Geophysical Research Abstracts
Vol. 15, EGU2013-9708, 2013
EGU General Assembly 2013

Finite mixture models for the computation of isotope ratios in mixed isotopic samples

Daniel Koffler (1), Gregor Laaha (1), Friedrich Leisch (1), Stefanie Kappel (2), and Thomas Prohaska (2)

(1) BOKU Vienna, Institute of Applied Statistics and Computing, Department of Landscape, Spatial and Infrastructure Sciences, Vienna, Austria (daniel.koffler@boku.ac.at), (2) Department of Chemistry, Division of Analytical Chemistry – VIRIS Laboratory, University of Natural Resources and Life Sciences, BOKU Vienna, Konrad Lorenz-Straße 24, 3430 Tulln, Austria (prohaska@mail.boku.ac.at)

Finite mixture models have been used for more than 100 years, but have seen a real boost in popularity over the last two decades due to the tremendous increase in available computing power. The areas of application of mixture models range from biology and medicine to physics, economics and marketing. These models can be applied to data where observations originate from various groups and where group affiliations are not known, as is the case for multiple isotope ratios present in mixed isotopic samples. Recently, the potential of finite mixture models for the computation of 235U/238U isotope ratios from transient signals measured in individual (sub-)μm-sized particles by laser ablation – multi-collector - inductively coupled plasma mass spectrometry (LA-MC-ICPMS) was demonstrated by Kappel et al. [1]. The particles, which were deposited on the same substrate, were certified with respect to their isotopic compositions. Here, we focus on the statistical model and its application to isotope data in ecogeochemistry.

Commonly applied evaluation approaches for mixed isotopic samples are time-consuming and are dependent on the judgement of the analyst. Thus, isotopic compositions may be overlooked due to the presence of more dominant constituents. Evaluation using finite mixture models can be accomplished unsupervised and automatically. The models try to fit several linear models (regression lines) to subgroups of data taking the respective slope as estimation for the isotope ratio. The finite mixture models are parameterised by:
• The number of different ratios.
• Number of points belonging to each ratio-group.
• The ratios (i.e. slopes) of each group.
Fitting of the parameters is done by maximising the log-likelihood function using an iterative expectation-maximisation (EM) algorithm. In each iteration step, groups of size smaller than a control parameter are dropped; thereby the number of different ratios is determined. The analyst only influences some control parameters of the algorithm, i.e. the maximum count of ratios, the minimum relative group-size of data points belonging to each ratio has to be defined. Computation of the models can be done with statistical software. In this study Leisch and Grün's flexmix package [2] for the statistical open-source software R was applied. A code example is available in the electronic supplementary material of Kappel et al. [1].

In order to demonstrate the usefulness of finite mixture models in fields dealing with the computation of multiple isotope ratios in mixed samples, a transparent example based on simulated data is presented and problems regarding small group-sizes are illustrated. In addition, the application of finite mixture models to isotope ratio data measured in uranium oxide particles is shown. The results indicate that finite mixture models perform well in computing isotope ratios relative to traditional estimation procedures and can be recommended for more objective and straightforward calculation of isotope ratios in geochemistry than it is current practice.

[1] S. Kappel, S. Boulyga, L. Dorta, D. Günther, B. Hattendorf, D. Koffler, G. Laaha, F. Leisch and T. Prohaska: Evaluation Strategies for Isotope Ratio Measurements of Single Particles by LA-MC-ICPMS, Analytical and Bioanalytical Chemistry, 2013, accepted for publication on 2012-12-18 (doi: 10.1007/s00216-012-6674-3)
[2] B. Grün and F. Leisch: Fitting finite mixtures of generalized linear regressions in R. Computational Statistics & Data Analysis, 51(11), 5247-5252, 2007. (doi:10.1016/j.csda.2006.08.014)