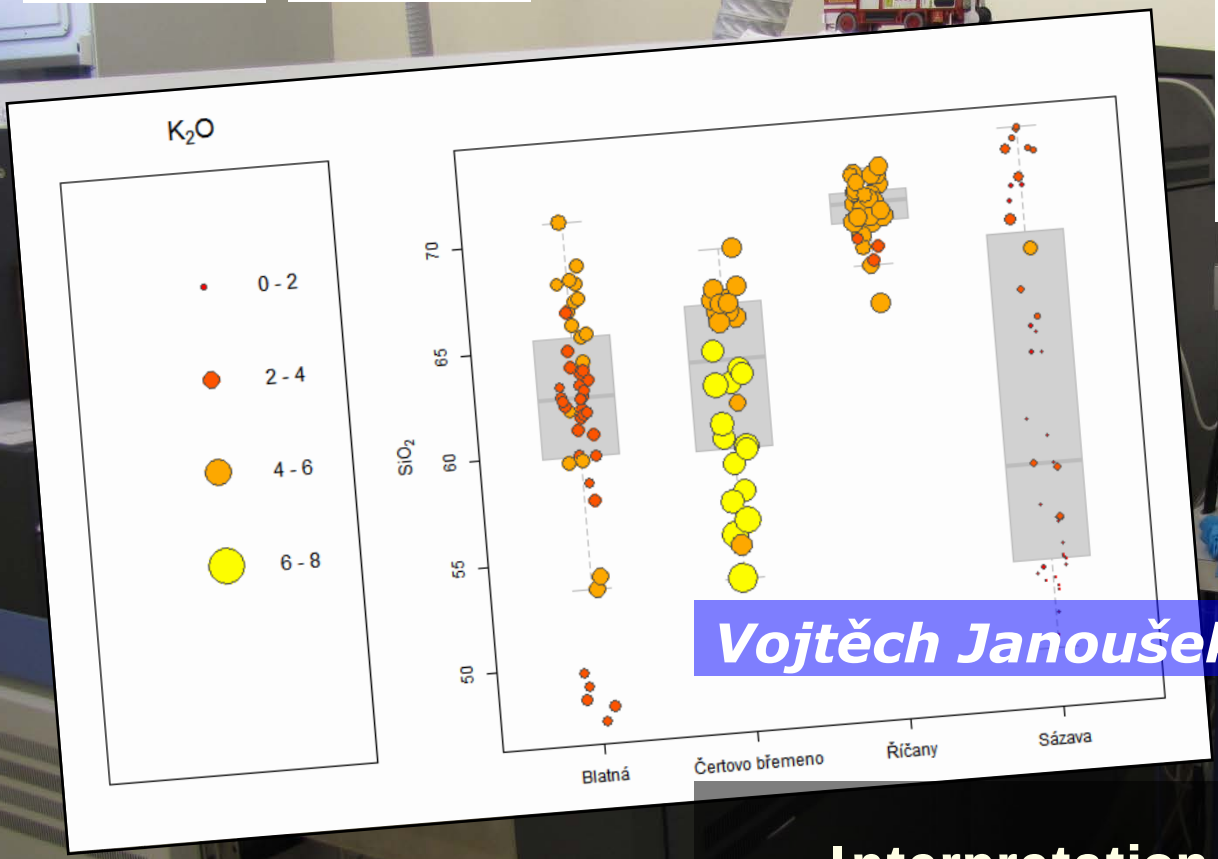
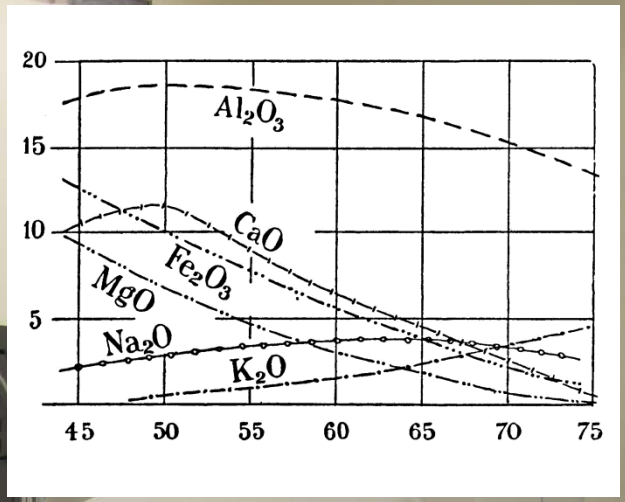
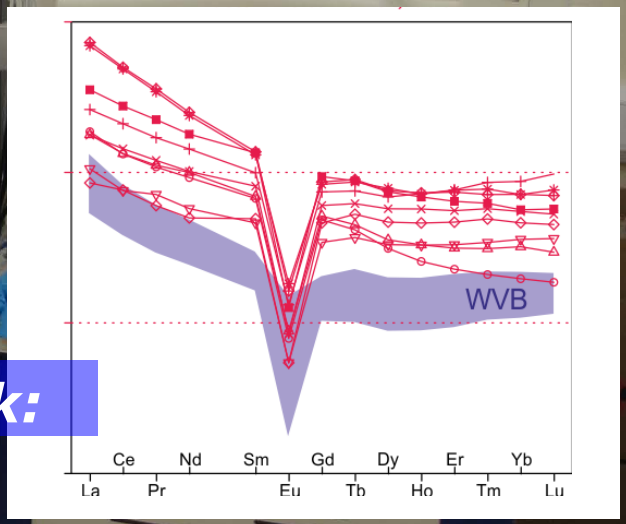


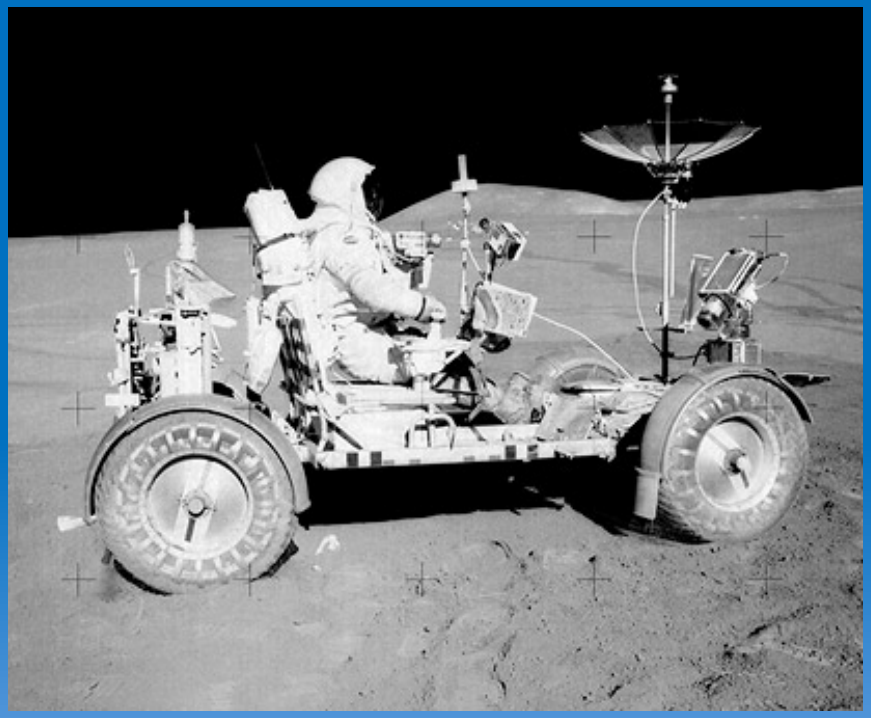




Vojtěch Janoušek:

Interpretation of the whole-rock geochemical data using the R language

The challenge: interpretation of whole-rock chemical data



- Lunar program in the late 1960's came with the requirement of precise and accurate chemical, and then isotopic, analyses of small samples.
- In 1970's appeared novel analytical techniques for trace-element determinations (e.g., XRF; INAA), later ICP-OES, ICP-MS.
- Advancement of radiogenic isotope methods (originally TIMS, then SIMS, ICP-MS).

- Downside: the current flood of precise geochemical data needs to be interpreted by a potent and widely available software tool.

A revolution? The R Language

- Originally designed by Ihaka and Gentleman (1996)
- Version 1.0 published on 29 Feb 2000
- Based on syntax of the S language (Becker *et al.* 1988) which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues.

- The commercial version, S-PLUS, is currently being distributed by TIBCO Software (*www.tibco.com*)
- Since 1997 the development of the R project is overseen by an open group of experts, the R Development Core Team (*http://www.r-project.org*)
- R is Free Software distributed under the Free Software Foundation's GNU General Public License
- R is available for all main OS. It runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

A revolution? The R Language

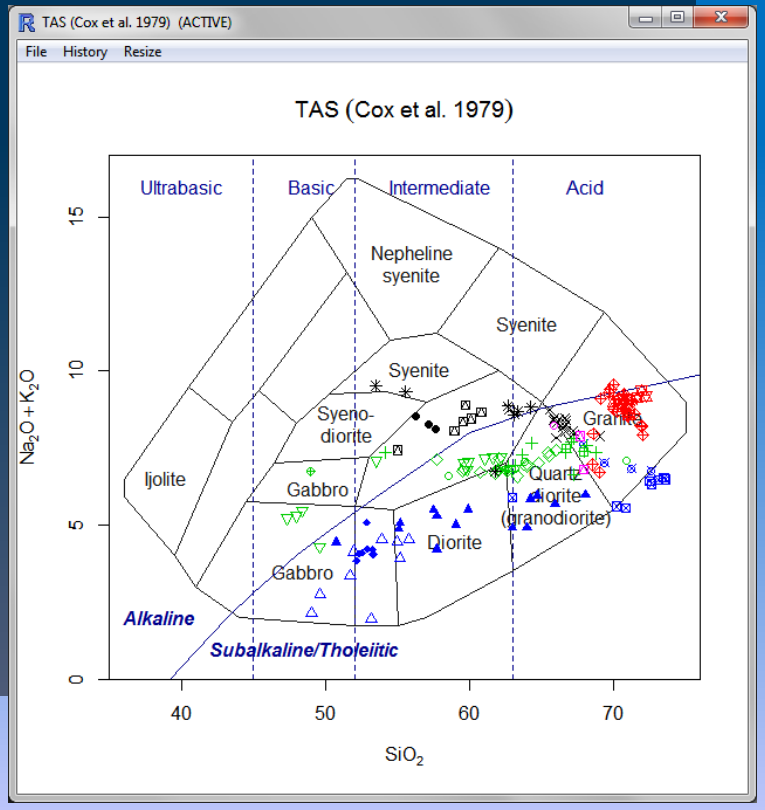
- Frequently updated, supported by a large and still growing community
- R is highly extensible by numerous additional packages
- Large collection of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and database tools
- Data import in many formats, also via SQL and ODBC
- Graphical facilities for data exploration and plotting
- High-level graphical output, well-designed publication-quality plots can include mathematical symbols and formulae where needed
- Graphics can be exported into many formats (*PostScript, WMF, PDF, TIFF, PNG...*), for incorporation into DTP programs, word processors or further editing
- Excellent control over individual functions [= power]
- R allows interactive as well as batch use (as a true programming language).
- Effective object-oriented programming language
- Most of the R system is written in R; for computationally intensive tasks, C, C++, and FORTRAN code can be linked and called at run time.

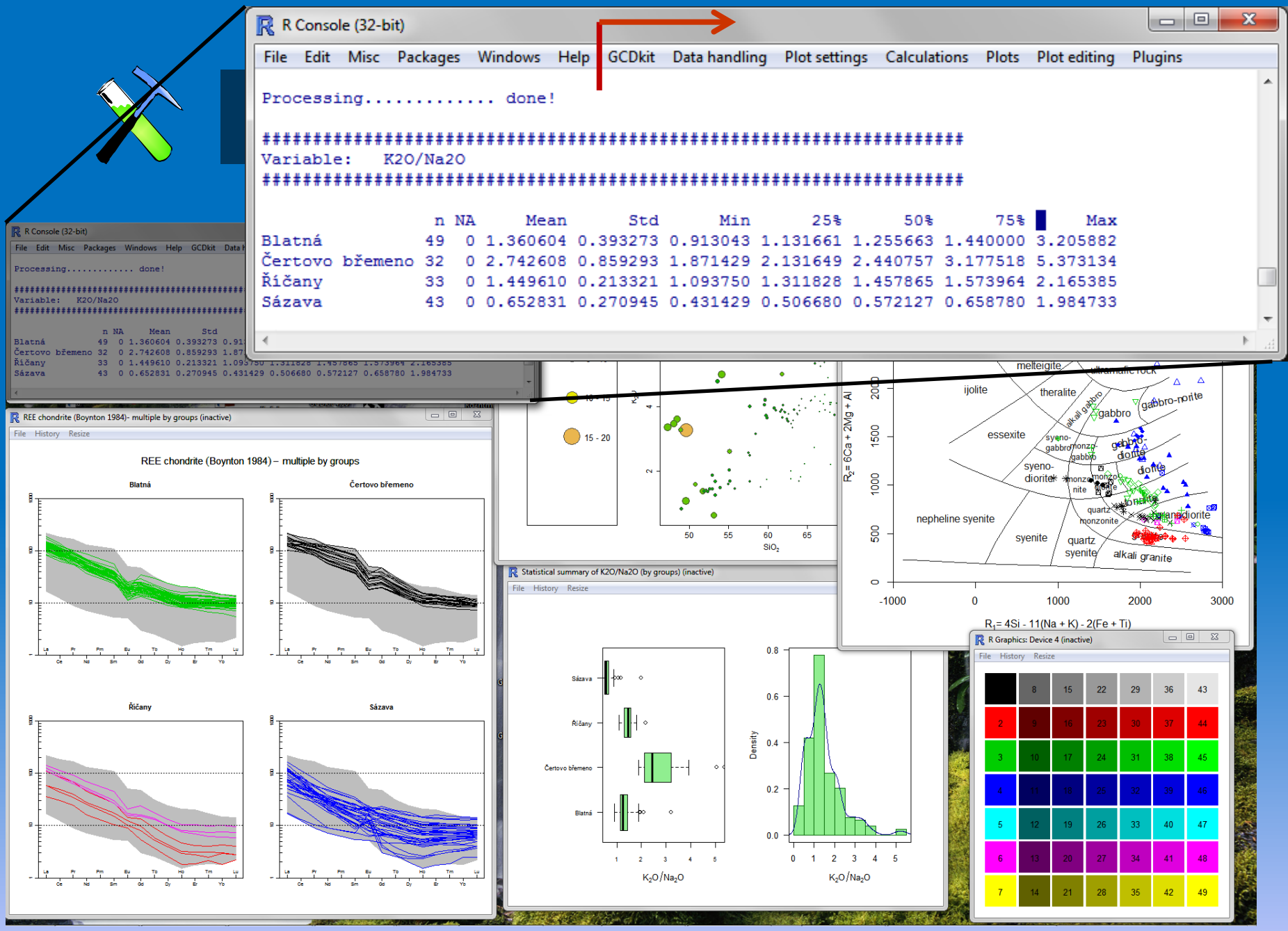
Geochemical Data Toolkit = GCDkit

http://www.gcdkit.org

MAIN FEATURES:

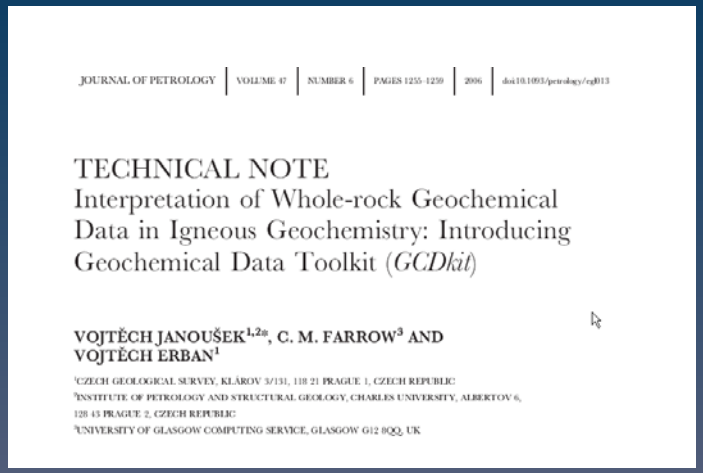
- A more human (less inhuman) interface to the wealth of functions in R
- Windows-like GUI = no programming necessary!!
- Data ready for further handling under R (dot prompt veterans)
- Standard geochemical calculations involving whole-rock major-, trace-element data and Sr-Nd isotopes
- Effective data management (searching, subsetting, grouping)
- Common plots (binary, ternary, spider, classification, geotectonic...)
- Publication quality graphic output




R Console (32-bit)
File Edit Misc Packages Windows Help GCDkit Data handling Plot settings Calculations Plots Plot editing Plugins
Processing..... done!
##########
Variable: K2O/Na2O
##########
n NA Mean Std Min 25% 50% 75% Max
Blatná 49 0 1.360604 0.393273 0.913043 1.131661 1.255663 1.440000 3.205882
Čertovo břemeno 32 0 2.742608 0.859293 1.871429 2.131649 2.440757 3.177518 5.373134
Říčany 33 0 1.449610 0.213321 1.093750 1.311828 1.457865 1.573964 2.165385
Sázava 43 0 0.652831 0.270945 0.431429 0.506680 0.572127 0.658780 1.984733
REE chondrite (Boynton 1984)- multiple by groups (inactive)
REE chondrite (Boynton 1984) – multiple by groups
Blatná
Čertovo břemeno
Říčany
Sázava
15 - 20
SiO2
R2= 6Ca + 2Mg + Al
R1= 4Si - 11(Na + K) - 2(Fe + Ti)
melteigite
ijolite
theralite
gabbro
gabbro-norite
essexite
syeno-gabbro
monzo-gabbro
diorite
syeno-diorite
monzonite
quartz monzonite
adamellite
granodiorite
nepheline syenite
syenite
quartz syenite
alkali granite
Statistical summary of K2O/Na2O (by groups) (inactive)
Density
K2O/Na2O
R Graphics: Device 4 (inactive)

Main ‘milestones‘

- 2000 – launched graduate-level courses on interpretation of geochemical data using R (Masaryk University in Brno & Charles University in Prague)
- 2003 – Goldschmidt Conference, Kurashiki, Japan – *GCDkit* 1.0 released
- 2006 key publication in *Journal of Petrology* (327 hits on WOS)
- October 2015 – Monograph on Geochemical modelling in R/GCDkit (Springer Verlag)
- October 2016 – brand new graduate-level course on data analysis using R and Python (Charles University in Prague)
- August 2019 – last stable version (6.0, *Il padrino sono yo)* for R 3.6.0 unleashed


JOURNAL OF PETROLOGY | VOLUME 47 | NUMBER 6 | PAGES 1255–1259 | 2006 | doi:10.1093/petrology/egl013

TECHNICAL NOTE

Interpretation of Whole-rock Geochemical Data in Igneous Geochemistry: Introducing Geochemical Data Toolkit (*GCDkit*)

VOJTĚCH JANOUŠEK[1,2*], C. M. FARROW[3] AND VOJTĚCH ERBAN[1]

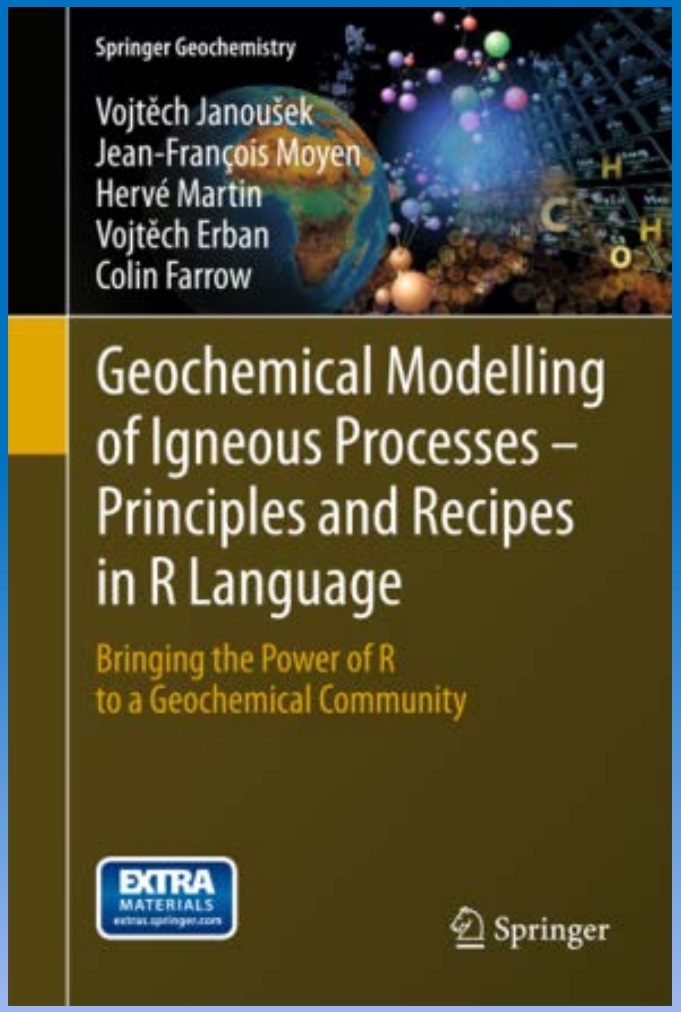
[1]CZECH GEOLOGICAL SURVEY, KLÁROV 3/131, 118 21 PRAGUE 1, CZECH REPUBLIC
[2]INSTITUTE OF PETROLOGY AND STRUCTURAL GEOLOGY, CHARLES UNIVERSITY, ALBERTOV 6, 128 43 PRAGUE 2, CZECH REPUBLIC
[3]UNIVERSITY OF GLASGOW COMPUTING SERVICE, GLASGOW G12 8QQ, UK

Invited workshops on *GCDkit* and/or R modeling

- Czech Geological Survey, Prague, (11 June 2004),
- TU Bergakademie Freiberg, Germany (16 Oct 2005),
- CGS/EOST, Université Louis Pasteur Strasbourg, France (23–24 Oct 2008),
- University of Tromsø, Norway (16–17 June 2010),

- Université Jean Monnet, Saint-Etienne, France (9–11 May 2011),
- University of Helsinki, Finland (7–11 Nov 2011),
- University of Stellenbosch, South Africa (19–23 Mar 2012),
- National Geophysical Research Institute, Hyderabad, India (12–15 Jan 2013),
- University of Arba Minch, Ethiopia (2–6 April 2015),
- Polish Academy of Sciences, Kraków, Poland (23–27 Nov 2015),
- University of Catania, Sicily, (26–30 Sept 2016, 24–27 Sept 2018),...

“The R Book” – philosophy

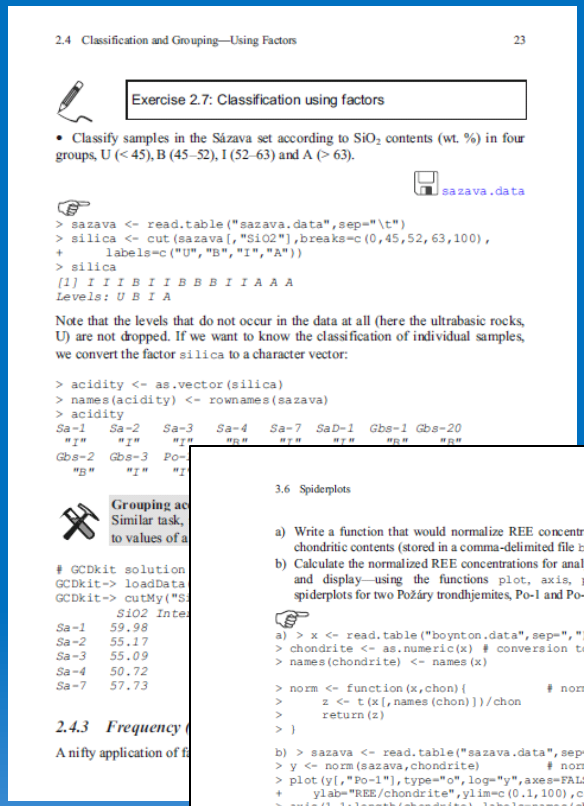


- Provides basics of R language and its application to geochemical problems,
- Gives the first comprehensive introduction to the *GCDkit* system,
- Explains fundamentals of numerical modelling of igneous processes,
- Shows not only formulae, but also the successful modelling strategies,
- Includes numerous worked examples how geochemical modelling helps us to understand geological problems.

Springer Geochemistry series, vol. 1
345 pp., 332 illus., 86 illus. in colour

“The R Book” – contents

- **Part I: R/GCDkit at work**
  - Loading and manipulating data
  - Linking Whole-Rock Chemistry with Mineral Stoichiometry
  - Statistics
  - Classification and Grouping
  - Classical Plots (binary, Harker, ternary, spider)
  - Specialized Plots (log–log, specialized spiderplots, contour plots, anomaly plots…)
  - Radiogenic isotopes (initial ratios, epsilon values, model ages, isochrons…)



2.4 Classification and Grouping—Using Factors 23

Exercise 2.7: Classification using factors

- Classify samples in the Sázava set according to SiO_2 contents (wt. %) in four groups, U (< 45), B (45–52), I (52–63) and A (> 63).

sazava.data

```
> sazava <- read.table("sazava.data",sep="\t")
> silica <- cut(sazava[,"SiO2"],breaks=c(0,45,52,63,100),
+       labels=c("U","B","I","A"))
> silica
[1] I I I B I I B B I I A A A
Levels: U B I A
```

Note that the levels that do not occur in the data at all (here the ultrabasic rocks, U) are not dropped. If we want to know the classification of individual samples, we convert the factor `silica` to a character vector:

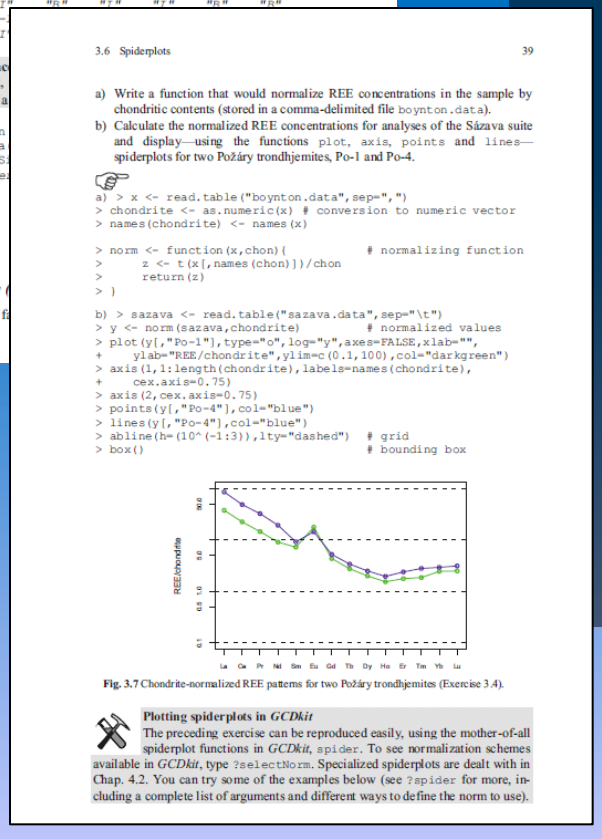
```
> acidity <- as.vector(silica)
> names(acidity) <- rownames(sazava)
> acidity
Sa-1  Sa-2  Sa-3  Sa-4  Sa-7 SaD-1 Gbs-1 Gbs-20
 "I"   "I"   "I"   "B"   "I"   "I"   "B"   "B"
Gbs-2 Gbs-3 Po-
 "B"   "I"  "I
```

Grouping ac
Similar task,
to values of a

```
# GCDkit solution
GCDkit-> loadData
GCDkit-> cutMy("S
         SiO2 Inte
Sa-1    59.98
Sa-2    55.17
Sa-3    55.09
Sa-4    50.72
Sa-7    57.73
```

2.4.3 Frequency

A nifty application of f



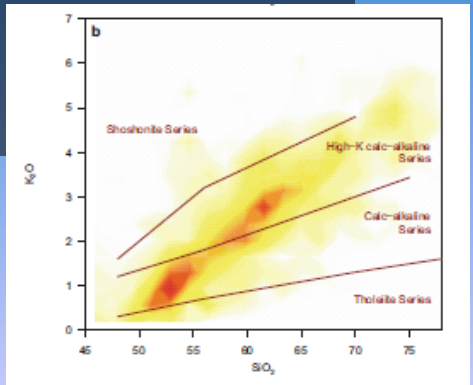
3.6 Spiderplots 39

a) Write a function that would normalize REE concentrations in the sample by chondritic contents (stored in a comma-delimited file `boynton.data`).
b) Calculate the normalized REE concentrations for analyses of the Sázava suite and display—using the functions `plot`, `axis`, `points` and `lines`—spiderplots for two Požáry trondhjemites, Po-1 and Po-4.

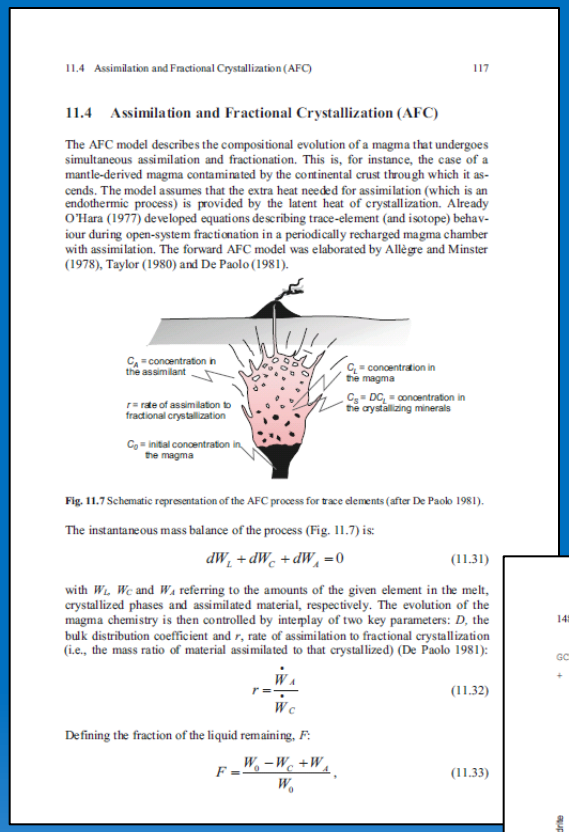
```
a) > x <- read.table("boynton.data",sep=",")
> chondrite <- as.numeric(x) # conversion to numeric vector
> names(chondrite) <- names(x)

> norm <- function(x,chon){                # normalizing function
>      z <- t(x[,names(chon)])/chon
>      return(z)
> }

b) > sazava <- read.table("sazava.data",sep="\t")
> y <- norm(sazava,chondrite)              # normalized values
> plot(y[,"Po-1"],type="o",log="y",axes=FALSE,xlab="",
+    ylab="REE/chondrite",ylim=c(0.1,100),col="darkgreen")
> axis(1,1:length(chondrite),labels=names(chondrite),
+      cex.axis=0.75)
> axis(2,cex.axis=0.75)
> points(y[,"Po-4"],col="blue")
> lines(y[,"Po-4"],col="blue")
> abline(h=(10^(-1:3)),lty="dashed")   # grid
> box()                                # bounding box
```


Fig. 3.7 Chondrite-normalized REE patterns for two Požáry trondhjemites (Exercise 3.4).

Plotting spiderplots in *GCDkit*
The preceding exercise can be reproduced easily, using the mother-of-all spiderplot functions in *GCDkit*, `spider`. To see normalization schemes available in *GCDkit*, type `?selectNorm`. Specialized spiderplots are dealt with in Chap. 4.2. You can try some of the examples below (see `?spider` for more, including a complete list of arguments and different ways to define the norm to use).



“The R Book” – contents



11.4 Assimilation and Fractional Crystallization (AFC) 117

11.4 Assimilation and Fractional Crystallization (AFC)

The AFC model describes the compositional evolution of a magma that undergoes simultaneous assimilation and fractionation. This is, for instance, the case of a mantle-derived magma contaminated by the continental crust through which it ascends. The model assumes that the extra heat needed for assimilation (which is an endothermic process) is provided by the latent heat of crystallization. Already O'Hara (1977) developed equations describing trace-element (and isotope) behaviour during open-system fractionation in a periodically recharged magma chamber with assimilation. The forward AFC model was elaborated by Allègre and Minster (1978), Taylor (1980) and De Paolo (1981).

Fig. 11.7 Schematic representation of the AFC process for trace elements (after De Paolo 1981).

The instantaneous mass balance of the process (Fig. 11.7) is:

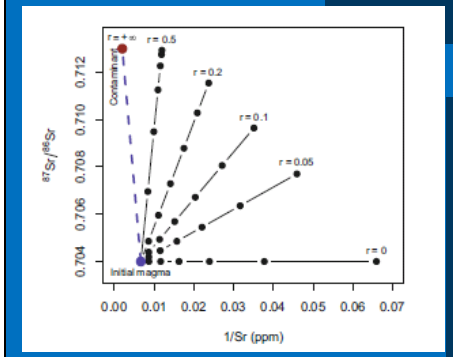
$$dW_L + dW_C + dW_A = 0 \qquad (11.31)$$

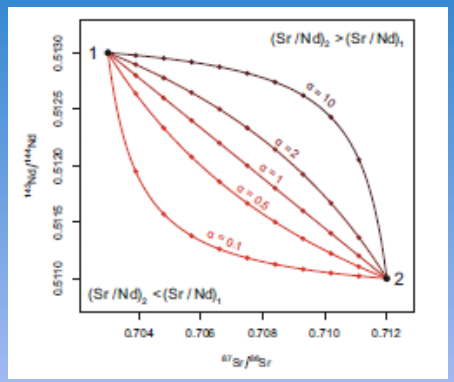
with W_L, W_C and W_A referring to the amounts of the given element in the melt, crystallized phases and assimilated material, respectively. The evolution of the magma chemistry is then controlled by interplay of two key parameters: D, the bulk distribution coefficient and r, rate of assimilation to fractional crystallization (i.e., the mass ratio of material assimilated to that crystallized) (De Paolo 1981):

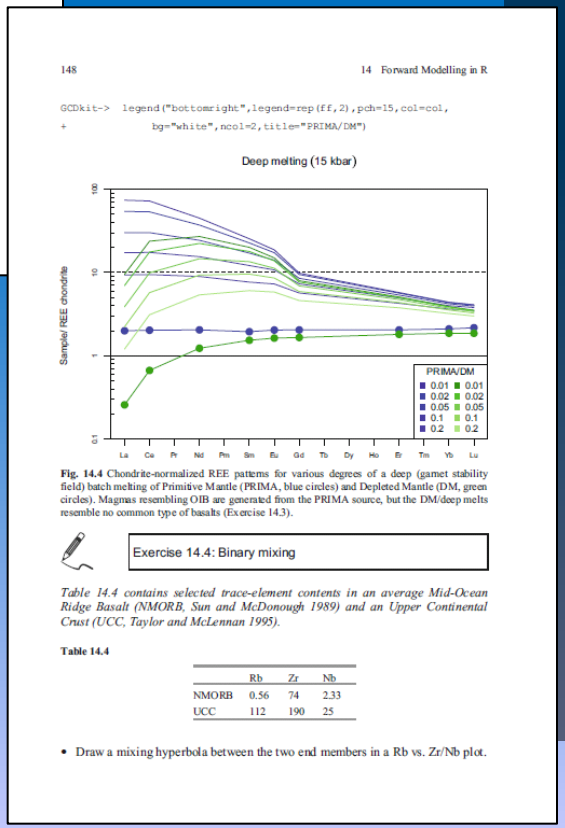
$$r = \frac{\dot{W}_A}{\dot{W}_C} \qquad (11.32)$$

Defining the fraction of the liquid remaining, F:

$$F = \frac{W_0 - W_C + W_A}{W_0}, \qquad (11.33)$$







148 14 Forward Modelling in R

```
GCDkit-> legend("bottomright",legend=rep(ff,2),pch=15,col=col,
+              bg="white",ncol=2,title="PRIMA/DM")
```


Fig. 14.4 Chondrite-normalized REE patterns for various degrees of a deep (garnet stability field) batch melting of Primitive Mantle (PRIMA, blue circles) and Depleted Mantle (DM, green circles). Magmas resembling OIB are generated from the PRIMA source, but the DM/deep melts resemble no common type of basalts (Exercise 14.3).

Exercise 14.4: Binary mixing

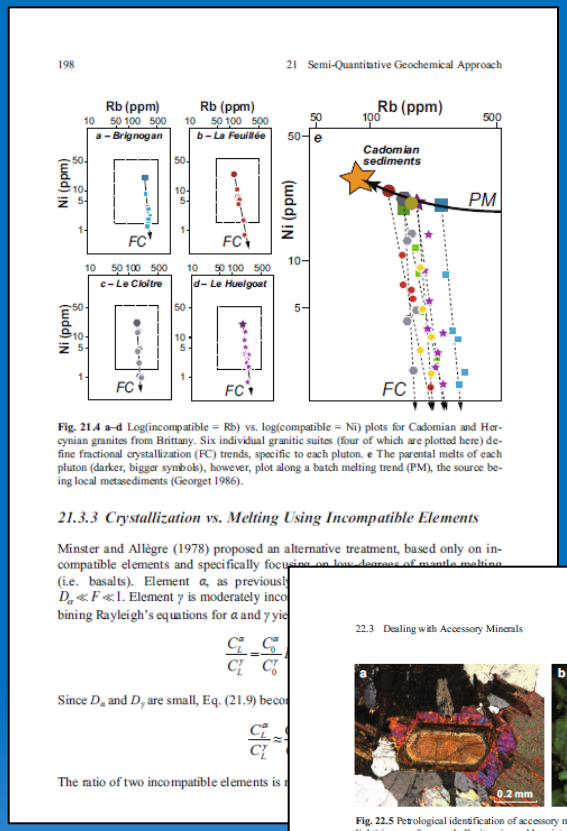
Table 14.4 contains selected trace-element contents in an average Mid-Ocean Ridge Basalt (NMORB, Sun and McDonough 1989) and an Upper Continental Crust (UCC, Taylor and McLennan 1995).

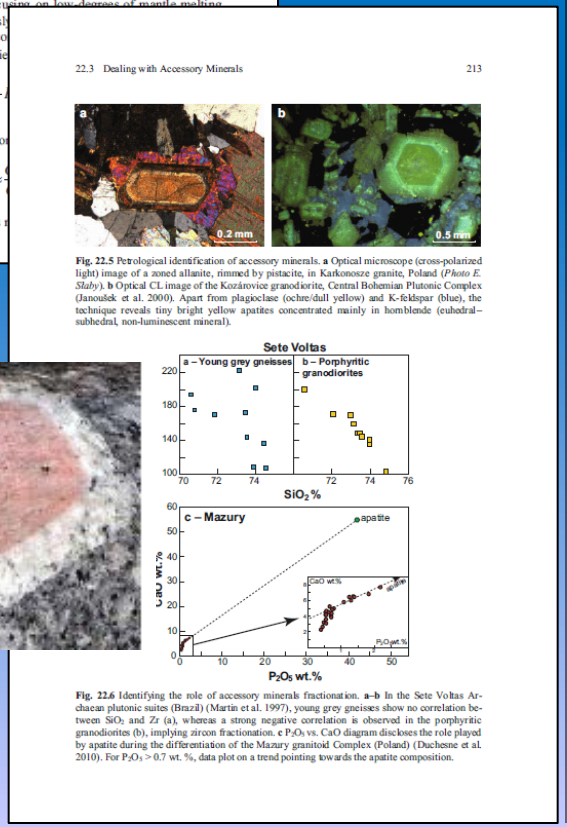
Table 14.4

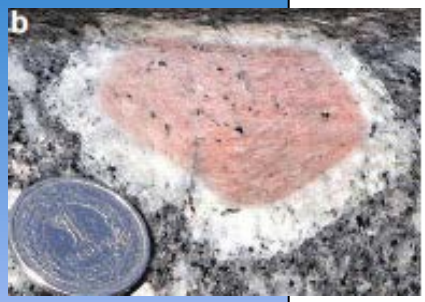
	Rb	Zr	Nb
NMORB	0.56	74	2.33
UCC	112	190	25

- Draw a mixing hyperbola between the two end members in a Rb vs. Zr/Nb plot.

- **Parts II–IV: Majors, traces, radiogenic isotopes**
  - Core of the book
  - Explains fundamentals of each direct and reverse modelling, including the relevant formulae
  - Then introduces the numerical solution and its implementation in the R language
  - Includes a number of real numerical problems
  - Each is presented as a numerical receipt with solution in R (± *GCDkit*)







"The R Book" – contents

- **Part V: Practical Modelling**
- Choosing an Appropriate Model (evidence for crystallization, partial melting, magma mixing and assimilation…)
- Semi-Quantitative Approach (assessing the trace-element compatibility, process identification, mixing test…)
- Constraining a Model (using appropriate strategy, input parameters for the model, partition coefficients, dealing with accessories…)
- Numerical Tips and Tricks (reducing system, colinearity, breaking minerals to end-members, coupling majors and traces…)
- Common Sense in Action (thermodynamic, rheological constraints, scale/speed of processes, how well can we distinguish between models, dangerous projections…)

“The R Book” – contents

- **Part VI: Worked Examples**
  - Differentiation of a Calc-Alkaline Series: Atacazo-Ninahuilca volcanoes, Ecuador
  - Progressive Melting of a Metasedimentary Sequence: the Saint-Malo Migmatitic Complex, France
- Appendix A: R Syntax in a Nutshell
- Appendix B: Introduction to *GCDkit*
- Appendix C: Solving Systems of Linear Algebraic Equations in R

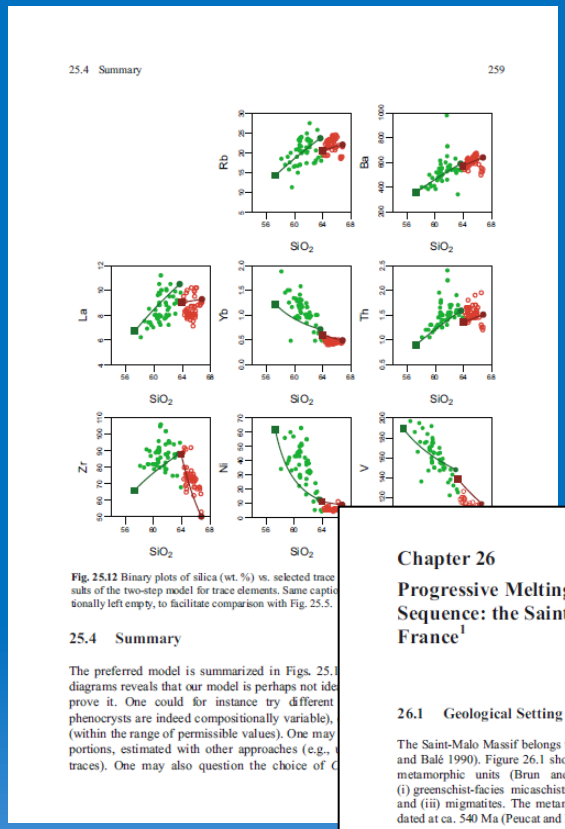
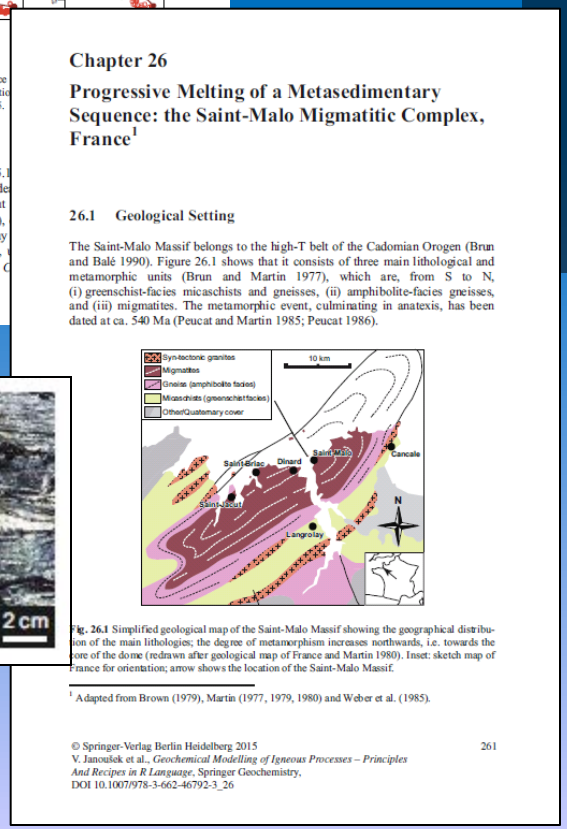


Fig. 25.12 Binary plots of silica (wt. %) vs. selected trace
sults of the two-step model for trace elements. Same captio
tionally left empty, to facilitate comparison with Fig. 25.5.

25.4 Summary

The preferred model is summarized in Figs. 25.1
diagrams reveals that our model is perhaps not ide
prove it. One could for instance try different
phenocrysts are indeed compositionally variable),
(within the range of permissible values). One may
portions, estimated with other approaches (e.g.,
traces). One may also question the choice of C



Chapter 26

Progressive Melting of a Metasedimentary Sequence: the Saint-Malo Migmatitic Complex, France[1]

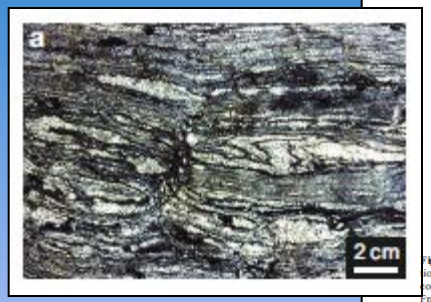
26.1 Geological Setting

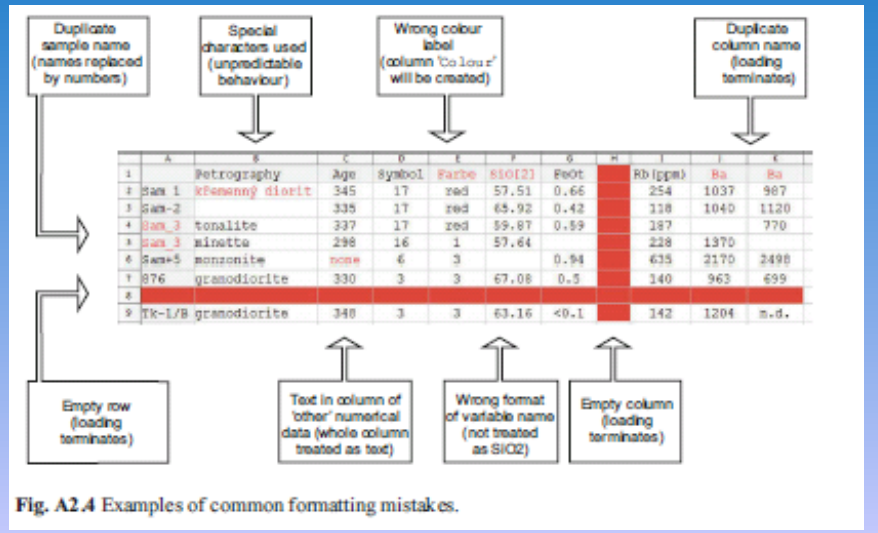
The Saint-Malo Massif belongs to the high-T belt of the Cadomian Orogen (Brun and Balé 1990). Figure 26.1 shows that it consists of three main lithological and metamorphic units (Brun and Martin 1977), which are, from S to N, (i) greenschist-facies micaschists and gneisses, (ii) amphibolite-facies gneisses, and (iii) migmatites. The metamorphic event, culminating in anatexis, has been dated at ca. 540 Ma (Peucat and Martin 1985; Peucat 1986).

Fig. 26.1 Simplified geological map of the Saint-Malo Massif showing the geographical distribution of the main lithologies; the degree of metamorphism increases northwards, i.e. towards the core of the dome (redrawn after geological map of France and Martin 1980). Inset: sketch map of France for orientation; arrow shows the location of the Saint-Malo Massif.

[1] Adapted from Brown (1979), Martin (1977, 1979, 1980) and Weber et al. (1985).

© Springer-Verlag Berlin Heidelberg 2015 261
V. Janoušek et al., *Geochemical Modelling of Igneous Processes – Principles And Recipes in R Language*, Springer Geochemistry,
DOI 10.1007/978-3-662-46792-3_26





	A	B	C	D	E	F	G	H	I	J	K
1		Petrography	Age	Symbol	Farbe	SiO[2]	FeOt		Rb (ppm)	Ba	Ba
2	Sam 1	kŘemenný diorit	345	17	red	57.51	0.66		254	1037	987
3	Sam-2		335	17	red	65.92	0.42		118	1040	1120
4	Sam_3	tonalite	337	17	red	59.87	0.59		187		770
5	Sam_3	minette	298	16	1	57.64			228	1370	
6	Sam+5	monzonite	none	6	3		0.94		635	2170	2498
7	876	granodiorite	330	3	3	67.08	0.5		140	963	699
8											
9	Tk-1/B	granodiorite	348	3	3	63.16	<0.1		142	1204	n.d.

Fig. A2.4 Examples of common formatting mistakes.

Web links

http://www.r-project.org
http://www.gcdkit.org
http://blog.gcdkit.org
http://book.gcdkit.org

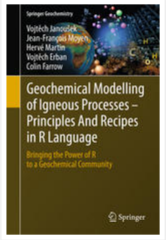
GCDkit forum

News, hints, comments and more from GCDkit team.

Tuesday, 27 October 2015

Book on geochemical modelling of magmatic processes in R/GCDkit

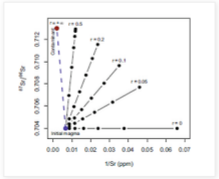
We break the two years of our silence with some news potentially interesting to anyone involved in interpretation of whole-rock geochemical data from igneous rocks. We are pleased to announce that, after two years of intense writing, we have finally published a book:



Janoušek, V., Moyen, J. F., Martin, H., Erban, V. & Farrow, C. M. (2016). *Geochemical Modelling of Igneous Processes – Principles and Recipes in R Language. Bringing the Power of R to a Geochemical Community*. Springer-Verlag, Berlin, Heidelberg, 346 pp.

doi:10.1007/978-3-662-46792-3

Abstract from the publisher:



"The aim of this book is to unlock the power of the freeware R language to advanced university students and researchers dealing with whole-rock geochemistry of (meta-) igneous rocks. The first part covers data input/output, calculation of commonly used indexes and plotting in R. The core of the book then focuses on the presentation and practical implementations of modelling techniques used for fingerprinting processes such as partial melting, fractional crystallization, binary mixing or AFC using major-, trace-element and radiogenic

Search This Blog

Search

Blog Archive

- 2015 (1)
  - October (1)
    - Book on geochemical modelling of magmatic processe...
- 2013 (8)

About Us

- Jeff
- Vojtech Janousek
- Vojta Erban
- Colin

Labels

how-tos (3) new features (3) release (3) batch mode (2) code (2) patch (2) book (1) fractional crystallization (1) graphic example (1) graphics (1) igneous rocks (1) localization (1) magma mixing (1) major elements (1) modelling (1) numerical modelling (1) partial melting (1) radiogenic isotopes (1) spiderplots (1) tables (1) trace elements (1) workshops (1)

The R Project for Statistical Computing

[Home]

Download

CRAN

R Project

About R
Logo
Contributors
What's New?
Reporting Bugs
Development Site
Conferences
Search

R Foundation

Foundation
Board
Members
Donors
Donate

Help With R

Getting Help

Documentation

Manuals
FAQs
The R Journal

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To **download R**, please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.

News

- **The R Journal Volume 8/1** is available.
- The **useR! 2017** conference will take place in Brussels, July 4 - 7, 2017, and details will be appear here in due course.
- **R version 3.3.1 (Bug in Your Hair)** has been released on Tuesday 2016-06-21.
- **R version 3.2.5 (Very, Very Secure Dishes)** has been released on 2016-04-14. This is a rebadging of the quick-fix release 3.2.4-revised.
- **Notice XQuartz users (Mac OS X)** A security issue has been detected with the Sparkle update mechanism used by XQuartz. Avoid updating over insecure channels.
- The **R Logo** is available for download in high-resolution PNG or SVG formats.
- **useR! 2016**, hase taken place at Stanford University, CA, USA, June 27 - June 30, 2016.

http://rseek.org/
http://r.789695.n4.nabble.com

References and further reading

BECKER R.A., CHAMBERS J.M., WILKS A.R. (1988) The new S language. Chapman & Hall, London ("Blue book")

CHAMBERS J.M. (1998) Programming with data. Springer, New York ("Green book")

CHAMBERS J.M. (2009) Software for data analysis: programming with R. Springer, Berlin

CHAMBERS, J. M., HASTIE, T. J. (1992) Statistical Models in S. Chapman & Hall, London ("White book").

CRAWLEY M.J. (2007) The R book. John Wiley & Sons, Chichester

IHAKA R., GENTLEMAN R. (1996) R: a language for data analysis and graphics. *J. Comp. Graph. Stat.* **5:** 299–344

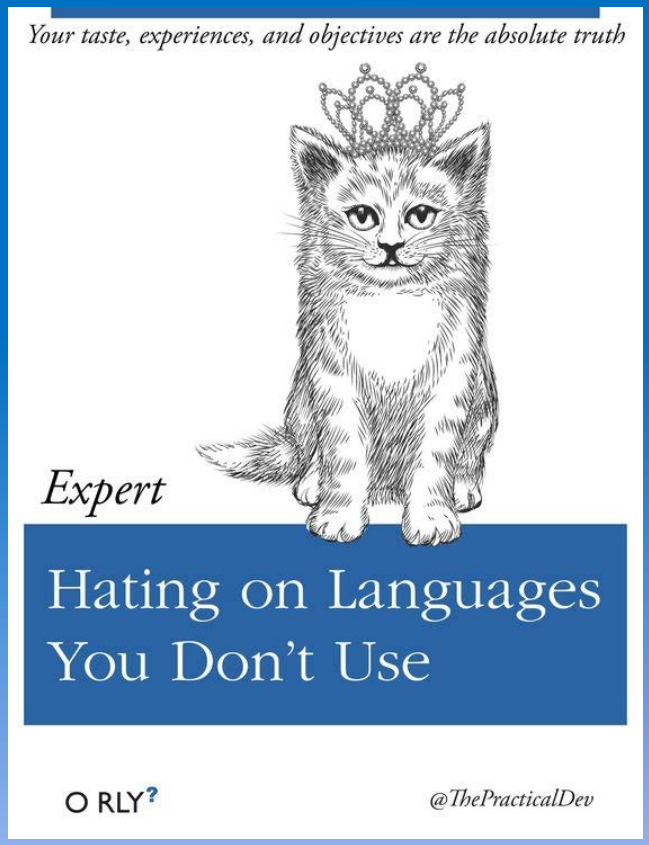
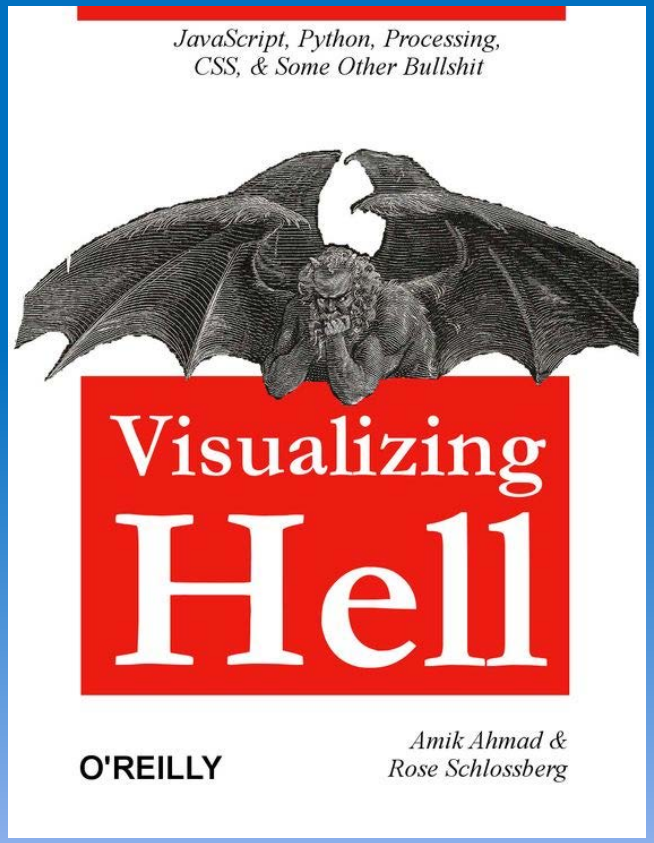
JANOUŠEK V., FARROW C.M., ERBAN V. (2006) Interpretation of whole-rock geochemical data in igneous geochemistry: introducing Geochemical Data Toolkit (GCDkit). *J. Petrol.* **47:** 1255–1259

JANOUŠEK V., MOYEN J.F., MARTIN H., ERBAN V., FARROW C.M. (2016) Geochemical modelling of igneous processes – principles and recipes in R language. Bringing the power of R to a geochemical community. Springer, Berlin.

MAINDONALD J., BRAUN J. (2003) Data analysis and graphics using R. Cambridge University Press, Cambridge

MURRELL P. (2005) R Graphics. Chapman & Hall/CRC, London

REIMANN C., FILZMOSER P., GARRETT R., DUTTER R. (2008) Statistical data analysis explained: applied environmental statistics with R. John Wiley & Sons, Chichester

VAN DEN BOGAARD P., TOLOSANA-DELGADO R. (2013) Analyzing compositional data with R. Springer, Berlin

References and further reading