x) = \sum_{k=1}^{n} \left((H/H_0)_{actual} - (H/H_0)_{predicted} \right)^2 \quad (1)$$

Where H and H_0 are the monthly average daily solar radiation and the monthly average daily extra-terrestrial solar radiation, respectively, and n is the number of observations. $(H/H_0)_{actual}$ and $(H/H_0)_{predicted}$ are the actual and predicted fraction of possible monthly average daily GSR, respectively. Calculation of the extra-terrestrial solar radiation (H_0) and maximum possible sunshine duration (S_0) are discussed in Appendix B.

STEP (c): Outputs of the algorithm (candidates of the best empirical coefficients) after each run are validated using a different series of data called herein the validation data series.

The convergence of the objective function and sensitivity analysis are examined for different combinations of important factors of GA (Population size, selection function, reproduction, crossover, mutation and generation).

It was found the fitness function was least sensitive to the population size.

The performance of GA is satisfactory using the following user-specified factors for all models:

Population :(Population size: 200)/Selection :(Selection function: Stochastic uniform)/ Reproduction :(Elite count: 2.0, Crossover fractions: 0.8)/ Crossover :(Crossover function:

Scattered)/Mutation :(Mutation function: Gaussian, Scale: 1.0, Shrink: 1.0)/ Stopping criteria :(Generation: 1000).

The performance of the models is accomplished by using well known statistical indices including: mean absolute percentage error (MAPE) and absolute fraction of variance (R^2). MAPE and R^2 are defined in Appendix C.
Figure 2 shows the overall description about determining empirical coefficients for each model using proposed algorithm.

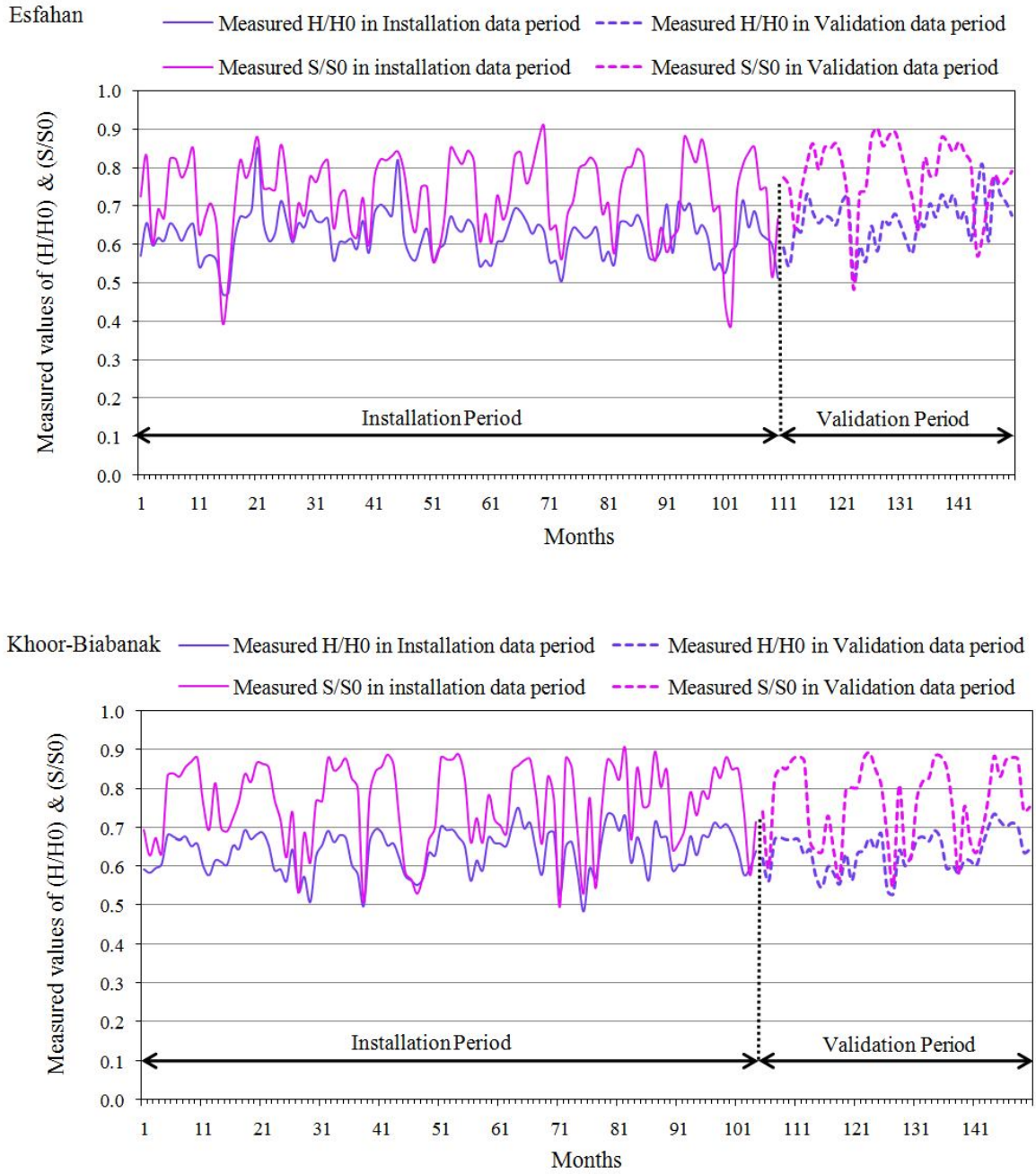

Figure 1. The measured values of H/H_0 and S/S_0, in both installation and validation periods for Esfahan and Khoor-Biabanak.

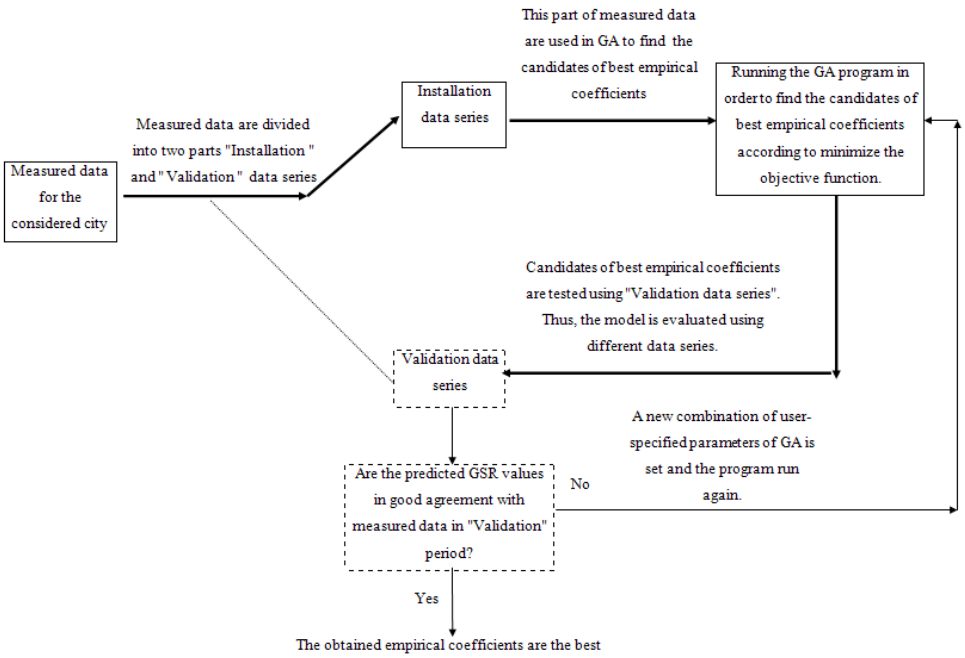

Figure 2. Description about determining empirical coefficients for each model using GA.

Table 3 shows the quantities of obtained empirical coefficients using GA for considered models in this study.

Table 3. The quantities of obtained empirical coefficients using GA for considered models in this study.

model no.	a	b	c	d	e	f	g	h	i	j
1	0.7457	-	-	-	-	-	-	-	-	-
2	0.2369	0.5091	-	-	-	-	-	-	-	-
3	0.4250	0.2250	0.0992	-	-	-	-	-	-	-
4	0.3161	0.1567	-	-	-	-	-	-	-	-
5	0.1809	-0.2718	0.3071	-	-	-	-	-	-	-
6	0.4220	0.2850	0.8821	-	-	-	-	-	-	-
7	0.7667	-1.1225	0.2868	0.4780	0.5019	0.8003	-	-	-	-
8	0.4109	0.3463	0.2464	0.4674	0.8866	0.1668	0.2603	-0.0915	0.4378	-
9	0.0348	1.2789	-2.0311	0.7115	0.1512	0.0682	0.5318	0.7479	0.2910	0.1435
10	-0.1028	0.2859	0.6442	-0.3141	0.3918	0.3141	0.4901	0.2232	-	-

MAPE and R^2 for developed GA models are shown in Table 4. All models gave the absolute fractions of variance (R^2) more than 98% (on both installation and validation series). So, the predicted GSR values from all models are in good agreements with the measured data.

The model giving the best results is the Model 9 which has an MAPE and R^2 about 8.28% and 98.99%, respectively.

Table 5 shows the performance of the best model (i.e. Model 9) for all selected cities in this study.

Table 4. MAPE and R^2 for the developed models using GA.

Model no.	Installation	Validation	
	R^2	R^2	MAPE
1	0.985	0.9888	0.0955
2	0.9813	0.9864	0.0880
3	0.9850	0.9902	0.0853
4	0.9847	0.9900	0.0871
5	0.9830	0.9892	0.0858
6	0.9832	0.9896	0.0897
7	0.9830	0.9893	0.0858
8	0.9839	0.9896	0.0831
9	**0.9844**	**0.9899**	**0.0828**
10	0.9832	0.9875	0.0844

Table 5. Performance of the best model (i.e. Model 9) for Birjand, Esfahan, Kerman, Khoor-Biabanak, Tabass, and Tehran.

Location	Validation	
	R^2	MAPE
Birjand	0.9855	0.0899
Esfahan	0.9830	0.0991
Kerman	0.9865	0.0831
Khoor-Biabanak	0.9971	0.0398
Tabass	0.8161	0.2784
Tehran	0.9859	0.1102

As it can be seen in this table, there is a good agreement between measured GSR and predicted GSR according to the best model (i.e. Model 9) for all considered cities in this study.
The measured data and predicted values of monthly average daily GSR for the considered cities based on the best model (Model 9) are compared in Figures 3 to 5.

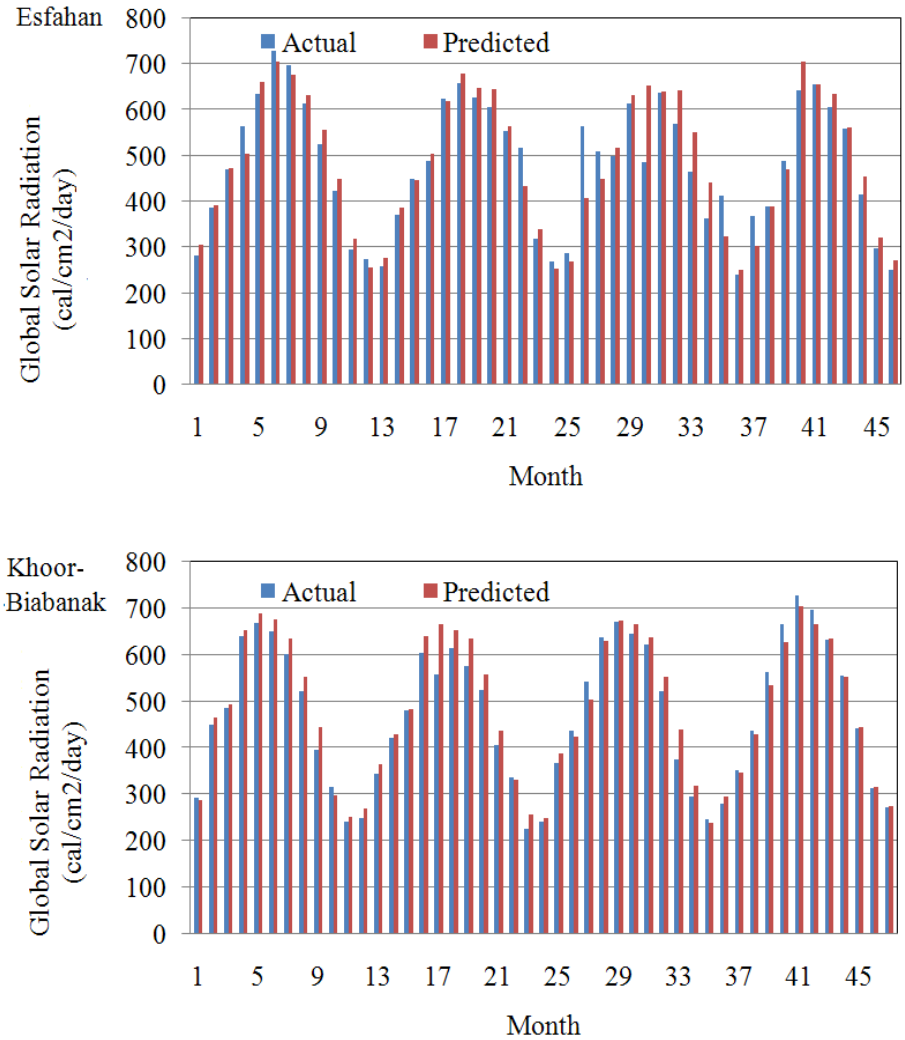

Figure 3. Comparison between measured data and predicted values of monthly average daily GSR based on the best model (Model 9) for Esfahan and Khoor-Biabanak.

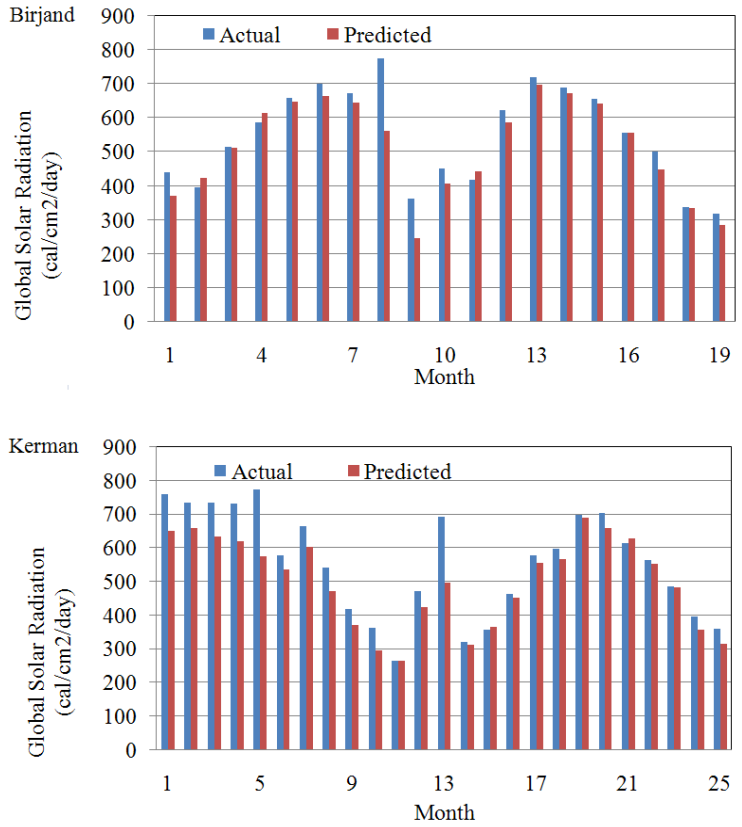

Figure 4. Comparison between measured data and predicted values of monthly average daily GSR based on the best model (Model 9) for Birjand and Kerman.

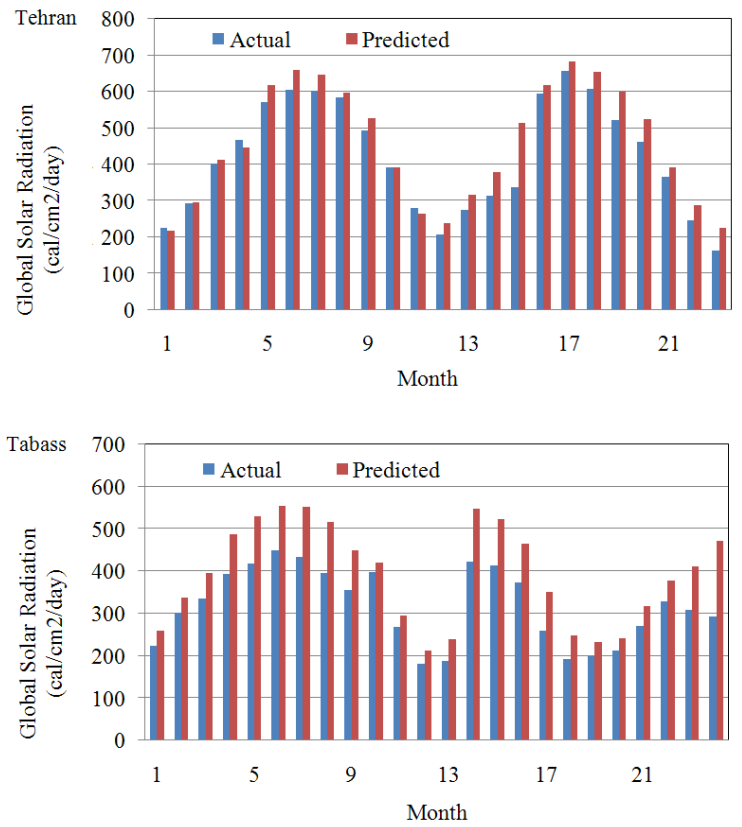

Figure 5. Comparison between measured data and predicted values of monthly average daily GSR based on the best model (Model 9) for Tehran and Tabass.

Comparison between GA and Statistical Regression Techniques (SRTs)

In this section same data are used to make a fair comparison between the performance of GA and SRTs on GSR modeling.

The empirical coefficients for all models are separately calculated using the SRT (the least- squared method) and shown in Table 6.

Table 6. The quantities of obtained empirical coefficients using the SRT for considered models in this study.

model number	a	b	c	d	e	f	g	h	i	j
1	0.5500	-	-	-	-	-	-	-	-	-
2	0.3778	0.3527	-	-	-	-	-	-	-	-
3	0.3538	0.425	-0.0524	-	-	-	-	-	-	-
4	0.27	0.1747	-	-	-	-	-	-	-	-
5	0.4218	0.4936	-0.0702	-	-	-	-	-	-	-
6	253.1	-252.4	-0.0008	-	-	-	-	-	-	-
7	6.609	0.4777	0.0548	0.434	-1.132	-2.635	-	-	-	-
8	0.487	-0.178	0.1289	0.6888	-0.009	0.6013	1.812	0.6317	0.487	-
9	0.5107	0.7419	-0.951	0.836	0.2666	0.5785	1.0825	0.8305	0.623	0.378
10	31.01	-34.67	-0.1182	-118.20	4.2690	59.03	-0.1044	-0.0168	-	-

MAPE and R^2 for developed models using the SRT are shown in Table 7. MAPE error improvements by GA for all considered models in this study (in comparison with the SRT) is given in Table 8.

Table 7. MAPE and R^2 for the developed models using the SRT.

Model no.	Installation	Validation	
	R^2	R^2	MAPE
1	0.9724	0.9617	0.1826
2	0.9889	0.9748	0.1265
3	0.9890	0.9747	0.1267
4	0.9890	0.9748	0.1263
5	0.9890	0.9747	0.1267
6	0.9888	0.9748	0.1271
7	0.9867	0.9734	0.1307
8	0.9869	0.9829	0.1311
9	0.9884	0.9844	0.1298
10	0.9870	0.9715	0.1357

Table 8. Error improvements by GA for all considered models in this study (in comparison with the SRT).

Model no.	MAPE reduction (%)
1	8.71
2	3.85
3	4.14
4	3.92
5	4.09
6	3.74
7	4.49
8	4.8
9	4.7
10	5.13
Average	**4.76**

The maximum and minimum MAPE error improvements by GA (in comparison with the SRT) were 8.71% and 3.74% for Model 1 and Model 6, respectively. An average MAPE error reduction of 4.67% was achieved when obtained empirical coefficients based on GA were used instead of using ones obtained from the SRT.

The results indicate that obtained empirical coefficients based on GA have better performance than the SRT for all 10 models.

CONCLUSION

This study showed application of GA to estimate monthly average daily GSR on horizontal surface over central arid deserts of Iran. For this purpose, 10 different sunshine-based models were considered. The quantities of empirical coefficients for all models were determined using GA technique. The models were validated using validation data series. Model 9 gave the best results with an MAPE and R^2 about 8.28% and 98.99%, respectively.

Finally, a fair comparison was made between the obtained results from GA and the SRT (i.e. the least-squared method). The results indicated that obtained empirical coefficients for all 10 models based on GA have more acceptable performance than the SRT.

Future work is focused on comparing the presented method with other available tools. Predicting of global solar radiation can also be investigated with other intelligent optimization techniques like Bees Algorithm, Gravitational Search Algorithm and etc. The results of the different techniques could be compared with available methods.

ACKNOWLEDGMENTS

The authors are grateful for the support provided for the present work by Ashkzar Branch, Islamic Azad University, Iran.

Appendix A

The limit check (i.e. higher limits of monthly average daily sunshine hours and monthly average daily limits of extra-terrestrial solar radiation of the location) is carried out on the monthly mean daily GSR and monthly mean daily sunshine duration to make sure that the data are homogeneous and the variations of monthly mean daily GSR are caused only by climatic influences and not by other sources of errors (e.g. systematic errors cased by instruments, calibration, data transferring, etc) (Behrang et al, 2011).

Number of valid data for each city is calculated from following equation:

Number of valid data =Number of total available data in the period - (Number of data in the period out of limit check).

Appendix B

Calculation of the extra-terrestrial solar radiation (H_0) and maximum possible sunshine duration (S_0) are as follow:

$$H_0 = \left(\frac{1}{\pi}\right).I_{sc}.E_0.\left(\cos\phi.\cos\delta.\sin\omega_s + \left(\frac{\pi}{180}\right).\sin\phi.\sin\delta.\omega_s\right)$$

Where $I_{sc}, E_0, \phi, \delta$ and ω_s are solar constant (=1367 W m^{-2}= 11810.4 J cm^{-2} day^{-1}= 2822.4 Cal cm^{-2} day^{-1}), eccentricity correction factor, latitude of site (deg), solar declination (deg) and sunrise hour angle (deg) (Spencer, 1971).

The eccentricity correction factor (E_0) and can be calculated by the expression

$$E_0 = 1.00011 + 0.034221.\cos\xi + 0.00128.\sin\xi + 0.000719.\cos 2\xi + 0.000077.\sin 2\xi$$

Solar declination can be calculated from following equation (Spencer, 1971):

$$\delta = \left(\frac{180}{\pi}\right).(0.006918 - 0.399912.\cos\xi + 0.070257.\sin\xi - 0.006758.\cos 2\xi + 0.000907.\sin 2\xi$$
$$- 0.002697.\cos 3\xi + 0.00148.\sin 3\xi)$$

where $\xi = 2\pi.(m-1)/365$ (radians), and m is the number of the day of the year, starting from first January.

sunrise hour angle (ω_s) can be calculated in degree from the following equation:

$$\omega_s = \cos^{-1}(-\tan\phi.\tan\delta)$$

maximum possible sunshine duration can be calculated in degree from the following equation (Spencer, 1971):

$$S_0 = (2/15).\omega_s$$

Appendix C

MAPE and R^2 are defined as follows:

$$\text{MAPE}=\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\left|(H/H_0)_{i(\text{predicted})}-(H/H_0)_{i(\text{actual})}\right|}{(H/H_0)_{i(\text{actual})}}\right)\times 100$$

$$R^2=\left(1-\left(\frac{\sum_{i=1}^{n}\left|(H/H_0)_{i(\text{predicted})}-(H/H_0)_{i(\text{actual})}\right|^2}{\sum_{i=1}^{n}(H/H_0)_{i(\text{actual})}}\right)\right)\times 100$$

where n is the number of observations.

REFERENCES

1. Bakirci, K. (2009). Correlations for estimation of daily global solar radiation with hours of bright sunshine in Turkey. Energy 34 (4): 485-501.
2. Azadeh, A., Maghsoudi, A., and Sohrabkhani, S. (2009). An integrated artificial neural networks approach for predicting global radiation. Energy Conversion and Management doi: 10.1016/j.enconman.2009.02.019.
3. Akinoglo, B.G., and Ecevit, A. (1990). Constraction of quadric model using modified Angstrom coefficients to estimate global solar radiation. Solar Energy 45: 85-92.
4. Angstrom, A. (1924). Solar and terrestrial radiation. Journal of the Royal Meteorological Society 50: 121–6.
5. Benghanem, M., and Mellit, A. (2010). Radial Basis Function Network-based prediction of global solar radiation data: Application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia. Energy 35 (9): 3751-3762.
6. Coskun, C., Oktay, Z., and Dincer, I. (2011). Estimation of monthly solar radiation distribution for solar energy system analysis. Energy 36: 1319-1323.
7. Katiyar, A.K., and Pandey, C.K. (2010). Simple correlation for estimating the global solar radiation on horizontal surfaces in India. Energy doi:10.1016/j.energy.2010.08.014.
8. Prescott, J.A. (1940). Evaporation from water surface in relation to solar radiation. Transaction of the Royal Society of Australia 46: 114-8.
9. Evseev, E.G., and Kudish, A.I. (2009). The assessment of different models to predict the global solar radiation on a surface tilted to the south. Solar Energy 83: 377–88.
10. Prieto, J.I., Martínez-García, J.C., and García, D. (2009). Correlation between global solar irradiation and air temperature in Asturias, Spain, Solar Energy. 83: 1076-1085.
11. Rehman, S., and Mohandas, M. (2008). Artificial neural network estimation of global solar radiation using air temperature and relative humidity. Energy Policy 63: 571–6.
12. Senkal, O. (2010). Modeling of solar radiation using remote sensing and artificial neural network in Turkey. Energy doi:10.1016/j.energy.2010.09.009.
13. Behrang, M.A., Assareh, E., Ghanbarzadeh, A., and Noghrehabadi, A.R. (2010). The potential of different Artificial Neural Network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Solar Energy 84: 1468–1480.
14. Holland, J.H., 1975. Adaptations in Natural Artificial Systems. University of Michigan Press, Michigan.
15. Michalewicz, Z., (1999). Genetic algorithms + data structures = evolution programs, 3rd rev. and extended ed. Springer-Verlag, Berlin.
16. Davis, L., 1991. Handbook of genetic algorithms. Van Nostrand Reinhold, New York.
17. Syswerda, G., (1991). Reproduction in generational and steady state Genetic Algorithm. In: Ratlines, G. ed. Foundations of Genetic Algorithms. Morgan Kaufmann, Los Altos, CA, p.94-101.
18. Goldberg, D.E., (1989). Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman, Reading.
19. Baker, J.E., (1985). Adaptive selection methods for genetic algorithms. Proceedings of Int. Conf. on Genetic Algorithms and Their Applications, p.101-111.
20. Goldberg, D.E., Deb, K., (1991). A comparison of selection schemes used in genetic algorithms. G.J. ed. Foundations of Genetic Algorithms (FOGA 1), Rawlins, p.69-93.
21. Grefenstette, J.J., (1986). Optimization of Control Parameters for Genetic Algorithms. Systems, Man and Cybernetics, *IEEE Transactions on*, 16(1), 122-128.

How to Cite this Article

Farzad F.(2013).Predicting Global Solar Radiation using Genetic Algorithm (GA). *Bull. Env. Pharmacol. Life Sci.,* Vol 2 (6): 54- 63.