Evaluation of M5 Tree Model Using Water Vapor to Estimate the Solar Radiation

Narges Kefayati[1], Ali rahimi khob[2], Aliakbar Noroozi[3]

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

Abstract: In this study, M5 tree model with input data of water vapor product MOD05_L2 MODIS were used to estimate the solar radiation. To determine the radiation, the six stations with different climates for a period of three-year were used as the actual data. The observed radiation was obtained by the Meteorology Organization and extraterrestrial radiation obtained mathematically. Water vapor was received from MODIS images of Terra satellite. The available images changed into input parameters using GIS and Envi softwares. The results showed that the tree model M5, Isfahan stations with the value of R^2, 0.94 has the highest accuracy in estimating solar radiation.
[Narges Kefayati, Ali rahimi khob, Aliakbar Noroozi. **Evaluation of M5 Tree Model Using Water Vapor to Estimate the Solar Radiation.** *World Rural Observ* 2016;8(3):43-47]. ISSN: 1944-6543 (Print); ISSN: 1944-6551 (Online). http://www.sciencepub.net/rural. 6. doi:10.7537/marswro080316.06.

Key words: Solar radiation, Water vapor, MODIS sensor, Tree model M5, Isfahan station

1. Introduction

The solar radiation received by Earth is an essential part in evapotranspiration models and hydrologic models that controls the exchange of water and heat (Keen et al, 2011). The part of the solar radiation that reaches the Earth's surface without scattering or absorption is called direct radiation and the part of the solar radiation that is absorbed or scattered by atmospheric components is called scattered radiation. The direct radiation from the sun comes in the form of a straight line, while the scatter edradiation enters in all directions from the sky. The total radiation reaching the Earth's surface, on a cloudy day is in the form of scattered radiation (Moradi, 2005). Normally in clear and cloudless weather radiation reaching the Earth is 75 percent of radiation reaching the upper atmosphere(Ra) and in complete cloudy weather it is about 25% Ra (Allen, 1998). In general, two simulation and experimental methods are provided to determine the attenuation effects of atmosphere and radiation reaching the Earth's surface. In simulation method ,the attenuation effect of atmosphere is analyzed physically from upper level to the Earth's surface and the radiation reaching the Earth is determined by solving it. This method is accurate but its implementation requires different layers of atmosphere data that are prepared using thrownradio sounds into the sky which their number is very limited. Using coefficient on extraterrestrial radiation of the atmosphere is a method that is used in experimental models (Trenka, 2005).

Despite the importance of this parameter, its measurement is done in a limited way and this deficiencycan be seen not only in developing countries but also in developed countries (Samani, 2000).In addition to being expensive, this parameter's measurement device must be calibrated regularly due to its high sensitivity and most weather stations lack it and its data are limited (Rahimi Khoub et al., 2011). For this reason, many researchers have presented the estimation techniques of this parameter. Because these methods use the type of point spatial of meteorological data, thus the amount of approximate radiation by them would also be point (Emamifar et al., 2013). In recent years, the models of remote sensing are used to determine some of the phenomena which its input data are obtained from satellite imagery. The advantage of satellite imagery is that it covers vast and widespread levels and spatial changes of phenomena can be studied in it (Rahimi Khoub et al., 2011). In a research, to estimate the radiation reaching the Earth in South East of Tehran, the statistical method of remote sensing was preceded. In this study, data from NOAA satellite images AVHRR sensor were used. This method is based on the linear correlation between the estimated cloud index from satellite images and measured air clearness index in meteorological station. Moreover, a multiple variable correlation model was developed to convert five bands of NOAA satellite data and extraterrestrial radiation to radiation reaching the Earth and was compared by statistical methods. The results showed that multiple variable correlation model with an R^2 determination coefficient and root mean square error RMSE is equal to 0.93and 5.8 % respectively and estimates the radiation reaching the earth more accurately compared to statistical method (Rahimi Khoub et al., 1390).

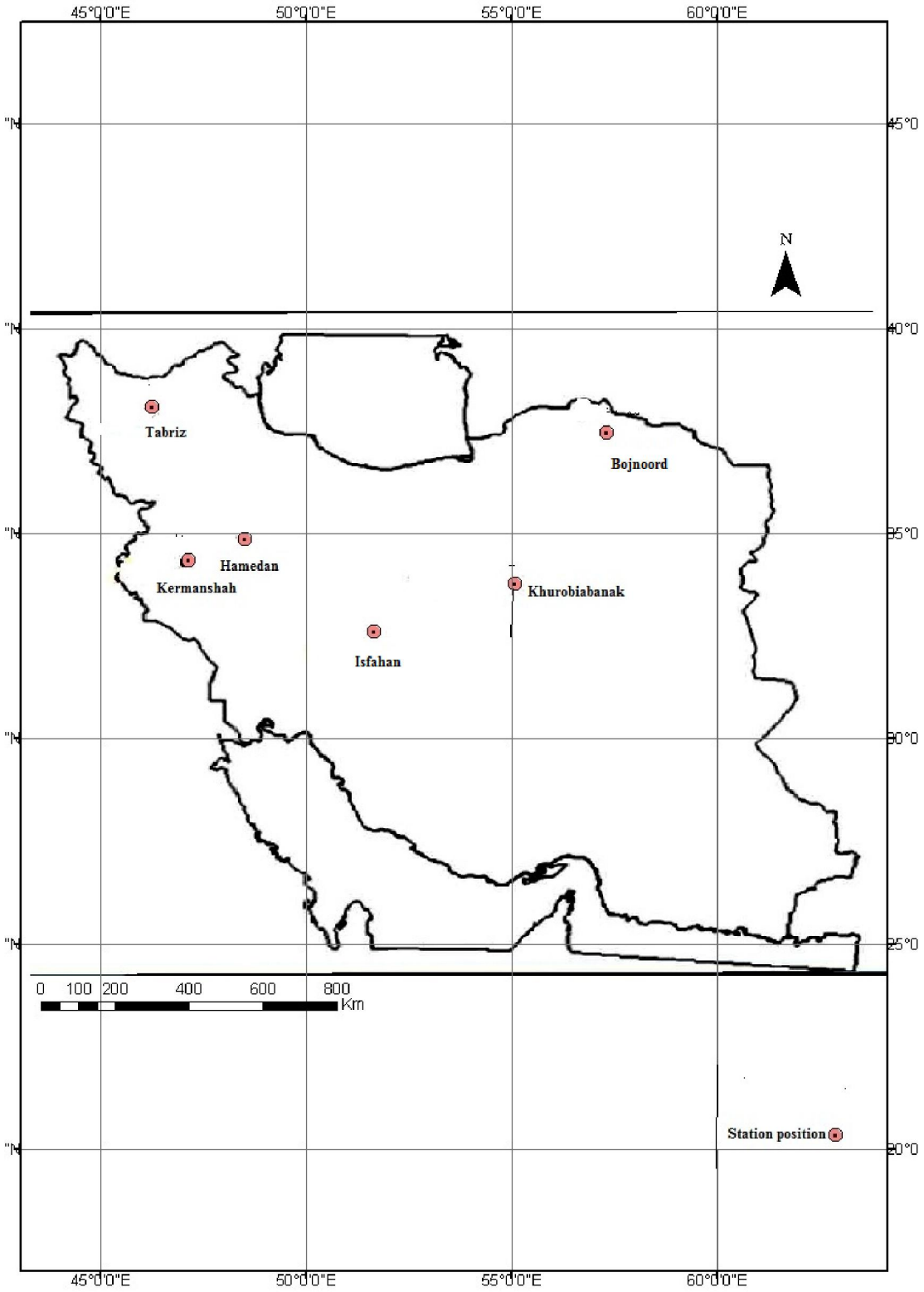

Fig 1. Geographical situation of studied stations

Table 1. Specifications of studied synoptic stations

Station	Climate type (Emberger classification)	Year	Latitude (Degree minute)	Longitude (Degree-minute)	Altitude from sea level (M)
Hamedan	Cold Semi-Dry	1976	34 52	48 32	1741.5
Kermanshah	Moderate	1951	34 21	47 09	1318.6
Tabriz	Dry cold	1951	38 05	46 17	1361
Isfahan	Dry	1951	32 27	51 40	1550.4
Khurobiabanak	Very dry	1986	33 47	55 05	845
Bojnoord	Semi-dry	1977	37 28	57 19	1091

The smart models are used as new methods for modeling of complex relationships in recent years. Among these models are networks of tree decision making. The new generation decision making trees are data mining and are used for knowledge extraction and discovery. In a study, tree model M5was used to change the Earth's surface temperature data and extraterrestrial radiation to reference evapotranspiration in Qazvin irrigation network. In this study, the separate outlet algorithm based on data from NOAA satellite AVHRR sensor was used to estimate the Earth's surface temperature. The results showed that the developed model based on dedicated data in the test phase can estimate the amount of ETO with the determination coefficient, percent of root mean square error and percentage of mean deviation error0.81, 8.5% and 2.5%respectively (Moradi and Rahimi Khoub, 2012). The aim of this study was to evaluate M5 tree model for estimating radiation reaching the Earth using water vapor of atmosphere.

2. Materials and Methods

For doing this study, the solar radiation data of six synoptic stations in Iran was used for a three-year period (2003 - 2005). Areas of study include: Tabriz, Hamadan, Kermanshah, Isfahan, Khurobiabanak and Bojnoord. The areas have different climatic and geographical situation. The specifications of stations studied are presented in Table and Fig 1.

Satellite Data

In this research, MODIS sensor data was used that are among Terra satellite sensors. MODIS accomplishes some daily observations on sea, land and atmosphere on a global scale and has a continuous and broad spectral and spatial covering with a resolution of 250, 500 and 1000 meters. MODIS data includes day and night data, which has a good resolution (Wen et al., 2002). The images used in this study were a subgroup of level two MODIS (L2) with a code characteristic MOD05_L2 related to the Terra satellite. Its pixels sizes are 1 × 1 km and 5 × 5 km. valid span for water vapor data from the MODIS is (0- 20) cm. To receive images of atmospheric water vapor column ,MODIS uses two algorithms of infrared and near-infrared. Infrared algorithm just takes images in the day and taken images are clearer and more vivid and near-infrared algorithm also has a day and night imaging. MODIS images were collected via the internet site http://modis.gsfc.nasa.gov for years (2003- 2005). system change of images coordinates from sinusoidal to UTM was conducted using the MRT (Modis Reprojection Tool) and data extraction of atmosphere water vapor at six stations was performed in the software ARCGIS 9.3 using the HAT tool (Hawth Analysis Tools).

M5 Tree Model

M5 tree model is based on tree classification method that was presented by Quinlan (1986) to create a relationship between independent and dependent variables. This model is a combination of linear and tree regression models and can be used for any type of quantitative and qualitative data (Quinlan, 1992).

Every decision-making tree has a structure similar to tree that is composed of roots, branches, nodes and leaves and is drawn evenly from top to bottom. In this structure the root is located at the top as first node and a chain of branches and nodes ends to leaves. Split in nodes is conducted by branches and each node represents a predictor variable.

In this study Weka software was used which is among the most reliable and widely used applications available in the field of data mining and especially in the field of producing tree models that is written and developed at the University of Waikato New Zealand. This comprehensive software includes solutions for all data mining problems such as regression, classification, clustering, and law propriety and traits selection.

The easiest way to use Weka software is through the graphical interfaces, which is called Explorer. To implement M5 Tree Model, only the first two parts is needed. In the pre-processing section, data are called from a file to csv format and, in the second part of the Explorer unit which is called classification; categorization was applied on pre-processing data. In this section, the user can evaluate the models obtained. To compile M5 Tree Model the data were divided into two groups. Data of 2003 and 2004 were used to

training and data of 2005 were used for testing. In the training phase, input parameters including (atmosphere water vapor and extraterrestrial radiation) for all six stations in 2003 and 2004 were changed into csv format and took to M5 tree model. M5 tree model gives an equation in terms of input parameters that its equation was used to test the data for 2005 in each station. It is worth noting that the amount of extraterrestrial radiation cannot be measured and isdependent on latitude, day of the year and the time of satellite passing. In this study, Allen and colleagues suggested relations were used to determine it (Allen et al., 1998). To check the results and quantitative comparison of them needs to choose appropriate indicators so as to analyze the factors affecting the results. In this study, three statistical indicators of determination coefficient R^2, MBE average deviation error and root mean error RMSE was were used.

3. Results and discussions

In Table 2, the results from the tree model are presented. As it can be seen in Table 2, the Isfahan station has the high coefficient of determination R^2.In this station water vapor has the least effect on the radiation reaching the Earth which corresponds with the physic of water vapor effects on solar radiation (increasing the water vapor decreases the amount of radiation reaching the earth). In Kermanshah station negative value of MBE indicates that the tree model has estimated the amount of radiation less than 1.68 MJ per square meter per day and reflects the underestimation of the model and in the rest of the stations the MBE value is positive and the amount of radiation is estimated higher than actual amount and represents that the model has over estimated. Tabriz station with RMSE value equal to 1.62 MJ per square meter per day has the least errors in the tree model.

Table 2. Results from tree model in different stations

Stations	R^2	RMSE (Mj m-2 d-1)	MBE (%)
Bojnoord	0.9	2.85	1.53
Isfahan	0.94	1.89	0.88
Hamedan	0.76	2.05	1.21
Kermanshah	0.92	2.05	-1.68
Khurobiabanak	0.92	1.83	1.36
Tabriz	0.88	1.62	0.57

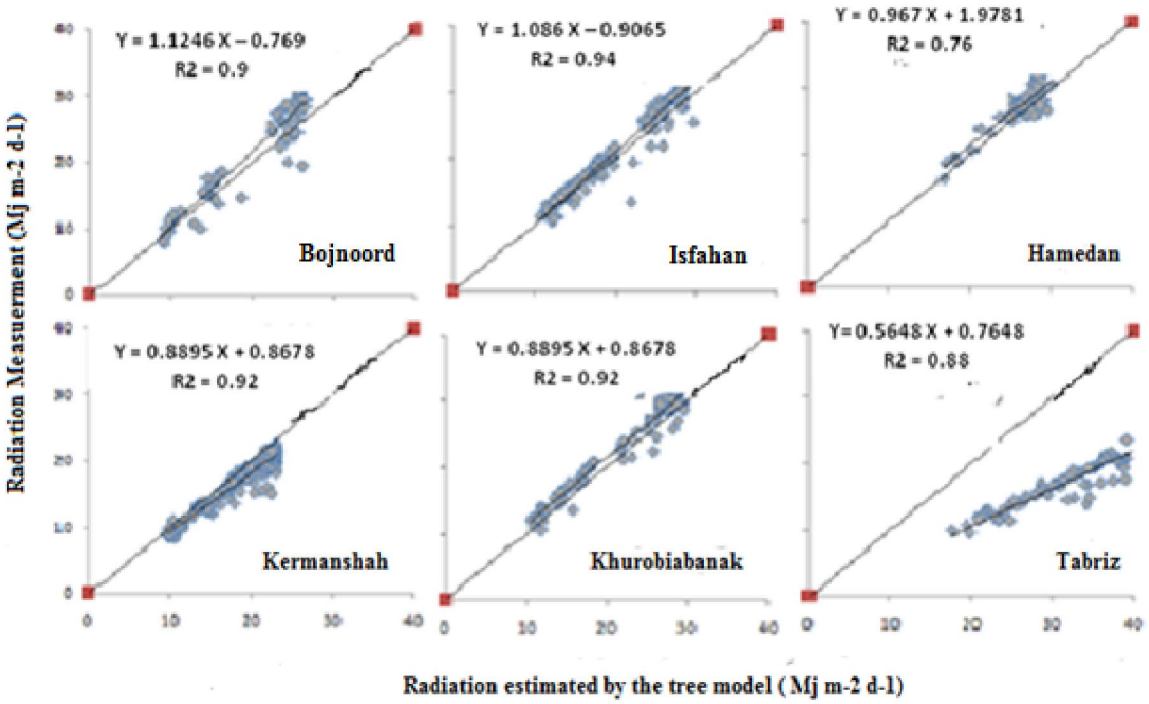

Figure 2. Dispersion of estimated radiation (atmospheric water vapor and extraterrestrial radiation) in comparison to measured radiation of Earth using the Tree Model

In Figure 2, the amount of correlation with the amount of measured radiation is shown in the tree model. Isfahan station has a high dispersion coefficient and there is a high correlation between the estimated radiations by model with the amount of measured radiation. But low dispersion coefficient is seen in Hamadan station. As can be seen, in Isfahan station compared to other stations, it is closer to the one by one line and it shows that the dispersion of points around the cohesion line has been distributed more compact and uniform. In stations (Bojnoord, Khurobiabanak and Hamadan) most of the points are over the one by one line, and the correlation line was over of the one by one line, so in named stations the radiation of sun is underestimated. In Tabriz station correlation line is located below the one by one line and is overestimated.

4. Conclusions

In this study, to estimate the amount of radiation reaching the earth, M5 tree model based on MODIS water vapor imagery was used. Using data from satellite images instead of ground parameters improves results, because satellite data are more accurate than ground-based measurements data. The results showed that images of Terra satellite, MODIS has adesirable ability to calculate different parameters and requires no calibration. The amount of water vapor in dry areas is low and water vapor has decreasing effect on the amount of radiation reaching the Earth. The results showed that the tree model has a high accuracy in estimating solar radiation using water vapor data products MODIS MOD05_L2 as an input. On the other hand ,the accuracy of radiation prediction in Isfahan station in the tree modelis better than five other stations.

References

1. Imamifar, S., Noroozi,. A. A, Hosseini, S. Vkarimzadeh, A.,2013, Evaluation ofTree Model and two empirical models based on air temperature to estimate solar radiation using LST in a semi-arid climate. Journal of Soil and Water Resources Protection 0.3 (1): 1-12.
2. Rahimi Khoub, A. Saberi, P. Behbahani, S. M. R., and Nazarifar, M. H.,2011. Estimating the solar radiation received by Earth using NOAA satellite images and statistical relationships in the South East of Tehran. Journal of Science and Technology of Agriculture and Natural Resources, Soil and Water Sciences, 5 (56): 79-88.
3. Moradi, S. 2005. Settingthe environmental conditions, Tehran, Ashian publications, 79- 104.
4. Moradi, M. and Rahimi Khoub, A., 2012. Estimation of evapotranspiration using NOAA Satellite images and tree model for irrigation networks. A Case Study of Qazvin irrigation network. Journal of Science and Technology ofAgriculture and Natural Resources, Soil and Water sciences, 4 (62): 123 136.
5. Allen, R. G., Pereira, L. S. D. Raes, and Smith M., 1998, *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage* Paper. No. 56, FAO, Rome.
6. Qin, J., Chen, Z., Yang, K., Liang, S. and Tang, W., 2011, *Estimation of monthly-mean daily global solar radiation based on MODIS and TRMM products*. Appl. Energy, 88(7):2480-2489.
7. Quinlan, J. R., 1986*, Introduction of decision trees*. Machine Learning 1: 81-106.
8. Quinlan, J. R., 1992, *Learning with continuous classes. Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence*, Hobart, Australia, 16-18 November, World Scientific, Singapore, pp. 343–348.
9. Samani, Z., 2000, *Estimation solar radiation and evapotranspiration using minimum climatological data*. J. Irrig. Drain. Eng., 126(4):265-267.
10. 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.

9/6/2016