The use of the Pareto distribution for fracture transmissivity assessment

Gunnar Gustafson · Å Fransson

Abstract For many applications data on the transmissivity distribution of individual fractures are necessary, i.e., discrete fracture network (DFN) modelling of groundwater flow and transport of solutes in fractured rock and design and performance of rock grouting. Using borehole data from the Äspö Hard Rock Laboratory, Sweden, it is shown that evaluated fracture transmissivities from three boreholes at Äspö can be well described by a Pareto or power-law distribution. Evaluated distribution parameters for the three boreholes are similar, which indicates that the Pareto distribution is a robust tool to assess three fracture transmissivity distributions. Using the evaluated distribution parameters random simulations of the original interval test data show that the approximate lognormal distributions of these are reproduced. This strengthens the credibility of the approach. It is shown how the distribution parameters can be assessed from incomplete data using the properties of the distribution. Finally, Pareto distribution transmissivities also imply Pareto distribution apertures of the fractures.

Résumé Pour beaucoup d'applications sont nécessaires des données sur la distribution des transmissivités des fractures individuelles; des exemples sont la modélisation de l'écoulement et du transport dans les roches fracturés ainsi que l'analyse de performance de la cimentation des roches. Les données des forages analysées dans le Laboratoire des roches dures d'Äsprö-Suède ont montré que la distribution des transmissivités des fractures estimée pour trois forages d'Äsprö est en accord avec les lois de Pareto ou de puissance. paramètres pour trois forages ont des valeurs semblables ce qu'indique que la loi de Pareto peut bien évaluer la distribution des transmissivités des fractures. Les simulations aléatoires réalisés sur l'intervalle original des données qui ont utilisé les paramètres estimés des distributions ont bien reproduit la distribution log-normale approximée. Ce résultat a consolidé la confiance dans l'approche faite. On a montré que les paramètres de la distribution peuvent être estimés à partir de données incomplètes si on utilise les propriétés de la distribution. Finalement, la distribution de Pareto pour les transmissivités implique la même loi pour la distribution de l'épaisseur des fractures.

Resumen Los datos sobre la distribución de la transmisividad en fracturas individuales, son necesarios para muchas aplicaciones, Ej. Modelación de redes de fracturamiento discreto (RFD) para flujo de agua subterránea y transporte de solutos en roca fracturada y para el diseño y ejecución de operaciones de cementación en roca. Mediante el uso del Laboratorio de Rocas Duras de Äspö en Suecia, se ha demostrado que las transmisividades por fractura evaluadas en tres perforaciones en Äspö pueden ser descritas bien por una distribución Pareto o Potencial. Los parámetros de distribución evaluados para las tres perforaciones son similares, lo cual indica que la distribución Pareto es una herramienta adecuada para evaluar tres distribuciones de transmisividad por fractura. Mediante el uso de simulaciones al azar, de los parámetros de distribución evaluados, a partir de los datos del intervalo de la prueba original, se muestra que las distribuciones lognormal así aproximadas de aquellos, pueden ser reproducidas. Esto refuerza la credibilidad del método. Se demuestra como los parámetros de distribución pueden ser evaluados a partir de datos incompletos, usando las propiedades de la distribución. Finalmente, las transmisividades con distribución Pareto, también implican aberturas de las fracturas con distribución Pareto.

Keywords Hydraulic properties · Groundwater hydraulics · Fractured rocks · Crystalline rock

Received: 11 March 2004 / Accepted: 3 January 2005
Published online: 10 June 2005

G. Gustafson (✉) · Å. Fransson
Department of GEO Engineering, Chalmers University of Technology,
S-416 80 Gothenburg, Sweden
e-mail: g2@geo.chalmers.se
Tel.: +46-31-772-2081
Fax: +46-31-772-2107

Introduction

For many applications data on the transmissivity distribution of individual fractures are necessary. The most obvious case is discrete fracture network (DFN) modelling of groundwater flow and transport of solutes in fractured rock (see a.o. Dershowitz et al. 1998). Other applications are design and performance of rock grouting (Fransson 2001a).

The transmissivity of a fracture is here, in analogy with the transmissivity of an aquifer, defined as the flux of

 DOI 10.1007/s10040-005-0440-y

groundwater parallel to the fracture plane through a unit width of the fracture and under unit gradient. The transmissivity of a borehole interval is in the context of this paper thus the sum of the transmissivities of the fractures penetrated by the borehole.

The transmissivities of fractures intersecting a borehole are not readily measured. The reason is the great range of values that typically can stretch over several orders of magnitude. The measurement resolution will thus not be good enough to identify the tightest fractures by normal measurement techniques. This is even more pronounced since a standard packer test normally straddles at least a few metres of the borehole, which makes the more pervious fractures to obscure the influence of the less pervious ones. Knowing transmissivities of hydraulic packer tests and the number of fractures in each test section an approximate transmissivity distribution can be derived (Osnes et al. 1988; Axelsson et al. 1990; Fransson 2002). Recently the development of the Posiva Flowlog (Rouhianen 1993) has made it possible to measure directly at least the transmissivities of the most conductive fractures.

The use of these methods is fairly complicated and there is a strong need for a method based on packer test data, that is simple to use, robust and based on few parameters, and that reproduces patterns and magnitudes of input data. It is believed that such a feasible method can be based on the Pareto statistical distribution. In the following, data will be used from three boreholes, KLX02, KLX01and KA2598A, in fractured granite at the Äspö Hard Rock laboratory (Stanfors et al. 1999), Sweden, to evaluate its use.

Borehole data

As an example, the cored borehole KLX 02 (Ludvigsson et al. 2002) between 206 and 341 m depth was chosen. For this interval there are sequential flow logging tests of 3-m intervals as well as detailed differential flow-logs of 0.5-m intervals taken overlapping with an 0.1 m displacement as well as cores where the fractures were mapped. A plot of the empirical cumulative distribution function (CDF) of the 3 m transmissivity measurements is shown in Fig. 1. The plot has a rather typical outlook. It is close to the shown lognormal distribution having the same statistics but it could also be argued that it is a combination of at least two distinct lognormal distributions. The frequencies of the number of fractures of the tested intervals are shown in Fig. 2. It can be noted that the correlation between the number of fractures per interval and the transmissivity or the log transmissivity is weak with correlation coefficients : $R(n_\mathrm{i}, T_\mathrm{i}) = 0.28$ and $R(n_\mathrm{i}, \log T_\mathrm{i}) = 0.52$. Here n_i and T_i are the number of fractures and measured transmissivity in the test interval, *i*.

The assessment of transmissivities to individual fractures is in principle not possible since each measurement in general straddles several fractures. The Posiva flowlog, however, in general, measures the major fracture in each measured interval, as will be explained later, since in a

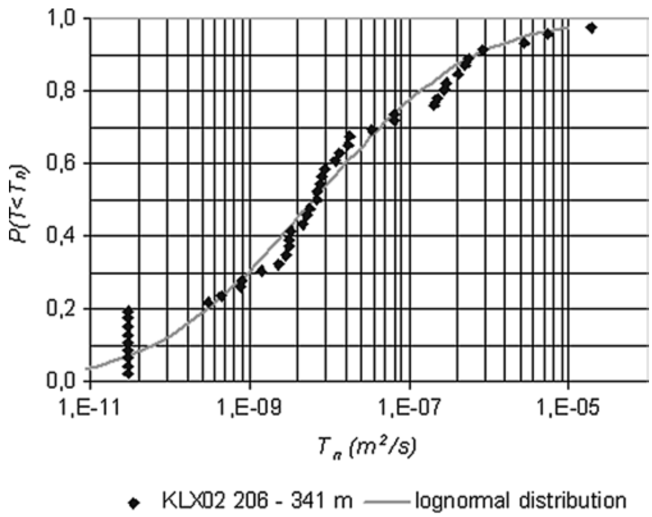

Fig. 1 Cumulative distribution plot of 3 m transmissivities for KLX02 206–341 m and the corresponding lognormal distribution, $m_{\log T} = -8.21$, $s_{\log T} = 1.59$

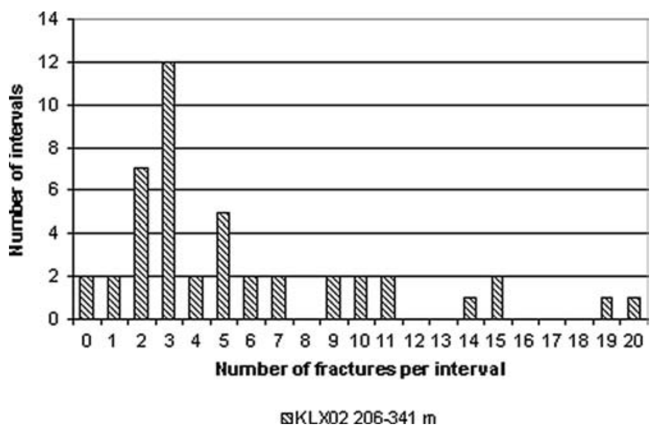

Fig. 2 Frequencies of 3-m intervals with different numbers of fractures in KLX 02 206–341 m. m = 5.5, s = 4.9

small sample the major fracture is most likely to dominate the transmissivity of the interval. This makes it possible, knowing the number of fractures, to construct a truncated distribution of the fracture transmissivities.

Several methods to derive fracture transmissivity distributions from fixed interval hydraulic tests have been devised (Osnes et al. 1988; Dershowitz et al. 1998; Fransson 2002). All of them assume that the fractures can be seen as independent sub-parallel features and that the sum of the fracture transmissivities adds up to the transmissivity of the tested section. The assumption of fractures being independent and not interconnected is a simplification, which follows the assumption of two-dimensional flow. The evaluated transmissivity value depends upon the tested length but is also influenced by the duration of the test. Early time transmissivities are likely to give a better representation of fractures found close to the borehole.

In this text Fransson's method is used since it does not assume a pre-chosen distribution function. A simple method, also devised by Fransson (2002), that gives points on the

 DOI 10.1007/s10040-005-0440-y

cumulative distribution curve is to use combinatorics to calculate the probability that all fractures in an interval is smaller than a preset value, T. Assuming that the probability of any fracture for having a transmissivity less than T is $P(T)$. The probability that all in an interval will be smaller is:

$$P_T = P(T)^{n_i} \tag{1}$$

Here n_i is the number of fractures in interval i. Again assuming that the dominating fracture determines the transmissivity of the interval, the expected number of intervals, I_T out of I, having a transmissivity smaller than T will be:

$$I_T \approx \sum_{n=1}^{N} P(T)^{n_i} \tag{2}$$

This equation can readily be solved for $P(T)$ by the Newton-Raphson iteration method using the ranked interval transmissivity and fracture frequency data. All three methods will be applied to the data from borehole KLX 02 to determine parameters of the Pareto distribution.

The Pareto distribution

The Pareto or Power-Law distribution has been used for ensembles where a few entries show very high values whereas the bulk shows low to extremely low values. An overview of the theory and use of the Pareto distribution is given by Adamic (2002). Examples of Pareto distributed parameters are annual personal income of individuals, magnitudes of earthquakes and number of hits at web sites. A related relation is Zipf's law that relates the frequency of occurrence of an event to its rank, originally applied to the frequency of use of words in English texts. The property of a few occurrences of large values together with a bulk of very small values is also typical for transmissivities of fractures in fractured crystalline rocks (Fransson 2001a).

Pareto's law can be given in terms of the CDF (Adamic 2002):

$$P(x) = P[X < x] = 1 - (m/x)^k, \quad m > 0,\ k > 0,\ x \geq m \tag{3}$$

Here m represents the smallest value in the sample and k is the Pareto distribution parameter. As a consequence the probability density function, PDF, will be:

$$p_X(x) = \mathrm{km}^k x^{-(k+1)} \tag{4}$$

Introducing a hypothetical minimum transmissivity $T_{\min}$ in Eq. (1) gives:

$$P(T_n) = P[T < T_n] = 1 - (T_{\min}/T_n)^k \tag{5}$$

In this case, a definition of the Pareto distribution based on the maximum fracture transmissivity value, $T_{\max}$, will be more practical to use, since normally very little is known of the transmissivities of the tightest fractures.

$$P(T_{\max}) = 1 - (T_{\min}/T_{\max})^k \tag{6}$$

Another way to assess the values of the CDF is to use a rank based plotting probability like Weibull's formula (see Chow 1964):

$$P(T_n) = \frac{n}{N+1} \tag{7}$$

Here T_n is the transmissivity with number n in size-sorted sample of total number N. Thus:

$$P(T_{\max}) = \frac{N}{N+1} \tag{8}$$

Equations (5) and (8) give:

$$T_{\min}^k = \frac{T_{\max}^k}{N+1} \tag{9}$$

Inserting Eqs. (9) in (6) gives:

$$P(T_n) = 1 - \frac{(T_{\max}/T_n)^k}{N+1} \tag{10}$$

Rearranging and taking the log of Eq. (10) gives:

$$\log[1 - P(T_n)] = \log\left[T_{\max}^k/(N+1)\right] - k\log(T_n) \tag{11}$$

The Pareto distribution is thus easily recognised in a log-log plot as a straight line with a slope $-k$, which makes it possible to determine the distribution coefficient readily. The $T_{\min}$-value can be assessed by the intersection of the line with the $1-P(T_n)-1$ line. $T_{\max}$ can then be calculated using Eq. (9) .

Evaluation of the parameters of the Pareto distribution

Figure 3 shows a plot of fracture transmissivities from KLX02 206–341 m evaluated by Fransson's nonparametric method, the Posiva flowlog and by combinatorics. Data from the different methods lie close to each other in the range where reasonably good field data can be obtained by hydraulic tests. For all data sets a power law trend-line has been fitted by the least square method giving the parameters k and $T_{\max}^k/(N+1)$.

The goodness of fit is measured by the calculated R^2 values. When fitting the function for the non-parametric evaluation data they were truncated at $T=10^{-10}$ m^2/s as evaluated values for the very low transmissivities are not very reliable, as pointed out by Fransson (2002).

 DOI 10.1007/s10040-005-0440-y

An interesting comparison can be made with other boreholes in the Äspö HRL. Figure 4 shows a comparison of Pareto distributions for the borehole KLX02 206–341 m, injection tests in 3-m intervals in KLX01 106–691 m and inflow measurements to 2-m intervals of a pump test of KA 2598A (Vidstrand 1999), all evaluated by combinatorics. All can be described well with the power-law regression lines. The evaluated parameters of the distributions are given in Table 1. As can be seen they are reasonably similar for all the boreholes.

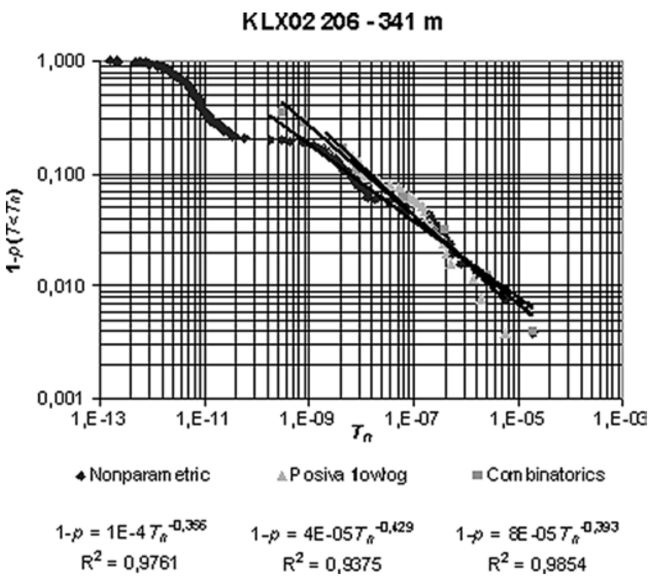

Fig. 3 Evaluation of the parameters of the Pareto distribution for fracture transmissivity data for KLX 02 206–341 m

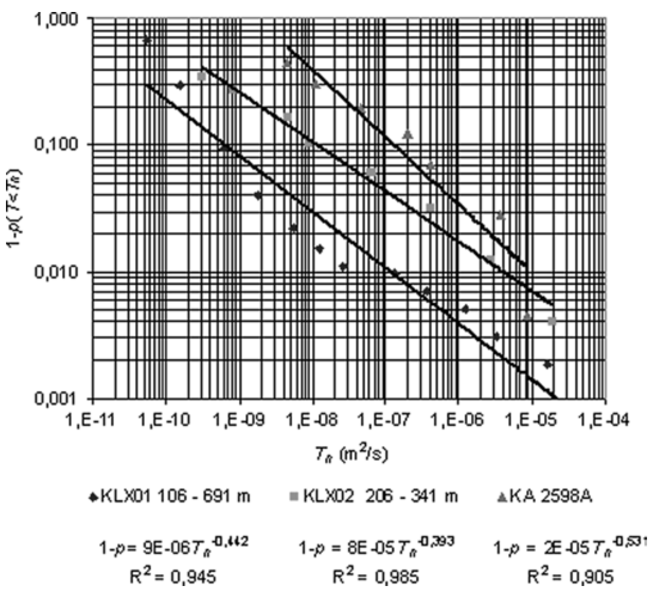

Fig. 4 Comparison of evaluated fracture transmissivities from three boreholes at Äspö HRL

Table 1 Parameters for the fracture transmissivity distributions for the boreholes KLX02 206–341 m, KLX01 106–691 m and KA 2598A at the Äspö HRL

Borehole	(m)	N	k	$T_{\max}$ (m^2/s)
KLX02	206–341	256	0.393	5.10E-05
KLX01	106–691	1637	0.442	7.19E-05
KA 2598A	1-90	315	0.531	7.21E-05

Some properties of the Pareto-Transmissivity distribution

T_n can be determined from Eqs. (7) and (10) to be:

$$T_n = T_{\max\,/}/[[1 - n/(N+1)] \cdot (N+1)]^{1/k} = T_{\max}/(N - n + 1)^{1/k} \quad (12)$$

or using the rank $r = N - n + 1$:

$$T_r = T_{\max}/r^{1/k} \quad (13)$$

The sum of all transmissivities, T_{tot}, will then be:

$$T_{\text{tot}} = T_{\max} \cdot \left[\frac{1}{1^{1/k}} + \frac{1}{2^{1/k}} + \frac{1}{3^{1/k}} \cdots\cdots\cdots + \frac{1}{N^{1/k}}\right] = T_{\max} \cdot S(k, N) \quad (14)$$

$$S(k, N) = \sum_{r=1}^{N} \frac{1}{r^{1/k}} \quad (15)$$

The infinite series has a finite sum for $k \leq 0{,}5$ where $S(1/2, \infty) = \pi^2/6$ (Råde and Westergren 1990). In general, only a few exact values for the infinite series can be calculated, i.e., $S(1/4, \infty) = \pi^4/90$ and $S(1/6, \infty) = \pi^6/945$ etc. However, the sum is readily calculated in a spreadsheet for any value of N and k and its magnitude can easily be read out from Fig. 5.

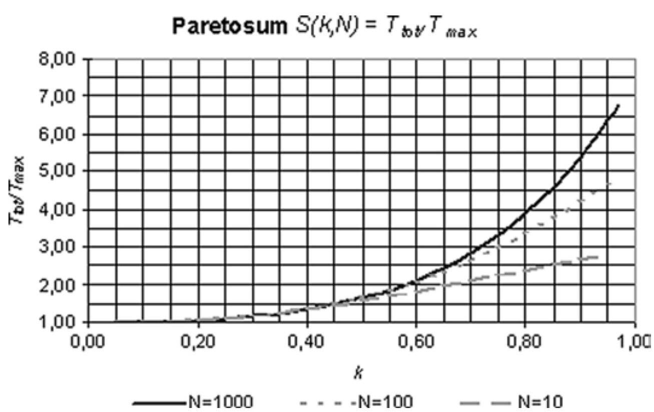

Fig. 5 $S(k, N)$ as a function of k and N

This gives some general indications on the relation between the most conductive fracture, T_{max} and the transmissivity of a tested section T_{tot}:

- If k and N are approximately known T_{max} can easily be estimated from T_{tot}.
- If the section contains few fractures T_{max} is always of the same order of magnitude as T_{tot}.
- If $k<0.5$, T_{max} is always of the same order of magnitude as T_{tot}.
- For very fractured sections where $k>0.5$ T_{max} is up to an order of magnitude smaller than T_{tot} but still makes up a significant portion of it.

Simulated hydraulic tests

A way to test the credibility of the Pareto distribution as a way to assess the fracture transmissivity distribution along a borehole is to simulate transmissivity values for hydraulic tests, i.e., intervals of a series of packer tests, and then compare the resulting distribution to patterns of real test sequences. Many authors (a.o. Gustafson and Krásny 1994) have pointed out that data from packer tests tend to be lognormal distributed and therefore a simple comparison to a lognormal distribution was made.

Using the Pareto distribution a random transmissivity value, T_n, for each fracture, n_i, in an interval, i, can be generated as:

$$T_{n_i} = T_{max} / (N \cdot \mathrm{rnd})^{(1/k)} \tag{16}$$

Here N is the total number of fractures and $\mathrm{rnd} \in (0, 1]$ is a random number. T_{max} is the transmissivity of the largest fracture. Figure 6 shows a cumulative distribution plot of packer interval transmissivities for intervals with the same numbers of fractures as the tested intervals of borehole KLX 02 206–341 m as shown in Fig. 2. The parameters for the Pareto distribution for this borehole are given in Table 1.

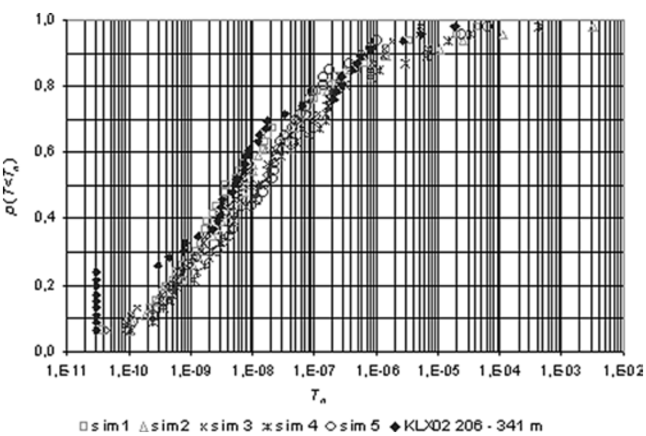

Fig. 6 Cumulative distribution plot of simulated transmissivities of test intervals with fracture frequencies as in borehole KLX02 206–341 m and Pareto distributed fracture transmissivities. Pareto parameters: k=0.393, T_{max}=5.1×10^{-5} m^2/s and N=256

As a comparison, data from borehole KLX02 are shown in the figure. Although the fit is not perfect, the CDFs are very similar and an assumption of lognormal distributed interval transmissivities cannot be rejected.

Fracture apertures

A general relation transmissivity and hydraulic aperture of a fracture is the so-called cubic law (Brown 1987). It can be directly derived from the laminar flow between parallel smooth plates. Knowing that fractures have no smooth surfaces, the hydraulic aperture can be seen as an equivalent measure of the hydraulic fracture width. It has also shown to be a realistic estimate of the effective aperture for grout penetration (Fransson 2001b). The hydraulic aperture of the fracture according to the cubic law will be:

$$b = \sqrt[3]{T \cdot \frac{12\mu_w}{\rho_w g}} = C \cdot T^{1/3} \tag{17}$$

Here ρ_w is the density of water and μ_w is the viscosity. C is a proportionality-constant. Inserted in Eq. (10) using Eq. (13) we obtain the aperture b_r of the fracture with rank r as:

$$b_r = C \cdot [T_{max}/r^{1/k}]^{1/3} = C \cdot T_{max}^{1/3}/(r^{1/k})^{1/3} \tag{18}$$

$$b_r = b_{max}/r^{1/3k} \tag{18b}$$

Here b_{max} is the hydraulic aperture of the largest fracture. The hydraulic apertures thus follow another Pareto-distribution with the parameter $3k$.

Conclusions

An analysis of field data from boreholes of the Äspö HRL have shown that the Pareto or power-law distribution is a feasible and robust tool to assess the transmissivity distribution of fractures penetrated by the borehole. The distribution accurately reproduces the common observation that a few conductive fractures give the major contribution to the borehole transmissivity and that there are many fractures with a very low transmissivity.

The parameters of the distribution can readily be evaluated from fixed-interval water pressure tests or detailed flow-logging. From the tested boreholes, however, one can suspect that the value of the distribution parameter, k, in the relatively sparsely fractured granitic rocks at Äspö has a rather narrow range. Given the value of k the other parameters of the distribution can be estimated from the borehole transmissivity and the fracture intensity. From the previous text, it is obvious that the order of magnitude of the value of T_{max} is very much determined by the total transmissivity

 DOI 10.1007/s10040-005-0440-y

of the tested section, i.e., the whole borehole. Principally N is known if if there is a core from the borehole mapped or if the fracture intensity can be estimated. It may be suspected, however, that there is a positive correlation between N and k since intuition says that the more fractures there are, the smaller will the most conductive fracture be as a portion of the totals.

Finally, the Pareto distribution should be used as an empirical tool until there is strong evidence that geological processes produce this kind of relationships.

Acknowledgment The authors thank Swedish Nuclear Fuel and Waste Management Co (SKB) for making the borehole data available and for economic support of the work.

Notation

b_{max}	The hydraulic aperture of the largest fracture (m)
b_r	The hydraulic aperture of the fracture with rank r (m)
C	Proportionality constant (−)
I	The number of tested intervals (−)
I_T	The number of intervals having a transmissivity smaller than T. (−)
i	Interval number (−)
k	Pareto distribution parameter (−)
m	Sample average, smallest value of a Pareto distributed value (x)
N	Total number of fractures in a tested borehole (−)
n	Order number of a size-sorted sample (−)
n_i	Number of fractures in the tested interval i (−)
$P(X)$	The probability that a member of a sample is smaller than X (−)
P_T	The probability that all fracture transmissivities of a tested interval are smaller than a pre-chosen value (−)
$p_X(x)$	The probability density of a Pareto distribution (−)
R	Correlation coefficient (−)
r	The rank of a value in an ordered sample (−)
rnd	A random number (−)
$S(k, N)$	Pareto sum (−)
s	Sample standard deviation (x)
T	Transmissivity (m^2/s)
T_i	Transmissivity of the tested interval i (m^2/s)
T_{max}	Transmissivity of the largest facture (m^2/s)
T_{tot}	Transmissivity of the whole borehole (m^2/s)
x	Dummy variable (x)
μ_w	The viscosity of water (Pa s)
ρ_w	The density of water (kg/m^3)

References

Adamic LA (2002) Zipf, Power-laws, and Pareto—a ranking tutorial. Internet Ecologies area, Xerox Palo Alto Research Center, Palo Alto, CA 94304 (http://ginger.hpl.hp.com/shl/papers/ranking.html)

Axelsson C-L, Jonsson E-K, Geier J, Dershowitz W (1990) Discrete fracture modelling. SKB Progress Report 25-89-21. Stockholm, Sweden

Brown SR (1987) Fluid flow through rock joints: the effect of surface roughness. J Geophys Res 92:1337–1347

Chow VT (1964) Handbook of applied hydrology. McGraw-Hill, New York

Dershowitz W, Lee G, Geier J, Foxford T, La Pointe P, Thomas A (1998) User Documentation, FracMan, Interactive discrete data analysis, geometric modelling, and exploration simulation, Version 2.6 Golder Associates Inc, Seattle

Fransson Å (2001a) Characterisation of fractured rock for grouting using hydrogeological methods. Department of Geology, Publ. A 97, Chalmers University of Technology, Göteborg, Sweden

Fransson Å (2001b) Characterisation of fractured rock for a grouting field experiment. Tunnel Undergr Space Technol 16:331–339

Fransson Å (2002) Nonparametric method for transmissivity distributions along boreholes. Ground Water 40(2):201–204

Gustafson G, Krásny J (1994) Crystalline rock aquifers: their occurrence, use and importance. Appl Hydrogeol 2:64–75

Ludvigsson J-E, Hansson K, Rouhiainen P (2002) Methodology study of Posiva difference flow meter in borehole KLX02 at laxemar. SKB Report R-01-52, Stockholm

Osnes JD, Winberg A, Andersson J-E (1988) Analysis of well test data—Application of probabilistic models to infer hydraulic properties of fractures. Topical Report RSI-0338. RE/SPEC INC. Rapid City, South Dakota

Rouhianen P (1993) TVO-Flowmeter, Report YJT-93-01, Nuclear Waste Commission of Finnish Nuclear Power Companies, Helsinki, Finland

Råde L, Westergren B (1990) BETA Mathematics Handbook, 2nd edn. Studentlitteratur, Lund

Stanfors R, Rhén I, Tullborg E-L, Wikberg P (1999) Overview of the geological and hydro-geological conditions of the Äspö hard rock laboratory site. Appl Geochem 14(7):819–834

Vidstrand P (1999) Hydrogeological scale effects in crystalline rocks. Licentiate thesis, Department of Geology. Chalmers University of Technology, Göteborg, Sweden

DOI 10.1007/s10040-005-0440-y