PERGAMON

Journal of Structural Geology 25 (2003) 819–822

JOURNAL OF STRUCTURAL GEOLOGY

www.elsevier.com/locate/jstrugeo

Discussion

Comment on "The geometric and statistical evolution of normal fault systems: an experimental study of the effects of mechanical layer thickness on scaling laws" by R.V. Ackermann, R.W. Schlische and M.O. Withjack☆

Paul Gillespie

Norsk Hydro Research Centre, Postboks 7190, N-5020 Bergen, Norway

Received 01 October 2001; accepted 5 March 2002

Keywords: Experimental modelling; Fault populations; Statistics; Normal faults; Scaling laws; Spacing

1. Introduction

The article by Ackermann et al. (2001) presents a detailed experimental study of normal faulting and a thorough analysis of the size and spatial distributions of the resulting faults. The conclusions that are drawn have significant consequences for our understanding of the controls on the scaling of fault systems. This discussion is not intended to challenge the main conclusions of the work, but rather to comment on one of the techniques used to measure the spatial distribution of the faults and to suggest a simple but more effective method.

In measuring the spatial distribution of the faults, Ackermann et al. (2001) used two techniques: one technique based on line sampling and another technique based on area sampling in which the positions of the centroids of the fault traces were used. This discussion deals mainly with the line sampling technique, in which Ackermann et al. (2001) took a series of parallel sample lines perpendicular to the faults. The spacings between adjacent faults were then calculated and the mean and standard deviation of the spacings were plotted as a function of the extension of the model (Fig. 1a and b). The standard deviation was calculated as the average for all sample lines and was interpreted as a measure of the regularity of the spacings. The results show a systematic decrease in standard deviation with increase in extension (Fig. 1b). (The increase in standard deviation with extension shown in their summary diagram, Fig. 14A, is presumably an error.) However, the standard deviation suffers the drawback that it is dependent on scale, as smaller spaces tend to be associated with smaller standard deviations. Therefore the decrease in standard deviation with extension is largely the result of the smaller spacing developed as more faults enter the model and does not in itself prove an increased regularity of the spatial distribution.

I illustrate this using a simulation in which random points were successively generated along a line of unit length and the spacings were calculated. This kind of mathematical model is described as a Poisson point process. The results (Fig. 1d to f) show a systematic decrease in the standard deviation with an increased number of points (Fig. 1e), with an overall form very similar to the curve for the experimental data (Fig. 1b). That the decrease in standard deviation occurs in a random simulation makes it clear that such a trend does not require any regularity of the fault spacing.

A more suitable measure of the variability of the spacing is given by the coefficient of variation (Cox and Lewis, 1966; Gillespie et al., 1999, 2001), which is defined as

$$C_v = \frac{r}{s_{avg}} \quad (1)$$

where, following the notation of Ackermann et al. (2001), r is the standard deviation of the spaces between adjacent faults along a line sample and s_{avg} is the mean. The meaning of C_v can be seen with reference to a Poisson point process, which produces a negative exponential distribution of spaces at the limit of an infinite number of points (Cox and Lewis, 1966). The negative exponential distribution has the property that the standard deviation is equal to the mean

☆ PII of original article: S0191-8141(01)00028-1.

PII: S0191-8141(02)00058-5

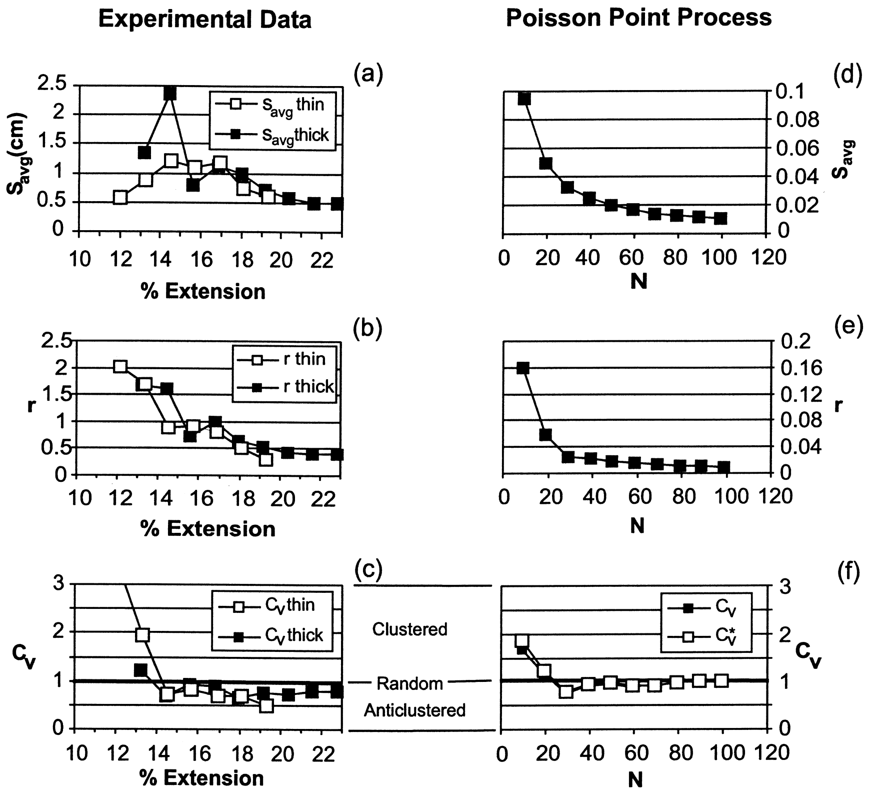

Fig. 1. (a) and (b) Re-digitised results of Ackermann et al. (2001) showing the development of the average spacing, s_{avg} and standard deviation, r, of spacing of the faults along line samples of the thick and the thin models. (c) Calculated coefficient of variation, C_v for the same data. (d)–(f) Results of a single simulation of the Poisson point process as a function of the number of spaces, N. C_v^* is the adjusted coefficient of variation (see text for details).

and so $C_v = 1.0$. Hence, in the simulation in Fig. 1d to f, the mean and the standard deviation decrease together and after a period of instability at low numbers of spaces, the C_v tends to a value of 1.0.

Clustered faults have spacings which are more variable than random, and so $C_v > 1.0$, whereas more regular (or anticlustered) faults have $C_v < 1.0$. Saturated fault systems are therefore characterised by $C_v < 1.0$. A perfectly regular set of faults will have standard deviation of zero and so $C_v = 0$. The C_v is therefore a statistic that can describe the full range of spatial distribution from anticlustering to clustering.

The data from Ackermann et al. (2001) were replotted as a graph of C_v vs. extension (Fig. 1c). After a period of instability at strains of less than 15%, the C_v values settle down to values of less than 1, indicating anticlustering. In the case of the thin model, the C_v decreases to a final value of 0.5. In the case of the thick model the C_v appears to stabilise at a value of 0.8.

2. Sources of error

Before discussing the results further, the sources of error and bias in the measurement of C_v will be discussed.

2.1. Sample size

As explained above the C_v tends to 1 in a Poisson process when there is an infinite number of points. However, when there is a finite number of points, the frequency distribution tends to a beta distribution with

$$s_{avg} = \frac{L}{N} \tag{2}$$

$$r^2 = L^2 \frac{N-1}{N^2(N+1)} \tag{3}$$

(Borgos 1997) where L is the sample line length and N is the number of spaces in the sample. Hence the C_v for a Poisson point process with a finite number of points is given by

$$C_v = \sqrt{\frac{N-1}{N+1}} \tag{4}$$

Therefore, in order that the C_v of a Poisson point process should tend to a value of 1.0, a modified version of the C_v should be calculated:

$$C_v^* = \frac{r}{s_{avg}}\sqrt{\frac{N+1}{N-1}} \tag{5}$$

In practice this causes only a minor change in the estimated C_v and the modification is only required for small samples.

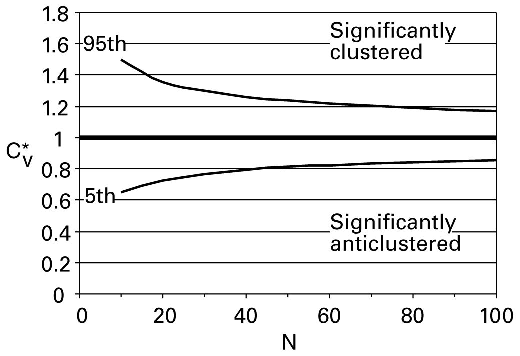

Fig. 2. Confidence intervals for the adjusted coefficient of variation, C_v^*, for different numbers of faults, N, calculated from the 5th and 95th percentiles of a Monte Carlo simulation. Samples falling outside the confidence intervals are clustered or anticlustered at the 95% confidence interval.

2.2. Statistical uncertainty

The question that arises is: are the faults in the models *significantly* anticlustered, i.e. are they significantly more anticlustered than a random sample? Confidence limits are therefore needed. These have been calculated here by multiple simulations of random samples (so-called Monte Carlo simulations).

C_v^* was calculated for 100,000 repeated simulations of the Poisson point process and the 5th and 95th percentiles of the probability density function were calculated. This was repeated for a series of different numbers of spaces, from 10 to 100 (Fig. 2). As the number of spaces increases, the 5th and 95th percentiles converge slowly towards a value of unity.

The percentiles can be used as confidence limits in the following way: if a sample of size N has a C_v that is greater than the 95th percentile for N then we can say that the sample is significantly more clustered than a random sample at the 95% confidence interval. Similarly, if the C_v is less than the value of the 5th percentile then we can say with 95% confidence that it is significantly more anticlustered than a random sample. Values of the calculated confidence limits are also given in Table 1.

As the number of spaces in the samples of Ackermann et al. (2001) is not known it is not possible to determine here whether the calculated C_v values are statistically significant. However, as long as there were more than 40 faults in the samples then $C_v = 0.8$ found in the thick model is significantly anticlustered. In the thin model, the final value of $C_v = 0.5$ demonstrates significant anticlustering even if N is as low as 10.

2.3. Multiple line samples

In the method of Ackermann et al. (2001), multiple parallel sample lines were used through the data rather than a single line. This method has the advantage of increasing the sample number but it can also introduce bias in the sample. If the sample lines are closely spaced, then multiple sample lines can intersect the same pair of neighbouring faults and so the same spacing is counted multiple times. This oversampling will tend to bias C_v. In order to minimise the effect, the sample lines should not be closer than the maximum length of the faults. In the case of Ackermann et al. (2001), the distance between the sample lines was not stated.

Table 1
Confidence limits for the adjusted coefficient of variation. N is the number of spaces in the sample, $c5$ and $c95$ are the 5th and 95th percentiles, corresponding to 95% confidence limits (cf. Fig. 2)

N	$c5$	$c95$
10	0.650	1.496
20	0.728	1.356
30	0.769	1.302
40	0.793	1.262
50	0.812	1.237
60	0.824	1.217
70	0.835	1.201
80	0.844	1.190
90	0.852	1.179
100	0.858	1.170

2.4. Data resolution

In any dataset there will be a spatial resolution below which the spacings cannot be measured. This introduces a bias that tends to increase the regularity of the sample and cause a lowering of C_v. Ackermann et al. (2001) scanned their models at 4000 dpi and so the effect of this bias should be minimal.

3. Discussion

Given that the data of Ackermann et al. (2001) are not badly affected by the sampling biases discussed above and that the sample sizes were adequate, the plots of C_v (Fig. 1c) indicate that anticlustering is developed in the models and that the thin model contains faults that are more anticlustered than the thick model. The development of anticlustering implies a repulsive mechanism such as stress shadowing. It happens that Ackermann et al. (2001) came to the same conclusions despite using an inappropriate technique.

The C_v has already been used in analysis of the spatial distributions of line samples through opening mode fractures, i.e. joints and veins (Gillespie et al., 1999, 2001). It has been shown that when the fractures are confined to mechanical units (stratabound) they are characterised by $C_v < 1$ and so are anticlustered. However, when the fractures are non-stratabound they are typically clustered and $C_v > 1$. This is in concurrence with Ackermann et al. (2001), who describe faults rather than opening mode fractures, but also recognise the importance of the layer thickness in controlling the spatial distribution of fractures.

Ackermann et al. (2001) use the nearest neighbour statistic, Ω, to analyse area samples of the experimental faulting. This clustering index is comparable with the C_v for line samples, although in the case of Ω, 0 is the most clustered value and values greater than 1 are anticlustered. Ackermann et al. (2001) found that at greater than 16% extension, Ω was less than 1 in both the thick and the thin models, indicating either a random spatial distribution or clustering of the faults. A statistical test is needed to determine whether or not the clustering is significant (e.g. Davis, 1973). It is difficult to resolve this result for the area sample with the anticlustering determined for the line samples. Ackermann et al. (2001) recognised that there was anticlustering in the direction perpendicular to the faults and that this can be attributed to stress reduction shadows either side of the faults. Taking this one step further, it may be that the faults are anticlustering in the direction perpendicular to the faults, but clustering along strike as a result of high stress concentrations at the lateral tips of the faults. In order to test for this anisotropic spatial distribution, the nearest neighbour vectors could be plotted on a polar plot.

4. Conclusions

In the analysis of spatial distributions of faults along line samples by Ackermann et al. (2001), an inappropriate statistic was used. The C_v is shown to be a more meaningful statistic, which is able to distinguish between clustered, random, anticlustered and regular fault spacings. As this technique is effective and simple to calculate, it can be used routinely for analysis of fractures and it may help to improve our description and understanding of the processes of fracture saturation and clustering.

Acknowledgements

I thank Brian Farrelly for help with programming the Monte Carlo simulations and Walter Wheeler for careful reading of the manuscript.

References

Ackermann, R.V., Schlische, R.W., Withjack, M.O., 2001. The geometric and statistical evolution of normal fault systems: an experimental study of the effects of mechanical layer thickness on scaling laws. Journal of Structural Geology 23, 1803–1819.

Borgos, H.G., 1997. Partitioning of a line segment. Statistics, 14, Norwegian University of Science and Technology, Trondheim, Norway.

Cox, D.R., Lewis, P.A.W., 1966. The Statistical Analysis of Series of Events, Chapman and Hall, London.

Davis, J.C., 1973. Statistics and Data Analysis in Geology, Wiley, New York.

Gillespie, P.A., Johnston, J.D., Loriga, M.A., McCaffrey, K.L.W, Walsh, J.J., Watterson, J., 1999. Influence of layering on vein systematics in line samples. In: McCaffrey, K.L.W., Lonergan, L., Wilkinson, J.J. (Eds.), Fractures, Fluid Flow and Mineralisation. Geological Society, London, Special Publication 155, pp. 35–56.

Gillespie, P.A., Walsh, J.J., Watterson, J., Bonson, C.G., Manzocchi, T., 2001. Scaling relationships of joints and vein arrays from The Burren, Co. Clare, Ireland. Journal of Structural Geology 23, 183–202.