PERGAMON

Journal of Structural Geology 25 (2003) 1587–1595

JOURNAL OF STRUCTURAL GEOLOGY

www.elsevier.com/locate/jsg

The effect of sample size on geological strain estimation from passively deformed clastic sedimentary rocks

Patrick A. Meere[a,*], Kieran F. Mulchrone[b]

[a]*Department of Geology, National University of Ireland, Cork, Ireland*
[b]*Department of Applied Mathematics, National University of Ireland, Cork, Ireland*

Received 11 May 2002; received in revised form 12 December 2002; accepted 12 December 2002

Abstract

The role of sample size in the estimation of geological strain, both finite strain (R_s) and that of the orientation of the finite strain ellipse (ϕ_s), is investigated for clastic sedimentary rocks. This study looks at four strain methods, the Robin method, the linearization method, the Mulchrone and Meere method and the mean radial length method that are initially tested using simulated strained data sets and subsequently by applying the methods to real data. It is found that the optimum strain analysis sample size for a clastic sedimentary rock is primarily dependant on the intensity of strain suffered by that rock because of the error behavior associated with R_s estimates. An iterative process is therefore recommended starting with a minimum sample size of 150, which can be maintained or reduced based on the initial R_s estimates.

Keywords: Strain analysis; Sample size; Clastic sedimentary rocks

1. Introduction

The estimation of finite strain in rocks is fundamental to a meaningful understanding of deformational processes and products on all scales from microscopic fabric development to regional structural analyses. The last four decades have seen structural geologists concentrate on using strain markers that approximate an ellipsoid, e.g. ooids, accretionary lapilli and sedimentary clasts in sandstones and conglomerates. The importance and effect of strain marker sample size has received relatively little attention with the exception of work by Shimamoto and Ikeda (1976), Sanderson (1977) and Robin and Torrance (1987). Previous work has tended to look at the effect of sample size on the finite strain values (R_s) while ignoring its effect on the estimation of the orientation of the finite strain ellipse (ϕ_s). Marker sample size is arguably one of the most important criterion a geologist will need to determine before commencing a given strain analysis. Issues such as the relationship of sample size to overall strain estimates and to the behavior of confidence brackets need to be addressed. This paper will initially test four strain methods, Robin (Robin, 1977), the linearization method (Yu and Zheng, 1984), Mulchrone and Meere (2001) and Mulchrone et al. (2003) using simulated strained data sets. The methods will then be applied to real data to again look at the effect of marker sample size on strain estimates and associated errors. These four methods are taken as being representative of techniques commonly used for the estimation of geological strain using R_f/ϕ data and will be used to develop a sample size strategy for such studies. Strain analyses of real data to date have very much relied on a number of crucial assumptions including the assumption that markers in a given deformed rock have deformed homogenously with the matrix, i.e. that there is no competency contrast between matrix and marker. While this may not always be the case, it is fair to assume that deformation of lapilli, ooids and for the most part sedimentary clasts can be taken as passive. This study will focus on the behavior of clasts in two deformed sandstones and a conglomerate and will make the assumption that this deformation is essentially passive.

2. Theoretical background

Most analyses of geological strain have evolved from the basic passive strain theory of Ramsay (1967). This R_f/ϕ

* Corresponding author. Tel.: +353-21-902533; fax: +353-21-271565.

0191-8141/03/$ - see front matter
PII: S0191-8141(03)00007-5

method has a number of important premises:

1. passive deformation of the markers with respect to the matrix,
2. definition of sectional strain axes before analysis,
3. strain event being analyzed must be coaxial throughout.

Ramsay's R_f/ϕ method is essentially based on two equations (Ramsay 1967, eqs. [5-22] and [5-27]). There have been several variants of this basic method since then and this study will utilize two relatively recent graphical adaptations of the technique, the linearization method of Yu and Zheng (1984) and an updated version of Lisle's (1977) 'Theta Curve' method (Mulchrone and Meere, 2001).

The linearization method (Yu and Zheng, 1984) arranges the terms of the original R_f/ϕ equations such that they take on the linear form:

$$\cosh 2\epsilon_f = (\tanh 2\epsilon_s)[(\sinh 2\epsilon_f)(\cos 2\phi)] + (\cosh 2\epsilon_i / \cosh 2\epsilon_s) \quad (1)$$

where ϵ_f is $\ln R_f$, the final marker aspect ratio; ϵ_i is $\ln R_i$, the initial marker aspect ratio; ϵ_s is $\ln R_s$, the strain ratio; and ϕ is the final orientation of the marker long axis with respect to a predefined direction.

This results in the typical onionskin distribution of R_f/ϕ curves being transformed into a series of lines representing different R_i values. Data tend to be concentrated towards the ends of these lines giving a final crescent shape distribution. Regression analysis can be used to determine the R_s value, which corresponds to the slope of the 'best fit' line. Mulchrone and Meere's (2001) updated version of Lisle's (1977) 'Theta' method involves destraining deformed markers perpendicular to a predetermined preferred long axis orientation in small increments. The strain value that brings about the most uniform orientation distribution, which is tested using the χ^2 test, is taken as the R_s value for that section.

The Robin (1977) method is not dependant on the markers approximating elliptical shapes and involves drawing a set of axes (a_j, c_j) through the center of a given marker parallel to a set of predefined reference axes. The logarithmic average of a_j/c_j is taken as the R_s values for the section. The Robin method also assumes homogenous coaxial strain.

Mulchrone et al. (2003) derive a method for calculating finite strain from distributions of elliptical objects, assuming that the orientations of the long axes are from a uniform distribution and that the distribution of the axial ratios is axially symmetric. The method is based on the conceptually simple fact that averaging the parameters of an initial ellipse distribution (i.e. the unstrained state) defines a circle, so that in the strained state the averaged parameters define an ellipse, namely the strain ellipse. For a set of n ellipses the following parameters are measured: R_i, the axial ratio and ϕ_i the orientation of the long axis and the following parameters are calculated:

$$m_i = \frac{1}{2}\left(R_i - \frac{1}{R_i}\right) \quad (2)$$

$$p_i = \frac{1}{2}\left(R_i + \frac{1}{R_i}\right) \quad (3)$$

Using this data the following averages are calculated:

$$p_s = \frac{1}{n}\sum_{i=1}^{n} p_i \quad (4)$$

$$q_s = \frac{1}{n}\sum_{i=1}^{n} m_i \cos(2\phi_i) \quad (5)$$

$$t_s = \frac{1}{n}\sum_{i=1}^{n} m_i \sin(2\phi_i) \quad (6)$$

and finally the axial ratio (R_s) and orientation of the long axis of the strain ellipse (ϕ_s) are calculated:

$$R_s = \sqrt{\frac{p_s\cos(2\phi_s) + q_s}{p_s\cos(2\phi_s) - q_s}} \quad (7)$$

$$\tan(2\phi_s) = \frac{t_s}{q_s} \quad (8)$$

Mulchrone et al. (2003) estimate the confidences associated with the strain ellipse parameters by applying the bootstrap method.

3. Method implementation

3.1. Software

A program was written in C++ using Microsoft Visual C++ version 6.0 to implement the methods for strain analysis investigated in this paper. The software allows direct input or importing of natural data for which finite strain is estimated using the methods of Robin (1977), Yu and Zheng (1984), Mulchrone and Meere (2001) and Mulchrone et al. (2003). In each case confidences can be calculated using the bootstrap at the 95% confidence bracket. Additionally the software can be used to generate simulated data sets, which can be mathematically strained, analyzed and the results written to file. Random numbers are generated using the algorithm of Press et al. (1988), p. 282.

3.2. Estimation of confidence in simulation studies

Simulation is a relatively new and important tool in the armory of the scientific investigator and has risen to prominence alongside cheap and powerful computational resources. Most of the strain-analysis techniques investigated in this paper do not have associated information about the sampling distribution derived from basic probability rules (however, Robin and Torrance (1987) do give such a formulation for the method of Robin (1977)) and it may be

extremely difficult to derive an analytical expression for the sampling distribution. Simulation is used here to get information about the sampling distribution of the strain ellipse calculated using the methods discussed above. For a particular data set we can calculate R_s and ϕ_s using any of the above methods and the sampling distribution describes how R_s and ϕ_s vary in value across all possible samples that might be selected from the full ellipse-object population. Without simulation, the full ellipse-object population would need to be measured and numerous samples selected and the variation of R_s and ϕ_s calculated. This is usually impractical. Because the methods of analysis assume knowledge of the initial distribution of elliptical objects, numerous data sets can be generated according to these assumptions and strained so that the variation in R_s and ϕ_s across these datasets can be used to model the sampling distribution and its confidence brackets. In the case of this study 100 data sets were generated for each of 10 data sets ranging in size (n) from 25 to 250. The initial marker aspect ratios (R_i) ranged in value from 1 to 10.

3.3. Estimation of confidence in results from real data sets

Although it may be possible to analytically derive a simple expression for the confidence associated with R_s and ϕ_s for the real data, it is felt that in the absence of analytical expressions for confidences associated with all of the studied methods (noting that Robin and Torrance (1987) do give such an expression for Robin (1977)) the easiest approach for the current problem is to apply the bootstrap method (Efron, 1979). Typically application of the bootstrap method involves generating another set of data by resampling the original dataset with replacements. Suppose there are n data in the original dataset, then the new set is generated by selecting $n$points from the original dataset each time putting the selected point back into the original dataset. Some data points may be repeated in the generated dataset. Another estimate R_s and ϕ_s may be calculated for the generated dataset. Resampling may be performed an arbitrary number of times such that the distributions of R_s and ϕ_s are derived and are referred to as the bootstrap distributions. The bootstrap distributions tend to coincide with the actual (but unknown) sampling distributions.

3.4. Geological material

Three types of clastic sedimentary rocks were selected for this study: a tectonically undeformed Upper Devonian quartz wacke from Dunmore East, west Waterford (Fig. 1a), a tectonically deformed quartz wacke for Ballycrovane, east Cork (Fig. 1b) and a deformed Silurian cobble conglomerate with quartzite and psammite clasts from Croagh Patrick, south Mayo, Ireland (Fig. 1c). This material was selected to represent a wide range of clast size and deformation regime. In the case of the sandstones the R_f/ϕ quartz clast data were collected from photomicrographs while the conglomerate quartzite and psammite clast data were taken from a field photograph.

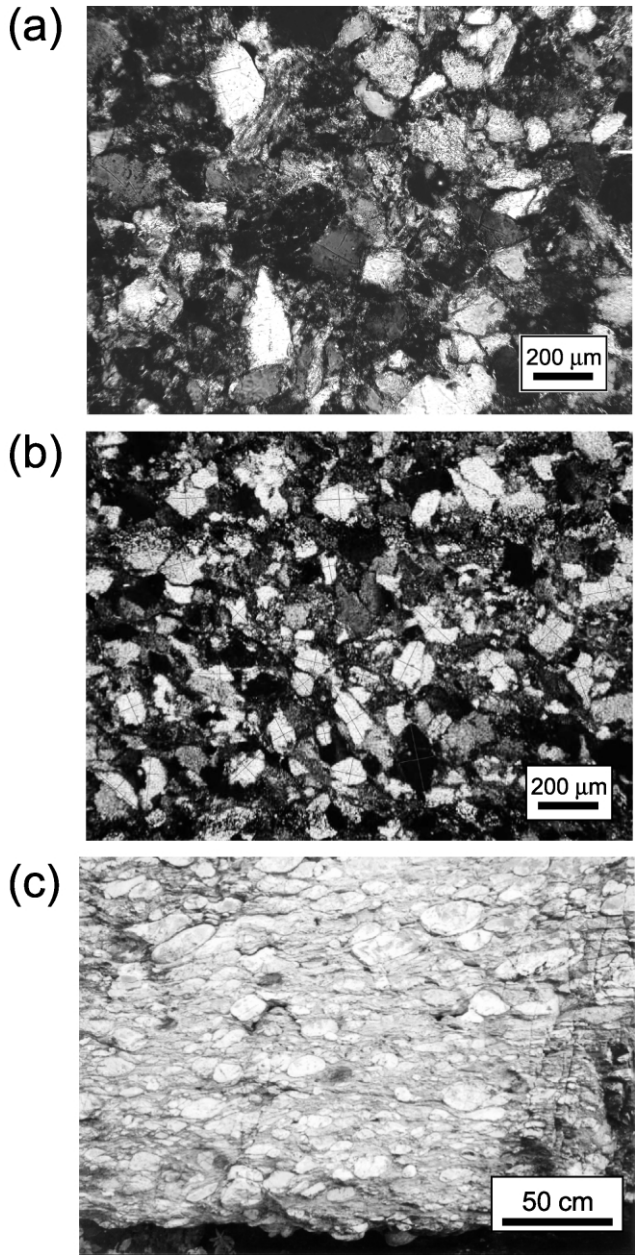

Fig. 1. Clastic sedimentary rocks used in this study, (a) undeformed quartz wacke for Dunmore East, County Waterford, (b) deformed quartz wacke from Ballycrovane, County Cork and (c) deformed boulder conglomerate from Croagh Patrick, County Mayo.

4. Simulation study results

4.1. Finite strain ratio R_s

The results of the estimates of finite strain rations with increasing sample size and applied strain values of $R_s = 1.2$, 3.2 and 4.8 for the four methods are presented in Fig. 2a. All

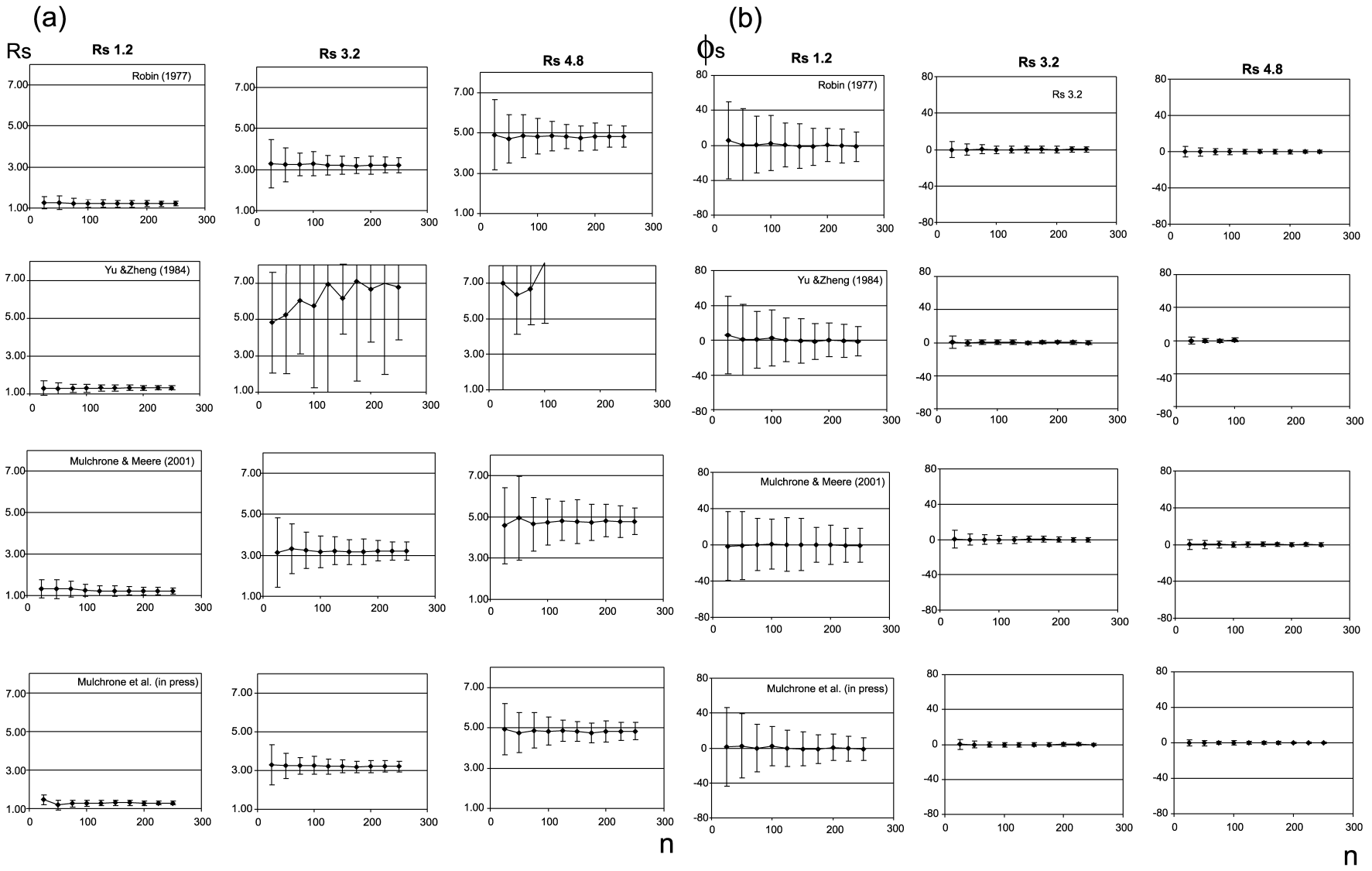

Fig. 2. Strain simulation results for the four methods studied; (a) variation in strain estimates (R_s) and bootstrapped error at the 95% confidence interval with sample number (n), (b) variation in estimates of strain ellipse long axis orientation (ϕ_s) and bootstrapped error at the 95% confidence interval with sample number (n).

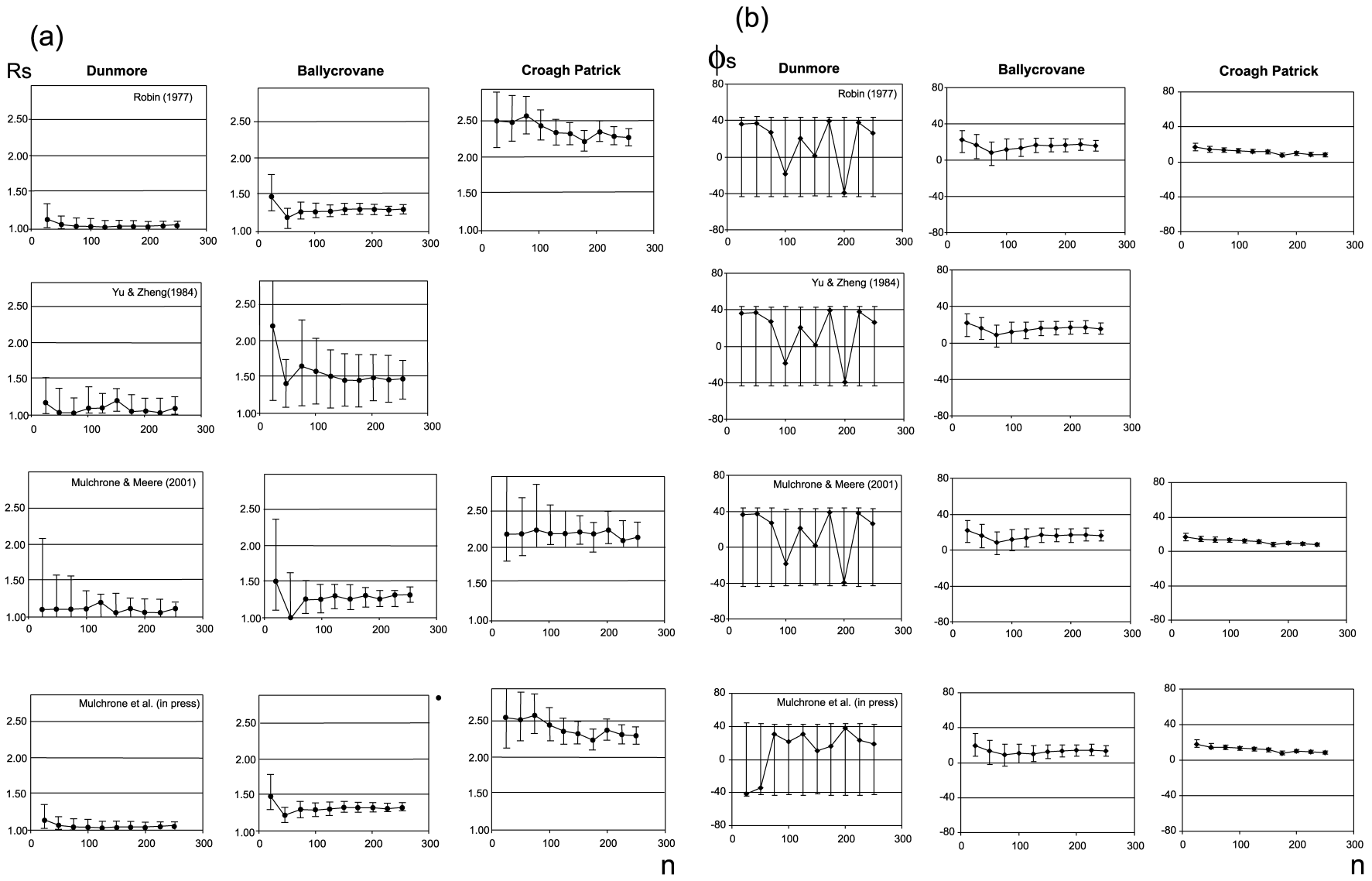

Fig. 3. Strain estimation results from real data for the four methods studied; (a) variation in strain estimates (R_s) and bootstrapped error at the 95% confidence interval with sample number (n), (b) variation in estimates of strain ellipse long axis orientation (ϕ_s) and bootstrapped error at the 95% confidence interval with sample number (n).

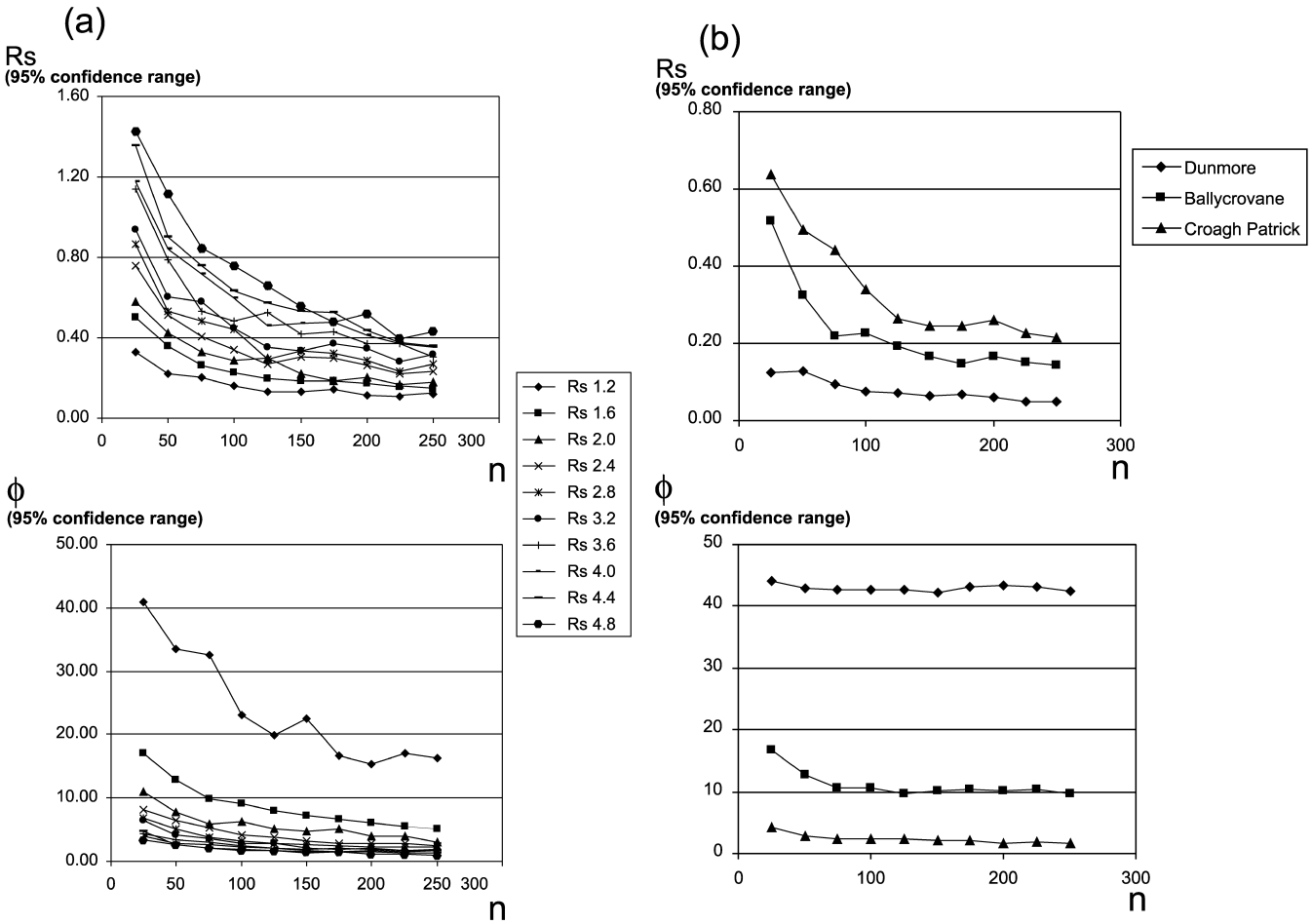

Fig. 4. (a) The simulation study variation in error at the 95% confidence interval with sample number (n) for R_s and ϕ_s. (b) The real data variation in error at the 95% confidence interval with sample number (n) for R_s and ϕ_s.

of the methods show excellent estimates and small 95% confidence brackets at low n values for low R_s values. At higher R_s values the confidence ranges increase with decreasing n values as expected. In the case of the linearization method of Yu and Zheng (1984) no meaningful estimates were obtained for the $R_s = 3.2$ and 4.8 simulations, the method simply breaks down. The Mulchrone et al. (2003) and Robin (1977) methods exhibit the best estimates across the full R_s and n value range with the Mulchrone et al. (2003) method showing the tightest confidence brackets.

4.2. Orientation of finite strain ellipse long axis ϕ_s

The behavior of the estimates of the orientation of the long axis of the finite strain ellipse ϕ_s is the converse of that for R_s (Fig. 2b). Very high confidence brackets are observed at low applied R_s values that rapidly decrease with increasing R_s. This is the case for the four methods studied and simply reflects the rotation of individual marker axes into parallelism with the direction of maximum sectional extension. It is interesting to note that in this, the ideal case, very small sample sizes will yield reasonable ϕ_s estimates for all but very low strain regimes, which effectively removes this variable from consideration when defining an optimum sample size for a given analysis.

5. Real data results

5.1. Finite strain ratio R_s

The behavior of R_s estimates from the real data sets mimics that of the simulation results (Fig. 3a). We see small 95% confidence brackets at low n values for the low strain regime of Dunmore East while in the higher strain regime of Croagh Patrick there is an overall increase in the confidence brackets and an increase with decreasing n values. The linearization method of Yu and Zheng (1984) again breaks down in the higher strain regimes bringing into question the applicability of this technique in all but low applied strain values. The Mulchrone and Meere (2001) method does not yield as tight confidence brackets as the Robin (1977) and Mulchrone et al. (2003) methods while the actual strain estimates are marginally lower. Overall the estimates of

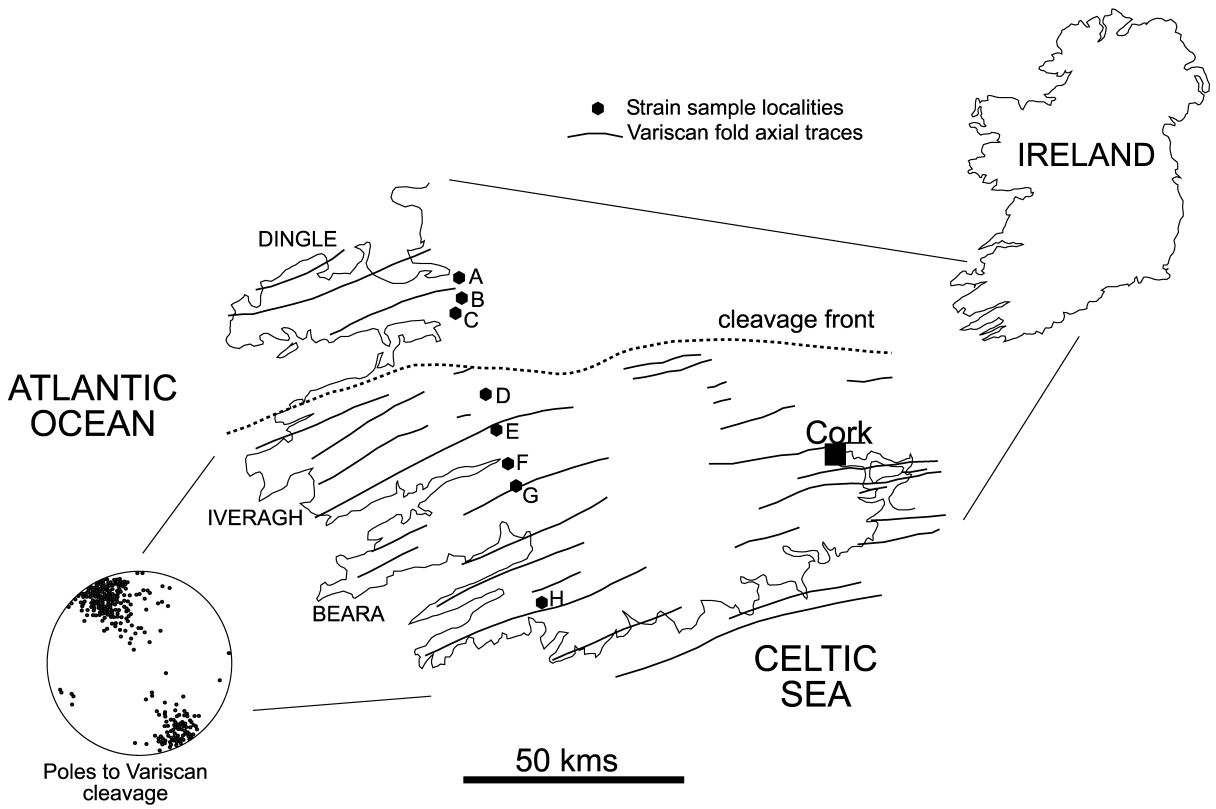

Fig. 5. Strain sample locality map for the Meere (1995) samples reanalyzed in this study. An equal area stereo plot of poles to cleavage south of the cleavage front is also included to show the regional consistency of the fabric orientation.

R_s are not as consistent as those from the simulation studies for low n values, which introduces an important additional criterion in optimum sample size selection. This behavior is to be expected in the non-ideal case. It has to be stated finally that all the R_s estimates and confidence brackets derived from the analyses of real data merely represent an estimate of the shape fabric on a given section and should not be taken as the actual R_s value for that section.

5.2. Orientation of finite strain ellipse long axis ϕ_s

The behavior of ϕ_s with increasing n values for real data again broadly mimics that of the estimates from the simulation studies (Fig. 3b). The sensitivity of estimates to the applied strain (R_s) is evidenced in the real data by the absence of meaningful results from the very low strain regime of Dunmore East. The ϕ_s estimates for the real data is more internally consistent than the equivalent R_s data and stabilizes at relatively low ($n = 100–150$) values.

6. Optimum sample size

The Mulchrone et al. (2003) method exhibits the most consistent R_s and ϕ_s estimates with the smallest confidence brackets and will be the method used to establish an optimum sample size (n) for strain analysis of clastic sedimentary rocks. If the strain estimate R_s and confidence bracket behavior with increasing n values is taken as a criterion to determine an optimum sample size for strain analysis it is clear from all the methods studied, with the clear exception of the linearization method, that a fairly consistent pattern emerges. In the case of the actual R_s estimates, the simulation studies show stable consistent estimates emerging at low ($n < 75$) sample size values (Fig. 4a). This criterion becomes an issue when real data are analyzed and where n values of ca. 125–150 are required for stabilization of the R_s estimates (Fig. 4b). These n values produce a 95% confidence interval of ca. ± 0.3 right across the full strain spectrum studied, which is geologically very acceptable. The overall stable behavior of ϕ_s estimates and confidence intervals, with the notable exception of low strain regimes ($R_s < 1.6$), below values of $n = 100$ negates the need to consider this criterion when selecting an optimum sample size for a given strain analysis (Fig. 4b).

The optimum sample size is therefore dependant of the strain regime being studied which, to a degree, presupposes an estimate of the strain involved. An iterative process is recommended, where initially a minimum n value of 150 is taken and subsequently maintained or reduced based on the

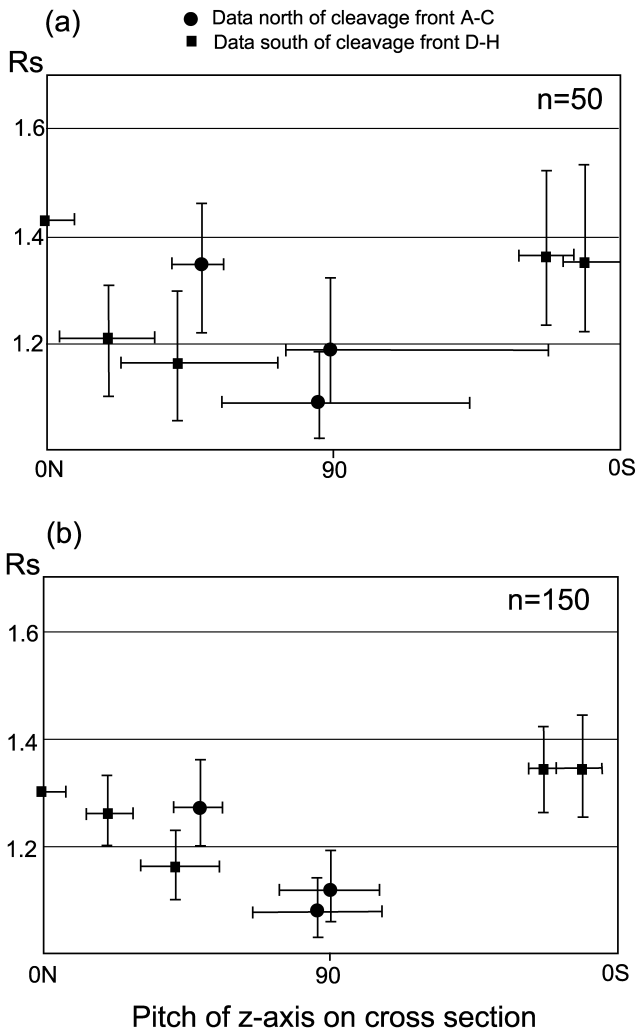

Fig. 6. R_s versus orientation of the strain ellipse z axis in the plane of a cross-section orthogonal to the structural grain of the south western Irish Variscides calculated for (a) a sample size of 50 and (b) a sample size of 150. The bootstrapped error is given at the 95% confidence interval.

initial R_s estimates. This approach will ensure the best results for low to moderate strain regimes for clastic sedimentary rocks.

Robin and Torrance (1987) also looked at the decrease in width of the 95% confidence bracket with increasing sample size of data from deformed anhydrite nodules from the Canadian Artic using the method of Robin (1977) to estimate strain. While their study was limited to R_s estimates and did not make a definitive recommendation on sample size, a similar stabilization of error was noted at $n > 150$ (see Fig. 2, p. 315).

7. Application of sample size criteria to a previous study

The newly established sample size criteria are now applied to a previous regional strain analysis study from the Variscides of southwest Ireland (Meere, 1995). This study analyzed 23 medium-grained sandstone samples across a regionally recognized cleavage deformation front at the western end of the Irish Variscan fold belt. Three mutually perpendicular sections were taken from each sample and a set of reference axes a, b and c defined with respect to an observable fabric plane. This study will reanalyze the data from the ac plane, which is orthogonal to the fabric and parallel to the fabric plane dip, i.e. data in a cross-sectional profile across the fold belt. It will focus on eight samples from the Slieve Mish–Killarney–Ballydehob section (Sections A and B of Meere, 1995). Three of these samples are from north of the deformation front and the remaining five from south of this structure (Fig. 5). While 50 R_f/ϕ measurements were taken from each section in the original analysis, this study will increase the sample size to the now established optimum value of 150. The results are presented with the Bootstrap errors in Fig. 6 as a plot of estimated R_s versus the pitch of the strain short axis in the cross-sectional plane. It is clear that results with $n = 50$ produce significant overlap of error bars from the eight analyses. Increasing n to 150 reduces error bars to a level that now allows a meaningful discrimination between the two strain regimes.

8. Conclusions

1. The optimum strain analysis sample size for a clastic sedimentary rock is dependant on the intensity of strain suffered by that rock, i.e. the higher the strain regime the larger the sample size.
2. This is due to behavior of error associated with R_s estimates making it the primary control on strain analysis sample size.
3. An iterative process is therefore recommended where a minimum sample size of 150 is taken and subsequently maintained or reduced based on the initial R_s estimates.

References

Efron, B., 1979. Bootstrap methods: another look at the Jackknife. Annals of Statistics 7, 1–26.

Lisle, R.J., 1977. Estimation of the tectonic strain ratio from the mean shape of deformed elliptical markers. Geologie en Mijnbouw 56, 140–144.

Meere, P.A., 1995. The structural evolution of the western Irish Variscides: an example of obstacle tectonics? Tectonophysics 246, 97–112.

Mulchrone, K.F., Meere, P.A., 2001. A windows program for the analysis of tectonic strain using deformed elliptical markers. Computer and Geosciences 27 (10), 1253–1257.

Mulchrone, K.F., O'Sullivan, F., Meere, P.A., 2003. Finite strain estimation using the mean radial length of elliptical objects with bootstrap confidence intervals. Journal of Structural Geology 25, 529–539.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1988. Numerical Recipes in C: the Art of Scientific Computing, Cambridge University Press, Cambridge.

Ramsay, J.G., 1967. Folding and Fracturing of Rocks, McGraw-Hill, New York.

Robin, P.F., 1977. Determination of geologic strain using randomly oriented strain markers of any shape. Tectonophysics 42, 7–16.

Robin, P.F., Torrance, J.G., 1987. Statistical analysis of the effect of sample size on palaeostrain calculation. I. Single face measurements. Tectonophysics 138, 311–317.

Sanderson, D.J., 1977. The analysis of finite strain using lines with an initial random orientation. Tectonophysics 43, 199–211.

Shimamoto, T., Ikeda, Y., 1976. A simple algebraic method for strain estimation from deformed eillipsoidal objects. 1. Basic theory. Tectonophysics 36, 315–337.

Yu, H., Zheng, Y., 1984. A statistical analysis applied to the R_f/ϕ method. Tectonophysics 110, 151–155.