SVERIGES GEOLOGISKA UNDERSÖKNING

SERIE C NR 727 AVHANDLINGAR OCH UPPSATSER ÅRSBOK 71 NR 2

LARS OLOF LARSSON

STATISTICAL TREATMENT OF IN-SITU MEASUREMENTS OF MAGNETIC SUSCEPTIBILITY

STOCKHOLM 1977

SVERIGES GEOLOGISKA UNDERSÖKNING

SERIE C NR 727 AVHANDLINGAR OCH UPPSATSER ÅRSBOK 71 NR 2

LARS OLOF LARSSON

STATISTICAL TREATMENT OF IN-SITU MEASUREMENTS OF MAGNETIC SUSCEPTIBILITY

STOCKHOLM 1977

ISBN 91-7158-110-3

Schmidts Boktryckeri AB
Helsingborg 1977

CONTENTS

ABSTRACT

The aim of this paper is to propose statistical methods for the division of complex rock magnetic susceptibility distributions. The investigated data include about 9000 in-situ susceptibility measurements on gabbroic rock massives in northern Sweden. For individual massifs most obtained distributions are complex multimodal in a logarithmic scale. A few monomodal distributions have been found. These monomodal distributions are assumed to be caused by different generating variables characterized by a certain variance, amplitude and position on the susceptibility scale. The generating variables are results of complex rock genetic processes including initial iron content, mode of differentiation, mode and depth of crystallization, and secondary processes acting on the iron containing minerals. The existence of monomodal distributions has led to the assumption that each complex distribution is composed of, at the most, five generating variables with constant variance. The division of complex distributions into subdistributions was finally carried out using statistical hypothesis testing. This method is described here and involves taking samples of neighbouring sites of measurements and testing these against the obtained generating variables.

TABLE ON SYMBOLS USED

M, m	mean value of a normally distributed variable
S^2	variance of a normally distributed variable
$\bar{c}$	sample mean
$\dot{S}^2$	sample variance
n	number of observations
p	significance level

INTRODUCTION

In the years 1972—1974 a large number of in-situ magnetic susceptibility measurements have been made by the Geological Survey of Sweden. In a first instance most larger gabbroic rock massives in Norrbotten in northern Sweden were investigated. The extreme variability of susceptibility in these rocks was known from laboratory measurements on rock samples, but the in-situ measurements showed even greater complexity. Among the obtained susceptibility spectra, a number of interesting findings appeared. The susceptibility seemed to constitute a discrete variable occurring sometimes as single distributions, but as a rule, as multimodal distributions. The occurrence of monomodal distributions encouraged the attempt to subdivide the complex distributions into a number of single subdistributions and, in the next step, try to correlate certain processes of formation or metamorphosis to the occurrence of these subdistributions. The first step, which is dealt with in this paper, is thus a preliminary statistical approach for the treatment of complex susceptibility distributions.

Before dealing with the susceptibility distributions obtained from measurements made by the Geological Survey, a number of the underlying assumptions for the construction of diagrams need to be elaborated.

SCALE TRANSFORMATIONS

In order to have nearly normally distributed sub-distributions, a transformation of the susceptibility (K) scale is necessary, as a linear scale means that certain sub-distributions will be strongly distorted. Four different scale transformations are illustrated in Fig. 1. The choice of a correct scale transformation is however dependent on how the spectrum is to be divided. A correct division requires in turn that we know the variables which generate susceptibility distributions. As these are unknown, it has been necessary to use the statistical properties of the observed monomodal distributions. Some rock massives have only a single frequency maximum (monomodal distributions) which is assumed to be generated by a single variable. The multimodal distributions are assumed to be generated by several such variables. If the effect of different scale transformations is studied on the monomodal distributions, it is found that a logarithmic transformation gives the best result. Among those transformations tried, the $^{10}\log$ tranformation gives the best stabilizing effect on the variance — one of M.S. Bartlett's (1947) conditions in finding a suitable transformation. In the ideal case:

(a) The variance of the transformed variate should be unaffected by changes in the mean level.

(b) The transformed variate should be normally distributed.

(c) The transformed scale should be one for which an arithmetic average is an efficient estimate of the true mean level for any particular group of measurements.

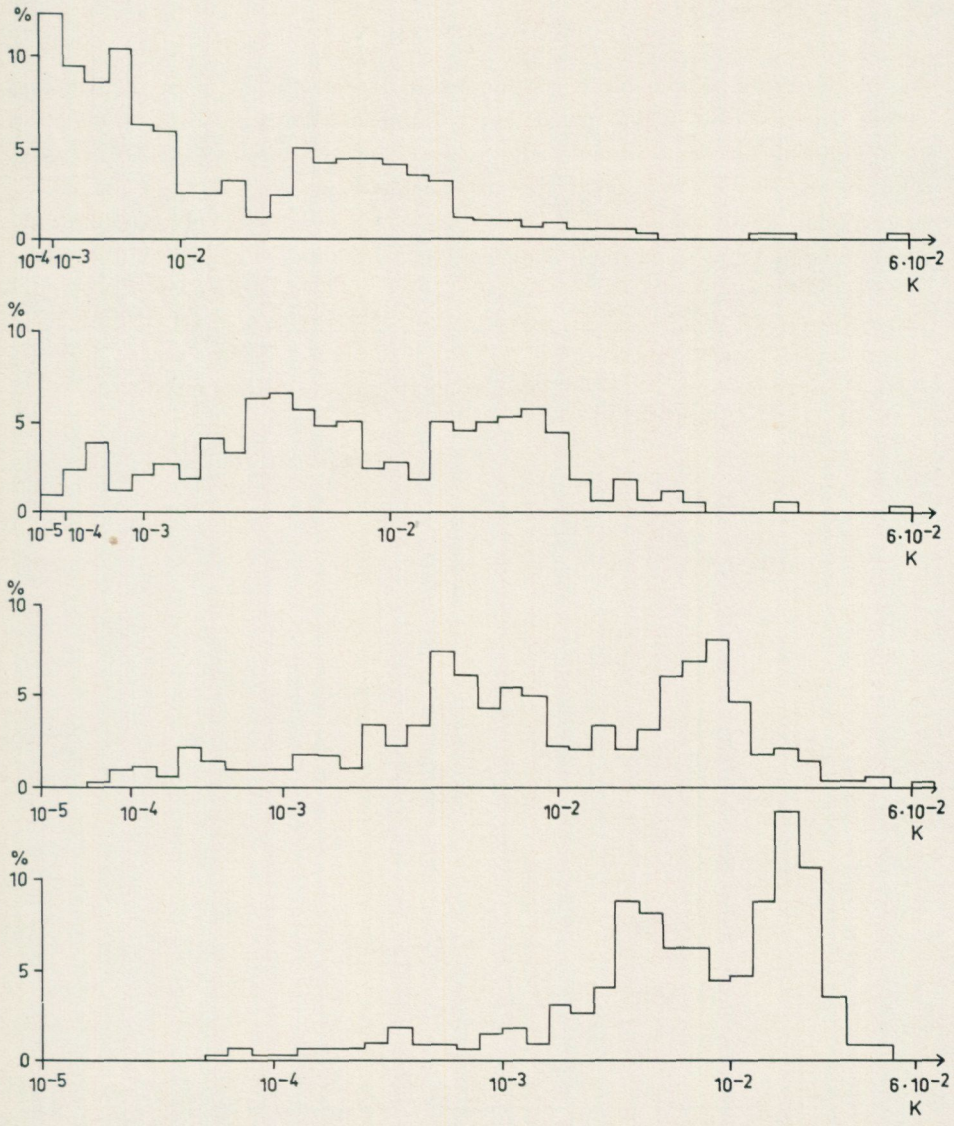

Fig. 1. Kalixfors gabbro, 340 measurements. Comparison of different scale transformations. From top: linear scale $\sqrt[3]{}$scale, $\sqrt[4]{}$scale, ^{10}log scale.

The variance of monodal distributions decreases somewhat when the mean occurs higher up on the K-scale (logarithmic transformation), but this effect (heteroscedasticity) is reversed and increases strongly with a square-root-transformation. The transformation $X=\sqrt[4]{K}$, also tried, means for instance that the variance of a distribution at the lower end of the scale (Skogsträsk, Fig. 5) is 1/5 of the variance of a distribution centered at 10^{-2} cgs.

CLASS DIVISION

As the susceptibility variable is continuous, the spectrum has to be divided into classes. In such cases commonly between 8 and 20 classes are used, and here a class width of 'one tenth of a decade' in the ^{10}log susceptibility axis is proposed. This means for example, that the measurements on the Vittangi gabbro (a monomodal distribution) occupy 10 classes. In a number of cases though, measurements from other rock complexes may require 30 classes; but in these

Fig. 2. Vittangi gabbro, 997 measurements. Class divisions used in this report.

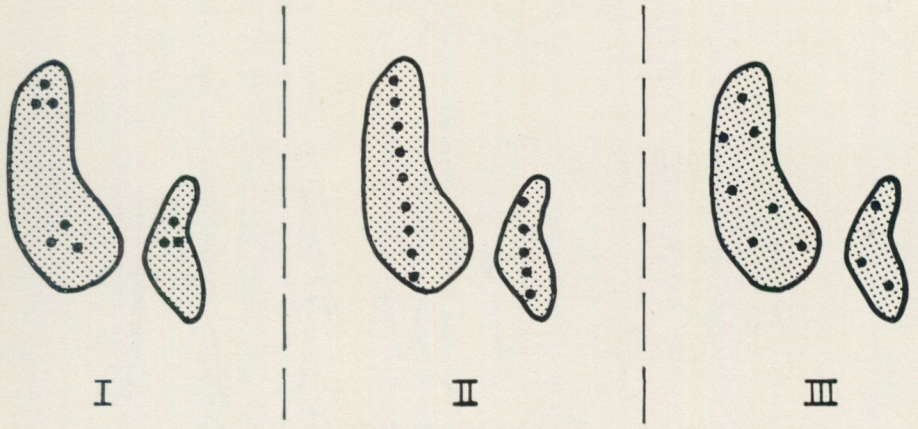

Fig. 3. Different sampling methods.

cases there are several underlying factors contributing to the final appearance of the total distribution.

After normalization the mean and variance can be expressed in classes according to Fig. 2, class 1 starts with $1 \cdot 10^{-5}$ cgs. (Vittangi, Figs. 2 and 5).

SAMPLE REPRESENTATIVITY

How should the measurements be made on the available outcrops so as to give a good estimate of the generating variables. Several different methods have been used and are illustrated in Fig. 3. The most common method has been to note points on the map, at distance of 50—100 metres and then making 3—5 measurements at each point, each measurement being 5—10 metres from the others (method 1). Another method is to draw profile lines across the outcrops with 20—100 measurements on each (method 2).

A better estimation of the susceptibility for a rock unit is possible however if the measurements are spread evenly over the entire area of outcrop (method 3). This method is more tedious than the other two. Each measurement must be marked on the map so as to be able to combine measurements from geographically neighbouring points. This method also requires that the greater part of the rock area is covered. With methods 1 and 2 the risk of "over representing" local pecularities increases.

When measurements for each test area are combined, these should contain at least 10 measurements to assure a good estimation of the generating variables. If the number of measurements made per area is too small, then the random variations may have an undesirably large influence.

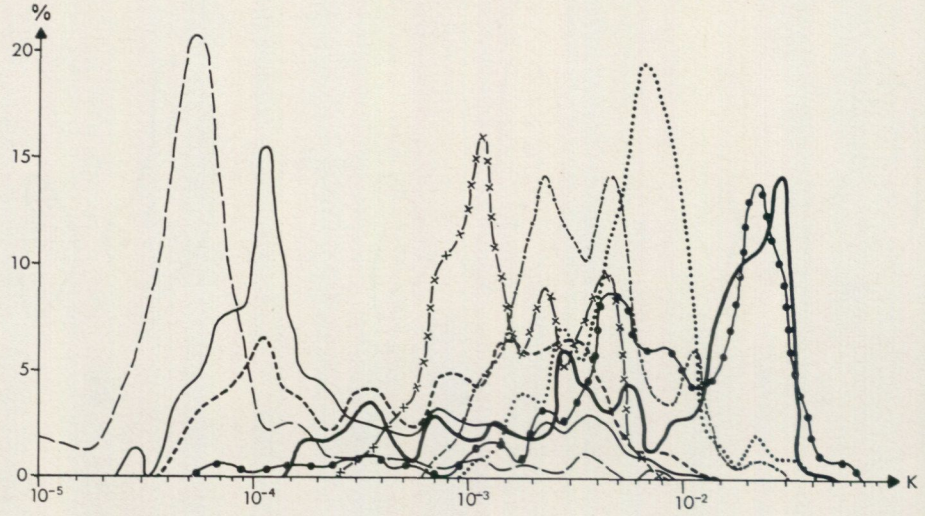

Fig. 4. Total distribution of susceptibility from 15 gabbro massifs in Norrbotten.

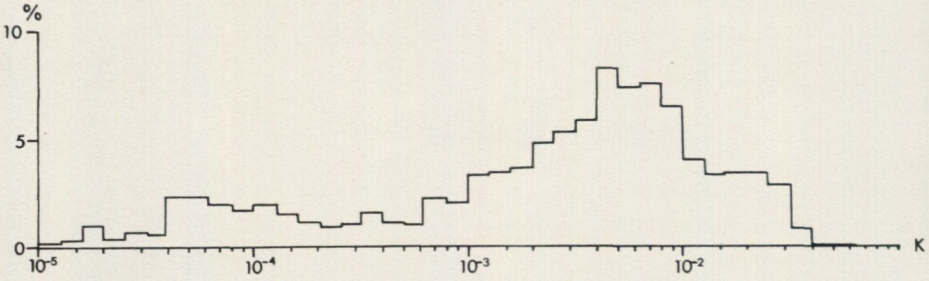

Fig. 5. Susceptibility distributions from some different gabbro massifs. Class width 0.1 decade.

SUSCEPTIBILITY MEASUREMENTS ON GABBROIC ROCKS IN NORRBOTTEN

The measurements on gabbro made by the Geological Survey have generally a complex spectrum compared with most reported distributions. Data from some areas are monomodal or bimodal, but the majority have multimodal distributions with up to 6 maxima. (Chernyuk 1971).

This article considers 17 areas of Norrbotten for which a minimum of 150 measurements are available. The measurements were made on well specified diorite-gabbro massifs where there are plenty of outcrops.

Fig. 4 shows the K-spectrum of all 17 areas with a total of 9 000 measurements. The areas all have the same weight in the distribution in order to minimize the effects of areas with a large number of measurements (in some cases over 1 000) which otherwise would dominate the final total distribution. The total distribution obtained finally in Fig. 4 follows Chernyuks (op. cit.) hypothesis that the susceptbility distributions are either mono- or bimodal, representing the paramagnetic and/or the ferromagnetic contributions. The frequency distributions for individual areals, however, do not agree with this hypothesis. In general the individual spectra have more than three frequency maxima. Fig. 5 shows the different distribution types obtained by the Geological Survey.

The following statements follow directly from Fig. 5:

1. Some massives show a distinct maximum in the region $K \leq 10^{-4}$ cgs.
2. Other areas show a maximum near $2 \cdot 10^{-2}$ cgs.
3. The majority of the measurements have maxima within the interval $10^{-3} \leq K \leq 10^{-2}$ cgs and can occur at all susceptibility levels within this interval.

The following hypothesis takes into account known facts which explain the

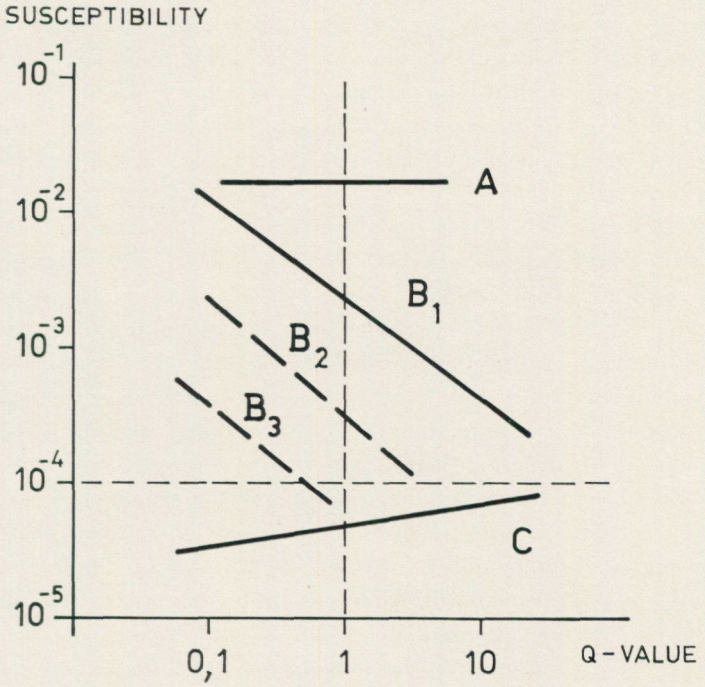

Fig. 6. Observed relations between susceptibility and q-value.

origin of the susceptibility distributions in a certain rock type (Nagata 1961). These facts show that susceptibility is:

(i) Due to the volume fraction of paramagnetic minerals and their Fe-content. This gives distributions with an upper limit at $K \leq 4 \cdot 10^{-4}$ cgs.

(ii) Due to the volume fraction of magnetite and its grain size. This leads to distributions with maxima in the range $3 \cdot 10^{-4}$ to $4 \cdot 10^{-2}$ cgs.

Distinct trends observed in susceptibility q-value (ratio of remanent to induced magnetization in rock samples) diagrams of grabbroic rocks show how different K-spectra are caused by various combinations of trends. (Fig. 6.)

Preliminary investigations show that these trends are caused mainly by the frequency of exsolution processes among the magnetic opaque minerals. (Henkel 1976). Projection on the K-axis gives the following distributions:

$A + B_1$	$K \leq 10^{-4}$ cgs.	low magnetic
$B_1 + B_2$	or some other combination	magnetic
$B_1 + B_2 + B_3$		
$B_2 + B_3$		
$B_3 + C$	$K = 2 \cdot 10^{-2}$ cgs.	high magnetic

The above facts and hypothesis are not sufficient to justify a division of the susceptibility distribution into component sub-distributions. For this, the generating variables for the distributions must also be known. These variables cannot easily be found, or approximated. Commonly used methods of fitting normal distributions to the experimental data are not possible for the complex distributions found here.

The problem centres on the complicated interval B (above). The low susceptibility distribution is known empirically and may be derived from theory. This shows the mean to be approximately located at $K \leq 10^{-4}$ cgs. An estimate of the variance $S^2 = 3.2$ class units is available from the calculated value for a monomodal distribution in this range (Skogsträsk, Fig. 5). The calculation of this variance ignores the right hand tail of the distribution (which is due to magnetite in the rock). Examination of data from the 17 regions reveals that half of these have a distribution within an almost constant interval centered at $2.2 \cdot 10^{-2}$ cgs. The variance for these distributions is smaller ($S^2 \approx 2.0$). When considering the distributions within interval B, it is observed that frequency maxima occur at all susceptibility levels within the interval. The Vittangi gabbro is an example of a monomodal distribution within this interval having its mean at $8.8 \cdot 10^{-3}$ cgs and the variance 2.5 class units.

We notice that the variance of the monomodal distributions seems to decrease when the mean values increase (heteroscedasticity). (Fig. 7.)

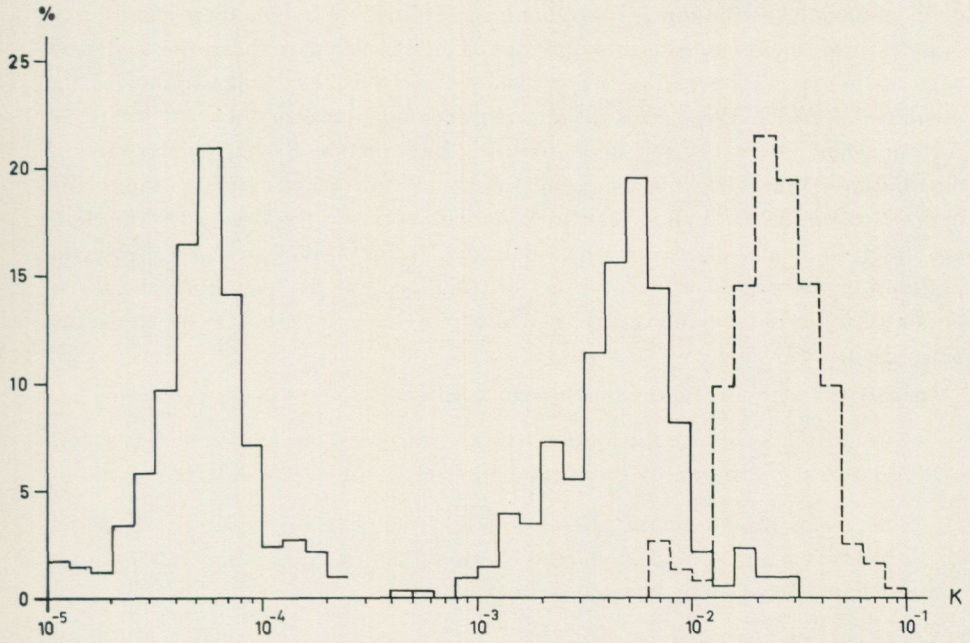

Fig. 7. Some different monomodal distributions.

SUBDIVISION OF SUSCEPTIBILITY DISTRIBUTIONS

Because of the lack of information on the generating variables it was decided to assume the more complex distributions to be a combination of a finite number

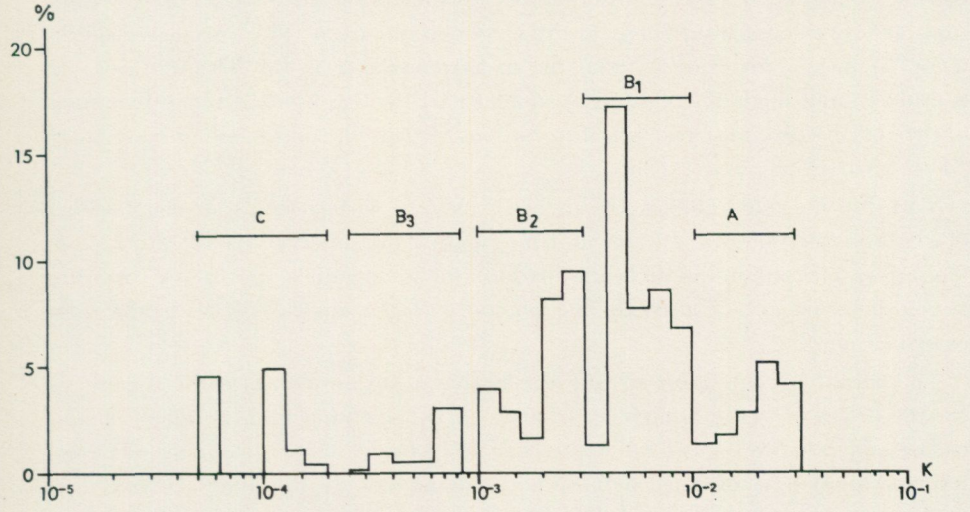

Fig. 8. Distribution of susceptibility modes.

of monomodal distribuitons. The monomodal distributions are then regarded as samples from ideal Gaussian distributions. This implies that the mean and variance of the generating variables are estimated by the mean and variance of the monomodal distributions. This assumption does not however solve all problems. An important interval of the susceptibility scale (interval B) has no monomodal distributions. When deciding on mean values for generating variables within this interval, the same methods as used by Chernyuk (op.cit.) has been selected. Here, the observed modes (frequency maxima) are assumed to represent roughly the positions of the means of each normal distribution. Fig. 8 then, shows the distribution of frequency maxima, each of which represents at least 2 % of a susceptibility spectrum.

In carrying out the subdivision the following rules were used;

(i) The variance of subdistributions should be constant at all K levels and equal to the variance of the monomodal distributions ($S^2 \approx 3$ class units).

(ii) The mean of the subdistributions should be near the observed mode for a certain massive, and preferably within the intervals where all massives generally have their frequency maxima.

It has been noticed (Fig. 7) that the variance of the monomodal distributions decreases when a distribution occur higher up on the K-scale. This indicates that the scale transformation is not perfect — the variance should be constant at all susceptibility levels.

This means that the variance requirement ought to vary according to the observed variance values at different susceptibility levels. The variance requirement (i) used here though is a fix value, partly because the hypothesis testing becomes much easier, and partly because the four observed monomodal distributions are considered insufficient to draw conclusions from. The variance requirement (i) means, for example, that the measurements from the Vittangi area are considered as a single distribution. A smaller variance implies that this distribution has to be divided into two sub-distributions, which is considered unreasonable. (This occurs if $S^2 = 1$.)

To fulfill the requirements (i) and (ii) above it is suggested that mean values of the sub-distributions are taken from within the intervals A—C in Fig. 8. This gives a maximum of five sub-distributions for complex spectra. A spectrum is then believed to be caused by five generating log normal variables in the most extreme case.

The validity of this proposal and the degree of agreement with the hypothesis for the origin of susceptibility must be tested by analysing rock samples (chemically + mineralogically) which represent the various sub-distributions. In this test, the division, as outlined here, should only be seen as a preliminary investigation.

GEOGRAPHICAL LOCATION OF SUB-DISTRIBUTIONS

For continuous studies of different kinds it is necessary to identify the sub-distributions geographically. This requires that each measurement can be assigned to a special sub-distribution. It is necessary therefore to have some pre-knowledge of the sub-distributions. The technique suggested is supposed to give approximations of the generating variables (mean values and variances). The simplest method would be to refer each measurement to the nearest variable. Observation 1 in Fig. 9 is then assigned to the normally distributed variable A and observation 2 to variable B.

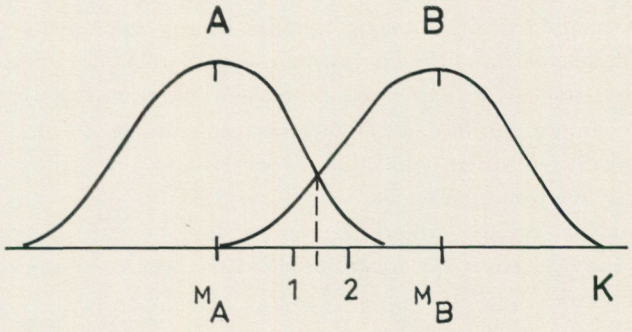

Fig. 9. Overlapping normal distributions.

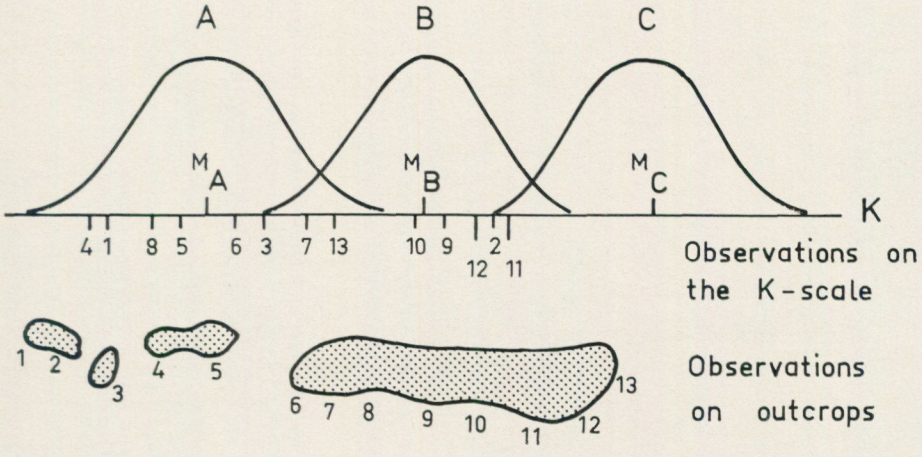

Fig. 10. Susceptibility distributions and their relation to the position of measurements.

The measurements assigned to a variable according to this simple method will however not be normally distributed. In the example above, a frequency distribution including all observations counted to variable A has no right hand tail. The method ignores at the same time the geographical dimension of the problem. The division of a total distribution should be aiming at consistent sub-distributions with respect to both susceptibility and geographic position. The method thus has to be completed with a rule saying that measurements assigned to a normally distributed variable must be close to one another geographically. The method described above leads to a great number of laborious choices. How much can, for instance observation 2 in Fig. 10 diverge from m_A (m_A = mean value of the normally distributed variable A) and still be assigned to the normally distributed variable A.

Difficult choices arise as soon as an observation occurs in an interval where two normally distributed variables overlap. If we instead use a statistical hypothesis testing method, then the division will be based on more objective criteria. In short, this method involves taking samples from neighbouring measurements and testing these against different normally distributed variables. How this is achieved in practice may be explained using the example in Fig. 10. The observations 1—8 are counted to the normally distributed variable A. The measurements are geographically consistent and this sample's mean ($\bar{c}$) and variance (S^2) differ only slightly from the mean (m_A) and variance (S^2_A) of the normally distributed variable A. The preceding requires that the null hypothesis (1) and (5) are accepted. (Normally distributed observation variable with known variance, Angsmark 1971.)

$$H_0 : S^2 \leq S^2_A \qquad (1)$$

$$H_1 : S^2 > S^2_A \qquad (2)$$

The test used is:

$$X^2 = \frac{(n-1)\,\dot{S}^2}{S^2_A} \qquad (3)$$

The critical region is:

$$X^2 > X^2 p\ (n-1) \qquad (4)$$

The null hypothesis is rejected for large X^2 values.

$$H_0 : m = m_A \qquad (5)$$

$$H_1 : m > m_A \qquad (6)$$

Test function:

$$Z = \frac{\bar{c} - m_A}{S_A} \sqrt{n} \qquad (7)$$

Critical region:

$$Z > Z_p \qquad (8)$$

The null hypothesis is rejected for large Z values.

If we count all thirteen measurements in Fig. 10 to the sample, then the null hypothesis (1) and (5) must be rejected. The sample is consequently defined by the hypothesis testing. In the sub-division of the susceptibility distribution given below, geographically consistent measurements are assigned to the sample as long as the variance requirement $S^2 \leq 3$ is satisfied.

The sample is then tested to find which normally distributed variable it could represent. Throughout this test the significance level p has been 1 %. This significance level specifies the risk taken that the null hypothesis is rejected when in fact it is correct. (A type I error.) When the risk of a type I error is small, as in this case, the risk of a type II error is large (i. e. the risk taken that the null hypothesis is accepted when in fact it is wrong). This problem will not be persued further here. It can just be noted that another significance level, for instance 5 %, often leads to samples with a smaller number of measurements. Samples including extreme observations accepted when the significance level 1 % was used, will sometimes be rejected with the significance level 5 %. The suitability of the significance level used can be controlled with better knowledge of the generating variables. It is then possible to investigate if the samples obtained fulfil requirements on samples from special variables.

Hypothesis (5) and (6) and the critical region (8) are variable and depend on the position of the normally distributed variable within the total distribution. This can also be demonstrated using Fig. 10. In order to test whether a sample is assigned to the normally distributed variable B, the hypotheses are as follows:

$$H_0 : m = m_B \qquad (9)$$

$$H_1 : m \neq m_B \qquad (10)$$

The critical region is:

$$|Z| > Z_{p/2} \qquad (11)$$

In this case the null hypothesis is rejected for both large and small deviations from Z.

Finally, in order to test whether a sample is to be assigned to the normally distributed variable C these new hypotheses apply:

$$H_0 : m = m_C \qquad (12)$$

$$H_1 : m < m_C \qquad (13)$$

and the critical region here is:

$$Z < -Z_p \qquad (14)$$

SUBDIVISION OF THE SUSCEPTIBILITY DISTRIBUTION OF THE TAAVINUNNANEN GABBRO COMPLEX

Taavinunnanen, Fig. 11, has a complex susceptibility distribution that includes nearly the whole range of values found for the Norrbotten gabbro measurements. In dividing the total distribution into sub-distributions, the first approximation is to take the means for five generating normally distributed variables according to the scheme outlined above.

Fig. 11 shows the mean values chosen. These all lie within the intervals presented in Fig. 8. The variance of the normally distributed variables is assumed to be 3 class units, and is approximately equal to the variance of the observed monomodal distributions. The next step is to make new sample regions of neighbouring measurements and to test to which generating variable the sample then belongs. This is made possible by the above hypothesis testing scheme.

Two types of anomaly in the expected pattern were noted when this was done:

(i) For small samples ($n = 3$ to 5) where the sample mean ($\bar{c}$) lies between two normally distributed variables it may happen that a sample could be assigned to either variable on the testing basis. In such cases, the sample is assigned to the nearest normally distributed variable ($|\bar{c} - m|_{min}$).

(ii) In other cases, it may be found that no normally distributed variable has been accepted. This happens for samples $n > 10$, where $\bar{c}$ lies between two normally distributed variables and where the null hypothesis $S^2 \leq 3$ has been accepted. In most cases the effect is due to an extreme observation displacing the sample mean. The sample mean of observations 1—8 in Fig. 10 has been displaced because observation 2 is relatively high. Problems of this kind are corrected by removal of the extreme tails of the sample.

A large number of anomalies of this type imply that the means of the normally distributed variables are wrongly located compared to observed frequency maxima. Fig. 12 and the corresponding map show the results of the subdivision for the Taavinunnanen area.

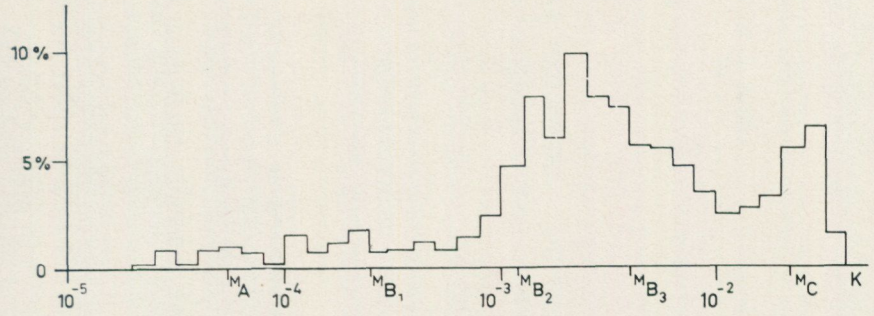

Fig. 11. Total distribution of the Taavinunnanen gabbro.

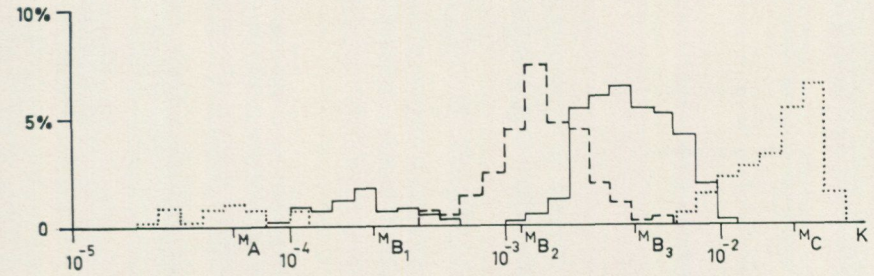

Fig. 12. Subdivision into monomodal distributions of the Taavinunnanen gabbro.

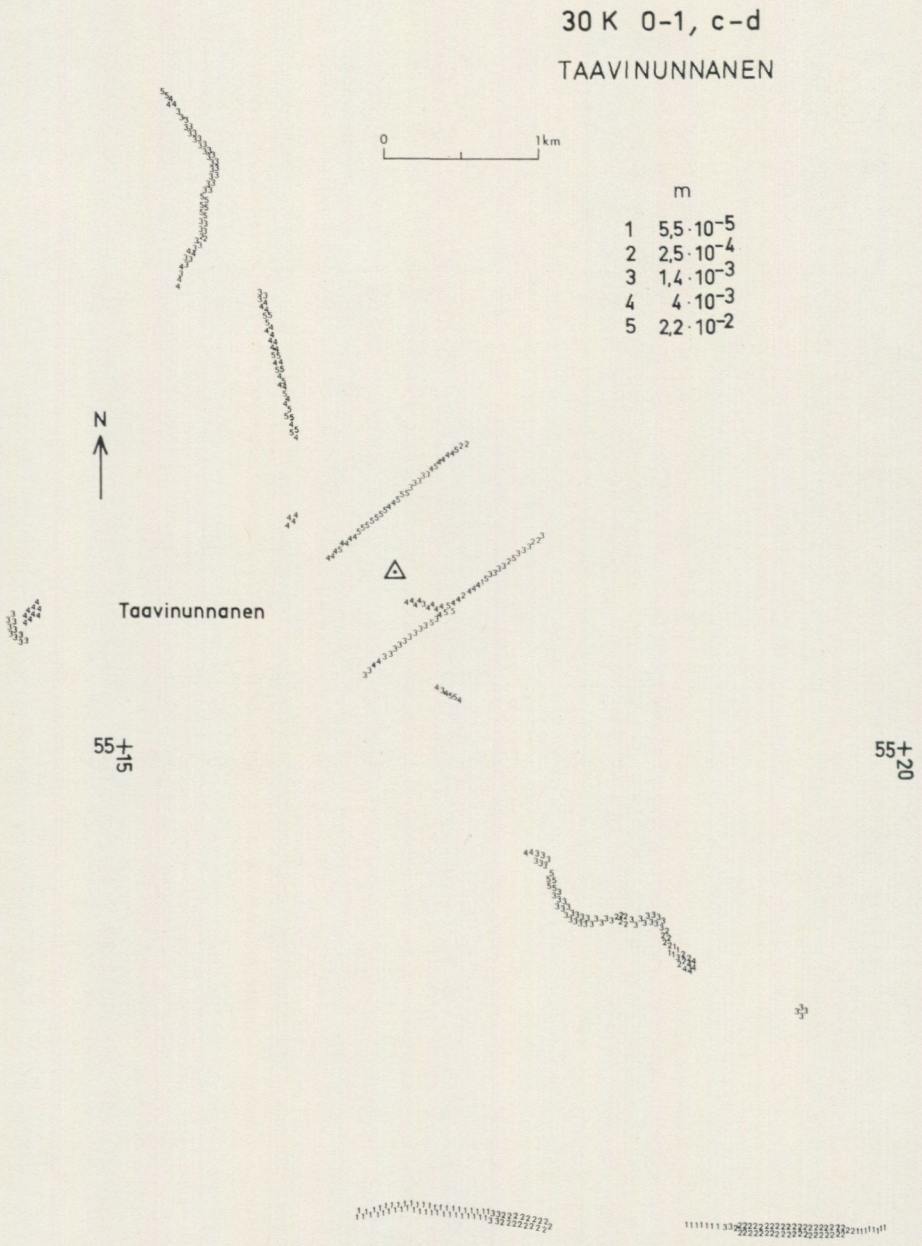

Fig. 13. Areal distribution of susceptibility sub-distributions of the Taavinunnanen gabbro.

SUBDIVISION OF THE SUSCEPTIBILITY DISTRIBUTION
OF THE KROKVIK—RAUTAS GABBRO COMPLEX

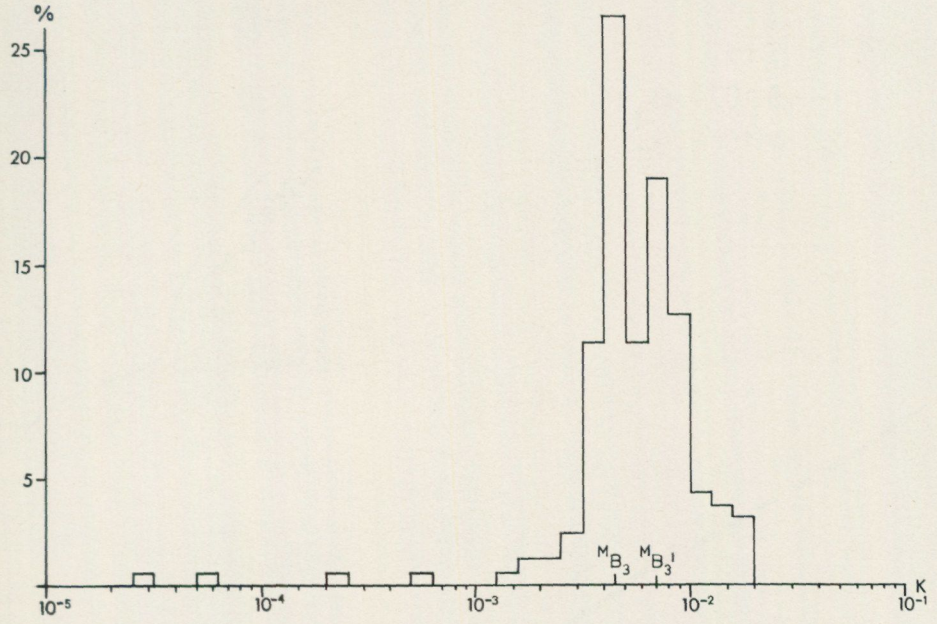

Fig. 14. Total distribution of the Krokvik—Rautas gabbro.

Krokvik—Rautas is a unique area, in that it has two frequency maxima lying close together, Fig. 14 (n.b. four extreme observations have been ignored). The variance is 3.5 class units, so the Krokvik—Rautas gabbro can probably be regarded as a monomodal distribution. The two maxima lie within the interval B_3 in Fig. 8.

The reason for the two adjacent frequency maxima found here is unclear. The complexity of the susceptibility distributions may be attributed to differences in both primary and secondary processes. This may lead to frequency maxima at almost all susceptibility levels. Differences in primary processes result in differences in rock composition. Chemical and mineralogical analyses of rock samples from adjacent distributions (as for the Krokvik—Rautas area) should answer the question of whether it is reasonable to divide a distribution which has only small susceptibility variations.

To divide the Krokvik—Rautas area, the variance was set to = 1 class unit to give good correspondence between a theoretical distribution and the observed

experimental one. The result of the subdivision is shown on the map. (M_{B_3} in Fig. 14 is marked with 1 on the map and M_{B_3}, with 2).

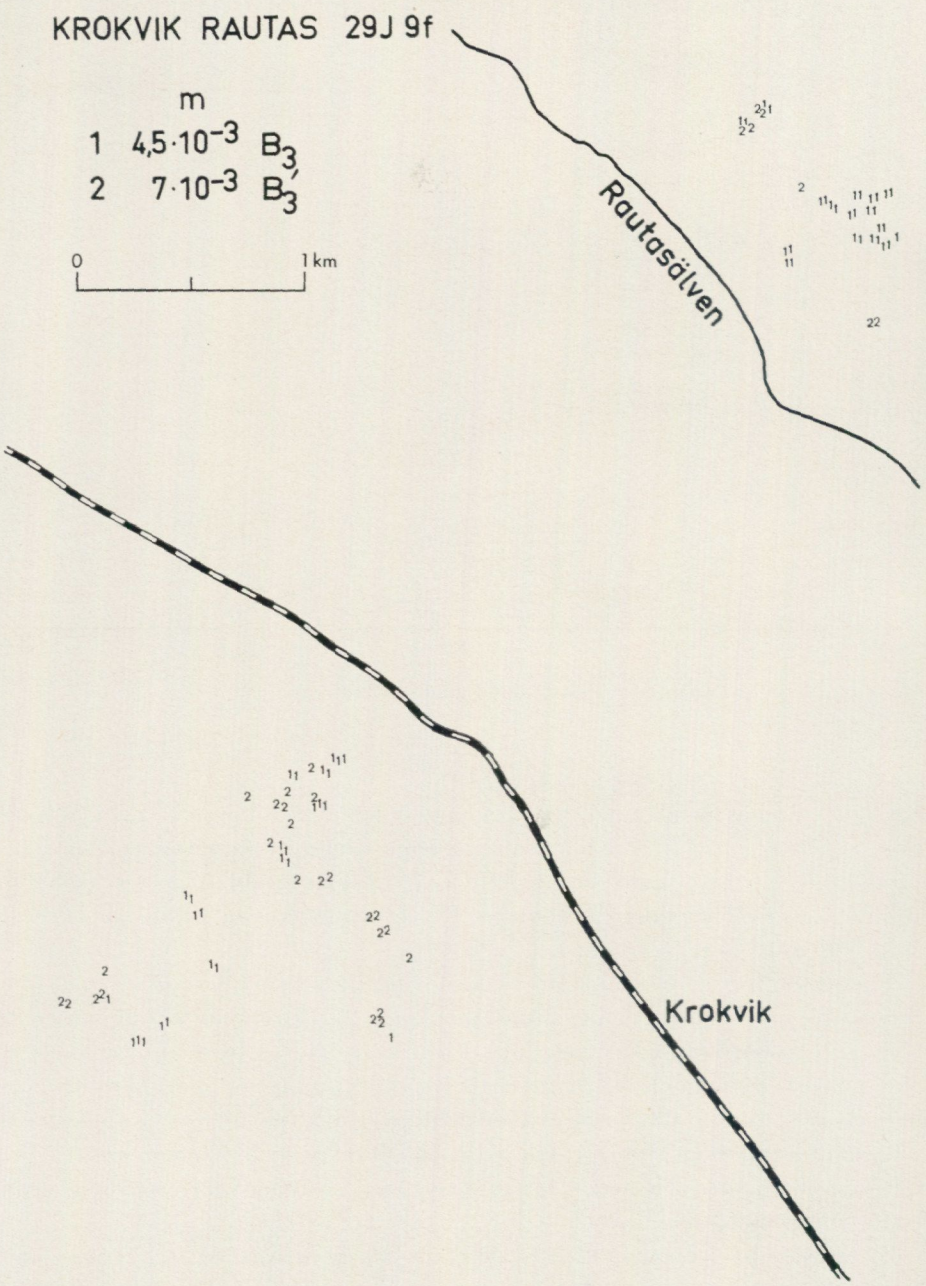

Fig. 15. Areal distribution of susceptibility sub-distributions of the Krokvik—Rautas gabbro.

DISCUSSION

Most of the discussion above is based on the fundamental assumption that the generating susceptibility variables can be estimated by the monomodal distributions. It has not been possible to derive the generating variables from theory.

The monomodal distributions were used first to find a suitable transformation. The transformation $X = {}^{10}\log K$ was preferred since it gave the best stabilizing effect on the variance. The monomodal distributions were used when estimating means and variances of the generating variables. These generating variables were finally the basis of the hypothesis testing which made it possible to show the location of each subdistribution on the map. In this context the hypothesis testing is supposed to represent an objective method for grouping samples from different generating variables. It need to be emphasized that the assumptions and methods used in order to divide complex susceptibility distributions must be tested in pracice. The Krokvik—Rautas gabbro has been included in order to make it possible to investigate whether it is reasonable to divide a susceptibility distribution into a larger number of sub-distributions. If the chemical and mineralogical investigations show that there are important differences between the two separated sub-distributions B_3 and B_3' for the Krokvik—Rautas gabbro, then the proposed general subdivision pattern probably has to be revised. Differences of this kind indicate that the interval B (above) is more complex than assumed in this paper.

The validity of the proposed subdivision for complex distributions could be tested the same way. Analyses of rock samples representing the five various sub-distributions should answer the question of wheter the subdivision is reasonable.

The lack of information on the generating variables prevent us from making general conclusions. Provided that the basic assumptions are correct, it is easy to realize the advantages of the methods proposed. The susceptibility data can be grouped and the sub-distributions shown geographically with a minimum of subjective decisions. This also makes it possible to use a computer, which in turn facilitates treatment of greater quantities of data.

REFERENCES

ANGSMARK, G., 1971: Statistisk inferens. — Studentlitteratur.
BARTLET, M. S., 1971: The use of transformations. — Biometrics.
CHERNYUK, M. V., 1971: Distribution of magnetic susceptibility on intrusive rocks. — Izvestia, Earth Physics, No. 3.
HENKEL, H., 1976: Studies of density and magnetic properties of rocks from northern Sweden. — PAGEOPH 114.
NAGATA, T., 1961: Rock magnetism. — Maruzen, Tokyo.

PRISKLASS C

Distribueras genom
LiberKartor
162 89 VÄLLINGBY

Schmidts Boktryckeri AB
Helsingborg 1977

ISBN 91-7158-110-3