Abstract

Chemom. Intell. Lab. Syst., **89**, 1-8 (2007)

Liebminger A., Papesch W., Haberhauer G., Varmuza K.:

Multivariate models for the concentration of oxygen-18 in precipitation based on meteorological and geographical features.

Concentrations of the stable isotope oxygen-18 in precipitation samples have been modeled by geographical and meteorological features of the sampling stations.

Precipitation samples are from 30 locations in Austria; oxygen-18 concentrations have been determined by isotope ratio mass spectrometry; each location has been characterized by three geographical features (longitude, latitude and elevation), and five meteorological features (relative humidity, fresh snow, wind speed, precipitation, and air temperature). All data are monthly means for the summer period April to September computed as long term averages for a time period of typically 20 years.

The basic feature set has been augmented by 49 nonlinear transformations of the original features giving a data set with 57 features. The number of objects was 180 corresponding to the 30 sampling stations times six months.

Different methods of feature selection (correlation coefficient, backward elimination, all subsets regression, and genetic algorithm) have been applied, and for the 20 best feature subsets the prediction errors of PLS models have been estimated from test sets.

The final best model for a prediction of oxygen-18 concentrations contains ten features selected by a genetic algorithm. This model has been applied to the computation of a geographical oxygen-18 distribution map from Austria based on computed oxygen-18 values for 200 locations, and experimental values of 30 locations.