Lithological mapping based on geochemical and remote sensing data using a hybrid method of random forest and metric learning

ZIYE WANG*, RENGUANG ZUO

State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China (*correspondence: *ziyewang@cug.edu.cn*)

Multi-source geoscience data (e.g., geology, geochemistry, geophysics, and remote sensing) provide abundant information for mineral exploration in different ways. For example, remote sensing images record the electromagnetic characteristics, and geochemical data represent the enrichment or depletion of geochemical elements, which reflect the physical and chemical attributes of lithological units, respectively. In this study, we developed a hybrid model that comprised of data fusion technology and metric learning method. Multi-source data fusion technology was first used to provide more accurate information for lithological identification by integrating geochemical data and ASTER remote sensing images, which retain the distribution characteristics of geochemical anomalies and clearer texture structure in remote sensing images. To address the redundant information in the fused data, a random forest metric learning (RFML) approach was introduced to promote classification and computing efficiency. RFML use random forest as the basic framework and applies metric learning to the classification of every decision tree, making full use of the advantages of random forest and metric learning. This process was illustrated through a case study of lithological mapping in Cuonadong dome, the northern part of the Himalayan, China. Seven target lithological units were effectively discriminated, with an 93% overall accuracy. This study provides an important scientific basis for further mineral exploration by detecting geological features under harsh natural conditions.