Estimation of fault activity by linear discriminant analysis using chemical composition of fault gouge

*Ryo Tateishi[1], Koji Shimada[2], Mayuko Shimizu[2], Shigeru Sueoka[2], Masakazu Niwa[2], Tsuneari Ishimaru[2]

1. University of Toyama, 2. Japan Atomic Energy Agency

Introduction: The identification of active faults is based on the displacement and deformation of the current topography and the late Quaternary strata. However, in the absence of them at the site, it is quite difficult to determine the fault activity. Tateishi et al. (2019) performed linear discriminant analysis using the chemical composition data of fault gouges of active faults and inactive faults in Japan, and found that the obtained discriminants can identify fault activity without error. It has been shown that the combinations of elements that represent are limited to six elements: TiO_2, Al_2O_3, MgO, Na_2O, P_2O_5, and Ba. In this study, we performed linear discriminant analysis with additional data for the purpose of further narrowing down the elements representing the difference in fault activity and obtaining a better discriminant.

Methods: This research was carried out in the following order; collection of papers, extraction of chemical composition data of fault gouge, addition of new chemical analysis data, formatting of input data, variable selection, linear discriminant analysis. First, chemical composition data of fault gouge of active faults and inactive faults in Japan was extracted from existing publications. The extraction condition is that the fault activity of the analysis target clearly describes and that the basement rock is diverse. In addition, a chemical analysis of the fault gouge sample possessed by the JAEA was carried out, then we created a database that integrates both of them. In this database, in addition to analytical value, 2 categories of fault activity were given to each sample. Next, we extract samples of fault gouges from the database, and select elements to maximize the number of samples. In order to further extract elements suitable for discrimination of fault activity, elements were selected by AIC (Akaike, 1973). Finally, linear discriminant analysis was performed using the combination of elements selected by AIC as explanatory variables.

Results: The chemical composition data of fault rocks was extracted from a database of a total of 276 samples, including the data from collection of 8 publishing papers, 7 JAEA reports, and chemical analysis of 36 fault rocks held by JAEA. Then we selected 15 elements as variable candidates. As a result, a total of 72 samples (48 samples for active faults and 24 samples for inactive faults) were chosen as input data. From these 15 elements, 11 elements were selected as optimum combination of elements in AIC. Then we performed linear discriminant analysis with a following combination of elements; (a) 11 elements selected by AIC, (b) 8 elements with p-value between 0 and 0.01, (c) 6 elements with p-value between 0 and 0.001. The discrimination rate between active faults and inactive faults is 100% in (a), (b) and 97% in (c).

Consideration about the elements representing the difference in fault activity: When the coefficients of the elements in these discriminants are arranged in descending order, the top 6 elements and the top 4 elements were common. Among them, the combination of TiO_2 and P_2O_5, and the combination of Al_2O_3 and Rb show high correlation coefficient with 2 groups. In addition, the distribution patterns of TiO_2, CaO, P_2O_5 and Ba are concentrated at low values on inactive faults, and show a wide distribution at higher values on active faults. Thus, the elements that represent the difference between the 2 groups are top 6 elements. Of these, TiO_2 and P_2O_5, and Al_2O_3 and Rb are considered important, including their respective combinations.

Consideration about better discriminant: Of the (a) and (b) with a discrimination rate of 100%, the number of elements is small in (b). However, (b) includes TiO_2 and P_2O_5, which are considered to be collinear as above. Therefore, we performed linear discriminant analysis with the combination of elements excluding TiO_2 and P_2O_5 from (b) as (c) and (d), and as a result, the discrimination rate was 99% in each case. From these, it is expected that the ability to discriminate a sample with known activity is the highest at (b), but that the discrimination performance for a sample with unknown activity is higher at (c) or (d). These results will greatly contribute to clarify the mechanism that creates the difference in chemical composition between active and inactive faults.

This study was carried out under a contract with METI (Ministry of Economy, Trade and Industry) as part of its R&D supporting program for developing geological disposal technology.

Referrence: Akaike, H., Proceedings of the 2nd International Symposium on Information Theory, Petrov, B. N., and Caski, F. (eds.), p. 267-281, 1973. ; Tateishi et al., JpGU2019, SSS15-P27, 2019.

Keywords: fault activity, chemical composition of fault gouge, linear discriminant analysis