

**R-04-66**

**Statistical analysis of fracture data,  
adapted for modelling Discrete  
Fracture Networks-Version 2**

Raymond Munier  
Svensk Kärnbränslehantering AB

August 2004

**Svensk Kärnbränslehantering AB**

Swedish Nuclear Fuel  
and Waste Management Co  
Box 5864  
SE-102 40 Stockholm Sweden  
Tel 08-459 84 00  
+46 8 459 84 00  
Fax 08-661 57 19  
+46 8 661 57 19



ISSN 1402-3091

SKB Rapport R-04-66

# **Statistical analysis of fracture data, adapted for modelling Discrete Fracture Networks-Version 2**

Raymond Munier

Svensk Kärnbränslehantering AB

August 2004

# Preface

This report originally occurred as an appendix to a strategy report for geological modelling /Munier et al, 2003/. Strategy reports were intended to be successively updated to include experience gained during site investigations and site modelling. Rather than updating the entire strategy report, we choose to present the update of the appendix as a stand-alone document. This document thus replaces Appendix A2 in /Munier et al, 2003/.

In short, the update consists of the following:

- The target audience has been broadened and as a consequence thereof, the purpose of the document.
- Correction of errors found in various formulae. All expressions have been rewritten.
- Inclusion of more worked examples in each section.
- A new section describing area normalisation.
- A new section on spatial correlation.
- A new section describing anisotropy.
- A new chapter describing the expected output from DFN modelling, within SKB projects.

# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Purpose</b>	<b>9</b>
<b>3</b>	<b>DFN parameters</b>	<b>11</b>
3.1	Orientation	11
3.1.1	Correction for orientation bias	14
3.1.2	Identification of fracture sets	16
3.1.3	Orientation distribution and its moments	17
3.1.4	Shape of the pole cluster	19
3.1.5	Comparison of stereograms	21
3.2	Fracture size distribution	24
3.2.1	Description of trace lengths	25
3.2.2	Computation of fracture sizes from trace lengths	30
3.2.3	Comparison of (mean) trace lengths	31
3.2.4	Area normalisation	33
3.3	Fracture density	36
3.3.1	Borehole and scanline	36
3.3.2	Outcrop	38
3.3.3	Computation of 3D fracture density	39
3.4	Evaluation of the spatial pattern of fractures	39
<b>4</b>	<b>Expected Output from DFN analyses</b>	<b>49</b>
<b>5</b>	<b>References</b>	<b>55</b>

# 1 Introduction

A geological model can be said to consist of three parts or aspects: A lithological model that describes the geometries and properties of various lithological units and/or domains and a structural model that equivalently describes the geometry and properties of deformation zones at various scales. The third aspect is the fracture array within lithological units. But fractures and small deformation zones that are too small to be described deterministically are therefore instead described statistically in terms of various distributions, and their relations. The fractures and small scale zones are described using DFN (discrete fracture network) parameters, which define the sizes, directions, and spatial distributions of the fractures, together with other characteristics such as mineralogy, transmissivity, etc.

RVS models, and their inherent DFN parameters, are subsequently used in DFN codes specially developed for that purpose, for analysing how the fractures influence, for example, stability, flow and transport in and through the rock mass.

## 2 Purpose

The present report describes the parameters which are necessary for DFN modelling, the way in which they can be extracted from the data base acquired during site investigations, and their assignment to geometrical objects in the geological model.

The purpose here is to present a methodology for use in SKB modelling projects. Though the methodology is deliberately tuned to facilitate subsequent DFN modelling with other tools, some of the recommendations presented here are applicable to other aspects of geo-modelling as well. For instance, we here recommend a nomenclature to be used within SKB modelling projects, which are truly multidisciplinary, to ease communications between scientific disciplines and avoid misunderstanding of common concepts.

DFN modelling can be broken down into two main tasks: One task consist of a block of exploratory data analyses, typically seeking various correlations, fitting distributions and testing hypotheses, another task, synthesis, consist of ennobling the result of the exploratory analysis to a self-contained DFN model. Exploratory data analysis, which mainly consists of basic statistics, can to some extent be standardised and we therefore provide guidelines for how results of such analysis could be presented in order to enable consistency in SKB reports, and to ensure that all receivers of DFN models within SKB projects base their DFN reasoning on a common ground. The synthesis of base statistics into a DFN model is, however, a complex task that requires skilled expertise and cannot be readily dictated by a guideline or methodology. The methodology must often be adapted to local conditions, and may also be tainted by style such that different DFN modellers may use disparate conceptual assumptions. Thus this report can only provide principles and guidelines aimed at promoting the integrative and iterative interpretation of fracture information. Accordingly, this report focuses mainly on the exploratory aspect of DFN modelling and since it is not intended as a tutorial, the reader is kindly expected to, when necessary, address cited references for technical details.

This report is structured as follows: In Chapter 3 we advocate for the need of a initial block of formal, exploratory data analysis using a worked example. The chapter also contains recommendations of a nomenclature to be used when reporting DFN-related work to SKB. In Chapter 4 we propose a structure for the various modelling steps described in Chapter 3 and how these can eventually be synthesized into a DFN description. A set of flowcharts briefly summarises the expected minimum output from each modelling step and from the synthesis.

### 3 DFN parameters

In order to create a fracture network, the following minimum amount of information is required:

- orientation of the fractures,
- size of the fractures,
- fracture density,
- fracture termination (how fractures cross each other), and
- the spatial distribution of the fractures (the fractures relative position in space, correlation).

This key information is preferably given in terms of statistical distributions, with appropriate moments and correlations. Other parameters of interest, such as mineralogy, water-conducting properties, rock mechanical properties, etc. are not treated here. Such parameters can be determined using conventional tools, if necessary, and, together with the geometrical aspects of the DFN parameters described here, can be used as a basis for more or less complex DFN models for different uses and users. Many of these parameters are estimated by similar techniques and we therefore also provide guidelines for such standard statistical analyses to endorse coherent terminologies, presentations, etc.

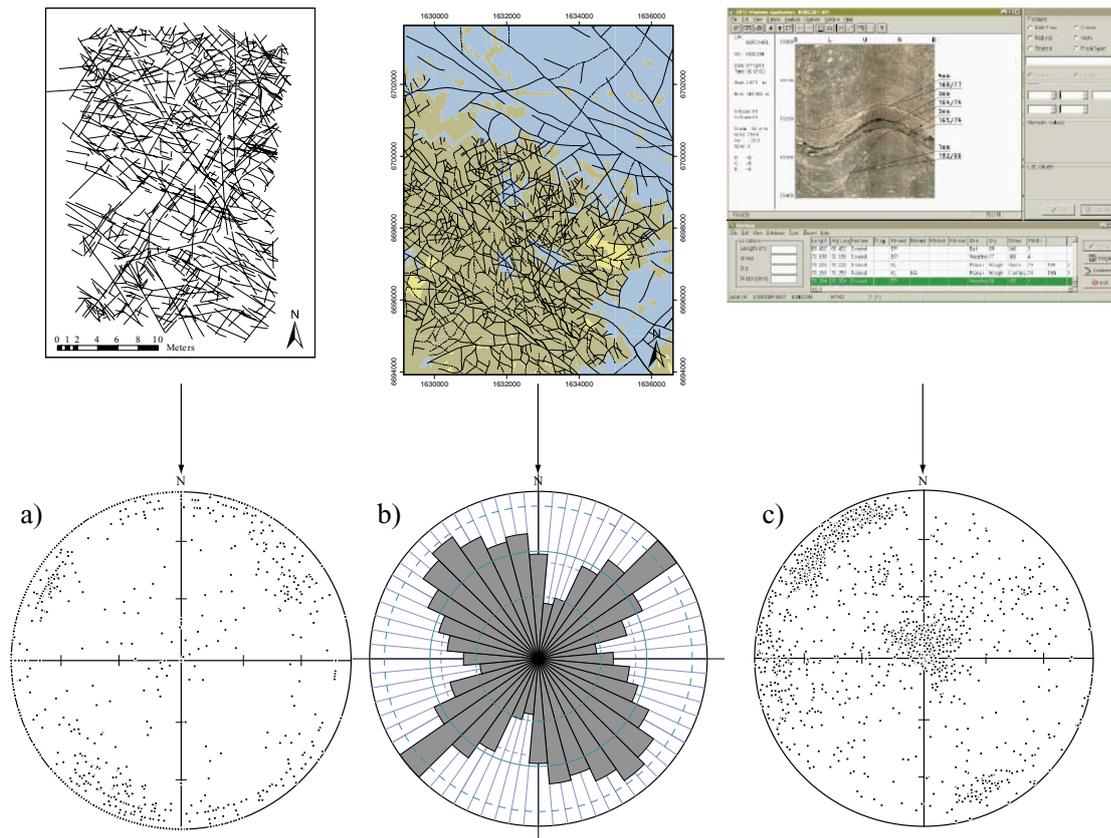
In the present section, we describe geometry-related parameters and how they could be determined and presented. Note that the analysis of fracture statistics is a very wide scientific subject area, and there exist many more analysis methods than those presented below. The methods presented here are, however, judged to provide adequate information for a proper construction of “base” DFN models.

#### 3.1 Orientation

Fractures are rarely randomly oriented in the bedrock. Commonly, one or several dominating orientations occur, hereafter called *fracture sets* /see, for example, Strähle, 2001 for nomenclature/. The *fracture array* is the entire ensemble of such sets, and their properties, i.e. the fracture population. The data for identification of fracture sets (Figure 3-1) stems from various sources e.g. outcrop mapping (scanline or area mapping), borehole logging and lineament analysis.

Sampling of fracture orientation is always more or less biased by the geometry and scale of the sampling domain; e.g. fractures that are subparallel to an outcrop are less likely to be sampled than fractures that make a high angle to the outcrop. There are methods to correct-, or take into account, biased orientation data, some of which are treated in the next section.

Regardless of whether bias correction has been applied or not, there are basically only two methods to present orientation data graphically: Rose diagrams are used when the information is essentially 2-dimensional e.g. lineament maps, whereas stereographic projections of poles to fracture planes are preferred when there is sufficient information also regarding the dip of the fractures.



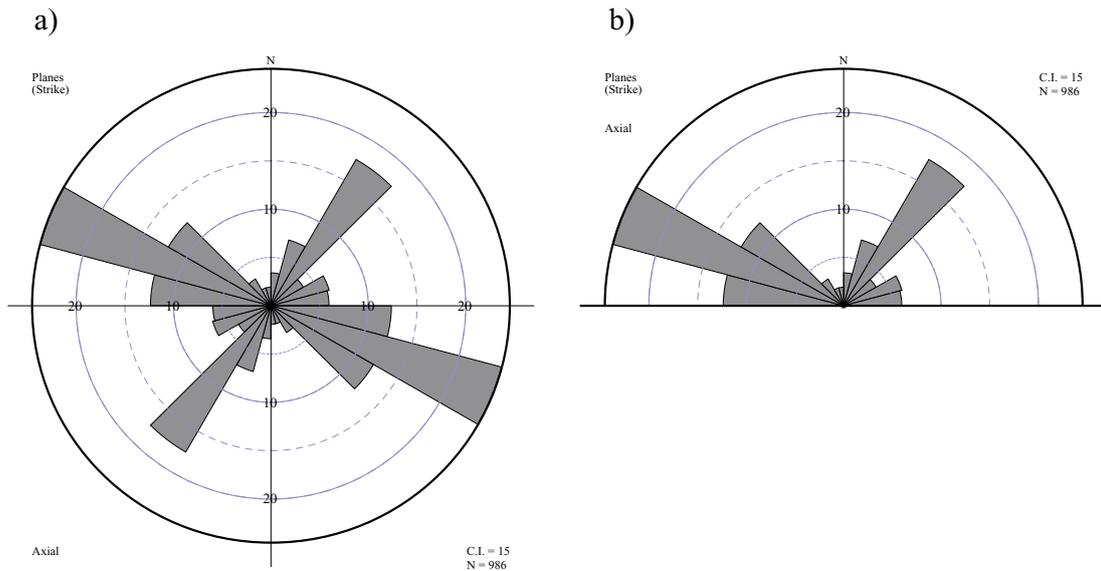
**Figure 3-1.** Data sources used for the estimation of fracture orientations a) outcrop mapping, b) lineament analysis and c) borehole logging.

Figure 3-2 shows examples of rose diagrams. Unless the rose diagram is intended to display vectors (Figure 3-2a), the bars are symmetric about the principal axes. It is therefore common practice to represent lineament trends in half-roses (Figure 3-2b). In either case, the following information should be clear in the graph, in the figure caption or elsewhere in the written documentation:

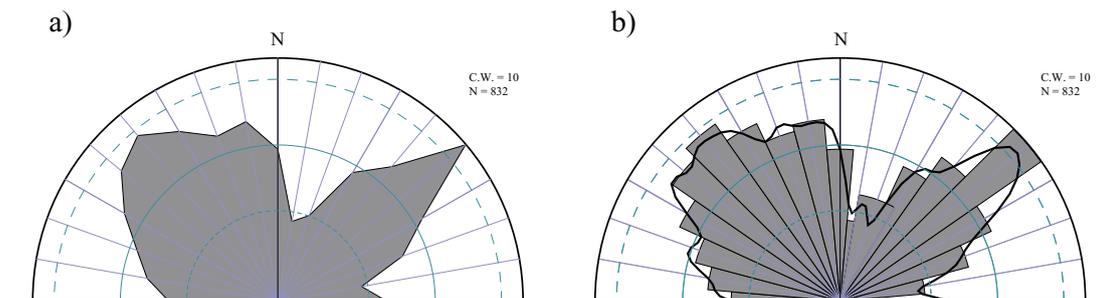
- The number of observations.
- The width of bars (class width).
- A scale (usually expressed as %).

An alternative representation is achieved by connecting the centre points of the bars, such as in Figure 3-3a, or to apply a filtering technique, such as in Figure 3-3b, to smooth and, in some circumstances, to highlight intricate details. In the latter case, we recommend to display also the histograms since smoothing can significantly distort the original data.

For visualising and analysing fracture data, the most commonly used tool in structural geology is the stereonet, an equal-area projection (sometimes referred to as the Schmidt net). The orientation data from the individual fractures (strike/dip) can either be plotted as great circles or as poles (plane normals). For large amounts of data, and for fractures in general, it is more appropriate to plot the fracture poles, since, otherwise, it is difficult to visually identify fracture sets (see 3.1.1).

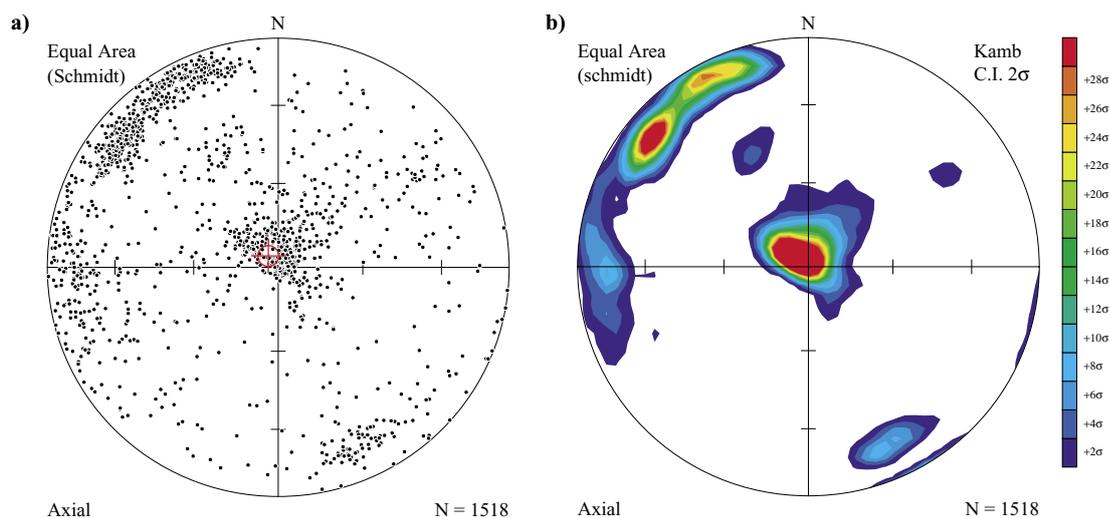


**Figure 3-2.** Example of rose diagram. A full rose a) is appropriate for (horizontal) vector data whereas a half rose is appropriate for axial data such as lineament orientations.



**Figure 3-3.** Example of alternative representations of 2D orientation data. In a) centre points are connected. In b) the data has been subjected to a spline technique to smooth the jagged appearance of the histograms, shown in the background. Both diagrams were scaled to the largest bar, using 10° bar widths.

Different contouring methods /Starkey, 1977; Lewis and Gray, 1985; Robin and Jowett, 1986; Adam, 1989; Vollmer, 1993; 1995/ can facilitate the identification of different fracture sets by visual inspection, and, at the same time, provide a basis for evaluating whether a pole concentration is statistically significant (Figure 3-4b). Of the standard methods available for contouring which are found in the literature, those which are based on the early work of /Kamb, 1959/ are preferred e.g. /Robin and Jowett, 1986/ or /Vollmer, 1993; 1995/. This is because these use the spatial orientation of the plane normals rather than the projected position of poles in the stereogram as a basis for data processing. Additionally, Kamb's use of a normalisation by an expected value /see Kamb, 1959, for details/ makes it possible to compare different stereograms on a statistically correct basis.



**Figure 3-4.** Fracture data from Forsmark site investigation, borehole KFM01A. (a) Equal area/ lower hemisphere stereogram of poles to fracture planes. The red symbol represents the mean direction of the borehole. (b) The same stereogram as (a), contoured according to the /Kamb, 1959/ method (contour interval  $2\sigma$ , significance level  $3\sigma$ ).

There are numerous softwares that aid in the plotting of poles and the computation of contours. Regardless of which software is used, the following should be made clear in the graph, in the figure caption or elsewhere in the written documentation:

- whether the data represents planes, poles or dip directions,
- the choice of contouring method, contouring intervals and significance level,
- the chosen projection, and
- the number of observations.

Additionally, it is good practice for borehole data to also plot the direction of the borehole. Unless otherwise motivated, we recommend the use of the poles to fracture planes and contouring according to the Kamb method. We further recommend that the plot of poles is separated from the graph of contours, as shown on Figure 3-4.

### 3.1.1 Correction for orientation bias

Fracture trace lengths are usually recorded on one-dimensional (e.g. scanline, borehole) or two-dimensional (e.g. outcrop, road cut, tunnel face) sampling domains. There are essentially two categories of bias affecting fracture arrays. One is governed by the size of the sampling volume and affects estimation of trace lengths. Another is governed by the orientation of the sampling domain in relation to the sampled structures and affects orientation analyses of the various fracture sets. This section discusses the latter.

On an infinite, perfectly flat sampling surface, the probability of intersecting an infinitely large fracture is directly proportional to the angle between the sampling surface and the fracture. As a consequence, outcrop mapping will generally yield fracture arrays depleted in subhorizontal fractures. Similarly, mapping of oriented, vertical boreholes will generally yield fracture arrays depleted in subvertical fractures. This was recognised early

by /Terzaghi, 1965/ who devised a simple correction procedure based on a trigonometric correction factor and assuming that the fracture spacings of the different sets are fairly equal. The density of each set,  $N_{obs}$ , is then weighted using the solid angle,  $\theta^i$ , between the mean pole of the fracture set and the orientation of the borehole, i.e.

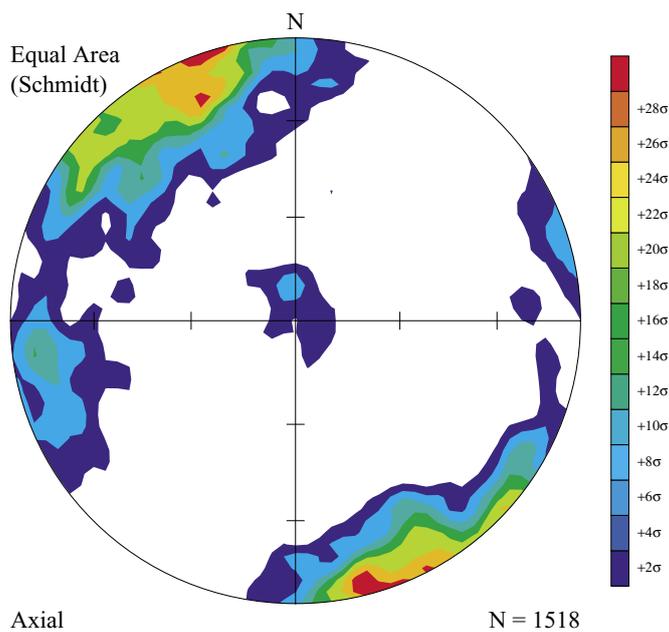
$$N_{adj}^i = N_{obs}^i \frac{1}{\cos(\theta^i)} \quad \text{Equation 1}$$

where the superscript  $i$  denotes the  $i$ :th fracture set  $N_{obs}$  and  $N_{adj}$  the observed and adjusted number of fractures, respectively.

In analogue, a similar expression might be formulated for outcrops:

$$N_{adj}^i = N_{obs}^i \frac{1}{\sin(\theta^i)} \quad \text{Equation 2}$$

It is also possible to weight each individual observation and take these weights into account when contouring the data. In either case, it should be noted though, that the weighting factor approaches infinity as the solid angle approaches  $90^\circ$ . To avoid artefacts it is common practice to limit the correction factor to some value, e.g. 5 or 10. For illustrative purposes, we present the result of such correction procedure in Figure 3-5. A comparison with the original data (Figure 3-4) shows that, as expected, the subvertical sets are enhanced whereas the subhorizontal set is relatively less dominant.



**Figure 3-5.** Fracture data from Forsmark site investigation, borehole KFM01A (Figure 3-4) corrected for orientation bias.

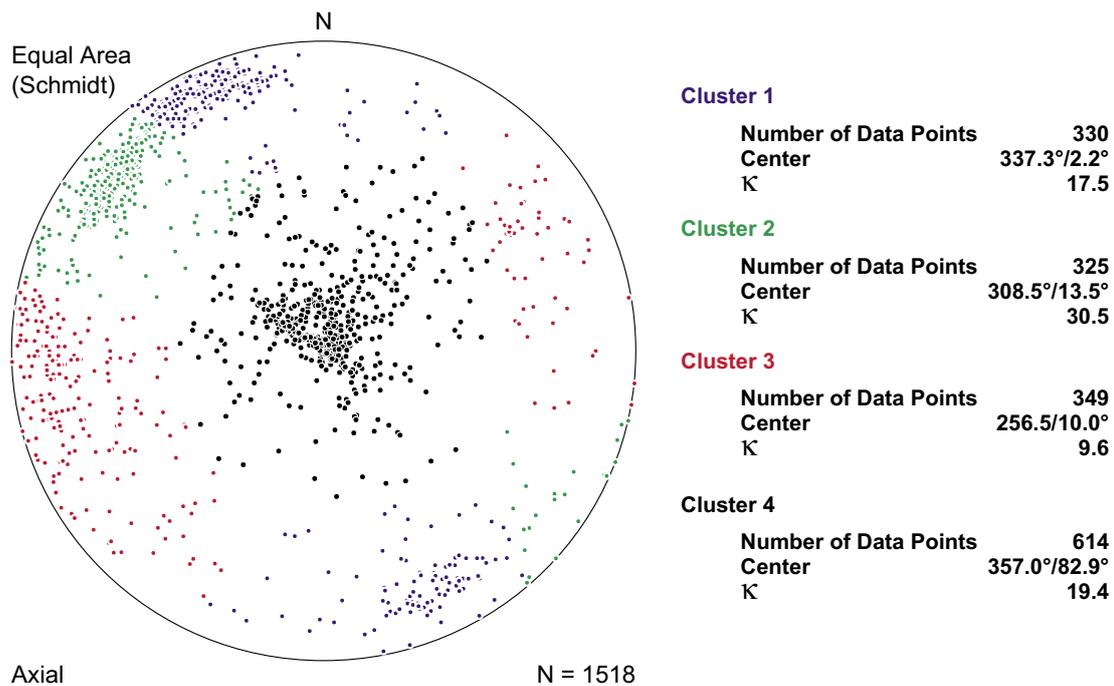
One problem is that outcrops and fractures have finite sizes and that any correction factor applied will be very much influenced by initial assumptions. The /Terzaghi, 1965/ correction, for instance, assumes that all fracture sets have the same spacing. In practice, though, the problems with orientation bias become less dominant if there is access to both boreholes and outcrop maps within the same rock volume, which is the case in SKB site investigations, assuming, of course, that the samples represent the same fracture population. Outcrop orientation data can be combined with borehole data to produce realistic estimates of the fracture orientations using spacings from both outcrops and boreholes. Though this is usually sufficient to define the *orientation* characteristics of the fracture sets, it is insufficient for the estimation of their *densities*. This because we assume that there is little difference in the size distributions between the vertical and horizontal fracture sets, which is not necessarily the case. However, as will be shown later, modern algorithms, such as those described in e.g. /Dershowitz et al, 1995/ and subsequent publications, enables iterative searches for the moments of the size distributions that best can reproduce the observed spacings in boreholes and trace lengths on outcrops. In summary, therefore, we do not anticipate any need for classical Terzaghi correction procedures on orientation data provided, of course, that the densities of the different sets are addressed by other means (see 3.2).

### 3.1.2 Identification of fracture sets

The example in Figure 3-4 shows that it is relatively easy to visually identify pole clusters on a contoured stereogram, given, of course, that the orientations are not random and that the number of observations is sufficient. However, the sensitivity of the contouring can be influenced by the choice of contour interval, so that the number of pole clusters can vary, which introduces a certain amount of subjectivity. It is not possible to give any general guidelines about how the algorithm should carry out the contouring, since the result of the contouring is dependent on the number of observations, the distribution of the data, the sensitivity, etc. Hence, it is very important that the contouring method and the parameters used are stated when describing stereonet or interpretations based on them, and that the same contouring method is used throughout in one and the same model.

A large number of methods have been published which, in a more objective way, help the modeller to identify fracture sets on a stereonet /e.g. Shanley and Mahtab, 1976; Miller, 1983a; Mahtab and Yegulalp, 1984; Schaeben, 1984; Kohlbeck and Scheidegger, 1985; Huang and Charlesworth, 1989; Vollmer, 1993; 1995/. It is difficult to recommend a particular one of these methods, since, in our opinion, they all have both advantages and disadvantages. Taking into account the uncertainties in the fracture orientation measurements, due to measurement errors, inadequate exposure, etc, the differences between the methods are thought to have relatively little significance for the present area of application. However, should the use of search algorithms be regarded necessary, it is important that the same clustering method be used throughout the modelling work.

Many methods for set identification /e.g. Pecher, 1989/ are based on the user or the code starting with an initial evaluation of the number of pole clusters visually identifiable in the contoured stereogram and the approximate orientation of the mean of each cluster. The algorithm then attempts to decide, by systematically working through all the data, whether a particular pole with a certain statistical significance can be regarded as belonging to one or other of the pole clusters (sets). An example of the result of such a search algorithm is shown in Figure 3-6.



**Figure 3-6.** Visual inspection of a contoured diagram (Figure 3-4b) suggests the presence of 4 pole clusters, one with a subvertical mean pole and three with subhorizontal mean poles, oriented approximately W, NW and NNW. Using these visual estimates, the algorithm of /Pecher, 1989/ assigns each data point to one of the three sets, and calculates the mean orientation and the maximum eigenvalue,  $E$ , of each set. Here we used Fisher  $\kappa$  to describe the dispersion.

This identification algorithm, however, is only apparently objective. As the example in Figure 3-6 shows, the algorithm assigns all data points uniquely to one of the pole clusters (sets). This is, of course, not correct, since the pole clusters (sets) are overlapping by definition. In other words, the algorithm can contribute to assessing to which pole cluster (set) a data point presumably belongs and to, recognising that the assignment to a cluster is always associated with some uncertainty. Sometimes, it is possible to obtain clarification from other geological information, such as mineral coatings, but such cases are generally fortunate exceptions.

The advantage of the use of a clustering algorithm is that the mean orientation of the fracture sets are always calculated in same way and that this process can be carried out consistently within a study area with many sampling points (outcrops, boreholes, etc). Also, a measure of the dispersion of each set is obtained.

### 3.1.3 Orientation distribution and its moments

The identification algorithm for fracture sets, discussed above, is mainly used to separate out the data belonging to each set from a fracture matrix with several sets. The next step in the process is to analyse each set separately in order to determine a representative distribution function in terms of a mean orientation (expected values, mean) and dispersion about the mean (e.g. variance).

There are a number of distribution functions which can be used to describe fracture orientations. The most common ones are the Dimroth-Watson /Watson, 1966/, /Bingham, 1964/ and /Fisher, 1953; Fisher et al, 1987/ functions.

The calculation of the eigenvector and eigenvalue of the normal axes gives directly the mean orientation and a measure of the dispersion. It is also possible to use Fisher statistics to calculate the mean vector and the dispersion,  $\kappa$  (kappa) for fracture poles, on condition that the fracture set is first rotated so that the mean fracture plane of the set is horizontal. This is due to the right-hand-rule convention and the standard of lower hemisphere projection. If this is not done, the mean orientation computed with the Fisher method will be incorrect, especially for steep fracture sets /see, e.g. Davis, 1986/.

Using different statistical tests /e.g. Hext, 1963; tex. Giné, 1975; Woodcock and Naylor, 1983/, one can test which of the distributions best fit the observed data. In practice, however, it is rarely necessary, since the measurement errors, the natural variations in orientation and the complexity of geological formations, etc makes it virtually impossible to choose one distribution in favour of another. On the other hand, it is important that the distribution is clearly stated in the description, with the mean orientation as well as the dispersion, independent of whether it was calculated with the aid of the eigenvalue method, Fisher, Bingham or Dimroth-Watson statistics. The aim is thus to express mathematically the orientation of the different fracture sets in terms of distribution functions with their moments, so that these can be used as parameters in a DFN model.

The following should be made clear in the documentation:

- Which distribution has been used and rationale for the choice.
- The moments of the distribution.
- The method by which the moments were derived.

Unless otherwise motivated, we suggest the use of the Fisher distribution for the description of orientations, despite being formulated for vectors, because it is common in the geologic literature, easy to compute, easy to simulate and fairly intuitive. In the following, we briefly describe the Fischer distribution and recommend modellers to adopt the nomenclature and notations used here to ease communication within the SKB modelling environment.

The Fisher distribution of vectors which are symmetrically distributed about the vector resultant can be expressed as:

$$f(\theta) = \frac{\kappa \cdot \sin \theta \cdot e^{\kappa \cdot \cos \theta}}{e^{\kappa} - e^{-\kappa}} \quad \text{Equation 3}$$

where  $\theta$  is the angle of divergence from the resultant vector, in degrees, and  $\kappa$  is the dispersion. An estimate of the dispersion factor,  $\kappa$ , is obtained from the relation /Fisher, 1953/:

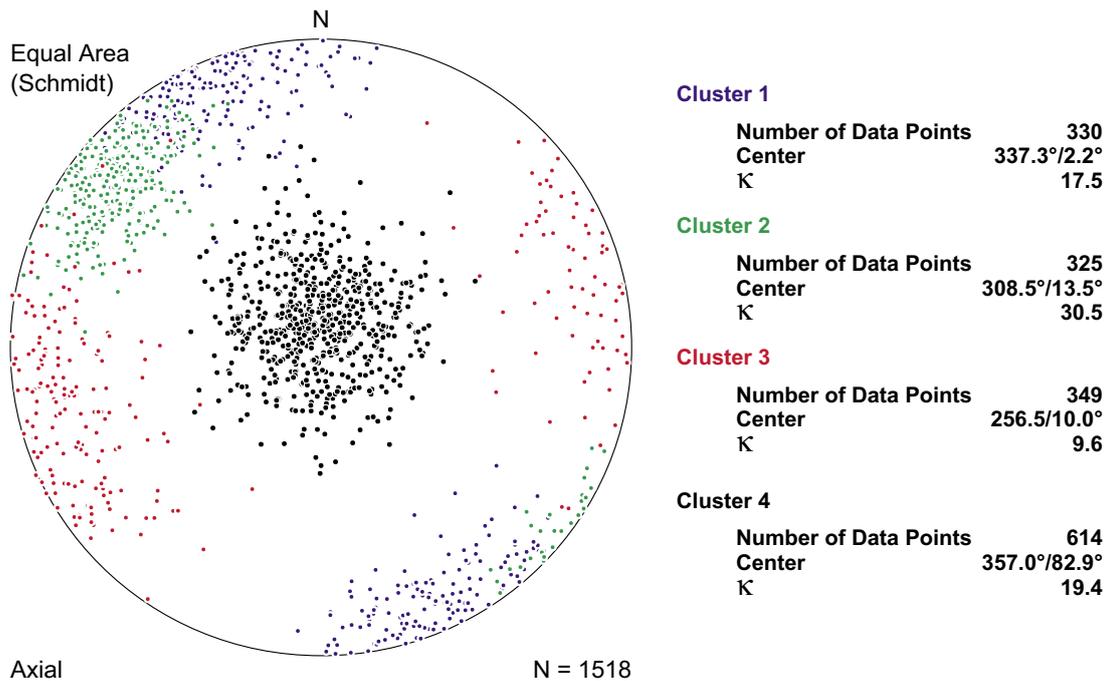
$$\frac{e^{\kappa} + e^{-\kappa}}{e^{\kappa} - e^{-\kappa}} - \frac{1}{\kappa} = \frac{|r_n|}{N} \quad \text{Equation 4}$$

where  $r_n$  is the resultant vector and  $N$  the number of vectors.

/Fisher, 1953/ showed that, for  $N > 30$ , the estimation of  $\kappa$  can be simplified to:

$$\kappa \approx \frac{N - 1}{N - |r_n|} \quad \text{Equation 5}$$

An example of simulation of fracture poles, using these statistics, is shown in Figure 3-7. A comparison with Figure 3-6 shows that the overall pattern of fracture poles are very similar, which is of course expected, but also that the clusters overlap which, again, illustrates the difficulties of denoting a particular fracture to a specific set with no other qualifiers than orientation.



**Figure 3-7.** Simulated fracture pole orientations based on the data shown in Figure 3-6.

### 3.1.4 Shape of the pole cluster

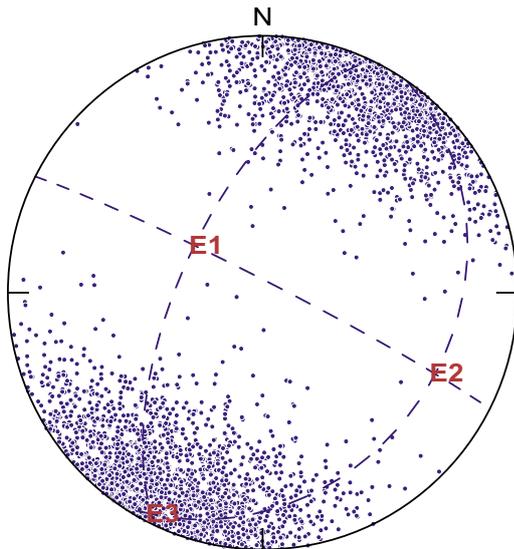
In crystalline bedrock fracture poles are typically symmetrically distributed about the mean pole, thus forming a roughly circular pole cluster on a stereogram (for example, cluster 4 on Figure 3-7). However, situations do occur, mainly in sedimentary rocks but sometimes also in igneous rocks, in which the fracture poles are distributed asymmetrically (in an ellipse rather than a circle) about the mean pole.

There are a number of methods of investigating and quantifying the degree of asymmetry, of which those of /Vollmer, 1989/ and /Woodcock, 1977/ (Figure 3-9) are most commonly used. Both, equally relevant, methods are based on different aspects of eigenvectors and eigenvalues (Figure 3-8) and give similar information when applied to fracture statistics.

If the deviation from a regular cluster shape can be shown to be significant using either of the described methods, the pole cluster could be defined using the bi-modal variants of the Fisher /Fisher et al, 1987/ or the Bingham distribution /Bingham, 1964/. The estimation of the dispersion,  $\kappa$ , then becomes much more complicated since  $\kappa$  varies in different directions about the mean pole. Instead of a single measure of the angular variation from the mean pole (cf Equation 1), two dispersions,  $\kappa_1$  and  $\kappa_2$ , are given, along the two principal axes of the elliptical pole cluster.

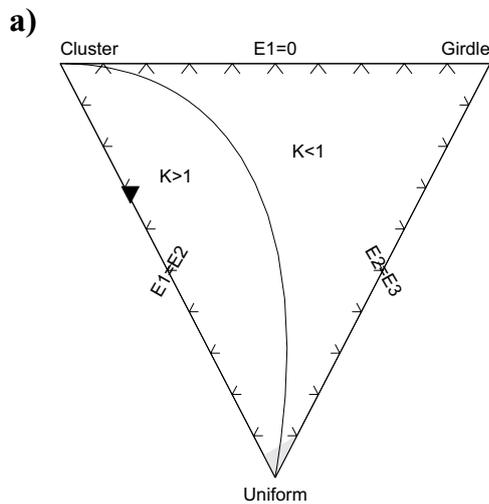
Analysis of cluster shapes should provide the following information, made clear in the graph, in the figure caption or elsewhere in the written documentation:

- The method used for the shape analysis
- The parameters derived from the method (fabric index, ratios)
- The three eigenvectors with their orientations and length, if applicable.
- $\kappa_1$  and  $\kappa_2$  of the bivariate Fisher distribution, if applicable.



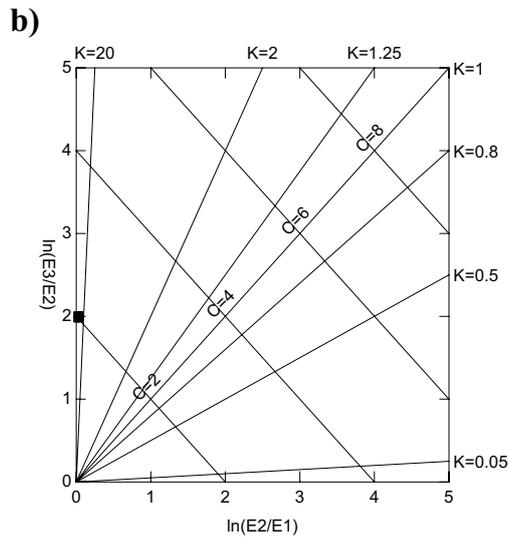
Eigenvectors			
Vector	Trend	Plunge	Eigenvalue
E1	304	63	330.05
E2	115	26	339.59
E3	207	4	2486.36

**Figure 3-8.** Example of the calculation of the centre of gravity of a pole cluster using eigenvectors. In addition to eigenvectors, which indicate direction, the inherent properties of eigenvalues can be used to analyse the cluster's shape (Figure 3-9). The centre of gravity is marked by the eigenvector with the largest eigenvalue (E3).



Vollmer's Fabric Indexes		
Index	Value	St.Dev.
Cluster	0.680	0.000
Girdle	0.006	0.000
Uniform	0.314	0.000
Cylindricity	0.686	0.000

Non-uniform distribution.  
Tested at the 95% confidence level.



Woodcock Ratios		
Ratio	Value	St.Dev.
K ratio [ln(E3/E2)/ln(E2/E1)]	69.914	1.692
C ratio [ln(E3/E1)]	2.019	0.001

Critical C value = 0.16  
Moderately developed uniaxial cluster.  
Tested at the 95% confidence level.

**Figure 3-9.** Illustration of two statistical methods for analysing the shape of pole clusters: the method of /Vollmer, 1989/, on the left, and the method of /Woodcock, 1977/, on the right. The plots show that the analysed cluster is symmetrical around the mean pole and that the fractures show relatively little dispersion about the mean orientation. In both methods, E1, E2 and E3 are the same as in Figure 3-8.

### 3.1.5 Comparison of stereograms

Even though mapped outcrops are located within the same rock type, the fracture array may differ so much that it might be judged necessary to subdivide the volume into smaller, more homogeneous domains or to describe the variation in terms of a transition, e.g. rotation, within the domain. However, since the creation of additional subdomains is afflicted with considerable modelling efforts, it is advisable, as a first step, to lump all outcrop observations within a rock domain defined in the geological model, and express the differences in orientation within that domain as a larger spread in orientation data unless, of course, the need of subdomains is strongly supported.

Changes in the orientation matrix of a fracture array, whether these are due to changes in relative densities between different fracture sets, due to changes in orientation of the sets or both, are commonly addressed by studies of stereograms. Comparison of stereograms is traditionally performed by visual inspection. Though the human brain is extremely capable of distinguishing similarities and dissimilarities between objects by visual comparison, in certain situations, however, visual comparison might be difficult. For instance, the orientations in two outcrops may be fairly identical whereas the densities of the clusters can differ significantly which is harder to detect by visual inspection. Differences in relative densities between the sets are important to address, especially if there is a significant difference in fracture sizes between the various sets, since this can be indicative of anisotropy within the studied volume. In this section we present a method slightly altered /Munier, 1995/ from /Miller, 1983a/ but recognise that many equally or more potent methods can be applied /e.g. Mahtab and Yegulalp, 1984; Kulatilake et al, 1990; Kulatilake et al, 1996/.

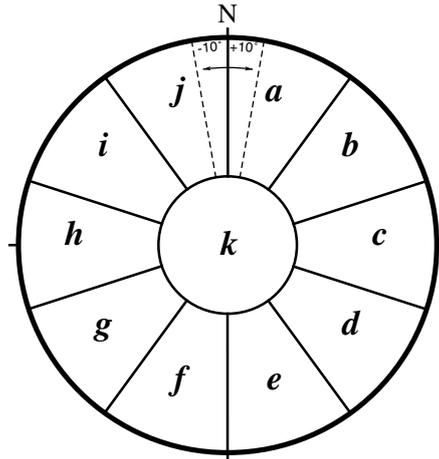
A commonly encountered problem of elementary statistics is that of comparing a distribution of sample observations to some pre-defined standard distribution. A test procedure known as  $\chi^2$  *goodness of fit* can be applied to the sample to evaluate whether the observed and the expected theoretical distributions are reasonably equal. The test statistic is calculated according to:

$$\chi^2 = \sum_{i=j}^k \frac{(obs_i - exp_i)^2}{exp_i} \quad \text{Equation 6}$$

where  $obs_i$  is the number of observations in class  $i$ ,  $exp_i$  the expected number in class  $i$  and  $k$  the number of classes. The degree of significance can be calculated from  $k$  and  $\chi^2$ .

In analogue, orientations of fractures can be compared to the expected, “theoretical” orientations and tested by  $\chi^2$  /Miller, 1983b/ for example, the orientations expected if two stereograms were equal. The problem reduces to a division of the stereograms into  $k$  patches of equal area (Figure 3-10), counting the poles within each patch and performing the test according to Equation 6. Care must be taken, however, because valid  $\chi^2$  tests cannot be performed unless the number of observations in each patch exceed 5 (Lancaster’s criterion). It is therefore necessary to adjust the number of patches or the patch network until this criterion is met. This can be performed in three ways:

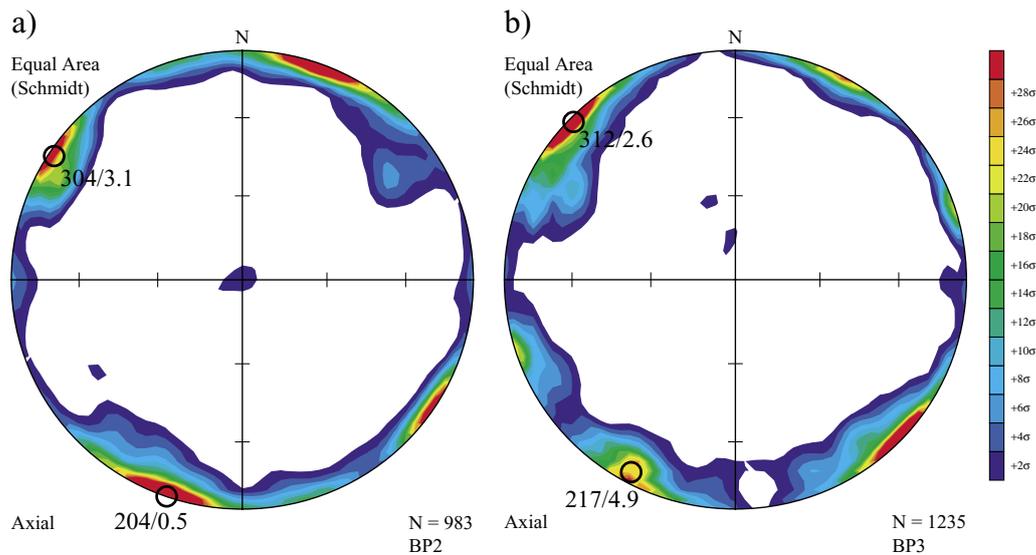
- The number of patches is decreased (their area increased) until all contain 5 or more poles. This will also result in a less discriminant, but weaker, test.
- The patch network is rotated until the criterion is eventually met.
- Patches with too few observations can be lumped together. This will also result in a less discriminant, but weaker, test.



**Figure 3-10.** The figure show the patch configuration used for stereonet comparison /from Munier, 1995/.

We have adopted a test that rotates the patches  $36^\circ$  (the width of one patch) about the vertical axis in steps of  $1^\circ$  to produce a mesh of  $\chi^2$  results from which the minimum (best) value is chosen. This to avoid that pole concentrations, are transected by a patch boundary by mere chance. To fulfil Lancaster's criterion, we lump all patches with less than 5 observations.

Figure 3-10 shows two contoured stereograms of poles to fracture planes mapped at two outcrops (BP2 and BP3 respectively) within the same rock domain at Forsmark (for simplicity we only display the contoured graphs). By visual inspection, the stereograms are strikingly similar. However, upon closer inspection, it is evident that the NW striking set is more intense in BP3 than in BP2. Further, BP2 displays a weak NNW set that appear to be absent from BP3. The question is, are these differences so large that the outcrops will fail to pass the test of similarity? A contingency table of frequencies and computed  $\chi^2$  values Table 3-1 shows that the stereonets can indeed not be regarded equal at the 95% confidence level, mainly due to differences in densities, as the sum of  $\chi^2$  by far exceeds the threshold value of 18.4 (95% confidence limit, 10 degrees of freedom). However, despite the visual similarities, there are significant differences in direction as well, upon closer inspection. In the example used, the mean pole direction of the most dominant sets in BP2 are  $304/3$  and  $204/0.5$  whereas the corresponding mean pole direction on BP3 are  $312/3$  and  $217/4.9$ . This corresponds to dihedral angles of  $8^\circ$  and  $14^\circ$  respectively.



**Figure 3-11.** Poles to fracture planes mapped at two outcrops in Forsmark contoured according to the /Kamb, 1959/ method (contour interval  $2\sigma$ , significance level  $3\sigma$ ). a) Drillsite 2 (BP2), b) drillsite 3 (BP3).

**Table 3-1.** Contingency table of  $\chi^2$  frequencies and calculated  $\chi^2$  values used to compare two stereonet (Figure 3-11) for equality.

Patch	BP2_obs	BP2_exp	BP3_obs	BP3_exp	BP2_ $\chi^2$	BP3_ $\chi^2$
a	160	152	184	192	0.421	0.333
b	116	182	295	229	23.930	19.020
c	72	46	32	58	14.690	11.650
d	79	51	35	63	15.370	12.440
e	43	29	22	36	6.758	5.444
f	18	23	34	29	1.086	0.862
g	38	73	127	92	16.780	13.310
h	146	107	96	135	14.210	11.260
i	232	213	248	267	1.694	1.352
j	43	80	137	100	17.110	13.690
k	33	25	24	32	2.560	2.000
<b>Sums</b>	<b>980</b>	<b>981</b>	<b>1234</b>	<b>1233</b>	<b>114.600</b>	<b>91.385</b>

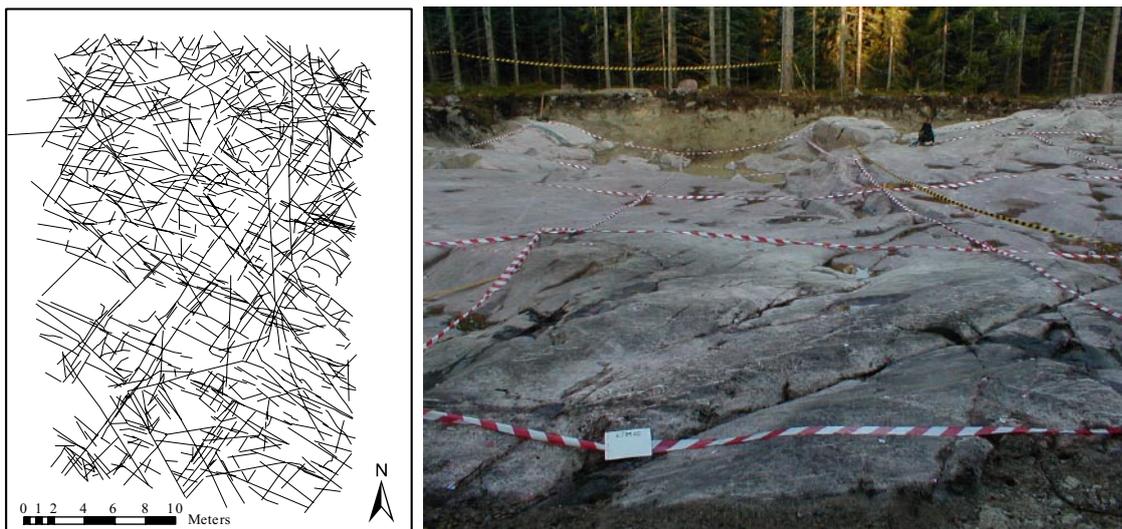
The parametric test described in this section is a powerful tool to objectively compare orientations on outcrops. However, the test is rather discriminant and might be perceived as counterintuitive. Yet, regardless of the outcome of the test, the main point is to present a *quantitative* value of similarity, or lack thereof, so that the issue of spatial continuity with respect to orientations can be properly addressed using supportive *qualitative* arguments.

## 3.2 Fracture size distribution

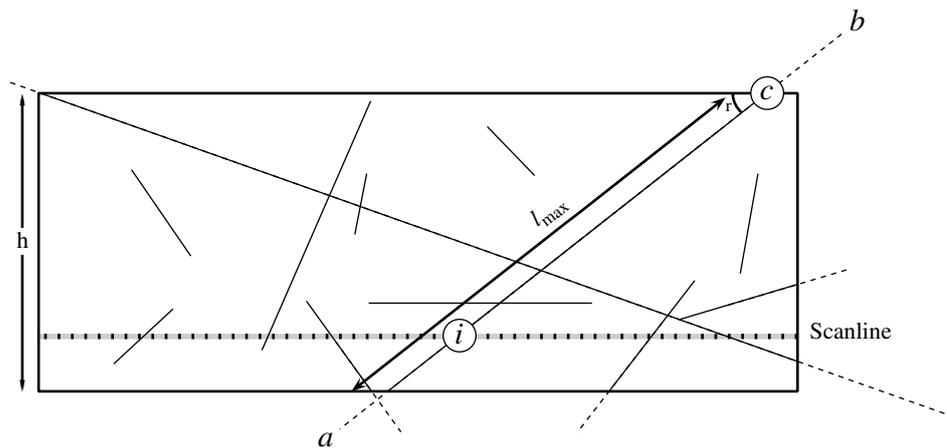
Although outcrop sampling of fractures are widely recognised to be length-biased, they are often the best source of data for inferring statistical parameters of fracture sets. Fracture trace lengths are mapped on outcrops, with, in the usual case, a lower censoring by policy of 1 m for area mapping and 0.1 m for scanline mapping (see SKB MD 132.003 for details). Lineament maps provide additional information on larger scale deformation zones. Here, we have adopted the terminology of /Jensen et al , 2000/ in which censoring denotes unmeasured information that is known, whereas truncation denotes unmeasured information that is unknown.

The size of fractures can rarely, if ever, be measured directly. Instead we measure the intersection of the fracture and the sampling domain, i.e. outcrop, road-cut or tunnel wall. In this section we discuss statistics of fracture trace lengths and means of obtaining the fracture sizes from these statistics.

It should be noted that, besides censoring due to sampling policy or inability to detect short traces, there are two sources of bias entirely due to the size of the sampling domain. The first bias is because a scanline or a wall will tend to intersect preferentially the larger fracture planes (see Figure 3-13 for schematic illustration). In fact, the probability of intersecting a fracture is directly proportional to its size. The second bias is that large fracture planes may extend beyond the rock exposure. Measurements are therefore truncated at some value dependent upon the size of the sampling domain. For clarity, the cut-off of the upper portion of the fracture trace is denoted “truncating” whereas cut-off of the lower portion is denoted “censoring”. Censoring can either be due to sampling policy (e.g. choice of sampling scale) or a physical inability to detect traces shorter than a specific threshold value. Techniques for correcting sampling bias on outcrop data have been widely reported and fall into two main categories: Those that assume an underlying parent distribution and those that do not. The former approach has been addressed by e.g. /Priest and Hudson, 1981/ and /La Pointe and Hudson, 1985/ whereas the latter approach was addressed by e.g. /Kulatilake and Wu, 1984/. The issue is further discussed below.



*Figure 3-12. Mapped fracture traces a) at BP2, an exhumed outcrop b) at the Forsmark site.*



**Figure 3-13.** Schematic cartoon of a tentative tunnel face. The intersection point (*i*) is randomly positioned along *a–b*. The maximum observable trace span from the points *a* to *c* and is denoted  $l_{max}$ . The length of  $l_{max}$  depend on the angle ( $r$ ), the intersection angle (rake) between the fracture plane and the sampling domain /from Munier, 1995/.

### 3.2.1 Description of trace lengths

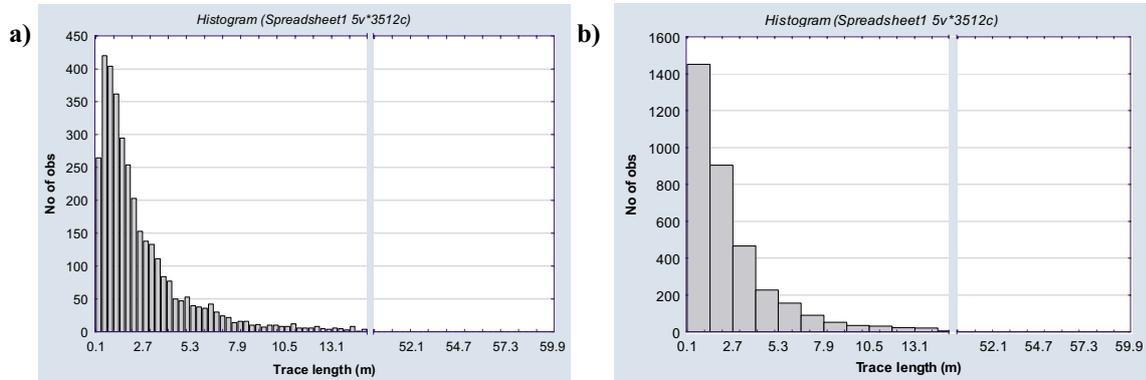
The most common distributions of fracture trace length reported in the literature are the lognormal- /e.g. Epstein, 1947; Baecher and Lanney, 1978; Priest and Hudson, 1981; Rouleau and Gale, 1985/, powerlaw- /e.g. Fujiwara et al, 1977; King, 1983; Turcotte, 1986a; b; 1990; An and Sammis, 1994; Dershowitz et al, 1995/ and exponential distributions /Call et al, 1976; Cruden and Ranalli, 1977; Priest and Hudson, 1981; Cowie et al, 1995; Kagan, 1997; Renshaw and Park, 1997/.

Though the powerlaw distribution has gained increasing attention, and is indeed both conceptually and computationally attractive, there are very strong empirical, experimental and theoretical arguments for each of these distributions as appropriate descriptors of fracture sizes. We therefore here advocate for the need of testing the fit of the distributions, and other distributions too if considered appropriate (e.g. gamma /Kagan, 1997/), prior to further analysis.

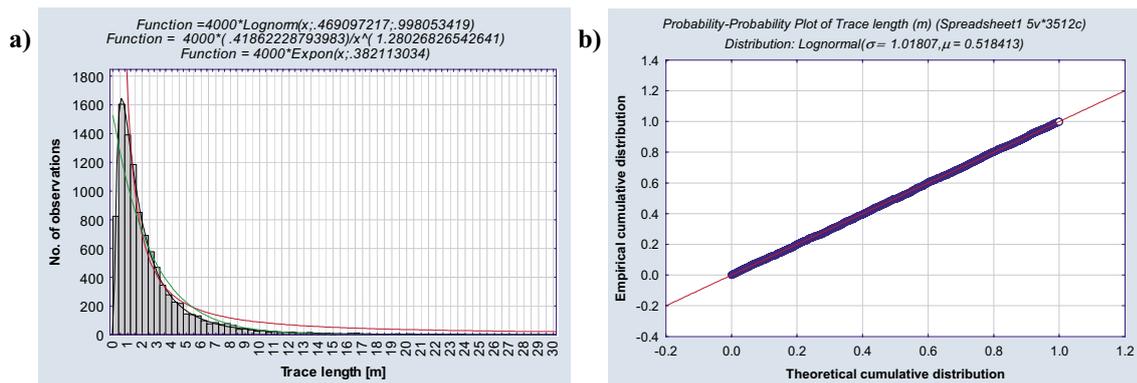
It is common practice to use histograms of the frequency distribution of the trace length in which the columns are drawn over the class intervals and the heights of the columns are proportional to the class frequencies. Histograms give a good overview of how the data is distributed and can, for instance, reveal the presence of several frequency maxima (e.g. bimodal distributions). Further, the shape of the histogram can give an indication of which type of distribution to test for the studied variable. The analytical power of the histogram is, however, fairly restricted and must be used with great care. For instance, the choice of class interval will, to a great extent, steer the appearance of the histograms. Unfortunately, it is not common practice to exploit the effect of altering the class widths; mostly, analyses contempt with standard class width, based on e.g. the range of the data, using built-in routines of the various softwares and wrong conclusions are often drawn regarding the nature of the distribution.

Figure 3-14 shows how the shape of the histogram changes with using different class width. Figure 3-14a rather clearly indicate a lognormal distribution by visual inspection whereas the default class width of the software, displayed in Figure 3-14b, indicates an exponential distribution using the same data. Unless a goodness of fit test is clearly articulated, we strongly recommend against the use of histograms for *choosing* distribution though we appreciate the strength of this graphical representation for visually displaying the fit.

When describing a trace length distribution, it is good practice to present the histogram, showing tested distributions, together with a probability plot (Figure 3-15b) of the chosen distribution and a table of the goodness-of-fit test performed Table 3-2.



**Figure 3-14.** Using the same data, histograms may appear quite differently depending on the choice of class width.



**Figure 3-15.** a) Histogram of trace lengths with fitted distributions. b) Probability plot of best fit distribution.

**Table 3-2.** Example of a table summarising test statistics for different distributions.

Distribution	$\chi^2 ; v = 11$	$P_{\alpha=0.05}$
Lognormal	10.48497	0.4874
Exponential	478.21602	0.0000
Powerlaw	1253.28746	0.0000

In order to promote a coherent nomenclature within SKB projects, we briefly present the distributions and their moments in the sections below with an appeal to comply with the terminology proposed.

## The Exponential Distribution

The variable  $x$  has an exponential (or negative exponential) distribution if it has a probability density function (swedish "t athetsfunktion") of the form /Johnson and Kotz, 1972/:

$$f(x) = \frac{\exp \frac{x_{\min} - x}{\sigma}}{\sigma}, \quad (x > x_{\min}; \sigma > 0) \quad \text{Equation 7}$$

The maximum likelihood estimators of  $x_{\min}$  and  $\sigma$  are respectively

$$x_{\min} = \min(x_1, x_2, \dots, x_n) \quad \text{Equation 8}$$

and

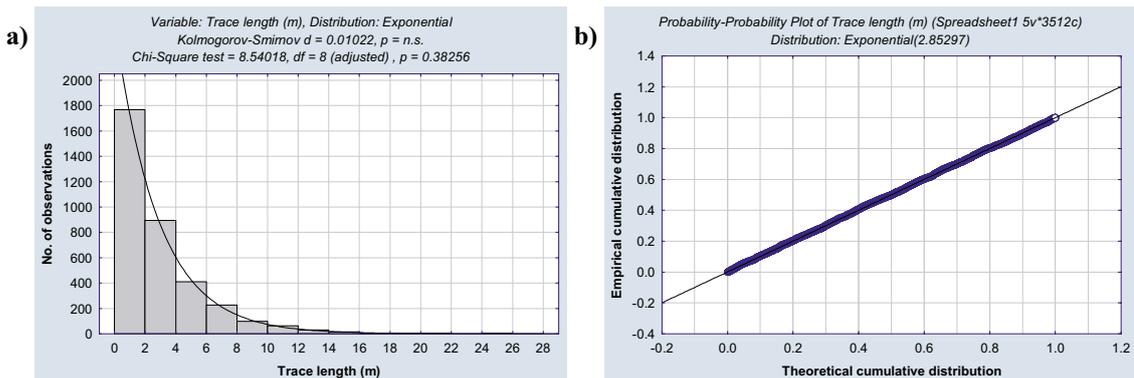
$$\sigma = \frac{\sum_{i=1}^n (x_i - x_{\min})}{n} \quad \text{Equation 9}$$

If the sample (fracture traces) is not censored, i.e. by policy or inability to detect values smaller than a certain threshold, then  $x_{\min} = 0$  which makes the distribution a one-parameter exponential distribution and Equation 7 may be reformulated in a more common form as:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} \quad \text{Equation 10}$$

with a mean and standard deviation which both equal  $\frac{1}{\lambda}$ .

In Figure 3-16 we illustrate how exponentially distributed trace lengths may be represented. Note that the histogram Figure 3-16a) displays goodness of fit tests. The moment (mean and standard deviation are equal) is displayed in the header of the probability plot Figure 3-16b).



**Figure 3-16.** Example of histogram (a) and probability plot (b) of a exponentially distributed trace lengths.

### The lognormal distribution

If there exists a number,  $x_{\min}$ , such that  $Z = \ln(x-x_{\min})$  is normally distributed, then the distribution of  $x$  is said to be lognormal. For this to be the case it is clearly necessary that  $x$  can take any value exceeding  $x_{\min}$ , but has zero probability of taking any value less than  $x_{\min}$ . For the case when  $x_{\min} \neq 0$  the distribution is “three-parameter lognormal”. If the sample (fracture traces) is not censored, i.e. by policy or inability to detect values smaller than a certain threshold, then  $x_{\min} = 0$  and the distribution is called “two-parameter lognormal” or simply lognormal. Commonly, fracture traces are censored by policy and it is therefore advisable to use the three parameter version formulation. The probability density function of a lognormal distribution has the general (three-parameter) form /Johnson and Kotz, 1972/:

$$f(x) = \frac{e^{-\frac{(\ln(x-x_{\min})-\mu)^2}{2\sigma^2}}}{\sigma(x-x_{\min})\sqrt{2\pi}}, \quad x > x_{\min} \quad \text{Equation 11}$$

The estimation of  $\mu$  and  $\sigma$  presents no difficulties beyond those associated to the normal distribution. The maximum likelihood estimator for  $\mu$  and  $\sigma$  are respectively:

$$\mu = \frac{1}{n} \sum_{i=1}^n \ln((x_i - x_{\min})_i) \quad \text{Equation 12}$$

and

$$\sigma = \left[ \frac{1}{n-1} \sum_{i=1}^n (\ln(x_i - x_{\min}) - \mu)^2 \right]^{1/2} \quad \text{Equation 13}$$

The mean, median, mode, and variance are expressed as:

$$\text{mean} = e^{\left(\mu + \frac{\sigma^2}{2}\right)}, \quad \text{Equation 14}$$

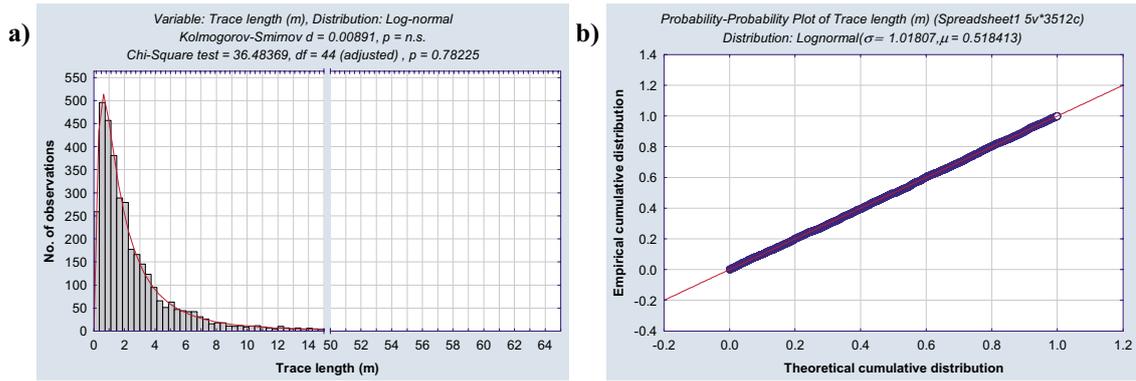
$$\text{median} = e^{\mu}, \quad \text{Equation 15}$$

$$\text{mode} = e^{(\mu - \sigma^2)}, \text{ and} \quad \text{Equation 16}$$

$$\text{variance} = e^{2\mu} e^{\sigma^2} (e^{\sigma^2} - 1), \text{ respectively.} \quad \text{Equation 17}$$

It is, unfortunately, not always made clear if the described parameters refer to the mean of  $x$ , or the mean of the logarithm of  $x$ ,  $\mu$ . We therefore recommend that any description of a lognormal variate is expressed in terms of the parameters,  $\mu$ ,  $\sigma$ , and  $x_{\min}$  (if  $\neq 0$ ), from which the mean and other moments always can be computed.

In Figure 3-17 we illustrate how lognormally distributed trace lengths may be represented. Note that the histogram (Figure 3-17a) displays goodness of fit tests. The moments are displayed in the header of the probability plot (Figure 3-17b).



**Figure 3-17.** Example of histogram (a) and probability plot (b) of lognormally distributed trace lengths.

### The Powerlaw distribution

The probability density function of a powerlaw distribution can be expressed as /Evans et al, 1993/:

$$f(x) = \frac{k \cdot x_{\min}^k}{x^{k+1}}, \quad x_{\min} \leq x < \infty \quad \text{Equation 18}$$

where  $x_{\min}$  is the location parameter (smallest value of  $x$ ), and  $k$  the shape parameter.

The mean and standard deviation of the powerlaw distribution can be expressed as:

$$\mu = \frac{k \cdot x_{\min}}{k-1}, \quad k > 1 \quad \text{Equation 19}$$

$$\sigma = \frac{x_{\min}}{k-1} \cdot \sqrt{\frac{k}{k-2}}, \quad k > 2 \quad \text{Equation 20}$$

The parameters of the distribution can be estimated using the maximum likelihood estimators:

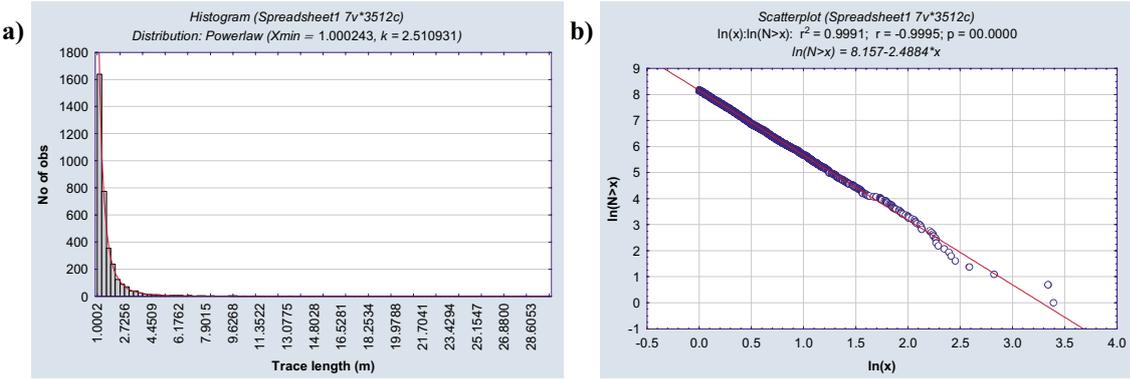
$$\frac{1}{\hat{k}} = \left( \frac{1}{N} \right) \sum_{i=1}^N \ln \left( \frac{x_i}{\hat{x}_{\min}} \right) \quad \text{Equation 21}$$

$$\hat{x}_{\min} = \min x_i \quad \text{Equation 22}$$

where  $N$  is the total number of data.

Since the advent of fractal geometry /Mandelbrot, 1982/ and its application to the earth sciences /e.g. Okubo and Aki, 1983; Scholz and Aviles, 1986; Turcotte, 1986a; Korvin, 1992/ it has been common practice, in geologic literature in particular, to denote the shape parameter (exponent) of the powerlaw distribution “D” implying the close relationship to the fractal dimension. This is misleading, because it is perfectly possible to have powerlaw distributed fracture sizes that do not form a fractal spatial pattern, and vice versa. Though the two concepts are related, they should under no circumstances be equated. We therefore strongly recommend against the use of “D” for the shape parameter, to avoid any such confusion within SKB projects.

In Figure 3-18 we illustrate how powerlaw distributed trace lengths may be represented. Figure 3-18b shows the complementary cumulative distribution (ccdf) of the trace lengths in log-log scale. Basically, the absolute value of the slope of the linear regression should equal the shape parameter,  $k$ , of the powerlaw distribution. However, since the powerlaw is heavily skewed, estimation by linear regression is less accurate, despite eventually high  $r^2$  values, than the estimator proposed in Equation 21; the linear fit does not take into consideration that almost all of the data are located on the first few point of the distribution and will be largely steered by the few, more “noisy” data in the tail of the distribution. We therefore call for caution when using the slope of the ccdf for estimation of  $k$  though we appreciate the strength of this graphical representation for visually displaying the fit.



**Figure 3-18.** Example of histogram (a) and probability plot (b) of powerlaw distributed trace lengths.

### 3.2.2 Computation of fracture sizes from trace lengths

Techniques for correcting sampling bias were early reported by, e.g. /Priest and Hudson, 1981/ and /Priest, 1993/ who devised formulae to identify the parent distribution from a biased sample. In /Priest and Hudson, 1981/, an exponential parent size distribution was assumed, in fact also the uniform and normal distributions, rarely reported for fractures in literature, which yielded a gamma distribution of the sampled trace length. Munier /Munier, 1995/ later applied the formulae to the lognormal distribution which yielded a trace length distribution that mimicked the exponential distribution and also to the powerlaw distribution which yielded a powerlaw trace length distribution with the same shape parameter. The latter has also been confirmed by others /e.g. LaPointe et al, 2000/. The drawback of these methods is that the deduced parent distribution can, in some circumstances, assume non-standard forms for which there are no standard analytical tools, and that might be conceived as counter-intuitive.

Another approach is based on the assumption that fractures can be represented by circular, perfectly flat, and infinitely thin discs with radii that are independent of both their position and orientation. With the additional assumption that fracture radii are distributed according to a Powerlaw distribution, /Marrett and Allmendinger, 1991/, /Berkowitz and Hadad, 1997/, /Piggot, 1997/and others /e.g. La Pointe, 2002/ have demonstrated that trace lengths on outcrop are also powerlaw distributed with a shape parameter  $k_{2D} = k_{3D} - 1$ . Additionally, it has be shown that a scanline intersecting traces will yield a sample of trace lengths that relates to the fracture radii as  $k_{1D} = k_{3D} - 2$ . In many, if not most, cases the assumptions required do, however, not prevail, especially the assumption of random positions in 3D

space and independence of size and position. In fact, assuming a truly fractal fragmentation mechanism requires a solid correlation between position and size and, arguably, also orientation.

Due to these uncertainties, we would recommend, unless otherwise motivated, another approach based on stochastic simulation. The basic idea is to simulate multiple realisations of a fracture network, and place tentative outcrops and boreholes within /e.g. Geier et al, 1988; Dershowitz et al, 1991a; Dershowitz et al, 1991b/. The artificial fracture traces and borehole intersections are then compared to the field observations. The parent distribution functions, and their moments, are systematically altered to cover a wide, yet realistic, range. The similarities, or lack thereof, between the mapped and simulated data are quantified using, for instance, Chi-Square or Kolmogoroff-Smirnov convergence criteria. Using such criteria, it is possible to find the size distribution, and its parameter(s), that best fit each fracture set of the array, simultaneously honouring statistics of mapped trace data and borehole intercepts. The iteration, which can be fairly tedious, can be made more efficient by the use of different search algorithms such as the “simulated annealing” /Kirkpatrick et al, 1983/ or “conjugate gradient” /Hackbusch, 1985/ methods.

Note, however, that trace length distributions cannot alone be used in such a procedure. Equally important is the number of fractures in the simulated system. In fact, both the amount and sizes of fractures must be simultaneously expressed in a measure called *fracture density* (or intensity) which is the subject of Section 3.2.4.

The process of determining the size distribution of the parent population of the different fracture sets should be articulated clearly. We recommend that such an analysis should contain at least the following information:

- The distributions that were tested.
- The range of moments that was tested for each distribution.
- If applicable, the search algorithm that was used for finding the best fit.
- The amount of realisations used for each test.
- The test statistic used.
- The chosen limit between deterministic and stochastic structures.
- The concepts/assumptions made if large- and small-scale features are modelled simultaneously (e.g. for borehole data, if the internal fracture array of a deformation zone is replaced by a single value representing the largescale feature).

### **3.2.3 Comparison of (mean) trace lengths**

The construction of a DFN initiates, as shown earlier, by the identification of fracture sets. This is a classification based on orientation and its importance cannot be sufficiently stressed. The fracture orientation combined with fracture size steer the degree of fracture anisotropy in the rock. Combined with yet other qualifiers, such as transmissivity, the anisotropy will more or less directly steer the flow and transport properties of the rock, and might also to some degree steer the repository layout.

During site investigations, the only means to obtain information of fracture sizes is by mapping their traces on outcrop. Naturally, a comparison of fracture sizes must be based on statistics on trace lengths or, alternatively, on the fracture radii once these have been determined (cf 3.2.2).

For small samples it might be convenient, to consider a non-parametric (also called distribution-free) test. The most commonly used test is the Friedman analysis of variance. This test should be used if one is not sure of the distribution of the variables used in the comparison. For example, the data may actually consist of rankings rather than precise measurements. However, it turns out that when the data set is large (e.g.  $n > 100$ ) it often makes little sense to use nonparametric statistics at all. The central limit theorem /Pólya, 1920/ states (in a nutshell) that when the samples become very large, the sample means will follow the normal distribution even if the respective variable is not normally distributed in the population, or is not measured very well. Thus, the parametric methods, which are usually much more sensitive (i.e. have more *statistical power*) are in most cases appropriate for large samples.

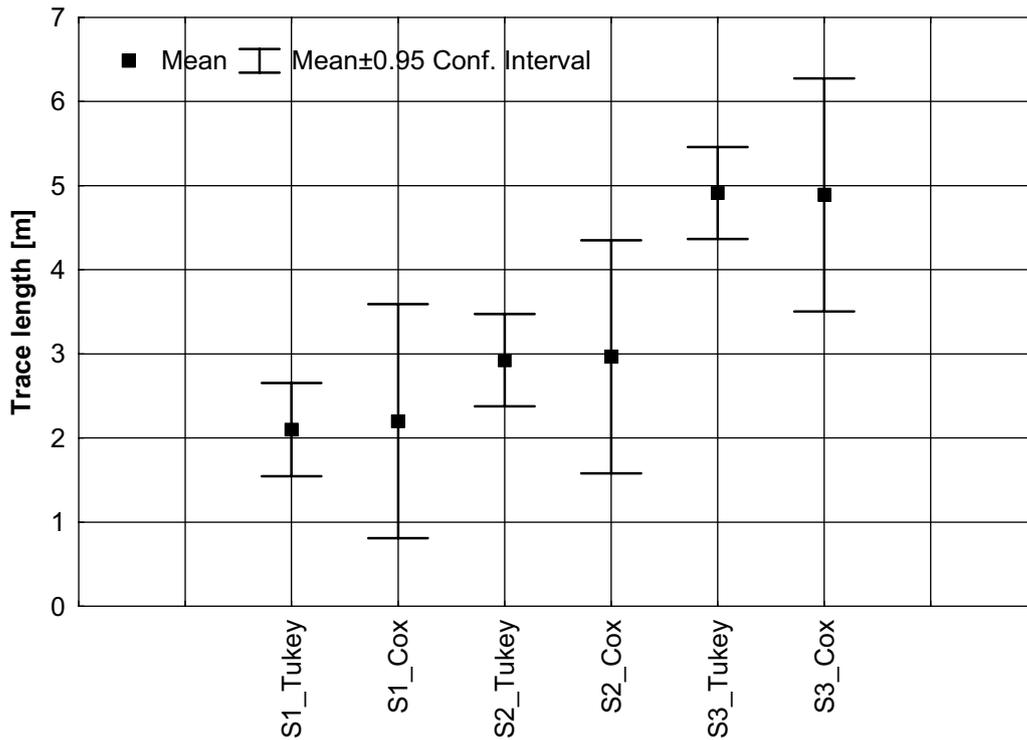
A way of comparing means makes use of confidence intervals which are graphically displayed together with the means in boxplots (Figure 3-19). The strength of this method is that any differences in means, under a certain confidence level, are immediately revealed; Significant differences in mean values are manifested as non-overlapping confidence intervals.

The problem is, for small samples, that computation of confidence intervals for the lognormal and powerlaw distributions are complex and there are different opinions in the literature of which method should be used /see Xiao-Hua and Sujuan, 1997, for a review/. For lognormally distributed variates, it is tempting to compute confidence intervals on the mean of the logarithm and then back-transform the calculations. Though this will indeed yield a statistically correct measure of differences, the calculated measure is the median and not the mean of the variates. Since lognormal distributions can be heavily skewed, the median is a poor measure for the expected value and this method is not recommended. Using simulations, /Xiao-Hua and Sujuan, 1997/ showed that a method of Cox /Land, 1972/ have the smallest coverage error, among the tested methods, for moderate sample sizes (as small as 50). The expression for computing the confidence interval of a lognormal variate is according to Cox:

$$e^{\left(\mu + \frac{\sigma^2}{2}\right)} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^4}{2(n-1)}} \quad \text{Equation 23}$$

where the Z value can be obtained from standard statistical tables. The three trace length samples in Figure 3-19 ( $n=30$  for all) are compared with two methods. Using the method of Cox reveals that there is indeed *not* any statistically significant difference between the sample means whereas the traditional way of computing confidence intervals /e.g. Tukey, 1977/ indicate that sample 3 differs from the other two samples. When the sample sizes were made to exceed about 200, the statistics were virtually identical (not shown in Figure 3-19).

In summary, we recommend the use of presentations as the one displayed in Figure 3-19, as the method is both intuitive and powerful. However, for small sample sizes ( $< 100$  /e.g. Xiao-Hua and Sujuan, 1997/), should parametric methods be used, it should be made clear by which method the confidence intervals were computed.



**Figure 3-19.** Boxplots displaying means values and confidence intervals for three samples ( $n=30$ ) of trace lengths that are lognormally distributed.

### 3.2.4 Area normalisation

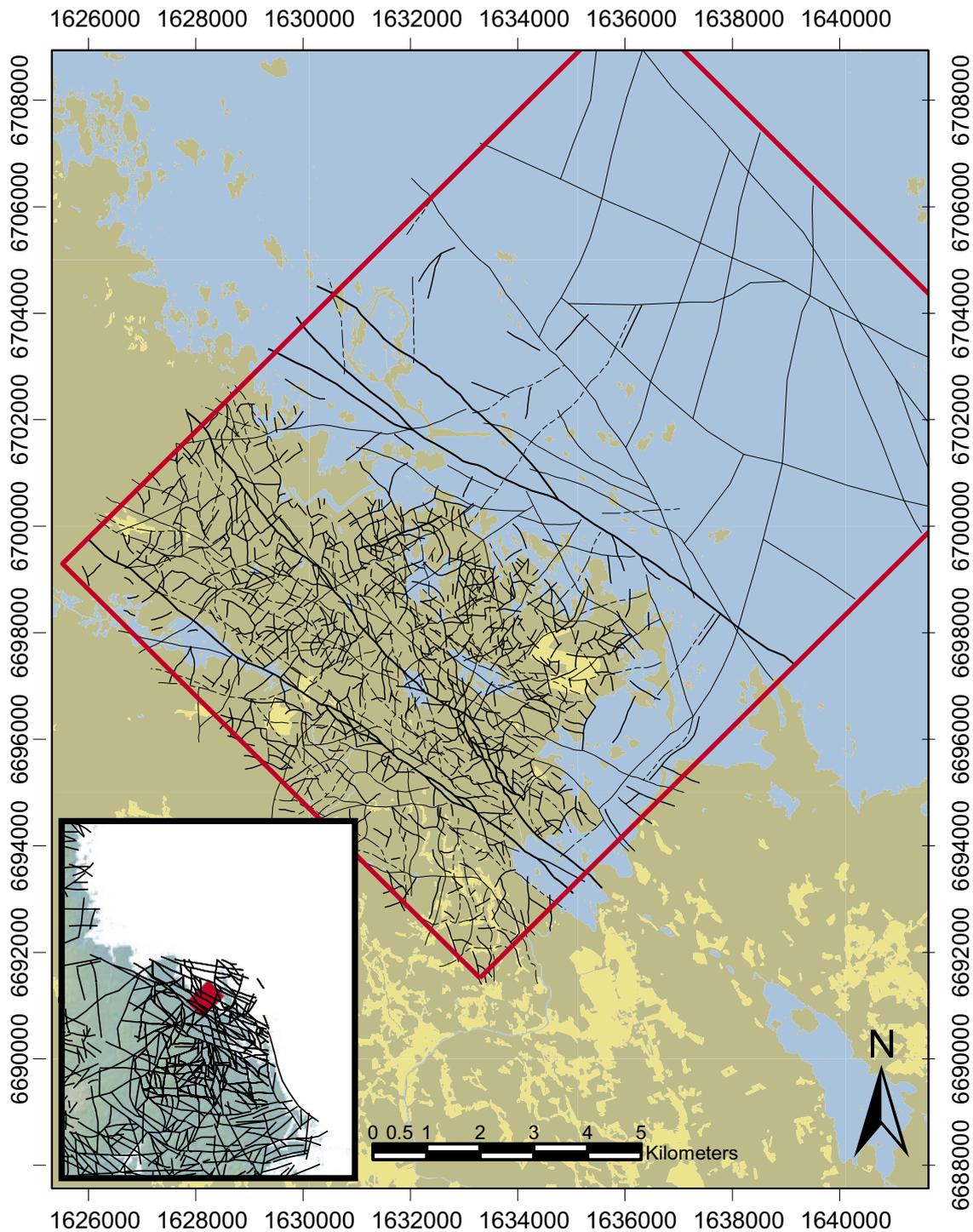
When fracture traces are sampled at many different scales, as outcrops, trenches and combined with for instance lineament data, it is for some analyses essential to perform a process called area normalisation. A fracture sample is area-normalised simply by multiplying the amount of mapped fractures with the ratio between the area of the outcrop, or sum of outcrop areas if more than one, with the area of the larger scale sampling domain, e.g. the lineament map.

The expression for normalising outcrop data to lineament data can be formulated as:

$$N_{adj} = \sum_1^n N_i \frac{\text{area}_{\text{map}}}{\sum_1^n \text{area}_i} \quad \text{Equation 24}$$

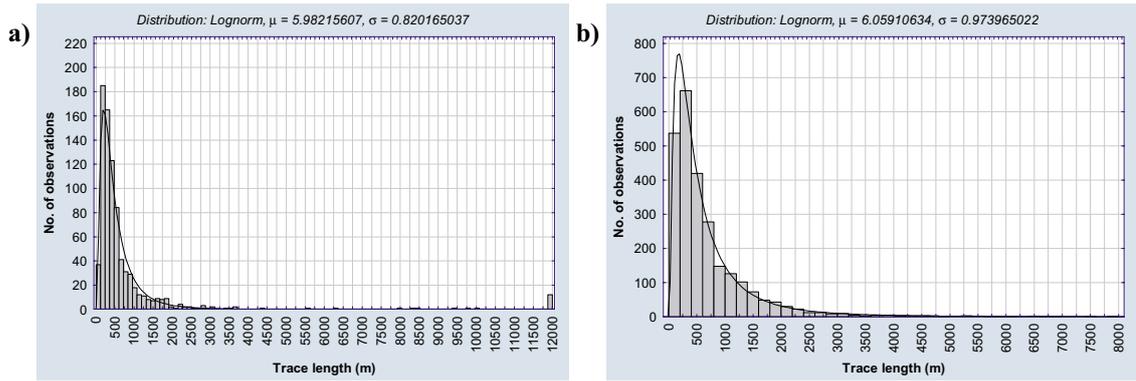
where  $N_{adj}$  is the number of outcrop-scale fractures, mapped on  $n$  outcrops, projected onto the larger scale map.

For illustrative purposes we here provide an example from Forsmark. As shown on Figure 3-20, there is a clear difference in lineament density between the land and sea areas. The main reason is that the geophysical signature of smaller geological structures is obscured by the water coverage and many lineaments are based solely, or heavily, on elevation data. Since large scale lineament show similar density over the model area, it is reasonable to assume that the density of small scale lineaments should also be similar over the entire model area unless, of course, contradicted by other geological data. To compute the lineament density over the entire model area, or the length distribution, (see Section 3.3.2), it is essential to area-normalise the number of small-scale lineaments.



**Figure 3-20.** Lineament map, Forsmark. The insert shows a section of a larger scale lineament interpretation (from /Fredén, 1994/).

The histogram displayed in Figure 3-21a shows the distribution of lineament lengths which are area normalised in Figure 3-21b. The portion of the model that has smaller lineaments interpreted (mainly the land area), constitutes roughly 32% of the entire model area, and the number of lineaments were adjusted accordingly.



**Figure 3-21.** Histogram of a) lineament lengths, b) area-normalised lineament lengths.

There is a minor complication in this procedure that we wish to stress. It is common practice when normalising for area, to superimpose interpretations performed on various scale. The insert in Figure 3-20 shows a portion of a lower resolution, nation-wide lineament interpretation from which the sub sample of those intersecting the model area has been extracted (Figure 3-22). If these lower resolution, but larger scale interpretation is to be incorporated into the analysis, then their higher resolution equivalents must be removed because they are not equal in amount. As magnification increases, it becomes clear from Figure 3-22 that the large scale, single lineament in fact consists of several shorter strands.

The area normalisation exemplified here, is based on the assumption that fracture densities scales linearly with model scale, i.e. Euclidian scaling. However, fracture densities often scales in a non-linear, i.e. fractal, manner.

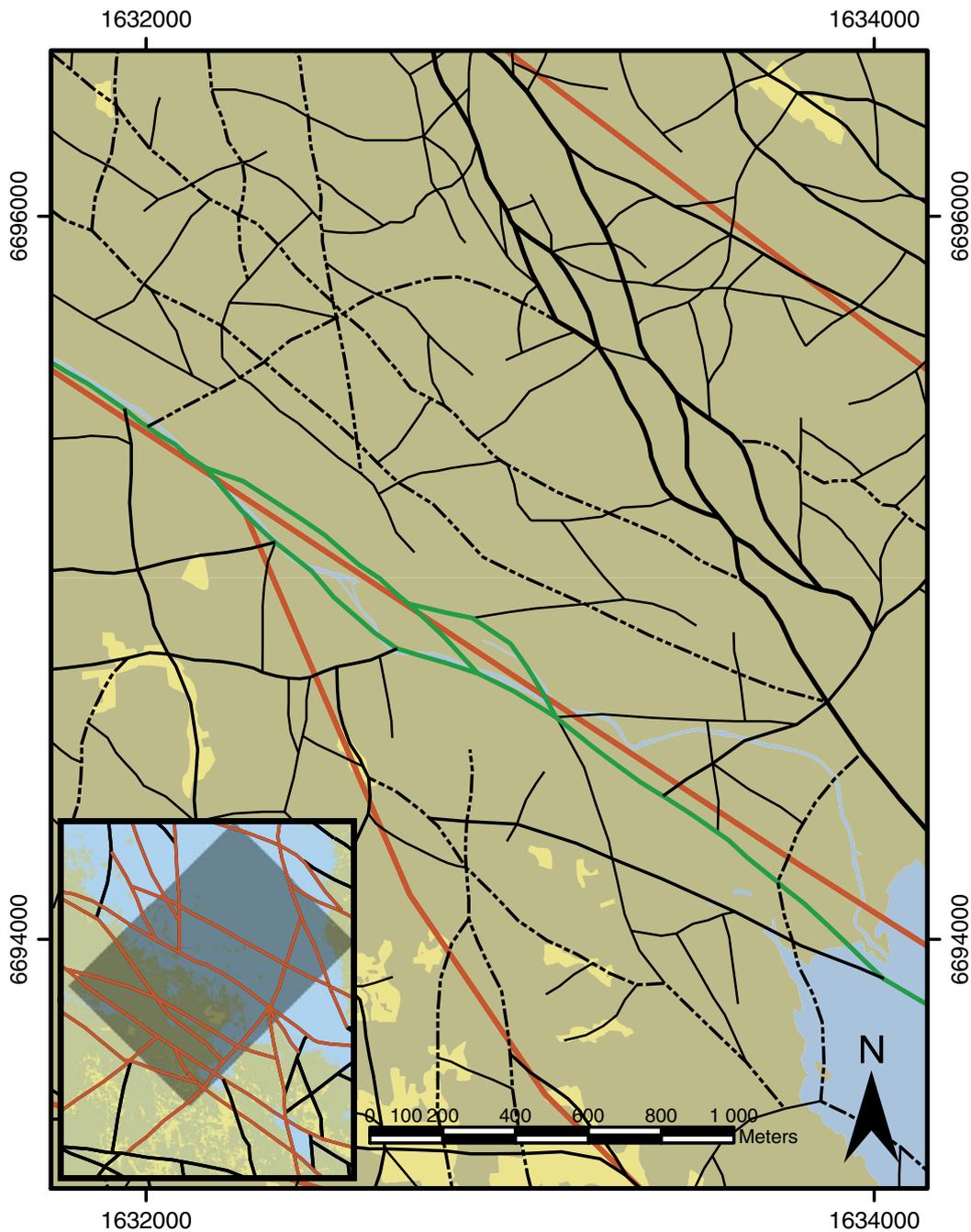
A non-linear area normalisation is commonly based on an estimate of the mass dimension. From Mandelbrot's definition of fractal dimension /Mandelbrot, 1982/:

$$D = \lim_{R \rightarrow \infty} \frac{\ln N(R)}{\ln(R)} \quad \text{Equation 25}$$

where  $N(R)$  is the number of objects within volume  $R$ , it is possible to express a mass dimension  $D_m$  such that:

$$L(r) \rightarrow r^{D_m} \quad \text{Equation 26}$$

in which the mass can be defined as the number of fractures or, alternatively, total fracture trace length  $L(r)$  included in a circular area of radius  $r$ . In a nutshell, Equation 26 predicts the  $P_{21}$  value for any model area knowing  $P_{21}$  for a particular area and assuming a fractal scaling behaviour. It is, however, beyond the scope of this report to elaborate further on this issue. The interested reader is therefore kindly referred to e.g. /Bonnet et al, 2001/ for details on theory and worked examples.



**Figure 3-22.** Detail of Figure 3-20 showing how a large scale lineament (red) is decomposed into several shorter strands (green) as magnification and resolution increases.

### 3.3 Fracture density

#### 3.3.1 Borehole and scanline

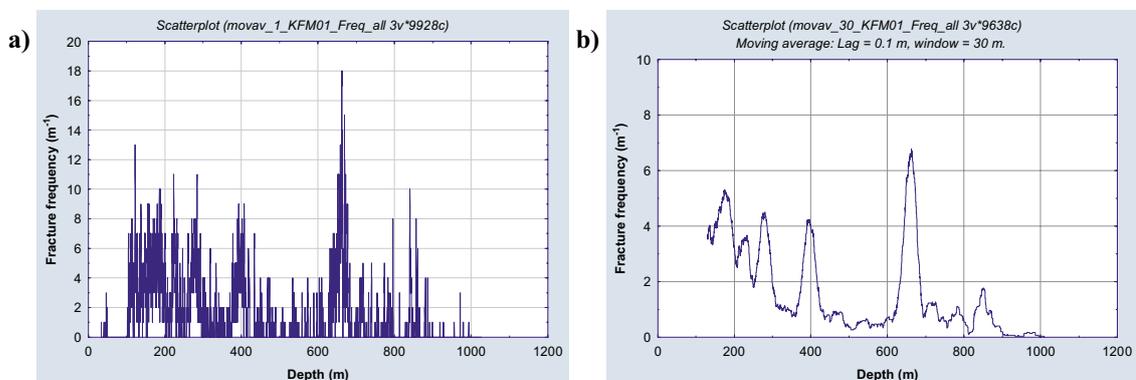
Fracture density can be measured in terms of the number of fractures per metre along a scanline or a borehole, in which case it is a one-dimensional parameter. This is traditionally denoted *fracture frequency*. It should be noted that, due to sampling bias, the measured frequency will be dependent on the orientation of the scanline or borehole (cf Section 3.1.1). There are numerous techniques reported to adjust for this bias /Priest and Hudson, 1981/ but it is out of the scope of this chapter to present details on such procedures.

It has become common practice within SKB projects to adopt the terminology of / Dershowitz and Herda, e.g. 1992/ whereby this parameter is termed  $P_{10}$ , where the subscript 1 stands for the dimensionality of the sampling domain and subscript 0 for the measurement unit (here, a number, i.e. dimensionless).

It is common practice to section borehole logs into 1-m intervals which directly yields the fracture frequency as defined (Figure 3-23a). It is, unfortunately, rarely recognised that even this sectioning constitutes an averaging over an interval, albeit small, with a lag that is equal to the window widths. Additionally, the class boundary, irrespective of the choice of class width, might be located within a local maximum which will be partly hidden. This effect can be avoided by the use of several different class width or, as advocated for here, filtering techniques such as the moving average /Kenney and Keeping, 1962; Whittaker and Robinson, 1967/. In addition, the use of moving average with varying window sizes can reveal properties of fracturing at different scales of interest. In Figure 3-23b we illustrate how a moving average of the data in Figure 3-23a might be displayed. We used a window size of 30 m, and a lag of 0.1 m. Spikes in the graph are indicative of fracture zones, which indeed were confirmed by other means, and the decreasing trend of frequency is made rather clear. The choice of window size will certainly depend on the length of the borehole and the local geology, but also reflect the purpose of the analysis. Often necessary to test various window sizes, e.g. 3 m, 10 m and 30 m it is not possible, therefore, to recommend any particular window or lag size. Clearly, the presence of fracture zones must be accounted for when computing the fracture density since *the DFN is intended to reflect fracture statistics between deterministically modelled zones*. The fracturing within zones is described in other instances of the site investigation. Plots of cumulative frequencies have the property of hiding or blurring “spikes”, making visual inspection difficult if not impossible.

Regardless of whether smoothing techniques have been applied or not, analysis of fracture frequencies in boreholes (and scanlines surveys if applicable) should contain at least the following information:

- Length of the studied series.
- Class width.
- For moving averages, window and lag size.
- If deterministic deformation zones are intercepted, the sections that are excluded from analysis.



**Figure 3-23.** a) Fracture frequency in KFM01, Forsmark site as traditionally computed. b) Moving average of data in a).

The reciprocal of  $P_{10}$  is the fracture spacing which possesses special qualities. It can be shown that for discs randomly distributed in space, a line cutting through will yield spacings of the intersections that are exponentially distributed /Gilvarry, 1964; Baecher, 1983/. Thus the spacing distribution in a borehole can be used to demonstrate spatial correlation (or lack thereof) in the direction of the borehole or scanline.

For illustrative purposes we present a histogram, with various tested distributions in Figure 3-24. The lognormal distribution appears to fit the data well. However, upon closer inspection, all of the tested distribution failed to pass formal, parametric, tests of goodness of fit (non-parametric tests, not shown here, accepted both the lognormal and exponential distributions). In cases where all distributions show equally poor (or good) fits, it is practical to assume an exponential distribution since this will significantly ease analysis and simulation of fractures. This is particularly true if the spatial correlation (Sections 3.2.4 and 3.4) is Euclidian.

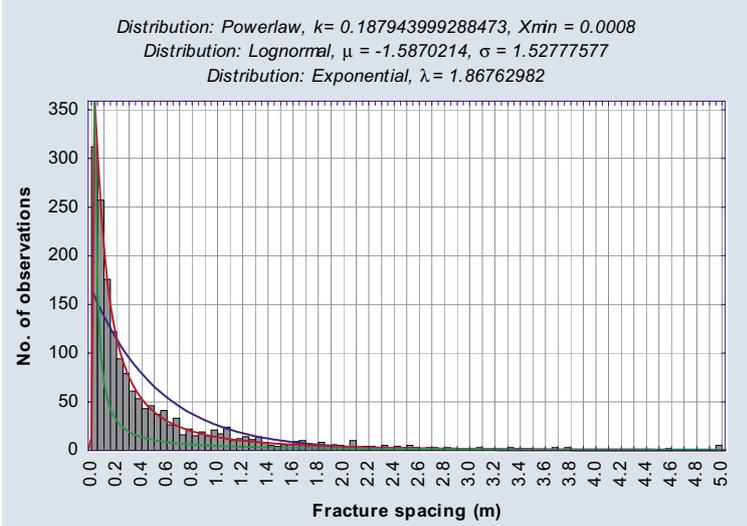


Figure 3-24. Histogram of fracture spacings in subvertical percussion holes, Forsmark.

### 3.3.2 Outcrop

The 2D equivalent of fracture frequency is expressed as  $m/m^2$ . Again, it need to be understood that the outcrop frequency is subject to sampling bias (cf Section 3.2). It is common practice in DFN terminology to denote this entity  $P_{21}$ , were 2 stands for the dimensionality of the sampling domain, and 1 stands for the dimensionality of the mapped feature (the fracture trace). The 2D fracture density,  $P_{21}$ , for an area,  $A$ , can be thus expressed as the sum of  $I$  fracture traces mapped on the outcrop, that is:

$$P_{21} = \frac{1}{A} \sum_{i=1}^I l_i = d_2 \bar{l} \tag{Equation 27.}$$

where  $d_2$  stands for the area centre density and  $\bar{l}$  for the mean trace length. Both  $P_{10}$  and  $P_{21}$  are governed by the orientation (and scale) of the sampling domain. The desired measure of fracture density, therefore, is  $P_{32}$  which is expressed as total fracture area per volume unit ( $V$ ), or:

$$P_{32} = \frac{1}{V} \sum_{j=1}^J a_j = d_3 \bar{a} \quad \text{Equation 28.}$$

where  $d_3$  stands for the bulk centre density,  $J$  the number of fractures and  $\bar{a}$  the mean fracture area:

$$\bar{a} = \frac{1}{J} \sum_{j=1}^J a_j \quad \text{Equation 29.}$$

### 3.3.3 Computation of 3D fracture density

In order to get a fracture density that is not dependent on sampling direction it has been proven very convenient to introduce yet another density measure; Fracture area per volume unit. In analogue to the reasoning above,  $P_{32}$  stands for a 2D measure in a 3D sampling domain with the unit  $m^2/m^3$ . It is, however, simply not possible to directly sample  $P_{32}$  which instead must be derived from the measures  $P_{10}$  and  $P_{21}$ .

There is, fortunately, a relation of proportionality /Dershowitz and Herda, 1992/ between  $P_{10}$ ,  $P_{21}$  and  $P_{32}$  which can be exploited to derive  $P_{32}$ . Again, this problem can be addressed by stochastic simulation. Having established the orientation and length distributions according to Sections 3.1.3 and 3.2.2, the following method can be used to calculate  $P_{32}$ .

1. Initially guess a  $P_{32}$  value,  $P_{32\text{mod}}$ , and simulate a DFN.
2. Sample the DFN with artificial boreholes and outcrops with the same orientations and sizes as the ones that have been mapped.
3. Calculate the fracture density parameters for the sampled boreholes,  $P_{10\text{mod}}$ , and surfaces  $P_{21\text{mod}}$ .
4. Calculate  $P_{32}$  using the relation of proportionality:

$$\frac{P_{10\text{field}}}{P_{10\text{mod}}} = \frac{P_{21\text{field}}}{P_{21\text{mod}}} = \frac{P_{32}}{P_{32\text{mod}}} \quad \text{Equation 30}$$

## 3.4 Evaluation of the spatial pattern of fractures

Fractures are often assumed to occur randomly in rock. Indeed the exact location of fractures is virtually impossible to predict whereas, for instance, the bulk fracture density ( $P_{32}$ ) can be rather accurately determined under favourable conditions. In this context, it is essential to recognise that the randomness can be structured and itself display patterns that are quantifiable. Randomness thus can be said to apply more to the individual fracture than the entire fracture array.

The simplest spatial model assumes that the locations of fracture centroids are completely independent, constant and uniform (isotropic) over the volume of interest. Such a *Poissonian* spatial model can be shown to possess the following properties: If a line cuts through the volume and the fracture locations (centroids) are purely independent, it can be shown /e.g. Gilvarry, 1964/ that their spacings are also independent with an exponential distribution. Naturally, the first step in assessing the spatial correlation or lack thereof is to test if fracture spacings are exponentially distributed. We would like to stress the point that spacings should be investigated in as many directions as possible to ensure isotropy.

In addition to the boreholes, it is possible to directly use a map of fracture midpoints to test for randomness. The map area is divided into a number of equally sized patches such that the patches contain a number of points. If the data are uniformly distributed over the study area, then we would expect each patch to contain the same number of points. The hypothesis of randomness can then be tested with the  $\chi^2$  method according to the following:

The expected number of points,  $E$ , in each patch is:

$$E = \frac{N}{m}, \quad \text{Equation 31}$$

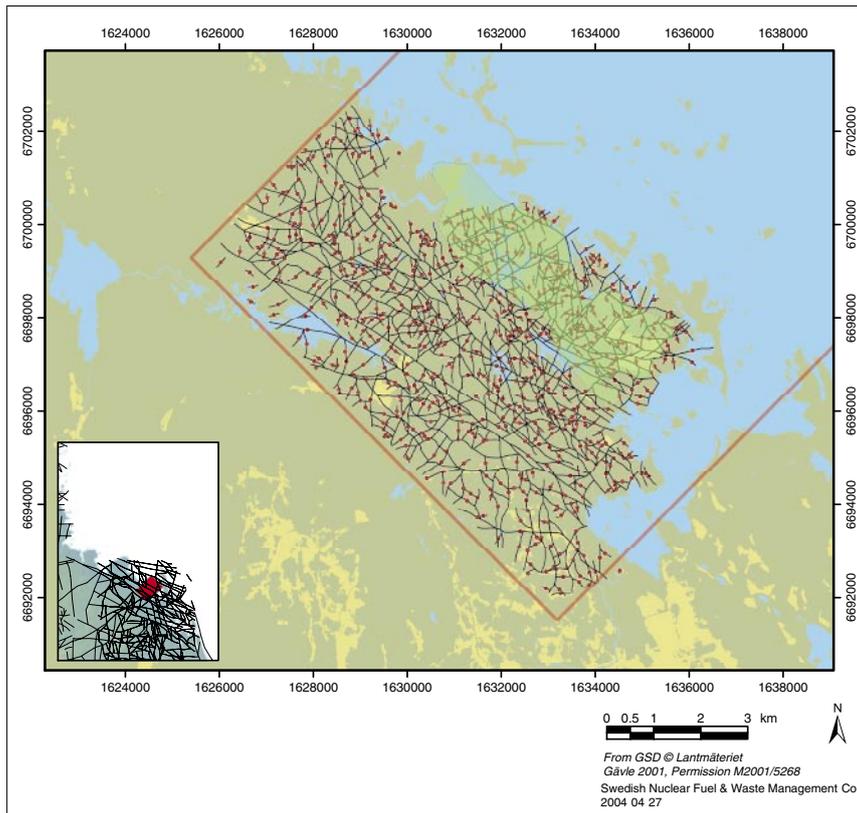
where  $N$  is the total number of points and  $m$  the number of patches.

A  $\chi^2$  test of goodness-of-fit of the expected (uniform) distribution to the observed is given by equation Equation 6, here replicated for convenience:

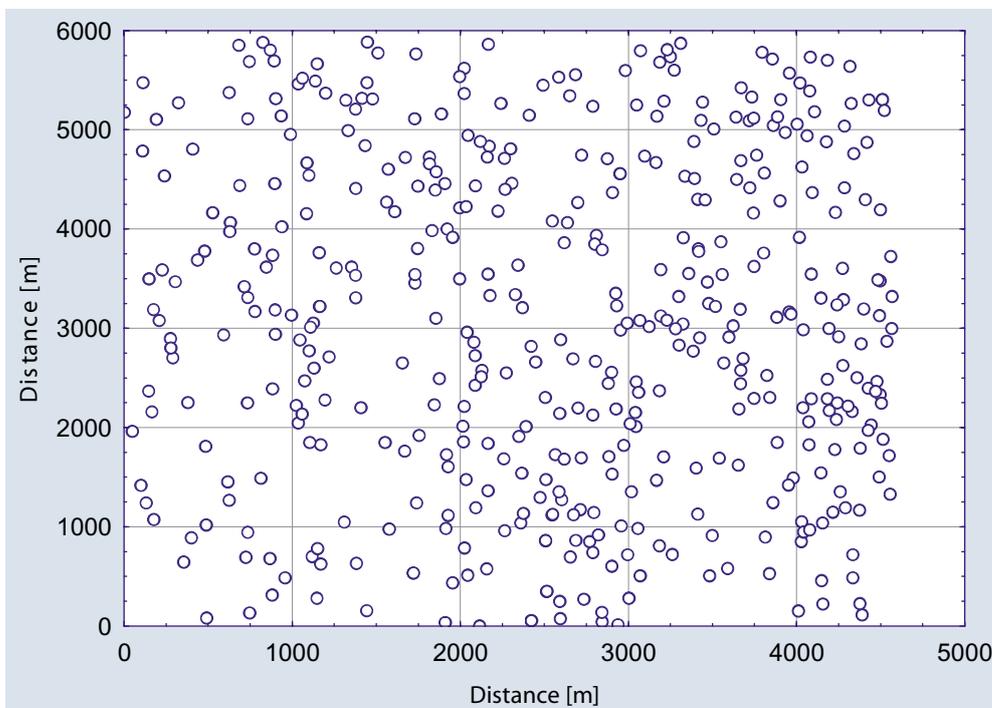
$$\chi^2 = \sum_{i=1}^k \frac{(obs_i - exp_i)^2}{exp_i}$$

with  $\nu = (m-2)$  degrees of freedom.

For illustrative purposes, we applied this technique to lineaments interpreted in the Forsmark model, version 1.1 (Figure 3-25). In Figure 3-26, a rectangular excerpt from Figure 3-25 has been rotated, for ease of computation, and divided into 30 equal patches, each of which we expect to contain about 15 points (451 midpoints in total). The observed number of points in each patch and the computations necessary to find the test value are displayed in Table 3-3. This test has  $\nu = 28$  degrees of freedom, so the critical value of  $\chi^2$  at the  $\alpha = 0.05$  significance level is 41.3. The computed  $\chi^2$  (61) exceed this value so we must conclude that the midpoints are probably not Poisson distributed in the investigated area or, formally, the null hypothesis that the midpoints are evenly distributed is rejected. Though this formalism appear fairly objective it is only apparently so. It is possible to *slightly* alter the result of the test by changing the patch sizes, using patches of unequal size, or even using sophisticated search algorithms aiming to find the patch configuration that yield the minimum  $\chi^2$  value. Thus a certain amount of subjectivity is inherent in the analysis. However, it is only possible to reverse the result of the test if the tested entity yield  $\chi^2$  that lie very close to the critical value at any given significance level. In our example, the calculated  $\chi^2$  value lies too far from the critical value and it was not possible to find, with reasonable efforts, any patch configuration that could reverse the test result.



**Figure 3-25.** The figure shows lineaments and centre points used for analysis. The lineaments are from the Forsmark model, version 1.1.



**Figure 3-26.** Lineament midpoints from Figure 3-25 were rotated for ease of computation.

**Table 3-3. Number of lineament midpoints in cells as defined Figure 3-26.**

Count	Num Of Cells	$\chi^2$
8	1	4
9	4	10
10	2	3
11	2	2
13	4	1
14	3	0
15	2	0
17	1	0
18	3	2
19	1	1
20	1	2
22	2	6
23	1	4
24	1	6
25	1	7
29	1	14
<b>Sum</b>	<b>30</b>	<b>61</b>

Another method based on points on fracture traces (e.g. centroids) is based on the so called correlation dimension,  $D_c$ , as defined by /Hentschel and Proccacia, 1983/. This measure is derived from the two-point correlation function and defined as:

$$C(r) = \frac{1}{N^2} N_d(r) \quad \text{Equation 32}$$

where  $N$  is the total number of points (fracture centroids),  $N_d$  is the number of pairs of points whose distance apart is less than  $r$  /Hentschel and Proccacia, 1983/. For a fractal population of points,  $C(r)$  scales with  $r$  as:

$$D_c = \lim_{r \rightarrow \infty} \frac{\ln[C(r)]}{\ln(r)} \quad \text{Equation 33}$$

Typically this measure range between 1.5 and 2 /see e.g. Bonnet et al, 2001 for an exposé/ and has been recently used within SKB projects /Darcel, 2003/ to demonstrate a fractal pattern of fractures at the Äspö HRL.

Using the same set of lineament centroids as displayed in Figure 3-26 and applying Equation 33 gives a  $D_c$  value of about 1.8 (2 would imply a perfectly poissonian pattern) which implies that the lineament pattern is non-poissonian, and thus confirms the result of the  $\chi^2$  method described above.

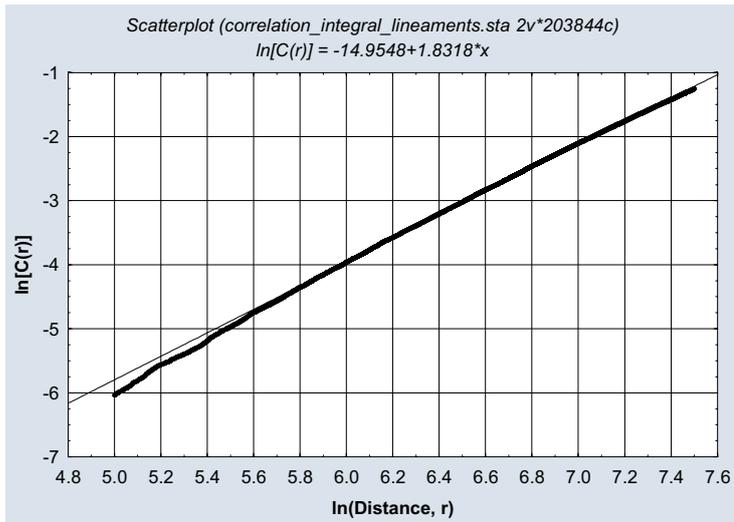


Figure 3-27. Correlation integral of lineament midpoints from Figure 3-25.

The two-point correlation function method is very similar to the box-counting method /see e.g. Kulatilake et al, 1996 for application on fracture networks/. In fact, the correlation dimension  $D_c$  is theoretically equivalent to the box dimension  $D_b$  /Hentschel and Proccacia, 1983/. However, /Bour, 1997/ have shown that the method using the two-point correlation function was better than the box-counting method at distinguishing natural- from purely random patterns. We would therefore, unless otherwise motivated, recommend the use of the correlation dimensions in favour for the box dimension.

A more traditional approach is to compute the density of lineament centroids using well-established algorithms such as nearest neighbour or inverse distance squared. Expanding the simple patch counting method discussed above, it is possible to take into account the number of centroids in neighbouring cells using a certain search radius. In Figure 3-28 we used a patch size of 500 m and a search radius around each point of 250 m. Though such analysis can highlight areas of high centroid density, the patch size in Figure 3-28a is too coarse to be useful to exploit any correlation of densities between the patches. However, using a finer patch configuration (50 m patches) and a larger search radius (500 m), clustering becomes apparent by visual inspection Figure 3-28b.

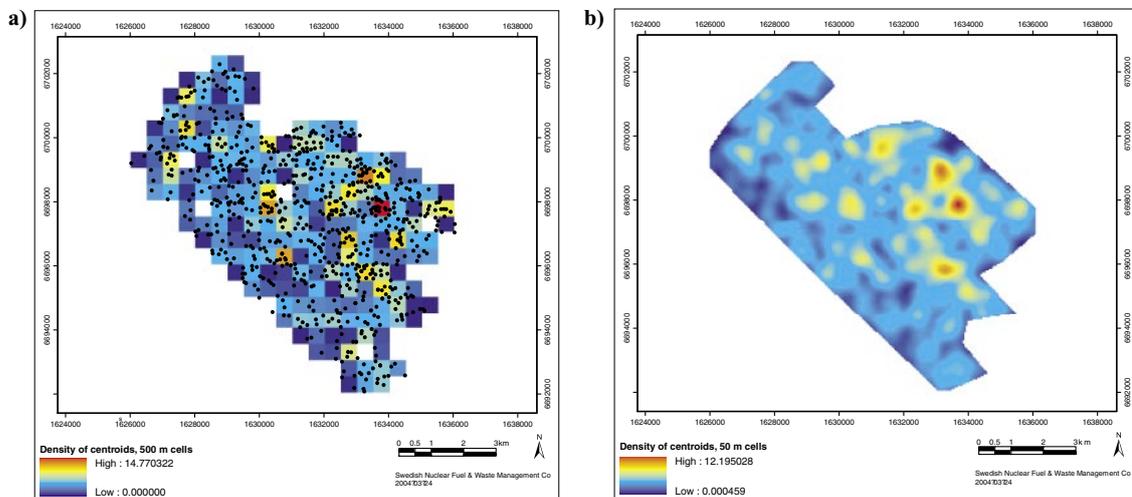
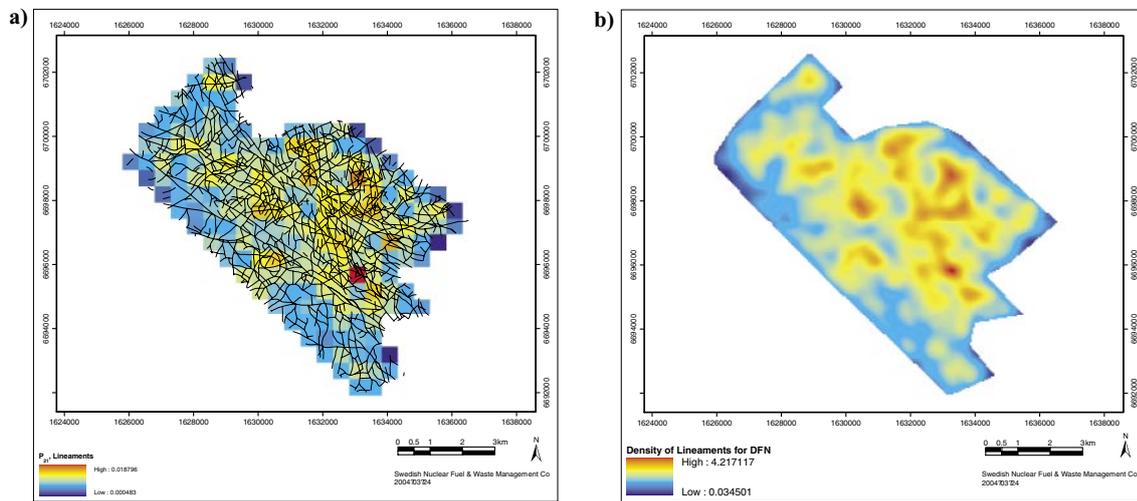


Figure 3-28. Density of lineament midpoints from Figure 3-25.

Using fracture centroids has one severe limitation; it does not take the fracture size or orientation into consideration. Both size and orientation govern the structural anisotropy; that is, when the spatial correlation is unequal in different directions. Another approach, that avoids this drawback, is to directly use one of the estimates of interest,  $P_{21}$ . In Figure 3-29 we computed  $P_{21}$  using a patch size of 50 m, and a smoothing filter using a 500 m wide search radius.

The spatial pattern of lineaments densities (Figure 3-29) can be further exploited using traditional geostatistics /e.g. Isaaks and Srivastava, 1989/. It is far beyond the scope of this report to provide an overview of the available techniques. However, we will for completeness present a simple analysis using a standard technique.

The variability of the lineament density over the map can be found by computing a measure of the self-similarity. That is, the map of densities can be compared to itself at successive positions and the degree of similarity between corresponding intervals computed. In a way, this procedure is very similar to the two-point correlation function with the difference that distances are computed in specified directions.



**Figure 3-29.**  $P_{21}$  values of lineaments. In a) we used a 500 m wide patch with no smoothing. In b) we used a 50 m wide patch and a 500 m wide search radius around each patch centre.

The autocovariance for lag  $h$  is the covariance between all observations  $X_i$  and  $X_{i+h}$ . That is, the covariance is calculated between a series and itself displaced by a lag of length  $h$ . The autocovariance,  $C_h$ , is defined as:

$$C_h = \frac{1}{N(h)} \sum X_i X_{i+h} - \bar{X}_i \bar{X}_{i+h} \quad \text{Equation 34}$$

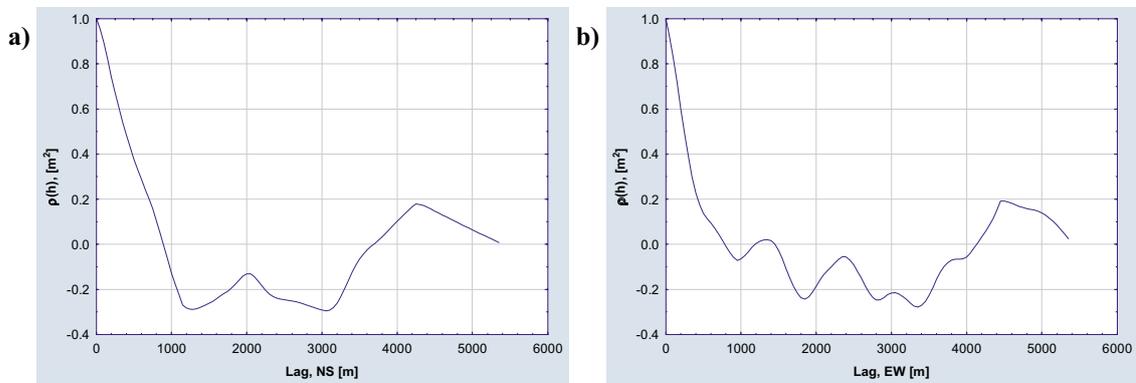
The summation is only over the  $N(h)$  pairs of data whose locations are separated by  $h$ ,  $\bar{X}_i$  and  $\bar{X}_j$  are the mean values of all the data whose locations are  $-h$  and  $+h$  away from some other locations respectively:

$$\bar{X}_i = \frac{1}{N(h)} \sum X_i; \quad \bar{X}_j = \frac{1}{N(h)} \sum X_j \quad \text{Equation 35}$$

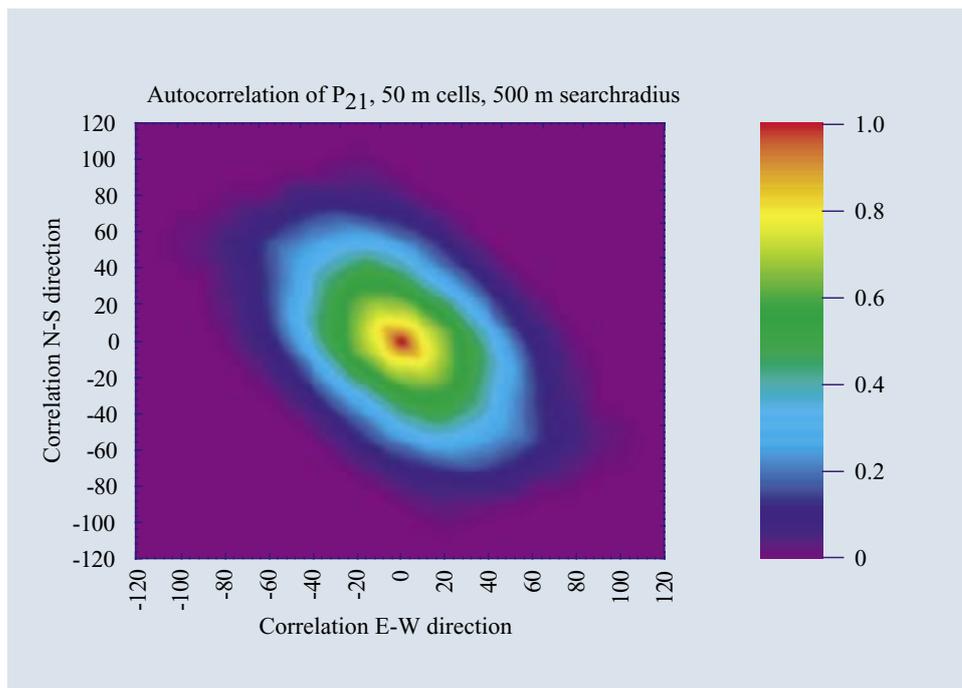
The correlation function,  $\rho(h)$ , is the covariance function standardised by the appropriate standard deviations:

$$\rho(h) = \frac{C(h)}{\sigma_i \sigma_j} \quad \text{Equation 36}$$

An average of all correlation functions computed in the NS direction through the computed  $P_{21}$  values (Figure 3-29), is displayed in Figure 3-30a. The correlation between neighbouring cells of  $P_{21}$  values drops smoothly and becomes anti-correlated at a lag distance of about 900 m. The correlation in NS direction is slightly different from the correlation in EW (Figure 3-29b), implying an anisotropy in correlation length. This can be exploited by calculating the correlation function in many directions and derive the axes of the ellipsoid that is formed when the correlograms are combined. Such a computation is displayed in Figure 3-31. The anisotropy ratio, approximately 1.93, is the ratio between the largest axis, which trends  $328^\circ$ , and the smallest axis of the ellipse.



**Figure 3-30.** Autocorrelogram of  $P_{21}$  values computed in NS (a) and EW (b) directions.



**Figure 3-31.** 2D Autocorrelogram of  $P_{21}$  values.

Certainly, this measure can be used to address the correlation and anisotropy in 3D. The extension of Equation 32 from 2D to 3D is, in principal, trivial though computationally extensive in practice. The complication lies in the general lack of underground data. We therefore do not provide any worked example on the fracture sample used above. However, we recommend the use of the following terminology, standard in structural geology, when describing the structural anisotropy:

- Isotropic or Spherical:  $\lambda_1 = \lambda_2 = \lambda_3$
- Oblate:  $\lambda_1 > \lambda_2 = \lambda_3$
- Prolate:  $\lambda_1 = \lambda_2 > \lambda_3$

where  $\lambda$  is the relative length of the principle axes of the anisotropy ellipsoid.

A closely related, alternative, measure which is favoured by some practitioners is the semivariance that can be expressed as /Isaaks and Srivastava, 1989/:

$$\gamma(h) = \frac{1}{2N} \sum_i^{N-h} (X_i - X_{i+h})^2 \quad \text{Equation 37}$$

In fact, if the studied variable ( $P_{21}$  in our case) is stationary, the semivariance for a lag  $h$  is equal to the difference between the variance and the spatial autocovariance for the same distance. If the variable is also standardised (mean = 0, variance = 1), the semivariogram is a mirror image of the autocorrelogram.

From an SKB modelling perspective, we do not consider it important which method is used to address spatial correlation and anisotropy since they are all closely related. We do, however, consider it important not to take isotropy for granted, but to explore the potential existence of anisotropy in 2D and, if possible, 3D using any of the methods mentioned here or, if properly argued for, any other suitable measure.

It is interesting to note, that while certain practitioners prefer to address fracturing patterns from a fractal point of view, others approach the problem from a geostatistical point of view. Yet, it can be shown that these, too, are very closely related; Consider the following:

If the values of  $P_{21}$  have a fractal spatial pattern, the matrix of  $P_{21}$  values can be regarded as a fractal surface. A vertical section through a fractal (self affine) surface  $X$  yields a fractal (self affine) curve which can be defined as /Peitgen et al, 1992/:

$$\text{variance} [X(t_2) - X(t_1)] \propto |t_2 - t_1|^{2H} \quad \text{Equation 38}$$

with a so called /Hurst, 1957/ exponent,  $H$ , such that  $0 < H < 1$ . An increment of  $X$  is said to be statistically self-affine with the parameter  $H$  or,

$$X(t) - X(t_0) \text{ and } \frac{X(rt) - X(t_0)}{r^H} \text{ are statistically identical. This can be shown by letting}$$

$t_0 = 0$  and  $X(t_0) = 0$  which yields:

$$X(t) \text{ and } \frac{X(rt)}{r^H}$$

A fractal curve,  $X(rt)$ , can thus be rescaled by dividing the amplitude by  $r^H$ . The lag,  $h$ , in Equation 37 corresponds to  $t_2 - t_1$  in Equation 38 and  $X_i$  in Equation 37 corresponds to  $X(t_2) - X(t_1)$  in Equation 38. Equation 37 can therefore be rewritten:

$$\gamma(h) = A(h)^{2H} \quad \text{Equation 39}$$

where  $A$  is a constant /Gallant, 1992; Gallant et al, 1993/

Since  $H$  is related to the fractal dimension,  $D$ , as :

$$D = 2 - H \quad \text{Equation 40}$$

we can relate the semivariance of a section or a surface to its fractal dimension as:

$$\gamma(h) = A(h)^{4-2D} \quad \text{Equation 41}$$

thereby bridging two apparently dissimilar modelling approaches.

## 4 Expected Output from DFN analyses

The procedure of defining a set of DFNs for a geological model is composed of several, somewhat dissimilar, components. Commonly, any DFN analysis is initiated by a thorough statistical treatment of the sampled data. The critical parameters to compute for all fractures, and also separately for the main fracture classes “open” and “sealed”, are:

- orientation of the fractures,
- size of the fractures,
- fracture density,
- fracture termination (how fractures cross each other),
- the spatial correlation of the fractures (both horizontally and vertically),
- structural control, e.g. correlation to neighbouring large scale deformation zones, large scale folding, etc.

Additionally, the following properties are regarded important:

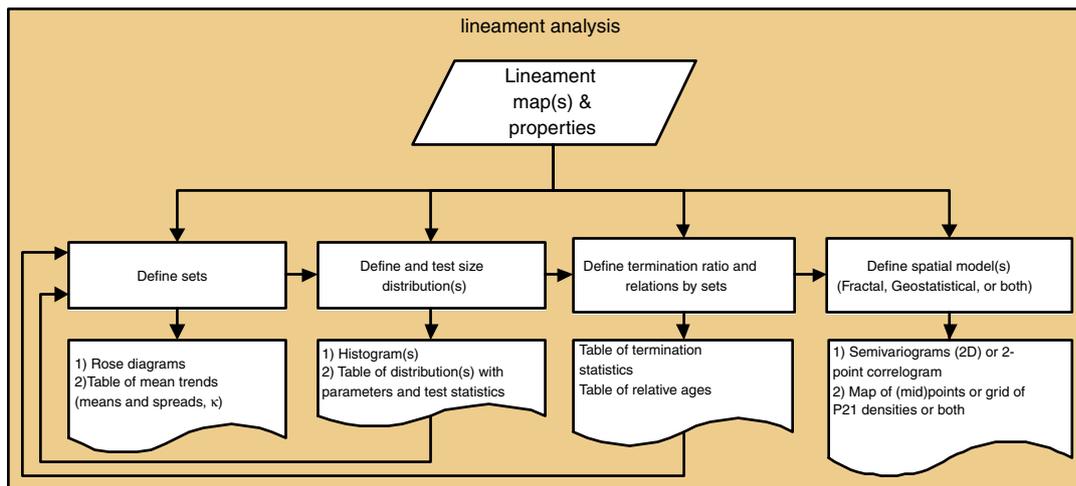
- The fracture surface roughness
- Mineralogy

Any potential correlation between these parameters, e.g. between length and orientation, mineralogy and orientation, etc, should be addressed, and analyses could be structured in a manner similar to the schematic correlation matrix proposed below:

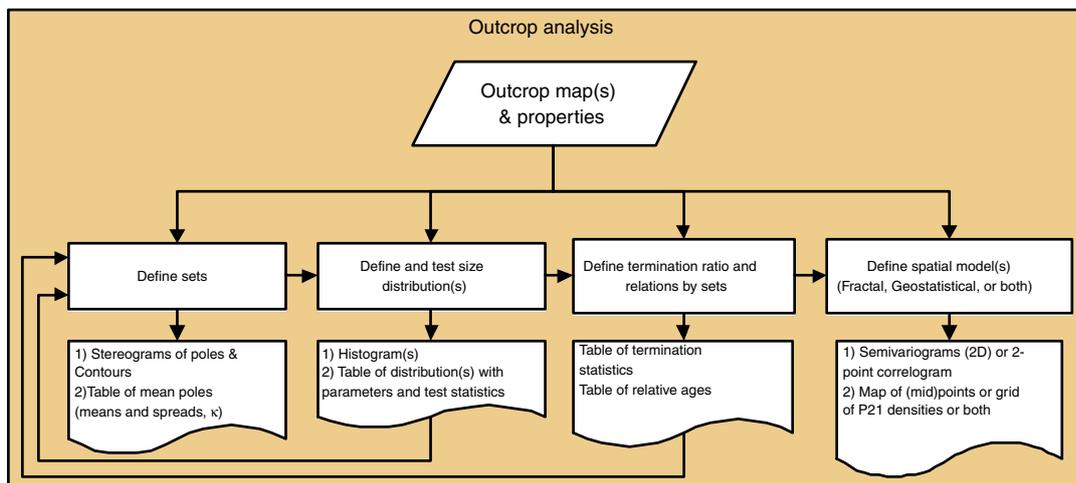
	Orientation	Size	Density	Termination	Mineralogy	Roughness
Orientation	-					
Size		-				
Density			-			
Termination				-		
Mineralogy					-	
Roughness						-

The analysis of lineaments provides information on the spatial correlation, length distribution and 2D orientation of large scale structures (Figure 4-1) over the entire model volume. The study of large scale structures can also provide qualitative information such as relative age relations and, above all, constitute the geometrical framework within which the DFNs are to be defined. Additionally, the impact of large scale structures on smaller scale fracturing is addressed by analysing the geometry of lineament in relation to the geometry of smaller scale fractures. Figure 4-1 summarises schematically the desired information, analyses and briefly the output of such analysis.

The analysis of outcrop fracture data is basically similar to the analysis of lineaments, with the main difference that small scale fractures are mapped with both strike and dip (Figure 4-2). Additionally, outcrop analysis provides input to define separate DFN models for each defined domain in the geological model.



**Figure 4-1.** Schematic workflow of lineament analyses and expected output for the construction of DFNs.

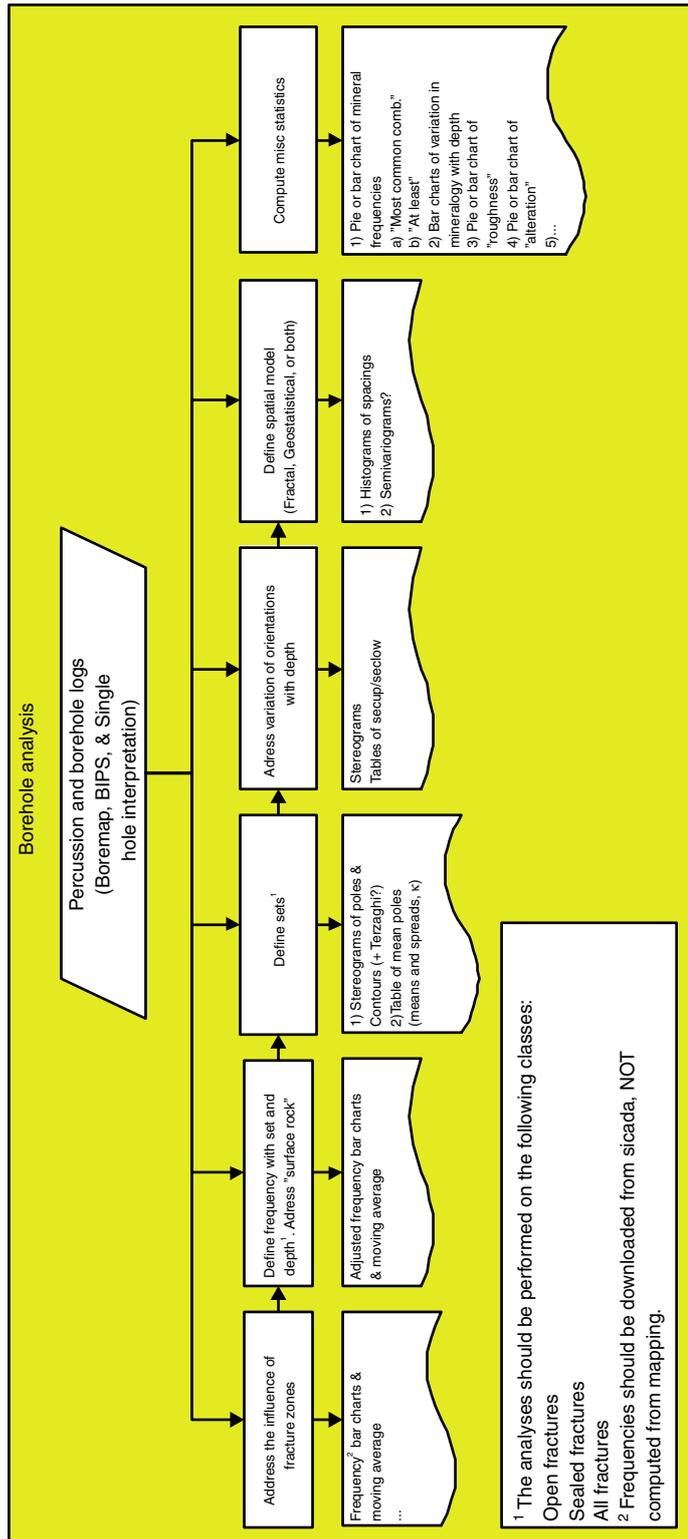


**Figure 4-2.** Schematic workflow of outcrop analyses and expected output for the construction of DFNs.

Analysis of borehole data constitutes a special challenge because:

- borehole data provide the highest possible level of fracture detail and accuracy in position,
- provides depth information, but
- suffers from orientation and sampling bias.

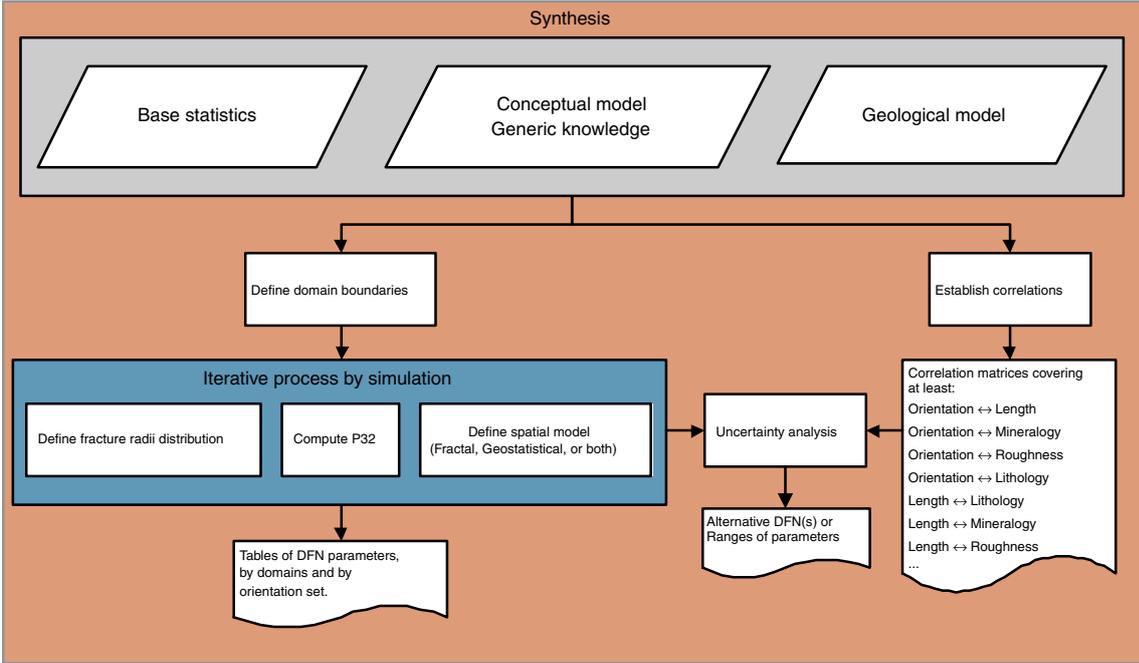
The borehole data is essential to compute the fracture density in a rock volume. Additionally, borehole data provide means to address any changes of fracture density with depth, boundary between different domains, rotation of fracture orientations, etc. Figure 4-3 illustrates schematically the minimum amount of analyses required and their expected output.



**Figure 4-3.** Schematic workflow of borehole analyses and expected output for the construction of DFNs.

Though most analyses briefly outlined above can be denoted “traditional statistics”, the synthesis of lineament-, outcrop-, borehole- and generic data paired with conceptual understanding of a geological environment is unique to DFN analyses. The procedure here called synthesis (Figure 4-4) can be rather heavily stamped by scientific discipline (cf fractal vs geostatistical approach), working environment (e.g. preferences or “styles”), etc, and it is therefore not possible for SKB to request any specific method of work or work plan. We do, however, request that the *output* of the analyses conforms to the expectations of the different “customers” within SKB and that the working procedure is transparent, traceable and reproducible. In this context, we find it extremely important that the conceptual models are clearly stated, how these were applied in the analysis and, finally, how these are manifested in the final outcome.

To achieve this goal, we find it practical to clearly separate the initial statistical treatment from subsequent analyses, as schematically illustrated in Figure 4-5. We suggest that base statistics constitute the bulk of an appendix in the DFN report, whereas the main report should focus on the ennoblement of the base statistics to address issues such as the estimation of the true fracture sizes from trace length information, the spatial correlation, variations with depth, arguments for subdomains, structural control, assumptions of modelling, etc.



**Figure 4-4.** Schematic workflow of the synthesis of analyses and expected output for the construction of DFNs.

Typically, the base statistics could be comprised of histograms, probability plots, contingency tables, etc for each sampling domain, i.e. borehole, outcrops, etc, separately. The purpose is to provide the receiver of the DFN model a clear set of background information that enables the user to judge the validity and possible consequences of the assumptions and conclusions made in subsequent analyses presented in the main DFN-model report. All data sets extracted from SKB databases should be properly referenced by filename, log, SQL, or equivalent so that the presented analysis can be easily and accurately reproduced.

The final outcome of the DFN analyses can, in fact, be boiled down to simple tables, examples of which are shown below (Table 4-1 and Table 4-2). Such tables should be produced for each rock domain, in the geological model, that has been judged internally homogeneous with respect to the DFN statistics. That does not exclude the possibility of describing a smooth transition of a parameter within a rock domain, e.g. an exponentially decreasing fracture density with depth or distance from any other structure, e.g. domain boundary or deterministically modelled deformation zone. Should it, however, be judged necessary to define subdomains with respect to DFN properties, in addition to those already provided by the geological model, it is important to consult with SKB at an early stage in this matter. This because discrete, horizontal domain boundaries at repository depth, at the moment 400–700 m, may impose severe implication on the projecting.

**Table 4-1. Tentative example of a summary table for DFN parameters.**

Set #	Orientation <sup>1</sup> (T/P, Fisher $\kappa$ )	Size distribution <sup>2</sup> (pdf, parameters)	Density (m <sup>2</sup> /m <sup>3</sup> )	Spatial model
1	219.0/83.7, 4.84	Lognormal, $\mu=2.0$ , $\sigma=2.0$	0.42	Poisson
2	127.0/84.2, 8.35	Lognormal, $\mu=2.1$ , $\sigma=2.5$ ,	0.34	Poisson
3	20.6/6.0, 8.33	Powerlaw, $k = 2.8$ , $x_{min} = 0.5$	1.2	Poisson

<sup>1</sup> The orientation may be given either as the trend/plunge of the pole, or as strike/dip of the plane.

<sup>2</sup> The lower censoring,  $X_{min}$ , and the boundary to discrete structures,  $X_{max}$ , must be stated regardless of the chosen distribution.

**Table 4-2. Tentative example of a summary table of additional qualifiers for DFN parameters.**

Set #	Termination (%)	Open (%)	Roughness JRC class	Mineralogy % containing at least indicated
1	25%, set#2 40%, set#3	10	4	Chl (80), Ca (60), Ep (20)
2	50%, set#1 20%, set#3	10	8	Chl (80), Ca (70), Lau (15)
3	30%, set#1 8%, set#2	15	10	Chl (60), Fe (10), Lau (10)

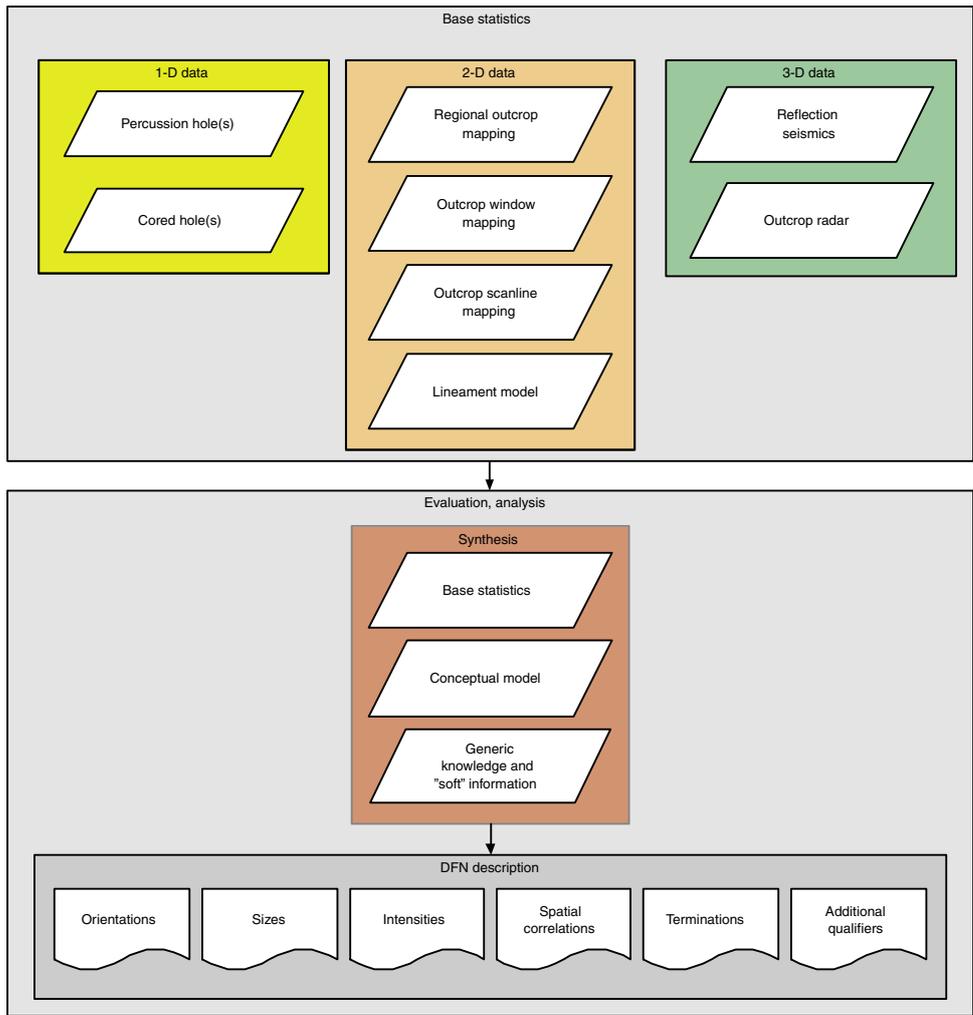


Figure 4-5. Schematic structure of DFN analysis and reporting.

## 5 References

- Adam J F, 1989.** New methods of counting and contouring densities of three-dimensional orientation data. International Geological Congress, Abstracts--Congres Geologique Internationale, Resumes. Vol 28(1): p 10.
- An L J, Sammis C G, 1994.** Particle size distribution of cataclastic fault materials from Southern California; a 3-D study. Pure and Applied Geophysics. Vol 143(1-3): p 203-227.
- Baecher G B, 1983.** Statistic analysis of rock mass fracturing. Mathematical geology. Vol 15: p 329-348.
- Baecher G B, Lanney N, 1978.** Trace length biases in joint surveys.
- Berkowitz B, Hadad A, 1997.** Fractal and multifractal measures of natural and synthetic fracture networks. J. Geoph. Res. Vol 102(12205-12218).
- Bingham C, 1964.** Distributions on the sphere and on the projective plane. Ph.D. Thesis. Yale University.
- Bonnet E, Bour O, Odling N, Main I, Berkowitz B, Davy P, Cowie P A, 2001.** Scaling of fracture systems in geological media. rev. Geophys. Vol 39: p 347-383.
- Bour O, 1997.** Transferts de fluides dans les milieux fracturés: effets d'échelle. PHD thesis. Université de Rennes.
- Call R D, Savely J P, Nicholas D E, 1976.** Estimation of joint set characteristics from surface mapping data.
- Cowie P A, Sornette D, Vanneste C, 1995.** Multifractal scaling properties of a growing fault population. Geophysical Journal International. Vol 122(2): p 457-469.
- Cruden D, Ranalli G, 1977.** Length distribution of strike-slip faults and the process of breakage in continental crust. Canadian Journal of Earth Sciences = Journal Canadien des Sciences de la Terre. Vol 14(3): p 508-510.
- Darcel C, 2003.** True Block Scale continuation project. Assessment of the feasibility of tracer tests with injection in "background fractures" using a model based on a power law fracture length distribution. IPR-03-41, Svensk Kärnbränslehantering AB.
- Davis J C, 1986.** Statistics and Data Analysis in Geology. 1986, New York: John Wiley & Sons. p 646. ISBN.
- Dershowitz W, Herda H H, 1992.** Interpretation of fracture spacing and intensity. in Proc. U.S. Symp Rock Mech. 1992.
- Dershowitz W, Lee G, Geier J, 1991a.** FracMan Version 2.3. Interactive discrete feature data analysis, geometric modeling, and exploration simulation. User documentation. Golder Associates Inc. Redmond Washington USA.

- Dershowitz W, Lee G, Geier J, Foxford T, Lapointe P, Thomas A, 1995.** FRACMAN-Interactive Discrete Feature Data Analysis, Geometric Modeling, and Exploration Simulation. User Documentation, version 2.5, Version Golder Associates Inc.
- Dershowitz W, Wallmann P, Kindred S, 1991b.** Discrete fracture modelling for the Stripa site characterization and validation drift inflow predictions. Stripa Project Technical Report 91-16, Svensk Kärnbränslehantering AB.
- Epstein H H, 1947.** The mathematical description of certain breakage mechanisms leading to the logarithmico-normal distribution, *J. Frank. Inst.* Vol 244: p 471–477.
- Evans M, Hastings N, Peacock B, 1993.** Statistical distributions. 2:nd ed. 1993, New York: John Wiley & Sons, Inc. ISBN 0-471-55951-2.
- Fisher N I, Lewis T, Embleton B J, 1987.** Statistical Analysis of Spherical Data. 1987, Cambridge, U. K.: Cambridge University Press. p 329. ISBN.
- Fisher R, 1953.** Dispersion on a sphere. Vol 217. 1953: Royal Society of London Proceedings. p 295–305. ISBN.
- Fredén C, 1994.** Sveriges National Atlas. Berg och jord. 1994, Sweden: Sveriges Nationalatlas Förlag (SNA)/LMV SSAG SCB. ISBN.
- Fujiwara A, Kamimoto G, Tsukamoto A, 1977.** Destruction of basaltic bodies by high-velocity impact. *Icarus.* Vol 31(2): p 277–288.
- Gallant J, pers. comm 1992.**
- Gallant J, Moore I, Gessler P, 1993.** Estimating fractal dimensions of profiles: A comparison of methods, *Mathematical geology.* Vol
- Geier J E, Lee K, Dershowitz W S, 1988.** Field validation of conceptual models for fracture geometry, in AGU 1988 fall meeting, Anonymous, Editor. 1988, American Geophysical Union: Washington, DC, United States. p 1177.
- Gilvarry J J, 1964.** Distribution of fragment size in repetitive fracture of brittle solids. *Solid State Communications.* Vol 2(1): p 9–11.
- Giné E M, 1975.** Invariant tests for uniformity on compact Riemannian manifold based on Sobolev norms. *Annals of Statistics.* Vol 3: p 1243–1266.
- Hackbusch W, 1985.** Multi-Grid Methods and Applications. 1985, Berlin, New York.: Springer-Verlag. ISBN.
- Hentschel H G E, Proccacia I, 1983.** The infinite number of generalised dimensions of fractals and strange attractors. *Physica.* Vol 8(435–444).
- Hext G R, 1963.** The estimation of second-order tensors, with related tests and designs, *Biometrika.* Vol 50: p 353–373.
- Huang Q, Charlesworth H, 1989.** A Fortran-77 program to separate a heterogeneous set of orientations into subsets. *Computers & Geosciences.* Vol 15(1): p 1–7.
- Hurst H, 1957.** A suggested statistical model of some time series which occur in Nature. *Nature.* Vol 180:494.

- Isaaks E H, Srivastava R M, 1989.** An introduction to applied geostatistics. 1989: Oxford University Press. ISBN 0-19-505013-4.
- Jensen J L, Lake L W, Corbett P W M, Goggin D J, 2000.** Statistics for petroleum engineers and geoscientists, Handbook of Petroleum Exploration and Production. 2:nd ed. 2000, Amsterdam: Elsevier. ISBN.
- Johnson N L, Kotz S, 1972.** Continuous Multivariate Distributions. Distributions in statistics, ed. Bradley, R.A. 1972: Wiley. ISBN 0-471-44370-0.
- Kagan Y Y, 1997.** Seismic moment-frequency relation for shallow earthquakes; regional comparison. Journal of Geophysical Research, B, Solid Earth and Planets. Vol 102(2): p 2835–2852.
- Kamb W B, 1959.** Ice petrofabric observations from Blue Glacier, Washington, in relation to theory and experiment, J. Geophys. Res. Vol 64: p 1891–1909.
- Kenney J F, Keeping E S, 1962.** Moving Averages, in Mathematics of Statistics. 1962, Van Nostrand: Princeton, NJ. p 221–223.
- King G, 1983.** The accommodation of large strains in the upper lithosphere of the Earth and other solids by self-similar fault systems; the geometrical origin of b-value. Pure and Applied Geophysics. Vol 121(5–6): p 761–815.
- Kirkpatrick S, Gelatt C D, Vecchi M P, 1983.** Optimization by Simulated Annealing. Science. Vol 220( 4598): p 671–680.
- Kohlbeck F, Scheidegger A E, 1985.** The power of parametric orientation statistics in the earth sciences. Mitteilungen der Oesterreichischen Geologischen Gesellschaft. Vol 78(2): p 251–265.
- Korvin G, 1992.** Fractal models in the earth sciences. 1992: Cambridge University Press. ISBN.
- Kulatilake P H S W, Fiedler R, Panda B B, Pan G, 1996.** Box fractal dimension and the first invariant of fracture tensor of fracture networks as measures of statistical homogeneity of jointed rock masses, in Proceedings of the 2nd North American rock mechanics symposium; NARMS '96, a regional conference of ISRM; Rock mechanics tools and techniques, Aubertin, M, Hassani, F, and Mitri Hani, S, Editors. 1996, A. A. Balkema: Rotterdam, Netherlands. p 1779–1786.
- Kulatilake P H S W, Wathugala D N, Poulton M, Stephansson O, 1990.** Analysis of structural homogeneity of rock masses. Engineering Geology. Vol 29(3): p 195–211.
- Kulatilake P H S W, Wu T H, 1984.** Estimation of mean trace length of discontinuities. Rock Mechanics and Rock Engineering. Vol 17: p 215–232.
- La Pointe P, 2002.** Derivation of parent fracture population statistics from trace length measurements for fractal fracture populations, International Journal of Rock Mechanics and Mining Sciences. Vol 39: p 381–388.
- La Pointe P, Hudson J A, 1985.** Characterization and interpretation of rock mass joint patterns, Geological Society of America. Vol Special Paper 199.
- Land C E, 1972.** An evaluation of approximate confidence interval estimation methods for lognormal means. Technometrics. Vol 14: p 145–158.

- LaPointe P R, Cladouhos T T, Outters N, Follin S, 2000.** Evaluation of the conservativeness of the methodology for estimating earthquake-induced movements of fractures intersecting canisters. SKB TR-00-08, Svensk Kärnbränslehantering AB.
- Lewis C, Gray N H, 1985.** Contouring orientation data; an approach using spherical Voronoi tessellations, in The Geological Society of America, 98th annual meeting. 1985, Geological Society of America (GSA): Boulder, CO, United States. p 643.
- Mahtab M A, Yegulalp T M, 1984.** A similarity test for grouping orientation data in rock mechanics, in Rock mechanics in productivity and protection, Dowding Charles, H. and Singh Madan, M, Editors. 1984, A.A. Balkema: [location varies], United States. p 495–502.
- Mandelbrot B, 1982.** The Fractal Geometry of Nature. 1982, San Fransisco: Freeman. ISBN.
- Marrett R, Allmendinger R W, 1991.** Estimates of strain due to brittle faulting: Sampling of fault populations, J. Struct. Geol. Vol 13: p 735–738.
- Miller S M, 1983a.** A statistical method to evaluate homogeneity of structural populations, in Geostatistics, geomathematics, and computer applications, Jones, T.A, Editor. 1983a, Plenum [for the] International Association for Mathematical Geology: New York–London, International. p 317–328.
- Miller S M, 1983b.** A statistical method to evaluate homogeneity of structural populations. Mathematical Geology. Vol 15: p 317–328.
- Munier R, 1995.** Studies of geological structures at Äspö. Comprehensive summary of results. SKB HRL Progress Report 25-95-21, Svensk Kärnbränslehantering AB.
- Munier R, Stenberg L, Stanfors R, Milnes A G, Hermanson J, Triumf C-A, 2003.** Geological site descriptive model. A strategy for the model development during site investigations. SKB R-03-07, Svensk Kärnbränslehantering AB.
- Okubo P G, Aki, K, 1983.** Fractal geometry in the San Andreas Fault system, in American Geophysical Union; 1983 fall meeting. 1983, American Geophysical Union: Washington, DC, United States. p 766.
- Pecher A, 1989.** SCHMIDTMAC; a program to display and analyze directional data. Computers & Geosciences. Vol 15(8): p 1315–1326.
- Peitgen H-O, Jürgens H, Saupe D, 1992.** Chaos and fractals, new frontiers of science. 1992, New York: Springer Verlag. ISBN 3-540-97903-4.
- Piggot A, 1997.** Fractal relations for the diameter and trace lengths of disc -shaped fractures, J. Geoph. Res. Vol 102: p 18121–18125.
- Pólya G, 1920.** Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem. Mathematische Zeitschrift. Vol 8: p 171–181.
- Priest S D, 1993.** Discontinuity analysis for rock engineering. 1st ed. 1993, London: Chapman & Hall. ISBN 0-412-47600-2.
- Priest S D, Hudson J A, 1981.** Estimation of discontinuity spacings and trace length using scanline surveys, Int. J. Rock Mech. Min. Sci. Vol 18.

- Renshaw C E, Park J C, 1997.** Effect of mechanical interactions on the scaling of fracture length and aperture. *Nature (London)*. Vol 386(6624): p 482–484.
- Robin P Y F, Jowett E C, 1986.** Computerized density contouring and statistical evaluation of orientation data using counting circles and continuous weighting functions. *Tectonophysics*. Vol 121(2–4): p 207–223.
- Rouleau A, Gale J E, 1985.** Characterization of the fracture system at Stripa with emphasis on the ventilation drift: Technical information report no. 52. LBL-14875 SAC-52, A joint project of Swedish Nuclear Fuel Supply Co. Stockholm, Sweden and Lawrence Berkeley Laboratory, Earth Sciences Division, University of California, Berkeley, California, USA.
- Schaeben H, 1984.** A new cluster algorithm for orientation data. *Journal of the International Association for Mathematical Geology*. Vol 16(2): p 139–153.
- Scholz C H, Aviles C A, 1986.** The fractal geometry of faults and faulting, in *Earthquake source mechanics*, Das, S, Boatwright, J, and Scholz, C.H, Editors. 1986, American Geophysical Union: Washington, DC, United States. p 147–155.
- Shanley R J, Mahtab M A, 1976.** Delineation and analysis of clusters in orientation data. *Journal of the International Association for Mathematical Geology*. Vol 8(1): p 9–23.
- Starkey J, 1977.** The contouring of orientation data represented in spherical projection. *Canadian Journal of Earth Sciences = Journal Canadien des Sciences de la Terre*. Vol 14(2): p 268–277.
- Stråhle A, 2001.** Definition och beskrivning av parametrar för geologisk, geofysisk och bergmekanisk kartering av berg. SKB R-01-19, Svensk Kärnbränslehantering AB.
- Terzaghi R, 1965.** Sources of error in joint surveys. *Geotechnique*. Vol 15(3): p 287–304.
- Tukey J W, 1977.** *Exploratory data analysis*. 1977, Reading, MA, USA: Addison-Wesley. ISBN.
- Turcotte D L, 1986a.** Fractals and fragmentation. *JGR. Journal of Geophysical Research*. B. Vol 91(2): p 1921–1926.
- Turcotte D L, 1986b.** A fractal model for crustal deformation, in *Intraplate deformation; characteristics, processes and causes*, Johnson, B. and Bally Albert, W, Editors. 1986b, Elsevier: Amsterdam, Netherlands. p 261–269.
- Turcotte D L, 1990.** Implications of chaos, scale-invariance, and fractal statistics in geology, in *Geosphere fluctuations; short term instabilities in the Earth's system*, Cloetingh, S, Editor. 1990, Elsevier: Amsterdam, Netherlands. p 301–308.
- Watson G S, 1966.** The statistics of orientation data. *Journal of Geology*. Vol 74(5, Part 2): p 786–797.
- Whittaker E T, Robinson G, 1967.** Graduation, or the Smoothing of Data, in *The Calculus of Observations: A Treatise on Numerical Mathematics*. 1967, Dover: New York. p 285–316.
- Vollmer F W, 1989.** A triangular fabric plot with applications for structural analysis. *Eos (American Geophysical Union Transactions)*. Vol 70: p 463.

**Vollmer F W, 1993.** A modified Kamb method for contouring spherical orientation data, in Geological Society of America, 1993 annual meeting., Anonymous, Editor. 1993, Geological Society of America (GSA): Boulder, CO, United States. p 170–171.

**Vollmer F W, 1995.** C Program for automatic contouring of spherical orientation data using a modified Kamb method. *Computers & Geosciences*. Vol 21(1): p 31–49.

**Woodcock N H, 1977.** Specification of fabric shape using an eigenvalue method. *Geological Society of America Bulletin*. Vol 88: p 1231–1236.

**Woodcock N H, Naylor M A, 1983.** Randomness testing in three-dimensional orientation data. *Journal of Structural Geology*. Vol 5(5): p 539–548.

**Xiao-Hua Z, Sujuan G, 1997.** Confidence intervals for the lognormal mean. *Statistics in Medicine*. Vol 16: p 783–790.