Discovering geochemical patterns by factor-based cluster analysis

Jian Wang, Renguang Zuo, Jef Caers

PII: S0375-6742(17)30179-6
DOI: doi: 10.1016/j.gexplo.2017.07.006
Reference: GEXPLO 5955

To appear in: *Journal of Geochemical Exploration*

Received date: 10 March 2017
Revised date: 22 June 2017
Accepted date: 9 July 2017

Please cite this article as: Jian Wang, Renguang Zuo, Jef Caers , Discovering geochemical patterns by factor-based cluster analysis, *Journal of Geochemical Exploration* (2017), doi: 10.1016/j.gexplo.2017.07.006

Discovering geochemical patterns by factor-based cluster analysis

Jian Wang[1, 2], Renguang Zuo[1*], Jef Caers[2*]

[1] State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China

[2] Department of Geological Sciences, Stanford University, California 94305, USA

Corresponding authors: zrguang@cug.edu.cn (R. Zuo); jcaers@stanford.edu (J. Caers)

Abstract

Exploring geochemical patterns based on statistical data analysis is a convenient and effective way to improve our understanding of the geochemical characteristics of the given study area. In this paper, a procedure combining factor analysis and cluster analysis in a novel way was proposed to discover geochemical patterns of interest, or more specifically, a cluster of observations most likely indicative of mineralization. Factor analysis on the entire dataset is firstly performed to obtain a baseline factor structure to be compared with. Secondly, cluster analysis based on factor scores is carried out to produce several clusters reflecting heterogeneity within the dataset. This type of heterogeneity is further investigated in more detail by implementing factor analysis upon each cluster, and comparing the factor structure of each cluster with the baseline factor structure by a defined measure. The randomization test was applied here to further test whether the difference of measures is significant or not. A case study of processing soil samples collected in Jilinbaolige district, Inner Mongolia of China, is presented to illustrate and validate this procedure.

Keywords: Geochemical patterns; Factor analysis; Cluster analysis; Compositional data;