SAMPLERR: A Program for Estimating Sampling Error in Whole-Rock Analyses

Reference to be Cited: Reimink, J.R. and Chacko, T. (2024) Geological sampling. In: *Treatise on Geochemistry (3rd edition),* in press.

Overview of the SAMPLERR Program

SAMPLERR**,** which is short for 'sampling error', is a program that calculates the effect of sample size on the precision of a whole-rock geochemical analysis. To put it another way, the program investigates the relationship between sample size and sampling error. The sampling error associated with a whole-rock analysis is a function of the modal abundances of the rock's constituent minerals, the grain sizes of those minerals and the size (mass) of the sample collected to be representative of the rock body being investigated. In general, the coarser the grain size of a rock, the larger the sample that is required to achieve a certain level of precision in the analysis.

The rationale for the program is described in detail in Reimink and Chacko (2024) but the interested user is encouraged to also look at earlier papers by Kleeman (1967) and, in particular, Stanley (2003), as the present program is largely based on the sampling error algorithms developed in the latter study.

The calculation of sampling error with SAMPLERR involves the following steps:

1. Use Stanley's (2003) algorithms to calculate the expected sampling error standard deviations in the modal abundances of the rock's constituent minerals. Specifically, for a given sample mass, mineral modal abundance and mineral grain size, these algorithms determine the number of grains of a particular mineral that will be present in the sample and then, using sampling theory equations, calculates the expected standard deviation on the modal abundance of that mineral resulting from sampling error.
2. Employ a Monte Carlo approach to generate 10,000 synthetic modal rock compositions in which the modal abundance of each mineral phase has a distribution with a mean corresponding to that estimated for the rock body as a whole and a standard deviation corresponding to that predicted by sampling theory in step 1.
3. Normalize each of the 10,000 synthetic rock modal compositions to 100 % (i.e., the sum of all the mineral modes in each rock composition must total 100 %). This normalization step imposes closure on the system.
4. Calculate whole-rock chemical compositions for each of the 10,000 trials based on the normalized modal compositions of step 3 and the chemical compositions of each mineral phase in the rock.
5. Calculate means, standard deviations and relative standard deviations for each element in the whole-rock compositions.

Step 3, the normalization step, is critical and was not included in Stanley's original approach to calculating sampling error, which was coded in an Excel spreadsheet. Omitting that step in our programming yields results in sampling error major- and minor-element standard deviations in whole-rock compositions that are virtually identical to those given by Stanley's (2003) spreadsheet for calculating sampling error. Importantly, the normalization step significantly changes the standard deviations in the modal abundance distributions of many of the mineral phases relative to those input in steps 1 and 2. More specifically, normalization increases the standard deviations of some minerals and decreases those of others. This in turn produces standard deviations in the

elemental compositions (steps 4 and 5) that are in some cases markedly different (both higher and lower) than those indicated by the Stanley spreadsheet. As discussed in Reimink and Chacko (2024), estimates of sampling error given by the SAMPLERR program are in much better agreement with empirical data obtained in a case study of a coarse grained, porphyritic granite than those given by the Stanley spreadsheet.

Input Data for SAMPLERR

The input data required by SAMPLERR include the modal abundances (in volume %) and grain sizes (in cm) of minerals in the rock, the densities of the minerals (g/cm^3), and mineral compositions (in wt. % oxide), as determined by EPMA for a particular rock or representative mineral compositions taken from a mineralogy reference text. Also required is the selection of the modal abundance distribution type for each mineral in the rock. The two distribution options are hypergeometric (h) and Poisson (p); hypergeometric is the preferred option for major and minor minerals and Poisson for trace minerals. The required data can be input directly on the SAMPLERRR app or through the uploading of a .csv Excel file (see instructions on the app for details). An example of the csv-formatted input table is given below.

X	Modes	Dimension1	Dimension2	Dimension3	Distribution	Densities	SiO2	TiO2	Al2O3	Cr2O3	Fe2O3	FeO	MnO	MgO	CaO	Na2O	K2O	O	H2O	P2O5	ZrO2	CO2	SO2	Li2O	ThO2	BaO	La2O3	Ce2O3	Pr2O3	Nd2O3	Sm2O3	Eu2O3	Gd2O3	Tb2O3	Dy2O3	Ho2O3	Er2O3	Tm2O3	Yb2O3	Lu2O3	Y2O3
K-feldspar	38.2	2.4	1.2	1	h	2.57	64.10	0.00	19.71	0.00	0.00	0.08	0.00	0.00	0.69	2.47	12.44	0.00	0.00	0.00	0	0	0	0	0	0.5001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Plagioclase	27.1	0.4	0.2	0.2	h	2.64	60.56	0.00	24.95	0.00	0.00	0.13	0.00	0.00	6.15	8.09	0.12	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Quartz	26.3	0.8	0.4	0.4	h	2.65	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Biotite	5.3	0.1	0.05	0.05	h	3.07	38.23	1.42	15.15	0.00	0.00	17.12	0.38	13.67	0.02	0.09	9.82	0.00	4.00	0.00	0	0	0	0	0	0.102	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Magnetite	1.4	0.1	0.1	0.1	h	5.2	0.00	0.00	0.04	0.07	68.86	30.97	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Titanite	1.09	0.05	0.05	0.05	p	3.53	31.86	35.17	1.96	0.00	0.00	2.12	0.00	0.18	0.06	28.54	0.03	0.00	0.00	0.00	0.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Apatite	0.52	0.05	0.02	0.02	p	3.16	0.16	0.00	0.00	0.00	0.05	0.00	0.04	0.01	55.63	0.00	0.00	0.00	1.79	42.32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.79	0
Epidote	0.08	0.05	0.02	0.02	h	3.45	37.44	0.03	23.22	0.00	13.97	0.00	0.19	0.01	23.23	0.00	0.02	0.00	1.88	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ilmenite	0.06	0.05	0.05	0.05	h	4.79	0.08	50.52	0.00	0.00	4.22	29.88	14.94	0.03	0.33	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Zircon	0.044	0.02	0.01	0.01	p	4.71	32.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	67.2221	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Thorite	0.0041	0.02	0.01	0.01	p	5.3	24.87	0.11	1.28	0.00	3.21	0.00	0.02	0.05	1.78	0.01	0.00	0.00	11.24	0.17	0	0	0	0	57.2572	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nepheline	0	0	0	0	h	2.64	41.82	0.00	35.49	0.00	0.00	0.00	0.00	0.00	0.00	19.41	3.28	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Olivine	0	0	0	0	h	3.5	40.87	0.00	0.00	0.00	0.00	9.77	0.00	49.35	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cpx	0	0	0	0	h	3.31	54.31	0.00	0.00	0.00	0.00	6.49	0.00	16.39	22.81	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Opx	0	0	0	0	h	3.34	56.31	0.00	0.00	0.00	0.00	13.47	0.00	30.22	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Amphibole	0	0.75	0.75	0.75	h	3.28	42.11	0.00	11.64	0.00	0.00	20.58	0.00	9.04	11.51	1.66	1.60	0.00	1.88	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Muscovite	0	0.02	0.02	0.002	h	2.83	45.44	0.00	38.55	0.00	0.00	0.00	0.00	0.00	0.00	0.78	10.69	0.00	4.54	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chlorite	0	0.3	0.3	0.03	h	2.78	29.89	0.00	16.91	0.00	0.00	17.87	0.00	23.39	0.00	0.00	0.00	0.00	11.95	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11.95	0
Cordierite	0	0.1	0.05	0.05	h	2.64	48.73	0.00	33.08	0.00	0.00	11.65	0.00	6.54	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Andalusite	0	0.05	0.05	0.05	h	3.15	37.08	0.00	62.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Garnet	0	0.01	0.01	0.01	h	4.06	38.66	0.00	21.87	0.00	0.00	30.82	0.00	8.65	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chromite	0	0	0	0	h	5.12	0.00	0.00	0.00	67.90	0.00	32.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rutile	0	0.005	0.005	0.005	h	4.25	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Monazite	0	0.005	0.005	0.005	h	5.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0	0	0	0	0	69.8103	0	0	0	0	0	0	0	0	0
Calcite	0	0	0	0	h	2.71	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	56.03	0.00	0.00	0.00	0.00	0.00	0	43.9713	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pyrrhotite	0	0	0	0	h	4.69	0.00	0.00	0.00	0.00	0.00	52.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	47.1379	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pyrite	0	0	0	0	h	5.01	0.00	0.00	0.00	0.00	0.00	35.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	64.0731	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Spodumene	0	2	2	2	h	3.15	64.60	0.00	27.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pentlandite	0	1	1	1	h	2.42	78.56	0.00	16.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	4.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lepidolite	0.5	0.5	0.5	0.5	h	2.84	61.89	0.00	13.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	12.13	0.00	0.00	0.00	0	0	0	7.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pollucite	0	1.007	1.007	1.007	h	2.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Columbite	0	1.007	1.007	1.007	h	5.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

SAMPLERR Output

SAMPLERR can perform calculations in two different modes: 1) Range of Sample Sizes, and 2) Defined Sample Size. In both modes, the calculation output comes out in tabular as well as graphical form. Examples of the tabular output for the two modes are given below. The 'Range of Sample Sizes' option carries out sampling error calculations for sample masses of 625, 1250, 2500, 5000, 10000 and 20000 grams. For each mass, the program outputs the mean whole-rock composition given by the 10,000 simulations (see steps 4 and 5 above). The program also gives the expected sampling error standard deviation and relative standard deviation (RSD) for each oxide or element in the whole-rock analysis. For example, in the table below, the mean SiO_2 content is 68.19 wt. % for a 625 g sample mass with a predicted sampling error standard deviation on the SiO_2 concentration of ± 0.39 wt. %, which corresponds to a RSD of 0.57 %. The sampling error standard deviations for all elements decrease as the sample mass increases.

Example Output for 'Range of Sample Sizes' Calculation

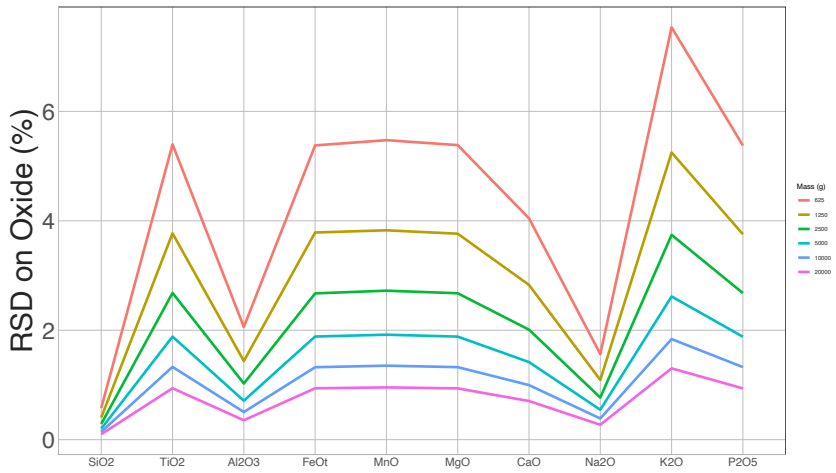
	value	mass	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	P2O5	Zr	Th	Ba
1	Mean	625	68.1876	0.6439	14.7755	3.6968	0.0437	0.829	2.6635	3.0587	5.1469	0.26	383.9866	40.7132	1682.1179
2	StDev	625	0.3899	0.0344	0.3016	0.1961	0.0024	0.0442	0.1069	0.0473	0.3854	0.0139	20.436	2.2381	147.3031
3	RSD	625	0.5718	5.3382	2.041	5.3058	5.4389	5.334	4.0152	1.5451	7.4871	5.3281	5.3221	5.4974	8.757
21	Mean	1250	68.1932	0.6433	14.7748	3.694	0.0436	0.8283	2.6616	3.0576	5.1499	0.2598	383.6725	40.6771	1683.3286
22	StDev	1250	0.2769	0.0243	0.2127	0.1389	0.0017	0.0311	0.0753	0.0332	0.2711	0.0098	14.4039	1.5824	103.6052
23	RSD	1250	0.406	3.772	1.4399	3.7608	3.8451	3.7568	2.8273	1.0871	5.2638	3.7571	3.7542	3.8902	6.1548
31	Mean	2500	68.1847	0.6429	14.7813	3.6915	0.0436	0.8278	2.6606	3.0578	5.1564	0.2596	383.4333	40.6502	1685.7774
32	StDev	2500	0.197	0.0171	0.151	0.0978	0.0012	0.0219	0.0531	0.0237	0.1913	0.0069	10.1601	1.111	73.0897
33	RSD	2500	0.289	2.6554	1.0214	2.6499	2.6846	2.6485	1.995	0.7742	3.709	2.6452	2.6498	2.733	4.3357
4	Mean	5000	68.1832	0.6425	14.7838	3.6891	0.0436	0.8272	2.6593	3.0573	5.1607	0.2595	383.1903	40.6263	1687.4398
41	StDev	5000	0.137	0.0122	0.1055	0.0697	8.00E-04	0.0156	0.0379	0.0168	0.1355	0.0049	7.2465	0.7937	51.7875
42	RSD	5000	0.201	1.8984	0.7134	1.8893	1.9324	1.8883	1.4262	0.5499	2.6248	1.8911	1.8911	1.9537	3.069
5	Mean	10000	68.1829	0.6423	14.7849	3.688	0.0435	0.827	2.6585	3.057	5.1629	0.2594	383.0527	40.6112	1688.3006
51	StDev	10000	0.0983	0.0086	0.0751	0.0491	6.00E-04	0.011	0.0268	0.0119	0.096	0.0035	5.1104	0.5611	36.6999
52	RSD	10000	0.1441	1.34	0.5082	1.3318	1.3632	1.3351	1.0081	0.3905	1.8597	1.3348	1.3341	1.3815	2.1738
6	Mean	20000	68.1823	0.6422	14.7856	3.6874	0.0435	0.8269	2.6583	3.0569	5.1638	0.2593	382.9999	40.6044	1688.6406
61	StDev	20000	0.0687	0.0061	0.0531	0.0351	4.00E-04	0.0079	0.0191	0.0085	0.0682	0.0025	3.657	0.4007	26.0944
62	RSD	20000	0.1008	0.9562	0.3592	0.9528	0.9678	0.9564	0.7192	0.2777	1.3213	0.9536	0.9548	0.9868	1.5453

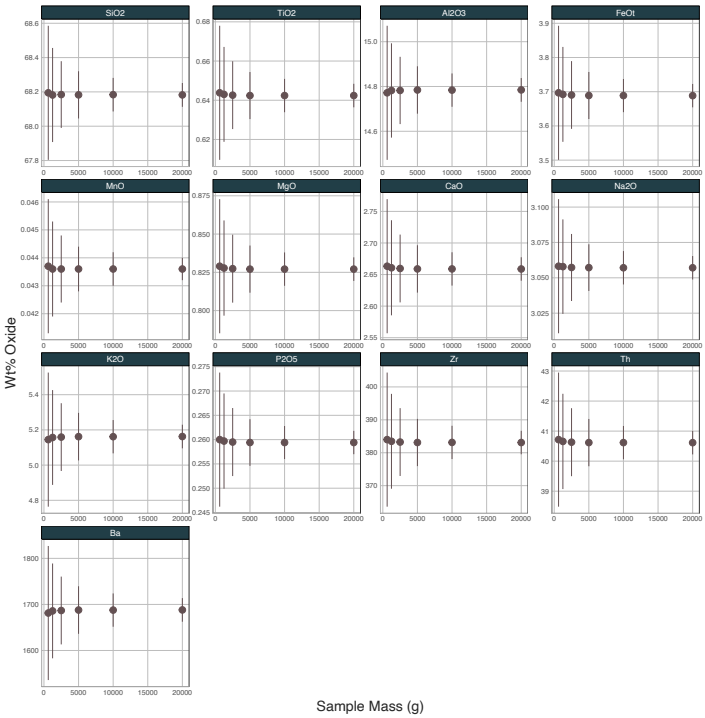
SAMPLERR's tabular output for a 'Defined Sample Size' calculation is similar to that for a 'Range of Sample Sizes' calculation but includes median as well as mean values for each elemental component from the 10,000 simulations. The output also lists the 97.5 and 2.5 % quantiles for each element from the 10,000 simulations. That is, 95% of the simulations for that element had values that fell between those two quantiles. The output from an example 'Defined Sample Size' calculation for a 100 gram sample mass is given below:

Example Output for 'Defined Sample Size' Calculation

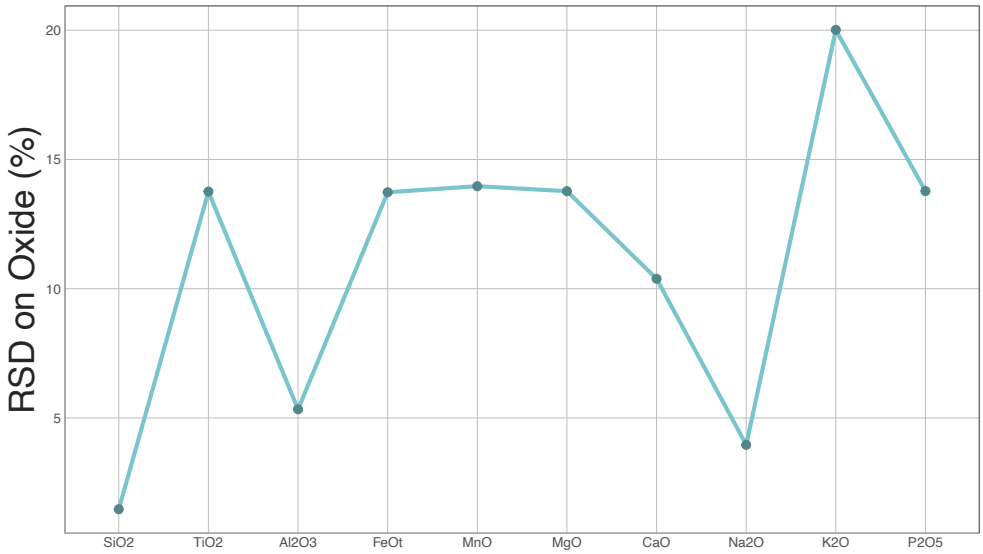
	mass	SiO2	TiO2	Al2O3	FeOt	MnO	MgO	CaO	Na2O	K2O	P2O5	Zr	Th	Ba
Mean	100	68.253	0.655	14.695	3.757	0.044	0.843	2.696	3.068	5.027	0.264	390.418	41.407	1636.176
StDev	100	1.008	0.092	0.791	0.526	0.006	0.118	0.285	0.124	1.021	0.037	54.652	5.945	390.43
RSD	100	1.476	14.055	5.381	14.001	14.234	13.987	10.569	4.043	20.306	13.983	13.998	14.358	23.862
Median	100	68.196	0.643	14.763	3.699	0.044	0.828	2.661	3.056	5.14	0.26	383.91	40.66	1679.27
Percentile_97.5	100	70.411	0.875	16.014	5.008	0.06	1.13	3.379	3.348	6.657	0.353	521.694	55.431	2260.322
Percentile_2.5	100	66.422	0.509	12.927	2.923	0.034	0.657	2.244	2.861	2.651	0.206	303.939	31.939	722.287

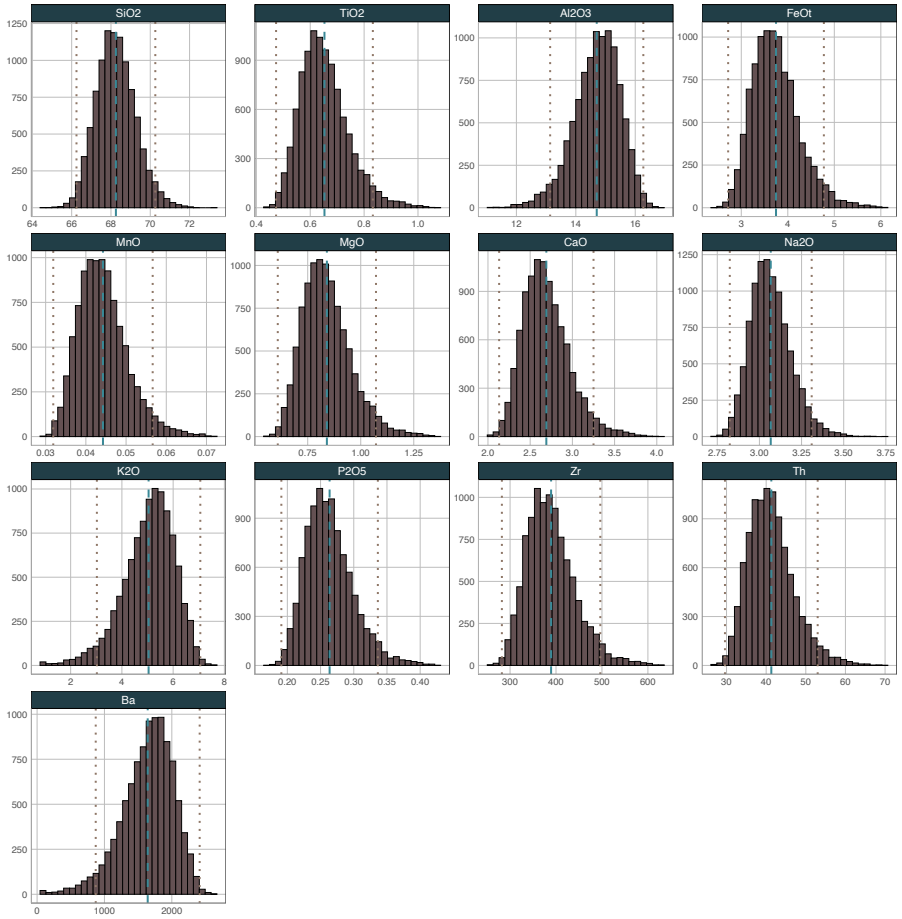
SAMPLERRR provides two types of plot outputs for the 'Range of Sample Sizes' calculation: 1) Sampling Error RSD vs Oxide, and 2) Sampling Error vs Sample Mass. Examples of these plots are given below:





SAMPLERR also provides two types of plot outputs for the 'Defined Sample Size' calculation: 1) Sampling Error RSD vs Oxide and 2) Histograms of Elemental Distributions from the 10,000 simulations. Note that the distributions in the latter type of plot can show significant skewness for small sample masses. Conversely, the distributions become more Guassian as sample mass increases. Example plots from a 100 gram sample mass calculation are given below:





References:

Kleeman, A.W. (1967) Sampling error in chemical analyses of rocks. *Journal of Geological Society of Australia*, 14:43-47.

Reimink, J.R. and Chacko, T. (2024) Geological sampling. In: *Treatise on Geochemistry (3rd edition),* in press.

Stanley, C.R. (2003) Estimating sampling errors from major and trace elements in geological materials using a propagation of variance approach. *Geochemistry: Exploration, Environment, Analysis*, 3:169-178.