Irrigation Sciences and Engineering (JISE) Vol. 42, No. 4, Winter 2019, p. 45-59

EXTENDED ABSTRACT

Regionalization of the Eastern Part of Urmia Lake Basin Based on the Rainfed Yield and Precipitation Using the Ward, K-Means and PCA Methods

P. Mohammadi[1*], A. Fakherifard[2], Y. Dinpazhoh[3] and E. Asadi[4]

1*- Corresponding Author, MSc, Water Resources Engineering, Faculty of Agriculture, Tabriz University, (mohammadi.parva@yahoo.com).
2- Professor, Water Resources Engineering, Faculty of Agriculture, Tabriz University.
3- Associate Professor, Water Resources Engineering, Faculty of Agriculture, Tabriz University.
4- Assistant Professor, Water Resources Engineering, Faculty of Agriculture, Tabriz University.

Received: 3 October 2016 Revised: 26 December 2017 Accepted: 30 December 2017

Keywords: Clustering, Precipitation, Regionalization, Rainfed yield. ***DOI:*** 10.22055/jise.2017.19506.1396.

Introduction

Rainfall is among the most important climatic factors affecting the rainfed cultivation. Thus, in order to maintain water consumption in current agriculture, with the view of water resources management, the country needs to convert some irrigated land areas to rainfed cultivation in the near future. Indeed, it is necessary to conduct an analytical study on rainfed agriculture and identify appropriate areas for rainfed agriculture in the country, especially in Urmia Lake basin. Principal component analysis (PCA), K-Means and Ward have been already used to assess climate regionalization in different regions such as Spain (Diaz and Rodrigo, 2004), Greece (Kitsara et al, 2005), central-northeastern region of Mexico (Pineda-Martinez et al, 2007), Luanhe basin (Hassan and Ping, 2012) and Iberian Peninsula (Parracho et al, 2015). This study was, thus, intended to study the regionalization of the eastern part of Urmia Lake basin on the basis of the precipitation and yield of rainfed wheat using PCA, K-Means and Ward methods. To that end, the maps were drawn in the GIS environment and three methods of clustering were compared. Finally, using the clustering of precipitation and rainfed yield, wheat cultivability was investigated in the eastern part of Urmia Lake. To that end, the daily rainfall dataset of 26 rain gauge stations were used and the yield of rainfed wheat was considered during the period. Then, PCA, K-Means and Ward clustering were performed and the results were compiled. The homogenousity of the resulting clusters were analyzed by H and S statistical tests and homogeneous clusters were drawn in the GIS environment. The analytical factor coefficients to the main components, through K-Means clustering method, showed that the clusters point of view, precipitation and rainfed yield were more consistent and the results were close to each other.

Methodology

The daily rainfall statistical data of 26 rain stations from the years 1992 to 2014 were gathered from the eastern part of Urmia Lake. For the regionalization of the eastern part of Urmia Lake, based on the precipitation and rainfed yield, wheat was selected as a strategic product and PCA, K-Means and Ward methods were used for its yield analysis. PCA is considered as a multivariate statistical method. Thus, initial variables were transformed into new and independent variables, as new components were linear combinations of the initial variables. The main purpose of PCA was to decrease the number of predictive variables and convert them into new variables, called principal components (PC) (Rencher, 2002).

However, in the Ward method, each object was considered as a separate cluster. Then, in each step, the two objects were merged together. This work was repeated so that a single cluster was formed at the end of the process. This method was adopted as it minimized errors in joining the clusters, as recommended by many scholars using hierarchical methods (Romero et al., 1999; Raziei, 2017). K-Means method was, in turn, the most practical method for data clustering and was first introduced by Macqueen (1967). The number of clusters in this method is constant and predetermined. This method is based on minimizing the squared error or variance within a group that is equivalent to maximizing the variance between clusters. This algorithm was, indeed, intended to obtain parts that minimized the square of the errors with the constant number of clusters. There are several tests used for the homogeneity of the clusters in terms of hydrology, but with regard to the dependence of most tests on the type of statistical distribution, H Hosking (1993) and S statistical tests (Wiltshire, 1986) were utilized in this research. S statistical test is based on the coefficient of variation. However, H statistics test operates based on the standardized observations of linear momentum based on the artificial data moments derived from the best distribution fitted to the observational data.

Results and Discussion

In this research, the regionalization of precipitation and rainfed yield in the eastern part of Urmia Lake basin was investigated using clustering methods such as K-Means, Ward and PCA in the GIS environment in two regions. Based on the factor coefficients with K-Means clustering, the first region included the northeast, northwest, south and center of the region, while the second region included parts of the north, northeast, and west of the study area.

In Ward clustering based on the factor coefficients of precipitation, the first region covered north, northwest, and parts of the northeast, center, and south of the region. However, the second part included the south, center and part of the northeast of the region.

In the regionalization of the rainfed yield based on the factor coefficients with clustering K-Means, the first region included parts of the northeast, northwest, west and southwest of the region. The second region, however, covered the northeastern and central parts of the region. Based on the factor coefficients, Ward clustering included the northeastern, southern, western, and central parts of the study area. In turn, the second region covered the northern, northwestern, and southern parts of the study area.

The regionalization of precipitation and rainfed yield was divided into two regions by Ward and K-Means methods. In the regionalization of precipitation, the first area covered the western, southwestern, and central parts. The second region of the northeast was part of the center and northwest of the study area.

The results of regionalization of the rainfed yield using the Ward method showed that the first region covered the northeastern, southern, western and central parts of the study area. The second region covered parts of the north, northwest, and northeast of the region.

In K-Means method, the precipitation regionalization covered the northeastern, central and northwestern parts. The second region was part of the north, south, and parts of the northeast and west of the region. Regionalization of the rainfed yield by K-Means method included the northeastern, southern, western and central parts of the study area. However, the second region covered parts of the north, northwest, north-east and south of the region.

Conclusions

Results showed that, the analytical factor coefficients to the main components, through K means clustering method, from the percent of area and average showed that clusters point of view, precipitation and rainfed yield were more consistent and the results were close to each other. On the other hand, from the place changing point of view, the yield lines were drawn along with the lines having the same precipitation. Increasing the crop yields and rainfall in the northwest and center of the study area showed results which were consistent with the physical nature of the rainfed yield process. However, some areas were found to be heterogeneous in other parts of the region. In effect, the heterogeneous and homogeneous areas respectively covered 24.47 percent and 75.53 percent of the study area. Given the

  DOI: 10.22055/jise.2017.19506.1396

importance of the rainfed cultivation, the introduction of such prone areas for optimal water use can play a significant role in preserving water resources.

Acknowledgment

We are grateful to the University of Tabriz and the East Azarbaijan Regional Water Company for providing the necessary facilities for conducting this research and preparing the relevant paper.

References

1- Diaz, D.M. and Rodrigo, F.S., 2004. Spatio-temporal patterns of seasonal rainfall in Spain (1912-2000) using cluster and principal component analysis: comparison. *Journal of Annales Geaphssicae*, 22, pp. 1435-1448.

2- Hassan, B.G.H. and Ping, F., 2012. Formation of homogenous regions for Luanhe basin by using L-moments and cluster techniques. *International Journal of Environmental Science and Development*, 3(2), pp. 205- 210.

3- Hosking, J.R.M. and Wallis, J.R., 1993. Some statistics useful in regional frequency analysis. *Water Resource Research*, 29 (2), pp. 281- 671.

4- Kitsara, G., Pappaioannou, G., Mitropoulo, A. and Markopoulos, P., 2005. Reference Crop evapotranspiration and agricultural rainfall index. *In the 9th International Conference on environmental Science and Technology*, Rhodes island, Greece.

5- Macqueen, J., 1967. Some methods for classification and analysis of multivariate observation. *In Proceeding of the 5th Berkeley symposium on mathematical statistics and probability*, Berkeley, CA: University of California.

6- Parracho, A.C., Melo-Goncalves, P. and Rocha, A., 2015. Regionalisation of precipitation for the Iberian Peninsula and climate change. *Physics and Chemistry of the Earth*, Parts A/B/C, 94, pp. 146-154.

7- Pineda-Martinez, L.F. and Carbajal, N. and Median –Roldan, E., 2007. Regionalization and classification of bioclimatic zones in the central- northeastern region of Mexico using principal component analysis. *Journal of Atomofera*, 20(2), pp. 133- 145.

8- Raziei, T. 2017. A precipitation regionalization and regime for Iran based on multivariate analysis. Theoretical and Applied Climatology, 131(3-4), pp. 1429-1448.

9- Rencher, A.C., 2002. Methods of multivariate analysis. John Wiley and Sons, INC publication.

10- Romero, R., Sumner, G., Ramis, C., Genovés, A. A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 19(7), pp. 765-785.

11- Wiltshire, S.E., 1986. Identification of homogeneous regions for flood frequency analysis. *Journal of Hydrology*, 84(3), pp.287-302.