

# GLOBAL SOLAR RADIATION PREDICTION USING GENETIC ALGORITHM

R.MEENAL A.IMMANUEL SELVAKUMAR

Department of Electrical Technology, Karunya University, Coimbatore - 641114, Tamil Nadu, India  
Email id: meenasekar5@gmail.com

**Abstract:** Solar radiation is the essential parameter required for all solar energy applications. In India, this radiation data is not available in all the locations due to the higher cost of measuring equipment and techniques involved. In this work, Genetic algorithm (GA) technique based global solar radiation (GSR) prediction is proposed. To achieve this, Angstrom model based on sunshine duration is employed. Generally sunshine-based models are very accurate. However sunshine records are not available in all the locations. For this purpose, temperature-based regression model is also developed to estimate the GSR. The regression coefficients for the two models are separately determined using the conventional statistical regression technique (SRT) and GA. The monthly average daily GSR on horizontal surface is estimated for three different regions of India namely Bhubaneswar, Nagpur and New Delhi using these techniques. The estimated GSR values are compared with measured GSR data of India Meteorological Department (IMD), Pune. The comparisons are based on statistical parameters namely MBE, RMSE, MAPE and Correlation coefficient R. In addition, the t-statistics is used to determine the statistical significance of the model estimates. The results show that the regression coefficients derived from the proposed GA, performs better than the conventional SRT in estimating the GSR in India. Hence GA technique can be effective in estimating solar radiation wherever the radiation data is not available.

**Key words:** Genetic algorithm [GA], Global solar radiation [GSR], temperature, sunshine duration, Statistical Regression technique [SRT]

## 1. Introduction

The accurate knowledge on solar radiation is very essential for all solar energy applications. The network of solar radiation measuring station is limited throughout the world due to expensive equipment and difficulties in measurement techniques. Hence it is economical to develop radiation models based on available meteorological parameters like sunshine hour, temperature, relative humidity, atmospheric pressure, cloud cover, wind speed and so on. Different methods have been developed by the researchers to predict solar

radiation. The most widely used parameter for estimating global solar radiation is sunshine duration. Angstrom [1] was the first to propose solar radiation model based on sunshine data later Prescott [2] and other researchers developed new regressive models to predict GSR using sunshine data. Gasser Hassan [3] established seventeen new temperature based models to estimate the GSR. Temperature based regression model [4] has good potential for use in predicting solar radiation values for the sites where the sunshine records are not available. In the above, statistical regression methods, higher level of mathematics has to be solved. This statistical approach is both time and mind consuming. Different Intelligent based techniques have also been used by the researchers to estimate GSR with improved prediction accuracy. Hatice Citakoglu [5] and Benghanem [6] applied ANN technique and Behrang et al. [7] has used Particle Swarm optimization technique to determine the empirical coefficients for Iran. Farzad Fathian [8] has applied GA technique to estimate GSR for sunshine models. GA has never been applied for temperature based model. Generally sunshine based model gives very accurate GSR prediction. However the sunshine records are not available in all the meteorological station in India. Therefore, it is difficult to predict GSR at locations where the sunshine hour data do not exist. Consequently, it is necessary to develop model based on temperature which is commonly available in all meteorological station. As a result, in addition to sunshine model, a regression model based on the temperature is also developed in this work and GA optimization technique is applied to estimate the GSR in India.

The objective of this work is to estimate GSR on horizontal surfaces in India with good prediction accuracy. For this purpose, the most widely used Angstrom Prescott sunshine model and temperature based regression models are employed. The conventional SRT (least square method) and GA

optimization technique is applied to obtain the regression coefficients of the model. The monthly average GSR for three locations of India is estimated by both SRT and GA technique. The estimated GSR values are compared with measured GSR data provided by IMD, Pune. This paper is organized as follows; the estimation of global solar radiation by regression models is covered in section 2. The data base used in this work is described in Section 3. The detailed theory about genetic algorithm and methodology is depicted in section 4. Section 5 contains the results and discussion. Finally section 6 covers the conclusion part of this work.

## 2. Regression models

Regression analysis is commonly used for prediction and forecasting of GSR [9-10]. The regression models based on sunshine record and temperature data for estimating monthly average daily GSR is given in Table.1  $H$  is the monthly average global solar radiation on horizontal surface,  $S$  is the monthly average daily bright sunshine hours,  $a$  and  $b$  are regression constants.  $S_0$  is the maximum possible monthly average daily sunshine hours or the day length.  $T_{max}$  and  $T_{min}$  are maximum and minimum temperature. Computation of Extra-terrestrial radiations ( $H_0$ ) and ( $S_0$ ) are defined in Appendix A. The geographical information of the study sites are available in Table.2

Table 1: Regression Models used in this study

Regression equation	Source
$\frac{H}{H_0} = a + b \left( \frac{S}{S_0} \right)$	Angstrom(1924) and Prescott(1940)
$\frac{H}{H_0} = a + b \left( \frac{T_{max}}{T_{min}} \right)$	Angstrom type temperature model

Table 2: Geographical parameters of the study sites

Locations	Latitude (N).	Longitude (E)	Altitude (m)	Climatic Classification	Time Period	Installation Data Set	Validation Data Set
Bhubaneswar	20.27	85.83	62	Tropical Savanna	2003-2008	2003-2006	2007-2008
New Delhi	28.65	77.22	227	Hot semi-arid	2003-2011	2003-2008	2009-2011
Nagpur	21.14	79.08	319	Tropical Savanna	2004-2010	2004-2007	2008-2010

## 3. Data base

India Meteorological Department has provided the monthly average daily global solar radiation in MJ/m<sup>2</sup>/day, bright sunshine hour, monthly average maximum temperature and minimum temperature for the study sites namely Bhubaneswar, Nagpur and New Delhi. The complete dataset is divided into two categories such as Installation data set (data subset 1) and Validation data set (data sub-set 2). It is further processed with the help of sunshine and temperature based regression models. The computer codes for determination of regression coefficients by SRT and GA technique are developed using the MATLAB R2010a computing software. The installation data set is used for determining the regression coefficients and the models are validated using validation dataset.

## 4. Genetic Algorithm

The GA technique is an optimum search technique based on the idea of natural selection in biological systems. The chief unique features of GA are three genetic operators namely selection, crossover and mutation. The GA approach is used for finding the optimum solutions in problems where a fitness function is present. This search technique uses a "fitness" measure to determine which of the candidates in the population survive and reproduce.

Under GA technique rather than considering only the present value of the function to be optimized, they examine a population of values. These values corresponding to various set of variable of the objective function. This feature of GA allows the minimization to consider many global solutions at the same time. In this study, the regression coefficients are separately determined using statistical regression and GA technique. Fig.1 shows the flow chart for determining regression coefficients of sunshine and temperature based Angstrom model using GA approach.

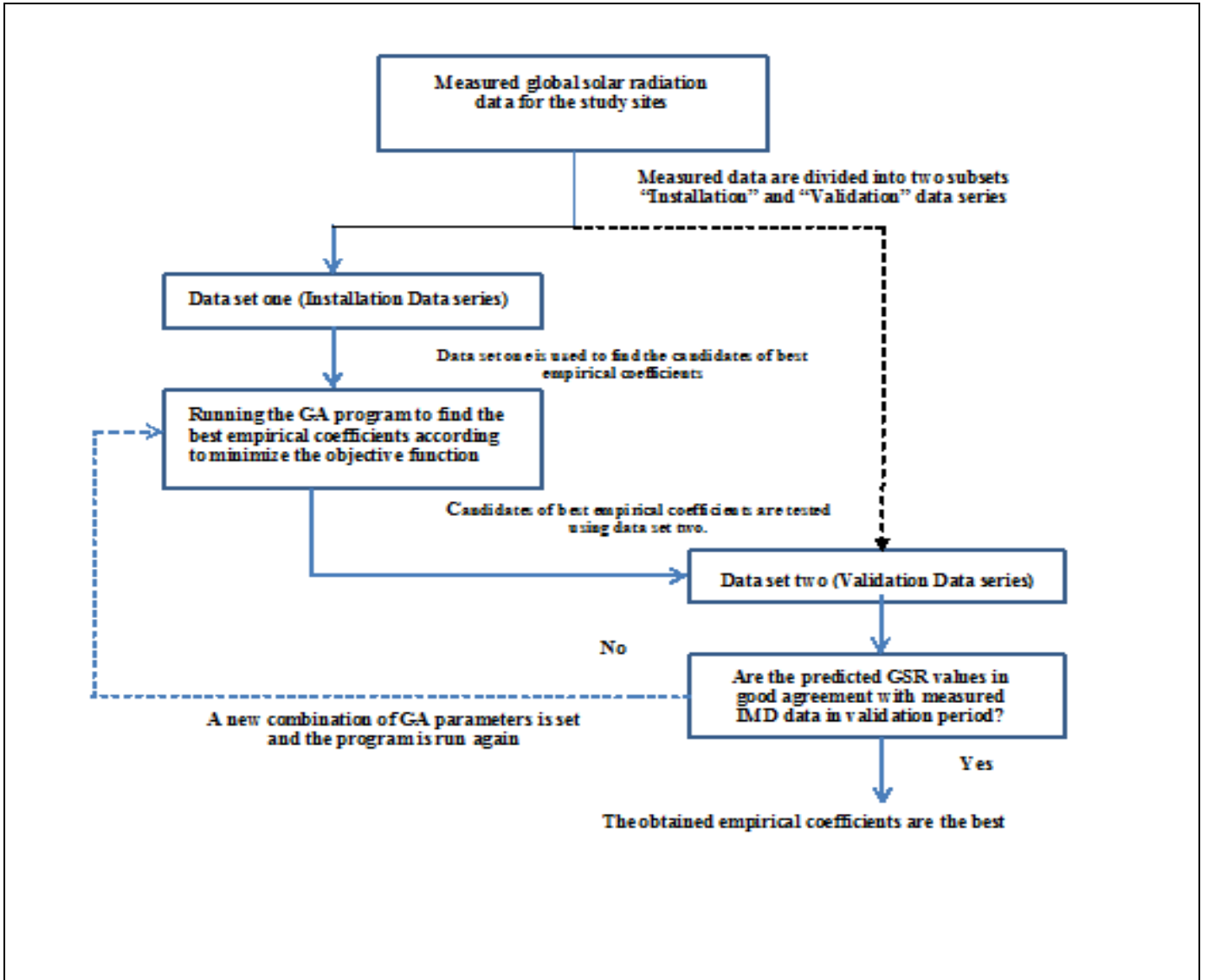


Fig.1. Flow chart for determining Regression coefficients using Genetic Algorithm [GA]

Table 3: Parameters used in GA

Parameter	Values
Population Size	100
Selection Function	Stochastic uniform
Reproduction	Elite count: 2.0, Cross over fractions: 0.8
Crossover Function	Scattered
Stopping criteria	Generation: 100

To find the optimal coefficients for sunshine and temperature-based regression models, the subsequent steps are carried out:

#### Step 1

Divide the data into two subsets:

- 1) Installation data set (data sub-set 1)
- 2) Validation data set (data sub-set 2)

Table 2 contains the information about installation and validation data set.

#### Step 2

Installation data sets are employed to find the candidates of the best regression coefficients. The minimum fitness function is given by

$$F(x) = \sum_{k=1}^n (H_{actual} - H_{predicted})^2 \quad (1)$$

Where  $H_{actual}$  the measured GSR data from IMD, Pune, is  $H_{predicted}$  is the predicted GSR using GA algorithm in  $MJ/m^2/day$  and  $n$  is the number of observations.

**Step 3**

The output of GA algorithm is the candidates of the best regression coefficients. The coefficients are validated after each run using the validation data set. The GA performance is found satisfactory with the user specified factors mentioned in Table 3.

**5. Results and Discussion**

The regression coefficients of the sunshine and temperature based angstrom models are separately determined by SRT and GA technique. The calculated regression coefficients using installation data set are summarized in Table 4. Fig.2. shows the measured bright sunshine hours and Fig.3 shows the monthly average GSR for the study sites. The measured and estimated GSR values for the study sites using the temperature and sunshine based models by SRT and GA technique are summarized in Table 5.

Table 4: Regression Coefficients

Regression Coefficients					
Location	Method	Sunshine based Model		Temperature based Model	
		a	b	a	b
Bhubaneswar	SRT	0.4788	0.2644	-0.6948	0.9784
	GA	0.4560	0.2790	-0.7840	1.0430
Nagpur	SRT	0.4515	0.2515	-0.6624	0.9366
	GA	0.4400	0.2590	-0.7310	0.9800
New Delhi	SRT	0.4886	0.2746	-0.1932	0.6386
	GA	0.4540	0.2880	-0.2420	0.6670

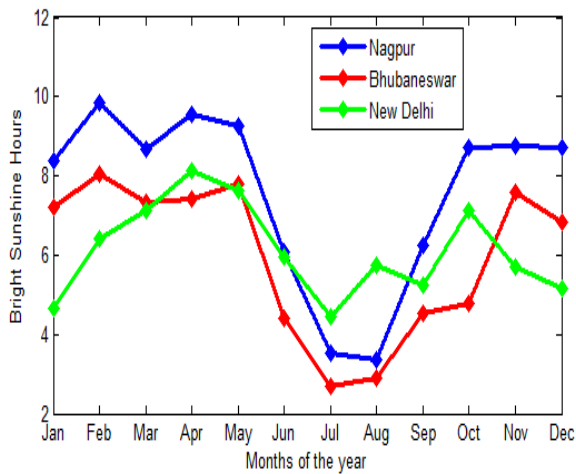


Fig. 2. Measured bright Sunshine Hours for the study sites provided by IMD,Pune

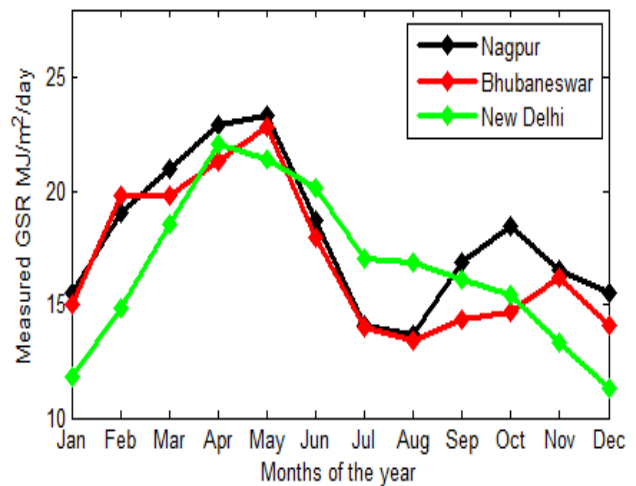


Fig. 3. Measured GSR for the study sites provided by IMD,Pune

Table 5: Estimated GSR in MJ/m<sup>2</sup>/day by sunshine and temperature based Angstrom Model using SRT and GA.

Sunshine based Angstrom Model ( SRT and GA)									
Location	Bhubaneswar			Nagpur			New Delhi		
	Estimated GSR			Estimated GSR			Estimated GSR		
Months	H <sub>Measured</sub>	SRT	GA	H <sub>Measured</sub>	SRT	GA	H <sub>Measured</sub>	SRT	GA
January	14.99	16.32	16.31	14.76	16.23	16.21	10.61	11.25	11.24
February	19.76	17.62	17.66	18.17	18.56	18.53	15.80	16.04	15.86
March	19.80	19.83	19.88	20.99	21.18	21.14	19.14	19.66	19.41
April	21.32	21.17	21.22	23.69	23.39	23.34	22.57	22.06	21.77
May	22.82	21.52	21.58	23.53	22.60	22.58	22.22	22.88	22.60
June	17.98	15.32	15.67	20.55	17.82	17.92	22.16	20.41	20.31
July	13.97	13.72	14.13	13.08	13.88	14.07	17.17	17.18	17.27
August	13.41	14.97	15.27	15.05	12.67	12.86	15.96	17.82	17.75
September	14.29	14.79	15.01	18.44	16.77	16.81	13.86	15.23	15.17
October	14.69	15.78	15.83	17.42	16.79	16.76	15.23	14.93	14.70
November	16.14	13.36	13.43	15.15	14.49	14.48	11.46	10.98	10.90
December	14.05	13.25	13.32	13.85	15.27	15.23	9.85	11.20	11.09

Temperature based Angstrom Model (SRT and GA)									
	Estimated GSR			Estimated GSR			Estimated GSR		
Months	H <sub>Measured</sub>	SRT	GA	H <sub>Measured</sub>	SRT	GA	H <sub>Measured</sub>	SRT	GA
January	14.99	17.37	17.87	14.76	17.95	18.25	10.61	13.49	13.71
February	19.76	18.23	18.59	18.17	19.43	19.68	15.80	16.01	16.20
March	19.80	19.27	19.52	20.99	21.10	21.29	19.14	18.64	18.73
April	21.32	19.45	19.57	23.69	21.57	21.70	22.57	20.30	20.27
May	22.82	19.40	19.48	23.53	20.17	20.14	22.22	20.82	20.66
June	17.98	17.50	17.33	20.55	19.25	19.13	22.16	20.59	20.35
July	13.97	17.07	16.88	13.08	16.57	16.25	17.17	19.33	18.96
August	13.41	15.68	15.42	15.05	15.96	15.68	15.96	17.62	17.27
September	14.29	14.46	14.26	18.44	15.80	15.66	13.86	15.33	15.08
October	14.69	14.37	14.42	17.42	15.44	15.47	15.23	13.18	13.14
November	16.14	14.29	14.53	15.15	13.89	13.95	11.46	11.50	11.57
December	14.05	14.99	15.34	13.85	15.51	15.72	9.85	11.85	12.08

The estimated GSR values are compared with measured GSR values from IMD, Pune. The performance of the models are evaluated by using selected statistical indicators, namely, Mean absolute percentage error (MAPE), Mean Bias Error (MBE), Root Mean square error (RMSE) and t-stat which are defined in Appendix B. The accepted range of the values of the errors should be less than 10%. RMSE values are always positive and smaller values indicate the better model performance. Maximum and minimum temperature records for the three regions are shown in Fig.4 and Fig.5. respectively. According to Fig.6. the t-stat values computed by

GA is lesser than SRT. Smaller the t-stat values better the model performance. As seen in Fig.7.the mean bias error values computed by GA are also lesser than SRT. Comparison of error statistics between SRT and GA are given in Table 6.The best results are obtained by GA technique .The error values are in the acceptable range.(<10%)The t-stat values computed for all the locations are less than the critical value of 2.2 which clearly shows that the models are statistically significant. The measured and predicted values of monthly average daily GSR by SRT and GA are shown in Fig.8.The predicted GSR values by GA technique are in good agreement with the measured values.

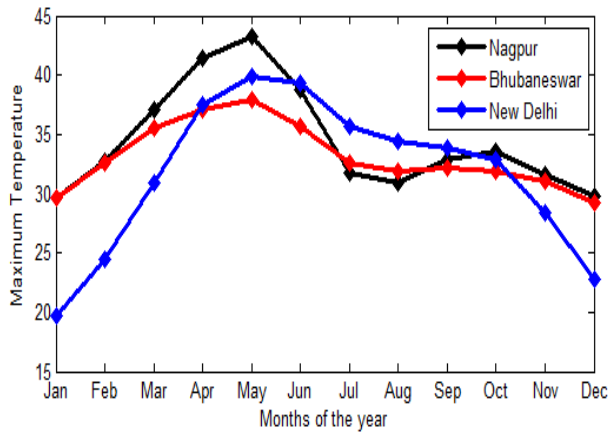


Fig. 4. Measured maximum temperature for the study sites provided by IMD,Pune

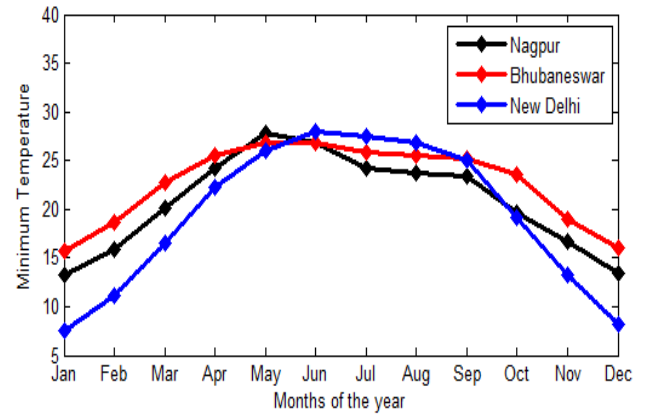


Fig. 5. Measured minimum temperature for the study sites provided by IMD,Pune

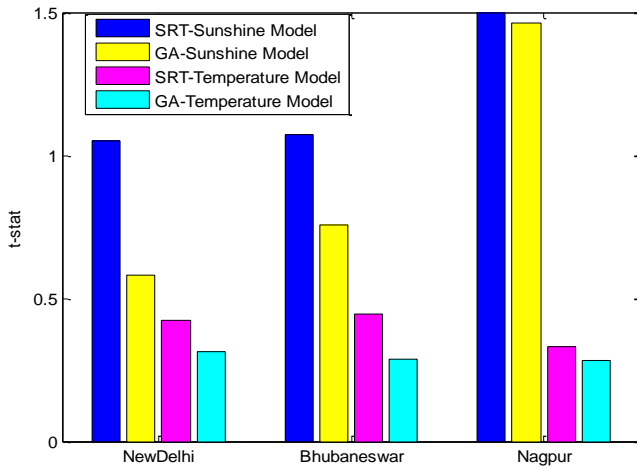


Fig. 6. Comparison of t-stat between SRT and GA technique

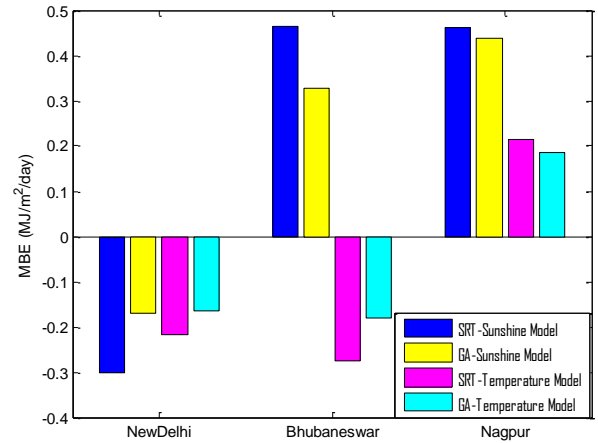


Fig. 7. Comparison of MBE between SRT and GA technique

Table 6: Comparison of error statistics between SRT and GA

Location	Technique	Temperature-based Model					Sunshine-based Model				
		R	MBE	MAPE	RMSE	t-stat	R	MBE	MAPE	RMSE	t-stat
Bhubaneswar	SRT	0.743	-0.276	3.152	2.067	0.447	0.926	0.466	3.322	1.513	1.073
	GA	0.725	-0.181	2.602	2.098	0.288	0.929	0.329	2.598	1.479	0.756
Nagpur	SRT	0.776	0.214	2.644	2.139	0.333	0.953	0.462	2.732	1.122	1.499
	GA	0.770	0.191	2.481	2.154	0.294	0.955	0.439	2.646	1.086	1.464
New Delhi	SRT	0.931	-0.218	0.589	1.733	0.421	0.977	-0.302	1.090	1.000	1.052
	GA	0.934	-0.165	1.039	1.749	0.314	0.976	-0.171	0.248	0.989	0.581

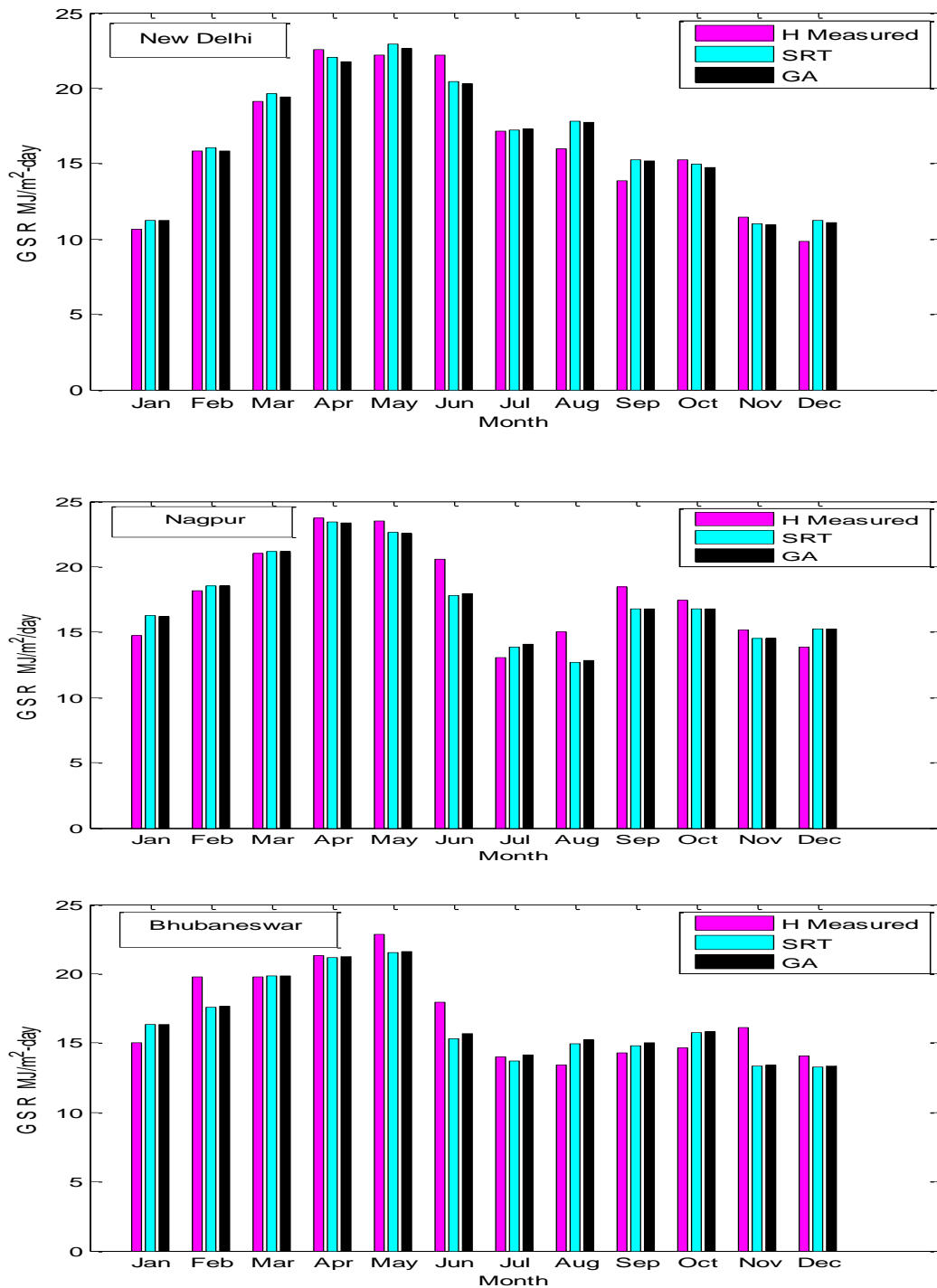


Fig. 8. The predicted values of monthly mean daily global solar radiation on testing data by GA and SRT for the study sites New Delhi, Nagpur and Bhubaneswar.

## 6. Conclusion

The aim of this work is to assess SRT and GA technique for the estimation of solar radiation on horizontal surface in India. For this purpose, angstrom type regression models based on sunshine

duration and temperature are considered. GSR is estimated at three locations of India namely Bhubaneswar, Nagpur and New Delhi. The regression coefficients of the models are determined by SRT and GA technique separately. The estimated GSR values

by SRT and GA are compared with measured GSR of India Meteorological Department (IMD), Pune. The models are validated against the measured GSR data based on statistical measures such as RMSE, MBE, MAPE, correlation coefficient (R) and t-statistics. The results indicated that obtained regression coefficients derived from GA performed better than the regression technique (least square method). The correlation coefficient R ranges from 0.92 to 0.97. The t-stat values and the statistical errors computed by GA are less when compared to SRT. Hence it is concluded that, GA technique can be effective in estimating GSR wherever the radiation data is not available. Future work will be focused on solar radiation prediction using other optimization techniques and machine learning approaches like SVM, regression tree, random forest and other models to improve the accuracy of the GSR prediction.

### Acknowledgments

The authors gratefully acknowledge India Meteorological Department, Pune for providing the meteorological data

### Appendix A

Calculation of Extra-terrestrial radiations and maximum possible sunshine duration are as follow;

$$H_0 = \frac{24}{\pi} I_{sc} \left[ 1 + 0.33 \cos \left( \frac{360 D_n}{365} \right) \right] \times \left[ \cos L \cos \delta \sin \omega_s + \frac{2\pi \omega_s}{360} \sin L \sin \delta \right] \quad (2)$$

where  $I_{sc}$  is the solar constant.  
 $D_n$  is the day of year starting from first January;  
 $L$  is the latitude of location under consideration;  
 $\delta$  is declination angle as given below

$$\delta = 23.45 \sin \left[ \frac{360(284 + D_n)}{365} \right] \quad (3)$$

and  $\omega_s$  is sunset hour angle in degree as given below

$$\omega_s = \cos^{-1}(-\tan L \tan \delta) \quad (4)$$

For a given month, the maximum possible sunshine duration can be computed by the following equation:

$$S_o = \frac{2}{15} \cos^{-1}(-\tan L \tan \delta) \quad (5)$$

### Appendix B

MAPE, MBE, RMSE and t-stat are defined as follows:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{H_m - H_c}{H_m} \right| \times 100 \% \quad (6)$$

Where  $H_m$  is measured and  $H_c$  is calculated GSR;  $n$  is the number of observations.

$$MBE = \frac{1}{n} \sum_{i=1}^n (H_{im} - H_{ic}) \times 100 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (H_{im} - H_{ic})^2} \quad (8)$$

$$t - stat = \sqrt{\frac{(n-1) MBE^2}{RMSE^2 - MBE^2}} \quad (9)$$

### References

- [1] Angstrom, A.: *Solar and terrestrial radiation*. In: Quarterly journal of the Royal Meteorological Society (1924), 50(4), p.121-26, Chichester, UK.
- [2] Prescott, J.A.: *Evaporation from a water surface in relation to solar radiation*. In: Transaction of royal society, (1940), 64, p.114-18, South Australia.
- [3] Gasser Hassan, E., Elsayed Youssef, M., Zahraa Mohamed, E., Mohamed Ali, A., Ahmed Hanafy, A.: *New temperature based models for predicting global solar radiation*. In: Applied Energy (2016), 179, p.437-50.
- [4] Meenal, R., Immanuel Selvakumar, A.: *Temperature based Radiation Models for the Estimation of Global Solar Radiation at Horizontal Surface in India*. In: IJST Vol 9/46, (2016), 101922, India.
- [5] Hatice Citakoglu: *Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation*. In: Computers and Electronics in Agriculture (2015), 115, p.28-37.
- [6] Benganem, M., Mellit, A., Alamri, S.N.: *ANN-based modelling and estimation of daily global solar radiation data: A case study*. In: Energy Conversion and Management, 50 (2009), p.1644-55.
- [7] Behrang, M.A., Assareh, E., Noghrehabadi, A.R., Ghanbarzadeh, A.: *Sunshine based models for predicting global solar radiation using PSO (particle swarm optimization) technique*. In: Energy, 36(2011), p.3036-49.
- [8] Farzad Fathian: *Predicting global solar radiation using Genetic algorithm [GA]*. In: Bull. Env. Pharmacol. Life Sci., Vol 2 (6) May (2013), p.54-63.
- [9] Fariba Beshrat, Ali Dehghan, A., Ahmad Faghieh R.: *Empirical Models for estimating global solar radiation: A review and case study*. In: Renew sustain Energy Rev. 21, (2013), p.798-821.
- [10] Rich Inman, H., Hugo Pedro, T.C., Carlos Coimbra, F.M.: *Review Solar forecasting methods for renewable energy integration*. In: Progress in Energy and Combustion Science, 39 (2013), p.535-76.