Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Results of geochemical analyses of paleovolcanic rocks in Uzbekistan using the Petrolog software

Zokir D. Yo'ldoshev

yuldoshevz1995@gmail.com

Institute of Geology and Geophysics named after Kh.M. Abdullaev

Usmonjon D. Mamarozikov

University of Geological Sciences

Ahmadali H. Hamrayev

Institute of Geology and Geophysics named after Kh.M. Abdullaev

Nuriddin E. Saitov

Institute of Geology and Geophysics named after Kh.M. Abdullaev

Utkir A. Normatov

Institute of Geology and Geophysics named after Kh.M. Abdullaev

Zilola A. Keldibekova

Institute of Geology and Geophysics named after Kh.M. Abdullaev

Research Article

Keywords: geochemical data processing, ICP-MS, rare earth elements, reproducibility, paleovolcanic rocks, Uzbekistan

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11269431/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Title page

Results of geochemical analyses of paleovolcanic rocks in Uzbekistan using the Petrolog software

Zokir D. Yoldoshev[1], Usmonjon D. Mamarozikov[2,1], Ahmadali H. Hamrayev[1], Nuriddin E. Saitov[1,3], Utkir A. Normatov[1], Zilola A. Keldibekova[1]

[1] Institute of Geology and Geophysics named after H.M. Abdullaev, Tashkent, Uzbekistan

[2] University of Geological Sciences, Tashkent, Uzbekistan

[3] Tashkent State Technical University named after Islam Karimov, Tashkent, Uzbekistan

Zokir D. Yoldoshev — ORCID: 0009-0004-2428-9002

Corresponding author: Zokir D. Yoldoshev, yuldoshevz1995@gmail.com

Abstract

Geochemical archives in regional institutes hold large numbers of analyses whose processing history is no longer recoverable: the reference values behind a diagram, the background and anomaly thresholds, and the treatment of below-limit values are rarely recorded. We present the trace-element module of Petrolog, an offline desktop system that reads laboratory tables, checks them against the working ranges declared in the protocol, and produces normalisation diagrams in which every computation carries the reference composition, citation and DOI that produced it. The module was applied to archive ICP-MS data from the Qorason ore field in southern Uzbekistan, a Lower Carboniferous paleovolcanic structure with hydrothermal sulphide mineralisation: 32 samples, 62 elements, 1823 determinations. The whole chain, from reading the table to three figures and three result tables, runs in 1.5 s on an ordinary laptop; seven determinations were flagged for the analyst. Chondrite-normalised patterns show light rare earth

enrichment ((La/Yb)N 1.95–36.4, median 12.6) and a negative europium anomaly in every sample (median 0.60), while primitive-mantle patterns show troughs at Nb, Ta, Sr, P and Ti. Clarke coefficients separate a field-wide enrichment in Mo and Se from a local component of Sb, As, Cd, Bi, Cu, Zn and Pb that matches the sulphide assemblage of the same samples. Two values — indium and zirconium — are reported as analytical questions rather than geochemical results, and an error found in the stored reference composition illustrates why reference databases need the same quality control as the analyses themselves.

Keywords: geochemical data processing · ICP-MS · rare earth elements · reproducibility · paleovolcanic rocks · Uzbekistan

1 Introduction

Geochemical analyses age badly. The rock is sampled once, the laboratory reports a column of numbers once, and from then on the data live in spreadsheets that pass from one researcher to the next. In regional geological institutes this archive is substantial: tens of thousands of determinations accumulated over decades of survey work. What usually does not survive is everything around the numbers — which normalisation values were used for a diagram, what counted as background and what as an anomaly, which values were below the determination limit and how they were treated. Two geologists can process the same table and obtain two different sets of diagrams, and neither of them can reconstruct, a year later, how exactly the figures were made.

Software has, of course, been written for this work. GCDkit (Janoušek et al. 2006) gives a comprehensive environment for whole-rock geochemistry in R; pyrolite (Williams et al. 2020) provides the same kind of functionality for Python users, with a well-maintained set of reference compositions; ImageJ and its distribution Fiji (Schneider et al. 2012) cover image measurement, and several specialised packages exist for thin-section images. These are capable tools, and the present work is not an attempt to replace them.

The difficulty is a practical one. All of them assume a user who is comfortable installing an R or Python environment, and most assume a network connection for installation and updates. In a regional laboratory the geologist who needs the diagram is rarely the person who writes code, the working language is not English, and the computer that holds the archive is often not connected to the internet at all. The result is predictable: processing returns to Excel, and the documentation of how a figure was made returns to the memory of whoever made it.

Petrolog was written for that situation. It is an offline desktop application that works with thin-section and polished-section images together with whole-rock major-element and trace-element analyses, in Uzbek, Russian and English, and keeps its data in a local database. It is not a replacement for the expert: the system measures, computes and drafts, and the geologist confirms or corrects. What it adds is bookkeeping — every number it produces carries the reference values, the thresholds and the source behind it.

This paper describes the trace-element part of the system and applies it to archive ICP-MS data from the Qorason ore field in southern Uzbekistan, a Lower Carboniferous paleovolcanic structure with hydrothermal sulphide mineralisation. Thirty-two samples, sixty-two elements. We had four aims:

1. to show how a heterogeneous laboratory table is read, checked against the working ranges declared in the protocol, and converted into a dataset that can be computed on without manual editing;

2. to document the computational chain — normalisation to stored reference compositions, rare earth parameters, clarke coefficients, enrichment factors and the concentration index — in enough detail that it can be reproduced by hand if necessary;

3. to state, for every diagram, which elements and which reference composition it was built from, since a normalisation diagram cannot be interpreted without that information;

4. to show what the resulting diagrams say about the Qorason volcanics, and equally, where the archive data do not support an interpretation.

The fourth point deserves a sentence of its own. Archive data carry no information about the instrument, the reference materials or the replicate precision, and two of the results reported

below — an extreme indium value and anomalously low zirconium — are more likely to reflect the analytical history of the samples than their geochemistry. We report them as the program reported them, with the reasons for caution, rather than quietly dropping them. A system that hides inconvenient numbers is of no use to the geologist who has to defend the interpretation.

The paper is organised as follows. Section 2 describes the ore field, the samples and the analytical data. Section 3 sets out the processing chain and the diagram definitions. Section 4 presents the architecture and the user interface of the system. Section 5 reports the results for Qorason, and Section 6 discusses what they mean for the suite and for the use of archive geochemical data in general.

2 Materials

2.1 The Qorason ore field

The samples analysed here come from the South Qorason ore field in southern Uzbekistan, where a volcano-tectonic structure developed at the junction of two major fault systems: the dislocation zone of the South Tien Shan geosyncline and the Qurama–Hisor transverse uplift. The eastern part of the field is deeply eroded, while the western part is buried under Meso-Cenozoic cover. In the central part of the field, independent volcanic edifices are preserved, among them the Qizilsoy shield volcano, which complicates the eastern limb of the Chaqchar–Qizilsoy syncline.

The Lower Carboniferous succession of the structure belongs to a rhyolite formation that Shoyoqubov (1978) divided into five subformations, from a rhyolite–terrigenous unit at the base, through rhyolite–limestone–dolomite and dacite–rhyolite–andesite–quartz–keratophyre units, to a rhyolite unit and an upper rhyolite–siliceous–limestone unit. In the central part of the Qorason–

Kharkush syncline, Upper Carboniferous (C2) siltstones, argillites and tuffaceous rocks are widespread. Hydrothermal sulphide mineralisation — pyrite, sphalerite, galena, chalcopyrite, molybdenite and arsenopyrite — is hosted by these volcanogenic and volcano-sedimentary rocks, which are variably silicified, sericitised, chloritised and carbonatised; the pyrite–polymetallic ores of southern Uzbekistan and the volcanogenic formations that host them were mapped and described in the regional surveys of the early 1970s (Kolchedanno-polimetallicheskoe orudenenie 1971; Vulkanogennye formatsii Uzbekistana 1971). The rock types sampled are rhyolites, tuff-siltstones, clastolavas, secondary quartzites and tuffisites (Fig. 1).

This is, therefore, a genuinely paleovolcanic object: an eroded Carboniferous volcanic structure in which the primary igneous signature has been partly overprinted by hydrothermal alteration. That combination is what makes the dataset useful for testing an automated workflow — the data are real archive data from an altered terrain, not a clean teaching example.

2.2 Analytical data

Trace elements (ICP-MS). Thirty-two samples of the QN series (QN-1 to QN-30, including QN-3/1, QN-10/1 and QN-26/1) were analysed for 62 elements, from Li to U, including the full rare earth series and the ore-related elements Cu, Pb, Zn, Ag, Mo, As, Sb, W and Bi (Table 1). Sample preparation followed the routine procedure of crushing, pulverising and fusion. The analyses are archive data: the laboratory protocol that accompanies them reports the element concentrations and the working range of the method for each element, but does not record the instrument, the reference materials or the results of replicate analyses. We state this explicitly rather than reconstruct it, and Section 5 discusses which conclusions that limitation does and does not permit. A second worksheet of the same protocol lists the working determination range

for 59 elements (for example 0.05–4000 g/t for Li and Be, 20–200000 g/t for Al), and these ranges are used below as the quality filter.

Of the 1823 determinations in the Qorason table, six fall outside the stated working range (four for Se, one for Ta, one for Ag) and one value is missing (Zn in QN-11). The dataset is therefore essentially complete, and no imputation was required for the diagrams presented here.

Major elements (silicate analysis). Fifteen samples of the same series were analysed for major oxides at the Central Laboratory of JSC Uzbekgeologorazvedka (Eshonguzar, Tashkent region) on commission from the Institute of Geology and Geophysics. The analyses combine photometric and ICP-OES determinations on an ICPE-9820 (Shimadzu) emission spectrometer and a UV-1800 (Shimadzu) spectrophotometer, with national standard procedures (O'z O'U 0178.01:2017 for SiO2, O'z O'U 0178.02:2002 for Al2O3 and related documents for the remaining oxides). Unlike the trace-element data, these analyses carry a full methodological record, and they are used to anchor the classification of the same samples.

Table 1 Datasets used in this study.

Dataset	Samples	Elements	Laboratory and method
ICP-MS, QN series	32	62	Archive data; crushing, pulverising and fusion; working ranges declared in the protocol; instrument and reference materials not recorded
Silicate analysis, QN series	15	major oxides	Central Laboratory of JSC Uzbekgeologorazvedka; photometry and ICP-OES (ICPE-9820 and UV-1800, Shimadzu); national standards O'z O'U 0178.01:2017, 0178.02:2002 and related

3 Methods

Petrolog processes geochemical data in four steps — reading, quality control, computation and plotting — and records, for every number it produces, which reference values it used. The

sequence below describes what the program actually does; every step is covered by the automated test suite.

3.1 Reading laboratory tables

Laboratory protocols are not uniform. Column headers mix Cyrillic and Latin scripts, element symbols carry footnote marks ("Zr *"), decimal separators may be commas, and values below the determination limit are written as text ("<0.05"). The reader therefore normalises headers to plain element symbols, converts decimal commas, and treats non-numeric and below-limit entries as missing rather than as zero. Zeros would be fatal here: all normalisation diagrams use a logarithmic axis.

3.2 Quality control against the working range

Each value is compared with the working range declared in the protocol for that element. Values outside the range are flagged and reported with the element and sample identifier, but are not silently removed; the analyst decides. For the Qorason dataset this step flagged the six values listed in Section 2.2.

3.3 Normalisation and computed parameters

Normalisation is a division by a reference composition, C(sample)/C(reference). Petrolog stores eight reference compositions — chondrite after Sun and McDonough (1989), McDonough and Sun (1995) and O'Neill (2016), primitive mantle and N-MORB after Sun and McDonough (1989), upper continental crust after Rudnick and Gao (2014), PAAS after Taylor and McLennan (1985) and NASC after Gromet et al. (1984) — and three clarke sets. Each entry carries its source, citation and DOI, and the chosen reference is written into every output, so a diagram can

be reproduced from the result file alone. Reference sets that have not been checked against the original publication are marked as such and raise a warning when used.

From the normalised values the program computes the parameters that are conventionally reported for volcanic suites:

- the sum of the rare earth elements and the light (La–Sm) and heavy (Gd–Lu) subtotals, computed on raw concentrations;
- the fractionation ratios (La/Yb)N, (La/Sm)N and (Gd/Yb)N, computed on normalised values;
- the europium anomaly Eu/Eu* = EuN / sqrt(SmN · GdN) and the cerium anomaly Ce/Ce* = CeN / sqrt(LaN · PrN).

As a test of the implementation, feeding the chondrite values themselves back into the program returns Eu/Eu* = 1.00 and (La/Yb)N = 1.00; this check is part of the test suite.

3.4 Clarke coefficient, enrichment factor and concentration index

Enrichment relative to average crust is expressed by the clarke coefficient Kc = C(sample)/C(clarke). Because alteration changes the mass of a sample, Kc may be normalised to an immobile element, giving the enrichment factor EF = Kc(element)/Kc(immobile). The immobile reference is an explicit parameter: zirconium is the default, and aluminium, titanium, thorium, scandium or yttrium can be chosen instead. The overall intensity of enrichment in a sample is summarised by the concentration index Zc = sum(Kc of anomalous elements) − (n − 1), where the anomaly threshold is an explicit parameter (1.5 by default for Zc, 3 for the anomaly list).

One correction to the reference base was required. In the stored table of Rudnick and Gao (2014) the value for sodium was 83 g/t, which is three orders of magnitude below the published upper-

crustal content; it was recomputed from Na2O = 3.27 wt% as 24260 g/t, and the change is recorded in the entry. Before the correction, sodium appeared as the strongest "anomaly" in every Qorason sample — a reminder that a reference database needs the same quality control as the data it is applied to.

3.5 Diagrams and the elements behind them

Each diagram uses a defined element list, a defined reference composition and a defined axis. Table 2 states them explicitly, because a normalisation diagram cannot be read without knowing what went into it.

Table 2 Definition of the diagrams: elements, reference composition and axis.

Diagram	Elements used	Reference	Axis
REE spectrum	La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu (14)	chondrite, Sun and McDonough (1989)	log, sample/chondrite
Multi-element (spider)	Cs, Rb, Ba, Th, U, Nb, Ta, K, La, Ce, Pb, Pr, Sr, P, Nd, Zr, Hf, Sm, Eu, Ti, Gd, Tb, Dy, Ho, Y, Er, Tm, Yb, Lu (29)	primitive mantle, Sun and McDonough (1989)	log, sample/primitive mantle
Clarke coefficient	all elements present in both sample and clarke set (56 for Qorason)	upper continental crust, Rudnick and Gao (2014)	log, Kc

Figures are produced at the resolution the journal requires, with the reference source and its DOI printed beneath each plot, and exported as images and as editable slides.

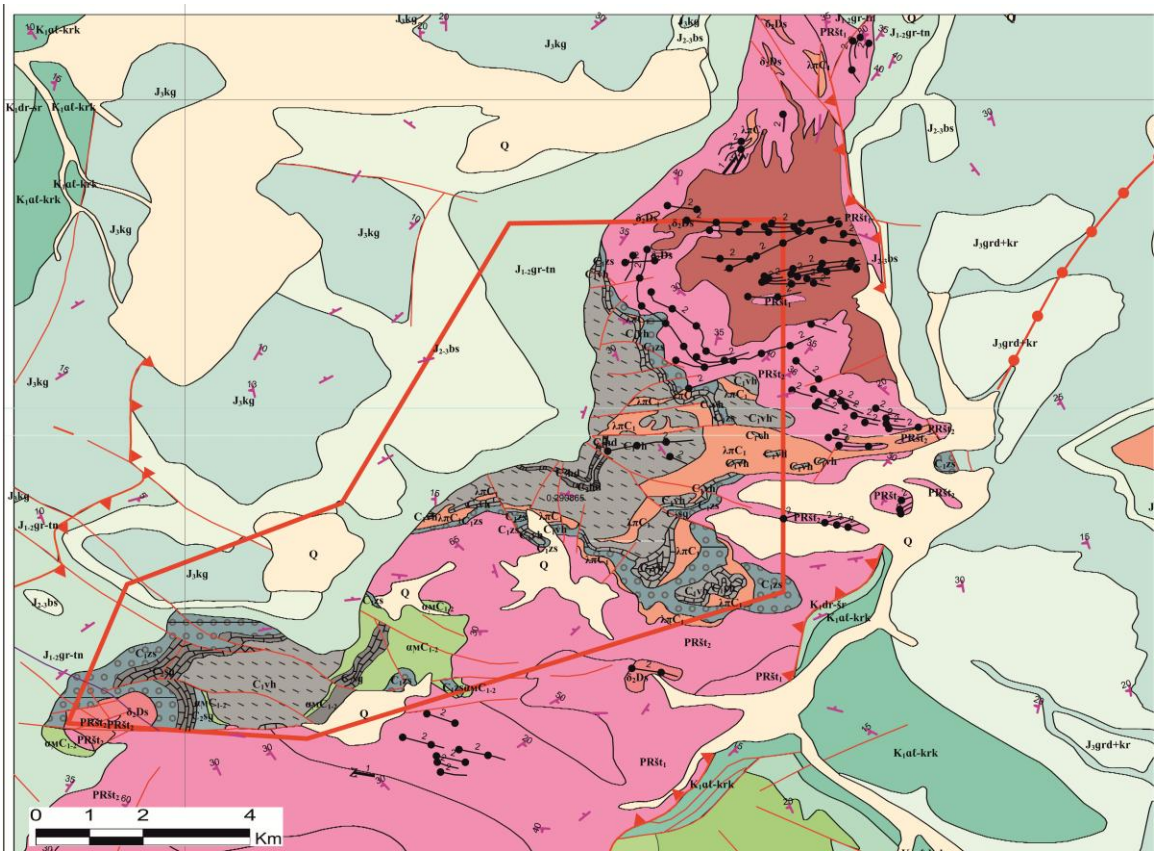


Fig. 1 Geological sketch map of the Qorason ore field (redrawn after E.D. Bezuglov).

4 The system

4.1 Architecture

Petrolog is distributed as a single Windows executable built with PyInstaller. Inside it, a FastAPI application serves a Jinja2 interface that is displayed in a pywebview window, so the program looks and behaves like a desktop application while the same code can also be deployed as a web server. Data are stored locally in an SQLite database under the user's application data folder. No part of the analysis requires a network connection.

The code base comprises about 21,300 lines in 80 modules, with a further 4,800 lines of tests; 539 automated tests are run on each build. Numerical work rests on NumPy, pandas, OpenCV and Matplotlib. The statistical routines that the geochemical modules need — Student's

distribution through the incomplete beta function, Ward clustering via the Lance–Williams update, the random forest used by the grain-level classifier — were implemented directly rather than taken from a larger library. That decision keeps the executable small and, more importantly, makes the output identical on every machine, which matters when a laboratory cannot reinstall its software stack.

The interface is available in Uzbek, Russian and English through 449 translation keys, with complete coverage in all three languages. Fig. 2 shows the start screen: four analyses are offered side by side — thin section, polished section, ICP-MS and silicate analysis — and each runs independently, requiring only its own kind of data. A geologist with a trace-element table alone can use the ICP-MS module without having images, and the reverse is equally true.

This paper is concerned with the trace-element module, but the literature base of the other three is stored in the same way, each rule carrying its source. The silicate module classifies an analysis on the TAS diagram in the form given by Middlemost (1994), with the subalkaline subdivision of Peccerillo and Taylor (1976) and the alkaline–subalkaline boundary of Irvine and Baragar (1971); it computes the alumina saturation index of Shand (1943), the geochemical indices of Frost et al. (2001), the immobile-element discrimination of Winchester and Floyd (1977) for altered rocks, and the CIPW norm in the formulation of Cross et al. (1902). The petrographic modules follow the nomenclature of Le Maitre (2002) and the Petrographic Code of the Republic of Uzbekistan (2024), the microstructural terms of Vernon (2018), the mineral abbreviations of Whitney and Evans (2010), and, for volcaniclastic rocks, the methodological guide of Mamarozikov et al. (2025).

4.2 The trace-element module

The ICP-MS section follows the sequence described in Section 3, with the analyst in control at each step.

Choosing the elements. After the table is read, the program lists every element it recognised and lets the analyst tick those that enter the analysis (Fig. 3). For the Qorason dataset this list contains 61 elements. The interface states the practical consequence of the choice directly: above forty elements the coefficients no longer fit inside the triangular correlation matrix and only the colour scale remains. Small notes of this kind are what make a tool usable by someone who is not a programmer.

Descriptive statistics and thresholds. Because geochemical data are closed — the components sum to a constant — correlations are computed after a centred log-ratio transformation (Aitchison 1986; Egozcue et al. 2003) rather than on raw concentrations. Values below the determination limit are handled explicitly and columns dominated by them are excluded, with the reason written into the report. Background and anomaly thresholds are derived with robust estimators, median + 2 MAD and the Tukey rule (Reimann et al. 2005), and are shown on ranked-value plots so that the analyst can see how many samples sit above the threshold and by how much (Fig. 4a).

Distributions. Histograms with kernel density estimates, Q–Q plots against the log-normal expectation and empirical cumulative distribution functions are produced for every selected element (Fig. 4b–d). These panels are diagnostic rather than decorative: a break in the ECDF curve or a departure from the Q–Q line indicates that the element has more than one population, a common situation in mineralised ground, and that a single threshold will not describe it.

Correlation and clustering. Spearman or Pearson coefficients are computed on the transformed data, p-values are corrected for multiple testing by the Benjamini–Hochberg procedure

(Benjamini and Hochberg 1995), and the elements are ordered by Ward clustering on the distance $1 - \rho$ (Ward 1963). The result is a triangular matrix in which the associations group visibly (Fig. 5); the strongest positive and negative pairs are listed beside it. In the Qorason data this ordering separates the rare earth block from the ore-element association without any instruction from the user.

Normalisation diagrams. The module added for this study computes the rare earth and multi-element patterns and the clarke coefficients described in Section 3.3, and plots them with the reference source and its DOI printed beneath the figure.

4.3 Reporting and export

Every analysis ends in a report that can be read, corrected and exported. The tables leave as CSV, the figures as JPEG at 150, 300 or 600 dpi, and the complete set as an editable PowerPoint file, so that a figure can be adjusted for a thesis chapter or a presentation without redrawing it. Each exported object carries the reference composition, the citation and the DOI that produced it. Reference sets that have not been verified against the original publication are marked in the database and raise a warning in the report whenever they are used.

The reference database itself is treated as data that needs checking. During this study one error was found in it and corrected (Section 3.4). We mention this not as an incidental detail but as part of the method: a system that stores reference values must also make them inspectable, or it simply moves the unverifiable step from the spreadsheet into the software.

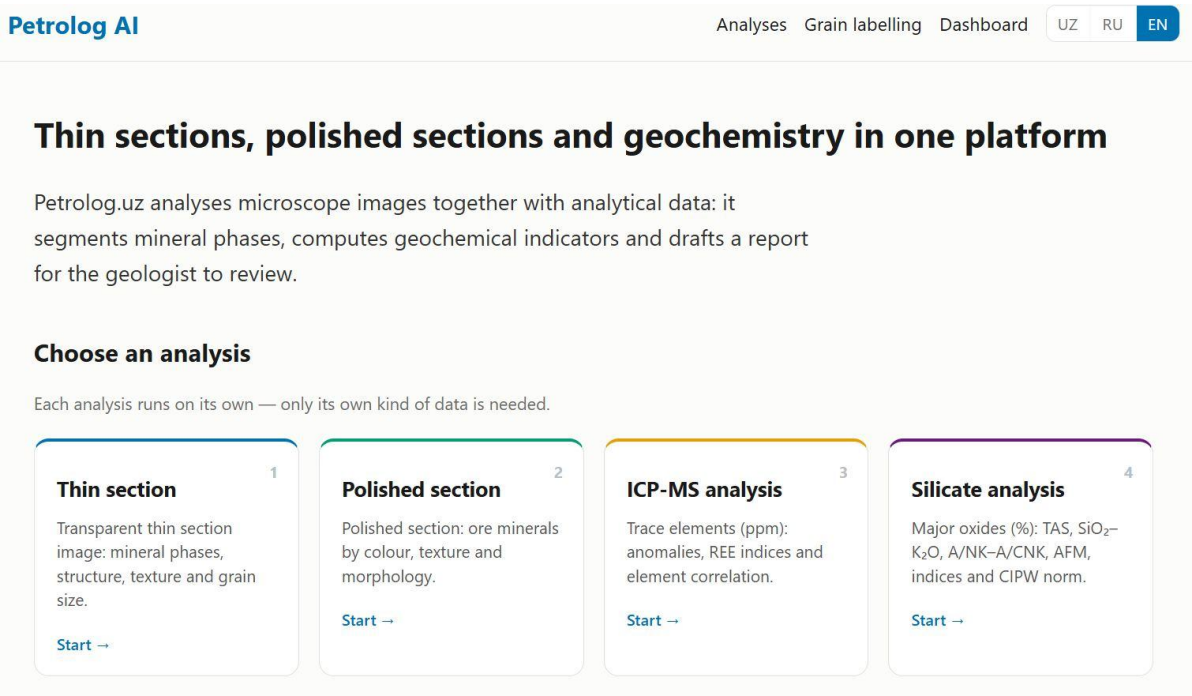


Fig. 2 Start screen of Petrolog. The four analyses — thin section, polished section, ICP-MS and silicate — run independently; the language switch (UZ / RU / EN) is in the upper right.

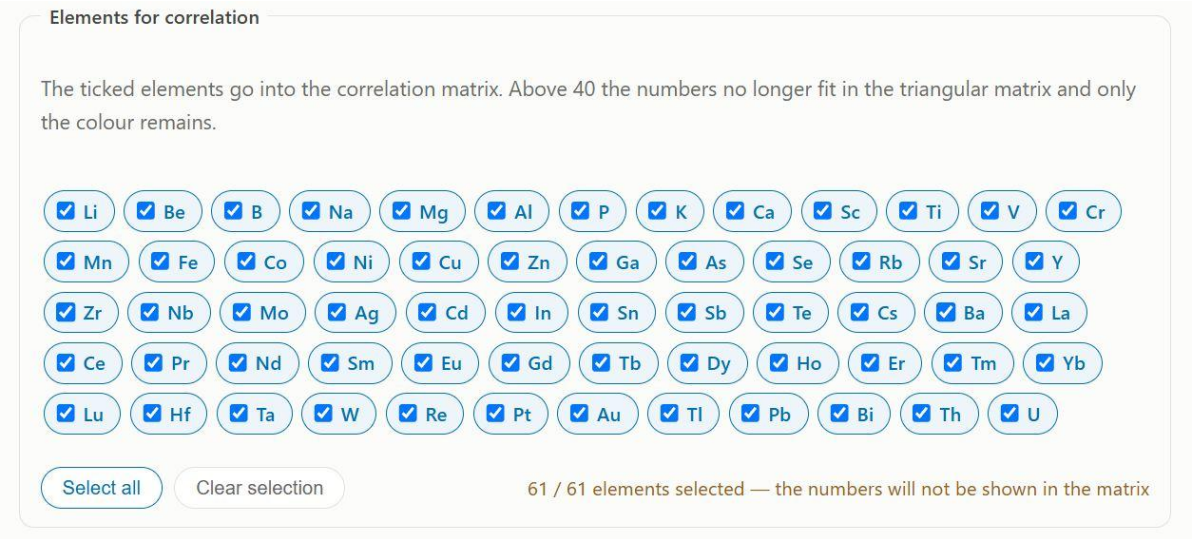


Fig. 3 Element selection panel of the ICP-MS module, with all 61 elements of the Qorason table selected.

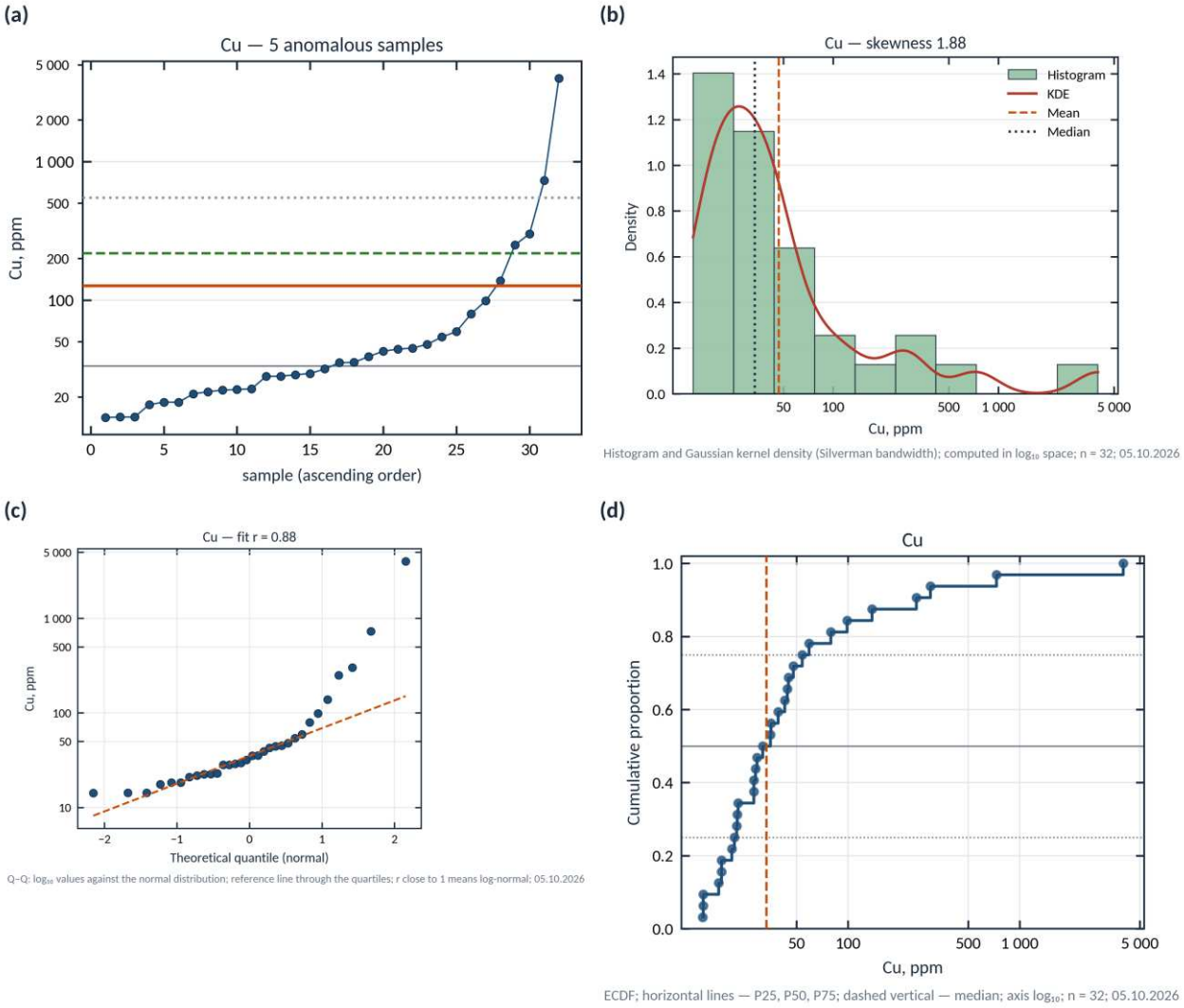


Fig. 4 The four diagnostic panels the ICP-MS module produces for every selected element, shown here for copper: (a) ranked values with the background (thin grey line) and three anomaly thresholds — median + 2 MAD (solid red), Tukey Q3 + 1.5·IQR (dashed green) and mean + 2σ (dotted grey), after Reimann et al. (2005); (b) histogram with a Gaussian kernel density estimate; (c) Q–Q plot against the log-normal expectation; (d) empirical cumulative distribution function. All panels are computed in $\log_{10}$ space.

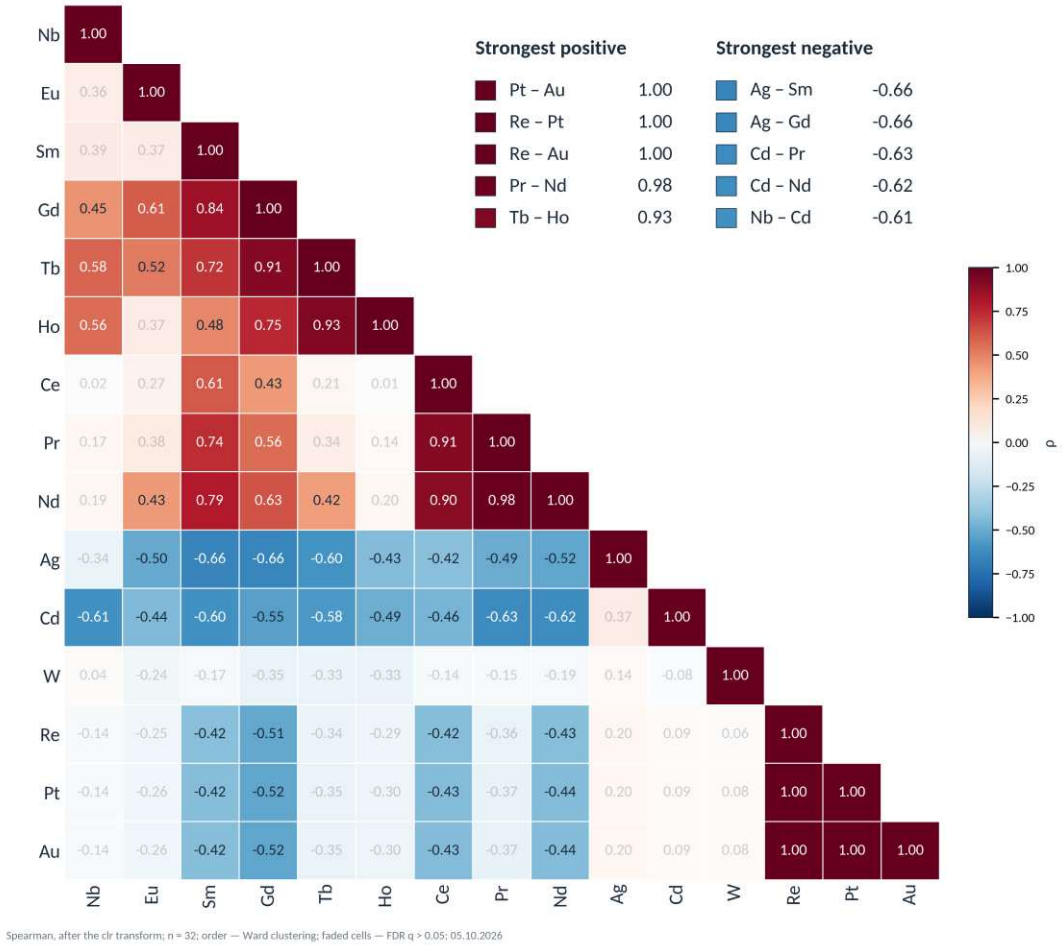


Fig. 5 Triangular correlation matrix for fifteen selected elements of the Qorason dataset: Spearman coefficients on centred log-ratio transformed data, elements ordered by Ward clustering; pairs with FDR-corrected $p > 0.05$ are shown in pale type.

5 Results

All numbers below come from a single run of the program on the Qorason dataset (32 samples of the QN series, 62 elements). The run is reproducible with one command; it takes 1.5 s on an ordinary laptop with no GPU, and it writes three tables and three figures.

5.1 Data preparation

The reader accepted the laboratory table without manual editing. Of 1823 determinations, six lie outside the working range declared in the protocol (four for Se, one each for Ta and Ag) and one value is missing (Zn in QN-11). These seven entries are flagged in the output and excluded from

the diagrams; everything else passed. Below-limit entries were kept as missing values rather than replaced by zeros, so no artificial point appears at the bottom of the logarithmic axes.

5.2 Rare earth elements

Chondrite-normalised patterns were obtained for all 32 samples using the full set of 14 rare earth elements (Fig. 6). The patterns share one shape: light rare earths strongly enriched, a pronounced negative europium anomaly, and a nearly flat heavy rare earth segment. The rare earth parameters computed from these patterns are listed in Table 3.

Table 3 Rare earth parameters of the 32 Qorason samples.

Parameter	Range	Median
ΣREE, g/t	9.5–342.8	111.0
LREE/HREE	2.6–21.5	9.9
(La/Yb)N	1.95–36.4	12.6
(La/Sm)N	1.35–8.94	3.66
(Gd/Yb)N	1.07–4.75	2.02
Eu/Eu*	0.27–0.93	0.60
Ce/Ce*	0.81–1.89	1.01

Two features are worth stating plainly. First, the europium anomaly is negative in every sample without exception, which points to plagioclase fractionation as a process common to the whole suite. Second, (La/Yb)N spans almost a factor of twenty, so the suite is not a single homogeneous magma batch: the samples record different degrees of fractionation, and in the most altered rocks part of that spread may be redistribution rather than magmatic evolution. The cerium anomaly stays close to unity (median 1.01), which is what one expects when weathering has not been severe.

5.3 Multi-element patterns

The spider diagram (Fig. 7) was built from 29 elements normalised to primitive mantle. The patterns show the troughs that are characteristic of subduction-related felsic rocks — niobium,

tantalum, strontium, phosphorus and titanium — together with enrichment in the large-ion lithophile elements caesium, rubidium, barium, thorium and uranium. The strontium trough runs parallel to the europium anomaly seen in the rare earth patterns, which is consistent with both being controlled by plagioclase.

5.4 Enrichment relative to average crust

Clarke coefficients were computed for the 56 elements present in both the samples and the upper-crustal reference. Taking Kc ≥ 3 as the reporting threshold, the program listed 119 anomalies across the 32 samples. Their distribution is clear and geologically meaningful (Fig. 8; Table 4).

Table 4 Elements with a clarke coefficient of three or more.

Element	Samples with Kc ≥ 3	Maximum Kc
Mo	32	34.7
Se	31	93.3
Sb	10	44.5
As	10	—
Cd	8	34.4
Bi	7	39.1
Cu	6	142.9
Zn	4	—
Pb	3	32.1

Molybdenum is enriched in every sample and selenium in all but one; the remaining elements — antimony, arsenic, cadmium, bismuth, copper, zinc, lead — appear in a third or less of the samples. That pattern matches the mineralogy recorded in the polished sections, where pyrite, sphalerite, galena, chalcopyrite, molybdenite and arsenopyrite were identified. The concentration index Zc ranges from 20.5 to 928.4 with a median of 47.1, and the number of elements above the background threshold (Kc ≥ 1.5) per sample ranges from 4 to 41, median 9.

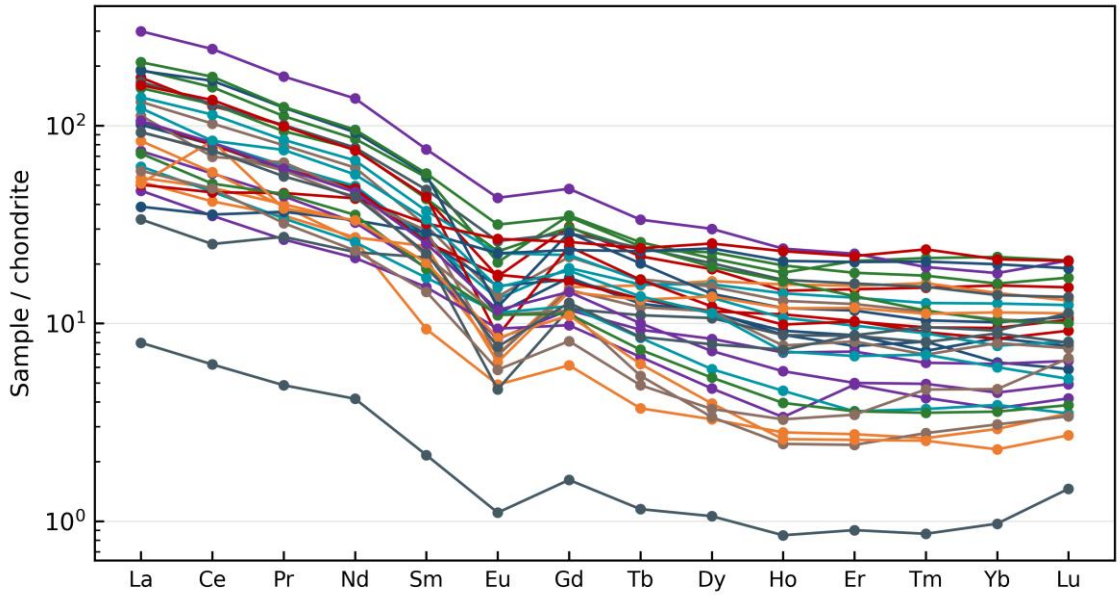
5.5 Two results that need caution, not silence

Indium. The single highest clarke coefficient in the dataset is indium (Kc = 516). In the laboratory protocol indium is one of the elements marked with an asterisk, and a value of that size is more likely to indicate a unit or reporting problem than a genuine enrichment of five hundred times. The program reports it as measured; we flag it here as a value that the laboratory should confirm before it is interpreted.

Zirconium. Zirconium contents are low for felsic volcanic rocks: the median is 14.4 g/t and 22 of 32 samples fall below 20 g/t, whereas rhyolites usually carry more than 100 g/t. Zirconium and hafnium are internally consistent (Zr/Hf = 16.7–36.6, median 26.7), so this is not a defect of one analytical channel; both elements behave as if the zircon fraction was not fully brought into solution. The practical consequence is methodological: zirconium is the program's default immobile reference for the enrichment factor, and in this dataset that choice is unsafe. For Qorason we therefore report clarke coefficients rather than enrichment factors; where an immobile normalisation is required, aluminium, titanium, thorium, scandium or yttrium can be selected in place of zirconium.

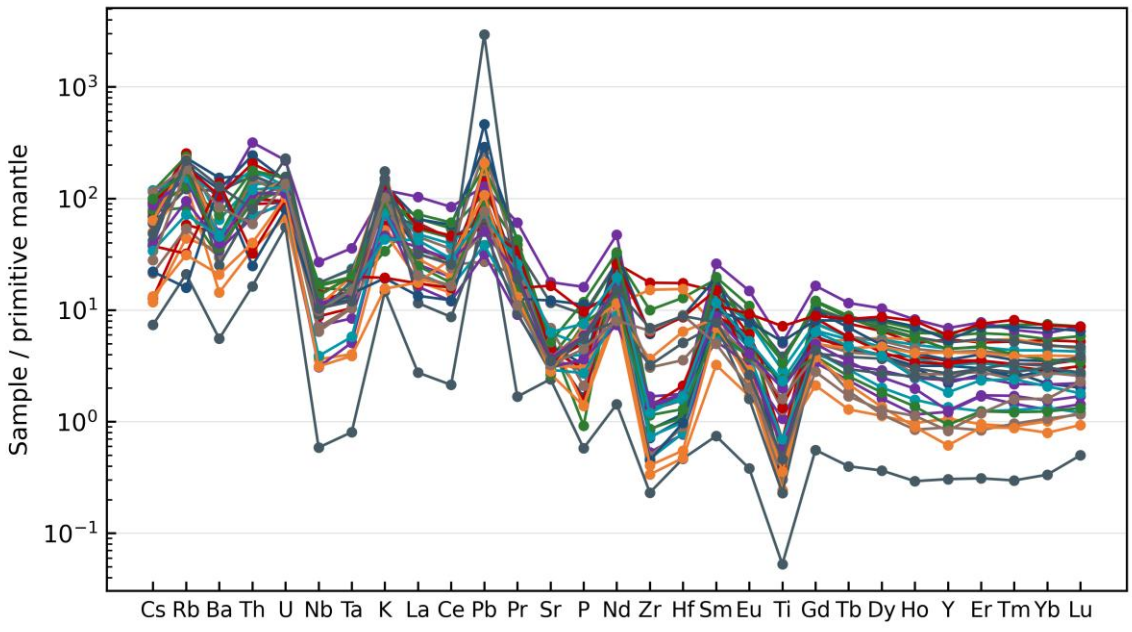
5.6 Reproducibility of the output

Every table and figure carries the reference composition it was computed against, with citation and DOI, and the figures print that source beneath the plot. Rerunning the script on the same input reproduces the numbers exactly, because no step depends on a random seed, and the eleven unit tests written for this module — including the check that feeding the chondrite values back through the program returns Eu/Eu* = 1.00 and (La/Yb)N = 1.00 — pass on each run.



Sun & McDonough (1989) · DOI 10.1144/GSL.SP.1989.042.01.19

Fig. 6 Chondrite-normalised rare earth element patterns of the 32 Qorason samples. Normalising values after Sun and McDonough (1989).



Sun & McDonough (1989) · DOI 10.1144/GSL.SP.1989.042.01.19

Fig. 7 Primitive-mantle normalised multi-element patterns of the 32 Qorason samples, 29 elements. Normalising values after Sun and McDonough (1989).

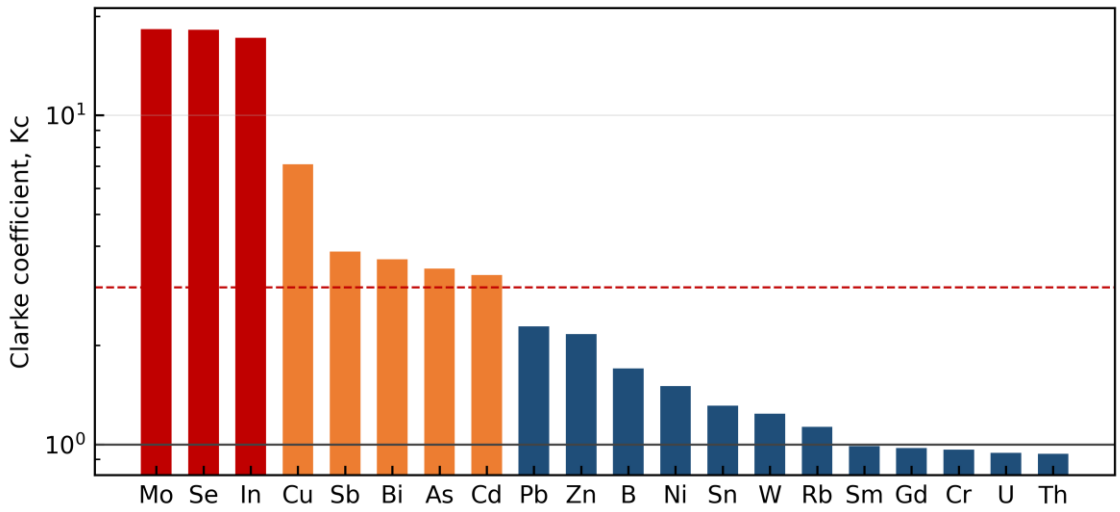


Fig. 8 Clarke coefficients of the mean Qorason composition for the 20 most enriched elements. Reference: upper continental crust after Rudnick and Gao (2014); the dashed line marks Kc = 3.

6 Discussion

6.1 What the diagrams say about the Qorason volcanics

Three features of the trace-element data are consistent across the 32 samples and, taken together, describe a single magmatic lineage rather than a set of unrelated rocks.

The first is the shape of the rare earth patterns: light rare earths enriched relative to heavy ones in every sample, with a median (La/Yb)N of 12.6 and a nearly flat heavy segment. The second is the europium anomaly, negative in all 32 samples, median Eu/Eu* = 0.60. The third is the set of troughs on the primitive-mantle diagram at niobium, tantalum, strontium, phosphorus and titanium, together with enrichment in caesium, rubidium, barium, thorium and uranium.

That combination is the standard signature of felsic magmas produced in a continental-margin setting with crustal involvement, and it agrees with the regional interpretation of the Lower Carboniferous rhyolite formation of southern Uzbekistan (Shoyoqubov 1978). The strontium trough and the europium anomaly track one another, as they should if both are controlled by

plagioclase: europium enters plagioclase in its divalent form and strontium substitutes for calcium, so fractional removal of plagioclase depletes the melt in both. The spread in (La/Yb)N, from 1.95 to 36.4, says that the suite was sampled at different stages of that process; it is not a single homogeneous batch of magma.

Two qualifications belong here. Degrees of fractionation and degrees of alteration are not easy to separate in rocks that are silicified, sericitised, chloritised and carbonatised, and part of the spread in the light rare earths may be hydrothermal redistribution rather than magmatic evolution. And although the cerium anomaly stays near unity (median 1.01), which argues against severe weathering, the archive data give no independent means of checking this.

6.2 Enrichment pattern and the mineralisation

The clarke coefficients separate the elements into two groups with a clarity that is worth noting. Molybdenum exceeds three times the upper-crustal value in every sample and selenium in 31 of 32, whereas antimony, arsenic, cadmium, bismuth, copper, zinc and lead exceed it in a third of the samples or fewer. The first group behaves as a field-wide background of enrichment; the second is local.

This matches what the polished sections show. Molybdenite, pyrite, sphalerite, galena, chalcopyrite and arsenopyrite were identified in the same sample series, and the elements that are locally enriched are precisely the ones hosted by those sulphides. The concentration index Zc, which summarises the total enrichment of a sample, has a median of 46.9 but reaches 927.9, and the number of elements above background per sample ranges from 4 to 35 — in other words, the field contains both weakly altered rock and strongly mineralised rock, and the index orders them without any interpretive input.

For exploration purposes this is the practical output of the workflow: a ranked list of samples and elements that identifies where the sulphide mineralisation is concentrated, produced in seconds from a table that otherwise sits unused in an archive.

6.3 What archive data cannot deliver

The limits of the dataset are as important as its results, and they are not limits of the software.

The ICP-MS analyses carry no record of the instrument, the international reference materials or the replicate precision. Without replicates there is no measure of analytical scatter, so the ranges reported in Section 5 contain both real variation between samples and an unknown analytical component. Any conclusion that depends on small differences between samples is therefore unsafe, and we have avoided drawing such conclusions.

Two specific values show how this plays out. The indium content of one sample gives a clarke coefficient of 516, which is implausible as a geochemical signal and far more likely to be a unit or transcription problem in a column that the protocol marks with an asterisk. Zirconium is low throughout — median 14.4 g/t, with 22 of 32 samples below 20 g/t, against the 100 g/t and more normally found in rhyolites — while the Zr/Hf ratio stays in its usual range (median 26.7). Both elements reside mainly in zircon, and both are low together, which is the pattern expected when the zircon fraction was not fully brought into solution during sample preparation. We cannot verify this from the archive record, but the practical consequence is clear: zirconium cannot serve as the immobile reference element for enrichment factors in this dataset, so we report clarke coefficients only and the software now offers thorium, scandium and yttrium as alternatives.

Reporting these two cases rather than removing them is a deliberate choice. Archive geochemical data are the main resource available to geologists in many regions, and a processing system that silently drops awkward values trains its users to trust numbers they should be questioning.

6.4 The reference database needs the same scrutiny as the data

One error found during this work deserves emphasis because it is a general risk. The stored upper-crustal composition contained 83 g/t for sodium, three orders of magnitude below the published value; every Qorason sample consequently appeared to be enriched in sodium by a factor of several hundred. The error was obvious once a geologist looked at the ranked list, and invisible as long as the number sat inside a JSON file.

Software that normalises data to stored reference compositions moves a judgement that used to be explicit — which values am I dividing by? — into a place where nobody looks. Our response is to keep the source, citation and DOI with every reference set, print them beneath every figure, mark unverified sets so that they raise a warning when used, and record corrections in the entry itself. This is bookkeeping rather than science, but it is the kind of bookkeeping that determines whether a diagram can be trusted ten years later.

6.5 Relation to existing software

GCDkit (Janoušek et al. 2006) and pyrolite (Williams et al. 2020) are more comprehensive than Petrolog in petrological modelling and are supported by larger communities; a user who works in R or Python will find them more flexible. Petrolog differs in three practical respects rather than in analytical ambition. It runs offline as a single executable, which matters where archive computers have no network access. It works in Uzbek, Russian and English, so the geologist who owns the data can use it in the language in which the data were described. And it keeps

petrography and geochemistry in one system, so that a thin-section description, a modal composition and a rare earth pattern for the same sample live in the same database.

The trade-off is equally plain. A fixed desktop application cannot match the flexibility of a scripting environment, and the diagram set is deliberately limited to what is covered by tests. For users who need an unusual plot, the exported CSV files remain the route to any other tool.

7 Conclusions

1. Archive ICP-MS data from a regional laboratory can be turned into a reproducible geochemical dataset without manual editing. For 32 samples and 62 elements the whole chain — reading, quality control against the declared working ranges, normalisation, computation of parameters and production of three publication-quality figures — took 1.5 s on an ordinary laptop, and seven of 1823 determinations were flagged for the analyst.

2. The Qorason paleovolcanic suite shows a consistent signature: light rare earth enrichment ((La/Yb)N median 12.6), a negative europium anomaly in every sample (median 0.60), and troughs at Nb, Ta, Sr, P and Ti. Plagioclase fractionation is common to the whole suite, and the twenty-fold spread in (La/Yb)N indicates that the samples record different stages of it.

3. Enrichment relative to average crust separates a field-wide component (Mo in all samples, Se in 31 of 32) from a local one (Sb, As, Cd, Bi, Cu, Zn, Pb), and the local component matches the sulphide assemblage identified in the polished sections of the same samples.

4. Two results — an indium value of 516 times clarke, and zirconium contents

an order of magnitude below those expected in rhyolites, with a normal Zr/Hf ratio — are reported as analytical questions rather than geochemical findings. They also removed zirconium from use as the immobile reference element for enrichment factors in this dataset.

5. Reference compositions stored inside software require the same quality control as the analyses they are applied to. The sodium error found in the upper-crustal set during this study produced a spurious anomaly in all 32 samples and was only visible because every computation carries its reference values with it.

6. For geologists working with archive material in regional institutes, the value of the system lies less in novel computation than in recording what was computed, from which reference values, and with which elements — which is what allows a figure to be defended, repeated and corrected.

Statements and Declarations

Competing interests. The authors declare no competing interests.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Author contributions. Z.D.Y. conceived the study, developed the Petrolog software, processed and analysed the ICP-MS data, prepared all figures and tables, and wrote the manuscript. U.D.M. supervised the work and contributed to the interpretation of the volcanogenic formations. A.H.H., N.E.S. and U.A.N. contributed to the sampling, to the geological characterisation of the Qorason ore field and to the verification of the petrographic descriptions. Z.A.K. contributed to the preparation and quality control of the analytical data. All authors reviewed and approved the final manuscript.

Ethics approval. Not applicable. This study did not involve human participants or animals.

Data availability. The ICP-MS table of the 32 Qorason samples is openly available, as it came from the laboratory, in the deposit at https://doi.org/10.5281/zenodo.23162644 under CC BY 4.0, together with the computed rare earth parameters, the clarke coefficients and the anomaly lists. A single command in that deposit recomputes every table and figure of this paper from the table. The raw laboratory protocols are the property of the H.M. Abdullaev Institute of Geology and Geophysics and are available from the corresponding author on reasonable request, subject to institutional permission.

Code availability. The geochemical calculation layer on which every number in this paper depends — the reference compositions with their citations and DOIs, the cleaning of raw ICP-MS cells, normalisation, the REE parameters, the clarke coefficient, the enrichment factor, the concentration index, correlation with false-discovery-rate control, and the figure routines — is released as open source under the MIT licence at https://doi.org/10.5281/zenodo.23162644, with its eleven unit tests and the reproduction script. The complete Petrolog desktop application (v1.0.0), of which this layer is one module, is distributed as a Windows x64 executable at https://doi.org/10.5281/zenodo.23161504, free of charge for research and teaching; commercial use requires the author's written permission. Both archives are published with their SHA-256 checksums.

Availability and requirements. Project name: Petrolog. Version: 1.0.0. Operating system: Windows 10 or 11, 64-bit, for the desktop application; the open calculation layer is platform-independent. Programming language: Python 3.12. Other requirements: 4 GB of memory and 3 GB of free disk space for the application, which needs no installer, no administrator rights, no account and no internet connection; NumPy, pandas and Matplotlib for the calculation layer. Licence: MIT for the calculation layer; free for research and teaching for the desktop application.

Restrictions for non-academic use: commercial use of the desktop application requires the written permission of the author.

Declaration of generative AI and AI-assisted technologies in the writing process. During the preparation of this work the authors used Claude (Anthropic) for translation, data analysis, software development and language editing. After using this tool the authors reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

Aitchison J (1986) The statistical analysis of compositional data. Chapman and Hall, London. https://doi.org/10.1007/978-94-009-4109-0

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Cross W, Iddings JP, Pirsson LV, Washington HS (1902) A quantitative chemico-mineralogical classification and nomenclature of igneous rocks. J Geol 10:555–690. https://doi.org/10.1086/621030

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003) Isometric logratio transformations for compositional data analysis. Math Geol 35:279–300. https://doi.org/10.1023/A:1023818214614

Frost BR, Barnes CG, Collins WJ, Arculus RJ, Ellis DJ, Frost CD (2001) A geochemical classification for granitic rocks. J Petrol 42:2033–2048. https://doi.org/10.1093/petrology/42.11.2033

Gromet LP, Dymek RF, Haskin LA, Korotev RL (1984) The "North American shale composite": its compilation, major and trace element characteristics. Geochim Cosmochim Acta 48:2469–2482. https://doi.org/10.1016/0016-7037(84)90298-9

Irvine TN, Baragar WRA (1971) A guide to the chemical classification of the common volcanic rocks. Can J Earth Sci 8:523–548. https://doi.org/10.1139/e71-055

Janoušek V, Farrow CM, Erban V (2006) Interpretation of whole-rock geochemical data in igneous geochemistry: introducing Geochemical Data Toolkit (GCDkit). J Petrol 47:1255–1259. https://doi.org/10.1093/petrology/egl013

Kolchedanno-polimetallicheskoe orudenenie Yuzhnogo Uzbekistana [Pyrite-polymetallic mineralisation of southern Uzbekistan] (1971) Fan, Tashkent, 185 pp (in Russian)

Le Maitre RW (ed) (2002) Igneous rocks: a classification and glossary of terms, 2nd edn. Cambridge University Press, Cambridge

Mamarozikov UD, Normatov O'A, Saitov NE, Hamrayev AX, Yo'ldoshev ZD (2025) Vulkanogen-bo'lakli jinslarini diagnostika qilish bo'yicha uslubiy qo'llanma [Methodological guide for the diagnostics of volcaniclastic rocks]. Science and Innovation, Tashkent, 108 pp (in Uzbek)

McDonough WF, Sun S-s (1995) The composition of the Earth. Chem Geol 120:223–253. https://doi.org/10.1016/0009-2541(94)00140-4

Middlemost EAK (1994) Naming materials in the magma/igneous rock system. Earth-Sci Rev 37:215–224. https://doi.org/10.1016/0012-8252(94)90029-9

O'Neill HSC (2016) The smoothness and shapes of chondrite-normalized rare earth element patterns in basalts. J Petrol 57:1463–1508. https://doi.org/10.1093/petrology/egw047

Peccerillo A, Taylor SR (1976) Geochemistry of Eocene calc-alkaline volcanic rocks from the Kastamonu area, northern Turkey. Contrib Mineral Petrol 58:63–81. https://doi.org/10.1007/BF00384745

Petrographic Code of the Republic of Uzbekistan (2024) Publishing High Future, Tashkent, 146 pp (in Uzbek)

Reimann C, Filzmoser P, Garrett RG (2005) Background and threshold: critical comparison of methods of determination. Sci Total Environ 346:1–16. https://doi.org/10.1016/j.scitotenv.2004.12.011

Rudnick RL, Gao S (2014) Composition of the continental crust. In: Holland HD, Turekian KK (eds) Treatise on geochemistry, 2nd edn. Elsevier, Oxford, pp 1–51. https://doi.org/10.1016/B978-0-08-095975-7.00301-6

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675. https://doi.org/10.1038/nmeth.2089

Shand SJ (1943) Eruptive rocks: their genesis, composition and classification, 2nd edn. Wiley, New York

Shoyoqubov T (1978) Geologiya i rudonosnost' paleozoyskikh vulkanogennykh formatsiy Yuzhnogo Uzbekistana [Geology and ore potential of Palaeozoic volcanogenic formations of southern Uzbekistan]. Fan, Tashkent, 126 pp (in Russian)

Sun S-s, McDonough WF (1989) Chemical and isotopic systematics of oceanic basalts: implications for mantle composition and processes. Geol Soc Lond Spec Publ 42:313–345. https://doi.org/10.1144/GSL.SP.1989.042.01.19

Taylor SR, McLennan SM (1985) The continental crust: its composition and evolution. Blackwell, Oxford

Vernon RH (2018) A practical guide to rock microstructure, 2nd edn. Cambridge University Press, Cambridge

Vulkanogennye formatsii Uzbekistana [Volcanogenic formations of Uzbekistan] (1971) Fan, Tashkent, 284 pp (in Russian)

Ward JH (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244. https://doi.org/10.1080/01621459.1963.10500845

Whitney DL, Evans BW (2010) Abbreviations for names of rock-forming minerals. Am Mineral 95:185–187. https://doi.org/10.2138/am.2010.3371

Williams M, Schoneveld L, Mao Y, Klump J, Gosses J, Dalton H, Bath A, Barnes S (2020) pyrolite: Python for geochemistry. J Open Source Softw 5:2314. https://doi.org/10.21105/joss.02314

Winchester JA, Floyd PA (1977) Geochemical discrimination of different magma series and their differentiation products using immobile elements. Chem Geol 20:325–343. https://doi.org/10.1016/0009-2541(77)90057-2