

AMD and PCA multivariate outlier detection: A case study of Chahar Gonbad area

MEHRAN FARAHMANDIAN^{1*}, SADEGH KIANPOURYAN¹
AND SIAM SAFAEI¹

¹Iranian Academic Center for Education, Culture and Research
IUT Branch, Isfahan, Iran

One of the main tasks in environmental and exploration geochemistry is the detection of outlier data and unusual structures. Outliers can greatly impact the results of the statistical methods especially separation of anomaly from background, thus recognition and decision about removal or correction of them is one of the first steps in the analysis of geochemical data processing. They can be identified from a univariate, bivariate, or multivariate perspective based on the number of variables considered, and the multivariate detection in Chahar Gonbad area is the aim of this study. Since geochemical data are compositional data, so the isometric log-ratio transformation were used for open them prior to any analysis. Consequently, Principal component analysis (PCA) and Adaptive Mahalanobis Distance (AMD) were applied on the stream sediment geochemical data of Chahar Gonbad area to recognize outliers. Based on the PCA method, about 8 percent of data were distinguished as outlier; however this value for AMD technique is approximately 6 percent. The later method not only separate outliers but also recognize extreme values, therefore the outliers which identify with it are lower than PCA. The results of these techniques indicated that both of them are powerful in outlier detection, nevertheless AMD is more powerful than PCA.