NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Model choice for fault size distribution

by

Hilde G. Borgos, Henning Omre and Chris Townsend

PREPRINT
STATISTICS NO. 3/2000

NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL `http://www.math.ntnu.no/preprint/statistics/2000/S3-2000.ps`

Hilde G. Borgos has homepage: `http://www.math.ntnu.no/~hgb`
E-mail: `hgb@stat.ntnu.no`
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology,
N-7491 Trondheim, Norway.

Model choice for fault size distribution

Hilde G. Borgos[1], Henning Omre[1] and Chris Townsend[2]

[1] Department of Mathematical Sciences, NTNU, Trondheim, Norway.
[2] Statoil Research Centre, Trondheim, Norway.

Abstract

Information about the number of faults and the fault size distribution is an important factor in reservoir modeling and fluid flow simulations. Inference is often made based on an assumption of a negative power law distribution of maximum displacements or fault widths. In this work it is shown how this assumption can be statistically validated, by comparing the negative power law distribution with an exponential distribution using Bayes factor. Inference about three dimensional fault size distributions are often based on observations obtained in lower dimension. Thus the relationship between the distribution in three and two dimensions must be known, for both the negative power law and the exponential distribution. Furthermore, observation errors, like bias, censoring and missing observations, are accounted for in the model choice problem.

1 Introduction

When tectonic forces, forces from overburdens, hydrostatic pressure etc. act on a subsurface rock, stress fields arise. Some of the stress can be released through the formation of faults and subsequent displacement of rock. A fracture in the rock forms, and if shear forces are present the rock in a volume on both sides of the fracture is displaced, creating a fault. See for instance Hatcher (1995) for an introduction to faults. The fault geometry is defined by the finite extension of the faulted volume. The displacement is typically at its maximum near the center of the faulted volume, descending to zero at the boundaries of the fault. Figure 1 gives a sketch of a vertical cross section through a fault, where the lines illustrate the displacement of rock.

When a stress field is present in a rock, a number of faults can be produced, resulting in a pattern of faults. The number of faults and the location, geometry and displacement of each fault in the fault pattern can have great influence on fluid flow in a reservoir. Faults

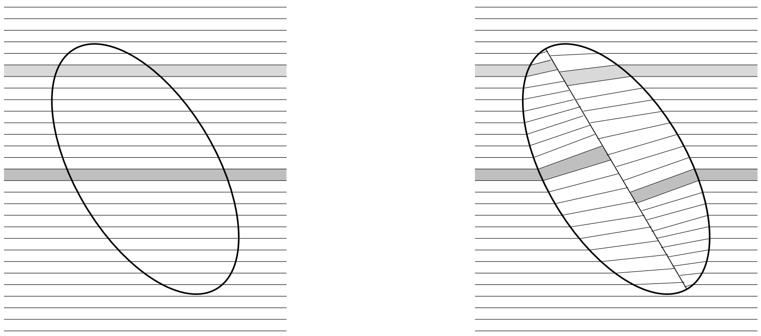


Figure 1: Vertical cross section through a fault. To the left is an illustration of the rock prior to faulting, with the fault geometry indicated. The right figure shows the rock after faulting. Only rock inside a bounded volume is displaced.

can act as either channels or barriers to fluid flow, and compartmentalizing of reservoirs can be related to the fault pattern, see for example Yielding et al. (1997) and Jones et al. (1998). Thus the effect of faults should be taken into account in reservoir modeling and fluid flow simulations. However, the fault pattern of a subsurface reservoir can not be observed exactly, but inference can be made based on available observations. For example, it is of interest to predict the number of faults in a reservoir and their size distribution.

Observations of faults often span only a few size scales. Seismic data include faults on a large scale, while core samples provide observations on a small scale and within a limited area. Extrapolation of fault size distributions obtained at seismic scales, combined with core observations if available, can be used to predict the distribution of fault sizes and the number of faults below seismic resolution, see for instance Badley et al. (1990), Childs et al. (1990), Walsh et al. (1991, 1994), Sassi et al. (1992), Gauthier and Lake (1993). Several uncertainties are related to the extrapolation of fault distributions, as discussed in the cited papers. For example, errors in the fault sampling may lead to great uncertainties in estimated model parameters. Furthermore, changes in geological properties beyond the scales of observations can not be predicted based on the observations, thus extrapolation of the estimated distribution can lead to unreliable results.

In this work the focus is on the size distribution of faults, where fault size is measured as either maximum displacement or width. The width represents the maximum horizontal extent of the fault. In order to make inference about the overall fault population based on available observations, an appropriate model should be chosen to represent the distribution of fault size, and the model parameters should be estimated. Model choice is important for predictive purposes. However, the model choice can also influence on the understanding of geological aspects related to the process of faulting.

Faults are objects embedded in a three dimensional space, while observations are often obtained in a lower dimension. Two dimensional observations are obtained from seismic maps or from surface outcrops of the rock, while one dimensional observations consist of measurements along a line. The size distribution observed in a lower dimension does not necessarily equal the original distribution, since large faults have a greater probability than small ones of being observed in lower dimensions and may thus be over-represented, see for example Heffer and Bevan (1990), Marrett and Allmendinger (1991). Hence, inference about the size distribution of faults in three dimensions can only be made if the relationship to the observations in lower dimensions is known. Another problem encountered when investigating the fault size distribution is observation errors, which should also be accounted for when choosing a proper model for the distribution.

If a fault size distribution is defined in three dimensions, corresponding unique distributions exist both in intersecting two dimensional planes and one dimensional lines. Under reasonable assumptions on fault geometry and spatial distribution of faults, these distributions can be obtained by simulations or forward calculations. Examples of forward calculations are found in the literature, mainly in the case of negative power law, or fractal, distributions of fault size. Heffer and Bevan (1990), Piggott (1997) and Berkowitz and Adler (1998) studied relationships between fault or fracture populations in three and two dimensions, assuming faults in three dimensions are shaped as circular discs. Marrett and Allmendinger (1991) presented a forward calculation performed without defining a specific model for the fault geometry, but using a general relationship $A \propto W^2$ where A is the surface area of the fault and W the fault width. A basic assumption in all the cited papers is that fault positions are independently, uniformly distributed in space.

Consider a hypothesis testing problem concerning fault size distributions in three dimensions. Inference can be made based on distributions in lower dimensions derived by forward calculations, and observations obtained in two or one dimensions can be used to decide which of two proposed distributions are most likely to describe the fault sizes in three dimensions. The conclusion only tells which of the two models is favored based on the observations, while there is still a possibility that none of the two distributions is the right one. For example, for a size distribution obtained in lower dimensions, the initial distribution in three dimensions need not be unique.

To the knowledge of the author, this work presents for the first time a statistical approach to dealing with ambiguities related to choosing a proper fault size distribution. Two possible distributions of maximum displacement of faults in three dimensions are considered: a fractal distribution and an exponential distribution. Corresponding distributions of fault width are obtained, and distributions of fault sizes observed in lower dimensions are derived under both models. A model choice problem is defined, comparing the two possible models for observations gathered in two dimensions. The uncertainty in the model choice is closely related to sampling errors, which are accounted for in the stochastic model. The conclusion of the model choice problem is based on Bayes factor, see Kass and Raftery (1995) for an overview. Bayes factor can not be analytically obtained for the stochastic model presented in this work, but is estimated using the Markov chain Monte Carlo (McMC) algorithm presented by Borgos and Omre (2000). Only if all observation errors are ignored can Bayes factor be calculated analytically.

The report is organized as follows. In Section 2 the fault model used in this work is presented. In Section 3 it is described how the distribution of fault size in lower dimensions can be derived from the distribution in three dimensions. This procedure is applied in Section 4, for both the fractal and the exponential model. Fault observations are discussed in Section 5, and the distribution of fault sizes and the number of faults, conditioned to the observations, is defined in a Bayesian framework. In Section 6 it is described how Bayes factor can be used to compare the two suggested models, and examples are provided in Sections 7 and 8.

2 Fault Representation

A single fault is a three dimensional object, and can be described through a number of fault characteristics. Throughout this work, fault geometry denotes the shape and extent of the faulted volume. Figure 1 shows a vertical cross section through a fault. A fault plane is a plane through the center of a fault, and the intersection of the faulted volume and the fault plane is called the fault surface. The fault surface forms a discontinuity in the rock, where rock on one side of the surface is displaced relative to the rock on the opposite side. Away from the fault surface, rock within the faulted volume is still continuous, and is only displaced relative to its original position. Consider two points infinitely close prior to faulting, but located on opposite sides of the fault surface. The distance between the two points after faulting, measured in the fault plane, is denoted the displacement. The displacement is typically at its maximum near the center of the fault surface, decreasing to zero along the fault surface and within the displaced volume (Walsh and Watterson, 1987). The angle between the fault plane and the horizontal plane is called the dip of the fault, and the direction of the line of intersection between the fault plane and the horizontal plane is commonly known as the strike. The strike is measured as the angle between the line and a reference direction, for example north.

In this study an individual fault is represented by its location, orientation, the fault geometry and the displacement within the faulted volume. The location of a fault is defined at its center point. To enable fault modeling, the complexity of fault geometries observed in nature is simplified, and the geometry is represented by a parametric model. The parameters include the volumetric extent and orientation of the fault. A simplified parametric model is also used for the displacement within the faulted volume.

2.1 Fault model

A fault pattern can be modeled as a marked point process where the points are the fault locations and the marks are sets of fault characteristics, see Chilès (1988), Munthe et al. (1993), Munthe et al. (1994), Stoyan and Stoyan (1994) and Stoyan et al. (1995). The marked point process is given as follows:

Location The location of a fault is defined as its center point, and is represented by the stochastic variable $\boldsymbol{\xi}$. In three dimensions the location is given by $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$.

Geometry The fault geometry $G = G(\Psi)$ is a parametric model representing the shape and extent of the faulted volume.

Parameterization The fault geometry is parameterized by a set of stochastic variables denoted Ψ. The parameterization includes the extents in three orthogonal directions: the horizontal extent W, the extent L in the dip direction and the extent R perpendicular to the fault plane. The angles of dip Θ and the strike Φ is also included, thus $\Psi = (W, L, R, \Theta, \Phi)$.

Maximum displacement The maximum displacement of the fault, obtained at the center point $\boldsymbol{\xi}$, is given by a stochastic variable D.

Displacement function For a fault with maximum displacement $D = d$, the displacement at any point u within the faulted volume is given by the displacement function $\delta(u; d)$. The function should satisfy the following conditions:

$$\begin{array}{rll} \text{(i)} & \delta(u;d) = d, & u = \boldsymbol{\xi} \\ \text{(ii)} & \delta(u;d) \leq d, & u \text{ inside } G(\Psi). \\ \text{(iii)} & \delta(u;d) \equiv 0, & u \text{ at the boundary or outside of } G(\Psi). \end{array}$$

Fault representation One fault can now be represented as a stochastic variable M, given as $M = (\boldsymbol{\xi}, \Psi, D, G(\Psi), \delta(\cdot))$, constituting a marked point in the marked point process.

In literature, the term size is used to represent both the width of a fault and the maximum displacement. In the first case size is a measure of the horizontal extent of the fault. In the latter case the maximum relative movement on opposite sides of the fault surface is measured. Fault width and maximum displacement are related, see Watterson (1986), and throughout this work the word size is used as a general expression for either quantity.

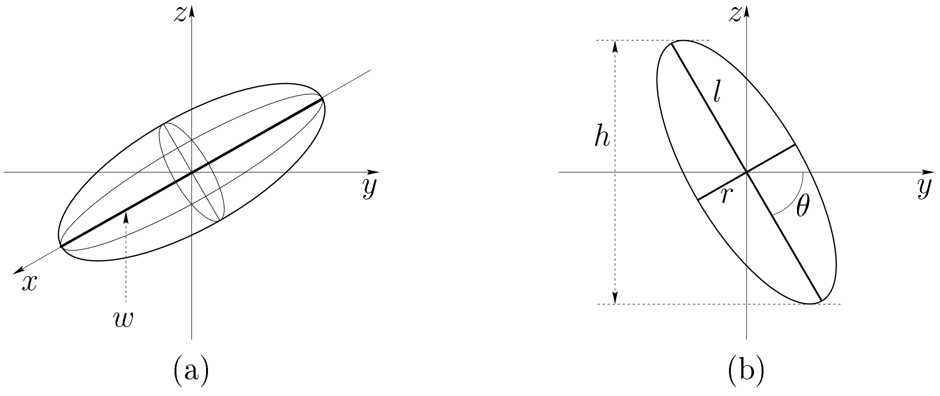


Figure 2: (a) Fault ellipsoid in three dimensions. The length of the horizontal major axis is $w/2$, measured from the center point to the ellipsoid boundary. The lengths of the two other axes are $l/2$ and $r/2$, and are illustrated in figure (b), which shows a vertical cross section of the fault ellipsoid perpendicular to fault strike. The dip of the fault is θ and the vertical extent is h.

2.2 Fault geometry

The fault geometry $G(\Psi)$ in three dimensions is represented by an ellipsoid (Walsh and Watterson, 1988, 1989), as illustrated in Figure 2a. A vertical cross section of the fault ellipsoid is shown in Figure 2b, where the intersection of the ellipsoid with a plane results in an ellipse. The ellipsoid is parameterized by the axis lengths $W/2$, $L/2$ and $R/2$ respectively, measured from the center point to the boundary of the ellipsoid, and by the dip Θ and strike Φ, see Figure 2. The strike is of no concern throughout this work, and all faults are assigned equal strike $\Phi = 0$ without loss of generality. For a fault located at the origin, the geometry $G(\psi)$ is then the ellipsoid described by the set of points in $\mathbb{R}^3$:

$$G(\psi) = G(w, l, r, \theta, 0) = \left\{ (x, y, z)\,;\, \frac{x^2}{w^2} + \frac{(y \sin\theta - z\cos\theta)^2}{l^2} + \frac{(y\cos\theta + z\sin\theta)^2}{r^2} = \frac{1}{4} \right\}.$$

When faulting occurs, only the volume inside of the ellipsoid is deformed by the fault, while the volume outside the ellipsoid is unaltered. The displacement inside the fault ellipsoid is given by the maximum displacement D and the displacement function $\delta(\cdot)$. Deterministic relationships between D, W, L and R are assumed:

$$W = (Dc_1)^{1/\tau}, \quad L = Wc_2^{-1}, \quad R = c_3, \tag{1}$$

where c_1, c_2 and c_3 and τ are constants, see Watterson (1986), Walsh and Watterson (1987, 1988, 1989), Cowie and Scholz (1992a, b), Gauthier and Lake (1993), Dawers et al. (1993). Several case studies have found values of the parameter τ in the range $[1, 2]$, and although the relationship $W \propto D^{1/\tau}$ is widely accepted, there has been some disagreement about the actual value of τ. See Cowie (1998) for a discussion and for further references.

The vertical extent, or height, of the ellipsoid is given by a function $h(\cdot)$ depending on the lengths of the axes and the dip, and is illustrated in Figure 2b. Since there is a deterministic relationship between D, W, L and R, the height can be expressed as a function of either of these, for example

$$\begin{aligned} h(w,\theta) &= \left((wc_2^{-1})^2 \sin^2\theta + c_3^2 \cos^2\theta\right)^{1/2} \approx wc_2^{-1}\sin\theta & (2)\\ h(d,\theta) &= \left(((dc_1)^{1/\tau}c_2^{-1})^2 \sin^2\theta + c_3^2 \cos^2\theta\right)^{1/2} \approx (dc_1)^{1/\tau}c_2^{-1}\sin\theta. & (3)\end{aligned}$$

The approximations of the height function correspond to the height of the axis of the fault ellipsoid in the dip direction. The length of the ellipsoid in the dip direction is L and the dip is Θ, resulting in a vertical extent of $L\sin\Theta$. Combining this with the expressions in (1) gives the approximations in Expressions (2) and (3). The approximations of the heights are used in later sections due to their simple analytical forms. A discussion of the approximations and the error terms is given in Appendix A.

Consider a fault of width $W = w$, measured through the center of the ellipsoid. At a vertical distance $uh(w,\theta)/2$ from the center, where $u \in [-1,1]$, the horizontal extent of the ellipsoid is

$$w'(u;w) = w\sqrt{1-u^2}, \qquad (4)$$

see Figure 3. Thus, if a horizontal plane intersects a fault of width $W = w$ at a vertical distance $uh(w,\theta)/2$ from the center of the fault, the width of the fault ellipse in this plane is $w'(u;w)$.

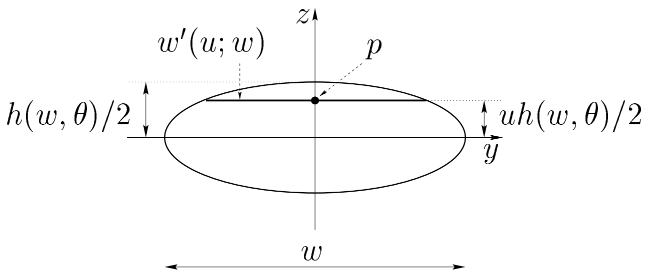


Figure 3: Vertical projection of the fault ellipsoid parallel to fault strike. At a vertical distance $uh(w,\theta)/2$ from the center of the fault ellipsoid, the horizontal extent is given by $w'(u;w)$ in Expression (4). At the point p the displacement is given by $\delta(u;d)$ in Expression (5).

2.3 Displacement function

The displacement function should satisfy the conditions (i)-(iii) given in the beginning of Section 2. In this work, a general displacement function of the form

$$\delta(u;d) = (1-|u|)^\gamma d, \qquad -1 \le u \le 1,\ \gamma > 0, \qquad (5)$$

is used. The parameter γ controls the shape of the displacement function along the axis in the dip direction, and some examples are illustrated in Figure 4. Throughout this work, γ is assumed to be a constant. See Walsh and Watterson (1987), Barnett et al. (1987), Dawers et al. (1993), and Dawers and Anders (1995) for discussions on displacement functions. If a horizontal plane intersects a fault with maximum displacement $D = d$ at a vertical distance $uh(d, \theta)/2$ from the center of the fault, see Figure 3, the maximum displacement of the fault in this plane is $\delta(u; d)$.

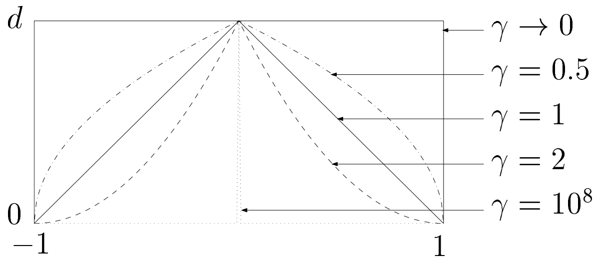


Figure 4: The displacement function $\delta(u; d)$ in Expression (5), with different values of γ.

3 Relationship Between Fault Distributions in One, Two and Three Dimensions

If a stochastic model for the fault geometry is well specified and a distribution of fault size in three dimensions is defined, the distribution in two dimensions can be found by forward calculations. The calculations may lead to expressions that are not analytically tractable, in which case numerical approximations can be used.

Consider faults spatially distributed within a three dimensional box $\mathcal{B}$ of vertical extent T_v and horizontal extent $T_{h1} \times T_{h2}$. The box is given by the set $\mathcal{B} = [0, T_v] \times [0, T_{h1}] \times [0, T_{h2}]$. If a two dimensional horizontal plane intersects $\mathcal{B}$ at an arbitrary depth $t \in [0, T_v]$, a number of faults inside $\mathcal{B}$ are intersected by the plane. The geometry of the intersection between the plane and an individual fault is a two dimensional ellipse, and within this intersection area the displacement is at a maximum at the center point of the ellipse. The stochastic variable X is now used as a common term for the size of a fault measured in three dimensions, being either maximum displacement or width. The corresponding size measured in two dimensions is denoted Y. The probability density functions (pdf) of the variables are denoted $f(x)$ and $f(y)$ respectively. The pdf $f(x)$ in three dimensions is assumed to be known, and the aim is to derive the corresponding pdf $f(y)$ in two dimensions.

Large faults or faults near the boundaries of the box $\mathcal{B}$ may intersect the edges of $\mathcal{B}$ and are thus not completely contained in the box. This is not accounted for in the calculations below. The probability of a horizontal plane intersecting a fault depends on the vertical thickness T_v. The actual value of T_v is of no interest in the calculations, and for simplicity T_v is assumed to exceed the vertical fault extents.

The variable Y is defined only for faults intersected by the plane, and an indicator variable I_2 of this event is introduced. The discrete stochastic variable I_2 is assigned the value 1 if a fault is intersected by the two dimensional plane, and thus is observed, or 0 otherwise:

$$I_2 = \begin{cases} 1, & \text{if the fault is intersected by the horizontal plane} \\ 0, & \text{else.} \end{cases}$$

The probability of intersection is proportional to the vertical extent of the fault, see Figure 2b, thus larger faults have greater probability of being intersected by the plane. The vertical extent, and hence also the probability of an intersection, are independent of the fault strike. The fault height $h(X, \Theta)$ is given in Expressions (2) and (3), and the probability of the fault intersecting the plane, conditioned on the size $X = x$ and the dip $\Theta = \theta$, is

$$\text{Prob}(I_2 = 1|X = x, \Theta = \theta) = f_{I_2}(1|x, \theta) = T_v^{-1} h(x, \theta).$$

The probability of the two dimensional plane intersecting an arbitrary fault of unknown size is then found by solving the integral

$$\begin{aligned} f_{I_2}(1) &= \iint f_{I_2}(1|x, \theta) f(x, \theta)\, dx\, d\theta \\ &= T_v^{-1} \iint h(x, \theta) f(x, \theta)\, dx\, d\theta \;=\; T_v^{-1} \mathrm{E}\left[h(X, \Theta)\right]. \end{aligned} \tag{6}$$

The pdf of the size X of a fault intersected by the two dimensional plane is not the same as the original pdf $f(x)$, since the larger faults are over-represented in the two dimensional sample. The new pdf can be found based on the joint pdf of X and Θ for an intersected fault, denoted $f(x, \theta|1)$, which is given by

$$f(x, \theta|1) = \frac{f_{I_2}(1|x, \theta) f(x, \theta)}{f_{I_2}(1)} = \frac{h(x, \theta) f(x, \theta)}{\mathrm{E}\left[h(X, \Theta)\right]}.$$

The fault height is assumed to be on the form $h(x, \theta) = h_1(x) h_2(\theta)$. This is fulfilled by the approximations in Expressions (2) and (3), where the height is given as $h(x, \theta) = cx^a \sin\theta$. Furthermore, X and Θ are assumed to be independent, giving $f(x, \theta) = f(x) f(\theta)$ and $\mathrm{E}\left[h(X, \Theta)\right] = \mathrm{E}\left[h_1(X)\right] \mathrm{E}\left[h_2(\Theta)\right]$. Under these assumptions, the marginal pdf of X of an intersected fault is found to be

$$f(x|1) = \int f(x, \theta|1)\, d\theta = \int \frac{h_1(x) h_2(\theta) f(x) f(\theta)}{\mathrm{E}\left[h_1(X)\right] \mathrm{E}\left[h_2(\Theta)\right]}\, d\theta = \frac{h_1(x) f(x)}{\mathrm{E}\left[h_1(X)\right]}, \tag{7}$$

which is independent of the dip.

The fault size X represents the maximum displacement D or the horizontal extent W. The displacement is at its maximum at the center of the fault, and decreases to zero at the border. The horizontal extent of the ellipsoid is also at its maximum at the center of the fault, and decreases to zero on the top an bottom of the faulted volume. Thus, if the horizontal plane intersects the fault at an arbitrary level away from the center, the observed displacement or horizontal extent at this level are less than D and W respectively. Thus the stochastic variable Y representing the size of the fault observed in the horizontal plane, is equal to or less than X. The variable Y is equal to X only if the plane intersects the fault through its center.

The size Y of a fault observed in an intersecting horizontal plane can be expressed as a function of the size X in three dimensions and the relative vertical distance U from the center of the fault to the horizontal plane, see Expressions (4) and (5). Given that a variable of size X is intersected by the two dimensional plane, the pdf $f(y|x,1)$ is found by a transformation $y = g(u;x)$, where the variable U has a uniform pdf $f(u|x,1)$. The unconditional pdf of Y is found by:

$$\begin{aligned} f(y) &= f(y|1) = \int_y^\infty f(x,y|1)\,dx = \int_y^\infty f(y|x,1)f(x|1)\,dx \\ &= \mathrm{E}\left[h_1(X)\right]^{-1} \int_y^\infty f(y|x,1)h_1(x)f(x)\,dx. \end{aligned} \tag{8}$$

The conditioning on $I_2 = 1$ in the pdf $f(y)$ is omitted, since the variable Y is defined only for intersected faults. Thus $I_2 = 1$ for all faults where the size Y is measured.

The procedure described above can be used to derive the distribution in one dimension, based on the distribution in two dimensions. Some of the faults in the two dimensional plane are then intersected by a line across the plane. In one dimension a fault is only observed as a point on a line, so the fault has no defined width. The size of the fault is therefore represented by the local displacement at the observed point.

Expression (8) gives the forward relationship from the distribution of fault size in three to two dimensions. Previous stereological analysis of ellipsoid or disc shaped faults obtain similar results, and have also focused on the problem of inverting the integral equation (8) in order to derive $f(x)$ based on a known pdf $f(y)$. See for instance Oakeshott and Edwards (1992) and further references therein, Piggott (1997) and Berkowitz and Adler (1998). As pointed out by Oakeshott and Edwards (1992), the inverse problem need not have a unique solution, but may be under- or over-determined. Furthermore, the inverse problem can only be solved analytically in special cases, but must in general be solved numerically. Berkowitz and Adler (1998) present an algorithm for solving the inverse problem for general, non-parametric, $f(x)$. In the current work, the inverse problem is not explicitly solved. The forward calculation in Expression (8) is performed under specified parametric models, and for the distributions $f(y)$ derived in two dimensions at least one solution of the inverse problem exists: the initially specified pdf $f(x)$. In most of the calculations below the distribution class from three dimensions is preserved in two dimensions, and there are reasons to believe that even if there exist other solutions to the inverse problem, the initial pdf $f(x)$ specified in three dimensions is the most plausible one. The results of forward calculations presented below also provide a one-to-one correspondence between the model parameters in two and three dimensions, thus the parameters of the model in three dimensions can be estimated based on observations from two dimensions.

4 Fault Size Distributions

The focus of this work is on fault size distributions. The distributions of interest are the distribution of the displacement D in three, two and one dimensions, and the horizontal extent W in three and two dimensions. First a distribution of D in three dimensions is specified. Through the deterministic relationship between D and W given in Expression (1), the distribution of W in three dimensions can be found based on the distribution of D. By solving the integral (8) for $X = D$ and $X = W$, the distributions of D and W in two and one dimensions can be found. The pdfs of the displacement in three, two and one dimensions are denoted

$$f_3(d), \quad f_2(d) \quad \text{and} \quad f_1(d)$$

respectively, and the pdfs of the horizontal extent, or width, in three and two dimensions are denoted

$$f_3(w) \quad \text{and} \quad f_2(w).$$

Before the fault size distributions are studied, a short discussion on the spatial distribution of faults is given in Section 4.1. In Section 4.2 the cumulative distribution of fault size is discussed in general, relating fault size distributions to common graphical presentations of observed fault sizes. Two possible models for the distributions of fault size in three, two and one dimensions are presented in Sections 4.3 and 4.4, a fractal model and an exponential model respectively.

4.1 Spatial distribution of fault locations

Locations of fault centers $\boldsymbol{\xi}$ are assumed to form a stationary Poisson point process of constant intensity $\lambda(x) = \lambda$, see Stoyan and Stoyan (1994), Stoyan et al. (1995). Conditioned on the number of points, n, the center points $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ located within the box $\mathcal{B}$ are independent uniformly distributed in $\mathcal{B}$. The locations of the fault centers are assumed to be independent of the marks of all faults, and the marks of different faults are assumed to be mutually independent. Thus there are no clustering or repulsion of faults, and the fault size is independent of the fault location. When a two dimensional plane intersects $\mathcal{B}$ at an arbitrary depth, the expected number of faults to be intersected by the plane is independent of the position $t \in [0, T_v]$ of the plane. Notice that for a fault with center point outside of the box $\mathcal{B}$, parts of the fault ellipsoid may still intersect $\mathcal{B}$. Thus also faults with center points in a volume surrounding $\mathcal{B}$ have influence on the number of faults in $\mathcal{B}$.

If the distance between the center points of two faults is small, the fault ellipsoids will overlap. In nature, the younger fault may terminate in the older fault, and the resulting ellipsoid should be truncated. The marked point process does not take any truncation into account. This can lead to a slight sampling bias in two dimensions, since the extents of faults that should be truncated are slightly too large in the specified model.

4.2 Cumulative fault size distribution

Consider the stochastic variable X representing the size of a fault, with pdf $f(x)$. The function $S(x)$ is defined in the following way:

$$S(x) = \text{Prob}(X \geq x) = \int_x^\infty f(u)\, du, \tag{9}$$

hence $S(x) = 1 - F(x)$ where $F(x)$ is the cumulative distribution function of X. If the number of observed faults is n, with a corresponding set of observations x_i, $i = 1, 2, \ldots, n$, an empirical estimate of the function $S(x)$ is given by

$$\hat{S}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}[x_i \geq x], \tag{10}$$

where $\mathbf{1}[\cdot]$ is the indicator function. The estimate of $S(x)$ is thus given as the relative number of faults with size x_i equal to or larger than x. Observations of fault sizes are often presented on a bilogarithmic plot, hereafter referred to as a log-log plot, giving empirical plots of $\log \hat{S}(x)$ versus $\log x$. An example of a log-log plot of observed displacements is shown in Figure 5. Instead of plotting the relative number of faults on log-log scale, the absolute number is often plotted, thus giving an estimate $n\hat{S}(x)$. In some cases the plots are normalized to give the number of faults per km^2 in two dimensions or per km in one dimension.

4.3 Fractal distributions

The term fractal was introduced by Mandelbrot (1983), and an introduction to fractals is also given by Feder (1988). The distribution of the stochastic variable X, representing

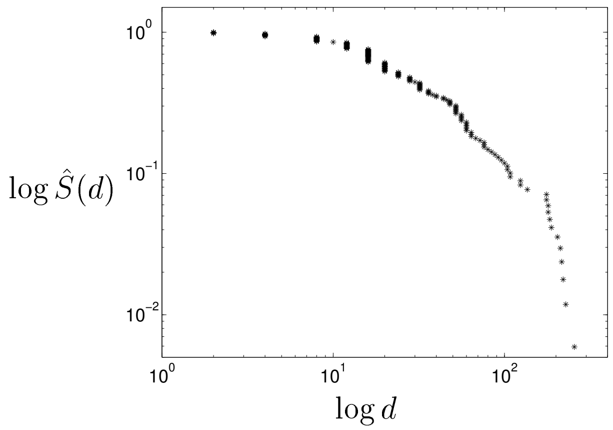


Figure 5: Log-log plot of relative number $\hat{S}(d)$ of faults with observed displacements larger than d, plotted at observations points d_i, $i = 1, \ldots, n$.

the maximum displacement or width of a fault, is called fractal if it follows a scaling relationship:

$$f(\lambda x) = \lambda^H f(x),$$

where H is the scaling exponent, see Feder (1988). An example of a scaling distribution is the negative power law distribution

$$S(x) = Cx^{-\beta}, \qquad \beta > 0, \tag{11}$$

where C is a positive constant and S(x) is defined in Expression (9). The parameter β is called the fractal dimension or the power law exponent. Negative power law distributions are frequently used in geology to model distributions of fault displacements and widths, see for example Childs et al. (1990), Heffer and Bevan (1990), Walsh et al. (1991), Marrett and Allmendinger (1992), Gauthier and Lake (1993) and Odling (1997). See also JSG (1996, Vol. 18).

The function $S(x)$ is often estimated as the relative number of observations larger than x, $\hat{S}(x)$, see Expression (10). If X follows the negative power law (11) and is observed exactly, a plot of $\hat{S}(x)$ versus x should produce a straight line of slope $-\beta$ on a log-log plot.

The relative number $\hat{S}(x)$ should be bounded above by 1. According to the negative power law (11) however, as $x \to 0$, the function $S(x)$ increases unlimited: $S(x) \to \infty$. Allowing $x = 0$, no constant C can be chosen to satisfy the upper bound condition. Thus it is not possible to define a negative power law distribution for all $X \geq 0$, but a lower limit $x_0 > 0$ for X must be chosen. Above this lower bound the distribution of X is fractal, while below the limit the distribution is left unspecified. Thus the need of a lower bound results in a model which is not able to reflect the true distribution on all scales. The choice of x_0 can be related to the geological properties of the rock, and a possible lower limit is the grain size of the minerals. In practice, an artificial lower bound is often used, for example the limit of resolution for a set of observations.

The pdf of a stochastic variable X following the negative power law distribution (11) is of the form

$$f(x; x_0, \beta) = \frac{\beta x_0^\beta}{x^{\beta+1}}, \qquad x \geq x_0,\ x_0 > 0,\ \beta > 0, \tag{12}$$

which is the pdf of the Pareto distribution (Johnson et al., 1994). The lower limit $x_0 > 0$, discussed above, ensures that the integral of the pdf is finite. The pdf is plotted in Figure 6a, while Figure 6b shows a log-log plot of the function $S(x)$, given as

$$S(x) = \left(\frac{x_0}{x}\right)^\beta, \qquad x \geq x_0. \tag{13}$$

Suppose a lower limit x_{03} is used in a negative power law distribution of the size of faults in three dimensions. When the corresponding distribution of fault size in two dimensions is studied, only faults of size above a lower limit $x_{02} \geq x_{03}$ in two dimensions can be considered. If a fault observed in two dimensions has a size smaller than x_{03}, it is uncertain if the size of the fault in three dimensions is smaller or larger than x_{03}, and thus the distribution of the fault size in three dimensions is unknown. By the same reasoning a lower limit x_{01} in one dimension must be defined, satisfying $x_{01} \geq x_{02}$.

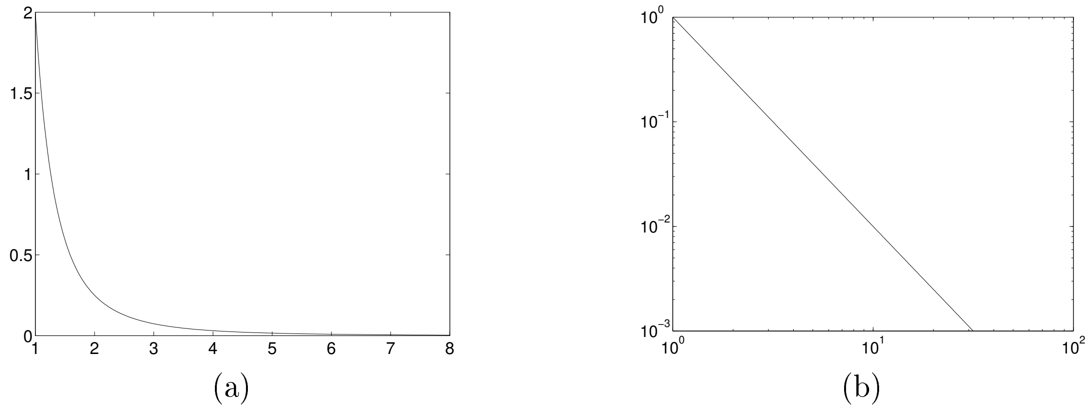


Figure 6: The Pareto distribution with parameters $\alpha = 1.0$ and $\beta = 2.0$. Figure (a) shows the pdf $f(x)$, Expression (12), and figure (b) the function $S(x)$, Expression (13), on log-log scale.

4.3.1 Displacement

The pdf of the displacement in three dimensions is given by

$$f_3(d) = \frac{\beta_3 d_{03}^{\beta_3}}{d^{\beta_3+1}}, \qquad d \geq d_{03}, \tag{14}$$

which is of the form given in Expression (12), where the lower bound and power law exponent are denoted d_{03} and β_3 respectively. Let D' denote the displacement obtained in two dimensions. The procedure described in Section 3 is used, with the approximation of the height of the fault ellipsoid given in Expression (3). Then, for $D' \geq d_{03}$ the unconditional pdf of D' is found to be

$$f_{D'}(d') = \frac{\beta_3 - 1/\tau}{\gamma(\beta_3 - 1/\tau) + 1} \cdot \frac{d_{03}^{\beta_3 - 1/\tau}}{(d')^{\beta_3 - 1/\tau + 1}}, \tag{15}$$

where γ is the parameter involved in the displacement function (5). The calculations are given in Appendix B. As pointed out above, a lower limit $d_{02} \geq d_{03}$ should be used for the displacement D' obtained in two dimensions. A distribution with this lower bound is derived by conditioning on $D' \geq d_{02}$. The conditional pdf can be found from Expression (15), without knowing the pdf for $D' < d_{03}$. The two parameters β_3 and τ only appear in the combination $\beta_3 - 1/\tau$ in Expression (15) and in the corresponding conditional pdf. Thus the pdf in two dimensions can be expressed as a two-parameter distributions with parameters $\beta_2 = \beta_3 - 1/\tau$ and d_{02}, and is found to be

$$f_2(d) = f_{D'}(d | d \geq d_{02}) = \frac{\beta_2 d_{02}^{\beta_2}}{d^{\beta_2 + 1}}, \qquad d \geq d_{02}. \tag{16}$$

This pdf is of the form given in Expression (12), and the distribution of the displacement is fractal also in two dimensions.

The distribution of maximum displacement in one dimension can only be found if the relationship between D and W in two dimensions is known. This relationship is discussed below, after obtaining the distribution of fault width in two dimensions. Under the assumption that the relationship in three dimensions, Expression (1), also holds in two dimensions, the distribution in one dimension is found by the same procedure as above. Assuming independence between fault strike and displacement, the pdf of D in one dimension is found as

$$f_1(d) = \frac{\beta_1 d_{01}^{\beta_1}}{d^{\beta_1 + 1}}, \qquad d \geq d_{01}, \tag{17}$$

where $\beta_1 = \beta_3 - 2/\tau$ and $d_{01} \geq d_{02}$. A negative power law is obtained also in one dimension. This result is only valid if $\beta_3 > 2/\tau$. If $\beta_3 \leq 2/\tau$, non-tractable expressions arise in the calculations.

4.3.2 Width

Expression (1) gives the relationship $W = (Dc_1)^{1/\tau}$ between fault width and displacement. Thus the pdf of W in three dimensions can be found from the pdf (14), by transformation of D, giving the fractal distribution

$$f_3(w) = \frac{\mu_3 w_{03}^{\mu_3}}{w^{\mu_3+1}}, \qquad w \geq w_{03}, \tag{18}$$

where $w_{03} = (d_{03}c_1)^{1/\tau}$ is the lower limit of W and $\mu_3 = \tau\beta_3$. Let W' denote the width obtained in two dimensions. The procedure described in Section 3, with the approximation (2) of the fault heights, gives the following pdf of W' for $W' \geq w_{03}$:

$$f_{W'}(w') = \frac{\Gamma((\mu_3+1)/2)}{\Gamma((\mu_3+2)/2)} \cdot \frac{\sqrt{\pi}}{2} \cdot \frac{(\mu_3-1)w_{03}^{\mu_3-1}}{(w')^{\mu_3}}. \tag{19}$$

The calculations are given in Appendix B. As for the displacement, a lower limit w_{02} in two dimensions should be defined for the width, satisfying $w_{02} \geq w_{03}$. The pdf of the width in two dimensions, now denoted W, is found from the pdf (19), conditioning on $W \geq w_{02}$:

$$f_2(w) = f_{W'}(w|w \geq w_{03}) = \frac{(\mu_3-1)w_{02}^{\mu_3-1}}{w^{\mu_3}} = \frac{\mu_2 w_{02}^{\mu_2}}{w^{\mu_2+1}}, \qquad w \geq w_{02}, \tag{20}$$

where $\mu_2 = \mu_3 - 1 = \tau\beta_3 - 1$. This pdf is of the form given in Expression (12), and thus also the width distribution is fractal in two dimensions.

As pointed out in Section 4.3.1, the relationship between W and D in two dimensions is needed to obtain $f_1(d)$ from $f_2(d)$. In three dimensions the relationship $W = (Dc_1)^{1/\tau}$ applies. The relationship between width W' and maximum displacement D' of the ellipse observed in an intersection plane is obtained by combining Expressions (4) and (5). If the plane intersects the fault ellipsoid at relative distance $Uh(d,\theta)/2$ from the center point, where U is uniformly distributed on $[-1,1]$, the relationship is given as

$$W' = \sqrt{1-U^2}(1-|U|)^{-\gamma/\tau}(D'c_1)^{1/\tau}.$$

If the mean value with respect to U is calculated, the relationship $W' = (D'c_1')^{1/\tau}$ is obtained, assuming the expression is integrable. This relationship is consistent with the transformation

$$f_2(d) = f_2(w(d)) \cdot |w'(d)|,$$

using $f_2(d)$ from Expression (16) and $f_2(w)$ from Expression (20), and is used above when the pdf of the displacement in one dimension is found.

4.4 Exponential distributions

The fractal distribution is widely used to model the size distribution of faults. In some cases however, fault sizes are observed to be better described by an exponential distribution. For example, exponential distribution of fault size are observed on mid-ocean ridges, see Cowie et al. (1994), and transformations from fractal distribution at early stages of deformation into exponential distributions are observed in analog experiments of fault growth, see Spryopoulos et al. (1999) and Ackermann et al. (1999). The pdf of the exponential distribution is

$$f(x;\lambda) = \lambda \exp(-\lambda x), \qquad x \geq 0,\ \lambda > 0, \tag{21}$$

and Figure 7a shows a plot of the function. For this distribution there is no need to specify a lower limit for x, as was the case for the fractal distribution. The function $S(x)$, defined in Expression (9), is

$$S(x) = \exp(-\lambda x), \qquad x \geq 0, \tag{22}$$

and is plotted on log-log scale in Figure 7b. For the exponential distribution, the integrals which have to be solved to derive the size distributions in lower dimensions, see Section 3, are not analytical tractable for general τ and γ. Some expressions for general parameter values are obtained below, but final results are obtained only under the assumption

$$\tau = \gamma = 1. \tag{23}$$

This value of τ corresponds to a linear relationship between maximum displacement and fault width, Expression (1), while the value of γ implies a liner displacement function, see Expression (5). Under this requirement, the integrals involved in the calculations below are analytically tractable.

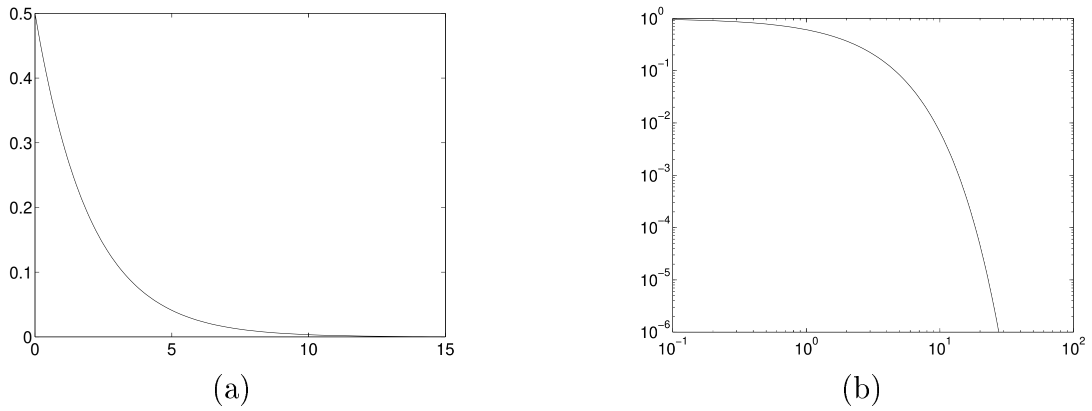


Figure 7: The exponential distribution with parameter $\lambda = 0.5$. Figure (a) shows the pdf $f(x)$ in Expression (21) and figure (b) the function $S(x)$ in Expression (22) on log-log scale.

4.4.1 Displacement

The pdf of the displacement in three dimensions is given by

$$f_3(d) = \lambda \exp(-\lambda d), \qquad d \geq 0. \tag{24}$$

Using the procedure described in Section 3, and the approximation (3) of the height of the fault ellipsoids, the pdf of D in two dimensions is found to be

$$f_2(d) = \gamma^{-1}\lambda^{1/\gamma}d^{1/\gamma-1}\frac{\Gamma(1/\tau - 1/\gamma + 1, \lambda d)}{\Gamma(1+1/\tau)}, \qquad d \geq 0, \tag{25}$$

where the parameters τ and γ must satisfy $\gamma > \tau/(1+\tau)$, see Appendix B. The incomplete gamma function $\Gamma(\cdot,\cdot)$ is described in Appendix C. Under the assumption in Expression (23), the pdf simplifies to

$$f_2(d) = \lambda\exp(-\lambda d), \qquad d \geq 0, \tag{26}$$

which is the pdf of the exponential distribution, see Expression (21). Thus under the specified assumptions, if the distribution of D in three dimensions is exponential, then the distribution is exponential also in two dimensions. The parameter is λ both in three and two dimensions.

As in Section 4.3.1, the relationship between W and D in two dimensions is required to be able to derive $f_1(d)$. This relationship is discussed below. If the relationship $W \propto D$ is assumed to be valid also for displacement and width in two dimensions, the displacement has the same exponential distribution in one dimension as in three and two:

$$f_1(d) = \lambda\exp(-\lambda d), \qquad d \geq 0. \tag{27}$$

4.4.2 Width

The pdf of the width in three dimensions is found from the pdf of D in Expression (24), using the transformation $W = (Dc_1)^{1/\tau}$:

$$f_3(w) = \rho\tau w^{\tau-1}\exp(-\rho w^\tau), \qquad w \geq 0, \tag{28}$$

where the parameter ρ is given by $\rho = \lambda c_1^{-1}$. This is the pdf of a Weibull distribution. Let W' denote the width obtained in two dimensions. Using the procedure in Section 3 and the approximation (2) of the fault height, the pdf of W' in two dimensions is found to be

$$f_{W'}(w') = \frac{\tau\rho^{1+1/\tau}}{\Gamma(1+1/\tau)}w'\int_{w'}^{\infty}\frac{w^{\tau-1}}{\sqrt{w^2-(w')^2}}\exp(-\rho w^\tau)\,dw, \tag{29}$$

see Appendix B. The integral does not have a general analytical solution for arbitrary τ. Returning to the assumption in Expression (23), the pdf of W in three dimensions, Expression (28), becomes $f_3(w) = \rho\exp(-\rho w)$, the pdf on an exponential distribution. The pdf of the width in two dimensions, now denoted W, is found from Expression (29):

$$f_2(w) = \rho^2 w K_0(\rho w), \qquad w \geq 0, \tag{30}$$

see Appendix B. The mean and variance of W are $\mathrm{E}[W] = \pi/2\rho$ and $\mathrm{Var}[W] = (16 - \pi^2)/4\rho^2$. The function $K_0(\cdot)$ is the modified Bessel function of order 0, see Appendix C, and the pdf $f_2(w)$ is shown in Figure 8a. Unlike the pdfs of the Pareto distribution and the exponential distribution, this function is not monotonely decreasing. The function $S(w)$ defined in Expression (9) is in this case given by

$$S(w) = \rho w K_1(\rho w), \qquad w \geq 0, \tag{31}$$

where $K_1(\cdot)$ is the modified Bessel function of order 1, see Appendix C. The function $S(w)$ is plotted on log-log scale in Figure 8b.

The relationship between W and D in two dimensions was needed in the derivation of $f_1(d)$ from $f_2(d)$, and the pdf in Expression (27) was derived under the assumption $W \propto D$. This relationship is however not consistent with the transformation

$$f_2(d) = f_2(w(d)) \cdot |w'(d)|, \tag{32}$$

using $f_2(d)$ from Expression (26) and $f_2(w)$ from Expression (30). It can be shown that this transformation is obtained for the relationship

$$D = -\lambda^{-1} \ln(\rho W K_1(\rho W)).$$

This relationship is slightly convex, but a relatively good approach to an asymptotic linearity is soon established as W increases.

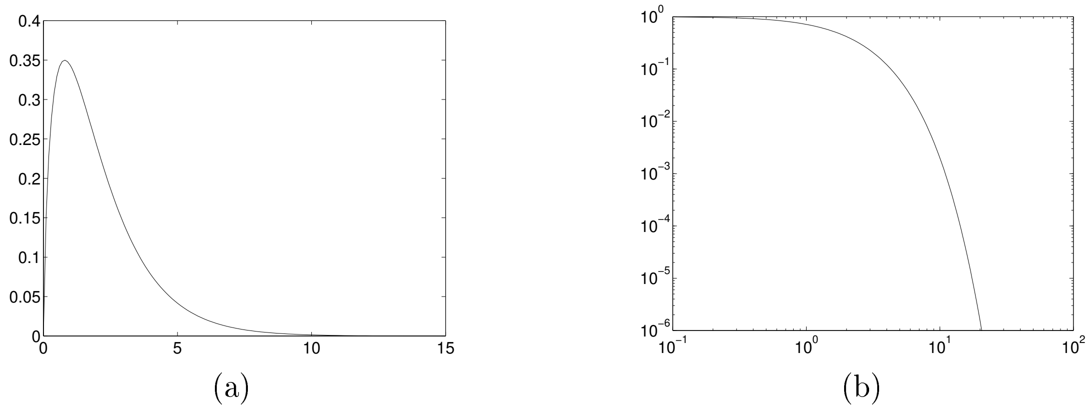


Figure 8: (a) The pdf $f_2(w)$ given in Expression (30) with parameter $\rho = 0.75$. (b) The function $S(w)$ in Expression (31) on log-log scale.

4.4.3 Exponential distributions with lower bounds

The pdf of a fractal distribution only exists if a lower limit of fault size is specified. This lower bound is not necessary for the distributions derived from an exponential distribution of displacement in three dimensions. However, to compare these distributions with the fractal distributions, the same lower bounds are introduced. Under the assumption in Expression (23), equal exponential distributions are obtained for the displacement in three, two and one dimensions, see Expressions (24), (26) and (27). Introducing a lower bound d_0, where $d_0 = d_{03}$ in three dimensions, d_{02} in two and d_{01} in one, the pdf becomes

$$f(d) = \lambda \exp(-\lambda(d - d_0)), \qquad d \geq d_0. \tag{33}$$

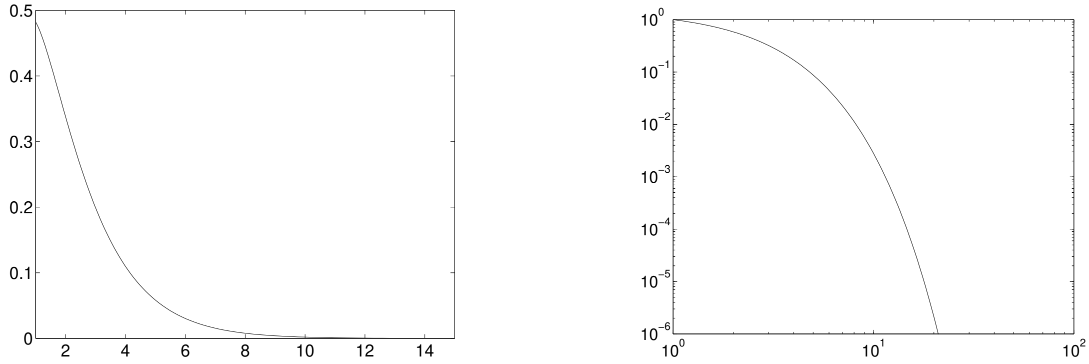
The pdfs of the width distributions in Expressions (28) and (30), with the same lower bounds as for the fractal distributions, are found to be

$$\begin{aligned} f_3(w) &= \rho \exp(-\rho(w - w_{03})), \qquad w \geq w_{03}, \\ f_2(w) &= \rho \cdot \frac{w K_0(\rho w)}{w_{02} K_1(\rho w_{02})}, \qquad w \geq w_{02}. \end{aligned} \tag{34}$$

The pdf (34) is shown in Figure 9a. For reasonably large values of the lower bound w_{02}, the truncated pdf is monotonely decreasing. When lower bounds are imposed on the width, the function $S(w)$ corresponding to the pdf (34) is

$$S(w) = \frac{w K_1(\rho w)}{w_{02} K_1(\rho w_{02})}, \tag{35}$$

and is illustrated on log-log scale in Figure 9b.



(a) (b)

Figure 9: (a) The pdf $f_2(w)$ given in Expression (34) width parameter $\rho = 0.75$ and lower bound $w_{02} = 1.0$. (b) The corresponding function $S(w)$ in Expression (35) on log-log scale.

4.5 Number of faults

In addition to the fault size distributions, the distribution of the number of faults is also of interest. The fault positions are assumed to follow a stationary Poisson point process, see Section 4.1, thus the distribution of the number of faults, K, is a Poisson distribution:

$$f(k;\omega) = \frac{\omega^k}{k!}\, e^{-\omega}, \qquad \omega > 0,\ k = 0, 1, 2, \ldots, \tag{36}$$

where ω is the intensity of the Poisson process. The number of faults in both three, two and one dimensions will follow a Poisson distribution, but with different intensity ω in different dimensions.

For seismic data, prior information about fault intensity can be obtained from wells, providing information about the number of faults along a one dimensional line. Thus if data is obtained in two dimensions, the fault intensity obtained from the well should be converted to an approximate number of faults in the area of observations. Denote by K_2 the number of faults in the two dimensional observation area and by K_1 the number intersected by a one dimensional line. The expected number of faults along a line through the area is then given as $\mathrm{E}\left[K_1|K_2\right] = f_{I_1}(1)K_2$, where $f_{I_1}(1)$ is the probability of a fault in the area being intersected by the line. If k_1 is obtained from well information, a rough estimate of K_2 given as $\hat{k}_2 = k_1/f_{I_1}(1)$. The probability $f_{I_1}(1)$ is calculated in a similar way as the probability f_{I_2} of a fault being intersected by a two dimensional plane, see Expression (6), and is proportional to the mean value of the fault width in two dimensions. Thus based on one dimensional well observations, the prior mean number of faults in two dimensions, $\mathrm{E}\left[K_2\right] = \omega$, should be chosen as

$$\omega = \mathrm{const} \times \frac{k_1}{c_1 \mathrm{E}_2\left[D\right]}, \tag{37}$$

under the assumption $\tau = 1$ in Expression (23), where $\mathrm{E}_2\left[D\right]$ is the expected maximum displacement in two dimensions. The constant term is related to the size of the observation area. Since $\mathrm{E}_2\left[D\right]$ differs under the fractal and exponential model, the mean number of faults ω should be chosen differently for each model, if based on the same fault intensity in one dimension. Note that the mean value of D in two dimensions in the case of a Pareto distribution is only finite for $\beta_2 > 1$, see Expression (16), thus Expression (37) only holds in this case. If $\beta_2 \leq 1$, the number of faults in two dimensions can be approximated by only considering faults of displacement below some specified size, $D \leq d_{\max}$, in which case the mean value is finite under both models.

4.6 Prior distribution

The distributions presented in Sections 4.3–4.5 can serve as prior distributions for a fault population, when a Bayesian framework is used to make inference based on observations of faults. Fault sizes in an observation area are now denoted $\boldsymbol{Y} = (Y_1, \ldots, Y_K)$, where Y_j represents either the maximum displacement or the width of a fault. The prior pdf of the number of faults in the observation area, K, and their sizes $\boldsymbol{Y}$ can be expressed as

$$f(\boldsymbol{y}, k, \theta) = f(\boldsymbol{y}|k, \theta) f(k;\omega) f(\theta), \tag{38}$$

where Θ is a model parameter involved in the pdf of $\boldsymbol{Y}$, while K has the prior pdf defined in Expression (36) and is independent of Θ. Assuming Y_j, $j = 1, \ldots, K$, are iid variables, the pdf $f(\boldsymbol{y}|k, \theta)$ is given as

$$f(\boldsymbol{y}|k, \theta) = \prod_{j=1}^{k} f(y_j|\theta) \tag{39}$$

where $f(y_j|\theta)$ is the marginal pdf of the fault size Y_j. If the observation area is two dimensional, the pdfs given in Expression (16), (26) or (33) can be used for the maximum displacement. For fault widths observed in two dimensions the pdfs in Expression (20), (30) or (34) can be used as prior pdfs. Some proper prior $f(\theta)$ is chosen for the model parameter. The prior pdf of $(\boldsymbol{Y}, K, \theta)$ is thus

$$f(\boldsymbol{y}, k, \theta) = \frac{\omega^k}{k!} e^{-\omega} f(\theta) \prod_{j=1}^{k} f(y_j|\theta). \tag{40}$$

5 Observations of Faults

Faults can be directly observed in outcrops or in core samples from wells, or indirectly observed from seismic data. In either case there is a sampling error connected to the observations, see for example Heffer and Bevan (1990) and Pickering et al. (1995). Different types of errors are of interest in this work: measurement errors, censoring and missing observations. A model for observation errors is defined through a likelihood function below. Qualitatively the model is based on the discussions of error types given by Heffer and Bevan (1990) and Pickering et al. (1995), although the cited papers do not use the same parametric models.

For seismic data, measurement errors mainly consist of a downward bias, where both fault widths and maximum displacements tend to be measured too small, due to the limit of resolution. Furthermore, the resolution of seismic data also results in a number of faults not being observed, even at scales above the limit of resolution, hence there are often missing observations. In geophysical literature, the term truncation is sometimes used to denote both the downward bias in measured fault width and the missing observations at some scales. Censoring of fault size observations is a result of the bounded observation area. For faults extending beyond the observation area, the observed sizes are censored and only give the size observed within the area.

5.1 Likelihood function

Consider the set of faults within an observation area, with sizes $\boldsymbol{Y} = (Y_1, \ldots, Y_K)$. Not all faults are observed, and the stochastic vector $\boldsymbol{S} = (S_1, \ldots, S_K)$ of indicator variables

$S_j \in \{0,1\}$ describes which faults that are observed. If fault j is observed, then $S_j = 1$, while $S_j = 0$ otherwise. The number of observed faults is thus given as $N = \sum_{j=1}^{K} S_j \leq K$. Furthermore, due to measurement errors, the observed values, denoted $\boldsymbol{Z}$, are not exact. The likelihood function of the observations accounts for the observation errors described above, and include information about the likelihood of observing the N faults defined by the indicator vector $\boldsymbol{S}$, with observed sizes $\boldsymbol{Z}$, when the actual number of faults is K and their sizes are $\boldsymbol{Y}$. The probability of observing a fault of size Y_j is given as $p(Y_j)$, depending on the fault size, but independent of all other faults. Thus $\mathrm{P}\{S_j = 1 | Y_j = y_j\} = p(y_j)$. If a fault is observed, i.e., $S_j = 1$, the observed value Z_j is described by the likelihood function $h(z_j|y_j)$, depending only on Y_j since Z_j is a measurement of this specific fault. The likelihood function of the observations $\boldsymbol{X} = (\boldsymbol{Z}, \boldsymbol{S})$ can thus be expressed as

$$h(\boldsymbol{x}|\boldsymbol{y}, k) = h(\boldsymbol{s}|\boldsymbol{y}, k) h(\boldsymbol{z}|\boldsymbol{s}, \boldsymbol{y}, k) = \prod_{j=1}^{k} p(y_j)^{s_j} (1 - p(y_j))^{1-s_j} \prod_{j; s_j = 1} h(z_j|y_j). \tag{41}$$

5.1.1 Observed number of faults

For a seismic dataset $\boldsymbol{x} = (\boldsymbol{z}, \boldsymbol{s})$ there are typically missing observations, and only a proportion of the true number of faults is observed. The probability of observing exactly the n faults observed, conditioned on K and $\boldsymbol{Y}$, is given as

$$h(\boldsymbol{s}|\boldsymbol{y}, k) = \prod_{j=1}^{k} p(y_j)^{s_j} (1 - p(y_j))^{1-s_j}$$

and is specified through the function $p(y)$. It is natural to assume that $p(y)$ is a monotonely increasing function, see Clark et al. (1999), since missing observations due to limited resolution of the measurement equipment is most dominant for small scale faults, while the presence of large scale faults in the observation area is more frequently detected. In this work $p(y)$ is assumed to be piecewise linear, given as

$$p(y) = \begin{cases} p_1, & y \leq \eta_1 \\ (\eta_2 - \eta_1)^{-1}((p_2 - p_1)y + p_1\eta_2 - p_2\eta_1), & \eta_1 \leq y \leq \eta_2 \\ p_2, & y > \eta_2, \end{cases} \tag{42}$$

and illustrated in Figure 10. The function gives a constant small probability p_1 of observing faults of size smaller than $y = \eta_1$, and a constant high probability p_2 of observing faults of size larger than $y = \eta_2$. Between $y = \eta_1$ and $y = \eta_2$ the probability of observation is linearly increasing. The parameter η_1 is related to the limit of resolution, and p_1 can be interpreted as the probability of observing a fault of size at the limit of resolution. The parameter η_2 can be interpreted as a value for which most faults of size larger than this value will be observed. In the special case $p_1 = p_2 = 1$ all faults have probability 1 of being observed, thus $S_j = 1$, $j = 1, \ldots, K$, and the observed number of faults equals the true number, $n = K$.

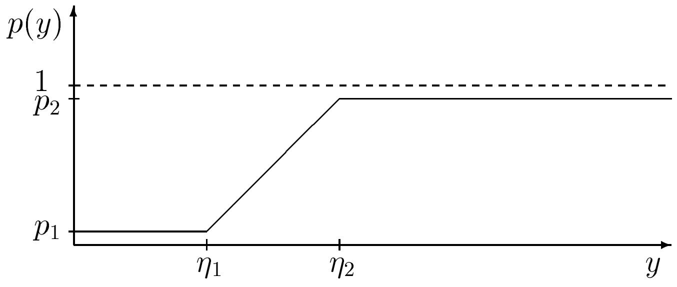


Figure 10: The function $p(y)$ in Expression (42).

5.1.2 Observed fault sizes

The likelihood function of observed fault sizes $\boldsymbol{Z}$,

$$h(\boldsymbol{z}|\boldsymbol{s}, \boldsymbol{y}, k) = \prod_{j; s_j = 1} h(z_j|y_j),$$

is used to model the measurement errors, which typically leads to a downward bias. For an observed fault, the measurement errors is given as the difference $Y_j - Z_j$. For seismic observations, the observed value tends to give a too small measurement of Y_j, due to the resolution of the seismic data, thus $Y_j - Z_j \geq 0$. For simplicity, it is assumed to be a lower and an upper limit of possible measurement error of this type, thus $\delta_1 \leq Y_j - Z_j \leq \delta_2$, and a uniform distribution of the error within this interval is chosen.

The dimension of the observation area of faults enforces an upper bound η on observable fault size Z. If fault widths are observed, η is directly related to the size of the observation area. There is however no strict upper bound of observable maximum displacement inside a bounded area, hence in this case η can be chosen as $\eta = \infty$. If the true fault size is only slightly larger than η, the downward bias in the observations gives an observation $Z \leq \eta$ anyway. As Y increases, the observed value is censored and remains below η. The observation Z now only measures the size of the fault within the observation area, and some bias is still allowed, with the same lower and upper limits as before.

If Y exceeds the upper bound η, censoring is guaranteed. However, also smaller faults, $Y < \eta$, can in practice be censored, if one or both fault tips cross the boundary of the observation area. This kind of censoring is not accounted for in the model.

The uniform likelihood function $h(z|y)$ is illustrated in Figure 11, and can be expressed as

$$h(z|y) = \frac{1 + (\delta_2 - y)\delta_z(0)}{\delta_2 - \delta_1}, \qquad \delta_1 \leq \min\{y, \eta + \delta_1\} - z \leq \delta_2 \text{ and } z \geq 0, \tag{43}$$

where the term $(\delta_2 - y)\delta_z(0)$; $\delta_z(\cdot)$ being the Dirac delta function, is included to account for boundary effects for small fault sizes. The choice of edge correction has only minor influence on the model.

The special case $\delta_1 = \delta_2 = 0$ implies no bias in the observations, and the fault size is observed exactly for all faults with size below the upper bound η. In this case the likelihood function is replaced by $h(z|y) = \delta_z(\min\{y, \eta\})$. If in addition the upper bound is $\eta = \infty$, all fault sizes are observed exactly.

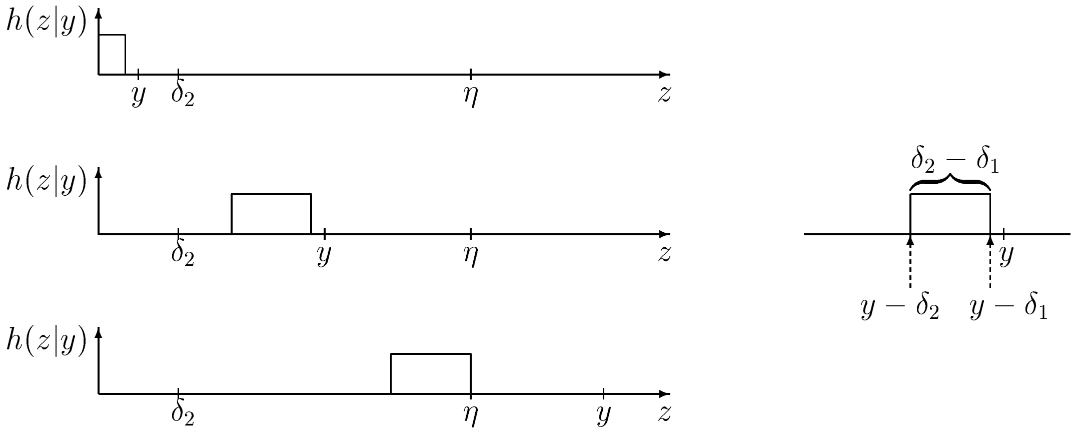


Figure 11: The likelihood function $h(z|y)$, Expression (43), for different values of Y. The left hand figures illustrate the uniform distribution of Z for small, intermediate and large values of Y. The right hand figure shows the interval boundaries of the distribution.

5.2 Posterior distribution

The focus of this work is on the posterior distribution of fault sizes, conditioned to observations, and inference about the underlying fault population is based on this distribution. In a Bayesian setting, the prior pdf $f(\boldsymbol{y}, k, \theta) = f(\boldsymbol{y}, k|\theta)f(\theta)$ of $(\boldsymbol{Y}, K, \Theta)$, see Section 4.6, includes general geological knowledge about the distribution of fault sizes and the number of faults in the observation area. The likelihood function $h(\boldsymbol{x}|\boldsymbol{y}, k)$ of the observations $\boldsymbol{X} = (\boldsymbol{Z}, \boldsymbol{S})$ models the data collection, see Section 5.1. The observations are assumed to be conditionally independent of the parameter Θ, given $(\boldsymbol{Y}, K)$. Combining the prior pdf (40) of the true fault population $(\boldsymbol{Y}, K, \theta)$ and the likelihood function (41) of the observations $\boldsymbol{X} = (\boldsymbol{Z}, \boldsymbol{S})$, the posterior pdf is obtained:

$$\begin{aligned} f(\boldsymbol{y}, k, \theta|\boldsymbol{x}) &= \text{const} \times \frac{\omega^k}{k!} e^{-\omega} f(\theta) \prod_{j=1}^{k} f(y_j|\theta) \\ &\quad \times \prod_{j=1}^{k} p(y_j)^{s_j} (1 - p(y_j))^{1-s_j} \prod_{j;s_j=1} h(z_j|y_j). \end{aligned} \tag{44}$$

6 Model Choice

The relationship between the observations $\boldsymbol{X}$ and the true variables $(\boldsymbol{Y}, K)$ is described by the likelihood function in Expression (41). Suppose now that the pdf $f(\boldsymbol{y}, k|\theta)$, Section 4.6, is unknown, but two possible models are suggested. The stochastic variable $M \in \Omega_m = \{m_1, m_2\} = \{1, 2\}$ indicates which model is the correct one, model 1 or 2 respectively. Under model m, the pdf of $(\boldsymbol{Y}, K)$ is denoted $f(\boldsymbol{y}, k|m, \theta_m) = f(k|m) f(\boldsymbol{y}|m, k, \theta_m)$, $m = 1, 2$, where different model parameters θ_m with prior pdfs $f(\theta_m|m)$ describe the two distributions. Based on the observations $\boldsymbol{X}$ it is of interest to determine which of the two pdfs of $(\boldsymbol{Y}, K)$ is most likely. This can be expressed as a model choice problem, and Bayes factor can be used as a criterion for choosing the model best fitted to the observations. See Kass and Raftery (1995) for an overview. The two competing models are assigned prior probabilities $f(m_1)$ and $f(m_2)$. Bayes factor $B_{12}(\boldsymbol{x})$ of the observations $\boldsymbol{X}$ is defined as

$$B_{12}(\boldsymbol{x}) = \frac{f(m_1|\boldsymbol{x})}{f(m_2|\boldsymbol{x})} \cdot \frac{f(m_2)}{f(m_1)} = \frac{f(\boldsymbol{x}|m_1)}{f(\boldsymbol{x}|m_2)}, \tag{45}$$

where $f(m|\boldsymbol{x})$ is the posterior probability of model m. If the value of Bayes factor exceeds one, $B_{12}(\boldsymbol{x}) > 1$, model 1 is favored, while for a value smaller than one model 2 is most likely. If the prior probabilities of model 1 and 2 are equal, Bayes factor coincides with the posterior odds ratio. The marginal pdf of $\boldsymbol{X}$ under model m is given as

$$f(\boldsymbol{x}|m) = \sum_{k=0}^{\infty} \iint h(\boldsymbol{x}|\boldsymbol{y}, k) f(\boldsymbol{y}, k|m, \theta_m) f(\theta_m|m) \, d\theta_m \, d\boldsymbol{y}. \tag{46}$$

In some cases Bayes factor can be found analytically, but in general the value must be estimated by the use of numerical integration techniques. Evans and Swartz (1995) and Kass and Raftery (1995) discuss different estimation techniques. In this work, Bayes factor is estimated by the use of McMC sampling, where the model M is included as a stochastic variable, see Carlin and Chib (1995). An introduction to McMC theory is given, for instance, by Besag et al. (1995) and Han and Carlin (2000) give a review of McMC methods for estimating Bayes factor.

Borgos and Omre (2000) describes an McMC sampling algorithm constructed to estimate Bayes factor based on non-exact observations. The McMC algorithm introduces auxiliary variables K_1 and K_2, representing the number of faults K under the two competing models. The posterior distribution of $(M, K_1, K_2, \Theta_1, \Theta_2)$ given $\boldsymbol{X}$ can be found, integrating over $\boldsymbol{Y}$. The McMC algorithm then generate samples from this posterior distribution, and the variable $\boldsymbol{Y}$ need not be included in the McMC simulations. Based on realizations $(m^j, k_1^j, k_2^j, \theta_1^j, \theta_2^j)$, $j = 1, 2, \ldots, r$, generated by the McMC algorithm, Bayes factor can be estimated as

$$\hat{B}_{12}(\boldsymbol{x}) = \frac{\sum_{j=1}^{r} \mathbf{1}[m^j = m_1]}{\sum_{j=1}^{r} \mathbf{1}[m^j = m_2]} \cdot \frac{f(m_2)}{f(m_1)}, \tag{47}$$

where $\mathbf{1}[\cdot]$ is the indicator function. Estimation of the model parameter Θ_m, the number of faults K and fault sizes $\boldsymbol{Y}$ under the favored model is discussed in Section 6.2.

6.1 Competing models

Two alternative fault size distributions are discussed in Section 4, a fractal and an exponential model. As illustrated in Figure 6b, for the fractal model a log-log plot of the function $S(x)$, Expression (9), gives a straight line. However, the log-log plot of the estimate $\hat{S}(x)$, Expression (10), based on observed fault sizes often shows a concave shaped curve. The concave shape can in some cases be explained by the typical sampling errors, especially missing observations can result in this shape of the curve. However, the concave curve can also be a result of a different underlying distribution, for example an exponential model. As shown in Figure 7b, an exponential distribution gives a concave curve on a log-log plot of $S(x)$. The aim of this work is to use a statistical framework to investigate which model is the most likely one, based on a set of fault observations.

In the calculations below, fault size Y is assumed to be measured as maximum displacement D. True fault sizes $\boldsymbol{Y}$ are assumed to be iid variables, see Expression (39), while the number of faults K follows a Poisson distribution, Expression (36). The fractal model with the pdf in Expression (16) as marginal distribution of Y_j and model parameter $\theta_1 = \beta_2$ is used as model 1. The assumption in Expression (23) is assumed to hold, and the truncated exponential model with the pdf in Expression (33) as marginal pdfs and model parameter $\theta_2 = \lambda$ is chosen as model 2. Thus

$$f(y|m_1, \theta_1) = \frac{\theta_1 y_0^{\theta_1}}{y^{\theta_1 + 1}} \qquad \text{and} \qquad f(y|m_2, \theta_2) = \theta_2 \exp(-\theta_2(y - y_0)), \qquad y \geq y_0. \tag{48}$$

A common lower bound y_0 is used in the two models.

The model choice problem is based on observations $\boldsymbol{X} = (\boldsymbol{Z}, \boldsymbol{S})$ of maximum displacements obtained in two dimensions, see Section 5, where the number of observed faults is given as $n = \sum_{j=1}^{K} S_j$ and the observed maximum displacements are given as Z_j for all $\{j; s_j = 1\}$. The two competing models are assigned equal prior probabilities, $f(m_1) = f(m_2) = 0.5$, thus there is no prior belief that one model is more likely than the other.

The variable K represents the same quantity under both models. As mentioned above, auxiliary variables K_1 and K_2 are introduced to represent this variable under the two competing models in the McMC algorithm described by Borgos and Omre (2000). The main reason for this is that the posterior distribution of K can be bimodal, with one mode corresponding to each model, and auxiliary variables are introduced to avoid this bimodality. Depending on how prior information about the number of faults is obtained, it may also be appropriate to define different prior distributions of K under the two models, as discussed in Section 4.5. This is obtained by choosing individual intensities ω_m in the Poisson distribution in Expression (36).

6.2 Inference under the favored model

When a model choice has been made based on a set of observations, the next aim is to estimate the corresponding model parameters. Estimation of the power law exponent in a fractal model is frequently found in the literature, and is used in prediction of the number of faults in the observation area or in extrapolation to sub-seismic faults, see Gauthier and Lake (1993), Walsh et al. (1994). The standard estimation procedure is to fit a straight line to the log-log plot of the cumulative number of faults versus fault size. For discussions and improvements on this estimation technique, see Pickering et al. (1996). Pickering et al. (1995) and Clark et al. (1999) compared different procedures for estimating the exponent in a negative power law distribution, for example maximum likelihood estimation.

In this work, the model parameters in the favored model is estimated based on McMC simulations from the posterior distribution. Since the algorithm generates realizations of the model indicator and the model parameters simultaneously, parameter estimation can be performed based on output from the same run of the algorithm. An estimate of the parameter value in model m is given by the estimated posterior mean

$$\hat{\theta}_m = \frac{\sum_{j=1}^{r} \theta_m^j \mathbf{1}[m^j = m]}{\sum_{j=1}^{r} \mathbf{1}[m^j = m]}. \tag{49}$$

The number of faults in the observation area can be estimated similarly based on the realizations of K_m under the favored model, and the unknown fault sizes $\boldsymbol{Y}$ can be generated from the posterior distribution $f(\boldsymbol{y}|\boldsymbol{x}, m, k_m, \theta_m)$, see Borgos and Omre (2000).

The number of faults in the three dimensional reservoir, denoted $K_{m,3}$ under the favored model, can be predicted based on the estimate of the number of faults in two dimensions, now denoted $K_{m,2}$. The expected number in two dimensions, given $K_{m,3}$, is

$$\mathrm{E}\left[K_{m,2}|K_{m,3}\right] = f_{I_2}(1|m) K_{m,3}$$

where $f_{I_2}(1|m)$ is the probability of intersecting a fault by a two dimensional plane, under model m, see Expression (6). In a reservoir of vertical thickness T_v, the probability is given as

$$f_{I_2}(1|m) = T_v^{-1}\mathrm{E}\left[h(D)|m\right] = \text{const} \times T_v^{-1}\mathrm{E}\left[D|m\right],$$

omitting the dip in the argument of the height function $h(\cdot)$, see Expression (3). The assumption $\tau = 1$ in Expression (23) is used, and the thickness of the reservoir is assumed to fulfill $T_v \gg h(D)$ for any maximum displacement D in the observation area, such that the vertical extent of all faults are completely contained in the region under study. The constant term depends on the distribution of the dip and on the constants c_1 and c_2, quantities which are not discussed in this work. When $K_{m,2}$ is observed or estimated, an estimate of the frequency $K_{m,3}/T_v$ is given by

$$\frac{\hat{K}_{m,3}}{T_v} = \frac{k_{m,2}}{f_{I_2}(1|m)T_v} = \text{const} \times \frac{k_{m,2}}{\mathrm{E}\left[D|m\right]}. \tag{50}$$

The mean value of D can be estimated for the chosen distribution, using the corresponding estimated model parameter $\hat{\theta}_m$. Under the fractal model, this is found to be

$$\hat{\mathrm{E}}\left[D|m_1\right] = \hat{\theta}_1^{-1}(\hat{\theta}_1 + 1)d_{03},$$

and in the exponential model

$$\hat{\mathrm{E}}\left[D|m_2\right] = \hat{\theta}_2^{-1} + d_{03}.$$

6.3 Exact observations

In this section the special case of no sampling errors is considered, where the exact number of faults is observed and fault sizes are observed without measurement errors. In this case the observations are $\boldsymbol{X} \equiv \boldsymbol{Y}$, and the likelihood function in Expression (41) is replaced by $h(\boldsymbol{x}|\boldsymbol{y}) = \delta_{\boldsymbol{x}}(\boldsymbol{y})$. The number of faults, K, is not treated as a stochastic variable in this case. The two proposed pdfs of Y under model 1 and 2, Expression (48), are now also the pdfs of X under the respective models. Bayes factor can in this situation be found analytically, using gamma priors $\mathrm{Gamma}(\alpha_m, \sigma_m)$ for the parameters θ_m, with mean $\mathrm{E}\left[\theta_m\right] = \alpha_m\sigma_m$ and variance $\mathrm{Var}\left[\theta_m\right] = \alpha_m\sigma_m^2$. The calculation of Bayes factor involves solving some integrals which are recognized as gamma functions, and is demonstrated by Borgos and Omre (2000). In this case Bayes factor is:

$$B_{12}(\boldsymbol{x}) = (\prod_{j=1}^{k} x_j)^{-1} \frac{\sigma_2^{\alpha_2}(\sigma_2^{-1} + \sum_{j=1}^{k}(x_j - x_0))^{\alpha_2+k}}{\sigma_1^{\alpha_1}(\sigma_1^{-1} + \sum_{j=1}^{k}(\ln x_j - \ln x_0))^{\alpha_1+k}} \cdot \frac{\Gamma(\alpha_2)\Gamma(\alpha_1+k)}{\Gamma(\alpha_1)\Gamma(\alpha_2+k)}. \tag{51}$$

The resulting posterior distributions of the parameters are gamma distributions with parameters $\alpha_m + k$ and $(\sigma_m^{-1} + t_m(\boldsymbol{x}))^{-1}$, where $t_1(\boldsymbol{x}) = \sum_{j=1}^{k}(\ln x_j - \ln x_0))$ and $t_2(\boldsymbol{x}) = \sum_{j=1}^{k}(x_j - x_0))$. The posterior mean and variance are given as

$$\mathrm{E}\left[\theta_m|\boldsymbol{x}\right] = \frac{\sigma_m(\alpha_m + k)}{1 + \sigma_m t_m(\boldsymbol{x})}, \qquad \mathrm{Var}\left[\theta_m|\boldsymbol{x}\right] = \frac{\sigma_m^2(\alpha_m + k)}{(1 + \sigma_m t_m(\boldsymbol{x}))^2}. \tag{52}$$

The model choice problem can thus be solved analytically, and the posterior mean can be used as estimators of the model parameters.

7 Example: Seismic Data

Figure 12 shows a dataset containing maximum displacements $\boldsymbol{x}$ observed on a two dimensional surface, interpreted from seismic data. The data are from the Gullfaks field in the North sea, see Fossen and Rørnes (1996). The number of observations is $n = 169$, with displacement values ranging from $2m$ to $256m$. Figure 12a shows a histogram of the observed displacements, corresponding to the pdf $f(x)$. Figure 12b shows a log-log plot of $\hat{S}(d)$, see Section 4.2 for an explanation.

The pdfs in Expression (48) are used as competing models, with $y_0 = 2m$ chosen as the minimum of the observations. The model parameters of the two competing models are assigned gamma priors with hyper-parameters $\alpha_1 = \alpha_2 = 4$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.005$. Thus the prior mean values are $\mathrm{E}\left[\theta_1|m_1\right] = 1.0$ and $\mathrm{E}\left[\theta_2|m_2\right] = 0.02$. If the observations are assumed to be exact, Bayes factor from Expression (51) is given as $B_{12}(\boldsymbol{x}) = \mathcal{O}(10^{-35}) \approx 0$ and the posterior mean values are $\mathrm{E}\left[\theta_1|\boldsymbol{x}\right] = 0.393$ and $\mathrm{E}\left[\theta_2|\boldsymbol{x}\right] = 0.0232$, see Expression (52). In this case the exponential model is undoubtedly favored. However, the seismic data are not exact and the estimation of Bayes factor should account for the uncertainty caused by sampling errors. These errors are modeled by the likelihood function. Different choices of parameters in the likelihood function and prior distributions can lead to different conclusions of the model choice problem. This is demonstrated through a number of simulations examples, using the McMC algorithm described by Borgos and Omre (2000). In all simulations reported below, the estimates of model parameters and the number of faults are based on 5000 realizations from the McMC algorithm, separated by 10 iterations, while all 50 000 iterations are used to estimate Bayes factor. A sensitivity analysis is first carried out, exploring the influence of the parameters in the prior pdf and the likelihood function. Finally, well information is used to determine which prior parameters are most realistic.

Only a few parameters are varied in the simulations, the parameters η_2 and p_1 in the likelihood function and the prior mean numbers of faults ω_m under the two models. The likelihood function of the observed displacements $\boldsymbol{x} = (\boldsymbol{z}, \boldsymbol{s})$, Expression (41), contains the functions $p(y)$ given in Expression (42) and $h(z|y)$ given in (43). In the simulation study the function $h(z|y)$ is held constant, with parameter values $\delta_1 = 1m$, $\delta_2 = 5m$ and $\eta = \infty$. Thus the measurement error of each fault displacement is assumed to be in the range $1m$ to $5m$, while no censoring is assumed to occur. The influence of the parameters η_2 and p_1 in the function $p(y)$ is studied. The other values are held constant with values $\eta_1 = 2m$, chosen as the minimum value of the observations, and $p_2 = 0.99$. For $y \geq \eta_2$ there is a constant probability of 0.99 of observing the fault, while for decreasing $y \leq \eta_2$

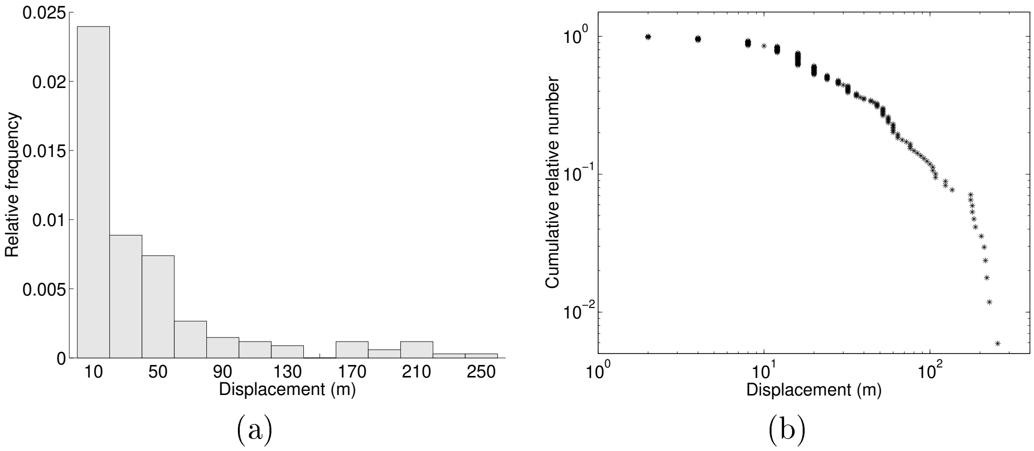


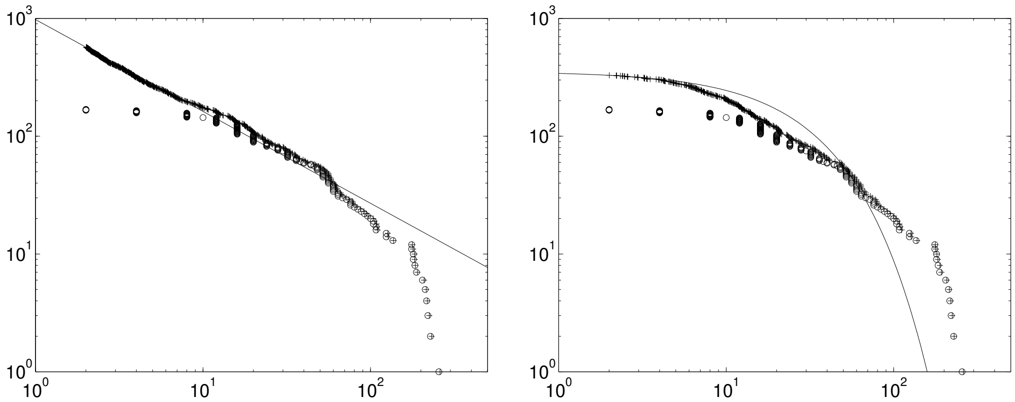
Figure 12: Maximum displacements observed on a two dimensional surface. (a) Histogram of the observations. (b) Log-log plot of $\hat{S}(d)$, see Expression (10).

the probability is decreasing to p_1 for $y = 2m$. Table 1 gives a summary of all parameter values held constant in the simulations, and the values of the hyper-parameters in the prior distributions of the model parameters. Preliminary estimates of the number of faults in two dimensions based on well observations, see Section 4.5, show that for typical values of Θ_1 and Θ_2 the number of faults in two dimensions under the fractal model exceeds the number of faults under the exponential model. Thus throughout the simulation study, $\omega_1 \geq \omega_2$ is chosen. Assessment of ω_m based on well observations from the Gullfaks field is discussed in Section 7.4.

Figure 13 shows two realizations of $(\boldsymbol{Y}, K)$ from the posterior distribution (44) conditioned to the observations $\boldsymbol{x}$. For these realizations, $\omega_1 = \omega_2 = 720$, $p_1 = 0.1$ and $\eta_2 = 20m$. Figure 13a shows a realization from the fractal model, while Figure 13b shows a realization from the exponential model. The figures show log-log plots of the estimates $k\hat{S}(y)$ based on the realizations, see Expression (10), plotting absolute number instead of relative number. The estimates are compared with the estimate $n\hat{S}(z)$ for the observed fault sizes and the functions $kS(y)$ obtained using Expression (48) with estimated model parameters. In both figures a number of unobserved small faults are present in the realization of $\boldsymbol{Y}$, and the left part of the estimate $\hat{S}(y)$ deviates to some extent from the corresponding estimate $\hat{S}(z)$ based on the observations. In the right part of the figures there are only minor differences between the curves based on the realizations and the observations. In this area almost all faults are observed, and since $\eta = \infty$ no censoring is adjusted for in the realizations.

Prior distribution	Likelihood function					Hyper-Parameters			
y_0	δ_1	δ_2	η	p_2	η_1	α_1	σ_1	α_2	σ_2
2	1	5	∞	0.99	2	4	0.25	4	0.005

Table 1: Parameter values held constant in all simulations.



(a) (b)

Figure 13: Realizations of $\boldsymbol{Y}$ from the posterior distribution $f(\boldsymbol{y}|\boldsymbol{x}, m, k_m, \theta_m)$ under (a) model 1 and (b) model 2. Estimates $\hat{\theta}_1 = 0.780$, $\bar{k}_1 = 570$, $\hat{\theta}_2 = 0.0370$ and $\bar{k}_2 = 329$ are obtained using parameters $\omega_1 = \omega_2 = 720$, $\eta_2 = 20m$ and $p_1 = 0.1$. The cumulative absolute number of faults, $k_m\hat{S}(y)$, is plotted on log-log scale (+), see Expression (10), and is compared to the observations $n\hat{S}(z)$ ($\circ$). The corresponding function $k_m S(y)$ for the prior distribution of Y is plotted with solid line, using the estimated model parameters.

7.1 Dependence on the prior distribution

Dependence of the model choice on the prior distribution is studied by varying ω_1 and ω_2, keeping all parameters of the likelihood function constant. Thus different results of the model choice problem is a consequence of the variations in ω_m, $m = 1, 2$. The parameter value $\eta_2 = 20m$ is used, while p_1 is held constant at the value $p_1 = 0.1$ or $p_1 = 0.01$. The remaining likelihood variables are given in Table 1. Table 2 shows the results of McMC simulations, reporting the estimated Bayes factor, $\hat{B}_{12}(\boldsymbol{x})$, for the observations in Figure 12, and the conclusion of the model choice problem. The table also lists the estimated model parameters, $\hat{\theta}_1$ and $\hat{\theta}_2$, and the estimated posterior mean number of faults, $\hat{\mathrm{E}}\left[K_m|\boldsymbol{x}, m\right] = \bar{K}_m$, for the favored model m.

From Table 2 it is clear that the prior mean numbers of faults have great influence on the result of the model choice problem. Especially, Table 2b illustrates how fast the conclusion of the model choice problem can change with changes in prior mean values. In most of the runs reported in the table, the exponential model is favored. However, with a relatively large, equal prior mean number of faults under both models, the fractal model is favored. From Table 2 it may seem that the model with the lowest prior mean number of faults in general is favored. However, comparing the posterior mean number of faults, the model with the higher mean number may be favored. In the simulation where $\omega_1 = \omega_2 = 720$ in Table 2b, no clear conclusion can be drawn. In this case, the posterior mean number of faults under model 1 is $\hat{k}_1 = 570$, while under model 2 it is $\hat{k}_2 = 329$, thus the posterior distribution of the number of faults alone is not crucial to the model choice.

In Table 2a,b the value of p_1 is $p_1 = 0.1$, while in Table 2c it is decreased to $p_1 = 0.01$. With a lower probability of observing faults at the limit of resolution, the posterior mean number of faults increases. Furthermore, different values of p_1 may result in different conclusions of the model choice, as is observed in the case $\omega_1 = \omega_2 = 700$. The dependence of the model choice on p_1 is studied below.

As ω_2, the prior mean number of faults under model 2, increases, the posterior mean number also increases, see Table 2. A corresponding change in parameter value θ_2 is observed. Similarly, the posterior mean number of faults and the parameter θ_1 under model 1 is observed to increase when the prior mean number of faults, ω_1, increases.

7.2 Dependence on the likelihood function

Dependence on the likelihood function is studied by assigning different values to η_2 and p_1, keeping ω_1 and ω_2 constant. Thus different results of the model choice are now a consequence of the variations in the values of η_2 or p_1. Both parameters η_2 and p_1 have influence on the number of unobserved faults. Table 3 shows the results of McMC simulations where η_2 is varied, with $p_1 = 0.1$ constant. In Table 4 the parameter p_1 is varied, keeping $\eta_2 = 20m$ constant. Estimates of Bayes factor, the conclusion of the model choice and the model parameter of the favored model are reported.

As η_2 increases, the fault size required for an almost certain observation increases. The overall probability of observation decreases, increasing the number of unobserved faults. The results of Table 3 show tendencies of an increasing Bayes factor as η_2 starts increasing, thus model 1 becomes more favorable. However, as η_2 increases further, Bayes factor decreases again. This effect is observed using prior mean values $\omega_1 = \omega_2 = 700$ and $\omega_1 = \omega_2 = 720$. For the dataset studied here, the posterior probability of the fractal model being the right model seems to have a maximum for values of η_2 in the range 20–40. The case $\omega_1 = 600$ and $\omega_2 = 400$ is also considered, in which case Bayes factor is approximately equal to zero for all reported values of η_2.

The number of unobserved faults also increases as p_1 decreases, since a small value of p_1 indicates a large proportion of unobserved faults at small scales. The effect of p_1 is illustrated in Table 4, for some choices of ω_1 and ω_2. In the case $\omega_1 = \omega_2 = 700$, model 2

ω_1	ω_2	$\hat{B}_{12}(\boldsymbol{x}, n)$	Model	$\hat{\theta}_1$	$\bar{k}_1$	$\hat{\theta}_2$	$\bar{k}_2$
400	400	0^*	2			0.0281	234
500	500	0^*	2			0.0303	256
600	600	$\mathcal{O}(10^{-17})$	2			0.0329	284
700	700	0.00538	2			0.0364	322
800	800	$\mathcal{O}(10^{12})$	1	0.822	624		
900	900	∞^*	1	0.874	702		
600	400	0^*	2			0.0283	235
700	500	0^*	2			0.0306	260
800	400	0^*	2			0.0283	235
800	600	0^*	2			0.0334	288
900	500	0^*	2			0.0311	265
900	700	0^*	2			0.0362	320
900	800	$3.25 \cdot 10^{-5}$	2			0.0387	349

(a)

ω_1	ω_2	$\hat{B}_{12}(\boldsymbol{x}, n)$	Model	$\hat{\theta}_1$	$\bar{k}_1$	$\hat{\theta}_2$	$\bar{k}_2$
700	700	0.00538	2			0.0364	322
710	710	0.115	2			0.0367	324
720	720	1.239	1/2	0.780	570	0.0370	329
730	730	169.8	1	0.770	559		

(b)

ω_1	ω_2	$\hat{B}_{12}(\boldsymbol{x}, n)$	Model	$\hat{\theta}_1$	$\bar{k}_1$	$\hat{\theta}_2$	$\bar{k}_2$
500	500	0^*	2			0.0317	271
700	500	0^*	2			0.0320	273
700	700	∞^*	1	0.795	589		
900	500	0^*	2			0.0322	273
900	700	0^*	2			0.0383	340
900	900	∞^*	1	0.915	773		

(c)

Table 2: Results of McMC simulations where $\eta_2 = 20m$ while various prior mean numbers ω_m are used. In (a) and (b) $p_1 = 0.1$ and in (c) this probability is reduced to $p_1 = 0.01$. Table (b) illustrates how the conclusion of the model choice can change for only minor changes of ω_1 and ω_2. The estimates of Bayes factor and the model parameters are obtained using Expressions (47) and (49). The estimates of the model parameters and the mean values $\bar{k}_m$ are based on 5000 realizations, separated by 10 iterations of the McMC algorithm. Bayes factor is estimated based on all 50 000 iterations. The reported values 0^* and ∞^* indicate that the estimate of Bayes factor is either very small or very large, without being identically 0 or ∞. $\mathcal{O}(u)$ gives the order of magnitude.

is favored when $p_1 = 0.1$. As p_1 decreases, the conclusion changes and model 1 is favored. The prior mean values $\omega_1 = 600$ and $\omega_2 = 400$ are also considered. In this case model 2 is favored for all values of p_1 considered, and a reduction of p_1 from 0.1 to 0.0001 does not give a change of the conclusion.

ω_1	ω_2	η_2	$\hat{B}_{12}(\boldsymbol{x}, n)$	Model	$\hat{\theta}_1$	$\bar{k}_1$	$\hat{\theta}_2$	$\bar{k}_2$
700	700	10	0^*	2			0.0275	225
700	700	20	0.00538	2			0.0364	322
700	700	40	0.129	2			0.0423	452
700	700	60	$\mathcal{O}(10^{-13})$	2			0.0420	527
700	700	80	0^*	2			0.0402	568
700	700	100	0^*	2			0.0382	585
720	720	10	0^*	2			0.0282	231
720	720	20	1.239	1/2	0.780	570	0.0370	329
720	720	30	$8.41 \cdot 10^4$	1	0.744	574		
720	720	40	0.743	1/2	0.743	621	0.0431	469
720	720	50	$5.25 \cdot 10^{-7}$	2			0.0430	507
600	400	20	0^*	2			0.0283	235
600	400	40	0^*	2			0.0320	295
600	400	60	0^*	2			0.0330	341
600	400	80	0^*	2			0.0326	369
600	400	100	0^*	2			0.0316	381

Table 3: Results of McMC simulations with $p_1 = 0.1$, while various parameter values of ω_1, ω_2 and η_2 are used. The estimates of Bayes factor and the model parameters are obtained using Expressions (47) and (49). The estimates of the model parameters and the mean values $\bar{k}_m$ are based on 5000 realizations, separated by 10 iterations of the McMC algorithm. Bayes factor is estimated based on all 50 000 iterations. The reported value 0^* indicates that the estimate of Bayes factor is very small, without being identically 0. $\mathcal{O}(u)$ gives the order of magnitude.

7.3 Parameter estimation and prediction of number of faults

Tables 2–4 show results of McMC simulations for observations with sampling errors. Both Bayes factor and the model parameter under the favored model are estimated, see Expressions (47) and (49). Figure 14 shows histograms of realizations of θ_1 and θ_2, for some of the simulation examples from Table 2. The histograms are compared with the corresponding prior pdf, with hyper parameters given in Table 1. The histograms give an estimate of the posterior pdf, showing how the modes of the posterior pdfs differ from the modes of the prior pdfs. Furthermore, the histograms have narrower ranges than the prior pdfs, thus the posterior distributions have a smaller variance than the prior distributions. As ω_m

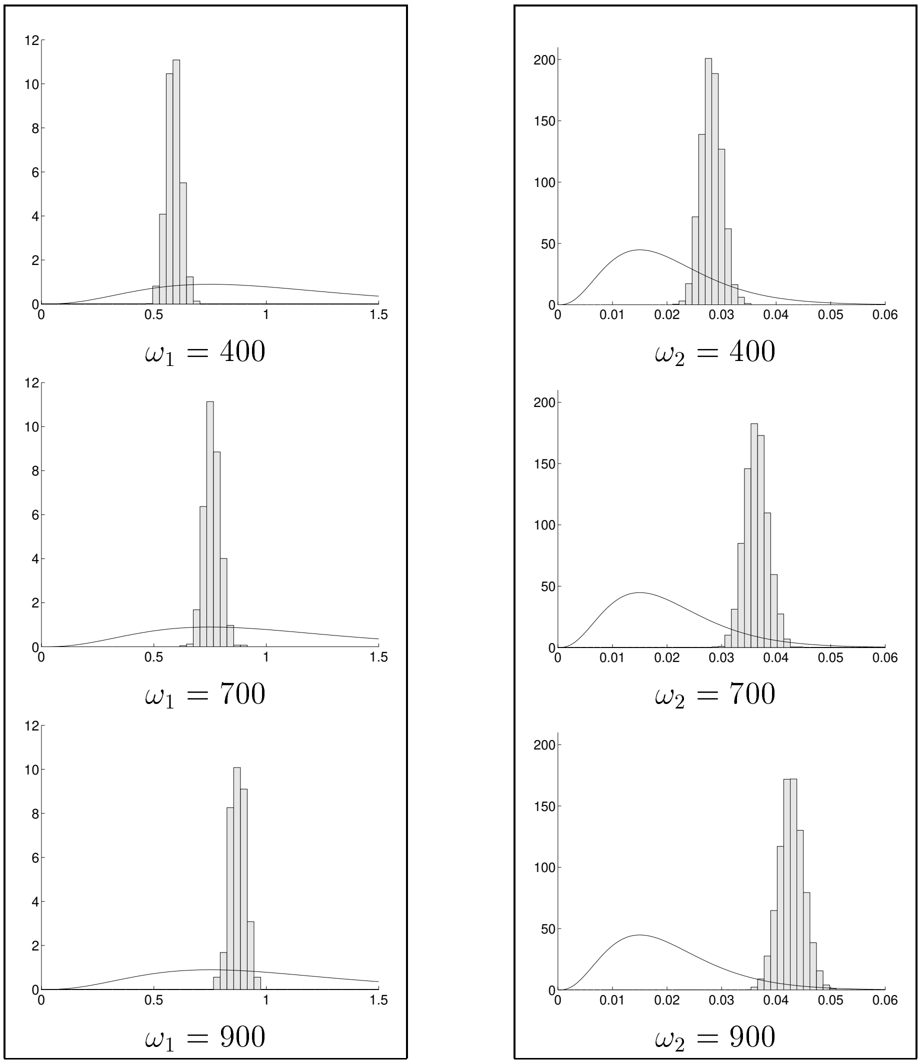
ω_1	ω_2	p_1	$\hat{B}_{12}(\boldsymbol{x}, n)$	Model	$\hat{\theta}_1$	$\bar{k}_1$	$\hat{\theta}_2$	$\bar{k}_2$
700	700	0.1	0.00538	2			0.0364	322
700	700	0.01	$2.40 \cdot 10^8$	1	0.791	583		
700	700	0.001	$9.00 \cdot 10^8$	1	0.805	601		
700	700	0.0001	$1.34 \cdot 10^9$	1	0.799	594		
600	400	0.1	0^*	2			0.0283	235
600	400	0.01	0^*	2			0.0292	245
600	400	0.001	0^*	2			0.0291	243
600	400	0.0001	0^*	2			0.0296	249

Table 4: Results of McMC simulations with $\eta_2 = 20m$, while various parameter values of ω_1, ω_2 and p_1 are used. The estimates of Bayes factor and the model parameters are obtained using Expressions (47) and (49). The estimates of the model parameters and the mean values $\bar{k}_m$ are based on 5000 realizations, separated by 10 iterations of the McMC algorithm. Bayes factor is estimated based on all 50 000 iterations. The reported value 0^* indicates that the estimate of Bayes factor is very small, without being identically 0.

increases, the mode of the posterior pdf $f(\theta_m|\boldsymbol{x})$ also increases. This coincides with the results of Table 2, where the posterior mean is observed to increase with ω_m.

Since the relationship between model parameters in three and two dimensions is known both for the fractal model and the exponential model, see Section 4, the parameters in three dimensions can be estimated based on the corresponding estimates in two dimensions. Thus, in cases where the fractal model is favored, the power law exponent in the displacement distribution in three dimensions is estimated by $\hat{\beta}_3 = \hat{\beta}_2 + 1 = \hat{\theta}_1 + 1$, under the assumption $\tau = 1$, see Expression (23). In the exponential model the distribution of displacement has the same parameter in all dimensions, with estimate $\hat{\lambda} = \hat{\theta}_2$.

When a proper model is chosen for the distribution of displacement in three dimensions, the final goal is to predict the number of faults in a three dimensional reservoir. The procedure is described in Section 6.2. To illustrate the consequence of the model choice on the predicted number of faults in three dimensions, the estimates of the model parameters Θ_m in Table 2b are used, with $\omega_1 = \omega_2 = 720$. This situation gives no conclusion to the model choice problem. The estimates of model parameters and fault numbers in two dimensions are $\hat{\theta}_1 = 0.780$, $\bar{k}_{1,2} = 570$, $\hat{\theta}_2 = 0.0370$ and $\bar{k}_{2,2} = 329$. The constant term in Expression (50) depends on the constants c_1 and c_2 from Expression (1) and the distribution of the dip. A typical dip of 60° is assumed, and c_1 and c_2 are chosen inspired by the work of Walsh and Watterson (1989) and Dawers et al. (1993). The result is a constant term of value 0.035. By the use of Expression (50), the frequencies of faults in three dimensions of displacement above $d_{03} = 2m$, vertically through the reservoir, are found to be $\hat{k}_{1,3} = 4370\, km^{-1}$ and $\hat{k}_{2,3} = 397\, km^{-1}$ under model 1 and 2 respectively. Thus in this case the estimates under the two models differ by a factor 11, illustrating the possible consequences of choosing a



(a) (b)

Figure 14: (a) Histograms of realizations of the model parameter θ_1 in model 1 from $f(\theta_1|\boldsymbol{x}, m_1)$, obtained from McMC simulations. The solid lines show the prior pdf $f_1(\theta_1|m_1)$. (b) Similar histograms of realizations of θ_2 from $f(\theta_2|\boldsymbol{x}, m_2)$, comparing with $f_2(\theta_2|m_2)$. The values of the prior means ω_1 and ω_2 are given below each figure, and $\eta_2 = 20m$ and $p_1 = 0.1$.

wrong model. No boundary effects are accounted for in these calculations.

7.4 Most favorable model based on well information

Hesthammer and Fossen (1997) studied fault intensities in wells in the Gullfaks fields, and found an estimate of 4.4 faults per km in the east-west direction, perpendicular to the dominating strike direction. This number only includes faults of displacement larger than approximately $10m$ and gives a lower limit of fault intensity, since also well observations can suffer from missing observations. This fault intensity can be converted into a number of faults in the two dimensional observation area, which can serve as a prior mean number. The procedure is described in Section 4.5. The simulations above show that the model parameters Θ_1 and Θ_2 typically lies in the intervals $\Theta_1 \in [0.8, 0.9]$ and $\Theta_2 \in [0.025, 0.04]$. Since Θ_1 seems to be less than 1, the mean value $\mathrm{E}_2\left[D|m_1\right]$ needed in Expression (37) does not exist. To deal with this problem, $\mathrm{E}_2\left[D|m\right]$ is replaced by $\mathrm{E}_2\left[D|m, D \leq d_{\max}\right]$, where $d_{\max}$ is chosen larger than the maximum observed displacement value. Assuming a value $c_1 = 75$ of the constant involved in Expressions (1) and (37), and using a rough approximation of $8km \times 12.8km$ for the observation area, see Fossen and Rørnes (1996), the number of faults of displacement above $2m$ are found to be 450–600 under model 1 and 150–250 under model 2. These numbers give an indication of how large the prior mean number of faults under the two models at least should be chosen.

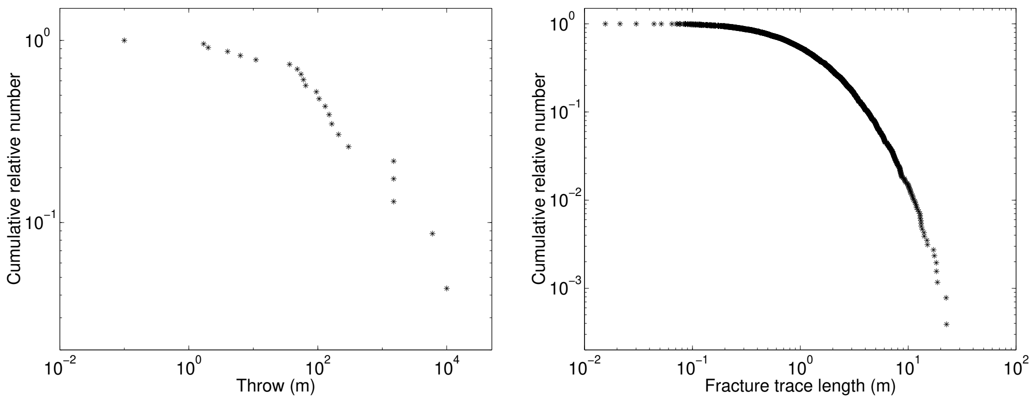
The sensitivity analysis presented above mainly concentrate on prior mean values ω_1 and ω_2 equal to or larger than the values found based on well information. Tables 2–4 show that the fractal model is only considered best in cases where the prior mean values are $\omega_m > 700$ for both m. Such a large value can be realistic for ω_1, but is clearly too large for ω_2 based on the well information. Among the simulations presented above, the ones with prior mean values $\omega_1 = 600$ and $\omega_2 = 400$ are more realistic. In this case the exponential model is clearly most favorable, regardless of the likelihood parameters included in the simulation study.

8 Example: Outcrop Data

Figure 15 shows log-log plots of faults observed in outcrops. Figure 15a shows $n = 23$ throw values obtained in a line sample in western Sinai, see Knott et al. (1996) and Beach et al. (1999). Throw is measured as the vertical component of displacement. Figure 15b shows $n = 2565$ fracture trace lengths observed in a two dimensional observation area in Hornelen, Norway, see Odling (1997). In contrast to faults, fractures have no displacement of rock.

The two models in Expression (48) are used as competing models for the dataset from western Sinai, where θ_1 and θ_2 are now the parameters of the distributions of displacements in one dimension, see Expressions (17) and (33). The common lower limit y_0 is chosen equal to the minimum observation. If the line sample observations from western Sinai are assumed to be exact, with no measurement errors or missing observations, maximum likelihood estimates of the parameters under the two models are given as $\hat{\theta}_1 = 0.151$ and $\hat{\theta}_2 = 0.00105$. Using these estimates as prior means in gamma distributions with shape parameters $\alpha_m = 4$, gives a Bayes factor of $B_{12}(\boldsymbol{x}) = 3.04 \cdot 10^6$. Other reasonable choices of prior mean values also give a large value of Bayes factor, and the fractal model is undoubtedly favored compared to the exponential model.

The observed fracture trace lengths from Hornelen are also assumed to be exact. The distributions of fault width in two dimensions are compared for this dataset, given in Expression (20) for the fractal model and in Expression (30) or (34) under the exponential model. Neither of the distributions from the exponential model give analytically tractable expressions of Bayes factor. If a first order approximation of the modified Bessel function of order 0 is used, see Expression (58), the distribution in Expression (30) can be approximated by a gamma distribution $[Y_j|m_2, \theta_2] \sim \text{Gamma}(1.5, \theta_2)$. In this case Bayes factor is analytically tractable, and reasonable choices of hyper-parameters in the prior pdfs of the model parameters give $B_{12}(\boldsymbol{x}) \approx 0$. Thus in this case the exponential model is far better than the fractal model. Maximum likelihood parameter estimation gives $\hat{\theta}_1 = 0.236$ and $\hat{\theta}_2 = 0.815$. The dataset is one of several datasets gathered at different scales, and Odling (1997) concludes with a log-normal distribution for each single dataset, while a fractal model is found to describe the coalescent fault population. The log-normal distribution is not considered as an alternative distribution in the model choice problem described in the current work, but the same approach as illustrated in this work could be applied to compare this distribution with either the exponential or the fractal model.



(a) (b)

Figure 15: (a) Fault displacements observed in one dimension in western Sinai. (b) Fracture trace lengths observed in two dimensions in Hornelen, Norway.

9 Conclusion

A proper model should be chosen and the model parameters estimated, in order to make inference about a fault population in three dimensions. The estimated number of faults and their sizes depend on which model is used to obtain the estimates. Observation errors should be accounted for in the model choice, and the relationship between the three dimensional fault population and the observations obtained in lower dimensions must be known.

Forward calculations are performed to derive the relationship between the distributions of maximum displacement in three, two and one dimensions, starting with both a fractal and an exponential distribution in three dimensions. Corresponding distributions of fault width are obtained using the relationship $W \propto D^{1/\tau}$. Using some basic assumptions on fault geometry and spatial distribution, it is shown that if the distribution of displacement in three dimensions is fractal, then so is the distribution of fault width and the corresponding distributions in two and one dimensions. The decrease in the power law exponent for the displacement both from three to two dimensions and from two to one is $1/\tau$, while for fault width the power law exponent decreases by 1 from three to two dimensions. The relationships found between power law exponents of displacement or fault width in one, two and three dimensions correspond to previously published results, see Heffer and Bevan (1990), Marrett and Allmendinger (1991), Piggott (1997) and Berkowitz and Adler (1998).

When an exponential distribution of displacement in three dimensions is assumed, the derived expressions of the other distributions are less analytically tractable than the corresponding ones under the fractal model. The expressions simplifies if $\tau = \gamma = 1$, where a linear relationship between width and displacement is assumed and a linear displacement function is applied, see Expressions (1) and (5). In this case the displacement follows an exponential distribution with the same parameter in three, two and one dimensions, thus observations of exponentially distributed displacements do not suffer from over-representation of large faults in lower dimensions. Under the assumption $\tau = 1$ the distribution of the width in three dimensions is also exponential, while in two dimensions the distribution is not exponential.

Section 6 describes a statistical approach for choosing between two suggested models for fault size distributions. The example in Section 7 shows how the conclusion of the model choice depends on the parameters both in the prior distribution of the number of faults and on the likelihood function. Parts of the parameter space is spanned to illustrate how the conclusion varies as the prior model or the assumptions on the sampling errors change. The prior distribution and likelihood function thus have great influence on the model choice, and in order to get the right conclusion the parameters of the prior pdf and the likelihood model should be chosen based on general geological experience and knowledge about the reservoir and the data collection.

If there is a strong belief in a negative power law distribution of fault sizes, based solely on prior experience, it can be tempting to define the prior and likelihood parameters based on the observations. For example, η_2 can be chosen as a break in the log-log curve and ω_1 can be chosen by extrapolating a line fitted to data with $d \geq \eta_2$ to find the total number of faults. This may lead to an optimal choice of parameter values, in the sense that the fractal model is more likely to be favored. However, this procedure uses the observations to define both the prior distribution and the likelihood function, thus reducing the reliability of the statistical analysis.

The prior mean numbers of faults under the two models should be related to the lithology and the tectonic history of the reservoir under study, and can also be calibrated with fault observations from wells. The likelihood functions should be related to the data acquisition. For example, for seismic data the likelihood parameters are related to the seismic resolution, and thus the seismic wavelength. The interpretation process can also have influence on the likelihood parameters, as the detection of small faults can depend both on the interpreter and on the time spent on the interpretation. If specific values for the likelihood parameters are hard to define, then alternatively the parameters can be assigned prior distributions. The statistical analysis presented illustrates the importance of choosing the prior and likelihood parameters correctly. For a given dataset from the Gullfaks field, it is shown how the fault intensity observed in wells is used to determine realistic prior mean number of faults in the observation area under both competing models, and based on these values the exponential model is found to describe the fault displacements displacements better than the fractal model.

Acknowledgements

We thank Noelle Odling at Nansen Environmental and Remote Sensing Center for permission to use data from Hornelen, western Norway. Hilde G. Borgos is funded by a PhD grant from VISTA, a research corporation between The Norwegian Academy of Science and Letters and Den norske stats oljeselskap a.s. (Statoil).

References

Abramowitz, M. and Stegun, I. A. (eds) (1965). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Dover books on advanced mathematics, Dover, New York.

Ackermann, R. V., Withjack, M. O. and Schlische, R. W. (1999). The geometric and statistical evolution of normal fault systems: An experimental study of the effects of mechanical layer thickness on scaling laws, submitted to *Journal of Structural Geology*.

Badley, M. E., Freeman, B., Roberts, A. M., Thatcher, J. S., Walsh, J., Watterson, J. and Yielding, G. (1990). Fault Interpretation During Seismic Interpretation and Reservoir Evaluation, *The Integration of Geology, Geophysics, Petrophysics and Petroleum Engineering in Reservoir Delineation, Description and Management*, Proceedings of the 1st Archie Conference, Houston, Texas, pp. 224–241. AAPG.

Barnett, J. A. M., Mortimer, J., Rippon, J. H., Walsh, J. J. and Watterson, J. (1987). Displacement Geometry in the Volume Containing a Single Normal Fault, *The American Association of Petroleum Geologists Bulletin* **71**(8): 925–937.

Beach, A., Welbon, A. I., Brockbank, P. J. and McCallum, J. E. (1999). Reservoir damage around faults: outcrop from the Suez rift, *Petroleum Geoscience* **5**: 109–116.

Berkowitz, B. and Adler, P. M. (1998). Stereological analysis of fracture network structure in geological formations, *Journal of Geophysical Research* **103**(B7): 15339–15360.

Besag, J., Green, P., Higdon, D. and Mengersen, K. (1995). Bayesian Computation and Stochastic Systems (with discussion), *Statistical Science* **10**(1): 3–66.

Borgos, H. G. and Omre, H. (2000). Sampling algorithm for estimating Bayes factor, *Preprint Statistics 4/2000*, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway.

Carlin, B. P. and Chib, S. (1995). Bayesian Model Choice via Markov Chain Monte Carlo Methods, *J. R. Statist. Soc. B* **57**(3): 473–484.

Childs, C., Walsh, J. J. and Watterson, J. (1990). A Method for Estimation of the Density of Fault Displacements below the Limits of Seismic Resolution in Reservoir Formations, *North Sea Oil and Gas Reservoirs - II*, Graham & Trotman, pp. 309–318.

Chilès, J. P. (1988). Fractal and Geostatistical Methods for Modeling of a Fracture Network, *Mathematical Geology* **20**(6): 631–654.

Clark, R. M., Cox, S. J. D. and Laslett, G. M. (1999). Generalization of power-law distributions applicable to sampled fault-trace lengths: model choice, parameter estimation and caveats, *Geophysical Journal International* **136**: 357–372.

Cowie, P. A. (1998). Normal Fault Growth in Three-Dimensions in Continental and Oceanic Crust, *Faulting and Magmatism at Mid-Ocean Ridges*, Geophysical Monograph 106, the American Geophysical Union, pp. 325–348.

Cowie, P. A. and Scholz, C. H. (1992a). Physical explanation for the displacement-length relationship of faults using a post-yield fracture mechanics model, *Journal of Structural Geology* **14**(10): 1133–1148.

Cowie, P. A. and Scholz, C. H. (1992b). Displacement-length scaling relationship for faults: data synthesis and discussion, *Journal of Structural Geology* **14**(10): 1149–1156.

Cowie, P. A., Malinverno, A., Ryan, W. B. F. and Edwards, M. H. (1994). Quantitative fault studies on the East Pacific Rise: A comparison of sonar image techniques, *Journal of Geophysical Research* **99**(B8): 15205–15218.

Dawers, N. H., Anders, M. H. and Scholz, C. H. (1993). Growth of normal faults: Displacement-length scaling, *Geology* **21**: 1107–1110.

Dawers, N. H. and Anders, M. H. (1995). Displacement-length scaling and fault linkage, *Journal of Structural Geology* **17**(5): 607–614.

Evans, M. and Swartz, T. (1995). Methods for Approximating Integrals in Statistics with Special Emphasis on Bayesian Integration Problems, *Statistical Science* **10**(3): 254–272.

Feder, J. (1988). *Fractals*, Plenum Press, New York.

Fossen, H. and Rørnes, A. (1996). Properties of fault populations in the Gullfaks Field, northern North Sea, *Journal of Structural Geology* **18**(2/3): 179–190.

Gauthier, B. D. M. and Lake, S. D. (1993). Probabilistic Modeling of Faults Below the Limit of Seismic Resolution in Pelican Field, North Sea, Offshore United Kingdom, *The American Association of Petroleum Geologists Bulletin* **77**(5): 761–777.

Han, C. and Carlin, B. P. (2000). MCMC Methods for Computing Bayes Factors: A Comparative Review, *Research Report 2000-001*, Division of Biostatistics, University of Minnesota.

Hatcher, R. D., Jr. (1995). *Structural Geology. Principles, Concepts, and Problems*, 2 edn, Prentice Hall, Englewood Cliffs, New Jersey.

Heffer, K. J. and Bevan, T. G. (1990). Scaling Relationships in Natural Fractures: Data, Theory, and Application, *Europec 90*, Society of Petroleum Engineers, pp. 367–376. SPE Paper 20981.

Hesthammer, J. and Fossen, H. (1997). The influence of seismic noise in structural interpretation of seismic attribute maps, *First Break* **15**(6): 209–219. (C. Townsend, pers.comm.).

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2 edn, Wiley.

Jones, G., Fisher, Q. J. and Knipe, R. J. (eds) (1998). *Faulting, Fault Sealing and Fluid Flow in Hydrocarbon Reservoirs*, number 147 in *Special Publications*, Geological Society, London.

JSG (1996). *Journal of Structural Geology*, Vol. 18, No. 2/3, Special Issue, Scaling Laws for Fault and Fracture Populations – Analyses and Applications.

KASS, R. E. AND RAFTERY, A. E. (1995). Bayes factors, *Journal of the American Statistical Association* **90**(430): 773–795.

KNOTT, S. D., BEACH, A., BROCKBANK, P. J., BROWN, J. L., MCCALLUM, J. E. AND WELBON, A. I. (1996). Spatial and mechanical controls on normal fault populations, *Journal of Structural Geology* **18**(2/3): 359–372.

MANDELBROT, B. B. (1983). *The Fractal Geometry of Nature*, W. H. Freeman and Company, New York.

MARRETT, R. AND ALLMENDINGER, R. W. (1991). Estimates of strain due to brittle faulting: sampling of fault populations, *Journal of Structural Geology* **13**(6): 735–738.

MARRETT, R. AND ALLMENDINGER, R. W. (1992). Amount of extension on "small" faults: An example from the Viking graben, *Geology* **20**: 47–50.

MUNTHE, K., HOLDEN, L., MOSTAD, P. AND TOWNSEND, C. (1994). Modelling Sub-seismic Fault Patterns Using a Marked Point Process, *Proceedings of the 4th European conference on the Mathematics of Oil Recovery, Røros, Norway.* Topic B: Heterogeneity Description and Assessment of Uncertainty.

MUNTHE, K. L., OMRE, H., HOLDEN, L., DAMSLETH, E., HEFFER, K., OLSEN, T. S. AND WATTERSON, J. (1993). Subseismic Faults in Reservoir Description and Simulation, *68th Annual Conference and Exhibition of the Society of Petroleum Engineers*, Society of Petroleum Engineers, pp. 843–850. SPE Paper 26500.

OAKESHOTT, R. B. S. AND EDWARDS, S. F. (1992). On the stereology of ellipsoids and cylinders, *Physica A* **189**(1/2): 208–233.

ODLING, N. E. (1997). Scaling and connectivity of joint systems in sandstones from western norway, *Journal of Structural Geology* **19**(10): 1257–1271.

PICKERING, G., BULL, J. M. AND SANDERSON, D. J. (1995). Sampling power-law distributions, *Tectonophysics* **248**: 1–20.

PICKERING, G., BULL, J. M. AND SANDERSON, D. J. (1996). Scaling of fault displacements and implications for the estimation of sub-seismic strain, *in* P. C. Buchanan and D. A. Nieuwland (eds), *Modern Developments in Structural Interpretation, Validation and Modeling*, Geological Society Special Publication No. 99, pp. 11–26.

PIGGOTT, A. R. (1997). Fractal relations for the diameter and trace length of disc-shaped fractures, *Journal of Geophysical Research* **102**(B8): 18121–18125.

SASSI, W., LIVERA, S. E. AND CALINE, B. P. R. (1992). Reservoir compartmentation by faults in Cormorant Block IV, U.K. northern North Sea, *in* R. M. Larsen, H. Brekke, B. T. Larsen and E. Talleraas (eds), *Structural and Tectonic Modelling and its Application to Petroleum Geology*, NPF Special Publication 1, Elsevier, Amsterdam, pp. 355–364.

Spryopoulos, C., Griffith, W. J., Scholz, C. H. and Shaw, B. E. (1999). Experimental evidence for different strain regimes of crack populations in a clay model, *Geophysical Research Letters* **26**(8): 1081–1084.

Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields*, Wiley, New York.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and its Applications*, 2 edn, Wiley, New York.

Walsh, J. J. and Watterson, J. (1987). Distribution of cumulative displacement and seismic slip on a single normal fault surface, *Journal of Structural Geology* **9**(8): 1039–1046.

Walsh, J. J. and Watterson, J. (1988). Analysis of the relationship between displacements and dimensions of faults, *Journal of Structural Geology* **10**(3): 239–247.

Walsh, J. J. and Watterson, J. (1989). Displacement gradients on fault surfaces, *Journal of Structural Geology* **11**(3): 307–316.

Walsh, J. J., Watterson, J. and Yielding, G. (1991). The importance of small-scale faulting in regional extension, *Nature* **351**: 391–393.

Walsh, J. J., Watterson, J. and Yielding, G. (1994). Determination and Interpretation of Fault Size Populations: Procedures and Problems, *North Sea Oil and Gas Reservoirs - III*, Norwegian Institute of Technology, Kluwer Academic Publishers, pp. 141–155.

Watterson, J. (1986). Fault dimensions, displacement and growth, *Pure and Applied Geophysics* **124**: 365–373.

Yielding, G., Freeman, B. and Needham, D. T. (1997). Quantitative fault seal prediction, *The American Association of Petroleum Geologists Bulletin* **81**(6): 897–917.

A Approximation of Fault Heights

The height of a fault ellipsoid can be expressed as a function $h(x) = Kx^{-1}\sqrt{1+ax^2}$, omitting the dependence on the dip θ, with

$$K = \sin\theta c_2^{-1}, \quad a = \left(\frac{c_2 c_3}{\tan\theta}\right)^2, \quad x = w^{-1} = (dc_1)^{-1/\tau}. \tag{53}$$

The function $h(x)$ can further be expressed as $h(x) = Kx^{-1}g(x)$, where $g(x) = \sqrt{1+ax^2}$. A Taylor expansion of $g(x)$ gives

$$g(x) = 1 + \frac{ax^2}{2(1+a\xi^2)^{3/2}} = 1 + R(x), \qquad \text{for some } \xi \in [0, x].$$

Since $a\xi^2 \geq 0$ it follows that $R(x) \leq ax^2/2$. Let

$$\hat{h}(x) = Kx^{-1} = \sin\theta(xc_2)^{-1}.$$

Substituting w or $(dc_1)^{1/\tau}$ for x, Expressions (2) and (3) are obtained respectively. The error term of the approximation is $h(x) - \hat{h}(x) = Kx^{-1}g(x) - Kx^{-1} = Kx^{-1}R(x)$, and the absolute error has an upper bound

$$\frac{h(x) - \hat{h}(x)}{h(x)} = \frac{Kx^{-1}R(x)}{Kx^{-1}g(x)} \leq \frac{ax^2}{2\sqrt{1+ax^2}} = \left(\frac{c_2c_3}{\tan\theta}\right)^2 \frac{x^2}{2}\left(1 + \left(\frac{c_2c_3}{\tan\theta}\right)^2 x^2\right)^{-1/2}.$$

This function can be shown to be monotonely increasing, thus substituting w or d from Expression (53), the relative error decreases monotonely for increasing fault size.

B Fault Distributions

B.1 Fractal distributions: displacement

The pdf of the maximum displacement D in three dimensions given in Expression (14) is:

$$f_3(d) = \frac{\beta_3 d_{03}^{\beta_3}}{d^{\beta_3+1}}, \qquad d \geq d_{03}.$$

Let D' be the displacement obtained in two dimensions. Using the approximation of the height of the fault ellipsoid given in (3) and the displacement function (5), the pdfs corresponding to Expressions (7) and (8) are found to be:

$$\begin{aligned} f(d|1) &= \frac{(\beta_3 - 1/\tau)d_{03}^{\beta_3-1/\tau}}{d^{\beta_3-1/\tau+1}} \\ f(d'|d,1) &= \frac{(d')^{1/\gamma-1}}{\gamma d^{1/\gamma}}, \quad 0 \leq d' \leq d \\ f(d'|1) &= \int_{d'}^{\infty} \frac{(d')^{1/\gamma-1}}{\gamma d^{1/\gamma}} \cdot \frac{(\beta_3 - 1/\tau)d_{03}^{\beta_3-1/\tau}}{d^{\beta_3-1/\tau+1}}\, dd \\ &= \frac{\beta_3 - 1/\tau}{\gamma(\beta_3 - 1/\tau)+1} \cdot \frac{d_{03}^{\beta_3-1/\tau}}{(d')^{\beta_3-1/\tau+1}}, \quad d' \geq d_{03}. \end{aligned} \tag{54}$$

The last expression is equal to the pdf in Expression (15), omitting conditioning on $I = 1$.

B.2 Fractal distributions: width

The pdf of the width W in three dimensions given in Expression (18) is

$$f_3(w) = \frac{\mu_3 w_{03}^{\mu_3}}{w^{\mu_3+1}}, \qquad w \geq w_{03}.$$

Let W' be the fault width obtained in two dimensions. Using the approximation of the height of the fault ellipsoid given in (2) and the relationship (4) between w' and w at a given position of intersection, the pdfs corresponding to Expressions (7) and (8) are found to be:

$$\begin{aligned} f(w|1) &= \frac{(\mu_3 - 1) w_{03}^{\mu_3 - 1}}{w^{\mu_3}} \\ f(w'|w,1) &= \frac{w'}{w\sqrt{w^2 - (w')^2}} \qquad (55) \\ f(w'|1) &= \int_{w'}^{\infty} \frac{w'}{w\sqrt{w^2 - (w')^2}} \cdot \frac{(\mu_3 - 1) w_{03}^{\mu_3 - 1}}{w^{\mu_3}} \, dw. \qquad (56) \end{aligned}$$

The latter integral can be solved using the substitution $w = w'/\cos u$, which gives:

$$f(w'|1) = \frac{w'(\mu_3 - 1) w_{03}^{\mu_3 - 1}}{(w')^{\mu_3+1}} \int_0^{\pi/2} (\cos u)^{\mu_3} \, du = \frac{w_{03}^{\mu_3 - 1}(\mu_3 - 1)}{(w')^{\mu_3}} \cdot \frac{\Gamma((\mu_3 + 1)/2)}{\Gamma((\mu_3 + 2)/2)} \cdot \frac{\sqrt{\pi}}{2}.$$

This is the pdf in Expression (19), omitting conditioning on $I = 1$.

B.3 Exponential distributions: displacement

The pdf of the maximum displacement D in three dimensions given in Expression (24) is

$$f_3(d) = \lambda \exp(-\lambda d), \qquad d \geq 0.$$

Let D' be the displacement obtained in two dimensions. Using the approximation of the height of the fault ellipsoid given in (3), the pdf corresponding to Expression (7) is found to be:

$$f(d|1) = \frac{\lambda^{1+1/\tau}}{\Gamma(1 + 1/\tau)} d^{1/\tau} \exp(-\lambda d).$$

This is the pdf of a gamma distribution. The function $f(d'|d,1)$ is the same as in the fractal case, given in Expression (54). The integral in Expression (8) gives the pdf of the displacement in two dimensions:

$$f(d'|1) = \gamma^{-1} \lambda^{1/\gamma} (d')^{1/\gamma - 1} \frac{\Gamma(1/\tau - 1/\gamma + 1, \lambda d')}{\Gamma(1 + 1/\tau)}, \qquad \gamma > \frac{\tau}{1 + \tau}.$$

Substituting D' by D and $f(d')$ by $f_2(d)$ and omitting conditioning on $I = 1$, Expression (25) is obtained.

B.4 Exponential distributions: width

The pdf of the width W in three dimensions given in Expression (28) is

$$f_3(w) = \rho\tau w^{\tau-1}\exp(-\rho w^{\tau}), \qquad w \geq 0.$$

Let W' be the width obtained in two dimensions. Using the approximation of the height of the fault ellipsoid given in (2), the pdf corresponding to Expression (7) is found to be:

$$f(w|1) = \frac{\tau\rho^{1+1/\tau}}{\Gamma(1+1/\tau)} w^{\tau}\exp(-\rho w^{\tau}),$$

while $f(w'|w,1)$ is the same as in the fractal case, Expression (55). The integral in Expression (8) gives the pdf of the width in two dimensions:

$$f(w'|1) = \frac{\tau\rho^{1+1/\tau}}{\Gamma(1+1/\tau)} w' \int_{w'}^{\infty} \frac{w^{\tau-1}}{\sqrt{w^2-(w')^2}}\exp(-\rho w^{\tau})\,dw, \tag{57}$$

which is the pdf given in (29). For $\tau = 1$ the integral becomes

$$f(w'|1) = \int_{w'}^{\infty} \frac{w'}{w\sqrt{w^2-(w')^2}}\rho^2 w\exp(-\rho w)\,dw = \rho^2 w' \int_{w'}^{\infty} \frac{1}{\sqrt{w^2-(w')^2}}\exp(-\rho w)\,dw.$$

Using the substitution $w = w'z$ the integral can be written as

$$f(w'|1) = \rho^2 w' \int_{1}^{\infty} \frac{1}{\sqrt{z^2-1}}\exp(-\rho w' z)\,dz = \rho^2 w' K_0(\rho w'),$$

where $K_0(\cdot)$ is the Bessel function described in Appendix C. Substituting W' by W and $f(w')$ by $f_2(w)$ and omitting conditioning on $I = 1$, Expression (30) is obtained.

C Special Functions

The gamma function, incomplete gamma function and the modified Bessel function are described below, see for example Abramowitz and Stegun (1965).

C.1 Gamma function and incomplete gamma function

The gamma function $\Gamma(\cdot)$ and the incomplete gamma function $\Gamma(\cdot,\cdot)$ is given by the integrals

$$\Gamma(n) = \int_0^{\infty} t^{n-1}e^{-t}\,dt, \qquad \Gamma(n,a) = \int_a^{\infty} t^{n-1}e^{-t}\,dt.$$

C.2 Modified Bessel function

The modified Bessel function of order ν, $K_\nu(\cdot)$, is the solution of the differential equation

$$t^2 \frac{d^2 w}{dt^2} + t\frac{dw}{dt} - (t^2 + \nu^2) w = 0,$$

and can be expressed by the integral

$$K_\nu(t) = \frac{\sqrt{\pi}(t/2)^\nu}{\Gamma(\nu + 1/2)} \int_1^\infty (z^2 - 1)^{\nu - 1/2} e^{-tz}\, dz.$$

For $\nu = 0, 1$ the expression simplifies to

$$\begin{aligned} K_0(t) &= \int_1^\infty \frac{1}{\sqrt{z^2 - 1}} \cdot e^{-tz}\, dz, \\ K_1(t) &= t \int_1^\infty \sqrt{z^2 - 1} \cdot e^{-tz}\, dz. \end{aligned}$$

For large $|t|$, the modified Bessel function of order 0 can be approximated as

$$K_0(t) \approx \sqrt{\frac{\pi}{2t}} e^{-t} \left(1 + \frac{(4-1)}{8t} + \frac{(4-1)(4-9)}{2!(8t)^2} + \frac{(4-1)(4-9)(4-25)}{3!(8t)^3} + \cdots \right) \qquad (58)$$

see Abramowitz and Stegun (1965).