

NEURAL NETWORK METHOD FOR DROUGHT MODELING USING SATELLITE DATA

Ramin Mokhtari^a, Mehdi Akhoondzadeh^{b,*}

^a M.Sc. Student, Remote Sensing Department, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, North Amirabad Ave., Tehran, Iran-ramin.mokhtari.d@ut.ac.ir

^b Remote Sensing Department, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, North Amirabad Ave., Tehran, Iran-makhonz@ut.ac.ir

KEY WORDS: Drought Model, Neural Network, MODIS, SMOS, TRMM, NDVI

ABSTRACT:

Drought is one of the natural crises in each region. Drought has a direct relationship with vegetation. Various factors affect vegetation. The relationship between these factors and vegetation can be expressed using methods of machine learning algorithms. Nowadays, using remote sensing images can be used to measure the factors affecting vegetation and investigate this phenomenon with high precision. In this research, vegetation and various factors affecting this factor, which can be measured using satellite imagery, are selected. The factors include land surface temperature (LST), evapotranspiration (ET), snow cover, rainfall, soil moisture that which are derived from the active and passive sensors of satellite sensors as the products of land surface temperature (LST), snow cover and vegetation, using images of products of the MODIS sensor and rainfall using the images of the TRMM satellite and soil moisture using the images of the SMOS satellite during a period from June 2010 to the end of 2018 for the central region of Iran has received and after that, primary processing was performed on these images. The vegetation index (NDVI) is modeled using artificial neural network algorithm for monthly periods. method have been able to achieve model with desirable accuracy. The average accuracy was RMSE=0.048 and R2=0.867.

1. INTRODUCTION

Droughts can be divided into agricultural drought, social drought, meteorological drought, and hydrological drought (Wilhite and Buchanan, 2005). Using remote sensing we can have newer ways to monitor drought. One of the most important advantages of remote sensing data over meteorological data is that the meteorological data is harvested on a point-by-point basis while the remote sensing data is harvested on a continuously. Also other advantages of this type of data can be acknowledged the spatial resolution, temporal resolution and availability of this type of information (Heumann, 2011). Agricultural drought is related to vegetation and occurs when soil moisture levels are lower than the amount of water needed for plant growth and health and vegetation is weaker than previous periods in the area. One of the most important parameters of agricultural drought is vegetation, which is one of the most important indicators to measure this factor is the Normalized Vegetation Difference Index (NDVI). Numerous studies have been conducted on a variety of drought indices using satellite data, including vegetation and thermal data in various regions of the world, but there are still challenges in increasing the accuracy of better prediction of this phenomenon. This phenomenon is more nonlinear while most studies use linear models (Bai et al., 2018). Various factors affect the drought, which can often be obtained by using remote sensing, such as rainfall, snow cover, land surface temperature, soil moisture, and vegetation cover (AghaKouchak et al., 2015). vegetation cover is very important in the production of energy and food sources. vegetation is affiliated and influenced by various factors such as natural factors such as seasonal changes, temperature, soil moisture, and also human factors, which ultimately can change ecosystems. The Normalized Vegetation Difference Index (NDVI) is based on the rule that the existing

structure of plants can absorb red light and reflect infrared light (Pettorelli et al., 2005). Seonyoung Park et al in 2017, using MODIS data including land surface temperature, NDVI, Evapotranspiration (ET) and precipitation rates using satellite data TRMM from 2000 to 2012 in the United States of America has been modeling the Standardized Precipitation Index (SPI) whit using three methods of machine learning such as Random Forest (RF), Boosted Regression Trees (BRT) and Cubist algorithm, and the results show that the method Random forest has better modeling than two other methods (Park et al., 2017). Alizade et al 2018, using data such as GPCP, CMAP, CHOMPS and TRMM, all relate to rainfall data, in order to model the SPI index from multilayer perceptron algorithms (MLP), Adjustable Neural Fuzzy Inference (ANFIS), Support Vector Regression (SVR), and finally evaluated with SPI obtained with meteorological data (Alizadeh and Nikoo, 2018). Barua et al, 2012, at first, it developed a time series of Nonlinear Aggregated Drought Index (NADI) using precipitation data in Australian weather stations, and then by using recursive multistep neural network (RMNN) and Direct Multistep Neural Network (DMNN) forecast this index for up to six months. The results have shown that both methods have the most accurate and similar predictions for the monthly. But for the prediction of the two-month and three-month periods, the RMNN method performs better than the DMNN method (Barua et al, 2012). Belayneh, 2013, at first, it obtained standardized precipitation index (SPI) using precipitation data at meteorological stations and then by using three machine learning methods including artificial neural network (ANN), support regression vector (SVR) and wavelet neural network (WA_ANN) predicted this time series over three and six months (SPI) periods, and the results showed that the wavelet neural network method was more accurate than the other two methods (Belayneh and Adamowski, 2013).

In most studies, used the low number of sensor and satellite products, while in the drought, various factors can influence, including precipitation, soil moisture, evapotranspiration, snow and land surface temperature, which today is accessible globally through satellite imagery. Nevertheless, we need to be able to involve more influential factors in drought modeling so that it can accurately model and predict this phenomenon. therefore, in this study, data from SMOS satellite soil moisture data, TRMM satellite rainfall data and vegetation cover, land surface temperature, snow cover of MODIS sensor and using Artificial Neural Network (ANN) methods were used to model the NDVI index has been executed on a monthly basis from June 2010 to the end of 2018.

2. STUDY AREA AND DATASETS

2.1 Study site

Esfahan province as an agricultural hub is located between 49° and 55° east longitude in center part of Iran. Fig. 1 shows the location of Esfahan province. Esfahan province has a total area of 107,029 Km² (almost 9% of the country's area) with about three distinct climatic regions including the mountainous area of the west and southwest with moderate winters and summers, the central regions with hot dry summers and the area located in the north and southeast with cold winters and hot summers.



Figure 1. The location of Esfahan province

2.2 datasets

In this study, the SMOS satellite soil moisture product, the TRMM satellite rainfall and MODIS sensor products have been used such as eight-day products of land surface temperature (MOD11A2), eight-day products of evapotranspiration (MOD16A2), the monthly product of snow cover (MOD10CM), and the monthly vegetation product (MOD13A3), using images of products of the MODIS sensor and rainfall, using monthly images (3B43) of the TRMM satellite and soil moisture content using monthly images of the satellite SMOS during the period from June 2010 to the end of 2018 for the region of Iran has received and after that, primary processing was performed on these images. The MODIS instrument is operating on both the Terra and Aqua spacecraft. It has a viewing swath width of 2,330 km and views the entire surface of the Earth every one to two days. Its detectors measure 36 spectral bands between 0.405 and 14.385 μm , and it acquires data at three spatial resolutions 250m, 500m, and 1,000m.

The (MOD11A2) Version 6 product provides an average 8-day per-pixel Land Surface Temperature and Emissivity with a 1 kilometer (km) spatial resolution in a 1,200 by 1,200 km grid. The MOD16A2 Version 6 Evapotranspiration/Latent Heat Flux

product is an 8-day composite dataset produced at 500 meter (m) pixel resolution. The (MOD10CM) provides an average monthly per-pixel average snow cover with a .05 degree. The (MOD13A3) Version 6 product provides an average monthly per-pixel Vegetation Indices with a 1 kilometer (km) spatial resolution in a 1,200 by 1,200 km grid.

Soil Moisture and Ocean Salinity, or SMOS, is a satellite which forms part of ESA's Living Planet Programme. It is intended to provide new insights into Earth's water cycle and climate. In addition, it is intended to provide improved weather forecasting and monitoring of snow and ice accumulation.

The Tropical Rainfall Measuring Mission (TRMM) was a joint space mission between NASA and the Japan Aerospace Exploration Agency designed to monitor and study tropical rainfall. The term refers to both the mission itself and the satellite that the mission used to collect data. The TRMM 3B43 data were obtained from the Goddard Earth Sciences Data and Information Service Center (GES DISC), USA; it combines three-hour TRMM 3B42 data, the Climate Anomaly Monitoring System (CAMS)'s global gridded precipitation, and Global Precipitation Climatology Center (GPCC)'s global gridded observed precipitation (Rhee and Im, 2017).

3. METHODOLOGY

3.1 Artificial Neural Networks

Artificial neural networks, behave similar to the behavior of natural neural networks. Artificial neural networks can be trained using a series of empirical examples and used for new data (Mas & Flores, 2008). Artificial neural networks can be classified according to the network structure for example, multi-layered perceptron neural networks (MLP-NNs) are one of the most used ANN. In these types of networks, the neurons are placed in a layer, and the signals flow from an input layer to an output layer along a network of communication lines (Dawson & Wilby, 2001). The ANN models used in this study have a feed forward Multi-layer perceptron (MLP). MLPs have often been MLPs consist of an input layer, one or more hidden layers, and an output layer. A usual three-layered neural network is composed of several elements namely nodes. These networks are made up of an input layer consisting of nodes representing different input variables, the hidden layer consisting of many hidden nodes, and an output layer consisting of output variables. The ANN models used in this study have a feed forward Multi-layer perceptron (MLP) architecture. MLPs have often been used in drought forecasting due to their simplicity. MLPs consist of an input layer, one or more hidden layers, and an output layer (Kim and Valdes, 2003). For ANN model development, the determination of the architecture of the model is very important this includes the number of neurons in the input layer, the number of neurons in the hidden layer, and the number of layers in the output layer. The number of neurons that provided the lowest RMSE values was chosen as the appropriate number. Generally, the number of hidden neurons for ANN models is selected via a trial and error method.

Figure. 2 shows the structure of the multilayer feed-forward back propagation algorithm.

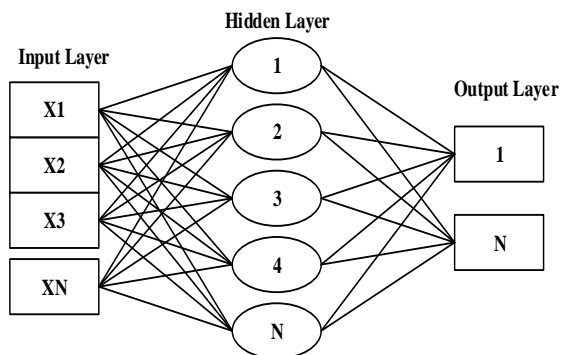


Figure.2 Structure of multilayer feed–forward back propagation algorithm.

4. IMPLEMENTATION

After performing the necessary pre-processing, the difference between land surface temperature data of the day and night are obtained from the land surface temperature images of the day and night.

In this research, in order to model the NDVI index, the first, time series of products of LST, rainfall, snow cover, ET, soil moisture and NDVI index were introduced from June 2010 to December 2018 and then, Using the neural network algorithm(MLP_NNs), vegetation modeling in the area have produce. Seventy percent of monthly data Monthly data from June 2010 to the end of 2017 as the training data and thirty percent data Monthly data from June 2010 to the end of 2017 as the test and finally after network training in the best way, the NDVI index has been modeled for 12 months of the year.

Performance Evaluation Criteria: Root mean square error (RMSE) and coefficient of determination (R2) were used as bench marks for the performance assessment of model (ANN). In this study, several different network structures have been tested. The best number of neurons in the hidden layer, that is, was determined 7 neurons. So the best network was (6 7 1) also the transfer function in the hidden layer was Tansig. Table 1. shows the values of RMSE and R2 for several different network architectures, with the results showing that the best number of neurons is for the hidden layer (7) and the best number of neurons for the input layer (6). Therefore, by increasing or decreasing the number of neurons of these two values in these two layers the accuracy decreased.

ANN models	RMSE	R2
6-5-1	0.035	0.85
6-6-1	0.032	0.89
6-7-1	0.028	0.92
6-8-1	0.033	0.90
5-7-1	0.041	0.81
7-7-1	0.037	0.83
8-7-1	0.032	0.88

Table 1. The accuracy for several different network architectures.

5. RESULTS AND DISCUSSION

First, after network training and network evaluation using existing data from year 2010 to year 2017, vegetation index modeling for the area concerned as well as twelve months from year 2018 has been discussed.

After creating the vegetation index (NDVI) model in the 12 months of 2018 in order to evaluate the accuracy using the actual vegetation index data measured by the MODIS Sensor. In order to evaluate the accuracy of the models developed for NDVI index in the 12 months of 2018, two methods including Root Mean Square Error (RMSE) and coefficient and coefficient of determination(R2) have been used.

The results show that the modeling has been done with a desirable accuracy, so that the highest accuracy is about July with accuracy of (RMSE=0.029, R2=0.921) and the average accuracy of the months in 2018 is (RMSE=0.042, R2=0.848).

Table 1 shows the accuracy for 12 months from 2018. According to Table 1 for each month of the year, a precision value is obtained and there is one evaluation for almost every season of the year. The least accuracy is about July with accuracy of (RMSE=0.051, R2=0.793) that this is probably due to the significant fluctuation of the data around the mean. It can be said that the accuracy in both summer and winter seasons is lower than in autumn and spring seasons.

Figure 1 shows the NDVI index in July 2018 and Figure 2 shows the NDVI index modeled by ANN in July 2018. Figure 5. Scatter plots for Observed NDVI and predicted NDVI.

Month	RMSE	R2	Month	RMSE	R2
January	0.033	0.874	July	0.029	0.921
February	0.037	0.869	August	0.031	0.902
March	0.040	0.843	September	0.032	0.897
April	0.053	0.776	October	0.049	0.794
May	0.051	0.793	November	0.045	0.826
June	0.037	0.871	December	0.046	0.813

Table 2. The accuracy for 12 months from 2018.

6. CONCLUSION

drought is one of the natural disasters that can have destructive and negative effects, including loss of vegetation, soil erosion and food shortages on an area. With the advancement of remote sensing methods, many of the parameters affecting this event can be estimated and calculated. In this research, modeling the NDVI index monthly was performed using TRMM rainfall data, SMOS satellite soil moisture, land surface temperature, snow cover, with machine learning algorithms such as artificial neural network algorithm. One of the problems in this study was the high volume of data also, the SMOS satellite soil moisture images since mid-2010, so all data has been used since mid-2010. Modeling done for twelve months of the year has shown that they can be used with remote sensing images and remote sensing products that can each measure the parameters affecting agricultural drought. However, one of the major challenges in this type of modeling is the recognition of the type of parameters affecting this phenomenon. Therefore, one of the factors of modeling accuracy can be the choice of the type of remote sensing products. Neural network method, which is one of the methods of machine learning, has been able to provide acceptable results in this research. However, other methods of

machine learning such as support vector regression, random Forest, can be used in future research.

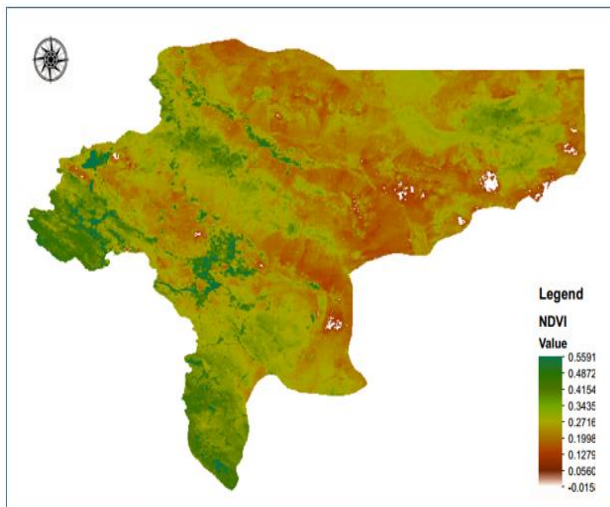


Figure3. NDVI index in July 2018

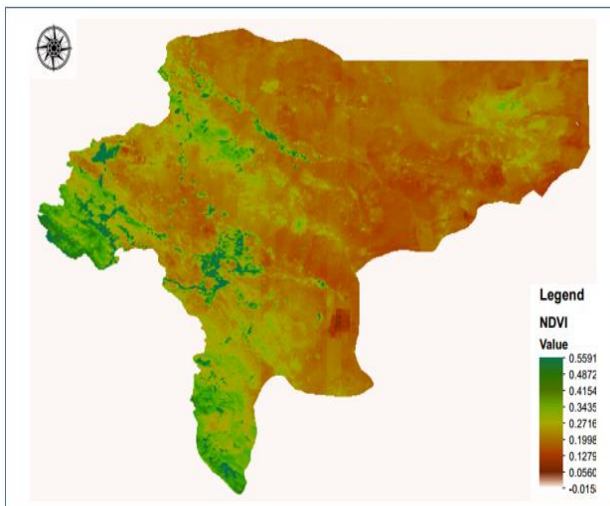


Figure4. NDVI index modelled by ANN in July 2018

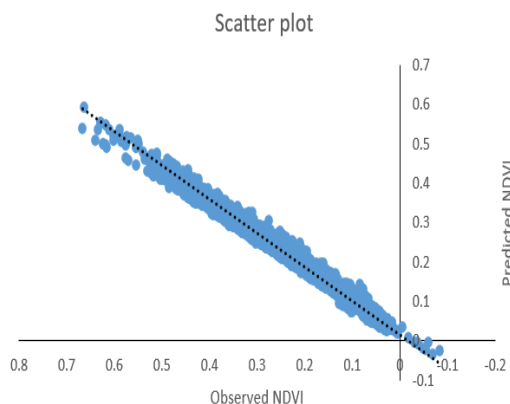


Figure 5. Scatter plots for Observed NDVI and predicted NDVI in July 2018

REFERENCES

- AghaKouchak, A., Farahmand, A., Melton, F., Teixeira, J., Anderson, M., Wardlow, B. D., & Hain, C. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, 53(2), 452-480.
- Alizadeh, M. R., & Nikoo, M. R. (2018). A fusion-based methodology for meteorological drought estimation using remote sensing data. *Remote Sensing of Environment*, 211, 229-247.
- Bai, J., Cui, Q., Chen, D., Yu, H., Mao, X., Meng, L., & Cai, Y. (2018). Assessment of the SMAP-Derived Soil Water Deficit Index (SWDI-SMAP) as an Agricultural Drought Index in China. *Remote Sensing*, 10(8), 1302.
- Barua, S., Ng, A., & Perera, B. (2012). Artificial neural network-based drought forecasting using a nonlinear aggregated drought index. *Journal of Hydrologic Engineering*, 17(12), 1408-1413.
- Belayneh, A., & Adamowski, J. (2013). Drought forecasting using new machine learning methods/Prognozowanie suszy z wykorzystaniem automatycznych samouczących się metod. *Journal of Water and Land Development*, 18(9), 3-12.
- Dawson, C., & Wilby, R. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1), 80-108.
- Heumann, B. W. (2011). Satellite remote sensing of mangrove forests: Recent advances and future opportunities. *Progress in Physical Geography*, 35(1), 87-108.
- Kim, T.-W., & Valdés, J. B. (2003). Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *Journal of Hydrologic Engineering*, 8(6), 319-328.
- Mas, J. F., & Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. *International journal of remote sensing*, 29(3), 617-663.
- Park, S., Im, J., Park, S., & Rhee, J. (2017). Drought monitoring using high resolution soil moisture through multi-sensor satellite data fusion over the Korean peninsula. *Agricultural and Forest Meteorology*, 237, 257-269.
- Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9), 503-510.
- Rhee, J., & Im, J. (2017). Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agricultural and Forest Meteorology*, 237, 105-122.
- Wilhite, D. A., & Buchanan-Smith, M. (2005). Drought as hazard: understanding the natural and social context. *Drought*

and water crises: science, technology, and management issues,
3-29.