Public-data File 88-12

DISCRIMINANT FUNCTION FOR SEPARATING TIN FROM NON-TIN GRANITES: APPLICATION TO THE HOPE GRANITE SUITE, EASTERN INTERIOR ALASKA

By

L.E. Burns, K.J. Johnson, and R.J. Newberry

Alaska Division of
Geological and Geophysical Surveys

March 1988

THIS REPORT HAS NOT BEEN REVIEWED FOR TECHNICAL CONTENT (EXCEPT AS NOTED IN TEXT) OR FOR CONFORMITY TO THE EDITORIAL STANDARDS OF DGGS.

794 University Avenue, Suite 200
Fairbanks, Alaska 99709

DISCRIMINANT FUNCTION FOR SEPARATING TIN FROM NON-TIN GRANITES: APPLICATION TO THE HOPE GRANITE SUITE, EASTERN INTERIOR ALASKA

by L.E. Burns, K.J. Johnson, and R.J. Newberry

OBJECTIVE

Historically, tin-granites have been difficult to distinguish from non-tin granites, such as those barren of mineralization or those giving rise to porphyry copper, molybdenum, or tungsten deposits, on the basis of mineralogy and composition. Granites associated with the various types of deposits and with no mineral deposits have broadly similar mineralogy and composition. If a granite can be established as having a high favorability for tin deposits, it is important to know whether those deposits are likely to be large or small. We have addressed these questions on worldwide granite data and have applied our "solutions" to the Hope granite suite (HGS) from interior Alaska (Burns and Newberry, 1987)[1].

To determine whether the Hope granite suite is likely to be associated with tin deposits, we first had to determine what compositional characteristics separate tin- from non-tin granites. We gathered a large database of tin- and non-tin-granites from around the world, and then computed discriminant functions based on major and trace element compositions to separate tin- from non-tin granites. Once the discriminant was functioning reasonably well on test data, the Hope granite suite analyses were run for classification.

To assess the potential of the Hope suite, the tin-granites from the previously constructed database were assigned low, moderate, and high deposit potential based on total tin production plus reserves (Laznicka, 1985). A new discriminant function was computed for these tin granites, based on major and trace element compositions. The resulting discriminant, which separated the three levels of tin deposit potential, was then applied to the Hope granite suite.

DATABASE

Two databases containing major and trace element compositional data of granites were compiled. The small database contained major oxide, trace element analyses, and CIPW normative mineralogy data from thirty-eight granites of the Hope granite suite. For the large database, about 1100 analyses of dominantly granite composition were extracted from the International Geologic Correlation Program database IGBA. The IGBA database is a compilation of compositional data for igneous rocks that have been published in about two/thirds of the major U.S. geologic periodicals (IGCP-163-IGBA). 310 additional analyses of tin- and non-tin-granites were added to the large database from published literature. CIPW normative mineralogy was computed for all of the data.

Because analyses of granodiorites, quartz monzonites, and other non-granites are commonly published with granite analyses, the compositional data in the large database were screened to eliminate non-granites. The classification scheme by Streckeisen and LaMaitre (1979) which labels plutonic rocks based on their CIPW normative mineralogy was used to define "granite composition" (fig. 1). Alkali granites, syenogranites, and monzogranites were kept in the database. The resulting large database had major oxide compositional data and CIPW normative mineralogy for 866 granites.

The most useful trace elements for which we had enough analytical data (about 175 analyses) to make the resulting discrimant functions worthwhile were Rb, Ba, and Zr. Other trace elements either were not good discriminators (as determined by analysis of ANOVA) or had too few analyses available to be included in the discriminant functions.

[1] This report complements "Probalistic estimation of mineral resources in the Lime Peak-Mt. Prindle area : Part 2" (Burns et al., 1988).

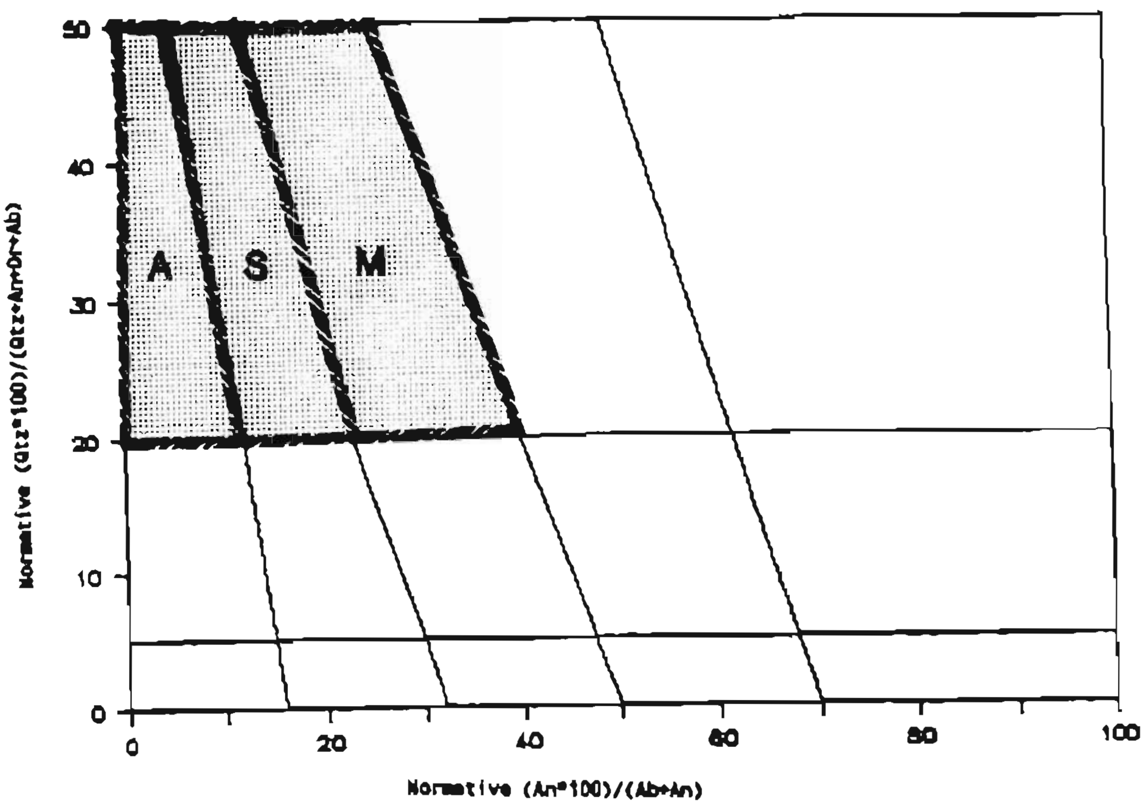

Fig. 1: Plot showing fields for normative compositions for alkali granite (A), syenogranite (S), and monzogranite (M). Symbols include: An - anorthite, Ab - albite, Or - orthoclase, Qtz - quartz. Plot modified from Streckeisen and LaMaitre (1979).

Three rocks from the Hope granite suite containing high normative corundum, acmite, and/or Na-metasilicate were deleted from the database because of assumed alteration. The rocks were aplite dikes from within and near the Quartz Creek pluton.

STATISTICAL METHODOLOGY

Correlation coefficients were computed for the possible discriminators. The trace elements Rb, Zr, and Ba were not highly correlated with major element or other trace element data and were therefore retained as possible discriminators. The major oxide compositions and the CIPW normative mineralogy largely correlated with one another; thus, the oxide composition and normative minerals were kept separately.

Discriminant functions were set up using 1) major oxide compositions, and 2) major oxide compositions plus Rb, Ba, and Zr. For brevity, these two types will be referred to throughout the rest of this report as TYPE M and TYPE T respectively. Discriminant functions were also computed based on the CIPW normative mineralogy and, as they were extremely similar to those computed from the major oxide composition, will not be discussed below.

As even the smallest data group had about 100 analyses, the data was randomly divided into a classifying group and a test group every time the discriminants functions were computed. The data was partitioned into groups of 80 and 20 percent for the classifying and test groups respectively.

The TYPE M discriminant was computed by a SAS discriminant program (DISCRIM: Sas Institute, 1982) using a quadratic equation[2]. Important discriminators for each type of discriminant function are shown in the tables.

For the TYPE T discriminant, the possible discriminators were run through a stepwise discriminant program by SAS Institute (STEPDISC; Sas Institute, 1982). We used the "stepwise" method of picking out useful discriminators; the method uses forward selection and backwards elimination. The data were then run through DISCRIM using the discriminators determined in the stepwise program. The data from the HGS was then run through the discriminant function. This method, of randomly picking classification and test groups, forming a discriminant function on the classification group, testing the discriminant function on the test group, and trying the discriminant function on the HGS data, was used four times on each TYPE T discriminant to get a more accurate estimate of the results.

RESULTS

Methods for evaluating the results

The results of the discriminant analyses are presented below. The success of the total discriminant function can be evaluated by the apparent error rate (APER) which is defined as

$$APER = \frac{n_{1m} + n_{2m} + \ldots + n_{im}}{n_1 + n_2 + \ldots + n_i}$$

where n_i = total # of observations in each group 1 - i
n_{im} = # of misclassified observations in each group.

[2] To reach the greatest accuracy concerning tin systems and the HGS, quadratic discriminant functions were developed. These are reported upon in tables 1-4. To make the discriminant functions useful to others, linear discriminant functions were computed resulting in slightly higher error rates. These are given, along with typical error rates, in table 5-7.

The total error rate, however, may not be the best number to base judgement of the discriminant function. Instead, minimizing a type II error (labeling fewer tin-granites as non-tin granites) so that fewer tin deposits will be overlooked would probably be the most important aspect for the exploration geologist. Type II errors are also summarized in the tables below.

<u>Quadratic discriminant analysis for differentiating tin-granite vs. non-tin granite:</u>

Four runs of both TYPE M and TYPE T were made. The discriminators used, APER, and type II errors for all eight runs are summarized in table 1. Typical "confusion" matrices for TYPE M and T are shown in table 2.

TYPE M - Major oxide composition

The advantages of this discriminant are that 1) a larger sample size existed to form the basic discriminant, and 2) all that is needed to run a published analysis is the major oxide composition. In the database used for this discriminant, 630 samples were non-tin and 236 were tin-related granites, yielding a total of 866 granites (table 2).

Although the apparent error rate (APER) is about the same as the APER one would get if assigning all the data to the non-tin-granite classification, the discriminant has a relatively low type II error rate. This suggests that even if one only has major oxide composition, the discriminant is a worthwhile endeavor, as tin-granites will be classified as non-tin granites about 18 percent of the time.

TYPE T - Major oxide composition, Rb, Ba, and Zr:

When Rb, Ba, and Zr were included as possible discriminators in addition to the major oxide data, STEPDISC always selected these trace elements as good discriminators. The inclusion of these trace elements in the discriminant function gave the best results achieved and improved the analysis significantly over the TYPE M discussed above (table 1). In the database used for this discriminant, 90 samples were non-tin and 47 were tin granites, yielding a total of 137 granites (table 1).

Hope Granite Suite

Classification of the Hope granite suite by use of the TYPE M discriminant resulted in an average of 93.5 percent of the rocks were tin-granites. Use of the TYPE T discriminant raised the percentage to 100. The average for the posterior probability for the data being tin-granites in the TYPE M and TYPE T runs respectively is 85.2 and 99.9 percent.

<u>Quadratic discriminant analysis for differentiating potential of tin deposits in a tin-granite system:</u>

TYPE M - Major oxide composition & TYPE T - Major oxide composition, Rb, Ba, and Zr:

Four runs for each a TYPE M and TYPE T discriminant function were made on samples from tin deposits having known low, moderate, and high tin potential. The summary data are shown in table 3, and typical confusion matrices are shown in table 4. The apparent error rate for the TYPE M discriminant is about 30 percent, while that for the TYPE T discriminant is about 15 percent. To minimize the chance of overlooking a major tin deposit, the TYPE T discriminant is best. As with differentiating between tin and non-tin granites, however, the TYPE M discriminant is more useful than assigning the data to random classes.

Increasing the number of analyses containing Rb, Ba, and Zr would greatly improve the database. Only 97 analyses were available for the TYPE T discriminant analysis. A nonparametric ANOVA showed

TABLE 1: SUMMARY RESULTS OF QUADRATIC DISCRIMINANT FUNCTION FOR PREDICTING TIN VERSUS NON-TIN GRANITES:

Part A: TYPE M - Major oxide composition: 866 cases

DISCRIMINATORS: All variables used, but major variables in order of importance are Fe_2O_3, CaO, P_2O_5, TiO_2, SiO_2, and Al_2O_3

	Classification Group		Test Group		HGS: Percent
	APER (%)	TYPE II error (%)	APER (%)	TYPE II error (%)	classified as tin-granite (42 total)
Run 1)	25.8	16.1	34.3	29.2	92.9
Run 2)	26.5	18.1	28.6	20.8	95.2
Run 3)	20.7	20.7	28.0	33.3	90.5
Run 4)	27.9	17.6	24.6	16.7	95.2
AVERAGE	25.2	18.1	28.9	25.0	93.5

Percent of known database correctly classified by following methods:

Discriminant based on all data	73 %
Prior proportions	73 %
Random classification	60.5 %

Part B: TYPE T - Stepwise: Major oxide composition, Rb, Ba, and Zr: 137 cases

DISCRIMINATORS in order of decreasing F-statistic:
Rb, Ba, Fe_2O_3, Zr, CaO, and TiO_2

	Classification Group		Test Group		HGS: Percent
	APER (%)	TYPE II Error (%)	APER (%)	TYPE II Error (%)	classified as tin-granite (38 total)
Run 1)	6.5	2.7	3.5	0.0	100
Run 2)	5.6	0.0	3.5	0.0	100
Run 3)	1.9	0.0	13.8	20.0	100
Run 4)	5.6	0.0	3.5	0.0	100
AVERAGE	4.9	0.7	6.1	5.0	100

Percent of known database correctly classified by following methods:

Discriminant based on all data	94.2 %
Prior proportions	65.7 %
Random classification	58.0 %

TABLE 2: Typical confusion matrices and APER (apparent error rate) for results of quadratic discriminant functions predicting tin vs. non-tin granites:

Part A: TYPE M - Major oxide composition

Example for Classification Group: Run 2

		predicted Non-tin granite	Tin-granite	Total	APER
actual	Non-tin granite	354	149	503	26.5 %
	Tin-granite	34	154	188	
	Percent	.56	.44		
	Priors	.73	.27		

Part B: TYPE T - Major oxide composition, Rb, Ba, and Zr

Classification Group example: Run 2

		predicted Non-tin granite	Tin-granite	Total	APER
actual	Non-tin granite	65	6	71	5.6 %
	Tin-granite	0	37	37	
	Percent	.60	.40		
	Priors	.66	.32		

TABLE 3: SUMMARY RESULTS OF QUADRATIC DISCRIMINANT FUNCTION FOR PREDICTING LOW, MODERATE, AND HIGH TIN POTENTIAL FOR TIN GRANITES:

Part A: TYPE M - Major oxide composition: 866 cases

DISCRIMINATORS: All variables input, but most important in order of decreasing importance are P_2O_5, TiO_2, Na_2O, Al_2O_3, and MgO

	APER for Classification Group	APER for Test Group
Run 1)	23.0	36.7
Run 2)	28.9	24.5
Run 3)	32.1	28.6
Run 4)	26.7	34.7
AVERAGE	27.7	31.1

Percent correctly classified by following methods:	
Discriminant based on all data	68 %
Prior proportions	46 %
Random classification	33 %

Part B: TYPE T - Stepwise: Major oxide composition, Rb, Ba, and Zr: 97 cases

DISCRIMINATORS in order of decreasing importance are P_2O_5, Rb, Na_2O, SiO_2

	APER for Classification GROUP	APER for Test GROUP
Run 1)	10.8	30.4
Run 2)	13.5	21.7
Run 3)	17.6	17.4
Run 4)	9.5	21.7
AVERAGE	12.9	22.8

Percent correctly classified by following methods:	
Discriminant based on all data	86.6 %
Prior proportions	60.8 %
Random classification	45.0 %

TABLE 4: Typical confusion matrices and APER (apparent error rate) for results of quadratic discriminant functions predicting low, moderate, or high tin potential:

Part A: TYPE M - Major oxide composition

Classification Group example: Run 2

		predicted				
		Low	Moderate	High	Total	APER
	Low	49	33	5	87	
actual	Moderate	5	54	3	62	28.9 %
	High	4	4	30	38	
	Percent	31.0	48.7	20.3	187	
	Priors	.47	.33	.20		

Part B: TYPE T - Major oxide composition, Rb, Zr, and Ba

Classification Group example: Run 2

		Low	Moderate	High	Total	APER
	Low	45	1	0	46	
actual	Moderate	5	9	0	14	13.5 %
	High	1	3	10	14	
	Percent	68.9	17.6	13.5	74	
	Priors	.62	.19	.19		

statistically important increases from the low to the high tin potential rocks in Rb (e.g. 319, 475, and 552 ppm) and corundum and decreases in Na_2O, magnetite, and albite. These changes are consistent with our geologic reasoning, and suggest that improving the number of analyses in the database may help the resolution of the tin potential.

Hope Granite Suite:

Using the low, medium, and high potential discriminant function on the Hope granite suite generally classifies the majority of the rocks as having low to moderate tin potential. An average of the four TYPE T runs places 73.1, 25.6, and 1.3 percent of the HGS samples in the low, medium, and high potential fields. TYPE M runs place the majority of the rocks in the medium tin potential field with percentages of 12, 86, and 2 for low, medium, and high fields. These runs compare favorably to the ROCKVAL analysis (Burns et al., 1988; this report; Pessel and Newberry, 1987), as the overall estimated contained tin resources of the area are most likely equivalent to a tin district with low to moderate reserves.

Linear discriminant functions

Results of linear discriminant functions and some of the classification functions are shown in tables 5, 6, and 7. The TYPE T discriminants had about an 8 percent APER rate for the tin vs. non-tin granite, and about a 25 percent APER rate for the tin potential classification problem. Our work suggests a linear discriminant function is not worthwhile for a dataset containing ony major oxide analyses (TYPE M).

SUMMARY

Our study has shown that tin-granites can be differentiated from non-tin granites by means of both linear and quadratic discriminant functions. The best discriminant functions include Ba, Rb, Zr, Fe_2O_3, TiO_2, and CaO as discriminators (TYPE T). The quadratic and linear functions yield apparent error rates of about 5 and 8 percent respectively. The type II error rate is similarly low. Low values of the type II error rate minimize missing a tin deposit.

Discriminating between the tin and non-tin granite using only major oxide analyses and a quadratic discriminant function yields poor apparent error rates, but lower type II error rates than assuming all the data belongs to non-tin granites. A linear discriminant function on major oxide analyses only is not worthwhile.

We have also shown that tin-granites can be delineated in terms of low, moderate, and high tin potential. Of the various discriminants formed, the discriminants using Rb, P_2O_5, Na_2O, and SiO_2 were the best. The linear discriminant function formed from these elements actually had a lower apparent error rate (about 23 as opposed to 27 percent) in our study than the quadratic discriminant function.

When the tin/non-tin granite discriminant function was applied to the HGS, the HGS highly resembled tin-granites. The low, moderate, and high tin potential discriminant function suggests the HGS has a low to moderate potential for total contained tin resources.

ACKNOWLEDGEMENTS

Support for this study came from the U.S. Bureau of Mines. We are grateful to Dr. D.B. Hawkins and Dr. Dana Thomas from the University of Alaska who provided much useful advice on statistics. We also appreciate the time spent by G.H. Pessel who read and revised the original report. Janice H. Leveille, aided by Gary Coulter and Rich Leveille, helped set up the original database and performed preliminary statistics.

TABLE 5: SUMMARY RESULTS OF LINEAR DISCRIMINANT FUNCTION

FOR PREDICTING TIN VERSUS NON-TIN GRANITES:

Part A: TYPE M - Major oxide composition: 866 cases

DISCRIMINATORS: All variables used, but major variables in order of importance are Fe_2O_3, CaO, P_2O_5, TiO_2, SiO_2, and Al_2O_3

	Classification Group APER (%)	Classification Group TYPE II error (%)	Test Group APER (%)	Test Group TYPE II error (%)
Run 1)	21.4	65.4	22.9	66.7
Run 2)	20.7	62.8	22.9	58.3
Run 3)	23.0	68.1	18.3	54.2
Run 4)	16.1	59.0	26.3	68.8
AVERAGE	20.3	63.8	22.6	62.0

Percent of known database correctly classified by following methods:

Prior proportions	73 %
Random classification	60.5 %

Part B: TYPE T - Stepwise: Major oxide composition, Rb, Ba, and Zr: 137 cases

DISCRIMINATORS: Rb, Ba, Fe_2O_3, Zr, CaO, TiO_2

	Classification Group APER (%)	Classification Group TYPE II Error (%)	Test Group APER (%)	Test Group TYPE II Error (%)
Run 1)	8.3	10.8	6.9	10.0
Run 2)	9.3	13.5	10.3	20.0
Run 3)	9.3	10.8	0.0	0.0
Run 4)	8.3	10.8	3.4	10.0
AVERAGE	8.8	11.5	5.2	10.0

Percent of known database correctly classified by following methods:

Prior proportions	65.7 %
Random classification	58.0 %

TABLE 7: LINEAR DISCRIMINANT FUNCTIONS

Tin versus non-tin: TYPE I

	Run 1		Run 2		Run 3		Run 4	
	Non-tin granite	Tin granite	Non-tin granite	Tin granite	Non-tin granite	Tin granite	Non-tin granite	Tin granite
Constant	-8.574	-6.962	-11.072	-9.338	-9.508	-6.491	-8.282	-6.624
CaO	4.137	2.354	7.935	4.341	7.242	3.781	4.857	2.541
Fe2O3	2.284	1.103	3.006	1.721	2.734	1.479	1.404	0.832
TiO2	-6.709	-3.173	-15.372	-8.056	-13.926	-6.817	-9.721	-4.716
Ba	0.004	0.002	0.005	0.003	0.004	0.002	0.003	0.001
Rb	0.0107	0.017	0.017	0.026	0.010	0.015	0.010	0.017
Zr	0.008	0.004	0.009	0.005	0.008	0.004	0.014	0.008

TABLE C7: LINEAR DISCRIMINANT FUNCTION

Tin potential: TYPE I

	Run 1			Run 2			Run 3			Run 4		
	Low	Medium	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High
Constant	-1681.982	-1791.450	-1747.313	-1712.287	-1742.454	-1763.459	-1413.510	-1458.378	-1477.577	1-1801.097	-1837.597	-1864.427
Al_2O_3	64.356	64.504	65.772	58.841	58.910	60.249	51.565	51.846	53.230	66.836	67.011	68.1377
Na_2O	-32.934	-33.997	-35.382	-27.445	-27.964	-29.884	-26.607	-27.863	-29.664	-38.553	-39.434	-41.273
P_2O_5	35.739	36.267	36.404	77.016	83.215	90.457	65.159	72.262	79.204	46.744	50.054	62.346
SiO_2	35.739	36.267	36.404	37.191	37.070	31.113	30.483	31.070	31.113	38.782	39.258	39.451
Rb	-0.037	-0.037	-0.034	-0.034	-0.033	-0.030	-0.030	-0.030	-0.027	-0.034	-0.033	0.030

REFERENCES CITED

Burns, L.E., and Newberry, R.J., 1987, Intrusive Rocks of the Lime Peak - Mt. Prindle area: Part A: Tin-related granites of the Lime Peak - Mt. Prindle area: *in* Smith, T.E., Pessel, G.H., and Wiltse, M.A. (eds), Mineral Assessment of the Lime Peak - Mt. Prindle Area, Alaska, Alaska Division of Geological and Geophysical Surveys, p. 3.1-3.62.

Burns, L.E., Newberry, R.J., and Pessel, G.H., 1988, The probabilistic estimation of mineral resources in the Lime Peak - Mt. Prindle area: Part 2: Alaska Division of Geological and Geophysical Surveys, Fairbanks, Alaska, Public data-file PDF 88-11.

IGCP-163-IGBA, 1984, An Introduction to Version 1 of IGBADAT: International Geological Correlation Program, National Oceanographic Atmospheric Association, National Geophysical Data Center, Boulder, Colo., 3 p.

Laznicka, P., 1985, Empirical Metallogeny, v. 1: Elsevier, Amsterdam, 1758 p.

Pessel, G.H., and Newberry, R.J., 1987, Pessel, G.H., and Newberry, R.J., 1987, Probabilistic estimation of mineral resources in the Lime Peak - Mt. Prindle area: *in* Smith, T.E., Pessel, G.H., and Wiltse, M.A. (eds), Mineral Assessment of the Lime Peak - Mt. Prindle Area, Alaska, Alaska Division of Geological and Geophysical Surveys, p. 7-1 - 7-17.

Sas Institute, Inc., 1982, Sas User's Guide: Statistics: Sas Institute, Inc., Cary, NC, 584 p.

Streckeisen, A.L., and LaMaitre, 1979, A chemical approximation to the model QAPF classification of the igneous rocks: Neues Jahrb. Mineral. Abh., v. 136, p. 169-206.