Journal of Structural Geology 32 (2010) 1978–1995

Contents lists available at ScienceDirect

Journal of Structural Geology

journal homepage: www.elsevier.com/locate/jsg

Determining brittle extension and shear strain using fault length and displacement systematics: Part II: Data evaluation and test of the theory

Robert J. Twiss [a,*], Randall Marrett [b]

[a] Geology Department, University of California at Davis, One Shields Ave., Davis, CA 95616-8605, USA
[b] Department of Geological Sciences, Jackson School of Geosciences, University of Texas at Austin, 1 University Station C1100, Austin, TX 78712-0254, USA

ARTICLE INFO

Article history:
Received 25 April 2009
Received in revised form
2 April 2010
Accepted 12 April 2010
Available online 25 May 2010

Keywords:
Brittle deformation
Strain
Extension
Shear strain
Fault systematics
Fault scaling

ABSTRACT

We use the theoretical relations developed in Part I of this work to evaluate the self-consistency of fault-length and fault-displacement data gathered in domains of one and two dimensions from the Yucca Mountain area and from the coalfields in south Yorkshire, U.K. These data sets are not all self-consistent. For the Yucca Mt. area, the theory shows that, the volume over which the sampling of the faults must occur should have a horizontal width no smaller than 2.4 times the horizontal length of the largest fault, and a depth no smaller than 1.6 times the vertical extent of the largest vertical-equivalent-fault. It also shows that the volumetric extension must be $\geq$95% of the extension of a two-dimensional domain and $\geq$80% of the extension of a one-dimensional domain. The theory successfully accounts for the observed cumulative extensional strain derived from fault-displacement data from a one-dimensional sampling domain at Yucca Mt., Nevada, U.S.A. Faults up to about four orders of magnitude smaller than the largest fault make a significant contribution to the strain. The most robust calculation of cumulative fractional strain requires the parameters inferred from sampling displacement in a one-dimensional domain. This sampling procedure therefore provides the most reliable results.

1. Introduction

In the companion to this paper (Twiss and Marrett, in this issue, referred to as Part I), we develop the theory by which we can infer the extensional and shear strain in one-, two-, and three-dimensional domains that have been deformed by brittle faulting. The inference requires a knowledge of an independent set of parameters that define the fault systematics and the measurements of either the length of, or displacement on, the largest fault in the domain.

In this paper (Part II), we apply the results of the theory to examine the fault-length and fault-displacement distributions from detailed one- and two-dimensional sampling of the Yucca Mountain area of southern Nevada, U.S.A. and from the coalfields of southern Yorkshire, U.K. (Watterson et al., 1996). We review the results of these studies in Section 2.

From the theory, we can calculate constraints on how large the sampling domain must be relative to the largest fault in the sample in order that we may include all the measured faults in this type of strain analysis. Conversely, we can calculate constraints on the size of the largest fault that can be included in a strain analysis of this type for a given size of sampling domain. We illustrate these constraints for the Yucca Mountain and south Yorkshire areas in Section 3.

We then use the results of the theory to test the self-consistency of the different data sets from these two areas (Section 4). The theory provides equations that relate the parameters in the equations that define the fault-length vs. fault-displacement systematics, and that define the cumulative frequencies of faults as a function of length and displacement for one-, two-, and three-dimensional domains. These equations permit quantitative tests for consistency among different data sets from the same area. We show that the different data sets from the same area commonly are not self-consistent, and that at least for the Yucca Mountain area, the most reliable measurements derive from one-dimensional sampling of fault-displacement. We conclude that some sampling bias distorts the results from two-dimensional sampling.

Determination of strain in one- and two-dimensional domains does not permit an exact determination of the three-dimensional strain, and the theory lets us calculate the lower bounds that such measurements impose on the volumetric strain. In Section 5 we illustrate these constraints for the Yucca Mountain area.

We then test the theory against a data set from a detailed one-dimensional sampling of the Yucca Mountain area and show that it successfully predicts the fractional cumulative extensional strain as a function of fault-displacement (Section 6).

* Corresponding author. Tel.: +1 530 752 0352/756 3326; fax: +1 530 752 0951.
E-mail addresses: rjtwiss@ucdavis.edu (R.J. Twiss), marrett@mail.utexas.edu (R. Marrett).

0191-8141/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsg.2010.04.006

In Section 7, we calculate for both these areas, the strains contributed by the largest fault and compare that with the total strains. For the Yucca Mountain area, we calculate the cumulative strain as a function of displacement and show that faults up to several orders of magnitude smaller than the largest fault in the area contribute significant amounts to the total strain.

For ease of reference, we collect in Table 1 the equations from Part I that are referenced in Part II.

2. Empirical determination of parameters

The equations we have derived in Part I provide specific relations among the parameters that characterize the power-law relationship between fault-length and fault-displacement, and the frequency distributions for both of these fault attributes. We also derived specific relations among the parameters in the equations describing the frequency distributions for sampling domains having different numbers of dimensions. These equations can be applied to infer strain accurately only if we know the values for these parameters and if the measurement techniques used to evaluate the parameters do not include any systematic biases that would limit the accuracy of the sampling. In this section we review two studies that provide some of the best available data for determining the parameters in the equations for fault systematics and for testing the self-consistency of measurements made on fault-length and fault-displacement in one- and two-dimensional domains.

2.1. Evaluation of the parameter p

Fig. 1 shows a log–log plot of fault-length vs. fault-displacement for nine different data sets from Clark and Cox (1996); it also includes one set from the Yucca Mountain area of southern Nevada, U.S.A. (compiled by Simonds et al. (1995); symbols labeled 'Y' in legend), and one set for the northeast-striking faults in the East Pennine coalfield of south Yorkshire, U.K from Watterson et al. (1996, their fig. 13f; symbols labeled 'W' in legend). We discuss the latter two data sets separately below.

Clark and Cox (1996) performed a statistical evaluation of the parameters in Eq. (I:3.1.2) relating fault-length to fault-displacement for nine different data sets from different areas. They concluded that individually, the data sets have slopes p in Eq. $(\text{I:3.1.2})_2$ that, at the 95% confidence level, are statistically indistinguishable from

$$p = 1.0 \tag{2.1}$$

(Fig. 1, Table 2, Clark and Cox, 1996, their table 1). Taken all together, the data span a range of more than six orders of magnitude in both length and displacement, although individually, the sets each span a range of at most only about 2 orders of magnitude. The best-fits to the individual sets of data actually give different values of p ranging from 0.63 to 1.4 (Clark and Cox, 1996, their table 1). Yielding et al. (1996), in their review of subsurface data, find values between 1.0 and about 2.0, with most results lying between 1.0 and 1.4 (their table 1 and Fig. 10C). For most of the individual data sets reviewed by Clark and Cox (1996), however, the span of values and the scatter of the data points are such that the slope is poorly defined, and the fact that a slope of $p = 1.0$ provides a good fit to each of the data sets individually, and to all of the data collectively suggests that this is a reasonable value to use in a general application of Eq. (I:3.1.2) to faults. According to Cowie and Scholz (1992), the value of p in Eq. (2.1) is expected theoretically for faults that originate by a fracture mechanics mechanism with residual friction.

Analysis of the nine data sets combined, also shows that the data cannot be fit well with a common intercept ($-\log B$); the range of

Table 1
Equations from Part I referenced in Part II.

Equations from Part I	Part I, Eq. No.	Sup. Mat., Eq. No.
$A^{(i)} = (\lambda^{(i)}L^{(i)})L^{(i)} = \lambda^{(i)}(L^{(i)})^2$	(I:2.1.13)	
$\lambda^{(i)} \equiv \dfrac{A^{(i)}}{(L^{(i)})^2}$		
$\lambda^{(i)} = 1\dfrac{l^{(i)}}{L^{(i)}}$ or $\lambda^{(i)} = \dfrac{\pi}{4}\dfrac{l^{(i)}}{L^{(i)}}$	(I:2.1.14)	
$L^p = B\delta$	(I:3.1.1)	
$p \log L = \log B + \log \delta$		
$\delta = \dfrac{1}{B}L^p$	(I:3.1.2)	
$\log \delta = -\log B + p \log L$		
$N_{(\zeta,v)}(L) = G_{(\zeta,v)}L^{-m_\zeta}$	(I:3.1.3)	
$\log N_{(\zeta,v)}(L) = \log G_{(\zeta,v)} - m_\zeta \log L$		
$f_{(\zeta,v)}(L) = g_{(\zeta,v)}L^{-m_\zeta}$	(I:3.1.4)	
$\log f_{(\zeta,v)}(L) = \log g_{(\zeta,v)} - m_\zeta \log L$		
$f_{(\zeta,v)}(L) \equiv N_{(\zeta,v)}(L)/\Delta_{(\zeta,v)}$	(I:3.1.5)	
$g_{(\zeta,v)} \equiv G_{(\zeta,v)}/\Delta_{(\zeta,v)}$		
$\Delta_{(3)} \equiv \mathcal{V} = \mathcal{TWH}$	(I:3.1.6)	
$\Delta_{(2,\perp h)} \equiv \mathcal{A}_h = \mathcal{TW}$		
$\Delta_{(1,v)} \equiv \left\{ \begin{matrix} \mathcal{T} & \text{for } v = t \\ \mathcal{W} & \text{for } v = w \end{matrix} \right\}$		
$N_{(\zeta,v)}(\delta) = G_{(\zeta,v)}(B\delta)^{-m_\zeta/p}$	(I:3.1.7)	
$\log N_{(\zeta,v)}(\delta) = \log(G_{(\zeta,v)}B^{-m_\zeta/p}) - \dfrac{m_\zeta}{p}\log\delta$		
$f_{(\zeta,v)}(\delta) = g_{(\zeta,v)}(B\delta)^{-m_\zeta/p}$	(I:3.1.8)	
$\log f_{(\zeta,v)}(\delta) = \log(g_{(\zeta,v)}B^{-m_\zeta/p}) - \dfrac{m_\zeta}{p}\log\delta$		
$s_\zeta \equiv \dfrac{m_\zeta}{p}$	(I:3.1.9)	
$R_{(\zeta,v)} \equiv G_{(\zeta,v)}B^{-s_\zeta} = G_{(\zeta,v)}B^{-m_\zeta/p}$		
$r_{(\zeta,v)} \equiv g_{(\zeta,v)}B^{-s_\zeta} = g_{(\zeta,v)}B^{-m_\zeta/p}$		
$N_{(\zeta,v)}(\delta) = R_{(\zeta,v)}\delta^{-s_\zeta}$	(I:3.1.10)	
$\log N_{(\zeta,v)}(\delta) = \log R_{(\zeta,v)} - s_\zeta \log\delta$		
$f_{(\zeta,v)}(\delta) = r_{(\zeta,v)}\delta^{-s_\zeta}$	(I:3.1.11)	
$\log f_{(\zeta,v)}(\delta) = \log r_{(\zeta,v)} - s_\zeta \log\delta$		
$r_{(\zeta,v)} = R_{(\zeta,v)}/\Delta_{(\zeta,v)}$	(I:3.1.12)	
$l^{(i)} = \mu^{(i)}L^{(i)}$	(I:3.1.21)	
$\mu^{(i)} \equiv \dfrac{l^{(i)}}{L^{(i)}}$		
$m_1 = m_3 - 2$	(I:3.2.2)	
$g_{(1,t)} = g_{(3)}\lambda \cos\theta \dfrac{m_3}{(m_3-2)}$	(I:3.2.3)	
$m_2 = m_3 - 1$	(I:3.2.4)	
$g_{(2,\perp h)} = g_{(3)}\mu \sin\rho \dfrac{m_3}{m_3-1}$	(I:3.2.5)	
$s_1 = s_3 - \dfrac{2}{p}$	(I:3.2.6)	
$r_{(1,t)} = r_{(3)}\lambda B^{2/p} \cos\theta \dfrac{s_3}{(s_3-2/p)}$	(I:3.2.7)	
$s_2 = s_3 - \dfrac{1}{p}$	(I:3.2.8)	
$r_{(2,\perp h)} = r_{(3)}\mu B^{1/p} \sin\rho \dfrac{s_3}{s_3-1/p}$	(I:3.2.9)	
$s_1 = \dfrac{m_1}{p} = \dfrac{m_2-1}{p} = \dfrac{m_3-2}{p}$	(I:3.2.10)	
$s_2 = \dfrac{m_2}{p} = \dfrac{m_3-1}{p}$		
$m_3 = s_1 p + 2 = s_2 p + 1$	(I:3.2.11)	
$r_{(1,t)} = g_{(3)}B^{-(m_3-2)/p}\lambda \cos\theta \dfrac{m_3}{(m_3-2)}$	(I:3.2.12)	
$r_{(2,\perp h)} = g_{(3)}\mu B^{-(m_3-1)/p} \sin\rho \dfrac{m_3}{m_3-1}$	(I:3.2.13)	
$\delta^{(\max)}_{(1,t)} \le \delta^{(\max)}_{(3)}$ so ${}_3\mathfrak{P}^{(1)}_t \dfrac{s_3}{s_1} \le 1$ and ${}_3\mathfrak{P}^{(1)}_t \le \dfrac{s_1}{s_3}$	(I:3.3.8)	
$\mathcal{WH} \ge \dfrac{s_3}{s_1}\lambda(L^{(1)})^2 \cos\theta$	(I:3.4.2)	
$\mathcal{H} \ge \dfrac{s_3}{s_2}\mu L^{(1)} \sin\rho$		

(continued on next page)

Table 1 (*continued*)

Equations from Part I	Part I, Eq. No.	Sup. Mat., Eq. No.
$\mathcal{W} \geq \frac{s_2}{s_1}\frac{\lambda}{\mu}\frac{\cos\theta}{\sin\rho}L^{(1)}$	(I:3.4.3)	
for $\frac{\lambda}{\mu}=1,\ \theta=0^\circ,$ and $\rho=90^\circ$	(I:3.4.4)	
$\mathcal{W} \geq \frac{s_2}{s_1}L^{(1)}$		
$\mathcal{H} \geq \frac{s_3}{s_2}\mu L^{(1)}$		
for $\frac{\lambda}{\mu}=1,\ \theta=0^\circ,$ and $\rho=90^\circ$	(I:3.4.5)	
$\mathcal{W} \geq \frac{s_2}{s_1}B^{\frac{1}{p}}(\delta_{(3)}^{(\max)})^{\frac{1}{p}}$		
$\mathcal{H} \geq \frac{s_3}{s_2}\mu B^{\frac{1}{p}}(\delta_{(3)}^{(\max)})^{\frac{1}{p}}$		
$e_{(3,t)}^{(i)} = {}_3\mathfrak{R}_t^{(i)}\left[\frac{\delta^{(i)}\cos\phi}{\mathcal{T}}\right]$	(I:4.1.1)	(S1.1.4)
$e_{(1,t)}^{(i)} = {}_1\mathfrak{R}_t^{(i)}\left[\frac{\delta^{(i)}\cos\phi}{\mathcal{T}}\right]$	(I:4.1.3)	(S3.1.3)
${}_3\mathfrak{R}_t^{(i)} = \frac{\lambda B^{2/p}\cos\theta}{\mathcal{A}_t}(\delta^{(i)})^{2/p}$	(I:4.1.7), [also (I:3.1.22)$_1$]	(S1.1.3)$_2$
${}_3\mathfrak{R}_w^{(i)} = \frac{\lambda B^{2/p}\cos\alpha}{\mathcal{A}_w}(\delta^{(i)})^{2/p}$		
${}_1\mathfrak{R}_t^{(i)} = 1$	(I:4.1.9)	(S3.1.5)
${}_1\mathfrak{R}_w^{(i)} = 1$		
$e_{(1,t)}^{(i)} = \left[\frac{\cos\phi}{\mathcal{T}}\right]\delta^{(i)}$	(I:4.1.13)	(S3.1.6)
$e_{(1,\gamma)}^{(i)} = \frac{1}{2}\left\{\left[\frac{\cos\beta}{\mathcal{T}}\right]\delta^{(i)} + \left[\frac{\cos\phi}{\mathcal{W}}\right]\delta^{(i)}\right\}$		
$e_{(3,x)}^{(\mathrm{tot})} = e_{(3,x)}^{(1)}\left(\frac{1+2/p}{1-(s_3-2/p)}\right)$	(I:4.1.40)	(S1.21)
$e_{(2,x)}^{(\mathrm{tot})} = e_{(2,x)}^{(1)}\left(\frac{1+1/p}{1-(s_2-1/p)}\right)$	(I:4.1.41)	(S2.3.10)
$e_{(1,t)}^{(\mathrm{tot})} = e_{(1,t)}^{(1)}\left[\frac{1}{1-s_1}\right]$	(I:4.1.42)	(S3.4.9), (S3.4.22), (S3.4.23)
$e_{(1,\gamma)}^{(\mathrm{tot})} = 0.5\sum_{v=t,w} e_{(1,\gamma v)}^{(1)}\frac{1}{(1-s_1)} = e_{(1,\gamma)}^{(1)}\frac{1}{(1-s_1)}$		
$\frac{e_{(3,x)}^{(1)}}{e_{(3,x)}^{(\mathrm{tot})}} = \frac{p+2-s_3p}{p+2} = \frac{p+2-m_3}{p+2}$	(I:4.1.43)	
$0 \leq s_3 < 1+2/p$ $0 \leq s_2 < 1+1/p$ $0 \leq s_1 < 1$	(I:4.1.45)	
$\frac{e_{(3,x)}^{(\mathrm{cum})}(\delta)}{e_{(3,x)}^{(\mathrm{tot})}} = 1-\frac{s_3p}{p+2}\left(\frac{\delta}{\delta_{(3)}^{(\max)}}\right)^{[1-(s_3-2/p)]}$	(I:4.1.48)	(S1.23)$_1$
$\frac{e_{(2,x)}^{(\mathrm{cum})}(\delta)}{e_{(2,x)}^{(\mathrm{tot})}} = 1-\frac{s_2p}{p+1}\left(\frac{\delta}{\delta_{(2,\perp h)}^{(\max)}}\right)^{[1-(s_2-1/p)]}$	(I:4.1.49)	(S2.3.12)$_1$
$\frac{e_{(1,t)}^{(\mathrm{cum})}(\delta)}{e_{(1,t)}^{(\mathrm{tot})}} = 1-s_1\left[\frac{\delta}{\delta_{(1,t)}^{(\max)}}\right]^{(1-s_1)}$	(I:4.1.50)	(S3.4.10), (S3.4.25)
$\frac{e_{(1,\gamma)}^{(\mathrm{cum})}(\delta)}{e_{(1,\gamma)}^{(\mathrm{tot})}} = 1-\frac{0.5}{e_{(1,\gamma)}^{(1)}}\sum_{v=t,w} e_{(1,\gamma v)}^{(1)}\left(s_1\left[\frac{\delta}{\delta_{(1,v)}^{(\max)}}\right]^{(1-s_1)}\right)$		
$\frac{e_{(2,x)}^{(\mathrm{tot})}}{e_{(3,x)}^{(\mathrm{tot})}} \leq \left[\frac{s_3}{s_2}\right]\left(\frac{p+1}{p+2}\right)$	(I:4.2.2)	(S4.2.16)
$\frac{e_{(1,t)}^{(\mathrm{tot})}}{e_{(3,t)}^{(\mathrm{tot})}} \leq \left[\frac{s_3}{s_1}\right]\left(\frac{p}{p+2}\right)$	(I:4.3.2)	(S4.3.4)
$\frac{e_{(1,\gamma)}^{(\mathrm{tot})}}{e_{(3,\gamma)}^{(\mathrm{tot})}} \leq \left[\frac{s_3}{s_1}\right]^{1/s_1}\left(\frac{p}{p+2}\right)$	(I:4.3.4)	(S4.4.8)

See Part I, Table 1 for definitions of symbols.

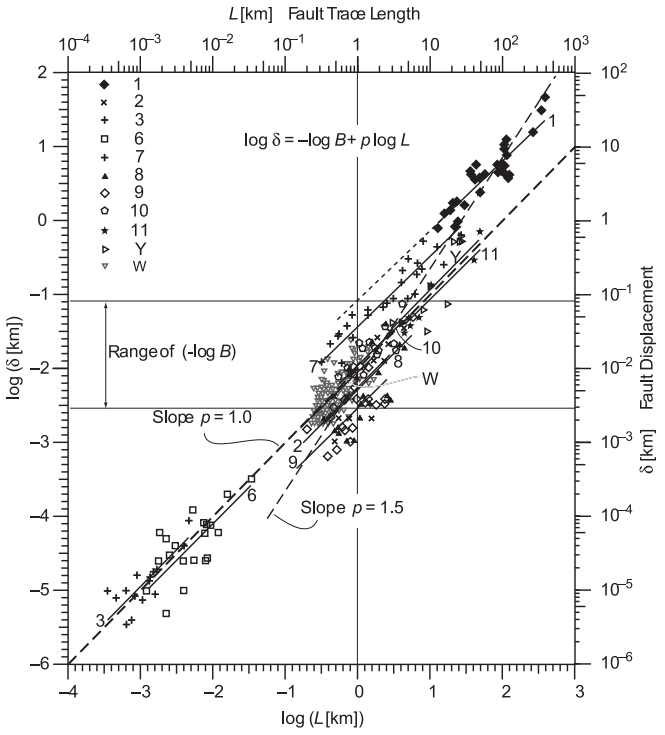

Fig. 1. Relationship between the logarithms of fault length [km] and displacement [km] determined for different sets of faults. Data plotted with the symbols identified in the legend by numbers are from Clark and Cox (1996, fig. 1): 1. Elliott, 1976; 2. Krantz (1988); 3. Muraoka and Kamata (1983); 6. Peacock and Sanderson (1991); 7. Villemin et al. (1995); 8. Walsh and Watterson (1987); 9. North Derby (Watterson, 1986); 10. Barnsley (Watterson, 1986); 11. Mid-ocean (Watterson, 1986). Data plotted with the symbols labeled 'Y' in the legend are for the Yucca Mountain area of southern Nevada, U.S.A. and were measured from the map of Simonds et al. (1995) (see Fig. 2A). Data plotted with the symbols labeled 'W' in the legend are for NE-striking faults in the coalfields in southern Yorkshire, U.K. from Watterson et al. (1996, their fig. 13f); these data are for those faults with tip-to-tip measured lengths and with corrections applied for truncation (see Fig. 3A) for which we have changed the initial measurements of fault-throw into displacement (Eq. (2.10)). Solid lines are fits with $p = 1.0$ to the individual data sets. Dashed lines show slopes of $p = 1.0$ fit approximately to the entire data set, and $p = 1.5$ fit approximately to the data with log (length) exceeding -1. Vertical gray line is at the abscissa value of 0; horizontal gray lines indicate the range of values for the zero-intercepts of the linear fits to the data.

values of B for the individual data sets, appropriate for the linear fits with the slope set to $p = 1.0$, is (Fig. 1; Table 2; Clark and Cox, 1996, their table 3)

$$-\log B = \{-1.11 : -2.557\},\ B \approx \{13 : 360\}\ [(\mathrm{km})^p/\mathrm{km}], \quad (2.2)$$

where the pairs of numbers enclosed in braces represent the minimum and maximum values from among the fits to the individual data sets. We evaluate the data sets, however, using different possible values of [p, B] permitted by the data (Table 2).

2.2. Parameters for the Yucca Mountain data

The faults in the Yucca Mountain area in southern Nevada, U.S.A. have been the subject of extensive study because of the importance of the area as a potential nuclear waste repository. The fault-length vs. fault-displacement data from the map of Simonds et al. (1995) fall almost exactly along the line for the combined data sets of Clark and Cox (1996) (Fig. 1), although the data are scattered and do not define a slope accurately (Fig. 2A). Fitting the line of Eq. (I:3.1.2)$_2$ with an unconstrained p to all the Yucca Mountain data by minimizing the sum of the squared orthogonal distances from each point to the line gives (Table 2)

Table 2
Empirically determined parameters for displacement vs. length cross-plots.

Location ⟹		Multiple data sets, various locations		Yucca Mt., Nevada, USA					S. Yorkshire, UK				
Reference ⟹		Clark and Cox (1996)		Cladouhos and Marrett (1996), Marrett et al. (1999)					Watterson et al. (1996)				
Data set ⟹		Min	Max	p constrained[a]			p unconstrained[a]		NE-striking normal faults[b,c]				
				All faults	Normal faults		All faults	Normal faults	p constrained[d]				p unconstrained[d]
Parameter ⇓	Units ⇓						⇓ Parameter values ⇓						
p	–	1.0	1.0	1.0	0.41	1.0	1.1	2.6	0.67	1.0	1.24	1.8	1.6
$\log B$	–	1.107	2.557	1.873	1.233	1.94	1.936	3.855	2.11	2.01	1.94	1.76	1.821
B	$\frac{(\text{km})^p}{\text{km}}$	13	360	74.7	17.11	87.9	86.3	7164	130	102	86.3	57.8	66.2

[a] Values minimize the sum of the squares of the orthogonal distances from points to the line.
[b] From Watterson et al. (1996, their fig. 13f).
[c] Original data for maximum throw (v) have been converted to maximum displacement (δ) using Eqs. (2.10).
[d] Values minimize the sum of the squares of the orthogonal distances from points to the line, using data digitized from original graph.

$$p = 1.1, \quad -\log B = -1.936, \quad B = 86.3\ (\text{km})^p/\text{km}. \tag{2.3}$$

Of the eleven data points, eight are for normal faults, and three are for strike-slip faults, the latter being the points for the shortest-length faults. If we use only the eight normal faults, and calculate the fit that minimizes the sum of the squared orthogonal distances from each point to the line, we find a considerably steeper line given by (Table 2, Fig. 2A)

$$p = 2.6, \quad -\log B = -3.855, \quad B = 7164\ (\text{km})^p/\text{km}. \tag{2.4}$$

The data are not numerous, however, and show considerable scatter, so these fits are not reliable. Given the consistency with the data from Clark and Cox (1996) and from Watterson et al. (1996) (Fig. 1), we consider a slope of $p = 1.0$ to be an acceptable fit to these data. Using least squares to find the intercept that minimizes the sum of the squared orthogonal distances from each data point to the line constrained to have the slope $p = 1.0$, we find nearly the same result for all the Yucca Mountain data and for only the normal faults (Table 2; Figs. 1, 2A)

$$\text{All faults}: \ p = 1.0, \ -\log B = -1.873, \ B = 74.7\ (\text{km})^p/\text{km}, \tag{2.5}$$

$$\text{Normal faults}: \ p = 1.0, \ -\log B = -1.94, \ B = 87.9\ (\text{km})^p/\text{km}. \tag{2.6}$$

These solutions are also nearly the same as that for all the faults when p is not constrained (Eq. (2.3)).

Finally, for comparison with self-consistency tests in Section 4, we calculate the intercept that minimizes the sum of the squared orthogonal distances of the normal fault data points from the line having a slope of $p = 0.41$ (Table 2; Fig. 2A):

$$\text{Normal faults}: \ p = 0.41, \ -\log B = -1.233, \ B = 17.11\ (\text{km})^p/\text{km}. \tag{2.7}$$

Fig. 2B shows data on the frequency distribution of fault-lengths determined from a two-dimensional sampling of a 116 km^2 area at Yucca Mountain (data from Scott and Castellanos, 1984, in Cladouhos and Marrett, 1996). The best-fit of Eq. $(\text{I:3.1.4})_2$ to the linear part of the curve gives (Table 3; Fig. 2B)

$$m_2 = 1.3, \quad \log g_{(2,\perp h)} \approx -0.41, \quad g_{(2,\perp h)} \approx 0.39\ (\text{km})^{m_2}/(\text{km})^2. \tag{2.8}$$

Fig. 2C shows the cumulative-frequency of fault-displacements in the Yucca Mountain area determined from a one-dimensional sampling along a line approximately normal to the strike of the faults (Marrett et al., 1999). The best-fit of Eq. $(\text{I:3.1.11})_2$ to the data is given by the parameters (Table 3; Fig. 2C)

$$s_1 = 0.73, \quad \log r_{(1,t)} = -1.28, \quad r_{(1,t)} = 0.052\ (\text{km})^{s_1}/\text{km}. \tag{2.9}$$

These data span approximately five orders of magnitude in displacement, and they are the concatenation of measurements made at three different scales: 1:100,000 for $3.2 \le \delta\ [\text{km}] \le 0.1$ using a 145.4 km transect length; 1:10,000 for $0.5 \le \delta\ [\text{km}] \le 0.03$ using a 14 km transect length; and $0.125 \le \delta\ [\text{km}] \le 6 \times 10^{-6}$ using a 4.8 km transect length in the Yucca Mountain Exploratory Studies Facility tunnel. The linearity of the data over such an unusually large range and over three different scales argues strongly that these parameters are particularly robust.

2.3. Parameters for the South Yorkshire data

Watterson et al. (1996) have made an exceptionally thorough analysis of the faults in the East Pennine coalfield in south Yorkshire, U.K. This area is cut by three sets of faults: a WNW-striking set of strike-slip faults and two sets of normal faults, one striking NE and the other striking NW. We use the results from their analysis of the NE-striking set of normal faults (Fig. 3), ignoring the other two sets because the strike-slip faults are younger and accommodate a different strain, and because many faults of the NW-striking set are constrained between faults of NE-striking set, thereby affecting the fault-length data. For Figs. 1 and 3, we converted the original data on maximum throw (v) on the faults (the vertical component of the displacement) to fault-displacements δ by assuming slip on the faults is dip-slip with an average fault-dip of 69° (Watterson et al., 1996), whereby

$$\left.\begin{aligned} \delta &= \frac{v}{\cos 69^\circ} = \frac{v}{0.3584}, \\ \log \delta &= \log v - \log(\cos 69^\circ) = \log v + 0.4457. \end{aligned}\right\} \tag{2.10}$$

This conversion simply shifts all the data points on the log–log plot by the constant amount of 0.4457 parallel to the displacement axis.

The relation between fault-length and fault-displacement from the data of Watterson et al. (1996, their fig. 13f) is plotted in Figs. 1 and 3A. These data include only faults with a length determined by a tip-to-tip measurement. The fault-lengths have been corrected for truncation errors (see Watterson et al., 1996), and all lengths have been converted to units of km. These data are compatible with Eq. $(\text{I:3.1.1})_2$ and lie very close to the line with $p = 1.0$ that is fit to all of the data of Clark and Cox (1996) (Fig. 1). They have a significant scatter, however, and do not resolve the slope well. Using data digitized from Watterson et al. (1996, fig. 13f), a regression minimizing the sum of the squared orthogonal distances from points to the line gives (Table 2, Fig. 3A)

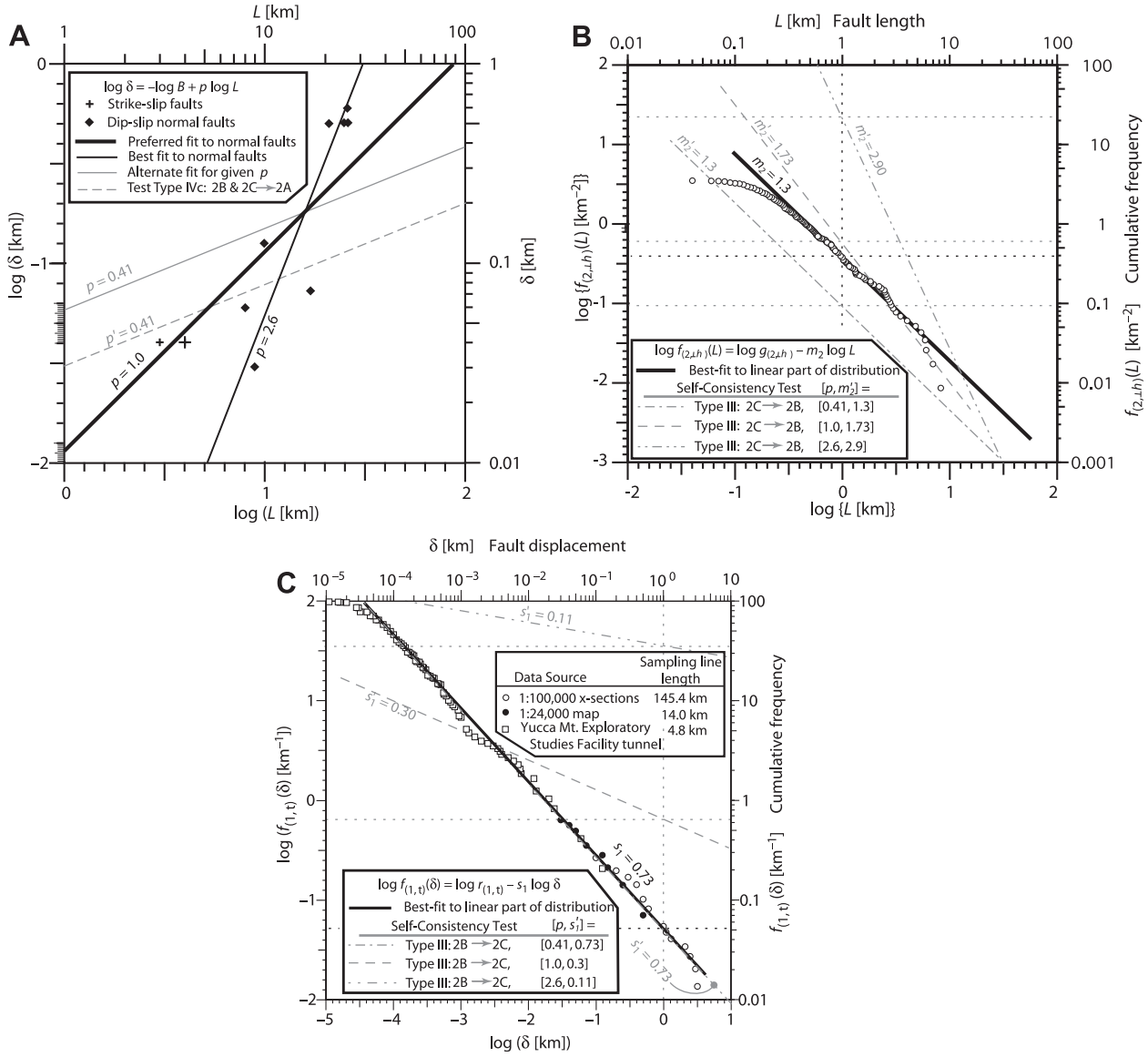

Fig. 2. Length vs. displacement, length vs. frequency and displacement vs. frequency scaling relations for normal faults in the Yucca Mountain area, southern Nevada, U.S.A. A. Length–displacement systematics measured from the map of Simonds et al. (1995). Solid lines show the linear best-fits to the normal fault data (diamonds) for p constrained to be 0.41 and 1.0, and for p unconstrained and equal to 2.6. The dashed line is the prediction of a Type IVc self-consistency test. The best-fits for constrained and unconstrained values of p were determined by minimizing the sum of the squared orthogonal distances from data points to the line. The black line with $p = 1.0$ is used as the preferred fit. The larger '+' sign marks two identical data points. B. Fault-length systematics, measured from two-dimensional sampling of an area of 116 km^2 in the Paintbrush tuff, plotted as the logarithm of fault-length [km] vs. the logarithm of cumulative-frequency [km^{-2}]. The solid black line with slope $m_2 = 1.3$ is the best-fit to the approximately linear part of the data distribution. The gray broken lines with the slopes of $m_2' = 1.3$, 1.73, and 2.90 are the self-consistency tests of Type (III: 2C → 2B) for different values of p using the value of $s_1 = 0.73$ for the displacement frequency data in Fig. 2C. Vertical dashed line is at the abscissa value of 0; horizontal dashed lines indicate the value of the intercepts for the plotted lines. (Data from Scott and Castellanos, 1984, in Cladouhos and Marrett, 1996). C. Logarithm of displacement [km] vs. the logarithm of cumulative-frequency, measured from one-dimensional sampling in the Paintbrush tuff, Yucca Mountain, Nevada. Sample line lengths are listed in the upper inset on the graph. The solid line with $s_1 = 0.73$ is the best-fit to the linear part of the data. Broken lines with $s_1' = 0.73$, 0.30, and 0.11 are the self-consistency tests of the Type (III: 2B → 2C) for different values of p for the data in Fig. 2B. Vertical gray dotted line is at the abscissa value of 0; horizontal gray dotted lines indicate the values of the intercepts for the fitted and the predicted lines. (Data from Marrett et al., 1999).

$$p = 1.6, \quad -\log B \approx -1.821, \quad B \approx 66.2\ (\text{km})^p/\text{km}. \tag{2.11}$$

Fitting the data with lines for which the slope p is constrained to have the values 0.67, 1.0, 1.24, and 1.8, we obtain (Table 2, Fig. 3A),

$$\text{For } p = 0.67: \quad -\log B \approx -2.11, \quad B \approx 130\ (\text{km})^p/\text{km}, \tag{2.12}$$

$$\text{For } p = 1.0: \quad -\log B \approx -2.01, \quad B \approx 102\ (\text{km})^p/\text{km}, \tag{2.13}$$

$$\text{For } p = 1.24: \quad -\log B \approx -1.94, \quad B \approx 86.3\ (\text{km})^p/\text{km}, \tag{2.14}$$

$$\text{For } p = 1.8: \quad -\log B \approx -1.76, \quad B \approx 57.8\ (\text{km})^p/\text{km}. \tag{2.15}$$

These values for B are within the range given in Eq. (2.2) (Table 2, Figs. 1, 3A).

Watterson et al. (1996) also determined the scaling systematics for fault-length using two-dimensional sampling, and for fault-throw using both one- and two-dimensional sampling. The number distribution of fault-lengths from a two-dimensional sampling within an 87 km^2 area, are fit by Eq. (I:3.1.3)$_2$ with $\zeta = 2$ (Fig. 3B, Table 3)

$$m_2 = 1.36 \pm 0.06, \quad \log G_{(2,\perp h)} \approx 1.67, \quad G_{(2,\perp h)} \approx 46.8\ \text{km}^{m_2}, \tag{2.16}$$

Table 3
Empirically determined parameters for displacement and length vs. frequency systematics.

Location ⟹		Yucca Mt. Nevada, USA	East Pennine coalfield, south Yorkshire, U.K.
Reference ⟹		Cladouhos and Marrett (1996), Marrett et al. (1999)	Watterson et al. (1996)
Data Set ⟹		All data	NE-striking normal faults
Parameter ⇓	Units ⇓	⇓ Parameter values ⇓	
m_2	–	1.3	$1.36 \pm .06$[a]
$\log G_{(2,\perp h)}$	–	–	1.67
$G_{(2,\perp h)}$	$[\mathrm{km}^{m_2}]$	–	46.8[a,b]
$\log g_{(2,\perp h)}$	–	−0.41	−0.269
$g_{(2,\perp h)}$	$(\mathrm{km})^{m_2}/\mathrm{km}^2$	0.39[b]	0.538[c]
s_1	–	0.73	$0.54 \pm .02$[d]
$\log r_{(1,t)}$	–	−1.28	−1.14[b,d,e]
$r_{(1,t)}$	$(\mathrm{km})^{s_1}/\mathrm{km}$	0.052[b]	0.072[b,d,e]
s_2	–	–	1.1 ± 0.1[f]
$\log R_{(2,\perp h)}$	–	–	−0.38[b,e,f]
$R_{(2,\perp h)}$	$(\mathrm{km})^{s_2}$	–	0.417[b,e,f]
$\log r_{(2,\perp h)}$	–	–	−2.32[g]
$r_{(2,\perp h)}$	$(\mathrm{km})^{s_2}/(\mathrm{km})^2$	–	0.0048[g]
$\Delta_{(1,t)}$	$=\mathcal{T}$ [km]	145.4: $_{3.2 \le \delta\ [\mathrm{km}] \le 0.1}$ 14: $_{0.5 \le \delta\ [\mathrm{km}] \le 0.03}$ 4.8: $_{0.125 \le \delta\ [\mathrm{km}] \le 6 \times 10^{-6}}$	2.7–11.8
$\Delta_{(2,\perp h)}$	$=\mathcal{TW}$ [km^2]	116	87

[a] From Watterson et al. (1996, fig. 10a).
[b] Values approximated from graphical construction.
[c] Calculated from $G_{(2,\perp h)}$ using Eq. (I:3.1.5)$_2$ with $(\zeta,v)=(2,\perp h)$.
[d] From Watterson et al. (1996, table 3 & fig. 8b).
[e] Original data for maximum throw (v) have been converted to maximum displacement (δ) using Eqs. (2.10).
[f] From Watterson et al. (1996, fig. 9a).
[g] Calculated from $R_{(2,\perp h)}$ using Eq. (I:3.1.12) with $(\zeta,v)=(2,\perp h)$.

where we estimated the value of $\log G_{(2,\perp h)}$ from the graph in Fig. 3B.

For one-dimensional sampling ($\zeta=1$) of fault-displacements along multiple sampling lines ranging in length from 2.9 km to 11.8 km, the data of Watterson et al. (1996, their table 2 and fig. 8b) give values for Eq. (I:3.1.11)$_2$ of (Table 3; Fig. 3C),

$$s_1 = 0.54 \pm .02, \quad \log r_{(1,t)} \approx -1.14, \quad r_{(1,t)} \approx 0.072\ (\mathrm{km})^{s_1}/\mathrm{km}, \tag{2.17}$$

where we estimated the value of $\log r_{(1,t)}$ from Fig. 3C by graphical construction.

For two-dimensional sampling of maximum throw ($\zeta=2$) within an 87 km^2 area (Fig. 3D), values of the parameters in Eq. (I:3.1.10)$_2$ are, from Watterson et al. (1996, their table 2 and fig. 9a),

$$s_2 = 1.1 \pm 0.1, \quad \log R_{(2,\perp h)} \approx -0.38, \quad R_{(2,\perp h)} \approx 0.417\ (\mathrm{km})^{s_2}, \tag{2.18}$$

where we converted throw to displacement using Eq. (2.10)$_2$ and estimated $\log R_{(2,\perp h)}$ graphically from Fig. 3D.

3. Constraints on the size of the sampling volume $\mathcal{V}$ relative to the largest fault within it

Equations (I:3.4.2)$_2$ and (I:3.4.3) prescribe constraints on the dimensions of the volume $\mathcal{V}$ relative to the size of the largest fault that can be considered in this type of analysis. For an order-of-magnitude estimate, we define a 'vertical-equivalent-fault' by assuming a rectangular ($\lambda/\mu=1$) vertical ($\rho=90^\circ$) fault and a traverse direction **t** perpendicular to that fault ($\theta=0^\circ$), whereby Eq. (I:3.4.4) gives simplified estimates for the relative sizes. We adopt values of p and s_ζ for the Yucca Mt. area from Eqs. (2.5)$_1$ and (2.9)$_1$, $p=1$, $s_1=0.73$. From Eqs. (I:3.2.6) and (I:3.2.8), we infer that $s_2=1.73$, $s_3=2.73$. Using these values in Eqs. (I:3.4.4) indicates that

$$\left.\begin{array}{l} \text{for } \dfrac{\lambda}{\mu}=1, \quad \theta=0^\circ, \text{ and } \rho=90^\circ, \\ p=1, \quad s_1=0.73, \\ \mathcal{W} \ge 2.4L^{(1)}, \quad \mathcal{H} \ge 1.6\mu L^{(1)} = 1.6l^{(1)}, \end{array}\right\} \tag{3.1}$$

where $\mu L^{(1)}=l^{(1)}$ is the down-dip width of the largest fault, which in this case we take to be the largest vertical-equivalent fault.

Thus, the volume $\mathcal{V}=\mathcal{TWH}$ over which the sampling of the faults must occur should have a horizontal width $\mathcal{W}$ is no smaller than 2.4 times the horizontal length $L^{(1)}$ of the largest fault, and a depth $\mathcal{H}$ no smaller than 1.6 times the vertical extent of the largest vertical-equivalent-fault $l^{(1)}$. In other words, the cross sectional area $\mathcal{WH}$ of the volume $\mathcal{V}$ must be greater than approximately 3.8 times the area of the largest fault projected onto a vertical plane that is parallel to the fault strike. If $p=2.0$, and $s_1=0.73$, then $s_3=1.73$ and $s_2=1.23$, which gives constraints for the size of the domain of $\mathcal{W} \ge 1.7L^{(1)}$, $\mathcal{H} \ge 1.4\mu L^{(1)} = 1.4l^{(1)}$, which are of comparable magnitude to the values in Eq. (3.1).

Constraints on the dimensions of the volume to which this strain analysis can be applied can also be obtained from the displacement on the largest fault in the analysis. For the Yucca Mt. area, for example, the best data available are the fault-displacements from one-dimensional sampling. Assuming the maximum measured displacement is the displacement on the largest fault $\delta_{(1)}^{(\max)}=\delta_{(3)}^{(\max)}$, we have from Eqs. (I:3.4.5), with Eq. (2.5), setting $L=L^{(1)}$ for which $\delta=\delta_{(3)}^{(\max)}$:

$$\left.\begin{array}{l} \text{for } \dfrac{\lambda}{\mu}=1, \quad \theta=0^\circ, \quad \rho=90^\circ, \\ p=1, \quad B=88.7\ (\mathrm{km})^p/\mathrm{km}, \quad s_1=0.73, \quad \delta_{(3)}^{(\max)}=0.125\ \mathrm{km}, \\ \mathcal{W} \ge 26.3\ \mathrm{km}, \quad \mathcal{H} \ge 17.5\mu\ \mathrm{km}. \end{array}\right\} \tag{3.2}$$

Thus the volume to which we can apply the estimates of the volumetric extension can have a cross section normal to the sampling line no smaller than about 26 km wide by 17.5μ km deep, where μ is the ratio of horizontal fault length to down-dip width (Eq. (I:3.1.21)$_2$).

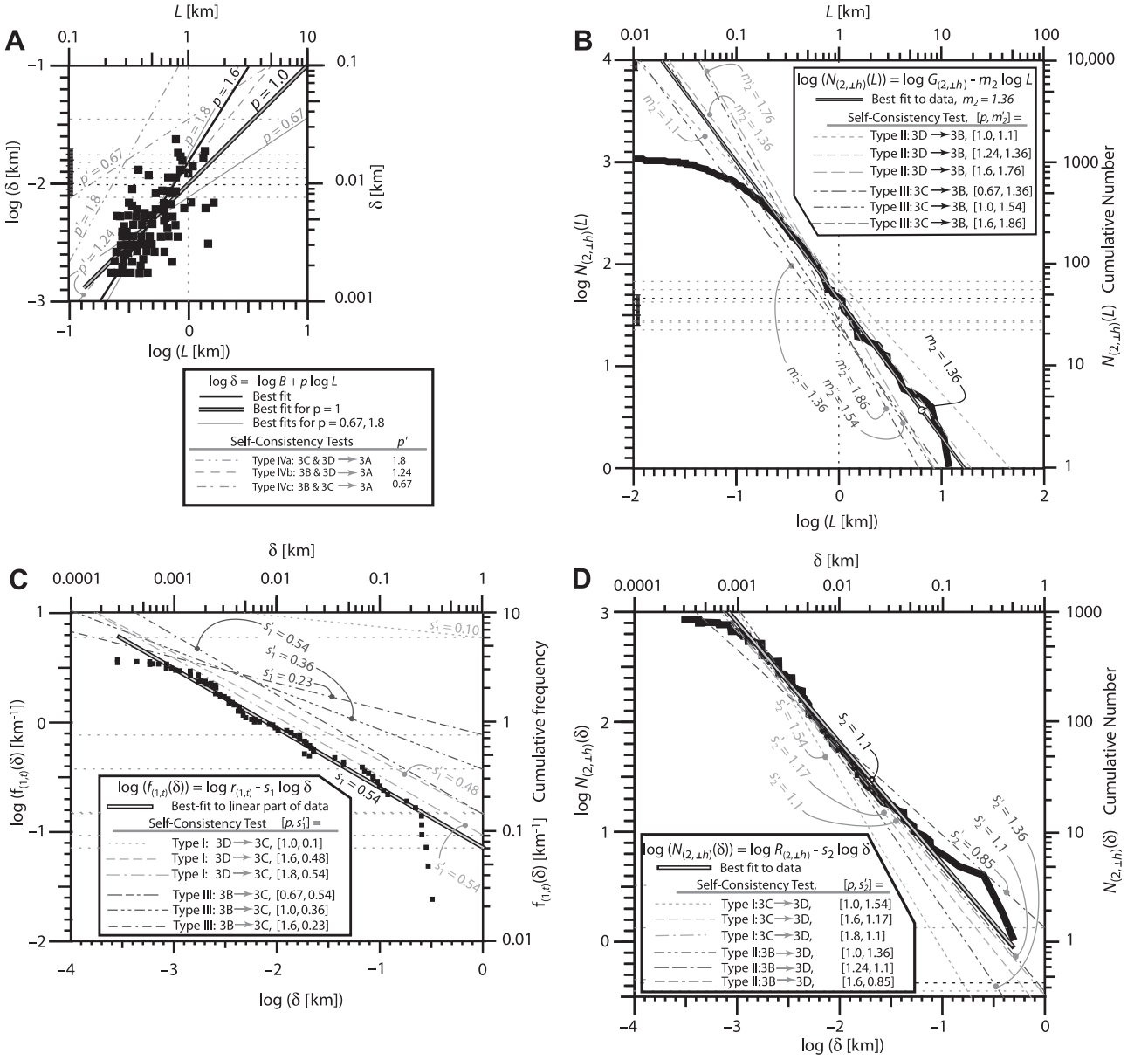

Fig. 3. Fault systematics data for the NE-striking set of normal faults in the south Yorkshire coalfields, from Watterson et al. (1996). Vertical dotted lines (Parts A, B) are at the abscissa value of 0; horizontal dotted lines show the intercept values for plotted lines. All plots of fault-displacement have been converted from the original measurements of throw using Eq. (2.10), and original units of meters have been converted to kilometers. A. Log–log plot of fault-length vs. fault-displacement (from Watterson et al., 1996, their fig. 13f). Solid lines are best linear fits to the data for p constrained to be 0.67, 1.0, and 1.8. Solid line for $p = 1.6$ is the best-fit line for unconstrained p. Fits are determined by minimizing the sum of the squared orthogonal distances from each point to the line. Broken lines are plots of self-consistency tests, as listed in the legend, for $p' = 0.67$, 1.24, and 1.8. B. Log–log plot of fault-length vs. cumulative-number measured in a two-dimensional sampling domain having an area of 87 km^2 (from Watterson et al., 1996, their fig. 10a). Solid double line with $m_2 = 1.36$ is the best-fit to the linear part of the data. Broken gray lines are self-consistency tests of Types II and III as listed in the inset legend. C. Log–log plot of fault-displacement vs. cumulative-frequency measured from multiple one-dimensional sampling domains with lengths ranging from 2.9 km to 11.8 km (from Watterson et al., 1996, their fig. 8b). Solid double line with $s_1 = 0.54$ is the best-fit to the linear part of the data. Broken gray lines are self-consistency tests of Types I and III, as listed in the inset legend. D. Log–log plot of fault-displacement vs. cumulative-number measured in a two-dimensional sampling domain having an area of 87 km^2 (from Watterson et al., 1996, their fig. 9a). The solid double line with $s_2 = 1.1$ is the best-fit to the linear part of the data. Broken gray lines are self-consistency tests of Types I and II, as listed in the inset legend.

Alternatively, these constraints can be taken to define the dimensions of the largest fault in a domain of a given size that logically can be included in an analysis of strain using the fault systematics. For the south Yorkshire area, for example, the map area of the analysis is $\mathcal{TW} = 87\ \text{km}^2$ (Table 3). From Watterson et al. (1996, fig. 2a), we can estimate that $\mathcal{W} \approx 10$ km, measured parallel to the NE-striking faults. We calculate the constraints for the length of the largest fault $L^{(1)}$ in terms of $\mathcal{W}$ and $\mathcal{H}$ from Equations (I:3.4.4) using Equations $(2.17)_1$ and (2.1) with Equations (I:3.2.6) and (I:3.2.8). If we assume, for the sake of an example, that the depth of the faulted volume is $\mathcal{H} \approx 15$ km, and that the vertical extent of the largest fault $\ell_v^{(1)}$ is half its length, we find,

$$\left.\begin{aligned} &\text{for } \frac{\lambda}{\mu} = 1, \quad \theta = 0^\circ, \quad \rho = 90^\circ, \\ &p = 1, \quad s_1 = 0.54, \\ &\mathcal{W} = 10\ \text{km}, \quad \mathcal{H} = 15\ \text{km}, \quad \ell_v^{(1)} = 0.5L^{(1)}, \\ &\mathcal{WH} \approx 150\ \text{km}^2 \geq 4.7L^{(1)}\ell_v^{(1)} = 4.7 \cdot 0.5\left(L^{(1)}\right)^2, \end{aligned}\right\} \tag{3.3}$$

from which we conclude that the length of the largest fault is constrained by

$$L^{(1)} \leq 8 \text{ km}. \tag{3.4}$$

From Fig. 3B, we see that this limit is close to the fault-length above which the cumulative-number curve departs from the linear approximation. Given the rough assumptions we made for this calculation, the estimate that the length of the largest fault admissible in the strain analysis for this area is about 8 km seems a reasonable approximation. In principle, then, the contributions of larger faults would have to be accounted for individually.

4. Tests for the self-consistency of empirical data

4.1. Types of self-consistency tests

The equations for the fault systematics that we have developed include the relation between fault-length and fault-displacement (Eqs. (I:3.1.1) and (I:3.1.2)), and the cumulative-number and cumulative-frequency distributions for fault-length (Eqs. (I:3.1.3), (I:3.1.4)) and for fault-displacement (Eqs. (I:3.1.7), (I:3.1.8), (I:3.1.10), and (I:3.1.11)). These equations are characterized by parameters defining the slopes of the log-linear fits to the data (p, m_ζ, and s_ζ) and the respective intercepts: $-\log B$; either $\log G_{(\zeta,v)}$ or $\log g_{(\zeta,v)}$; and either $\log R_{(\zeta,v)}$ or $\log r_{(\zeta,v)}$, where the subscript (ζ,v) indicates the number of dimensions ζ of the sampling domain and the orientation v of that domain. The pairs of alternative intercepts listed here are for data plotted as cumulative-number (upper case symbol) and cumulative-frequency (lower case symbol), and these quantities are related by Eqs. $(I:3.1.5)_2$ and (I:3.1.12) with Eqs. (I:3.1.6). Although these parameters are constrained empirically by the data plotted in Figs. 1–3 (the parameter values are collected in Tables 2 and 3), in theory they are not all independent, and this interdependence provides the opportunities for a number of different types of self-consistency test, which we define below and apply in Sections 4.2–4.4.

In subsequent equations, we adopt the convention that a prime on a quantity identifies it as having been calculated from other measured parameters, as opposed to being measured directly by fitting lines to the data.

4.1.1. The Type I test: same variable; different dimensionality

The Type I test is a check on the self-consistency of the *same variable* determined from measurements in domains of *different dimensionality*. In practice, the only variable to which this test can be applied is the displacement, because practical measurements of fault length over the necessary range of lengths cannot be made in one or three dimensions. The slopes and intercepts for the log-linear fits to displacement data are related by Eqs. (I:3.2.6) through (I:3.2.9). Solving both Eqs. (I:3.2.6) and (I:3.2.8) for s_3 and equating the results gives the relation between s_1 and s_2 as a function of p. Solving Eqs. (I:3.2.7) and (I:3.2.9) for $r_{(3)}$, equating the results, and using Eq. (I:3.1.12) with $(\zeta,v)=(2,\perp h)$ gives the relation between $r_{(1,t)}$ and $R_{(2,\perp h)}$:

$$s_2' = s_1 + \frac{1}{p},\quad R'_{(2,\perp h)} = r_{(1,t)}\Delta_{(2,\perp h)}\frac{s_1}{(s_1+1/p)}B^{-1/p}\frac{\mu}{\lambda}\frac{\sin\rho}{\cos\theta}, \tag{4.1.1}$$

$$s_1' = s_2 - \frac{1}{p},\quad r'_{(1,t)} = \frac{R_{(2,\perp h)}}{\Delta_{(2,\perp h)}}\frac{s_2}{(s_2-1/p)}B^{1/p}\frac{\lambda}{\mu}\frac{\cos\theta}{\sin\rho}. \tag{4.1.2}$$

4.1.2. The Type II test: different variables; same dimensionality

The Type II test is a check on the self-consistency of the two *different variables*, displacement and length, from measurements in a domain of the *same dimensionality* ζ. In practice, we must adopt $\zeta=2$ for this test, because fault-length systematics can realistically be measured only in two dimensions. The slopes and intercepts of the log-linear fits to data, $[m_2, g_{(2,\perp h)}]$ and $[s_2, r_{(2,\perp h)}]$, are related to one another as defined by Eqs. (I:3.1.9). Thus, from Eq. (I:3.1.9) with $(I:3.1.5)_2$ and (I:3.1.12), we have,

$$\begin{aligned} s_2' &= \frac{m_2}{p}, \\ R'_{(2,\perp h)} &= G_{(2,\perp h)}B^{-m_2/p},\quad r'_{(2,\perp h)} = g_{(2,\perp h)}B^{-m_2/p}, \end{aligned} \tag{4.1.3}$$

$$\begin{aligned} m_2' &= s_2 p, \\ G'_{(2,\perp h)} &= R_{(2,\perp h)}B^{s_2},\quad g'_{(2,\perp h)} = r_{(2,\perp h)}B^{s_2}. \end{aligned} \tag{4.1.4}$$

4.1.3. The Type III test: different variables; different dimensionality

The Type III test is a check on the self-consistency of the two *different variables*, displacement and length, from measurements in domains of *different dimensionality*. We use Eqs. (I:3.2.10) through (I:3.2.13) to check the self-consistency of the cumulative-frequency data for fault-length sampled in a two-dimensional domain, characterized by the parameters $[m_2, g_{(2,\perp h)}]$, against the cumulative-frequency data for fault-displacement sampled in a one-dimensional domain, characterized by the parameters $[s_1, r_{(1,t)}]$.

To that end, we recast Eq. $(I:3.2.10)_2$ to calculate a value of m_2' from s_1 and of s_1' from m_2.

$$m_2' = s_1 p + 1,\quad s_1' = \frac{m_2 - 1}{p}. \tag{4.1.5}$$

To calculate the value of $g'_{(2,\perp h)}$ from the empirical values of s_1, $r_{(1,t)}$, B and p, we use Eq. $(I:3.1.9)_5$ with $\zeta=1$, which we rewrite as,

$$g_{(1,t)} = r_{(1,t)}B^{m_1/p}. \tag{4.1.6}$$

From Eq. (I:3.2.5) we solve for $g_{(3)}$, substitute the result into Eq. (I:3.2.3), and use Eq. (I:3.2.4) to find a relation between $g_{(1,t)}$ and $g_{(2,\perp h)}$,

$$g_{(1,t)} = g_{(2,\perp h)}\frac{m_2}{(m_2-1)}\frac{\lambda}{\mu}\frac{\cos\theta}{\sin\rho}. \tag{4.1.7}$$

Then we can equate the right sides of Eqs. (4.1.7) and (4.1.6), substitute for m_2 from Eq. $(I:3.2.10)_4$, and solve for either $g_{(2,\perp h)}$ or $r_{(1,t)}$ to obtain,

$$\begin{aligned} g'_{(2,\perp h)} &= r_{(1,t)}\frac{s_1 p}{s_1 p + 1}B^{s_1}\frac{\mu}{\lambda}\frac{\sin\rho}{\cos\theta}, \\ G'_{(2,\perp h)} &= r_{(1,t)}\Delta_{(2,\perp h)}\frac{s_1 p}{s_1 p + 1}B^{s_1}\frac{\mu}{\lambda}\frac{\sin\rho}{\cos\theta}, \end{aligned} \tag{4.1.8}$$

$$\begin{aligned} r'_{(1,t)} &= g_{(2,\perp h)}\frac{m_2}{(m_2-1)}B^{-(m_2-1)/p}\frac{\lambda}{\mu}\frac{\cos\theta}{\sin\rho}, \\ r'_{(1,t)} &= \frac{G_{(2,\perp h)}}{\Delta_{(2,\perp h)}}\frac{m_2}{(m_2-1)}B^{-(m_2-1)/p}\frac{\lambda}{\mu}\frac{\cos\theta}{\sin\rho}, \end{aligned} \tag{4.1.9}$$

where we have used Eq. $(I:3.1.5)_2$ to express the equations in terms of either $G_{(2,\perp h)}$ or $g_{(2,\perp h)}$.

4.1.4. The Type IV test: self-consistency of p with $m_{(\zeta)}$ and $s_{(\zeta)}$

The Type IV test is a check on the self-consistency of the empirical value of the parameter pair $[p, B]$ with its value as determined independently by the parameter pairs $[m_2, g_{(2,\perp h)}]$ and $[s_\zeta, r_{(\zeta,v)}]$ $(\zeta=1,2)$, which characterize the log–log plots of length vs. frequency (Eq. $(I:3.1.4)_2$) and displacement vs. frequency (Eq. $(I:3.1.11)_2$), respectively (Yielding et al., 1996). Three subtypes of this consistency test can be used.

4.1.4.1. IVa: same variable, different dimensionality. We can solve for p' and B' using the frequency distributions for displacement that have been measured in domains of dimensionality $\zeta=1$ and 2. We solve each of Eqs. (I:3.2.6) and (I:3.2.8) for s_3, equate the results, and rearrange to find,

$$p' = \frac{1}{(s_2 - s_1)}. \tag{4.1.10}$$

We then solve both Eqs. (I:3.2.12) and (I:3.2.13) for $g_{(3)}$, equate the results, solve the resulting equation for B, and use Eqs. $(I:3.2.10)_3$ and $(I:3.2.10)_5$ with (I:3.1.12) to find,

$$\begin{aligned} B' &= \frac{r_{(1,t)}s_1}{r_{(2,\perp h)}s_2}\frac{\mu \sin\rho}{\lambda\cos\theta}, \\ B' &= \frac{r_{(1,t)}s_1 \Delta_{(2,\perp h)}}{R_{(2,\perp h)}s_2}\frac{\mu \sin\rho}{\lambda\cos\theta}. \end{aligned} \tag{4.1.11}$$

4.1.4.2. IVb: different variables, same dimensionality. We can also solve for p' and B' using the frequency distributions for both displacement and length that have been measured in domains of the same dimensionality, which in practice is restricted to $\zeta = 2$ because fault-length can be measured only in a two-dimensional domain. The parameters are related by Eqs. $(I:3.1.9)_1$, which we solve for p. Setting $\zeta = 2$, we find,

$$p' = \frac{m_2}{s_2}. \tag{4.1.12}$$

From Eqs. $(I:3.1.9)_2$ and $(I:3.1.9)_4$, we solve for B and set $(\zeta,\nu) = (2,\perp h)$ to find,

$$B' = \left[\frac{g_{(2,\perp h)}}{r_{(2,\perp h)}}\right]^{1/s_2} = \left[\frac{G_{(2,\perp h)}}{R_{(2,\perp h)}}\right]^{1/s_2}, \tag{4.1.13}$$

where we used Eq. $(I:3.1.5)_2$ and (I:3.1.12) to convert from the lower case to the upper case parameters.

4.1.4.3. IVc: different variables, different dimensionality. Finally, we can solve for p' and B' using the frequency distribution for displacement measured in a one-dimensional domain ($\zeta = 1$) and that for fault length measured in a two-dimensional domain ($\zeta = 2$). From Eqs. $(I:3.2.10)_2$, we solve for p to find,

$$p' = \frac{m_2 - 1}{s_1}. \tag{4.1.14}$$

From Eq. (4.1.9) with Eq. $(I:3.2.10)_2$, we find,

$$\begin{aligned} B' &= \left[\frac{g_{(2,\perp h)}}{r_{(1,t)}}\frac{m_2}{(m_2-1)}\frac{\lambda}{\mu}\frac{\cos\theta}{\sin\rho}\right]^{\frac{1}{s_1}}, \\ B' &= \left[\frac{G_{(2,\perp h)}}{\Delta_{(2,\perp h)}r_{(1,t)}}\frac{m_2}{(m_2-1)}\frac{\lambda}{\mu}\frac{\cos\theta}{\sin\rho}\right]^{\frac{1}{s_1}}. \end{aligned} \tag{4.1.15}$$

4.1.5. Assumed parameter values

The equations for fault systematics and some of the equations for the self-consistency tests are dependent on parameters that define the geometric characteristics of the faults, namely θ, the angle between the normal to a fault plane and the sampling line **t**; ρ, the dip of a fault; λ, a geometrical shape factor that defines the shape of a fault tip line; and μ, the ratio of the down-dip width to the horizontal length of a fault. In applying the tests for self-consistency in Sections 4.2–4.4, we assume from Eqs. (I:2.1.13), $(I:2.1.14)_1$, and $(I:3.1.21)_2$ that,

$$\frac{\mu}{\lambda} \approx 1, \tag{4.1.16}$$

which applies for faults with a rectangular tip line. We assume in addition that

$$\rho = 90^\circ - \theta, \tag{4.1.17}$$

which implies that the sampling line must be normal to the average fault strike. These assumptions give,

$$\frac{\sin\rho}{\cos\theta} = \frac{\sin(90^\circ - \theta)}{\cos\theta} = 1. \tag{4.1.18}$$

These are values for idealized faults, but our level of knowledge of the fault geometry and the accuracy of the other parameters involved in these calculations does not justify the refinement that adjusting these parameters would provide.

The analysis by Clark and Cox (1996) of fault-length vs. fault-displacement data, (Fig. 1), and the consistency with that analysis of the Yucca Mountain data (Figs. 1, 2A) and the south Yorkshire data (Figs. 1, 3A) (Watterson et al., 1996), indicate that the value of $p = 1$ (Eq. (2.1)) provides a reasonable description of the data. Individual data sets, however, give best-fit values for p that can differ substantially from this generalized value (Table 2; Figs. 1–3), but the scatter in the data and the insufficient range of the data preclude a definitive rejection of the value $p = 1$ (Clark and Cox, 1996). Nevertheless, in developing the following self-consistency tests, we evaluate the viability of different values for p that are permitted by the data.

4.1.6. Notation

In Sections 4.2–4.4 that follow, we use the Yucca Mountain data (Fig. 2) and the south Yorkshire data (Fig. 3) to illustrate the application of the self-consistency tests. Because we have nine self-consistency tests (two tests each for Types I, II, and III, and three Type IV tests) and seven plots of data to consider, we adopt a shorthand notation to help clarify the presentation. In this notation, we first write the type of self-consistency test we are using; then the figure number and letter that identifies the source graph from which we determine the values of the parameters that describe the distribution of the data; then an arrow, followed by the figure number and letter of the graph on which the calculated distribution is plotted. For example, (II: 3D → 3B) indicates we apply a Type II self-consistency test using the parameter values determined from Fig. 3D to calculate the theoretically expected distribution of data that is plotted in Fig. 3B. This same notation is used in the figure legends to identify the calculated distributions that are plotted, as well as in the first columns of Tables 4A, B and 5A, B.

4.2. Self-consistency tests for the Yucca Mountain data

The data available for the Yucca Mountain area (Fig. 2) permit the application of three of the self-consistency tests described in Section 4.1. For each test, we plot the calculated lines on the graph that allows us to compare the calculated relationship with the observed data and thereby to evaluate the self-consistency of the different measured fault characteristics.

4.2.1. Type III test: different variables, different dimensionality

The data from the Yucca Mountain area allow two Type III self-consistency tests: (III: 2C → 2B), for which we use the [slope, intercept] pair $[s_1, \log r_{(1,t)}]$ to calculate $[m_2', \log g_{(2,\perp h)}']$; and (III: 2B → 2C), for which we use $[m_2, \log g_{(2,\perp h)}]$ to calculate $[s_1', \log r_{(1,t)}']$. We evaluate the results for different values of $[p, B]$, where for a selected value of p, we find B by minimizing the sum of the squared orthogonal distances of the points in Fig. 2A to the line having the given slope. We choose the values of $p = 0.41$, 1.0, and 2.6 (Table 2). The first is the value that makes the slopes of the best-fit lines to the data in Fig. 2B and C appear self-consistent; the second is the value consistent with the analysis of Clark and Cox (1996); and the third is the slope of the best-fit line to the normal fault data in Fig. 2A. We then compare the calculated lines with the observed data to evaluate the self-consistency of the parameters.

Table 4
Yucca Mt. normal faults.

A: Calculated parameters for type I, II, and III self-consistency tests

Test type: source → plot[a]	Variables; domain dimensions	Calculated parameter	Units	[p,B] = [0.41, 17.1]	[p,B] = [1, 87.9]	[p,B] = [2.6, 7164]	Equations[a]
III: 2C → 2B	δ, L; 1 & 2	m'_2	–	1.3	1.73	2.90	$(4.1.5)_1$
		$g'_{(2,\perp h)}$	$\frac{(\mathrm{km})^{m_2}}{\mathrm{km}^2}$	0.095	0.576	22.2	$(4.1.8)_1$
		$\log g'_{(2,\perp h)}$	–	−1.022	−0.240	1.35	–
III: 2B → 2C	δ, L; 1 & 2	s'_1	–	0.73	0.3	0.11	$(4.1.5)_2$
		$r'_{(1,t)}$	$\frac{(\mathrm{km})^{s_1}}{\mathrm{km}}$	0.052	0.652	34.6	$(4.1.9)_1$
		$\log r'_{(1,t)}$	–	−1.29	−0.186	1.539	–

B: Calculated parameters for type IV self-consistency tests

Test type: source → plot	Calculated parameter	Units	Value	Equations[a]
IVc: 2B & 2C → 2A	p'	–	0.41	(4.1.14)
	B'	$(\mathrm{km})^p/\mathrm{km}$	32.5	$(4.1.15)_1$
	$\log B'$	–	1.512	–

[a] Implicitly includes Eqs. (4.1.16)–(4.1.18) where appropriate.

First, we use the parameter values from Fig. 2C (Table 3) to calculate theoretical fits to the data in Fig. 2B (III: 2C → 2B). We use Eqs. $(4.1.5)_1$ and $(4.1.8)_1$ with the empirical values for the normal faults from Tables 2 and 3 to find (Table 4A, Fig. 2B),

$$\text{For } [p,B] = [0.41, 17.1]: \quad m'_2 = 1.3,$$
$$g'_{(2,\perp h)} = 0.095\left[(\mathrm{km})^{m_2}/\mathrm{km}^2\right], \quad \log g'_{(2,\perp h)} = -1.022, \quad (4.2.1)$$

$$\text{For } [p,B] = [1.0, 87.9]: \quad m'_2 = 1.73,$$
$$g'_{(2,\perp h)} = 0.576\left[(\mathrm{km})^{m_2}/\mathrm{km}^2\right], \quad \log g'_{(2,\perp h)} = -0.240, \quad (4.2.2)$$

$$\text{For } [p,B] = [2.6, 7164]: \quad m'_2 = 2.90,$$
$$g'_{(2,\perp h)} = 22.2\left[(\mathrm{km})^{m_2}/\mathrm{km}^2\right], \quad \log g'_{(2,\perp h)} = 1.35. \quad (4.2.3)$$

Although the value $p = 0.41$ makes the slopes of the lines in Fig. 2B and C appear self-consistent, it does not even remotely reflect the distribution of the data in Fig. 2A. Thus this solution must be rejected, and the length and displacement distributions in Fig. 2B and C must be considered inconsistent. It is evident that the calculated lines using the values of $p = 2.6$, which gives a best-fit to the normal fault data in Fig. 2A, also do not even come close to fitting the data in Fig. 2B (gray dash–dot–dot line). The calculated line using $p = 1$ (gray dashed line), however, provides a reasonable fit, especially for fault-lengths exceeding approximately 2 km. Nevertheless, the calculated slope of $m'_2 = 1.73$ is 1.3 times the empirical value of $m_2 = 1.3$ (Fig. 2B; Table 3), and the observed data fall significantly below the calculated line at fault-lengths below about 2 km.

We can also do the reverse test (III: 2B → 2C) and use the parameters $[m_2, \log g_{(2,\perp h)}]$ that define the best linear fit to the distribution of fault-length data from a two-dimensional sampling domain (Fig. 2B) to calculate the parameters $[s'_1, \log r'_{(1,t)}]$ that should describe the fault-displacement distribution in a one-dimensional sampling domain (Fig. 2C). In this case we use Eqs. $(4.1.5)_2$ and $(4.1.9)_1$ with the empirical values for the normal faults from Tables 2 and 3 to find (Table 4A; Fig. 2C),

$$\text{For } [p,B] = [0.41, 17.1]: \quad s'_1 = 0.73,$$
$$r'_{(1,t)} = 0.052\left[(\mathrm{km})^{s_1}/\mathrm{km}\right], \quad \log r = -1.29, \quad (4.2.4)$$

$$\text{For } [p,B] = [1.0, 87.9]: \quad s'_1 = 0.3,$$
$$r'_{(1,t)} = 0.652\left[(\mathrm{km})^{s_1}/\mathrm{km}\right], \quad \log r'_{(1,t)} = -0.186, \quad (4.2.5)$$

$$\text{For } [p,B] = [2.6, 7164]: \quad s'_1 = 0.11,$$
$$r'_{(1,t)} = 34.6\left[(\mathrm{km})^{s_1}/\mathrm{km}\right], \quad \log r'_{(1,t)} = 1.539. \quad (4.2.6)$$

These lines are plotted in Fig. 2C as, respectively, the gray dash–dot, dashed, and dash–dot–dot lines. Although the solution in Eqs. (4.2.4) fits the data in Fig. 2C, the value of $p = 0.41$ is unacceptable as a representation of the data in Fig. 2A. The two other calculated lines for the distribution of the displacement data in Fig. 2C clearly do not fit the data, and the fit becomes worse as the value of p increases.

Thus if we use the displacement-distribution parameters from a one-dimensional sampling (Fig. 2C) and a value of $p = 1$ to calculate a line for the fault-length distribution (Fig. 2B), then the test (III: 2C → 2B) provides a reasonable fit, at least to the high-length part of the observed distribution. Other values of p, however, are either unacceptable representations of the data in Fig. 2A, or provide unacceptable fits to the data in Fig. 2B. If we use the parameters that describe the fault-length distribution from a two-dimensional sampling (Fig. 2B) to calculate a line to fit the observed one-dimensional displacement distribution (Fig. 2C), the test (III: 2B → 2C) is unsuccessful for any reasonable value of p. This result indicates that the empirically determined values of $[m_2, g_{(2,\perp h)}]$ are not self-consistent with the displacement data, and it suggests that there may be a systematic bias in measuring the fault systematics toward underdetermining either the lengths or the frequencies of faults at the mid- to low-length parts of the data set.

For the Yucca Mountain area, the determinations of $[s_1, \log r_{(1,t)}]$ for the displacement-cumulative-frequency relation (Fig. 2C) are particularly robust because the range of displacement values spanned by the data, almost five orders of magnitude, is exceptionally large, and because the one-dimensional sampling is unambiguous in terms of the displacements measured on the faults intersected by the sampling line. The values of $[m_2, \log g_{(2,\perp h)}]$ from the two-dimensional sampling, on the other hand, are determined over only about one order of magnitude of fault-length data (Fig. 2B). We conclude that the parameters $[s_1, \log r_{(1,t)}]$ are more robust and more reliable than $[m_2, \log g_{(2,\perp h)}]$. We discuss the problems of sampling in different dimensions in Section 4.4, but note here that a sampling of fault trace lengths in a two-dimensional domain may be problematic.

4.2.2. Type IVc test: self-consistency of the parameters p with m_2 and s_1

We can apply the test Type (IVc: 2B & 2C → 2A) using Eqs. (4.1.14) and $(4.1.15)_1$ with the parameter values from Table 3. This tests the self-consistency of the frequency distributions of fault-lengths

Table 5
S. Yorkshire, NE normal faults.

A: Calculated parameters for type I, II, and III self-consistency tests

Test type: Source → Plot[a]	Variables; domain dimensions	Calculated parameter	Units	[p,B] = [0.67, 130]	[p,B] = [1.0, 102]	[p,B] = [1.24, 86.3]	[p,B] = [1.6, 66.2]	[p,B] = [1.8, 57.8]	Equations[a]
I: 3C → 3D	δ; 1 & 2	s_2'	–	–	1.54	–	1.17	1.1	$(4.1.1)_1$
		$R'_{(2,\perp h)}$	$(\mathrm{km})^{s_2}$	–	0.022	–	0.211	0.324	$(4.1.1)_2$
		$\log R'_{(2,\perp h)}$	–	–	−1.67	–	−0.675	−0.490	–
I: 3D → 3C	δ; 1 & 2	s_1'	–	–	0.1	–	0.48	0.54	$(4.1.2)_1$
		$r'_{(1,t)}$	$\frac{(\mathrm{km})^{s_1}}{\mathrm{km}}$	–	5.38	–	0.153	0.092	$(4.1.2)_2$
		$\log r'_{(1,t)}$	–	–	0.731	–	−0.817	−1.035	–
II: 3B → 3D	δ, L; 2	s_2'	–	–	1.36	1.1	0.85	–	$(4.1.3)_1$
		$R'_{(2,\perp h)}$	$(\mathrm{km})^{s_2}$	–	0.087	0.352	1.33	–	$(4.1.3)_2$
		$\log R'_{(2,\perp h)}$	–	–	−1.062	−0.453	0.123	–	–
II: 3D → 3B	δ, L; 2	m_2'	–	–	1.1	1.36	1.76	–	$(4.1.4)_1$
		$G'_{(2,\perp h)}$	$\frac{(\mathrm{km})^{m_2}}{\mathrm{km}^2}$	–	67.5	56.2	42.0	–	$(4.1.4)_2$
		$\log G'_{(2,\perp h)}$	–	–	1.83	1.75	1.62	–	–
III: 3B → 3C	δ, L: 1 & 2	s_1'	–	0.54	0.36	–	0.23	–	$(4.1.5)_2$
		$r'_{(1,t)}$	$\frac{(\mathrm{km})^{s_1}}{\mathrm{km}}$	0.149	0.384	–	0.790	–	$(4.1.9)_2$
		$\log r'_{(1,t)}$	–	−0.828	−0.415	–	−0.103	–	–
III: 3C → 3B	δ, L: 1 & 2	m_2'	–	1.36	1.54	–	1.86	–	$(4.1.5)_1$
		$G'_{(2,\perp h)}$	$\frac{(\mathrm{km})^{m_2}}{\mathrm{km}^2}$	23.1	26.7	–	28.0	–	$(4.1.8)_2$
		$\log G'_{(2,\perp h)}$	–	1.363	1.426	–	1.448	–	–

B: Calculated parameters for type IV self-consistency tests

Test Type: Source → Plot[a]	Calculated parameter	Units	Value	Equations[a]
IVa: 3C & 3D → 3A	p'	–	1.79	(4.1.10)
	B'	$(\mathrm{km})^p/\mathrm{km}$	7.36	$(4.1.11)_2$
	$\log B'$	–	0.867	–
IVb: 3B & 3D → 3A	p'	–	1.24	(4.1.12)
	B'	$(\mathrm{km})^p/\mathrm{km}$	73.1	$(4.1.13)_2$
	$\log B'$	–	1.864	–
IVc: 3B & 3C → 3A	p'	–	0.67	(4.1.14)
	B'	$(\mathrm{km})^p/\mathrm{km}$	28.2	$(4.1.15)_2$
	$\log B'$	–	1.451	–

[a] Implicitly includes Eqs. (4.1.16)–(4.1.18) where appropriate.

measured in two-dimensions [m_2, $\log g_{(2,\perp h)}$] (Fig. 2B) and the displacement measured in one dimension [s_1, $\log r_{(1,t)}$] (Fig. 2C), with the parameters [p, $\log B$] from Fig. 2A. We find (Table 4B; Fig. 2A)

$$p' = 0.41,\ B' = 32.5\left[(\mathrm{km})^p/\mathrm{km}\right],\ \log B' = 1.512. \tag{4.2.7}$$

This line is plotted as the dashed line in Fig. 2A, and it is obviously a very poor representation of the data distribution. This is not surprising, given the preceding conclusion that the fault-length frequency-distribution parameters seem to be unreliable, but it does reinforce that conclusion, and using (Eq. (4.1.14)), it shows that the value of m_2 is too small.

4.3. Self-consistency tests for the south Yorkshire data

The data available from Watterson et al. (1996) for the south Yorkshire area (Fig. 3) permit self-consistency tests of all nine types, including two tests for each of the Types I through III and the three Type IV tests. For each test, we plot the calculated lines on the appropriate graph in Fig. 3, which allows us to compare the calculated relationship with the observed data and to evaluate the self-consistency of the different measured fault characteristics.

4.3.1. Type I test: same variable, different dimensionality

Watterson et al. (1996) determined the throw systematics from sampling in both one and two dimensions. We converted the throws to displacements using Eq. (2.10) and expressed them in units of km. We use test Type (I: 3C → 3D) given by Eqs. (4.1.1) to test the self-consistency of the parameters [s_1, $\log r_{(1,t)}$], which describe the displacements measured in a one-dimensional domain (Fig. 3C), with the parameters [s_2, $\log R_{(2,\perp h)}$], which describe the displacements measured in a two-dimensional domain (Fig. 3D). We introduce parameter values from Tables 2 and 3 and examine solutions for values of $p = 1.0$, 1.6, and 1.8, which are, respectively, the value consistent with the analysis of Clark and Cox (1996), the best-fit value to the data in Fig. 3A, based on minimizing the sum of the squared orthogonal distances from each point to the line, and a value that makes the lines in Fig. 3C and D appear self-consistent. We find (Table 5A),

$$\begin{aligned}\text{For } [p,B] &= [1.0, 102]: \quad s_2' = 1.54,\\ R'_{(2,\perp h)} &= .022\ \left[\mathrm{km}^{s_2}\right],\ \log R'_{(2,\perp h)} = -1.67,\end{aligned} \tag{4.3.1}$$

$$\begin{aligned}\text{For } [p,B] &= [1.6, 66.2]: \quad s_2' = 1.17,\\ R'_{(2,\perp h)} &= .211\ \left[\mathrm{km}^{s_2}\right],\ \log R'_{(2,\perp h)} = -0.675,\end{aligned} \tag{4.3.2}$$

$$\begin{aligned}\text{For } [p,B] &= [1.8, 57.8]: \quad s_2' = 1.1,\\ R'_{(2,\perp h)} &= 0.324\ \left[\mathrm{km}^{s_2}\right],\ \log R'_{(2,\perp h)} = -0.490.\end{aligned} \tag{4.3.3}$$

These lines are plotted in Fig. 3D as, respectively, the gray short-dash, medium-dash, and dash–dot lines. The line using $p = 1$ is only a fair representation of the data and is worst at the higher

values of displacement. The line using $p = 1.6$ is a relatively good fit to the data, and the line for $p = 1.8$ is a very close fit.

The converse Type I test (I: 3D → 3C) is to use the best-fit results for the two-dimensional displacement data in Fig. 3D, $[s_2, \log R_{(2,\perp h)}]$ (Tables 2 and 3) with Eq. (4.1.2) to calculate the lines that should fit the one-dimensional displacement data in Fig. 3C, $[s_1', \log r'_{(1,t)}]$, (Table 5A).

$$\text{For } [p,B] = [1.0,102]: \quad s_1' = 0.1, \quad r'_{(1,t)} = 5.38\ [\text{km}^{s_1}/\text{km}],\ \log r'_{(1,t)} = 0.731, \tag{4.3.4}$$

$$\text{For } [p,B] = [1.6,66.7]: \quad s_1' = 0.48, \quad r'_{(1,t)} = 0.153\ [\text{km}^{s_1}/\text{km}],\ \log r'_{(1,t)} = -0.815, \tag{4.3.5}$$

$$\text{For } [p,B] = [1.8,57.8]: \quad s_1' = 0.54, \quad r'_{(1,t)} = 0.092\ [\text{km}^{s_1}/\text{km}],\ \log r'_{(1,t)} = -1.035. \tag{4.3.6}$$

These lines are plotted in Fig. 3C as, respectively, the gray short-dash, medium-dash, and dash–dot lines. The first is completely inconsistent with the observed data. The second and third are respectively fair and good representations of the data. A better fit to the data could be obtained from a slightly higher value of p, although the slope of the line then becomes steeper than the best-fit line.

The two Type I tests are self-consistent only if the value of p is about 1.8. If p is close to 1, however, the two sets of data are not self-consistent.

4.3.2. Type II test: different variables, same dimensionality

We next apply the test (II: 3B → 3D) by using the [slope, intercept] pair $[m_2, \log G_{(2,\perp h)}]$ describing two-dimensional fault-length data (Fig. 3B) to calculate $[s_2', \log R'_{(2,\perp h)}]$ describing two-dimensional displacement data (Fig. 3D). We substitute into the right side of Eqs. $(4.1.3)_1$ and $(4.1.3)_2$ from the empirical values in Tables 2 and 3, and we use the three values of $p = 1.0$, 1.24, and 1.6 consistent, respectively, with the analysis of Clark and Cox (1996), a value that makes the lines in Fig. 3C and D appear self-consistent, and the best-fit value to the data in Fig. 3A, based on minimizing the sum of the squared orthogonal distances from each point to the line. We find (Table 5A),

$$\text{For } [p,B] = [1.0,102]: \quad s_2' = 1.36, \quad R'_{(2,\perp h)} = 0.087\ [(\text{km})^{s_2}],\ \log R'_{(2,\perp h)} = -1.062, \tag{4.3.7}$$

$$\text{For } [p,B] = [1.24,86.3]: \quad s_2' = 1.1, \quad R'_{(2,\perp h)} = 0.352\ [(\text{km})^{s_2}],\ \log R'_{(2,\perp h)} = -0.453, \tag{4.3.8}$$

$$\text{For } [p,B] = [1.6,66.7]: \quad s_2' = 0.85, \quad R'_{(2,\perp h)} = 1.33\ [(\text{km})^{s_2}],\ \log R'_{(2,\perp h)} = 0.123. \tag{4.3.9}$$

These solutions are plotted in Fig. 3D as, respectively, the gray dash–dot–dot, long-dash–long-dash–dot, and dash–dash–dot lines. All these lines fall reasonably close to at least part of the data, but the value of $p = 1.24$, for which the calculated and the best-fit slopes are the same, gives a very close approximation to the best-fit line (cf. Table 3).

The converse of this Type II self-consistency test is to calculate the lines for the fault-length distribution $[m_2', \log G'_{(2,\perp h)}]$ (Fig. 3B) from the distribution of fault-displacement $[s_2, \log R_{(2,\perp h)}]$ (Fig. 3D), where both are measured in a two-dimensional domain, test (II: 3D → 3B). Substituting into the right side of Eqs. $(4.1.4)_1$ and $(4.1.4)_2$ from the values in Tables 2 and 3, we find (Table 5A),

$$\text{For } [p,B] = [1.0,102]: \quad m_2' = 1.1, \quad G'_{(2,\perp h)} = 67.5\ [(\text{km})^{m_2}],\ \log G'_{(2,\perp h)} = 1.830, \tag{4.3.10}$$

$$\text{For } [p,B] = [1.24,86.3]: \quad m_2' = 1.36, \quad G'_{(2,\perp h)} = 56.2\ [(\text{km})^{m_2}],\ \log G'_{(2,\perp h)} = 1.750, \tag{4.3.11}$$

$$\text{For } [p,B] = [1.6,66.7]: \quad m_2' = 1.76, \quad G'_{(2,\perp h)} = 42.0\ [(\text{km})^{m_2}],\ \log G'_{(2,\perp h)} = 1.62. \tag{4.3.12}$$

These lines are plotted in Fig. 3B as, respectively, the gray short-dash, medium-dash, and long-dash lines. All the lines plot relatively close to the data, but the value of $p = 1.24$, for which the calculated slope m_2 is equal to the best-fit slope, is the closest fit, with an intercept that is only slightly too high (cf. Table 3). Thus the data measured in a two-dimensional domain for both fault-length and fault-displacement are self-consistent, especially with a value of $p = 1.24$.

4.3.3. Type III test: different variables, different dimensionality

We next test the compatibility of the [slope, intercept] pair $[m_2, \log G_{(2,\perp h)}]$ (Fig. 3B) with $[s_1, \log r_{(1,t)}]$ (Fig. 3C). For the test (III: 3B → 3C), we substitute for the parameters on the right sides of Eqs. $(4.1.5)_2$ and $(4.1.9)_2$ from the empirical values in Tables 2 and 3. We take $p = 0.67$, 1.0, and 1.6, which are, respectively, the value for which the slopes of the data in Fig. 3B and C appear to be self-consistent, the value consistent with the Clark and Cox (1996) analysis, and the value that is the best-fit slope for the data in Fig. 3A (Table 5A).

$$\text{For } [p,B] = [0.67,130]: \quad s_1' = 0.54, \quad r'_{(1,t)} = 0.149\ [(\text{km})^{s_1}/\text{km}],\ \log r'_{(1,t)} = -0.828, \tag{4.3.13}$$

$$\text{For } [p,B] = [1.0,102]: \quad s_1' = 0.36, \quad r'_{(1,t)} = 0.384\ [(\text{km})^{s_1}/\text{km}],\ \log r'_{(1,t)} = -0.415, \tag{4.3.14}$$

$$\text{For } [p,B] = [1.6,66.2]: \quad s_1' = 0.23, \quad r'_{(1,t)} = 0.790\ [(\text{km})^{s_1}/\text{km}],\ \log r'_{(1,t)} = -0.103. \tag{4.3.15}$$

These results are plotted in Fig. 3C as the gray dot–dash–long-dash, dot–dot–dash and dash–dash–dot lines, respectively. All provide a poor approximation to the one-dimensional displacement data, with the approximation becoming worse with increasing values of p.

The converse of this test (III: 3C → 3B) is to calculate the lines for the fault-length distribution, $[m_2', \log G'_{(2,\perp h)}]$(Fig. 3B), from the distribution of fault-displacement $[s_1, \log r_{(1,t)}]$ (Fig. 3C). Substituting into the right sides of Eqs. $(4.1.5)_1$, and $(4.1.8)_2$ from the values in Table 3, we find,

$$\text{For } [p,B] = [0.67,129.7]: \quad m_2' = 1.36, \quad G'_{(2,\perp h)} = 23.1\ [(\text{km})^{m_2}],\ \log G'_{(2,\perp h)} = 1.363, \tag{4.3.16}$$

$$\text{For } [p,B] = [1.0,102]: \quad m_2' = 1.54, \quad G'_{(2,\perp h)} = 26.7\ [(\text{km})^{m_2}],\ \log G'_{(2,\perp h)} = 1.426, \tag{4.3.17}$$

$$\text{For } [p,B] = [1.6,66.7]: \quad m_2' = 1.86, \quad G'_{(2,\perp h)} = 28.0\ [(\text{km})^{m_2}],\ \log G'_{(2,\perp h)} = 1.448. \tag{4.3.18}$$

These results are plotted in Fig. 3B as, respectively, the a gray dot–dash–long-dash, dash–dot–dot, and dash–dash–dot lines.

All lie relatively close to the plotted data, but generally at frequencies that are low relative to the linear part of the data. The value of $p = 0.67$ is an unacceptably low value for the data in Fig. 3A, even though the slope m_2' in Fig. 3B is the same as that for the length data. The slope m_2' for $p = 1.6$ is too high. Thus, of these three possibilities, the line for $p = 1.0$ provides the closest, although not a particularly good, fit to the data.

4.3.4. Type IV tests: self-consistency tests for the parameters p, m_ζ and s_ζ

IVa: We use the parameters from the fits to displacement data measured in one and two dimensions, [s_1, log $r_{(1,t)}$] and [s_2, log $R_{(2,\perp h)}$] (Fig. 3C and D), to test the self-consistency of p' and B' (IVa: 3C & 3D → 3A). From Eqs. (4.1.10) and (4.1.11)$_2$ and the results in Table 3, we find,

$$p' = 1.8, \quad B' = 7.36, \quad \log B' = 0.867. \tag{4.3.19}$$

IVb: We can use Eqs. (4.1.12) and (4.1.13)$_2$ with the parameters from fault length, [m_2, log $G_{(2,\perp h)}$] (Fig. 3B), and displacement, [s_2, log $R_{(2,\perp h)}$] (Fig. 3D), both measured in two-dimensional domains, to apply the test (IVb:3B & 3D → 3A). Using the parameter values from Table 3, we find,

$$p' = 1.24, \quad B' = 73.1, \quad \log B' = 1.864. \tag{4.3.20}$$

IVc: Finally, we use Eq. (4.1.14) and (4.1.15)$_2$ with the parameters from the fault length data [m_2, log $G_{(2,\perp h)}$] (Fig. 3B) and the displacement data [s_1, log $r_{(1,t)}$] (Fig. 3C) to apply the test (IVc: 3B & 3C → 3A). We use parameter values from Table 3 to find,

$$p = 0.67, \quad B' = 28.2, \quad \log B' = 1.451. \tag{4.3.21}$$

The first and third of these solutions, Eqs. (4.3.19) and (4.3.21) plotted as the gray dash–dot–dot and dash–dot lines, respectively, provide unacceptable fits to the data in Fig. 3A. The second solution (Eq. (4.3.6) (gray dashed line)) provides an acceptable fit to the data in Fig. 3A. This result is consistent with the two-dimensional data for fault-length and displacement being self-consistent.

4.4. Evaluation of the self-consistency tests

The self-consistency tests demonstrate the lack of internal consistency among the parameters in the equations describing the fault-length and fault-displacement systematics as measured in one- and two-dimensional domains in both the Yucca Mountain and the south Yorkshire areas.

4.4.1. Yucca Mountain

For the Yucca Mountain area, the displacement vs. frequency data measured from a one-dimensional domain (Fig. 2C) constitute by far the most robust data set among those we consider in this paper. The data result from the combination of measurements at three different scales, and they span a remarkable five orders of magnitude in displacement, compared to spans of one to three orders of magnitude for the other data sets (Figs. 2 and 3). The empirical parameters that describe this displacement-frequency distribution are therefore particularly robust.

The one-dimensional displacement data are reasonably successful in accounting for the length data measured in a two-dimensional domain if p has a value of about 1 (test Type (III: 2C → 2B)). Higher values of p do not give as good a fit (Fig. 2B). The parameters from the two-dimensional length data, however, are not successful in accounting for the one-dimensional displacement data for any value of p (Fig. 2C). Moreover, for the test Type (IVc: 2B & 2C → 2A), the combination of the parameters from the two-dimensional length data and the one-dimensional displacement data is not successful in accounting for the displacement–length distribution (Fig. 2A).

These results are consistent with a value of $p = 1$ and with the inference that there is a bias in the two-dimensional length data such that at the shorter fault-lengths, measurements underdetermine either the fault-lengths or the fault frequencies, or both.

4.4.2. South Yorkshire

There is no single value of p that makes all four data sets in Fig. 3 appear self-consistent. The calculation of the self-consistency tests (Eqs. (4.1.1)–(4.1.5), (4.1.8), (4.1.9)) all depend on the values of p and B that relate the fault-length to the fault-displacement. The difficulty in drawing firm conclusions from these tests is at least in part a result of the ambiguity in the values of [p, B]. Although it is clear that the parameters describing these data sets are not mutually consistent, without a clear and unique value for [p, B], it is not obvious which of the data sets is unreliable.

We can summarize the results as follows:

1) The two-dimensional length and displacement data (II: 3B → 3D), (II: 3D → 3B), (IVb: 3B & 3D → 3A) are self-consistent for $p = 1.24$.
2) The displacement data from one-dimensional and two-dimensional domains appear self-consistent if $p = 1.8$ (test Types (I: 3C → 3D) and (I: 3D → 3C)); but these two data sets together are not consistent with the displacement–length systematics (test Type IVa: 3C & 3D → 3A), indicating one of the sets at least is unreliable.
3) The one-dimensional displacement data is moderately successful in accounting for the distribution of the two-dimensional length data (III: 3C → 3B), although the intercepts are somewhat low. The parameters for the two-dimensional length data, however, do not account well for the one-dimensional displacement data for any value of p (III: 3B → 3C). This asymmetry in the self-consistency of the one-dimensional displacement data and the two-dimensional length data suggests that the length data may be affected by some measurement bias.
4) Despite the self-consistency noted in items 2 and 3 above, there is failure of the self-consistency tests that combine empirical parameters describing displacement data from one-dimensional domains with parameters describing both length and displacement data from two-dimensional domains (IVa: 3C & 3D → 3A) and (IVc: 3B & 3C → 3A). The predicted values of [p, B] do not account well for the data in Fig. 3A.

The analysis of Clark and Cox (1996) indicates that within the limits of statistical inference, a value of $p = 1$ can account for all the fault-length vs. fault-displacement distributions. Nevertheless, the possibility that $p > 1$ cannot be completely dismissed, as the individual data sets are not necessarily best-fit by a slope of $p = 1$. There is no value of p, however, that allows all the data sets in Fig. 3 to be mutually consistent, and different pairs of data sets require different values of [p, B] to make them appear self-consistent. We conclude, therefore, that one or more of the data sets must be unreliable.

The reliable determination of fault-lengths in a two-dimensional analysis is inherently more difficult than measurement of displacement, because the splaying and anastomosing structure of faults, as well as the incompleteness of exposure, can affect the inference as to the number and length of faults (see Kim and Sanderson, 2005). This difficulty introduces the potential for a systematic bias in these measurements, which would render unreliable the parameters in the equations used to fit the data for

fault-length systematics. This same problem could make it difficult reliably to identify separate faults for which to measure the displacement, thereby introducing bias into the displacement frequency distribution measured in two-dimensional domains.

We tentatively conclude, therefore, that the parameters derived from a one-dimensional analysis of displacement, with the theoretical relations among the parameters that we have derived, provide the most reliable descriptions of the fault length and displacement systematics. This also is consistent with the conclusion from the analysis of the Yucca Mountain data.

5. Constraints on the volumetric strains from sampling in one- and two-dimensional domains

In the companion paper (Part I), we showed that the strains inferred from sampling in domains of one and two dimensions in general provide constraints, but not necessarily exact measures, of the three-dimensional strains of the associated volume (Eqs. (I:4.2.2), (I:4.3.2), and (I:4.3.4)). As an example of the constraints that two-dimensional sampling places on the three-dimensional strains, we use parameter values for the Yucca Mountain area from Eqs. $(2.6)_1$ and $(2.9)_1$ in Eqs. (I:3.2.6) and (I:3.2.8), and substitute those values in Eqs. (I:4.2.2) to find:

$$\left.\begin{array}{l}\text{for } p = 1 \text{ and } s_1 = 0.73, \text{ then } s_2 = 1.73 \text{ and} \\ e^{(\mathrm{tot})}_{(3,x)} \geq 0.95 e^{(\mathrm{tot})}_{(2,x)}, \text{ or } e^{(\mathrm{tot})}_{(2,x)} \leq 1.05 e^{(\mathrm{tot})}_{(3,x)},\end{array}\right\} \tag{5.1}$$

where the extensional or shear strains are given by setting the subscript "x" equal to either "t" or "γ", respectively. The equality holds if the two-dimensional sampling actually includes the largest fault in the volume, $\delta^{(\max)}_{(2,\perp n)} = \delta^{(\max)}_{(3)}$. In this case, the extensional and shear strains of the volume $\mathcal{V}$ are 95% of the strains determined for the two-dimensional domain (Eq. $(5.1)_4$). If the two-dimensional sampling does not include the largest fault, then the inequality in Eq. $(5.1)_4$ applies, and 95% of the strain determined for the two-dimensional domain is a lower bound for the actual strain of the volume $\mathcal{V}$.

As an example of the constraints that one-dimensional sampling places on the three-dimensional strains, we use values for the Yucca Mountain area from Eqs. $(2.6)_1$ and $(2.9)_1$ in Eqs. (I:3.2.6) and (I:3.2.8), and then substitute the results into Eqs. (I:4.3.2) and (I:4.3.4) to find:

$$\left.\begin{array}{l}\text{for } p = 1 \text{ and } s_1 = 0.73, \text{ then, } s_3 = 2.73, \text{ and} \\ e^{(\mathrm{tot})}_{(3,t)} \geq 0.8 e^{(\mathrm{tot})}_{(1,t)}, \quad e^{(\mathrm{tot})}_{(1,t)} \leq 1.25 e^{(\mathrm{tot})}_{(3,t)}, \\ e^{(\mathrm{tot})}_{(3,\gamma)} \geq 0.5 e^{(\mathrm{tot})}_{(1,\gamma)}, \quad e^{(\mathrm{tot})}_{(1,\gamma)} \leq 2.0 e^{(\mathrm{tot})}_{(3,\gamma)}.\end{array}\right\} \tag{5.2}$$

If the one-dimensional sampling actually includes the largest fault in the volume ($\delta^{(\max)}_{(1,t)} = \delta^{(\max)}_{(3)}$), then the equality holds for the extension (Eqs. $(5.2)_{4,5}$), and the total extension of the volume $\mathcal{V}$ is 80% of the extension measured in the one-dimensional domain. In general, however, the sampling of a one-dimensional domain will not include the largest fault, so for these particular parameter values, 80% of the one-dimensional total extension is a lower bound for the actual extension of the volume. The constraint imposed by Eq. (I:4.3.4) on the shear strain, however, is less stringent, and regardless of whether or not the largest fault sampled in the one-dimensional domain is the largest fault in the volume, we can only conclude that the actual shear strain for the volume $\mathcal{V}$ is greater than half the total shear strain inferred from two orthogonal one-dimensional samplings (Eq. $(5.2)_6$).

6. Test of the theory for extensional strain

Eqs. (I:4.1.48) through (I:4.1.50) allow us to calculate the cumulative fractional extension (setting subscript 'x' = 't') or shear strain (setting subscript 'x' = 'γ') as a function of the displacement magnitude on faults in a faulted domain. In this section, we test the theory by comparing the displacement data gathered from a 4777 m-long linear transect at Yucca Mountain, Nevada, with the theoretical prediction of Eq. $(I:4.1.50)_1$. We also compare the measured extension of the one-dimensional domain to the extension predicted by Eq. (I:4.1.48) for the volume sampled by that domain.

Along the transect at Yucca Mountain, the displacements were measured on all 376 faults intersected by the transect for which the displacement was greater than or equal to 4 cm. The faults generally dip at $\rho \approx 75^\circ$; the faults are normal faults with a slip direction that is approximately down-dip; and the transect orientation $\mathbf{t}$ is perpendicular to the average strike of the faults, so $\kappa \approx 0^\circ$ whereby $\phi \approx \rho \approx 75^\circ$ (Figure I:1A, B).

The largest fault has a displacement of 125 m, so,

$$\delta^{(\max)}_{(1,t)} = 125 \text{ m}. \tag{6.1}$$

The one-dimensional extension contributed by the largest fault is then, from Eq. (I:4.1.3) with $(I:4.1.9)_1$ and with $i = 1$,

$$e^{(1)}_{(1,t)} = \frac{(125 \text{ m})\cos 75^\circ}{4777 \text{ m}} = 0.007. \tag{6.2}$$

Summing all the observed displacements gives a total of

$$\delta^{(\mathrm{tot})}_{(1,t)} = \sum_{i=1}^{N^{(\max)}_{(1,t)}=376} \delta^{(i)}_{(1,t)} = 422 \text{ m}, \tag{6.3}$$

which gives a total one-dimensional extension (from Eqs. (I:4.1.3) with $(I:4.1.9)_1$) of

$$\begin{aligned} e^{(\mathrm{tot})}_{(1,t)} &= \sum_{i=1}^{N^{(\max)}_{(1,t)}=376} e^{(i)}_{(1,t)} = \sum_{i=1}^{N^{(\max)}_{(1,t)}=376} \frac{\delta^{(i)}_{(1,t)} \cos\phi}{\mathcal{T}} \\ &= \frac{\cos\phi}{\mathcal{T}} \sum_{i=1}^{N^{(\max)}_{(1,t)}=376} \delta^{(i)}_{(1,t)}, \end{aligned} \tag{6.4}$$

$$e^{(\mathrm{tot})}_{(1,t)} = \frac{\cos 75^\circ}{4777 \text{ m}} 422 \text{ m} = 0.023. \tag{6.5}$$

The theoretical total extension for one-dimensional sampling can be calculated from the extension contributed by the largest-displacement fault and from the parameters fit to the data distribution in Fig. 2C, using Eqs. $(I:4.1.42)_1$, $(2.9)_1$, and (6.2):

$$e'^{(\mathrm{tot})}_{(1,t)} = e^{(1)}_{(1,t)} \left[\frac{1}{1 - s_1}\right] = .007 \left[\frac{1}{1 - .073}\right] = 0.025, \tag{6.6}$$

where we use a prime on the symbol for the total extension to distinguish calculated quantities from those determined directly from field measurements. To be consistent with this total strain, the total displacement should be,

$$e'^{(\mathrm{tot})}_{(1,t)} = \frac{\cos 75^\circ}{4777 \text{ m}} \delta'^{(\mathrm{tot})}_{(1,t)} = 0.025, \quad \delta'^{(\mathrm{tot})}_{(1,t)} = 461 \text{ m}. \tag{6.7}$$

These values for the theoretical one-dimensional extension and theoretical total displacement exceed that observed from the transect by 0.002 and 39 m, respectively. This excess can be

ascribed to faults with a displacement of less than 4 cm, which is the cutoff used for measurements along the transect, and it reflects the accumulation of extension and displacement that was not sampled in the measurements.

The theoretical total extension for the volume can be calculated from Eqs. (I:4.1.40), and (I:4.1.1) with $i = 1$,

$$e'^{(\mathrm{tot})}_{(3,t)} = e^{(1)}_{(3,t)}\left(\frac{1+2/p}{1-(s_3-2/p)}\right) = {}_3\mathfrak{P}^{(1)}_t\left[\frac{\delta^{(\max)}_{(3)}\cos\phi}{\mathcal{T}}\right]\left(\frac{1+2/p}{1-s_1}\right), \tag{6.8}$$

where we used Eq. (I:3.2.6) to get the last expression in Eq. (6.8). To evaluate Eq. (6.8), we use the empirical values for p and s_1 given in Eqs. (2.1) and $(2.9)_1$, and we assume that the largest observed fault intersected by the transect is the largest fault in the volume, which implies, from Eqs. (I:3.3.8),

$$\delta^{(\max)}_{(3)} = \delta^{(\max)}_{(1,t)}, \tag{6.9}$$

$${}_3\mathfrak{P}^{(1)}_t = \frac{s_1}{s_3} = \frac{0.73}{2.73} = 0.27, \tag{6.10}$$

$$e'^{(\mathrm{tot})}_{(3,t)} = 0.27\left[\frac{125\cos 75^\circ}{4777}\right]\left(\frac{1+2}{1-0.73}\right) = 0.020. \tag{6.11}$$

Thus if the largest fault observed in the transect is also the largest fault in the volume, the total extension for the three-dimensional volume (0.020) is predicted to be 80% of the total extension (0.025) predicted for the one-dimensional transect (the equality in Eq. $(5.2)_1$). There is a finite probability, however, that the transect does not intersect the largest fault in the volume, in which case the inequality in Eq. $(5.2)_1$ applies, and the actual total extension of the volume could be greater than that given by Eq. (6.11).

Both the total extension determined from the observed displacements on an incomplete set of the faults (0.023; Eq. (6.5)) and the predicted total extension calculated for one-dimensional sampling from the fault systematics (0.025; Eq. (6.6)) are greater than the predicted total extension for the volume calculated from the fault systematics (0.020; Eq. (6.11)). This difference is the result of the different values of the weighting parameter ${}_\zeta\mathfrak{P}^{(i)}_t$ for domains of different dimensionality (Eqs. (I:4.1.1) and (I:4.1.3)). For the extension of the volume, the contribution of each fault to the extension of the volume must be weighted by the size of the fault, as defined by the weighting factor ${}_3\mathfrak{P}^{(i)}_t$ in Eqs. (I:4.1.1) and $(I:4.1.7)_1$. In the calculation of the extension for one-dimensional sampling (Eq. (I:4.1.3)), all faults, no matter what the size, have the same weight (Eq. (I:4.1.9)). Thus in the one-dimensional sampling, the effect of the different-sized faults on the extension of the whole volume is not accounted for, and it therefore only provides a limit on the real extension of the volume.

We plot the observed data from the Yucca Mt. area on a graph of the cumulative fractional one-dimensional extension vs. the displacement for the set of observed faults, which are ordered in the sequence of decreasing displacement (Fig. 4). To this end, we sum Eq. $(I:4.1.13)_1$ to obtain the measured cumulative extension, and divide by the theoretical total extension from Eq. (6.6):

$$\frac{e^{(\mathrm{cum})}_{(1,t)}}{e'^{(\mathrm{tot})}_{(1,t)}} = \frac{1}{0.025}\sum_{i=1}^{N}\left[\frac{\delta^{(i)}\cos\phi}{\mathcal{T}}\right] = \frac{\cos 75^\circ}{0.025{*}4777}\sum_{i=1}^{N}\delta^{(i)}, \tag{6.12}$$

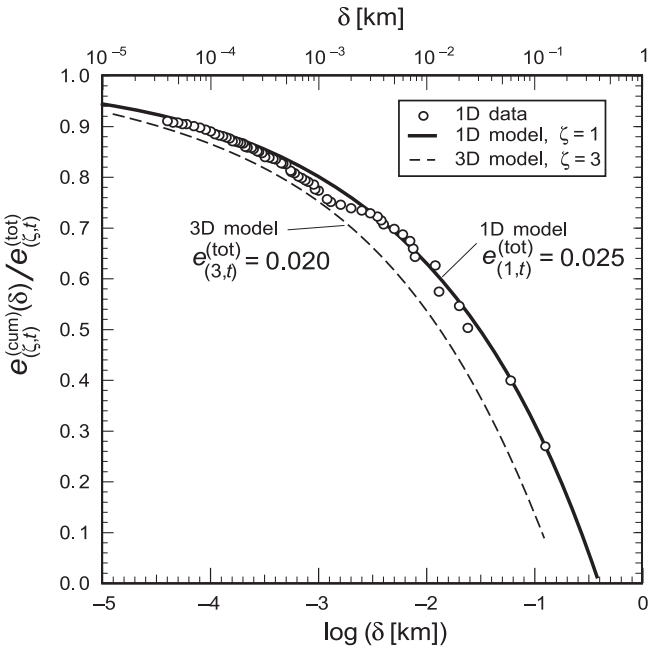

Fig. 4. Logarithm of the displacement plotted against the cumulative strain expressed as a fraction of the total strain for a 4777 meter-long fault-displacement survey along a one-dimensional domain normal to fault strike at Yucca Mountain, Nevada. The solid and broken curves, respectively, are the theoretical predictions for the fractional cumulative extension of a one-dimensional ($\zeta = 1$) and a three-dimensional ($\zeta = 3$) domain. Data include all the faults ($N = 376$) that have a displacement of at least 4 cm. Smaller displacements were not recorded. Faults have an average dip of 75°. The maximum observed displacement on a single fault is $\delta^{(\max)}_{(1,t)} = 125$ m; the total observed displacement is $\sum_{i=1}^{376}\delta^{(i)}_{(1,t)} = 422$ m. The theoretical total extension is calculated for the volume from Eq. (I:4.1.40) with $x = t$ and for the one-dimensional domain from Eq. $(I:4.1.42)_1$, assuming for both cases that the largest observed fault is the largest in the domain. The subscript "x" on the strain symbols $e_{(3,x)}$ stands for either the subscript "t" indicating the extensional strain or "γ" indicating the shear strain. For $\zeta = 3$, we find $e^{(\mathrm{tot})}_{(3,x)} = 0.020$, which is a minimum value, given the probability that the largest observed fault in the one-dimensional sampling domain is not the largest fault in the volume. For $\zeta = 1$, we find $e^{(\mathrm{tot})}_{(1,x)} = 0.025$, which is the number used in normalizing the observed data for plotting. The theoretical curves are from Eqs. (I:4.1.48) with $x = t$, and Eq. $(I:4.1.50)_1$, using parameter values of $p = 1.0$ (Eqs. $(2.5)_1$) and $s_1 = 0.73$ (Eq. $(2.9)_1$). The dashed curve for the three-dimensional sampling ($\zeta = 3$) is the same as the solid black curvefor $s_3 = 2.73$ in Fig. 5, except that in this plot, the displacement on the abscissa is not normalized by the maximum displacement.

$$\frac{e^{(\mathrm{cum})}_{(1,t)}}{e'^{(\mathrm{tot})}_{(1,t)}} = 0.0022\sum_{i=1}^{N}\delta^{(i)}, \quad N \le N^{(\max)} = 376. \tag{6.13}$$

In Fig. 4, we compare the plotted data to two curves, one calculated for one-dimensional cumulative fractional extension (Eq. $(I:4.1.50)_1$) and the other calculated for three-dimensional cumulative fractional extension (Eq. (I:4.1.48) with $x = t$). For both calculations we use the values for p and s_1 measured for the displacement distribution in the Yucca Mountain faulted terrane (Fig. 2C; Eqs. (2.1) and $(2.9)_1$), and we use the values for the total extension from Eq. (6.6) for the equation for one-dimensional sampling, and from Eq. (6.11) for the equation for three-dimensional sampling,

$$\frac{e'^{(\mathrm{cum})}_{(1,t)}(\delta)}{e'^{(\mathrm{tot})}_{(1,t)}} = 1 - s_1\left[\frac{\delta}{\delta^{(\max)}_{(1,t)}}\right]^{(1-s_1)} = 1 - 0.73\left[\frac{\delta}{125}\right]^{0.27},$$

$$\frac{e'^{(\mathrm{cum})}_{(1,t)}(\delta)}{e'^{(\mathrm{tot})}_{(1,t)}} = 1 - 0.20\,\delta^{0.27}, \tag{6.14}$$

$$\frac{e'^{(\mathrm{cum})}_{(3,t)}(\delta)}{e'^{(\mathrm{tot})}_{(3,t)}} = 1 - \frac{s_3 p}{p+2}\left(\frac{\delta}{\delta^{(\mathrm{max})}_{(3)}}\right)^{(1-s_1)} = 1 - \frac{(0.73+2)1}{1+2}\left[\frac{\delta}{125}\right]^{0.27},$$

$$\frac{e'^{(\mathrm{cum})}_{(3,t)}(\delta)}{e'^{(\mathrm{tot})}_{(3,t)}} = 1 - 0.25\,\delta^{0.27}. \tag{6.15}$$

Both curves clearly must approach a fractional cumulative extension of 1 as the displacement approaches 0, but the total extension represented by that value is smaller for the lower bound to the total extension of the volume (=0.020) than for the total extension of the one-dimensional domain (=0.025). The curve for the one-dimensional model (Eq. (6.14)) provides an excellent fit for the data sampled from a one-dimensional domain. Thus we conclude that our theoretical model is an accurate representation of the distribution of fault-displacement. The curve for the three-dimensional model is, at every value of δ, a slightly smaller fraction of the total extension than the curve for the one-dimensional model. This implies that in the three-dimensional model, the faults with smaller displacements contribute a larger fraction of the total extension than in the one-dimensional model.

7. Implications of the empirical data for strain calculations

The ultimate goal of this analysis is to provide a sound theoretical foundation for estimating the total extension or shear strain magnitudes in a domain of distributed faulting without having to measure the displacement on every fault in the domain, to evaluate the relative contribution to that strain of small and large faults, and to illustrate the tests that can be applied to evaluate the self-consistency of different data sets sampled in one- and two-dimensional domains. In this section, we discuss the implications of the parameters that we have examined for the Yucca Mountain and the south Yorkshire areas for inferring strains in faulted domains.

The equations for determining the total extension and shear strain in terms of the strain contributed by the largest fault are given by Eqs. (I:4.1.40) through (I:4.1.42). We can find the values for the parameters s_3 and s_2 from the empirical values of s_1 (Table 3; Eqs. $(2.9)_1$ and $(2.17)_1$) by using Eqs. (I:3.2.6) and (I:3.2.8) and assuming for present purposes that $p = 1$ (Table 2; Eq. (2.1)).

$$\left.\begin{array}{l}\text{For Yucca Mt. (Fig. 5B),}\quad s_2 = 1.73,\ s_3 = 2.73,\\ \text{For S. Yorkshire (Fig. 6C),}\quad s_2 = 1.54,\ s_3 = 2.54.\end{array}\right\} \tag{7.1}$$

These values of s_ζ are all consistent with the theoretically necessary constraints given in Eq. (I:4.1.45).

We can calculate the fraction of the total extension or shear strain contributed by the largest fault in a domain from Eq. $(\mathrm{I{:}4.1.43})_1$. For the Yucca Mountain and the south Yorkshire areas, respectively, we adopt the parameter values $p = 1$, and $s_3 = 2.73$ and 2.54 (Eq. (2.1), and Eqs. $(7.1)_2$ and $(7.1)_4$). Using these values in Eq. $(\mathrm{I{:}4.1.43})_1$ gives

$$\frac{e^{(1)}_{(3,x)}}{e^{(\mathrm{tot})}_{(3,x)}} = \left\{\begin{array}{l}0.09\ \text{(Yucca Mt.) and}\\ 0.15\ \text{(S. Yorkshire)}\end{array}\right\}, \tag{7.2}$$

which implies the largest fault contributes only 9% to the total extension in the Yucca Mountain area and 15% in the south Yorkshire area. Thus for higher values of s_3, there is a smaller contribution to the extension or the shear strain from the largest fault. In both of these cases, the smaller faults would be essential to an accurate determination of extension or shear strain.

The fractional cumulative extension of a volume $\mathcal{V}$ as a function of the fractional displacement is given by Eq. (I:4.1.48). From Eq. (I:3.2.6), we see that the exponent on the variable δ in this equation is independent of p,

$$1 - (s_3 - 2/p) = 1 - s_1, \tag{7.3}$$

whereby Eq. (I:4.1.48), expressed in terms of s_1, becomes,

$$\frac{e^{(\mathrm{cum})}_{(3,x)}(\delta)}{e^{(\mathrm{tot})}_{(3,x)}} = 1 - \frac{s_1 p + 2}{p+2}\left(\frac{\delta}{\delta^{(\mathrm{max})}_{(3)}}\right)^{[1-s_1]}, \tag{7.4}$$

Eq. (7.4) shows that the particular value of p has only a minor influence on the form of the curve, and this minor influence originates from the coefficient of the displacement term. Taking $s_1 = 0.73$, we see that Eq. (I:3.2.6) or Eq. (7.3) can be rewritten as

$$s_3 = 0.73 + \frac{2}{p}. \tag{7.5}$$

We illustrate these relations in Fig. 5, in which we plot Eq. (7.4), or Eq. (I:4.1.48) for s_3 given by Eq. (7.5), for three values of p. The minor differences in this set of curves for the three values of p result from the fact that if we know the value for s_1, then from Eq. (7.3), the exponent on the displacement term in Eq. (I:4.1.48) becomes constant and independent of p. From Eq. (7.3) with $s_1 = 0.73$,

$$[1 - (s_3 - 2/p)] = [1 - s_1] = [1 - 0.73] = 0.27. \tag{7.6}$$

The coefficient of the displacement term in Eq. (7.4) becomes $(0.73p+2)/(p+2)$, which causes the curve to change very little with changes in p when it is plotted on a semi-logarithmic plot.

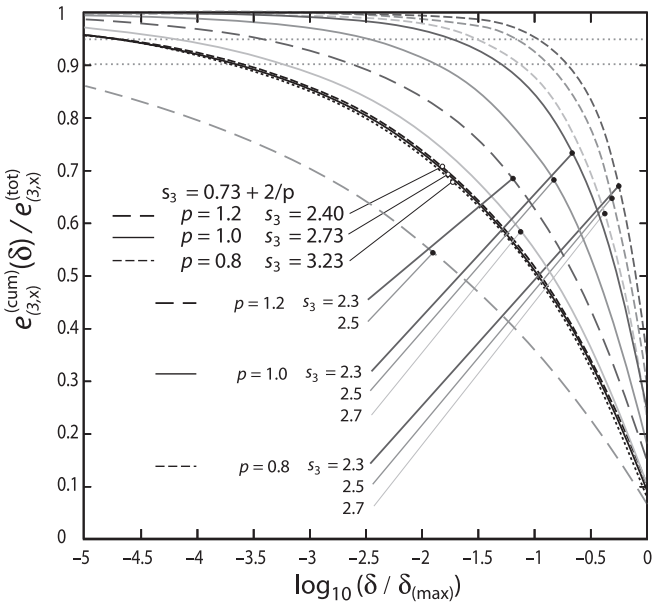

Fig. 5. Logarithm of the fractional displacement plotted against the fractional cumulative strain for three-dimensional sampling ($\zeta = 3$). The subscript "x" on the strain symbols $e_{(3,x)}$ stands for either the subscript "t" indicating the extensional strain or "γ" indicating the shear strain. For the three very closely clustered curves, Eq. (7.5) defines the value of s_3 for the given value of p. The solid line of these three is the same as the dashed line in Fig. 4, except that here the displacement is plotted as a fraction of the total displacement. Horizontal gray dotted lines define the points at which the cumulative strain reaches 90% and 95% of the total strain. The value of $s_3 = 2.7$ with $p = 1.2$ does not satisfy the constraint in Eq. (I:4.1.45) and so is not plotted.

These results indicate that, in order to calculate the cumulative strain as a function of displacement, the best parameter to determine empirically is s_1 because the results are then almost independent of p.

We also plot Eq. (I:4.1.48) in Fig. 5 for a variety of parameter values p and s_3 that are chosen independent of each other and that do not satisfy Eq. (7.5). Clearly in these cases there can be a large range of shapes to the curve, which in turn suggest a large range in the magnitude of faults required to account for 95% of the extension or shear strain. Thus small errors in the value of either p or s_3 can lead to vast differences in the inferred importance of small faults to the total strain, especially if these two parameters do not satisfy the relation in Eq. (I:3.2.6), or in the case of Yucca Mountain area, Eq. (7.5).

Fig. 5 shows that, for the preferred parameter values for the Yucca Mountain area, as included in Eq. (7.5), in order to account for 95% (or 90%), of the extension in the volume, we must account for all the faults having displacements from the maximum down to approximately 4.7 (or 3.6) orders of magnitude smaller than the maximum displacement. The relations, however, are very sensitive to the values of p and s_3, as is evident from the curves for the selected range of parameter values that do not satisfy Eq. (7.5). Thus if s_3 is determined from any two-dimensional data, for example through Eq. (I:3.2.8) or Eq. $(\text{I:3.2.10})_2$ with (I:3.2.6), then the exponent in Eq. (I:4.1.48) will depend on p, and any uncertainty in the value of p will strongly affect how the importance of small faults on the total strain would be interpreted. This constitutes a strong argument in favor of directly determining s_1 from the measurement of displacement in a one-dimensional domain, and we conclude that this should be the preferred method for assessing strain from fault systematics.

Much discussion exists in the literature concerning the best values to use for the parameters such as p, B, m_3, and s_3, and the statistics of sampling is an important part of the discussion. The estimates of the parameter values given by Eqs. (2.1)–(2.9), and Eqs. (2.11)–(2.18) from Figs. 1–3 illustrate the uncertainty of these values. The different values of s_3 for the Yucca Mountain and south Yorkshire areas, for example, indicate that this parameter should be determined for each specific fault set for which the strain is to be determined. In fact, the different fault sets in the south Yorkshire area alone give different values of s_3 (Watterson et al., 1996).

The self-consistency tests that we have done on data collected in the Yucca Mountain and the south Yorkshire areas for fault-displacement in one and two-dimensional sampling, and for fault-length in two-dimensional sampling, show that the values inferred for the parameters in the equations for fault systematics are not self-consistent. For the Yucca Mountain area, the one-dimensional sampling of fault-displacement proves to be the most self-consistent, because the fault-length distribution in a two-dimensional domain predicted from these parameters with $p = 1$ fall near the observed data, but the parameters inferred from the plots of the length data from sampling in a two-dimensional domain provides a bad approximation to displacements observed from one-dimensional sampling for any value of p. This result suggests that there is a systematic bias in the collection of fault-length data in two-dimensional domains.

The inferences for the south Yorkshire area are less clear-cut, because the best value to use for p is ambiguous. It is clear, however, that the slope of the one-dimensional displacement data (Fig. 3C) is inconsistent with the slopes for the two-dimensional length data (Fig. 3B) and displacement data (Fig. 3D), because the self-consistency tests of Type (IVa: 3C & 3D → 3A) and Type (IVc: 3B & 3C → 3A) are not satisfied (Fig. 3A). It is difficult to infer objectively, however, which of the data sets is unreliable.

Despite the lack of self-consistency among the parameters, the results based on Eq. (7.4) indicate the importance of including the contributions of smaller faults in the determination of strain in a faulted terrane.

8. Summary

A review of some recent determinations of the parameters defining the fault population systematics from Yucca Mountain, Nevada, U.S.A. and from south Yorkshire, U.K. shows that many of the sets of parameter values are not self-consistent, based on the theoretical relations we have derived among these parameters. Evaluation of which data sets are most reliable, however, depends on knowing a unique value for the parameter pair [p, B], and the available data leave room for debate about the best values to choose. We conclude, however, that the most robust results come from a one-dimensional sampling of the fault-displacements in the domain if only because the equation for the cumulative fractional extension, when expressed in terms of the slope s_1 of the displacement data from a one-dimensional sampling, is only slightly affected by the value of p Eq. (7.4).

We test our theory against a detailed set of displacement data sampled in one dimension in the Yucca Mountain area, and we show that the predicted distribution of cumulative strain provides an excellent fit to the observed data.

Based on the Yucca Mt. data, sampling in a two-dimensional domain would overestimate the extensional and shear strains in the volume by 5% (Eq. (5.1)). Sampling in a one-dimensional domain overestimates the extension in the volume by approximately 25% (Eqs. $(5.2)_{1-5}$), but the constraint on the shear strain from one-dimensional sampling is less restrictive; we can infer only that the total shear strain in a volume is greater than half the shear strain determined from sampling in one-dimensional domains along two orthogonal lines (Eqs. $(5.2)_{1-3}$, $(5.2)_{6,\,7}$).

The results of this analysis show that small faults, up to several orders of magnitude smaller than the largest fault in a volume, contribute significantly to the total strain in the volume, because the smaller individual displacements on progressively smaller faults are counterbalanced by the geometric increase in the numbers of these faults.

Acknowledgements

This research was partially supported by grant DE-FG02-03ER15430 from Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy (Marrett). We are most grateful to Graham Yielding for his perceptive review and helpful comments on this paper, which significantly improved the analysis.

References

Cladouhos, T.T., Marrett, R., 1996. Are fault growth and linkage models consistent with power-law distributions of fault lengths? Journal of Structural Geology 18 (2/3), 282–293.

Clark, R.M., Cox, S.J.D., 1996. A modern regression approach to determining fault displacement–length scaling relationships. Journal of Structural Geology 18 (2/3), 147–152.

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

Elliott, D., 1976. The energy balance and deformation mechanisms of thrust sheets. Philosophical Transactions of the Royal Society of London A203, 289–312.

Kim, Y.-S., Sanderson, D.J., 2005. The relationship between displacement and length of faults: a review. Earth-Science Reviews 68, 317–334.

Krantz, R.W., 1988. Multiple fault sets and three-dimensional strain: theory and application. Journal of Structural Geology 10, 225–237.

Marrett, R., Ortega, O.J., Kelsey, C.M., 1999. Extent of power-law scaling for natural fractures in rock. Geology 27 (9), 799–802.

Muraoka, H., Kamata, H., 1983. Displacement distribution along minor fault traces. Journal of Structural Geology 5, 483–495.

Peacock, D.C.P., Sanderson, D.J., 1991. Displacement and segment linkage and relay ramps in normal fault zones. Journal of Structural Geology 13, 721–733.
Scott, R.B., Castellanos, M., 1984. Stratigraphic and structural relations of volcanic rocks in drill holes USW GU-3 and USW G-3, Yucca Mountain, Nye County, Nevada. U.S. Geological Survey, Open-file Report 84-491.
Simonds, F.W., Whitney, J.W., Fox, K.F., Ramelli, A.R., Yount, J.C., Carr, M.D., Menges, C.M., Dickerson, R.P., Scott, R.B., 1995. Map showing fault activity in the Yucca Mountain area, Nye County, Nevada. U.S. Geological Survey Miscellaneous Investigations Series Map I-2520, scale 1:24000.
Twiss, R.J., Marrett, R. Determining brittle extension and shear strain using fault length and displacement systematics: Part I: theory. Journal of Structural Geology, in this issue.
Villemin, T., Angelier, J., Sunwoo, C., 1995. Fractal distribution of fault length and offsets: implications of brittle deformation evaluation – the Lorraine Coal Basin. In: Barton, C., LaPointe, P. (Eds.), Fractals in the Earth Sciences. Plenum Press, New York, pp. 205–226.
Walsh, J.J., Watterson, J., 1987. Distribution of cumulative displacement and of seismic slip on a single normal fault surface. Journal of Structural Geology 9, 1039–1046.
Watterson, J., 1986. Fault dimensions, displacements and growth. Pure & Applied Geophysics 124, 366–373.
Watterson, J., Walsh, J.J., Gillespie, P.A., Easton, S., 1996. Scaling systematics of fault sizes on a large-scale range fault map. Journal of Structural Geology 18 (2/3), 199–214.
Yielding, G., Needham, T., Jones, H., 1996. Journal of Structural Geology 18 (2/3), 135–146.