A PETROGRAPHIC CRITERION FOR THE POSSIBLE REPLACEMENT ORIGIN OF ROCKS

F. CHAYES

ABSTRACT. In rocks formed by replacement the amounts of residual and introduced materials ought to vary inversely. This implies negative correlation between the two types of material, and it is argued that unless the negative correlation is significant the sample from which it is computed may not be considered compatible with the replacement hypothesis in question. On this basis it is shown that variations in the chemical composition of the rapakivi granites are incompatible with Backlund's suggestion that these rocks have been formed by reaction between the Jotnian sandstone and alkali-aluminate emanations.

INTRODUCTION

IT must be evident to an increasing number of geologists that in our discussions of granitization we have reached something of an *impasse.* The Ottawa symposium was a useful and forceful reminder that even an *impasse* may be discussed reasonably and amicably. It is to be hoped, however, that the general air of good feeling which characterized the symposium will not tempt us to rest content with compromise rather than press for genuine resolution of our differences.

If we put the argument in very general or in very specific terms, it is obvious at once that we can not expect to reach the truth by taking a vote or casting up a sort of weighted average of opinion. Either granites are *generally* of magmatic origin, or they are not, and it is only of social interest that at the present moment about half the world's geologists hold for the "either," and the other half for the "or." Either *this particular granite* is magmatic, or it is not; the circumstance that it has been studied by five emanationists and three magmatists may affect the history of geology but has little to do with the history of the rock.

The substance of the granitization controversy is of the utmost importance to petrology, but its form is also typical of a great many quarrels of less significance, and perhaps of petrographic reasoning as a whole. Generally we seem content to reach hypotheses that can not be shown to be impossible, leaving to the dictates of "common sense," previous experience, or authority the decision as to which of these not impossible hypotheses is to be preferred in a particular situation.

Wherever there is any occasion for doubt we are now usually provided with a number of hypotheses, some admittedly old, others supposedly new. With the methods and mental habits at our disposal we are commonly unable to offer convincing reasons for preferring one or the other of these. A somewhat closer inspection is likely to reveal that even the supposedly new suggestions were familiar if not hackneyed before most of us were born. We urgently require a method of analysis that will help us weed out misleading or unlikely hypotheses so that we may devote our attention to modification, extension, and elaboration of the more fruitful traditional concepts.

At present we have no such method, but I believe we might obtain one by a careful, systematic adoption and adaptation of the simpler, basic procedures of small sample statistics, as now practiced chiefly by biologists. The proper utilization of statistical methods will require considerable re-orientation and re-education for all of us, and busy people are not likely to submit to either of these painful procedures without some reasonable prospect of reward. This note is designed to show by a specific example that the prospect is sufficiently pleasant to make the pain worth bearing.

At the outset certain general and almost self-evident consequences of extensive replacement unaccompanied by volume change are deduced. It is then shown that these effects are not found in analyses of the Rapakivi granites, which Backlund has suggested were formed by feldspathization of sandstone. I have tried to reduce technical jargon to a minimum, and to keep before the reader the purpose and advantage of the statistical mode of argument. Enough definition and detail are provided so that the argument can be followed without outside reference; but there are frequent text citations, and the paper will have served its purpose if many readers abandon it in favor of the texts.

THE GENERAL PROBLEM

In any rock formed by replacement of one group of minerals by another, the quantity (by weight, volume or area) of original minerals ought to vary inversely with the quantity of replacement or reaction products. Where the two types of material differ markedly in chemical composition this mineralogical relationship will be reflected in chemical analyses. Under the proper circumstances we ought to be able to use variations in chemical composition (or modal or normative statements) as an index of the possibility that some particular group of rocks has been formed by one or another of the replacement processes. When such general variations are suspected in most other fields of natural science, and there seems no reason to doubt that the measurements of each variate are more or less symmetrically concentrated about their mean, the Pearson coefficient of correlation is commonly used as a test of the conformity between hypothesis and evidence. Where, as here, the hypothesis requires one variable to increase as another decreases, we may say that it calls for significant negative correlation between the two, and a sample in which the variables concerned do not exhibit significant negative correlation is not in conformity with the hypothesis. Depending on our opinion about the size and quality of the sample, we may conclude either that it offers no support to, or that it provides sufficient grounds for rejection of, the hypothesis.

It is unfortunately true that the tables that form so much of the petrographer's data (tables of full chemical analyses, or of modes or norms) differ in one important respect from those usually subjected to statistical examination. In general, the sum of the variates is itself a variable, while in most of our tables it is actually, or practically, constant. The effect of this condition is to generate negative correlation. Although the statistics computed from such tables offer a more adequate description than can usually be obtained graphically or by inspection, their utility as probability statements is limited unless proper account is taken of this tendency. I hope to discuss the question at greater length in a subsequent report; here it will be sufficient to establish that a permissible though inefficient use of these statistics is possible if hypotheses may be stated so that failure to achieve significant negative correlation is cause for rejection.

CORRELATION IN A CLOSED TABLE

For convenience we may designate as "closed" any table formed according to the rule that $X + Y + \ldots + N = K$, where X, Y, etc. are measurements on a single specimen. Such tables undoubtedly occur in other subjects, but I have found no treatment of them and in most statistics texts they are ignored. A "closed" table is to be distinguished from an "open" one, in which the sum of variates in a single measurement is itself a variable. Statistics ordinarily deals with open tables; closed tables are the mainstay of chemical petrography.

Representing different minerals by subscripts, we may describe the weight composition of any specimen as

$$X_1 + X_2 + \ldots . + X_n = 1 \qquad (1)$$

The average for a group of specimens will be:

$$\bar{x}_1 + \bar{x}_2 + \ldots . + \bar{x}_n = 1 \qquad (2)$$

The deviation of X_1 from its mean $\bar{x}_1$, or $(X_1 - \bar{x}_1)$, is represented by x_1, and substracting (2) from (1) gives

$$x_1 + x_2 + \ldots . + x_n = 0 \qquad (3)$$

For total, or zero order, correlation, we consider only two variables at a time.

Forming a new variable,

$$Z = X_3 + X_4 + \ldots . X_n \qquad (4)$$

we may rewrite equation (1) as

$$X_1 + X_2 + Z = 1 \qquad (5)$$

and equation (3) as

$$x_1 + x_2 = -(x_3 + x_4 + \ldots x_n) = -z \qquad (6a)$$
$$x_1 + z = -x_2 \qquad (6b)$$
$$x_2 + z = -x_1 \qquad (6c)$$

Squaring both sides of each line of (6) and summing the results gives:

$$Sx_1^2 + Sx_2^2 + 2\,Sx_1x_2 = Sz^2 \qquad (7a)$$
$$Sx_1^2 + Sz^2 + 2\,Sx_1z = Sx_2^2 \qquad (7b)$$
$$Sx_2^2 + Sz^2 + 2\,Sx_2z = Sx_1^2 \qquad (7c)$$

The coefficient of correlation [(1)] of x_1 with x_2 is defined as

$$r_{12} = \frac{Sx_1x_2}{\sqrt{Sx_1^2 \cdot Sx_2^2}} \qquad (8)$$

(1) For a discussion of r and of correlation in general see Yule and Kendall, *An Introduction to the Theory of Statistics*, 13th Ed., 1946, Chapters 11-17, or G. W. Snedecor, *Statistical Methods*, 4th Ed., 1946, Chapter 7.

and since the denominator is taken as positive, the sign of r_{12} is determined by the sign of Sx_1x_2. The sign of Sx_1x_2 is in turn determined by the size relations of Sx_1^2, Sx_2^2, and Sz^2. (More precisely, and the difference may be quite important in interpreting results, the fourth quantity in any line of equation (7) is fixed by the other three, but the square terms are always positive.) If we arbitrarily set

$$Sx_1^2 > Sx_2^2 > Sz^2$$

it is clear from equation (7) that Sx_1x_2 and Sx_1z will be negative and that Sx_2z will also be negative unless $(Sx_2^2 + Sz^2) < Sx_1^2$. If the square terms in equation (7) all differ, as is commonly the case, then the chance that any given correlation coefficient will be positive will certainly be no more than the product of the chance that each of two square terms will be less than a third, times the chance that their sum will be less than the third, or $(1/2)^3$. The *a priori* chance that a given correlation coefficient will be negative is thus at least 7/8 or 87%, and this is probably an underestimate, for if the square terms are comparable in size, as will usually be the case, the sum of any two will likely exceed the third, and positive correlation will be excluded. It is clear from this that we can not attribute the usual significance to high negative correlation (or low positive correlation) unless we make allowance for cloture of the table. But it is equally clear that if any hypothesis calls for a significant negative correlation *not exhibited by the data*, we are justified, perhaps more justified than in the case of open tables, in asserting that the data do not support the hypothesis. The next section discusses a practical illustration in which it seems fair on hypothetical grounds to require significant negative correlation.

CORRELATION OF NORMATIVE QUARTZ AND ALKALI FELDSPAR IN SOME RAPAKIVI ANALYSES

In 1938 H. Backlund[2] offered an excellent summary of rapakivi petrography together with some notes on field occurrence, and some generalizations and speculations about chemical composition and history that I propose to examine here.

From a study of nearly the same analyses used in computing Table I of this note Backlund concludes "The figures cited show that there seems to be no formal difficulty in transform-

(2) Backlund, H., 1938, Jour. Geol., XLVI, p. 337.

ing, without essential volume change, a Jotnian sandstone suite into a rapakivi granite by addition of material approximating in composition to sodium, potassium (and calcium?) aluminates." Backlund further assures us of the " . . . nearly constant contents of $Fe_2O_3 + FeO + MgO + TiO_2$ throughout the rocks of each association" so that the granitization process envisioned requires reciprocal reduction in free quartz and increase in alkalis and alumina, on a weight basis; unless this relationship attains numerical significance it is hard to see how variations in the bulk composition of the rapakivis offer evidence in support of the hypothesis.

In Table I are listed normative Q and (Or + Ab) values for 40 rapakivi analyses, arranged in order of decreasing Q. From inspection some would conclude that Q varied inversely with (Or + Ab), others might be hesitant. Certainly the relationship is not a striking one, and if the nine lowest Q values were deleted the remaining array would be far from convincing. As a matter of fact the correlation coefficient for the entire group is –0.71; this high a value is likely to be drawn far less than

TABLE I

Normative Q and (Or + Ab) for 40 Rapakivi Analyses arranged in order of decreasing Q.

Q	Or + Ab	Q	Or + Ab
39.04	53.52	25.98	62.34
37.84	57.83	25.56	53.01
35.43	60.44	25.50	56.32
34.76	56.48	24.00	50.73
34.61	56.85	22.92	49.71
34.14	56.80	22.86	60.34
34.09	55.07	21.84	55.08
34.08	55.73	21.24	54.69
33.54	57.00	21.02	71.93
33.52	59.77	20.94	54.20
33.46	58.36	20.52	58.88
32.28	54.75	20.34	65.74
31.83	59.31	18.68	54.54
30.30	51.42	16.04	69.82
29.40	55.47	14.42	76.52
28.88	63.87	14.30	71.93
28.68	60.05	12.29	75.81
28.02	54.16	9.19	79.14
27.51	60.79	6.61	76.64
26.70	58.84	5.89	75.19

once in a hundred samplings of an uncorrelated population. Barring further inquiry about the homogeneity of the sample we would conclude that the data in Table I were clearly in agreement with hypotheses calling for inverse variation of Q and (Or + Ab).

Before going further I should make it clear that Dr. Backlund is not responsible for Table I, since in his work the identity of individual analyses is carefully retained. But if the hypothesis finally reached covers the entire group, as Dr. Backlund maintains, and there is no way of weighting the data, then Table I represents a necessary step in or result of the reasoning, and may not be avoided by being left implicit. We next give reasons for supposing that the data are inhomogeneous, and that "lumping" of the sort portrayed in Table I is not justifiable.

In Table II the order of decreasing Q content is retained,

TABLE II

Normative Q, (X), and Normative Or + Ab, (Y), for Suites of Rapakivis.

	Vehmaa		Laitila		Lappeenranta		Loos-Hamra Granites		Loos-Hamra Porphyries	
	X	Y	X	Y	X	Y	X	Y	X	Y
	33.54	57.00	34.61	56.85	34.76	56.48	37.84	57.73	39.04	53.52
	29.40	55.47	34.14	56.80	34.08	55.73	35.43	60.44	33.52	59.77
	28.68	60.05	34.09	55.07	32.28	54.75	33.46	58.36	31.83	59.31
	26.70	58.84	28.02	54.16	30.30	51.42	21.02	71.93	28.88	63.87
	25.98	62.34	20.34	65.74	25.50	56.32	16.04	69.82	27.51	60.79
	25.56	53.01	18.68	54.54	24.00	50.73	14.30	70.93	14.42	76.52
					22.92	49.71	9.19	79.14	12.29	75.81
					22.86	60.34			6.61	76.64
					21.84	55.08			5.89	75.19
					21.24	54.69				
					20.94	54.20				
					20.52	58.88				
Sums	169.86	346.71	169.88	343.16	311.24	658.33	167.28	469.35	199.99	601.42
SX²	4,853.00		5,073.30		8,391.63		4,794.79		5,689.08	
SY²		20,090.40		19,720.53		36,224.17		31,877.46		40,859.01
SXY		9,810.04		9,657.59		17,068.01		10,666.40		12,482.23

Totals: $SX = 1018.15$ $SY = 2418.97$ $SXY = 59{,}684.57$
$SX^2 = 28{,}801.30$ $SY^2 = 148{,}771.57$

but the analyses are broken into geographic suites.[3] The Loos-Hamra area is further subdivided by texture, as in von Eckermann's account.

In Table III the mean normative (Or + Ab) contents are shown to differ significantly between suites. Specifically, the differences between means are shown to be far larger than would reasonably be expected on the basis of intra-suite variations, and a homogeneous parent population[4]

TABLE III

Analysis of Variance of Normative (Or + Ab).

Source of Variation	Degrees of Freedom	Mean Square
Between Suite Means	4	287.17**
Within Suites	35	38.13

$$F = \frac{287.17}{38.13} = 7.53$$

In Table IV the same analysis is made for normative quartz, with opposite results. The differences between suite means are not significant in relation to within-suite variability. Comparison of the two tables shows that the variance of (Or + Ab)

(3) Analyses have been drawn from the following sources:
Vehmaa—Ilmari Kanerva, *Fennia*, 50, N:o 40, 1928.
Laitila—Th. G. Sahama, *C. R. Soc. Geol. de Finlande* N:o XVIII 1945
Lappeenranta—V. Hackman, *Bull. Comm. Geol. de Finlande*, N:o 106, 1934
Loos-Hamra—H. von Eckermann, *Geol. För*, 58, N:o 405, 1936, p. 130.
Tables IV and V of Sahama's useful paper give weight percentages and Niggli symbols for all the analyses shown here in Table II, as well as several more. Original references for the Laitila analyses will be found at the bottom of Sahama's Table V. All analyses used here have been checked against original references; norms were computed where lacking in the original.

(4) A description of the variance analysis used here and below will be found in G. W. Snedecor, *Statistical Methods*, 4th Ed., 1946. Procedure leading to tables III and IV is described in Chapter 10, and those who care to follow computations may do so with the sums, squares and products shown at the foot of Table II. Briefly, the method expresses the variability of data in terms of the sum of squared deviations; by suitable computations a portion of this "sum of squares" is associated with each source of variation, division by the degrees of freedom allotted to each source reduces the associated sum of squares to a mean square, or variance, and the significance test is finally made in terms of the ratios of pairs of these mean squares to each other. Tables for the comparison will be found in Snedecor, op. cit., pp. 222-225. For a concise outline of the method see Yule and Kendall, *An Introduction to Elementary Statistics*, 13th Ed., 1946. Pp. 444-449.

means is more than five times as great as the variance of Q means, but the within-suite variance of (Or + Ab) is just half that of Q. The normative alkali-feldspar content of rapakivi analyses differs much more than the normative quartz content from area to area, but within each area it is at least as stable and possibly more so.

TABLE IV

Analysis of Variance of Normative Q.

Source of Variation	Degrees of Freedom	Mean Square
Between Suite Means	4	52.96
Within Suites	35	76.25

$$F = \frac{52.96}{76.25} = 0.694$$

Now this is a very curious result if the relation between quartz and alkali-feldspar is nearly as close as it seems to me Dr. Backlund's hypothesis requires,(5) and it is particularly puzzling in view of the large negative correlation already drawn from Table I. This correlation makes it clear that there is a significant inverse variation between Q and (Or + Ab) in some part or parts of the data, and from Table III it is clear that the geographic grouping has generated markedly different subsamples. We may guess that the desired relation will be found in certain of the sub-samples and not in others. If it has not done so already, a moment's inspection of Table II will suggest regrouping the data into two broad geographic subdivisions. For those who care to follow the computations, the necessary sums, squares and products are given below in Table V, and Table VI gives correlation coefficients and computed values required to reach them. In Figure 1 the data of Table II are plotted according to the grouping of Table VI. Regression lines are shown for the Swedish rapakivis because

(5) Without passing judgment on whether or not there should be significant differences in either Q or (Or + Ab) between suites, it is obvious that the variance of Q ought not differ significantly from that of (Or + Ab). Since other constituents are allegedly constant, Q and (Or + Ab) are in effect coded values of each other, for Q + (Or + Ab) = 100 — C = K, where C represents other constituents. In theory the variances of Q and (Or + Ab) ought to be identical, and in practice they should not differ significantly. Hence if (Or + Ab) varies significantly from area to area, Q ought to do so also, and if Q does not then neither should (Or + Ab).

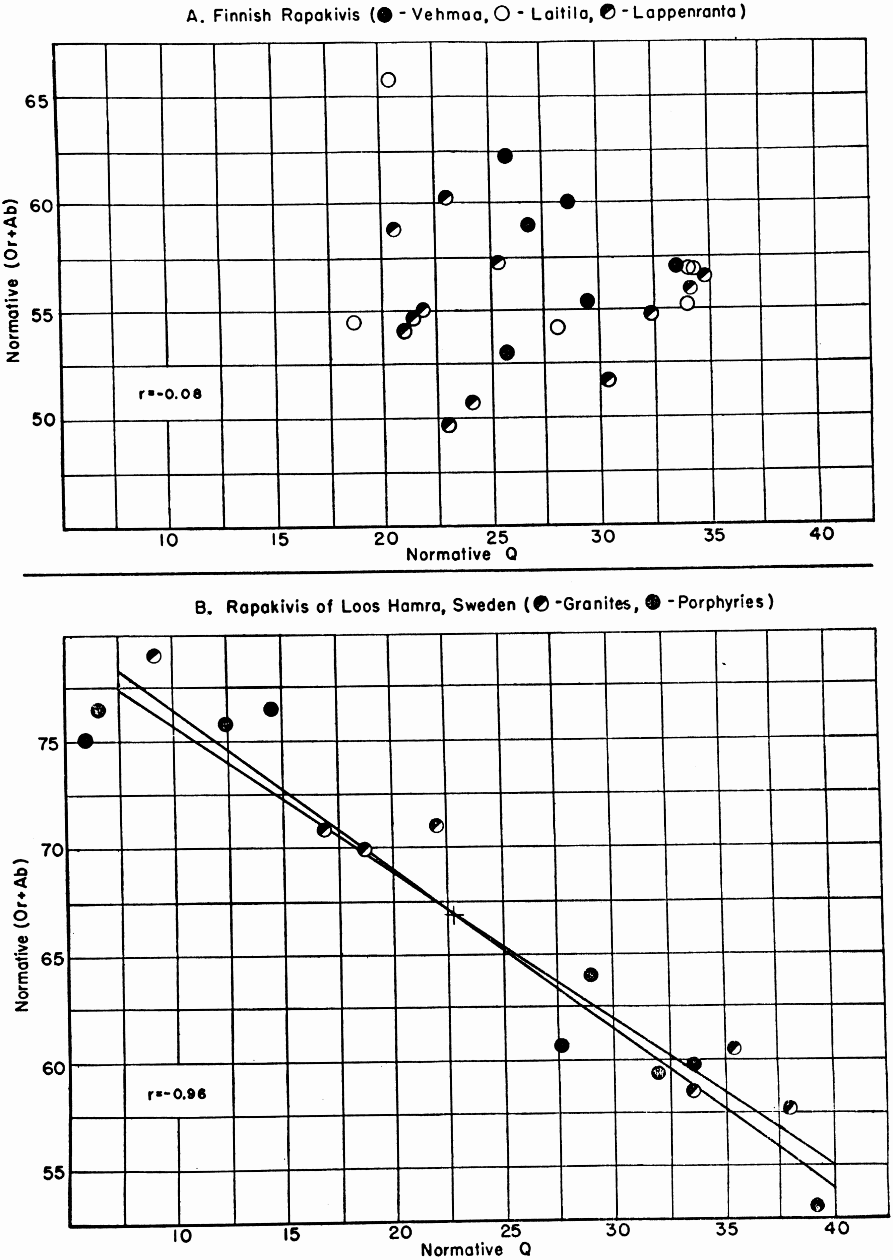

Figure 1. Normative Q and (Or + Ab) in Swedish and Finnish rapakivis. (Data from Table II.)

the correlation is easily high enough to render either line useful as a summary or reduction of the data. They have been omitted from the graph of the Finnish analyses because in these the correlation is so poor that the observed amount of one normative molecule in a given analysis gives no effective index of the observed amount of the other.

TABLE V
Data on Q and (Or + Ab) in Two Groups.

Area	SX	SY	SX^2	SXY	SY^2	N
Finland	650.98	1,348.20	18,317.93	36,535.64	76,035.10	24
Sweden	367.27	1,070.77	10,483.87	23,148.63	72,736.47	16
Total	1,018.25	2,418.97	28,801.80	59,684.27	148,771.57	40

First confining the interpretation of the correlation coefficients of Table VI to the samples themselves, it may be shown that r^2 is the proportion of the variance of one variable which may be interpreted as a linear function of the variations in a second variable, where r is the correlation coefficient connecting the two. For the Finnish rapakivi analyses less than 1 percent of the variance of normative (Or + Ab) is associated in this fashion with normative Q variations (and conversely), while for the Swedish rapakivi analyses the association is 93%.

TABLE VI
Correlation Data in Two Groups.

		Sums of Squares and Products of Deviations			Correlation (3) Coefficient
Area	Degrees of Freedom	(1) Sx^2	(2) Sxy	(1) Sy^2	
Finland	23	660.75	−36.42	293.22	−0.0827
Sweden	15	2051.35	−1436.25	1080.54	−0.9647

(1) $Sx^2 = SX^2 - \frac{(SX)^2}{N}$ (2) $Sxy = SXY - \frac{SXSY}{N}$

(3) $r_{xy} = \frac{Sxy}{(Sx^2 \cdot Sy^2)^{1/2}}$

Finally, passing to inferences about the parent populations (see Snedecor, op. cit. ftn. 5, pp. 148-153) from which these analyses were drawn:

1. The Finnish rapakivi analyses offer no reason for abandoning the null hypothesis (e.g., the hypothesis that norma-

tive Q and (Or + Ab) are uncorrelated in the parent population), and are consequently opposed to an explanation requiring negative correlation.

2. There is excellent reason to suppose that in the parent population from which the Swedish analyses were drawn, normative Q and (Or + Ab) are negatively correlated, and to this extent the Swedish sample is therefore in accord with the Backlund hypothesis, or any other hypothesis requiring or permitting such a relation.[6]

We have so far used a mineralogical rather than a stratigraphic definition of "sandstone." The Jotnian is very often, and perhaps dominantly, arkosic, and it may be well to point out that this does not affect the argument. A sandstone with some fixed amount of feldspar would serve as well as a pure sandstone, and in a sandstone of variable feldspar content there would already be a negative correlation of alkalis and uncombined silica. In the first case the feldspathization of quartz by alkali-aluminate emanations would generate the required negative correlation while in the second it would preserve or intensify an already existing relationship of this type. Thus the absence of significant negative correlation between alkali feldspars and quartz is quite as fatal to Backlund's argument whether the Jotnian is a sandstone, an arkose, or a variably arkosic sandstone.

SUMMARY

Where rocks are formed by replacement, the amounts of residual and introduced materials must vary inversely. If the two can be distinguished mineralogically or chemically, statistical analysis of the data ought to yield significant negative correlation coefficients.

(6) It may be useful to stress here that although rejection may be final, acceptance may not. Negative correlation of "residual" and "replacement" products is a necessary but not a sufficient condition for acceptance of the alkali emanation hypothesis.

Although it has no place in this argument, I should like to point out a very peculiar character of the Loos-Hamra rapakivis. For this suite correlations were computed for all pairs of the following normative minerals: Q, or, ab, an, px, ores. *All* correlations involving Q were negative, and despite the odds against positive correlation *all other correlations were positive*. It is very easy to believe that a large part of the quartz in these rocks stands in some replacement relation to all other constituents. Can it be that the Loos-Hamra rapakivis, initially syenitic, either have picked up quartz from the Jotnian sandstone prior to solidification, or have been silicified hydrothermally since solidification?

The condition that the sum of the variates is a constant insures that major constituents of rocks, whether stated as modes, norms, or oxides, will generally show negative correlation. A permissible but rather wasteful application of sample statistics may nevertheless be made if any hypothesis may be stated so that failure to achieve negative correlation is sufficient cause for rejection.

Forty analyses of rapakivi granites are used as a practical example, after reasons are given for requiring data in accord with the alkali-emanation hypothesis to exhibit significant negative correlation between normative Q and (Or + Ab).

The Swedish rapakivi analyses are shown to be in accord with the hypothesis, but in the Finnish analyses the correlation fails to achieve significance by a very large margin. The alkali-emanation hypothesis as advanced by Backlund does not afford a satisfactory explanation of variations in the bulk chemical composition of the rapakivi granites.

Geophysical Laboratory
Carnegie Institution of Washington
Washington, D. C.