Geophysical Research Abstracts
Vol. 20, EGU2018-12667, 2018
EGU General Assembly 2018

User-assisted processing of GC-MS data in R

Johanna Menges, David Maas, Micha Dietze, and Dirk Sachse
GFZ German Research Centre for Geosciences, section 5.1 Geomorphology, Germany (menges@gfz-potsdam.de)

Organic geochemical data measured either on a Gas Chromatograph – Mass Spectrometer (GC-MS) or a High Performance Liquid Chromatograph (HPLC) -MS are widely used in the field of Paleoclimatology.
The data received from the mass-spectrometers is usually manually treated in spreadsheets. While spreadsheets have the advantage of being familiar to everyone from basic education, it can be time consuming and prone to errors to copy and paste data manually. Additionally, in large spreadsheets the overview of formulas and relations between cells is complicated especially when inherited templates are used. Thus, by this practice, a principal of open and reproducible science is undermined.
In our lab we have developed an R-based solution utilising the R-package GCalignR to process n-alkane data measured on a GC-MS. The script enables us to treat all samples measured under the same conditions, simultaneously. For each sample a text-file is created from the GC-FID signal, with a-priori integrated compound peaks. These files are then transformed in R and peaks are aligned, allowing for robust and clear assessment of the similarity between samples. Peaks are then identified, quantified and ratios are calculated for further investigations.
Members of the group have started using this work flow and we are now discussing how to step further and implement version control via GitHub and converting the scripts into documented functions. We want to give an example how anyone could start to automate repetitive tasks in order to use time efficiently and gain consistency in data treatment.