Websites:
http://www.sciencepub.net
http://www.sciencepub.net/report

Emails:
editor@sciencepub.net
reportopinion@gmail.com

Report and Opinion

MARSLAND PRESS
Multidisciplinary Academic Journal Publisher

Evaluation of M5 Model Tree and Neural Network Model in Estimating the Radiation Reaching the Earth

Narges Kefayati, Ali rahimi khob , Aliakbar Noroozi

1 *Aburayhan Campus, Tehran University, Iran
2 Associate professor, Department of irrigation and drainage, Aburayhan Campus, Tehran University, Iran
. 3 Institute of Soil and Water Conservation. Tehran, Iran
narges.kefayati@gmail.com

Abstract: Daily solar radiation reaching the land surface is one of the most common parameters used in modeling and projects of hydrology, agriculture, meteorology and climatology. Most estimating models of radiation reaching the land surface using satellite data are based on the land surface temperature. In this study, the accuracy of solar radiation estimation was studied using artificial neural network models and M5 model tree, with the inputs including land surface temperature and water vapor of MODIS sensor in combination with extraterrestrial radiation. The solar radiation data were collected from six stations in three consecutive years (2003- 2005) and used as actual data. The stations include (Bojnoord, Isfahan, Hamadan, Kermanshah, Khurobiabanak and Tabriz). Based on the results, the neural network model of Khurobiabanak station with R^2 0 /97 and Hamadan station with the R^2 0/71, have the highest and the lowest accuracy in estimating solar radiation, respectively. The results showed that both models estimate the radiation reaching the Earth's surface with good accuracy and provide almost identical results.

[Narges Kefayati, Ali rahimi khob , Aliakbar Noroozi. **Evaluation of M5 Model Tree and Neural Network Model in Estimating the Radiation Reaching the Earth**. *Rep Opinion*2021;13(2):72-82]. ISSN 1553-9873 (print); ISSN 2375-7205 (online). http://www.sciencepub.net/report.. 9. doi:10.7537/marsroj130421.09.

Key words: solar radiation, land surface temperature; water vapor; MODIS

Introduction

The radiation reaching the Earth's surface is one of the most important parameters in energy balance models, plant growth, crop production and potential and actual evapotranspiration. Despite the importance of this parameter, the measurement was conducted only in limited areas due to the expensive measurement equipment, high maintenance costs and difficulty of calibration procedure. (The devices must be calibrated regularly and need to be maintained (Samani, 2000). On the other hand, station measurements are not enough for large surfaces because they are spotted (Rahimikhoob et al., 2010). Therefore, different methods are provided for estimating the radiation in different regions of the world including the empirical and regression correlations, remote sensing, random meteorological models, linear interpolation and artificial neural networks (Trnka et al., 2005). The solar radiation reaches the earth in two phases. In the first phase, some solar energy reaches the upper atmosphere vertically which is called extraterrestrial radiation that is a function of latitude and year calendar and does not need to be measured and is obtained through a series of mathematical and analytical relations. In the second phase, the radiation which passes through the atmosphere is knocked up with water vapor molecules and particles in atmosphere and eventually reaches the earth's surface that its estimation is intricate due to the atmospheric attenuation and is a function of cloudiness and moisture content of the air (Allen, 1998). In recent years, the models of remote sensing that its input data are obtained from satellite imagery are used to determine some phenomena. The advantage of satellite images is that they cover broad and wide areas in which spatial changes of phenomena can be investigated. But many terrestrial phenomena cannot be determined directly by satellite imagery such as radiation reaching the Earth and the models with satellite data input should be developed (Rahimikhoob et al., 2010). Some studies have been conducted so far in estimating the radiation reaching the Earth's surface using satellite imagery data combined with ground data. For example, Ozan (2010) used the artificial neural network model with inputs including the surface emissivity of bands 4 and 5 of AVHR images of NOAA satellite, height level, longitude, latitude and temperature of the earth's surface that is estimated with Split Window Method using Coll et al. algorithm (1994), the results had the correlation coefficients 0/95and 0/ 93 for data intended for test and learning phase respectively. In a research, statistical model for estimating solar radiation received by the Earth was calibrated and evaluated in Kermanshah. This model is based on the determination of cloud covering index for each pixel of satellite images and linear correlation between the cloud covering index and air cloudless indicator. The results showed that there is a good compatibility between the radiation from model and the radiation measured at studied station so that the determination coefficient R^2 was equal to 0/86 and root mean square error RMSE was equal to 0/8 Mj m^{-2} d^{-1} (Saberi et al., 2010). In another study, a simple physical model was presented for measuring solar radiation using satellite data in southern Canada. In this study, the effects of water vapor absorption and clouds mass were evaluated on radiation scattering. The results showed that the standard error in clear days is less than 5% of average and in very cloudy days it is 15-14% less than the average (Gautier et al., 1980). Recently, machine learning methods (Soft Computing Techniques) are used as new methods for modeling complex relationships. Among them are artificial neural network model and M5 model tree. In a research, the ability of M5 model tree was evaluated for estimating the land surface temperature in Tabriz station with an experimental model of Angstrom. For this purpose, they used meteorological data measured for land surface temperature in climatic period (1995-2000). The results showed that the M5 model tree had better function to Angstrom equation in estimating the land surface temperature (Imamifar and Rahimikhoob, 2011). In another study, the LST data of MOD11A1 products of MODIS were used to estimate the average daily air temperature in Khuzestan Province in combination with geographical variables (day of the year and extraterrestrial radiation) as inputs of M5 model tree. The results showed that the accuracy of developed model in estimating the average daily air temperature has a determination coefficient equal to 0/96 and the mean squares error equal to 2/3 ° C (Imamifar et al., 2012). In another study, neural network models and M5 model tree were compared to convert land surface temperature in day and night times of MODIS products of Terra satellite for Khuzestan Province. The input data of models includes land surface temperature in day and night and extraterrestrial radiation. The results showed that the determination coefficient of both models is more than 0/ 96. However, the neural network model estimates the temperature more accurately (Imamifar et al., 2014). In a study, Hargreaves-Samani and neural network models were compared for estimating radiation in Ahvaz. The maximum and minimum temperatures data were used as inputs to these models. The results of this study showed that among the studied models, neural network model provides better results and absolute mean percentage of error is equal to 2/53% (Rahimikhoob, 2010). In another study, the neural network model with input of different products of MODIS include a monthly average land surface temperature and vegetation index combined with ground-based data were used to estimate solar radiation. The results showed that the developed model has high accuracy (Qin et al., 2011). In a research, five experimental models were compared with neural network models for the East Southern District of Tehran. The neural network models were designed with four different input combinations which ANN1 model has three input parameters, ANN2 has four parameters, ANN3 has five parameters and ANN4 has six parameters. The results showed that the neural networks estimate the solar radiation with high accuracy in general, and models based on sunshine hours are better than models based on air temperature. ANN4 model with one hidden layer arrangement and structure of six input parameters with 14 neurons and instructional Levenberg algorithm and sigmoid transfer function Exxon, the determination coefficient equal to 0/96 and root mean square error RMSE equal to 0/93 Mj m^{-2} d^{-1} gave the best result (Rahimikhoob et al., 2009). In another study, the accuracy of estimating solar radiation was studied using four different models of neural networks with input of MODIS land surface temperature products in combination with external radiation and sunshine hour's ratio. The results showed that all four neural network models were able to estimate radiation reaching the earth with good correlation $R^2 > 0/85$ (Imamifar and Alizadeh, 2015). According to the good results of M5 model tree and neural network model, especially in remote sensing studies (Rahimikhoob, 2012 and Imamifar, 2013), the comparison of the performance of two neural network model and M5 model tree is one of the goals of this research for conversion of land surface temperature in day and night and atmosphere water vapor from MODIS images to radiation reaching the Earth. Finding the best combination among the input variables of land surface temperature in day and land surface temperature in night, extraterrestrial radiation and atmosphere water vapor in each of the neural networks and M5 model tree is the second objective of this research.

Materials and Methods

To do this research, the solar radiation data of six synoptic stations were used for a three-year period (2003 - 2005) in Iran. Areas of study include: Tabriz, Hamadan, Kermanshah, Isfahan, Khurobiabanak and Bojnoord. These areas have different climatic and geographical location. It should be noted that the climate of stations was based on Emberger classification. The specification of stations studied is presented in Table and Fig 1.

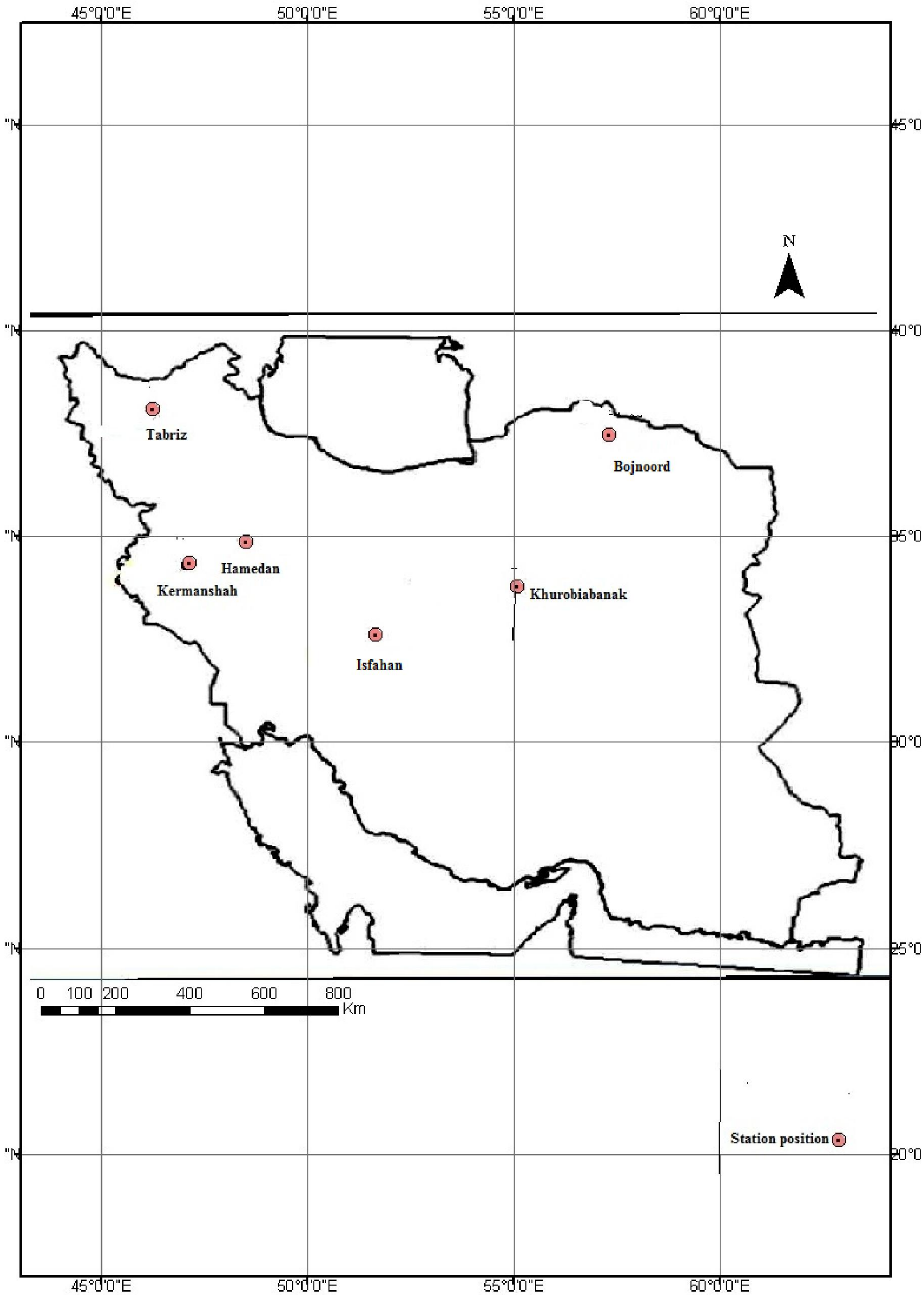


Fig 1: Geographical location of the stations studied

Table 1- details of the synoptic stations studied

Station name	Climate type Emberger Classification)	Established year	latitude (Degree-minute)	Longitude (Degree-minute)	Above sea level (M)
Hamadan	Cold semi-arid	1976	3452	4832	1741.5
Kermanshah	Mild	1951	3421	4709	1318.6
Tabriz	Cold and dry	1951	3805	4617	1361
Isfahan	Dry	1951	3237	5140	1550.4
Khurobiabanak	Very dry	1986	3347	5505	845
Bojnoord	Semi-arid	1977	3728	5719	1091

MODIS Images of Terra Satellite

In this research, the MODIS sensor data was used which is from Terra satellite. The design of elements of this sensor is applied to high signal noise that provides its capabilities in sciences such as agriculture, meteorology, geology and oceanography. The MODIS does a daily series of observations on a global scale on sea, land and atmosphere. The MODIS data includes day and night data which enjoys a good resolution (Wan et al., 2002).The pictures used in this study are MODIS Terra satellite products, land surface temperature and atmosphere water vapor that the images are from the http://modis.gsfc.nasa.gov// website. The land surface temperature data is a part of sub-group of level three (L3) data of MODIS with a characteristic code MOD11A1 related to Terra satellite. From the features of this product is that it has a spatial resolution capability of one kilometer and daily temporal resolution power. The data of atmosphere water vapor is a part of the subgroup of level two (L2) data of MODIS images with a characteristic code MOD05_L2 related to Terra satellite. Its pixel size is 1 × 1 km and 5 × 5 km. The valid range for water vapor data from the MODIS is (20-0) cm. The MODIS uses two algorithms of infrared and near-infrared for receiving the images of atmosphere water vapor column. The infrared algorithm takes images just in day and taken pictures are clearer and more transparent and near-infrared algorithm takes images in day and night. The conversion of coordinate system of images from sinus to UTM was done using the MCT tools and the data extraction at each station was done in the software environment ARCGIS 9.3 using HAT tools. The LST values extracted from the images are in terms of Kelvin degree, so before using them as input of models, these values were altered to Celsius degree. It is worth noting that each MODIS image loses a number of information due to the air cloudiness, high concentration of atmospheric aerosols, and the gap between two satellite tracks. So each station lacks data in picture on some days.

Artificial Neural Networks

The neural networks are computational models that can learn the relationship between inputs and outputs. The structure of the neural network includes an input layer, a middle layer and an output layer. In each layer there are one or more processing elements that are referred to as neurons. The vector of input data to model are entered to first layer neurons, in this layer no any type of processing is done and the output layer neurons are entered to the output vector of model . The number of neurons in input and output layers depends on the number of model's input and output variables, but the number of neurons in the middle layer is determined by trial and error (Bakhshoudeh and Rahimikhoob, 2014). In this study, Nero Solution Software version 5 is used to prepare an artificial neural network model for estimating the solar radiation reaching the earth's surface. The network that is used for modeling of solar radiation is multilayer Perceptron networks with instructional back propagation algorithm. The Levenberg-Marquardt algorithm is the most common back propagation algorithm. The transition function used in neurons is sigmoid function. In this study, we used a hidden layer. The number of neurons in the middle layer was specified by trial and error. Thus, for each station, the model was instructed 10 times with 1 to 10 neurons in the middle layer, and then the statistical index of root mean square error (RMSE) was estimated using the data assigned to the model test. The lowest amount of this index for each station was the basis for the selection of the number of neurons in the middle layer, while among the instruction and evaluation data, 70% was allocated to instruction and remaining 30 per cent to network assessment. The design and instruction of neural network requires input output data until we can extract nonlinear relationships between them by logical analysis and simulate simulation operation for possible similar cases (Bayat, 2008). The input data includes land surface temperature (day), land surface temperature (night), the atmosphere water vapor and the extraterrestrial radiation that the first three variables are obtained respectively from MODIS images (Terra satellite) and the extraterrestrial radiation was calculated from Allen and colleagues (1998) equation and the output data is the radiation reaching the earth. Figure 2 shows a scheme of a multilayer Perceptron network.

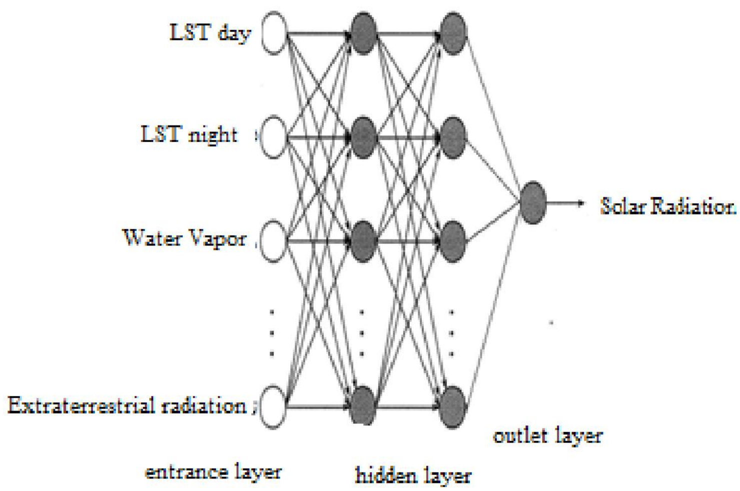


Figure 2- Scheme of a Multilayer Perceptron Network

Basically, inserting raw data reduces network speed and accuracy. To avoid such situations, as well as to equalize the value of data to the network, normalizing action takes place. The data normalizing takes place before instructing and during the span (0-1) (Rahimikhoob, 2010). In this study all data including input and output were normalized before applying to the network using Equation 1.

$$xn = \frac{(x - xmin)}{(xmax - xmin)} \quad (1)$$

Where x_n is the normalized value of each parameter, x is actual value of each parameter, x_{min} is the lowest value and x_{max} is highest value of each parameter. The network with different input is considered to estimate the radiation in neural networks.

1. The network with two inputs (extraterrestrial radiation and atmosphere water vapor)
2. The network with three inputs (extraterrestrial radiation, land surface temperature of day, land surface temperature of night)
3. The network with four inputs (extraterrestrial radiation, land surface temperature of day, land surface temperature of night and atmosphere water vapor)

Extraterrestrial Radiation (R_a)

The extraterrestrial radiation introduces solar radiative energy that reaches the top of the atmosphere vertically and exhibits the reflection of daily changes in air temperature. It is the function of yearly calendar, latitude and time of passing satellite. This radiation was estimated from equation 2 (Allen et al., 1998).

$$Ra = \frac{24(60)}{\pi} Gsc\, dr[ws \sin\varphi \sin\delta + \cos\varphi \cos\delta \sin ws] \quad (2)$$

In the above equation, R_a is extraterrestrial radiation in MJ per square meter per day, G_{sc} is the solar constant (equal to 1367 watts per square meter), d_r is the inverse relative distance of Earth and the Sun, j is the year calendar, and δ ،φ and ω_s *are* the declination angle of the sun, latitude and time angle in radians. The suggested relations of Allen and colleagues (Allen et al., 1998) were used to set d_r, φ andω_s.

$$\mathrm{dr} = 1 + 0.033 \cos\left(\frac{2\pi}{365}\,\mathrm{j}\right) \quad (3)$$

$$\delta = 0.409 \sin\left(\frac{2\pi}{365}\,\mathrm{j} - 1.39\right) \quad (4)$$

$$\mathrm{ws} = \cos^{-1}[-\tan(\varphi)\tan(\delta)] \quad (5)$$

M5 Model Tree

The M5 model tree is based on tree classification method that was presented by Quinlan (1986) to create a relationship between independent and dependent variables (Imamifar et al., 2013).The M5 tree is a binary decision tree and follows the classification rules of input data based on the irregularities measuring. Finally, allocates a linear relationship to the resulting referrals which can predict numeric values. The structure of decision tree is similar to a tree that is composed of roots, branches, nodes and leaves. The decision tree can be drawn from top to bottom. The root as the first node is located at the top and a chain of branches and nodes ends to leaf. Each node is related to a predictive variable and branching operation is performed in the node by branches. The branches include numerical span which is split from the parent node and reaches a child node (Quinlan, 1992). In M5 model two branches are branched from each parent node. The decision tree making is done in two phases. In the first phase, the decision tree is formed with a branching of data, the selection criterion in M5 model tree is the maximizing of deviation reduction of data standard in the child node. Reducing the standard deviation is calculated from equation (6) (Pal et al., 2009).

$$SDR = sd(t) - \sum \frac{Ti}{T} sd(Ti) \qquad (6)$$

Where, T is a batch of samples that reaches to the node, Ti is a referral of samples which has i^{th} output from potential batch and sd is the standard deviation. In the M5 model, the branch which produces the maximum standard deviation is selected after examining all the possible ramifications of a node. This branching often leads to creation of a great tree and over fit occurs on the training data (Quinlan, 1986). The over fit causes a decrease in the overall part of the model so that the model is valid only for data that were used to build it and has not sufficient accuracy for new data. So the second stage of M5 model tree designing contains minifying the over large tree by pruning the branches and replacing them by linear regression functions (Quinlan, 1992). In this survey, the weka software was used which is among the most reliable and widely used software written in the field of data mining and especially in the field of model trees making and developed by the University of Waikato in New Zealand. This comprehensive software includes solutions for all data mining problems such as regression, classification, clustering, rule congruence and selection of traits. The data were divided into two categories to develop a M5 model tree. Data from 2003 and 2004 were used to train and data from 2005 were used to test. In the training phase, the input variables include: (land surface temperature of day, land surface temperature of night, atmospheric water vapor and extraterrestrial radiation) for all six stations in 2003 and 2004 were given to M5 model tree. The M5 model tree gives an equation in terms of input parameters, which are used to test the data for 2005 in each station.

Statistical Indicators

It needs to choose appropriate indicators to check the results and compare them quantitatively to be able to analyze the effective factors and the results from phenomenon. In this study, three statistical indicators of determination coefficient R^2, mean bias error MBE and root mean square error RMSE was used.

$$\mathrm{R}^2 = \frac{\left[(\sum \mathrm{xy}) - \frac{(\sum \mathrm{x})(\sum \mathrm{y})}{\mathrm{n}}\right]^2}{\sqrt{\left[\sum \mathrm{x}^2 - \frac{(\sum \mathrm{x})^2}{\mathrm{n}}\right]\left[\sum \mathrm{y}^2 - \frac{(\sum \mathrm{y})^2}{\mathrm{n}}\right]}} \qquad (7)$$

$$\mathrm{MBE} = \left[\frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{xi} - \mathrm{yi})}{\mathrm{n}}\right] \qquad (8)$$

$$RMSE = \left[\frac{\sum_{i=1}^{n} (xi - yi)^2}{n}\right] \qquad (9)$$

Where,

x_i is estimated value, y_i is observed value and n is the number of samples.

Results and Discussion

The Results from Neural Network Model

As it can be seen in Table 2, atmospheric water vapor variable as an input for stations (Hamadan, Kermanshah, Tabriz and Bojnoord) had higher accuracy than the land surface temperature (day and night). For (Khurobiabanak and Isfahan) stations, land surface temperature variable is more accurate than atmosphere water vapor. At the above stations water vapor has the lowest effect on the radiation reaching the Earth which corresponds with physics of water vapor effects on solar radiation (increasing the amount of water vapor reduces the radiation reaching the Earth.) If the variables of land surface temperature and atmosphere water vapor be considered as input together, the amount of determination coefficient R^2 is reduced and lowers the accuracy of the model. Therefore, the neural network model provided the best results by combining extraterrestrial radiation input and atmosphere water vapor and the accuracy of model ascends by applying water vapor variable as the input. In Kermanshah station, the negative value of MBE indicates that the neural network model estimates the amount of radiation less than 5/52 MJ per square meters per day and represents low-estimation of the model and the MBE value is positive at the rest of stations and the amount of radiation is higher than estimated realistic amount and represents the over estimation of the model. Hamadan station with the RMSE of 52.6 MJ per square meter per day has the lowest error in neural network model.

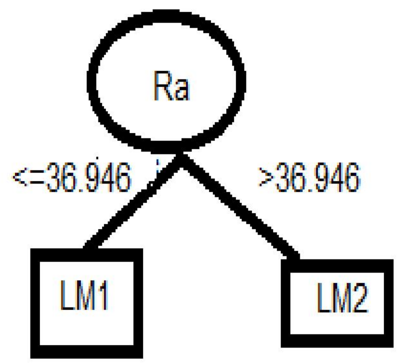
Table 2- results obtained from the neural network in different stations

Station name	The number of independent variables	The number of neurons	Extraterrestrial radiation	Land surface temperature (day)	Land surface temperature (night)	Water vapor	R^2	MBE (%)	RMSE (%)
Bojnoord	2	3	*			*	0.9	4.76	11.99
	3	7		*	*	*	0.89	9.25	14.53
	4	10	*	*	*	*	0.9	6.48	12.47
Isfahan	2	3	*			*	0.95	5.42	9.36
	3	8		*	*	*	0.95	7.89	12.03
	4	2	*	*	*	*	0.94	9.08	15.36
Hamadan	2	2	*			*	0.77	2.82	6.52
	3	4		*	*	*	0.71	3.65	7.54
	4	8	*	*	*	*	0.78	5.76	8.02
Kermanshah	2	1	*			*	0.91	-8.58	11.02
	3	6		*	*	*	0.8	-3.08	10.72
	4	1	*	*	*	*	0.83	-5.52	11.06
Khurobiabanak	2	5	*			*	0.96	3.85	6.74
	3	10		*	*	*	0.97	5.82	14.73
	4	10	*	*	*	*	0.95	7	9.25
Tabriz	2	9	*			*	0.88	5.61	11.52
	3	3		*	*	*	0.87	4.31	18.01
	4	1	*	*	*	*	0.81	10.77	14.76

Parameters which are considered as model input

The Results from M5 Model Tree

Figure 3 shows an example of M5 tree network model that is provided for the conversion of data of the land surface temperature in day and land surface temperature in night, atmosphere water vapor and extraterrestrial radiation reaching the earth. It is considered that the extraterrestrial radiation is determined as a decision variable and leads to two leaf based on two conditional propositions in the model tree.




```
(using smoothed linear models)

Ra             <= 36.946 : LM1 (89/41.094%)
Ra             >  36.946 : LM2 (96/38.481%)

LM num: 1
Rs =
        -1.3145 * pw
        + 0.5195 * Ra
        + 1.7247

LM num: 2
Rs =
        -1.329 * pw
        + 0.9379 * Ra
        - 13.6478

Number of Rules : 2
```

Figure 3. An overview of the M5 model tree fitted to the data

As it is observed, if extraterrestrial radiation is less than or equal to 36/946 watt per square meter, the linear model 1 is used, otherwise the linear model 2 is used for estimating the radiation. In this case the amount of radiation obtained from meteorological data was used as the basis and results from the model tree as the estimation values. As you can see in Table 3 , the atmosphere water vapor variable as an input had more accuracy than the land surface temperature (day and night)for stations of (Hamadan, Kermanshah, Tabriz and Bojnoord). For stations (Khurobiabanak and Isfahan) , the land surface temperature variable is more accurate than the atmosphere water vapor. It seems that the high temperature and humidity shortage in these stations have role in such a result. Because the low amount of moisture in the air causes the domination of temperature role in the amount of radiation, therefore reduces the computational error. If the variables of land surface temperature and atmosphere water vapor to be considered as input, the amount of determination coefficient R^2 reduces and lowers the model accuracy. Therefore, the combination of extraterrestrial radiation and atmosphere water vapor provided the best result in the M5 model tree and the accuracy of model increase by applying atmospheric water vapor. The Khurobiabanak station has the lowest error in the M5 model tree with the RMSE equal to 4/40 MJ per square meter per in day.

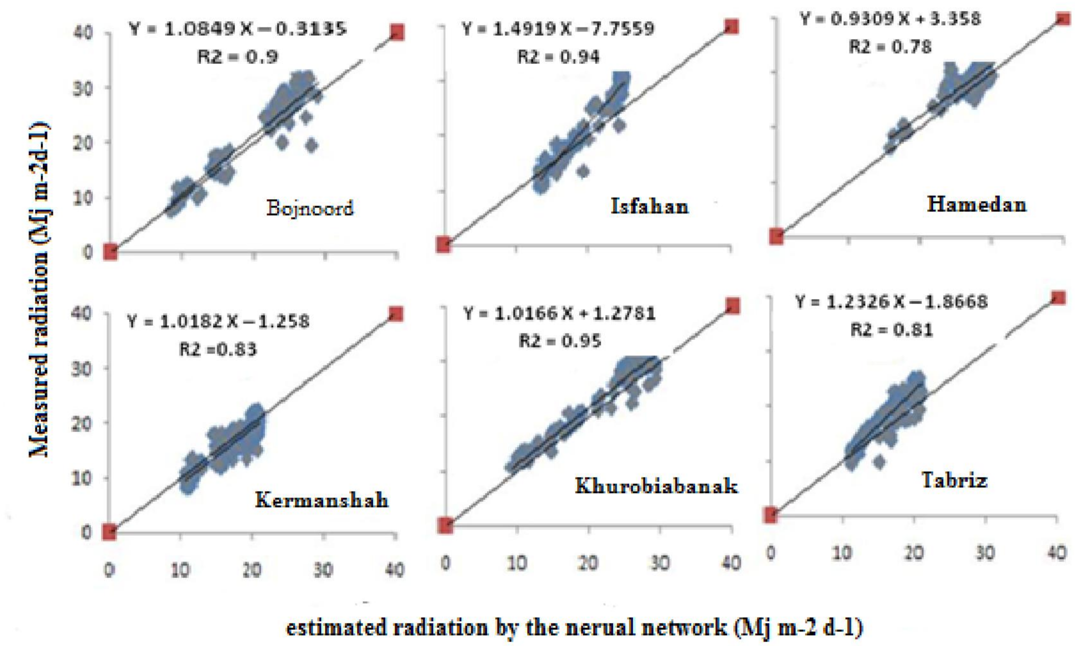
Table 3- results obtained from the M5 model tree at different stations

Station name	The number of independent variables	Extraterrestrial radiation	Land surface temperature (day)	Land surface temperature (night)	Water vapor	R^2	MBE (%)	RMSE (%)
Bojnoord	2	*			*	0.9	7.09	13.15
	3		*	*	*	0.89	7.36	13.64
	4	*	*	*	*	0.9	7.47	13.15
Isfahan	2	*			*	0.94	4.08	8.69
	3		*	*	*	0.95	3.35	7.8
	4	*	*	*	*	0.94	4.08	8.69
Hamadan	2	*			*	0.78	4.07	7.07
	3		*	*	*	0.76	5.22	7.97
	4	*	*	*	*	0.74	5.53	8.41
Kermanshah	2	*			*	0.92	-9.97	12.16
	3		*	*	*	0.9	-9.1	12.32
	4	*	*	*	*	0.9	-9.09	12
	2	*			*	0.96	5.82	7.86
Khurobiabanak	3		*	*	*	0.97	5.97	7.48
	4	*	*	*	*	0.97	6	4.04
	2	*			*	0.86	2.46	8.08
	3		*	*	*	0.8	2.8	9.93
Tabriz	4	*	*		*	0.81	2.37	9.67

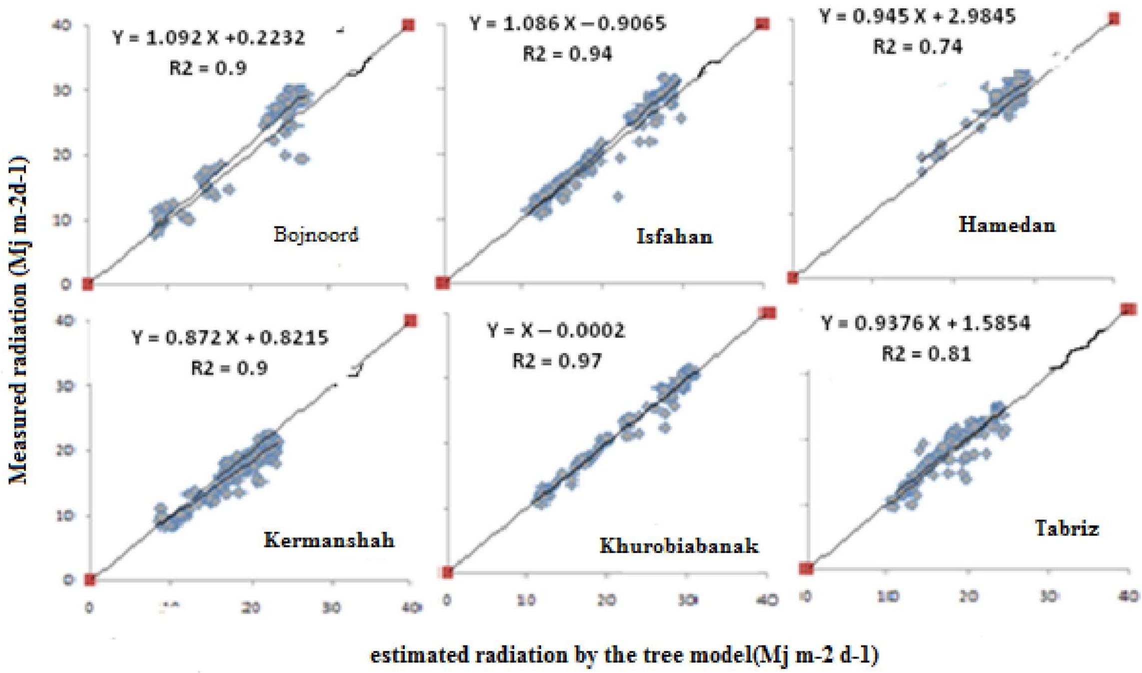
In Table 4, the values of statistical indicators R^2 , MBE and RMSE are shown at various stations that are obtained as a result of comparing the estimated values of solar radiation reaching the Earth's surface using two models presented and also the actual values from meteorological data (2003- 2005). In general, the lower the value of error detection parameters and the higher the determination coefficient, the model is more accurate and more relevant. As it can be seen in Table 4, in Isfahan and Khurobiabanak stations we can see the highest correlation with actual amounts of radiation in both models according to the value of determination coefficient. At Kermanshah station the MBE negative amount shows that the model used estimated the radiation less than actual amount and indicates that the model is under-estimation and in other stations the value of MBE is positive and the radiation is estimated more than actual amount and indicates that the model is over-estimation. In Figure 4, the correlation of these two models has been shown with the measured value of radiation. The stations of (Bojnoord, Isfahan, Kermanshah and Khurobiabanak) enjoy high dispersion coefficient and there is a high correlation between the radiations estimated by the two models with the amount of measured radiation. But Hamadan station has a low dispersion coefficient .In stations (Bojnoord, Isfahan, Hamadan, and Tabriz) correlation line stands top of the one by one line, so low-estimation is seen in these stations. In Kermanshah station the correlation line is under one by one line, so overestimation is considered. It is near to one by one line in Khurobiabanak station comparison to other stations and indicates that the dispersion of points around the correlation line is distributed more monotonous and compressed and correlation line is in accordance with one by one line, therefore it is very accurate.

Table 4- results obtained from artificial neural network model and model tree in different stations

Station name	Artificial Neural Network			Model Tree		
	R^2	MBE (%)	RMSE (%)	R^2	MBE (%)	RMSE (%)
	0.9	4.76	11.99	0.9	7.09	13.15
Bojnoord	0.89	9.25	14.53	0.89	7.36	13.14
	0.9	6.48	12.47	0.9	7.47	13.15
	0.95	5.42	9.63	0.94	4.08	8.69
Isfahan	0.95	7.89	12.03	0.95	3.35	7.8
	0.94	9.08	15.36	0.94	4.08	8.69
	0.77	2.82	6.52	0.71	4.07	7.07
Hamadan	0.71	3.65	7.54	0.76	5.22	7.97
	0.78	5.76	8.02	0.74	5.53	8.41
	0.91	-8.58	11.02	0.92	-9.97	12.16
Kermanshah	0.8	-3.08	10.72	0.9	-9.1	12.32
	0.83	-5.52	11.06	0.9	-9.09	12
	0.96	3.85	6.74	0.96	5.82	7.86
Khurobiabanak	0.97	5.82	14.73	0.97	5.97	7.48
	0.95	7	9.25	0.97	6	4.04
	0.88	5.61	11.52	0.88	2.46	8.08
Tabriz	0.87	14.3	18.01	0.8	2.8	9.93
	0.81	10.77	14.76	0.81	2.37	9.67



(A): The estimated radiation by neural network model with four input variable



(B)The estimated radiation by the model tree with four input variables

Figure* 4. The dispersion of estimated radiation with four input variables (day land surface temperature, night land surface temperature, atmosphere water vapor and extraterrestrial radiation) versus measured radiation of Earth *(a) Neural Network Model and (b) the Model Tree

Results

In this study, the neural network model and the M5 tree model are used to estimate the amount of radiation reaching the Earth based on the land surface temperature and atmosphere water vapor of MODIS images. The results showed that the satellite images have better application for regional studies against meteorological stations data due to the spatial coherence of data. Also the neural network model and M5 model tree comparison showed that there is not much difference between the two models and both models estimate the radiation reaching the Earth with high accuracy. On the other hand, the best combination of input variables (day and night land surface temperature, atmospheric water vapor and extraterrestrial radiation) is the combination of extraterrestrial radiation with atmosphere water vapor, in other words, by applying water vapor variable as an input both models̉ accuracy increase. The amount of water vapor is low in dry areas and water vapor has reducing effect on the amount of radiation reaching the ground, so the temperature has more effect on the radiation towards the water vapor in arid regions. If the variables of land surface temperature and atmosphere water vapor are considered together as input, the R^2 value dropped down and decreases the accuracy of the model.

References

[1]Emamifar, S., Rahimikhoob, A . 2011.The evaluation of M5 model tree and Angstrom empirical model for estimating radiation reaching the Earth's surface , the first National Conference on Meteorology and Agricultural Water Management, 10 -1.

[2]Emamifar, S., Rahimikhoob, A ., Noroozi, A. A. 2014. The evaluation of M5 model tree and artificial neural network to estimate the average daily air temperature based on the data of the Earth's surface temperature of MODIS, Iran Soil and Water Research, 45 (4): 423-433.

[3]Emamifar, S., Alizadeh, A.. 2015. Estimating the amount of solar radiation using the products of MODIS land surface temperature and neural network model, water and soil Journal, 28 (3): 617-625.

[4]Bakhshoudeh M., Rahimikhoob, A. 2014. The comparison of regression tree model, neural networks and Hargreaves-Samani in estimating evapotranspiration, reference in arid areas., Journal of Irrigation and Water Management, 4 (2): 149-160.

[5]Bayat, K., Mirlatify S. M. 2009.The estimation of total daily solar radiation using regression and artificial neural networks models, Journal of Science of Agricultural and Natural Resources, 16 (3): 1. 11.

[6]Rahimikhoob, A., Behbahani, M., Jamshidi, M. 2009.The assessment of experimental method and neural network models to estimate solar radiation received by Earth, Journal of Science and Technology of Agriculture and Natural Resources, 13 (50): 53-62.

[7]Rahimikhoob, A., Behbahani, S.M.R. Nazarifar,M.H.2010. The prediction of maximum temperature of Khuzestan according to NOAA satellite data and artificial neural network model , Journal of Science and Technology of Agriculture and Natural Resources, 11 (42): 357- 364.

[8]Rahimikhoob, A., Saberi., Behbahani, S.M.R. Nazarifar,M.H.2011.Estimating the amount of solar radiation reaching to the ground using NOAA satellite images and statistical relationships in South East of Tehran, Journal of Science and Technology of Agriculture and Natural Resources, Soil and Water Sciences, 15 (56): 79-88. .

[9]Saberi, P., Rahimikhoob, A.S., Saberi, F., Zebardast, S. 2010.The estimation of solar radiation received by the Earth using AVHRR images of NOAA satellite (Case Study : Kermanshah), Third National Conference on the Management of Irrigation and Drainage Networks, 1-9.

[10]Allen, RG1996. Assessing integrity of weather data for reference evapotranspiration estimation. J. Irrig. Drain. Eng. ASCE.:122.97-106.

[11]Allen, RG, Pereira, LS Raes, D. Smith, M.1998. Crop evapotranspiration guidelines for computing cropwater requirements. Irrigation and Drainage Paper No. 56. FAO, Rome.

[12]Beyer, HG, Costanzo, C. Heinemann, D. 1996. Modifications of the heliosat procedure for irradiance estimates from satellite images. Solar Energy. 56 (3): 207-212.

[13]Coll C., Caselles V., Sobrino J.A and Valor E. 1994.On the atmospheric dependence of the split-window equation for land surface temperature. International Journal of Remote Sensing; 15: 102-105

[14]Emamifar, S. Rahimikhoob, A. Noroozi, A.A 2013. Daily mean air temperature estimation from MODIS land surface temperature products based on M5 model tree. International Journal of Climatology, 33 (15): 3174-3181.

[15]Gautier, C., Diek, G., Masse, S. 1980. A simple physical model to estimate incident solar radiation at the surface from Goes satellite data. J. Apple. Meteor.19: 1005-1012.

[16]Hornik, K., Stinchcombe, M., White, H. 1989. Multilayer feed forward networks are universal approximators. Neural Netw. 2: 359-366.

[17]Ozan S¸ Enkal .2010. Modeling of solar radiation using remote sensing and artificial neural network in Turkey . Energy. 35: 4795-4801.

[18]Pal, M., Deswal, S. 2009. M5 model tree based on modeling reference evapotranspiration. Hydrol. Process. 23: 1437-1443.

[19]Qin, J., Zhuoqi, C., Yang, k., Liang, SH ., Tang, W. 2011. Estimate of monthly mean daily global solar radiation based on Modis. Appl. energy, 88 (7): 2480-2489.

[20]Quinlan, JR 1986. Induction of decision trees. Machine learning, 1 (1): 181-186.

[21]Quinlan, JR 1992.Learning with continuous classes. Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence, Hobart, Australia, 16-18November, World Scientific, Singapore, pp. 343-348

[22]Samani, Z. 2000. Estimating solar radiation and evapotranspiration using minimum climatological data. J.Irrig. Drain. Eng. ASCE, 126 (4): 265-267.

[23]Trnka, M., Zalud, Z., Eitzinger, J. and Dubrovsky, M., 2005, Global solar radiation in Central European lowlands estimated by various empirical formulas. Agric. and Forest Meteorol. 131 (1-2): 45-76.

[24]Wan, ZM, Zhang, YL, Zhang, QC, and Li, ZL 2002. Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectro-radiometer data. Remote Sensing of Environment, 83 (1): 163-180.

4/21/2021