R — an alternative to spreadsheets and special software for geochemical calculations and plotting

Vojtěch JANOUŠEK

Czech Geological Survey, Klárov 3, 118 21 Praha 1, Czech Republic

The interpretation of whole-rock geochemical data from igneous and metamorphic rocks often requires complex and time-consuming calculations. Nowadays, these are commonly performed on a personal computer using dedicated software (e.g. MinCalc — Melín & Kunst 1992, NewPet — Clarke 1993, and MinPet — Richard 1995). However, unless either detailed documentation is provided, or the user is skilled in the particular programming language, it is complicated to figure out exactly which algorithm has been employed by the author. Moreover, any modifications to the original program are usually difficult or even impossible as it is not a common practice to make the source code available to the public.

To eliminate these major drawbacks, a freeware QuickBasic 4.5 package for recalculation of major-element analyses, named NORMAN, has been developed (Janoušek 2000). It is capable of computing parameters according to the most common normative calculation schemes and consists of several core modules designed for importing and editing of the input data, selecting an appropriate calculation scheme, display, printing and saving the output data. The calculation algorithms are stored in nearly independent modules with a simple structure. The advantage of the chosen approach is the simplicity, availability and open architecture that enables an average user not only to follow, but also, if need be, to modify the algorithms applied to his data.

However, NORMAN is essentially limited to DOS/WIN 95/98 and QuickBasic is difficult and time-consuming to program, especially for graphic outputs — hence NORMAN lacks these. Writing a system based on a spreadsheet such as MS Excel would not eliminate these setbacks either, as for more complicated calculations the spreadsheet tends to become too complex and prone to errors; moreover, Excel's plotting capabilities are limited and far from being of a publication quality.

A viable alternative seems to be R, a system for statistical computation and graphics (Ihaka and Gentleman 1996) that is based on statistical programming language termed S (Becker *et al*. 1988). R is a very high level language, which means that the generated code is short and — compared with other languages — relatively easily understandable. It is also completely platform independent, as R implementations are available for the most common operating systems (Unix, WIN 95/98/NT). Additionally, R produces high-quality graphic output (e.g., PostScript, HPGL, WMF) and has tools for interaction with the plots that, for instance, make possible identification of the plotted points or their interactive selection. Apart from built-in basic arithmetic, matrix, database and statistical functions, there are numerous add-on packages available (see Hornik 2000 for a complete up-to-date list and further details on R). Last but not least, R is free ☺ .

To date, a great majority of the NORMAN modules have been ported into the new environment. In addition, most of the common geochemical plots used for interpretation of igneous rocks (e.g., Harker diagrams, AFM, spider diagrams, REE plots) as well as numerous classification diagrams (e.g., TAS, R_1–R_2, geotectonic diagrams for granitoids and basaltoids) are now ready to use. Functions for forward and reverse modelling of main petrogenetic igneous petrogenetic processes (fractional crystallization, binary mixing, AFC) utilising major- and trace-element data and/or radiogenic (Sr–Nd) isotopic compositions

have been also written as were basic modules to interpret these (calculating initial Sr and Nd compositions, Nd model ages etc.). The aim is to build, — partly on the basis of analogous S-based software written for Sun compatible computers by Farrow (1991) — a single, coherent and platform-independent system for interpretation of whole-rock geochemical data that would have high-level plotting capabilities. It should be straightforward to use by ordinary geochemists but, at the same time, easily expandable by the more demanding ones.

This work was supported by the Czech Grant Agency Grant 205/97/P113; the author is indebted to C. M. Farrow (Computing Service, University of Glasgow) for intoxicating with the idea using S / R in geochemical calculations and for invaluable discussions.

BECKER R.A., CHAMBERS J.M. and WILKS A.R., 1988. The New S Language. Chapman & Hall, London.

CLARKE D., 1993. NewPet for DOS. Accessed November 15, 1999, at URL ftp://www.esd.mun.ca/pub/geoprogs/np940107.exe.

FARROW C.M., 1991. GCDX (Geochemical Data System) v. 2.0 — Reference manual. Dept. of Geology & Applied Geology, University of Glasgow.

HORNIK K., 2000. The R FAQ. Accessed January 28, 2000, at URL http://www.r-project.org/doc/FAQ/R-FAQ.html.

IHAKA R. and GENTLEMAN R., 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics,* 5: 299–344.

JANOUŠEK V., 2000. NORMAN, a QuickBasic programme for petrochemical re-calculation of whole-rock major-element analyses on IBM PC. *J. Czech Geol. Soc.,* in print.

MELÍN M. and KUNST M., 1992. MinCalc development kit 2.1. Geol. ústav Akademie věd, Praha.

RICHARD L.R., 1995. MinPet: Mineralogical and petrological data processing system, version 2.02. MinPet Geological Software, Québec, Canada.