Abstract

Chemom. Intell. Lab. Syst., **138**, 64-71 (2014)

Varmuza K., Filzmoser P., Hilchenbach M., Krüger H., Silén J.:

KNN classification - evaluated by repeated double cross validation: Recognition of minerals relevant for comet dust.

Repeated double cross validation (rdCV) has recently been suggested as a careful and conservative strategy for optimizing and evaluating empirical multivariate calibration models.

This evaluation strategy is adapted in this work for *k*-nearest neighbor (KNN) classification. The basics of rdCV are described, including the search for an optimum *k*, and tests with Italian Olive Oil Data.

KNN-rdCV is applied to classify 17 mineral groups, relevant for the composition of comet dust particles, characterized by the peak heights at 20 selected masses in time-of-flight secondary ion mass spectra (TOF-SIMS). Predictive abilities for 15 mineral classes are >95%, for two classes 75 and 85%.