American Mineralogist, Volume 109, pages 1598–1607, 2024

GCDkit.Mineral: A customizable, platform-independent R-language environment for recalculation, plotting, and classification of electron probe microanalyses of common rock-forming minerals

Vojtěch Janoušek[1,2,*], Colin M. Farrow[3], and Vojtěch Erban[1,†]

[1]Czech Geological Survey, Klárov 3, 118 21 Prague 1, Czech Republic
[2]Institute of Petrology and Structural Geology, Charles University, Albertov 6, 128 43 Prague 2, Czech Republic
[3]Glasgow, Scotland

Abstract

GCDkit.Mineral is a platform-independent (Windows/Mac/Linux) freeware for recalculation, plotting, and statistical treatment of mineral data obtained by microbeam techniques, typically an electron microprobe. It is written in R, a language providing a feature-rich environment for statistics and data visualization.

This new program imports compositional data in various commonly used file formats or retrieves them from the clipboard. Routines are available for data management, i.e., grouping, searching, and generation of subsets, using regular expressions and Boolean logic. Raw compositional data (wt%) are recalculated to atoms per formula unit (apfu) based on a required number of O equivalents, atoms, or charges, with or without Fe^{II}/Fe^{III} estimation by various methods. Analyses may then be recast to structural formulae, i.e., the atoms are distributed into appropriate crystallographic sites. For minerals forming solid solutions, the molar percentages of end-members are computed. All the data may be treated statistically, either by built-in functions for descriptive and multivariate statistics or using the wealth of tools provided by the wide R community.

Raw and recalculated mineral data may be plotted on assorted binary and ternary plots and boxplots. Most are defined as internal templates that provide a means to make later changes to the plot (zooming and scaling, adding comments or legend, identifying data points, altering the size or color of the plotting symbols, etc.). The publication-ready graphics may be saved into several vector- (PostScript, PDF, and WMF) and bitmap-based (e.g., PNG, TIF, and JPG) formats, ready to be imported into a professional graphical, presentation, or desktop publishing software.

Importantly, the graphical templates are used as a basis for classification. The general classification routine looks for the name of the polygon within the diagram (= graphical template), into which the analysis falls according to its x–y coordinates. The outcome may be either the name of a mineral or a link to another diagram in the case of more complex classification schemes. Following the rules of the International Mineralogical Association (IMA), in some cases, the classification is not done graphically but using prescribed algorithms.

The class mechanism in R provides an elegant solution to the computational problems presented by the differing requirements of each mineral group. By assigning each mineral species to a particular class, all algorithms may be implemented as mutually independent but mineral group-specific methods. The default recalculation options for each mineral class are stored externally in a small and simple text file.

The program is designed to cater to three potential user groups. For users with no familiarity with R, the program is fully menu-driven and contains embedded default recalculation options for many common rock-forming minerals. More experienced users may easily tweak these parameters, as they are saved in a logically structured plain text file. Seasoned R users may invoke GCDkit.Mineral in command line mode, use batch scripts or Python-driven notebooks (e.g., of project Jupyter), or modify and develop new recalculations or plugins.

The lucid, open, and modular design thus makes GCDkit.Mineral a versatile workbench for everyday use, as well as a promising platform for community-driven development. The GCDkit family of R tools, including GCDkit.Mineral, is distributed through the WWW. The current version may be downloaded from http://mineral.gcdkit.org.

Keywords: Mineral composition, software, R language, mineral formula calculation, free and open-source software (FOSS)