1. Data collection

1.1. Sampling (according to Röllig et al., 1990)

The stream sediment samples were taken as an average sample from the stream or riverbed, in the case of standing water from the bottom of the water body close to the bank. In a relatively time-consuming process, the sediment obtained was wet sieved at the sampling point using water from the sampled water body to a grain size of ≤ 0.2 mm and filled into polyethylene bottles with a capacity of 500 cm^3. The supernatant water was decanted after at least 24 h after sampling.

Water samples were collected separately from the respective stream sediment as a scoop sample and filled into 100 ml polyethylene bottles.

1.2. Sample preparation (according to Röllig et al., 1990)

The preparation of the collected stream sediments started with drying the samples in a drying oven at 80 °C. Subsequently, the sample material was divided using a riffle divider. One half of the sample material was finely ground (< 0.063 mm) for analysis using a vibrating disk mill, the other half of the sample material using a planetary ball mill. After dividing the respective subsamples again using a riffle divider, the samples were analysed. The analysis was mainly carried out from the partial sample ground with the planetary ball mill. Only for the mercury and fluorine analyses was the partial sample ground with the vibrating disk mill used.

1.3. Analytics (according to Röllig et al., 1990)

The multi-element analysis of the stream sediments was carried out at the Central Geological Institute of the GDR (ZGI) by means of emission spectral analysis (ESA) and X-ray fluorescence analysis (XRF). For the ESA, a PGS 2 planar grating spectrograph (VEB Carl Zeiss Jena, evaluation: photographic registration with subsequent semi-automatic measurement of the spectral lines on the MD 100 microdensitometer and registration on punched tape) was used. The XRF was carried out with an X-ray fluorescence analyser VRA-20 with Rh and Au anode (VEB Präzisionsmechanik Freiberg, evaluation: fully automatic recording of the measured values with subsequent concentration calculation in the online procedure and output of the results as printer protocol and on punched tape). Deviating from this, an X-ray fluorescence analyser ARF-6 was used to determine the tungsten and zinc contents. The determination of fluorine was carried out by pyrohydrolysis with complexometric final determination using Toron, the determination of mercury by flameless AAS after prior amalgamation on a gold trap. Certified rock standards (clay slate TB ST RGW 2301) were used for quality assurance of the multi-element determinations.

The pH value in the water samples was determined potentiometrically, the specific electrical conductivity (in µS/cm) conductometrically. The fluoride contents in the water samples were analysed potentiometrically using a fluoride-insensitive electrode.

The analytical methods used for the individual elements and their limits of determination are summarised in Table 1. Values outside these determination limits are to be regarded as purely calculated values. See also the remarks in chapter 2.3.

Table 1 Analytical methods and limits of determination.

Element/ Parameter	Method	Unit	Limit of determination lower	upper
pH	potentiometric			
EC*	conductometric	µS/cm		
F_{aqua}	potentiometric	mg/l	0.04	1.0
Ag	ESA[1]	mg/kg	1	1000
As	XRF[2]	mg/kg	15	1000
B	ESA[1]	mg/kg	20	1000
Ba	XRF[2, 3]	mg/kg	100	2000 (5000)
Be	ESA[1]	mg/kg	1	100
Bi	ESA[1]	mg/kg	10	1000
Co	ESA[1]	mg/kg	5	500
Cr	ESA[1]	mg/kg	5	500
Cu	ESA[1]	mg/kg	2	2000
Fe	XRF[2]	mg/kg	100	100000
F_{sed}	Pyrolysis	mg/kg	200	40000
Ga	ESA[1]	mg/kg	5	500
Hg	AAS	ng/kg	5	5000
La	ESA[1]	mg/kg	30	1000
Li	ESA[1]	mg/kg	30	5000
Mn	XRF[2, 4]	mg/kg	80 (100)	3000 (10000)
Mo	ESA[1]	mg/kg	2	1000
Nb	XRF[2]	mg/kg	5	300
Ni	ESA[1]	mg/kg	5	1000
Pb	ESA[1]	mg/kg	5	500
Rb	XRF[2]	mg/kg	25	1000
Sc	ESA[1]	mg/kg	10	500
Sn	ESA[1]	mg/kg	5	2000
Sr	XRF[2]	mg/kg	25	1000
Ti	XRF[2, 3]	mg/kg	100	10000
V	ESA[1]	mg/kg	6	1000
W	XRF[5]	mg/kg	10	1000
Y	XRF[2]	mg/kg	5	300
Yb	ESA[1]	mg/kg	2	100
Zn	XRF[5]	mg/kg	10	2000
Zr	XRF[2, 4]	mg/kg	30 (100)	1000 (5000)

* Specific electrical conductivity, [1] ESA: emission spectral analysis with planar grating spectrograph PGS-2 (VEB Carl Zeiss Jena), [2] XRF: X-ray fluorescence analysis with VRA 20 (VEB Präzisionsmechanik Freiberg), [3] until 1990 ESA[1] (Limits of determination in brackets), [4] until 1980 ARF 6 (Limits of determination in brackets), [5] XRF: X-ray fluorescence analysis (ARF 6)

2. Preparation of the data for digital provision 2023

2.1. Corrections in the identifier sample type

The coding of the identifier sample type was partially contradictory to the available analytical data. In 35 cases samples were coded as “Water sample”, but analytical data on stream sediment are also available. In 18 cases the sample type “Dry sample” was coded, but there are also analytical data for water. In these cases, the sample type was recoded to -9 (erroneous statement).

Furthermore, in 69 cases there are two or more samples with identical coordinates (355 samples in total). In 7 cases, these are precipitation samples (sample type = 7, total 16 samples). In all other cases, it can be assumed that the second samples (as well as any other samples that may be present) are duplicate samples. But only in one case was the sample coded as a duplicate sample. Since it is not possible to decide whether these are duplicate samples and if so, which of the samples was taken as a duplicate sample, the original coding was retained.

2.2. Coordinate correction and coordinate transformation

In a few cases, the coordinates in the original data stored until 1990 were incorrect, resulting in a location outside the study area. In the majority of cases, the coordinates could be corrected using existing documentation (sampling protocols, sampling maps). In a small number of samples (< 10) this was not possible. The corresponding data records were removed from the database.

In the original data stored until 1990, the coordinates were stored in the georeferencing systems EPSG:31467, EPSG:31468 and EPSG:31469 (Gauss-Krüger representation, ellipsoid Bessel 1841, Potsdam datum (central point Rauenberg), 3° meridian strips, central meridians 9°, 12° or 15°), depending on the location. These coordinates were transferred unchanged into the file for the digital provision of the data (columns RW & HW). For the map representations, the coordinates were transformed into the system EPSG:31468 (Gauss-Krüger representation, ellipsoid Bessel 1841, Potsdam datum (central point Rauenberg) (DHDN), mean meridian 12°) as well as the more modern EPSG:35832 (UTM, ellipsoid/datum ETRS 1989, zone 32N (mean meridian 9°)). In the geochemical data provided in the geoportal of the BGR, the coordinates are included in both georeferences (columns RW_GK4 & HW_GK4 respectively RW_UTM_32N & HW_UTM_32N).

2.3. Corrections in the identifiers

The identifiers (sample type, water body order, anthropogenic influence, land use, geochemistry, geology, stratigraphy, petrology and formation) were checked for plausibility of the information and corrected if necessary. Some inconsistencies in the coding of the stratigraphy, petrology and formation features were eliminated. In the stratigraphy feature, for example, there were code numbers 100 and 199 for "Precambrian" without any further specification. This was certainly the result of merging the partial files generated from the investigation of the individual bedrock units without first comparing the code numbers. Such inconsistencies were corrected. In the case described, for example, the code number 100 is generally used for "Precambrian" without further specification in the digitally provided file; the data records with code number 199 were recoded accordingly.

In some cases, code numbers were used in the original data that are not available in the present code tables. For example, in the petrology feature there were a large number of samples with the code number "0000", which was not explained. In such cases, the coding was changed to "-1" (no information).

2.4. Corrections in the analytical data

Like the coordinates and identifiers, the analytical data were also checked for plausibility for the digital provision of the data in the geoportal of the BGR.

A frequently occurring error was the specification of "0" as the analysed value. In some cases, this was used to code missing analysis values, but in other cases it was used to code values below the limit of determination. It is possible that some of the values coded as "0" are also due to the incorrect use of number formats, whereby decimal places were lost. In a large number of cases, the reason for coding as "0" could be traced, e.g. when all contents within a bedrock unit were coded as "0". These were then recoded as "-9999999" (missing value). On the other hand, values coded as "0" also occurred in areas of otherwise very low contents. In this case, the values in question were recoded as -1 * limit of determination (lower then limit of determination). To prevent falsification of the data, in case of doubt the values were always recoded as "-9999999" (missing value).

A further problem proved to be the indication of analytical results below the limits of determination given in Röllig et al. (1990). According to Röllig et al. (1990), these are to be regarded as purely calculated values. However, investigations of the distribution of these values in the area showed partly quite plausible areal distributions with a clear relation to the geology of the study area, or the contents below the limit of determination occurred in areas that already had low contents of the corresponding element. In order to avoid a general discarding of these values and thus the loss of any information they might contain, a plausibility check of the distribution in the lower content range was carried out using CP plots. While for most elements the distribution of the values below the limit of determination shows clear breaks, for the elements As, Bi, Co, Sc, Sn and W a quite plausible distribution of the values in a certain range below the limit of determination could be observed. In these cases, the determination limit was lowered to the smallest plausible value, deviating from that given in Röllig et al. (1990). Subsequently, it was examined whether the areal distribution of the values is plausible using the lowered limit of determination. In some cases, analytical differences between the subareas of the investigated bedrock units came to light, so that the determination limit had to be raised again to a plausible value for the subarea with the lowest analytical quality (e.g. for Bi). An overview of the units of measurement, the limits of determination from Röllig et al. (1990) as well as the limits of determination used for the data preparation can be found in the files made available via the geoportal of the BGR. The limits of determination used, which deviate from Röllig et al (1990) (Table 1), are summarised in Table 2.

All values below the specified limits of determination were set to -1 * limit of determination.

Table 2 Limits of determination (mg/kg) used in deviation from Röllig et al. (1990)

Element	Limit of determination according to Röllig et al. (1990)	Limit of determination used
As	15	2
Bi	10	5
Co	5	1
Sc	10	5
Sn	5	3
W	10	2

The rounding of the analytical values, which varied partly between the sub-files of the individual bedrock units and partly also arbitrarily, also proved to be problematic. This led to an accumulation on certain values in the CP plots (e.g. integer values or x.0 and x.5) and a clearly reduced distribution density between these values. In order to eliminate these discontinuities in the distribution, the complete data set for these elements was then rounded uniformly.

For the purpose of comprehensibility of the described adjustments of the data, the download of the geochemical data in the geoportal of the BGR contains the described CP plots, whereby the respective plots reflect the distribution before as well as after the data adjustment. At the same time, tables with statistical data are included to illustrate the data quality (number of samples, number of determined values, number of values below the limit of quantification) as well as the most important univariate statistical parameters. For the calculation of the statistical parameters, the values below the limit of determination were set to the value of half the limit of determination. Statistical parameters below the limit of quantification are indicated as < limit of quantification.

References

Röllig, G., Berger, W., Birke, M., Enderlein, F., Grosche, G., Kaberdin, B., Kampe, A., Krull, P., Kruse, B., Kunz, I., Rauch, U., Rentzsch, J., Schwandtke, E., Sehm, K., Söllig, A., Thomas, U. & Wünsch, K. (1990): Vergleichende Bewertung der Rohstofführung in den Grundgebirgseinheiten im Südteil der DDR. Gesellschaft für Umwelt- und Wirtschaftsgeologie GmbH: S. 807; Berlin (Archiv der Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, Archivnummer 1004927).