Journal of Technical Science & Technologies; ISSN: 2298-0032; e-ISSN: 2346-8270; Volume 5, Issue 2, 2016

Solar Radiation Modeling Using M5 Model Tree

Ozgur KISI*
Okan ERAY**
Cihan MERT***

Abstract

The study investigates the ability of M5 model tree in modeling monthly solar radiation. Monthly data of maximum and minimum air temperatures, wind speed, relative humidity and solar radiation from Antalya, Turkey were used in the application. The effect of each climatic parameter on solar radiation is investigated. Periodic models are also developed to see the effect of periodicity component on models' accuracy. It was found that using periodicity as input to the applied models significantly increases their accuracy in modeling monthly solar radiation.

Keywords: Solar radiation, M5 model tree, modeling

Introduction

Solar engineers, agriculturists and architects need solar radiation for many applications such as solar heating, cooking, drying and interior illumination of buildings (Jiang, 2009). Solar radiation is also important for the growth period and development of plants (Citakoglu, 2015).

In the last decades, Model tree (MT) was used for modelling rainfall-runoff by (Solomatine & Dulal, 2003), in forecasting flood by (Solomatine and Xue, 2004), for determining stage discharge relationship by (Bhattacharya & Solomatine, 2005) , for modelling sediment by (Bhattacharya & Solomatine, 2006), for estimating reference evapotranspiration by (Pal & Deswal, 2009). MT based regression approach was used by (Pal & Deswal, 2009) for modelling daily reference evapotranspiration based on air temperature, solar radiation, wind speed and relative humidity. The accuracy of support vector machine and MT was compared in forecasting streamflow by (Sattari, Pal, Apaydin, Ozturk,2013) and later one was found to be better than the first method.

The purpose of the present study is to estimate solar radiation of Antalya City, Turkey using M5 model tree method.

M5 model tree

M5 model tree (M5tree), first presented by (Quinlan, 1992), is based on a binary decision tree comprising linear functions at the leaf nodes and can also be utilized for quantitative data (Rahimikhoob, 2014). M5tree needs two different phases: i) A decision tree is created by dividing data into subsets in the first phase, ii) Overgrown tree is pruned and sub-trees are replaced with linear regression functions in second phase. Division of the input space $X_1 \times X_2$ (where X_1 and X_2 are the independent parameters) into 4 subspaces is shown in Figure 1. LM1, …, LM4 at the leaves denote the linear regression functions. Tree structure of M5tree is shown in Figure 2 (Rahimikhoob, 2014). More information about M5tree can be obtained from (Quinlan, 1992).

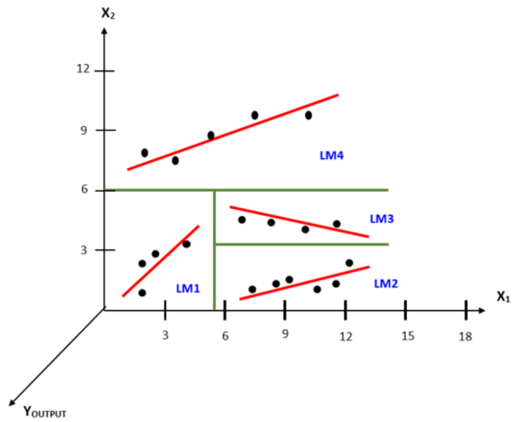

Figure 1. *Division of the input space X1×X2 into 4 subspaces in M5Tree.*

*Ph.D., Faculty of Computer Technologies and Engineering, International Black Sea University, Tbilisi, Georgia.
E-Mail: okisi@ibsu.edu.ge
**Assoc. Prof. Dr., Faculty of Computer Technologies and Engineering, International Black Sea University, Tbilisi, Georgia.
E-Mail: okaneray@ibsu.edu.ge
***Assoc. Prof. Dr., Faculty of Computer Technologies and Engineering, International Black Sea University, Tbilisi, Georgia.
E-Mail: cmert@ibsu.edu.ge

Journal of Technical Science & Technologies; ISSN: 2298-0032; e-ISSN: 2346-8270; Volume 5, Issue 2, 2016

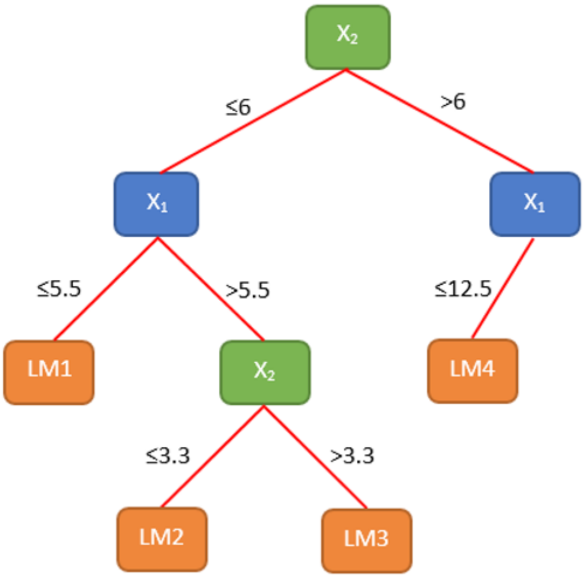

Figure 2. *Tree structure of M5tree*

Application and results

In the current study, the ability of M5tree was investigated by using data from Adana station, Turkey. Data cover 391 monthly values with a period of 1981-2016. 293 values (75% of the whole data) were used for calibrating M5tree models and remaining 98 values (25%) were used for evaluating the accuracy of the developed models. Following input scenarios were tried in the application:

i) Tmax,

ii) Tmax and Tmin

iii) Tmax, Tmin and W,

iv) Tmax, Tmin, W and RH.

where Tmax, Tmin, W and RH indicate the maximum air temperature, minimum air temperature, wind speed and relative humidity. The output of the models is monthly SR. The evaluation criteria, root mean square errors (RMSE), mean absolute relative errors (MARE), mean absolute errors (MAE) and determination coefficient (R^2) used in the present study are:

$$(1) RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(SRi_{model} - SRi_{observed})^2}$$

$$(2) MARE = \frac{1}{N}\sum_{i=1}^{N}\frac{|SRi_{model} - SRi_{observed}| \cdot 100}{SRi_{model}}$$

$$(3) MAE = \frac{1}{N}\sum_{i=1}^{N}|SRi_{model} - SRi_{observed}|$$

where N denotes the number of data.

The accuracy comparison of the M5tree models are provided in Table 1 with respect to four different statistics. It is clear from the table that the model with three inputs (Tmax, Tmin and W) has the best accuracy (RMSE=25.46, MARE=20.30, MAE=19.06 and R^2=0.697). The observed and estimated SR values using four different input scenarios are graphically compared in Figure 3 for the test stage. As seen from the figure that the M5tree3 model using the inputs, Tmax, Tmin and W has less scattered estimates than the other models.

$$(4) R^2 = \left[\frac{\sum_{i=1}^{N}(SRi_{observed} - SRi_{observedmean})(SRi_{model} - SRi_{modelmean})}{\sum_{i=1}^{N}(SRi_{observed} - SRi_{observedmean})^2 \sum_{i=1}^{N}(SRi_{model} - SRi_{modelmean})^2}\right]^2$$

Table 1. *The accuracy of M5tree models in estimating monthly SR*

Inputs	Model	RMSE	MARE	MAE	R^2
Tmax	M5tree1	25.33	21.11	20.87	0.642
Tmax and Tmin	M5tree1	26.54	22.42	21.39	0.611
Tmax, Tmin and W	M5tree3	23.39	20.30	19.06	0.697
Tmax, Tmin, W and RH	M5tree4	25.46	20.43	19.24	0.657

Journal of Technical Science & Technologies; ISSN: 2298-0032; e-ISSN: 2346-8270; Volume 5, Issue 2, 2016

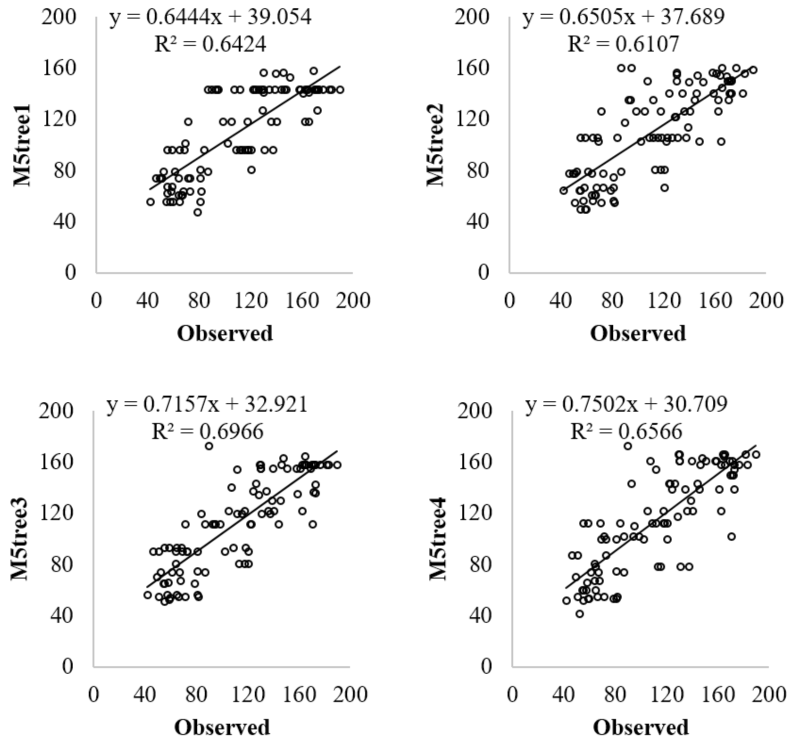

Figure 3. *The scatterplots of observed and estimated SR by M5Tree models in test stage.*

The effect of periodicity component on models' accuracy was also investigated by adding month of the year as input to the applied models. The test results of the periodic models are reported in Table 2 in respect of RMSE, MARE, MAE and R^2. In this table α denotes the periodicity and takes values from 1 to 12. From the table, It is apparent that the Periodic-M5tree1 with two inputs (Tmax and α) has the best accuracy from the MARE and R^2 statistics. The Periodic-M5tree2 model has slightly better RSME and MAE values than the first model. First model should be preferred because it has less number of input parameters. Comparison of Table 1 and 2 shows that adding periodicity into the model inputs significantly improves M5tree accuracy in estimating monthly SR. For example, Periodic-M5tree1 increase the RMSE, MARE and MAE accuracy of the M5tree1 model by 48%, 56% and 51%, respectively. The observed and estimated SR values by periodic M5tree models are graphically compared in Figure 4 for the test stage. Periodic-M5tree3 model seems to have less scattered estimates than the other models especially for the peak SR values. Comparison with the Figure 1 clearly indicates the significant positive effect of periodicity component on models' accuracies in modeling monthly SR.

Table 2. *The accurancy of Periodic-M5tree models in estimating monthly SR*

Inputs	Model	RMSE	MARE	MAE	R^2
Tmax and α	Periodic-M5tree1	13.25	9.40	10.18	0.909
Tmax, Tmin and α	Periodic-M5tree1	13.23	9.86	10.16	0.907
Tmax, Tmin, W and α	Periodic-M5tree3	14.17	10.73	11.16	0.896
Tmax, Tmin, W, RH and α	Periodic-M5tree4	13.94	10.41	11.03	0.904

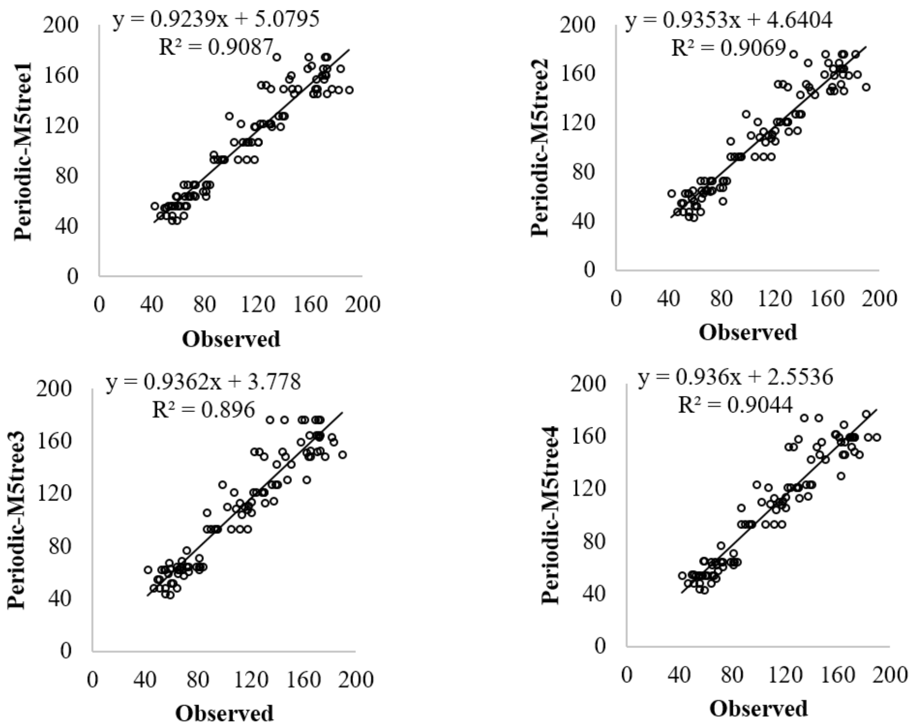

Figure 4. The scatterplots of observed and estimated SR by periodic M5Tree models.

Model tree obtained from the Periodic-M5tree1 model comprising temperature and periodicity component as inputs is given in Table 3. It is clear from the table that model tree involves if-then statements and can be simply used in practical applications.

Table 3. *The model tree of the Periodic-M5tree1*

if x2 <= 31.45	if x1 <= 11.5	if x2 <= 37.95
if x1 <= 10.5	y = 63.683182 (22)	if x2 <= 37.55
if x1 <= 2.5	else	y = 135.69172
if x1 <= 1.5	if x2 <= 21.55	+4.2441904*x1 (30)
y = 56.300833 (24)	y = 44.477778 (9)	else
else	else	y = 148.336 (5)
if x2 <= 21.6	if x2 <= 23.1	else
y = 67.589231 (13)	y = 54.81375 (8)	if x2 <= 39.25
else	else	y = 174.576 (10)
y = 73.331818 (11)	y = 48.378571 (7)	else
else	else	if x2 <= 40.5
if x1 <= 3.5	if x1 <= 8.5	y = 159.92333 (9)
if x2 <= 24.1	if x2 <= 35.1	else
y = 97.071429 (7)	if x1 <= 4.5	y = 167.9625 (4)
else	y = 128.21385 (13)	else
y = 106.81824 (17)	else	y = 149.05348 (23)
else	y = 145.1815 (20)	else
y = 119.66727 (11)	else	y = 375.17354 -
else	if x1 <= 7.5	28.207101*x1 (50)

Journal of Technical Science & Technologies; ISSN: 2298-0032; e-ISSN: 2346-8270; Volume 5, Issue 2, 2016

Conclusions

In the current study, the accuracy of a popular soft computing method, M5 model tree, in estimating monthly solar radiation was investigated. Monthly maximum and minimum air temperatures, wind speed and relative humidity data obtained from Antalya Station, Turkey were used as inputs to the models. The study examined the effect of each climatic parameter on solar radiation and maximum temperature was found to be the most effective parameter on solar radiation for the studied station. The effect of periodicity component on models' accuracy was also investigated and it was found that adding this component to the M5 model tree inputs considerably increase its accuracy in monthly solar radiation modeling. It can be concluded that solar radiation of Antalya Station can be successfully modeled by using only maximum temperature data.

References

Bhattacharya B, Solomatine DP (2005) Neural networks and M5 model trees in modeling water level–discharge relationship. Neurocomputing 63,381–396

Bhattacharya B, Solomatine DP (2006) Machine learning in sedimentation modelling. Neural Netw 19(2),208– 214

Citakoglu, H. (2015). Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation, Computers and Electronics in Agriculture, 118, 28-37.

Jiang, Y. (2009). Computation of monthly mean daily global solar radiation in China using artificial neural networks and comparison with other empirical models, Energy, 34, 1276-1283.

Pal M, Deswal S (2009) M5 model tree based modelling of reference evapotranspiration. Hydrol. Process 23, 1437–1443

Quinlan J.R. (1992). Learning with continuous classes. In Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence, Hobart, Australia, 16–18 November, World Scientific, Singapore: 343–348.

Rahimikhoob, A. (2014). Comparison between M5 Model Tree and Neural Networks for Estimating Reference Evapotranspiration in an Arid Environment, Water Resource Manage, 28: 657–669.

Sattari MT, Pal M, Apaydin H, Ozturk F (2013) M5 model tree application in daily river flow forecasting in Sohu stream, Turkey. Water Resources 40(3),233–242

Solomatine DP, Dulal KN (2003) Model trees as an alternative to neural networks in rainfall-runoff modelling. Hydrol Sci J 48(3),399–411

Solomatine DP, Xue Y (2004) M5 model trees compared to neural networks: application to flood forecasting in the upper reach of the Huai River in China. J Hydr Engrg 9(6),491–501