Fusion of Markov Chain and SAX Method for Drought Probability Analysis (Case Study: Eastern District of Isfahan, Iran)

Yaser Jouybari Moghaddam[1*], Seyyed Qasem Rostami[2]

1. PhD Student, Remote Sensing Department, Faculty of Surveying and Geoinformation Engineering, College of Engineering, University of Tehran
2. Department of Surveying Engineering, Faculty of Engineering, University of Bojnord, Bojnord, Iran

(Resaved 12October 2018 – Accept 11November 2018)

Abstract

Drought is one of the most powerful natural disasters, which are affected on different aspects of the environment. Most of the time this phenomenon is immense in the arid and semi-arid area. Monitoring and prediction the severity of the drought can be useful in the management of the natural disaster caused by drought. Many indices were used in predicting droughts such as SPI, VCI, and TVX. In this paper, based on three data sets (rainfall, NDVI, and land surface temperature) which are acquired from MODIS satellite imagery, time series of SPI, VCI, and TVX in time-limited between winters 2000 to summer 2015 for the east region of Isfahan province were created. Using these indices and fusion of symbolic aggregation approximation and hidden Markov chain drought was predicted for fall 2015. For this purpose, at first, each time series was transformed into the set of quality data based on the state of drought (5 group) by using SAX. Algorithm then the probability matrix for the future state was created by using Markov hidden chain. The fall drought severity was predicted by fusion the probability matrix and state of drought severity in summer 2015. The prediction based on the likelihood for each state of drought includes severe drought, middle drought, normal drought, severe wet and middle wet. The analysis and experimental result from proposed algorithm show that the product of this algorithm is acceptable and the proposed algorithm is appropriate and efficient for predicting drought using remote sensor data.

Introduction

Drought is such natural disasters that usually covers a large area and have long-term effects. Due to the impact of this phenomenon on weather, agriculture, water and socio-economic issues, it can have an infrastructural and destructive effect on the environment. In general, due to drought dependence on multiple

* Corresponding Author, Email: yaserjoibari@gmail.com

parameters and its complexity, a definition for this phenomenon is no easy task [1]. Drought forecasting can have a useful role in mitigation of this phenomenon's damages, which depends on the exact definition of drought and linking drought with a series of associated indices. Several parameters have been defined on this basis to be modeled during the period of drought forecasting.

Based on studies in the field, these indicators can be divided into two general categories meteorological indicators and satellite remote sensing indicators [6]. The most common weather indices are Standardized Precipitation Index (SPI) and the Palmer Drought Severity Index (PDSI).Generally, satellite indices are vegetation index (VI) and land surface temperature (LST) and its derivatives [3]. Currently, the analysis of time series of drought indicators used to predict drought which is forecast the absolute numerical value based on an extrapolation of the function fitted to the time series.

Firstly, if the drought is a phenomenon with qualitative nature, so even if we express this phenomenon numerically, ultimate results must be expressed qualitatively. Secondly, the nature of the predictions is always probabilistic thus providing a fixed amount is not meaningful. Another problem of existing methods is in determining the communicational interval of any data with previous data. Due to the uncertainty in determining these ranges (delay), an error entered into the prediction process. In this study, prediction carried out in a way that the preceding be considered in it.

Material and Methods:

This research study area is eastern Isfahan Province where has five sub-regional. The study area has semi-desert climate and is located in the range of latitude N "40 '29 ° 32 and N" 47 '45 ° 32 and longitude E "29 '42 ° 51 to" E52 '59 ° 51. Figure 1 shows the study area. The data used in this research is land surface temperature (LST), and normal differential vegetation index (NDVI) from MODIS satellite products that are free and downloaded from the NASA Earth Observations (NEO) Other data were also used is precipitation data from TRMM. The data for a period of 16 years from winter 2000 to summer 2015 were downloaded.

A. Symbolic Aggregate Approximation method

Symbolic Aggregate approximation method is one of the approaches to show time series offered by Lin et al. in 2003. This process took a time series as input and turned it into a set of strings as output [15]. By the use of Symbolic Aggregate Approximation method, a time series of arbitrary length n can be converted to an arbitrary string with length w ($w < n$ and commonly $w \ll n$). This method is based on the fact that normalized time series follow Gaussian distribution (Larsen and Marx 1986).

Symbolic Aggregate approximation method consists of two main stages. First, convert the Piecewise Aggregation Approximation (PAA) to reduce the time series dimension and second discrete time series obtained from the previous step to convert it to the string.

B. Markov Chain

A Markov chain is kind of modeling in which the current state of the system depends on its previous state. Determining the state of the system (projected) by using Markov model needs previous state of the system and the possibility of changing in system state to other possible states, the so-called transition probabilities to be known [16].

According to the current state of a system, a square matrix P formed and matrix elements P_{ij} has represented the transition probability.

In this matrix, the likelihood of early states in the left column and the possibility of cases where the system passes them along the lines of the matrix are shown.

$$P = \begin{bmatrix} p_{r1} & \cdots & P_{1j} \\ \vdots & \ddots & \vdots \\ P_{j1} & \cdots & P_{jj} \end{bmatrix} \qquad j=1,2,3\ldots r \tag{1}$$

Discussion and Results

First, by using data from TRMM sensor and the standardized precipitation index SPI and via Equation 2 for monthly time series of the winter of 2000 until the end of 2015 summer was calculated. By using time series data, land surface temperature (LST) and the normalized vegetation index (NDVI) for the same period with precipitation data, two indices VCI TVX were calculated using Equations 3 and 4.

$$SPI = \frac{P - \mu(P)}{\delta(P)} \tag{4}$$

$$TVX = \frac{LST}{NDVI} \tag{5}$$

$$VCI = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \tag{6}$$

In the second stage, the normal time series were used as input of SAX methods. First, since environmental changes are more noticeable in seasons the amount of W considered 63 for the PAA convert to reduce seasonally adjusted time series Figure 4 Showing PPA time series of indices used in this research. The next step, assuming a Gaussian distribution for each indicator, and the values of δ and δ3 as breakpoints for SAX method were selected.

These values were based on probability levels. Moreover, strings intended to convert SAX have been chosen as follows: 1- SW: extreme wet 2- MW: wet 3-

N: Normal 4- MD: drought 5- SD: severe drought.

According to these rules SAX conversion implemented and time series converted to the set of strings which are indicators of drought. Figure 1 showing the transformation for each series of the time.

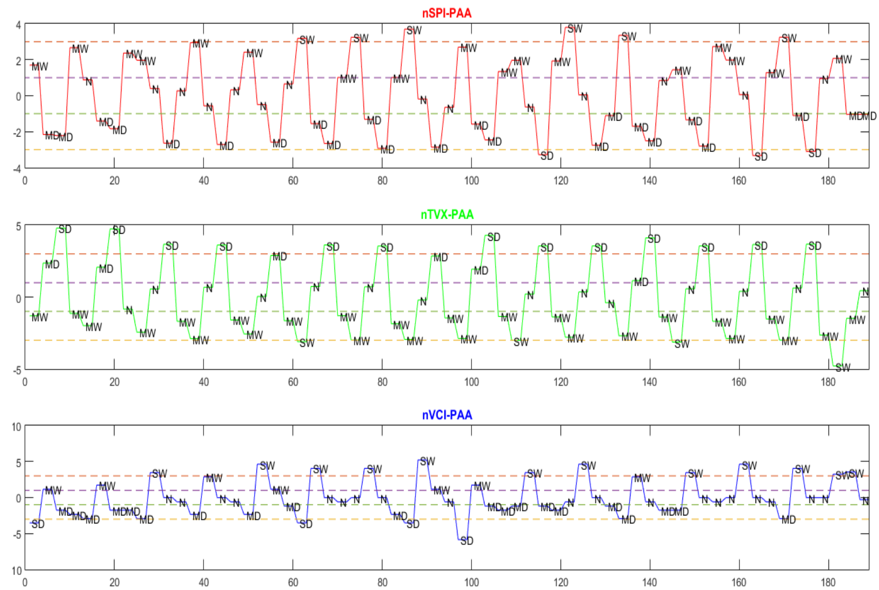

Fig. 1. SAX presentation for drought index

The horizontal axis represents time, the vertical axis represents the amount of PAA for each indicator, and the number written on each of intervals represents the new time series value. In other words, the input time series after SAX conversion converted to a set of strings of qualitative drought values. After conversion of each time series into qualitative data through the SAX method, in the next stage, the collection of qualitative data by using Markov chain method were prepared to predict the probability of the next state. In other words, the transition state matrix for each index was determined. Figure 2 shows probabilistic values for each index through Markov chain.

$$\begin{array}{c|ccccc} & SD & MD & N & MW & SW \\ \hline SD & 0.67 & 0 & 0 & 0.17 & 0.16 \\ MD & 0.03 & 0.81 & 0.01 & 0.05 & 0.1 \\ N & 0.02 & 0.10 & 0.82 & 0 & 0.05 \\ MW & 0 & 0.19 & 0.14 & 0.66 & 0 \\ SW & 0 & 0.03 & 0.23 & 0.66 & 0 \end{array} \tag{1}$$

	SD	MD	N	MW	SW
SD	0.66	0	0.12	0.22	0
MD	0.02	0.79	0.08	0.1	0.01
N	0.04	0.14	0.68	0.09	0.03
MW	0	0.07	0.12	0.73	0.18
SW	0	0.22	0.11	0	0.67

(2)

	SD	MD	N	MW	SW
SD	0.66	0	0.33	0	0
MD	0.04	0.79	0.10	0.06	0.01
N	0	0.05	0.66	0.05	0.23
MW	0	0.12	0	0.66	0.22
SW	0	0.28	0.06	0	0.66

(3)

Fig. 2. Probability model for drought based on (1) VCI (2) SPI (3) TVX

The correlation coefficient between the values of the likelihood of each index with other indices to evaluate accurately calculated as follows: 1 .The correlation coefficient between SPI and TVX equal to 89. 2. The correlation coefficient between SPI and VCI equal to 95. 3. The correlation coefficient between VCI and TVX index equal to 89.

High amounts of correlation between the probability values indicate that resulted probability values are acceptable.

In the final step, after calculating the probability values by using the last status of the three indicators, drought conditions for the next month predicted. The forecast shows that in the fall of 2016 in the Eastern region of Isfahan Province by the possibility of 2% severe drought, 12% drought, 50% typical situation (normal), 25% wet and 10 % extreme wet will occur. For validation of this probability model the dataset of October, November and December for 2015 are analyzed the result shows the probability model is matched with this dataset.

Conclusion

This study was conducted to investigate, modeling and forecast drought, one of the world's natural hazards and controversial issue. For this purpose, time series of three indices SPI, VCI, and TVX between winter 2000 and summer 2015 were designed and used, and in this study, it was done by incorporation of SAX and Markov chain. The advantage of using a combination of these methods compared to other methods is it provide a probabilistic qualitative model of drought.

In fact, due to the nature of drought that is qualitative and on the other hand, fuzzy and probabilistic nature of the predictions, this method seems more reasonable than other modeling methods. On the contrary, due to the high correlation between probabilistic models obtained as well as implementing the method for certain modes and the logical outcome of this case, the accuracy of the proposed method was acceptable.

The authors examining the impact of changes in the value of break points of SAX method on the result of probabilistic model and also determining these values on the basis of conditions of each climate left for the future works. In addition, using an analysis of other rules in order to qualify the indices in SAX method or using another method to display time series can be helpful to advance this research.

Keywords: SAX, Markov Chain, Drought, Remote Sensing, Isfahan.

Reference

[1].S. Javanmard, J. Badagh, J. Ahmadian and N. Javdani, “Care system for severity and extent of drought based on Palmer index segmentation”, 1st National Conference on Drought Mitigation and Water, Kerman, 1379

[2].I. Javadnia, M.R. Mobasheri, “Agricultural Drought assessment using remote sensing techniques”, 6th national Conference of Agricultural economic, Mashhad, 1386.

[3].F. Mahmodi, A. Esmaeli and M. Hani, “study of rainfall effect on Drought using Remote Sensing Index”, Geomatics 90, NCC, 1390.

[4].Y.J. Moghaddam, Akhoondzadeh, M. and Saradjian, M.R., 2015. A Split-Window Algorithm for Estimating LST from Landsat-8 Satellite Images. Journal of Geomatics Science and Technology, 5(1), pp.215-226.

[5].R. Moazzenzadeh, S.Arshad, B.Ghahraman, and K. Davari, “Drought monitoring in unirrigated lands based on the remote sensing technique”, Journal of Water and Irrigation Management, 2, 39:52, 1391.

[6].Solimani, “Analysis of Drought Risk”, MSc thesis, University of Tehran, 1389.

[7].R.R. Heim, “A review of twentieth-century drought indices used in the United States." Bulletin of the American Meteorological Society. 83: 1149-1165, 2002.

[8].G. Berhan, S. Hill, T. Tadesse, and S. Atnafu, “Using satellite images for drought monitoring: A knowledge discovery approach.” Journal of Strategic Innovation and Sustainability, 7:1, 135–15, 2011.

[9].T. Gebrehiwot, A. Van der Veen, and B. Maathuis, “Spatial and temporal assessment of drought in the northern highlands of Ethiopia.” International Journal of Applied Earth Observation and Geoinformation, 13, 309–32, 2011

[10]. M. T. Dastorani, H. Afkhami, and B. Borroni, “Application of artificial neural networks on drought prediction in Yazd (Central Iran)”. Desert, 16, 39–4, and 2011.

[11]. M. E. Keskin, O. Terzi, E. D. Taylan, and D. Kucukyaman, “Meteorological drought analysis using artificial neural networks.” Scientific Research and Essays, 6, 4469–447, 2009.

[12]. A. Nikhbakht Shahbazi, M. Heidarnejhad, “Meteorological drought prediction in Karoon watershed using meteorological variables.” International Research Journal of Applied and Basic Sciences, 3:9, 1760–1768, 2012.

[13]. B. Shirmohammadi, H. Moradi, V. Moosavi, M. Taei Semiromi, and A. Zeinali, “Forecasting of meteorological drought using wavelet-ANFIS hybrid model for different time steps (Case study: southeastern part of East Azerbaijan province, Iran).” Natural Hazards, 69, 389–402, 2013.

[14]. M. Jalili, J. Gharibshah, S. M. Ghavami, M. R. Beheshtifar, and R. Farshi, “Nationwide prediction of drought conditions in Iran based on remote sensing data.” IEEE Transaction on Computers, 63:1, 90–101, 2014.

[15]. J. Lin, E. Keogh, S. Lonardi, and B. Chiu, A symbolic representation of time series, with implications for streaming algorithms. InProceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery 2003 Jun 13 (pp. 2-11). ACM.

[16]. L. E. Baum, T. Petrie, “Statistical Inference for Probabilistic Functions of Finite State Markov Chains”, the Annals of Mathematical Statistics 37 (6): 1554–1563., 1996.

[17]. W. Palmer, "Meteorological Drought". Research paper no.45, U.S. Department of Commerce Weather Bureau, February 1965 (58 pgs)

[18]. R. J. Larsen, M. L. Marx, “An Introduction to Mathematical Statistics and Its Applications.” Prentice Hall, Englewood, Cliffs, N.J. 2nd Edition, 1986.

[19]. M. Rulinda Coco, Mining Drought from Remote Sensing Images, MSc thesis, Geo-information Science and Earth Observation, 2007.

[20]. D. A .Wilhite, “Drought as a natural hazard.” In D. A. Wilhite (Ed.), Drought. A Global Assessment, vol. I (pp. 1– 18). London: Routledge.

[21]. Y.J. Moghaddam, Akhoondzadeh, M. and Saradjian, M.R., 2015. A Novel Method for Retrieving Land surface Emissivity from Landsat-8 Satellite Data Based on Vegetation Index. Journal of Geomatics Science and Technology, 5(1), pp.175-187.