

On estimating crystal shape for crystal size distribution analysis

Dan J. Morgan, Dougal A. Jerram *

Department of Earth Sciences, University of Durham, South Road, Durham DH1 3LE, UK

Received 30 December 2004; accepted 30 September 2005

Available online 20 February 2006

Abstract

The accurate determination of three-dimensional crystal habit (long–intermediate–short axis) is a requirement to enable correct stereological conversion of 2D crystal size measurements to true 3D crystal size distributions (CSDs). In this contribution, we introduce a database and spreadsheet program which provides an objective estimate of true crystal habit from raw 2D measurements. The database compares the sample's 2D measurements with 2D shape curves for random sections through 703 different habits. The output gives the five best-match curves and corresponding crystal habits based on a least-squares fit between sample and database. The minimum sample size required to give good shape estimates from 2D data is tested using random runs of sections through known shapes with different population densities. *At minimum*, 75 crystal sections are required to robustly determine crystal habit for CSD measurements if crystals are tabular in shape, with more acicular shapes requiring a minimum of 250 sections, suggesting a sample size of >250 sections to be used for shape determination in natural examples where the true 3D shape is unknown. The crystal habit of the best-match shape can then be used to provide a more robust CSD analysis, which in turn can be used to investigate magmatic processes with more certainty.

© 2006 Elsevier B.V. All rights reserved.

Keywords: CSD; crystal habit; crystal shape; CSDslice; CSDcorrections; Stromboli

1. Introduction

Crystal populations provide a vital window into magmatic processes. Crystal size distributions (CSDs) of volcanic rocks, for example, depend on the cooling and crystallisation history of the magmatic system and provide valuable information about magma residence, mixing events, and magma chamber dynamics (Marsh, 1988, 1998; Jerram et al., 2003). Accurate quantification of crystal populations, however, is a non-trivial process.

In most cases the crystal population, which is a three-dimensional (3D) entity, can only be viewed in two-dimensional (2D) thin sections or on polished slabs due to the indurated nature of the rock. In such cases, the 2D representation of the crystal population must be corrected stereologically to provide a true 3D population. The 3D population can then be used to investigate the magmatic processes which shaped its formation.

Correction from 2D 'apparent' crystal sizes to true 3D crystal populations has been made more straightforward in recent years by the 'CSDcorrections' software (Higgins, 2000) created for this purpose. This program requires the 2D length measurements of the sampled crystals and, given information on crystal habit, crystal roundness and sample foliation, performs a stereological correction of this 2D data array providing a corrected 3D CSD. Thus, an accurate estimate of the 3D

* Corresponding author. Tel.: +44 191 3342281; fax: +44 191 3342301.

E-mail addresses: Daniel.Morgan@dur.ac.uk (D.J. Morgan), D.A.Jerram@durham.ac.uk (D.A. Jerram).

crystal habit is a key element of this correction procedure. This contribution describes a database of discrete crystal shapes and generating software which can be used to give an objective estimate of the 3D habit of crystals from normal 2D short axis/long axis measurements of non-foliated samples, defined by a least-squares fit of an ellipse. The database is contained in an easy-to-use spreadsheet (written for Microsoft® Excel®) which compares raw data with information concerning known sections and outputs the five best-matching crystal shapes for the input crystal population. Guidelines are also given for the minimum sample sizes required to get accurate crystal habits from 2D sections. These shape estimates can then be used in *CSDcorrections* to convert the 2D CSD to a true 3D distribution, providing a simple and objective way towards a more robust CSD measurement.

2. The crystal shape database

Estimating true 3D crystal habit from 2D sections was investigated by Higgins (1994) who showed that certain ‘apparent’ distributions of short axis/long axis measurements can be recognised from random sections through known 3D shapes. Fig. 1 shows random 2D sections through known 3D shapes and statistics on 2D shape (short axis/long axis) examples from this study but directly comparable to Higgins (1994). For the purposes of this paper, 25 bins are used in the 2D shape curves. This is a compromise between having enough points to define the curve well, which requires more bins, and

having few enough bins such that a typical number of data points from a sample (≥ 100) will give a relatively smooth distribution.

3. Method

The sections shown in Fig. 1 are drawn from the output of the program *CSDcut2* which rotates the shape of interest through a random angle around the x , y and z axes sequentially, repeated ten times to remove effects of preferential orientation and the effect of pseudo-randomness in the random number generator routine. The rotated shape is then sectioned perpendicular to the z axis at a randomly determined level between $\pm l$, the length of the body diagonal. This does mean that there is a chance the shape is not actually intersected by the cut plane. If this occurs, the rotated shape is discarded and the whole procedure is repeated (in effect, the thin section has ‘missed’ the crystal). Resulting data for successful intersections are then compared with an ellipse of the same area, centred upon the centroid of area of the intersected shape with the ellipse long and short axes and ellipse orientation adjusted until the root-mean-square (rms) residual mismatch is minimised. The program can operate in one of two modes; in the first mode, only the integrated database of binned long-axis to short-axis ratio is output (50 bins) ready for modification and incorporation into the spreadsheet. In the second mode, all of the calculated data are output, including numbers and co-ordinates of vertices, polygon area, perimeter, best-fit ellipse with major and minor axis

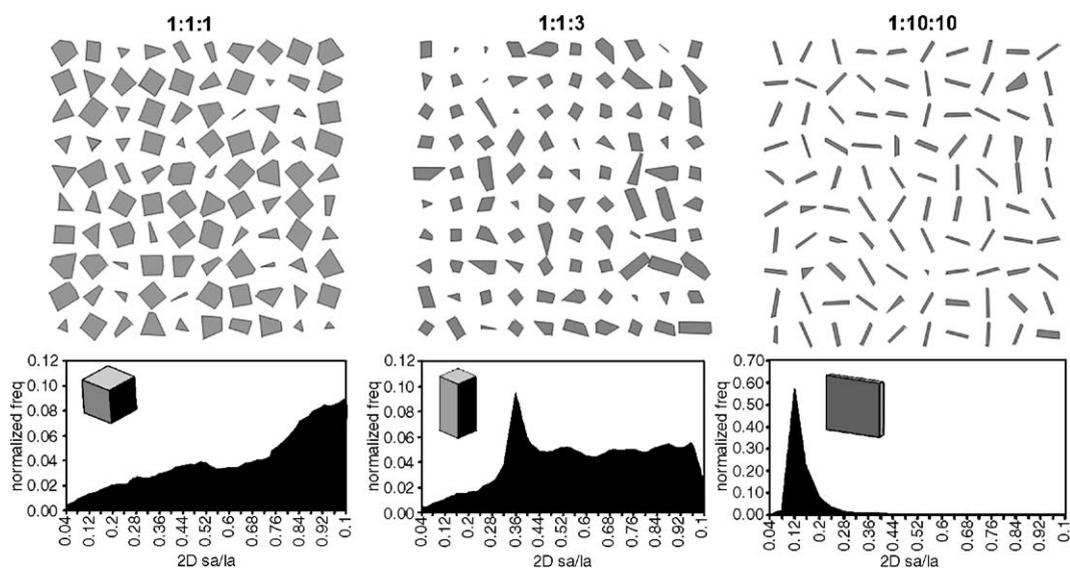


Fig. 1. Random sections through known 3D shapes with population ‘shape curves’ of normalised frequency vs. short axis:long axis ratio used to estimate 3D shape from 2D data (similar to Higgins, 1994).

lengths and ellipse orientation. This latter option is, however, not recommended for large databases or runs involving large numbers (> 10,000) of sections per shape both as the program is slightly slowed and the storage requirements for the resulting files are high to prohibitive. The program does give warnings and a projected estimate of required disk space, however.

In order to provide an objective basis for crystal habit identification, this study uses a database of some 703 discrete crystal habits. If required, the *CSDcut2* program can be used to generate a new database. In this case a batch file needs to be constructed; instructions and an example accompany the program. When the new database has been compiled (which can take some hours; the current database requires over 7 million sectioning and fitting routines to be carried out, for instance), the new database can be pasted over the old one in Excel® or the current database spreadsheet used as a template for a new database.

Each crystal habit has been randomly sectioned 10,000 times to produce the representative 2D shape curves. Greater numbers of sections were not used due to the computation time required and also because 10,000 sections are already far greater than that in a typical measured CSD. The resulting database contains the characteristic 2D curve for each crystal shape in the database, which can be compared with the data from the sample. A spreadsheet called *CSDslice* incorporates the

crystal shape database and through comparison with 703 crystal habits finds the closest-matching shape curve to the 2D shape data based on a linear regression comparing the frequency of occurrence in the sample shape curve and in the database shape curve data. The spreadsheet outputs the five best-matching crystal habits with a graphical representation of the raw data compared to the best-matching characteristic shape curve (Fig. 2).

4. Implementation

Using such methods it is possible to estimate the 3D crystal habit from 2D data given enough sections and assuming the 3D habit of all the crystals in the 2D section is the same. The various peaks on the short axis/long axis plot relating to the long–short and long–intermediate divisions allow the relative habit to be known, where the shortest dimension is one. In real data, however, the shape curves are not always clear-cut which can lead to more subjective estimates of the crystal habit and less robust CSD calculations. The effect of changes in the short-axis to intermediate-axis ratio and the intermediate-axis to long-axis ratio on the calculated volume percentage of that crystal for a natural dataset (best-matching habit 1:2.8:4) is investigated using *CSDcorrections* and the results are shown in Fig. 3. Fig. 3a shows that an increase in the intermediate axis whilst keeping the same ratio between long and

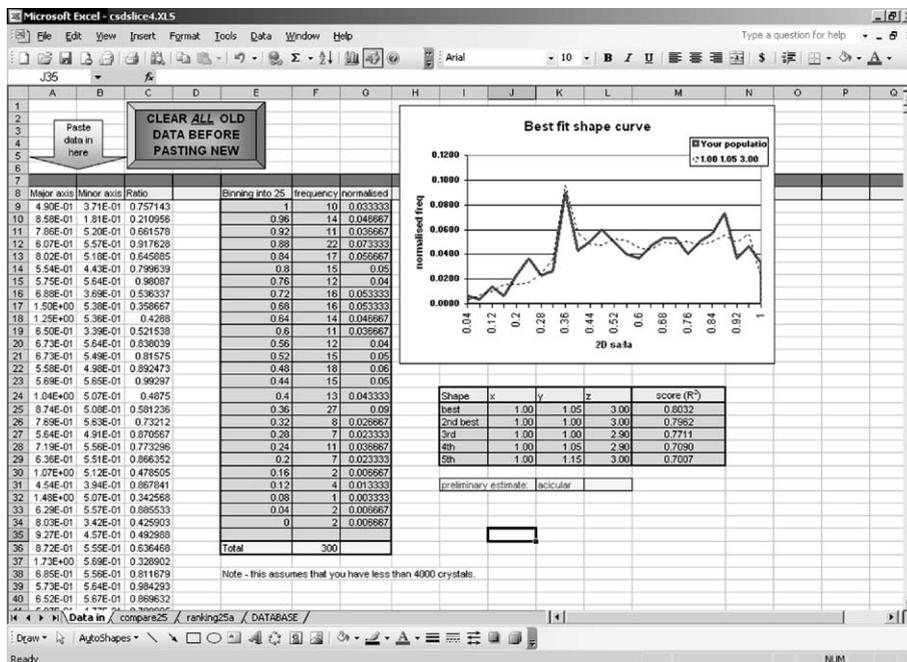


Fig. 2. Screen shot of *CSDslice* database (version 4).

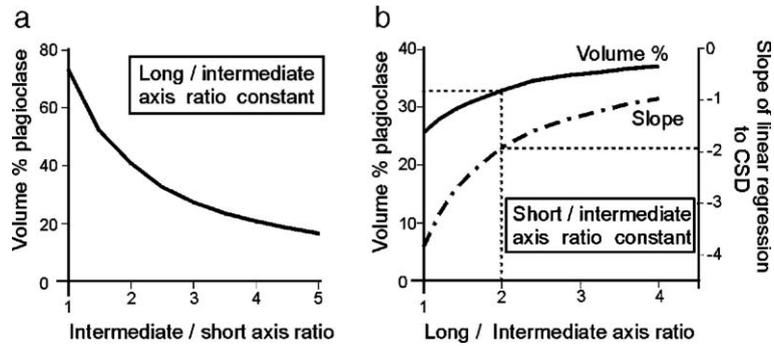


Fig. 3. Graphs showing effect of changes in (a) short axis : intermediate axis and (b) intermediate axis : long axis on the calculated volume percentage of crystals and the regression of the CSD slope.

intermediate axis (effectively a decrease in the short axis length) has a strong effect of decreasing the calculated crystal fraction. In addition, there is a small effect on the slope of the linear regression to the CSD, although this effect operates at the sub-percent level. In Fig. 3b, it can be seen that increasing the length of the crystal whilst holding the ratio of the short and intermediate axes

constant (i.e. increasing acicularity) increases the calculated crystal volume fraction. In addition, there is a strong effect of shallowing the slope of the linear regression to the CSD as crystal sizes are proportionately increased. Therefore, the CSD is sensitive to the relative lengths of all three axes of the crystal shape and shape may therefore have effects on parameters such as

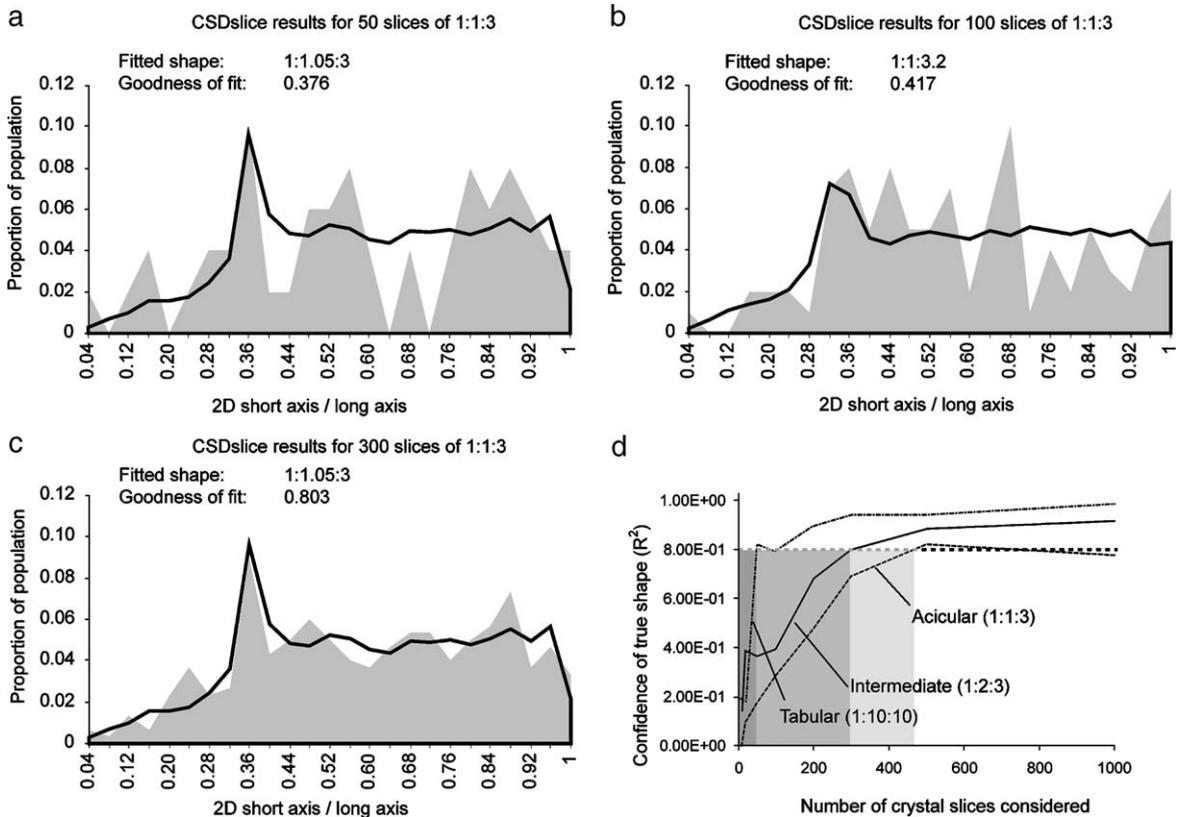


Fig. 4. Testing sample size and shape fit data with the shape 1 : 1 : 3. a) 50 sections, b) 100 sections c) 300 sections. d) Envelope showing the improved level of fit with sample size bounded by acicular and tabular shapes. Note that tabular shapes home in quickly to high scores, but acicular shapes are more widely variable due to the statistics of sectioning behaviour. Sample sets of 200–250 crystals routinely give very close agreements in shape with requisitely high scores ($R^2 > 0.8$).

crystal residence times calculated from the slope of the linear regression to the CSD.

5. Estimating the 3D crystal habit from 2D long and short axis measurements

Minimum sample size is an important question when calculating CSDs for natural populations of crystals from 2D thin sections. Mock and Jerram (2005) provided a test of this using a 3D reconstruction of a true CSD from a porphyritic rock. They used this to test the effect of sample size on the CSD and found that a sample size greater than 200 faithfully reproduced the CSD. In the case of shape estimates from 2D data it is

also important to know the minimum sample size that will provide robust 3D shape estimates from 2D section data. Here, we test this by comparing random sections through known shapes with different population densities, and use the database to estimate the shape.

As a test of the sample size required to estimate crystal habit accurately, different, random selections of 2D measurements were taken from the crystal section data produced along with the dataset for each of the crystal habits: 1:1:3, 1:2:3, and 1:10:10, examples of acicular, blocky and tabular crystals, respectively. The reliability of a shape is quoted as R^2 , the fractional measure of the variation in the sample explained by the best-fit shape of the database.

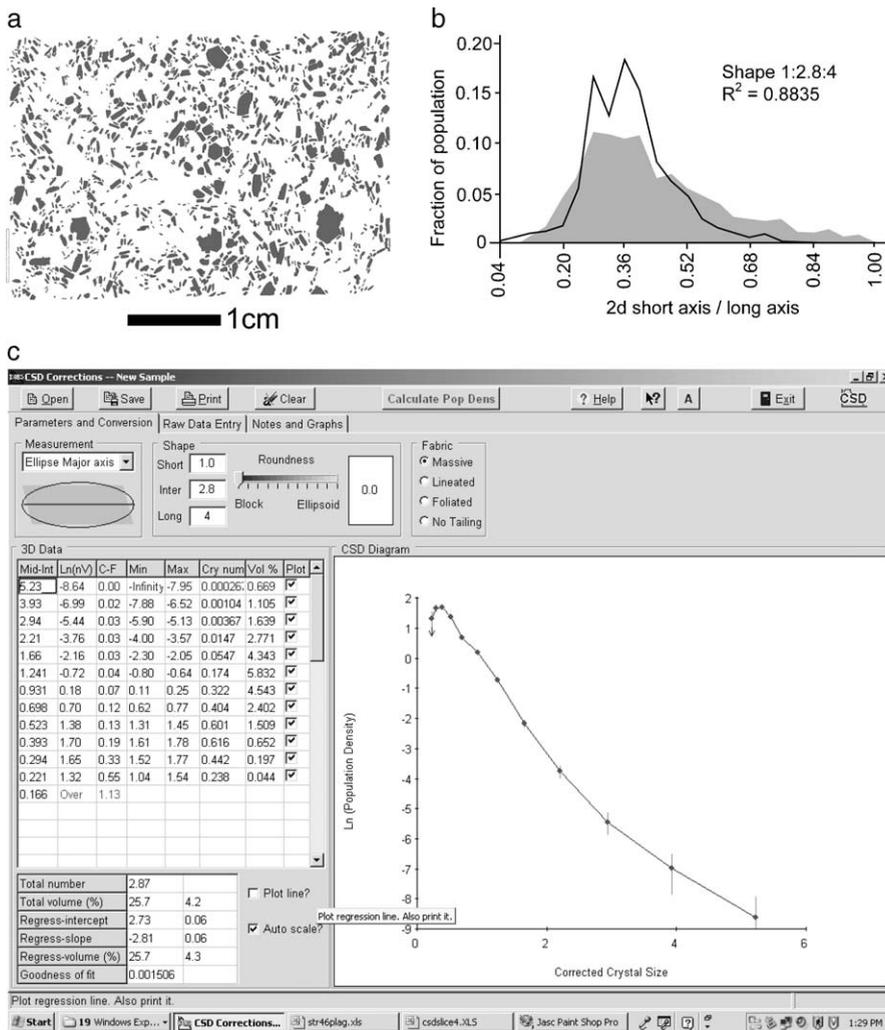


Fig. 5. Example of CSDslice shape estimate for natural example STR46 from Stromboli volcano, Italy; a) raw data for slide STR46 containing 1465 crystal outlines; b) short axis:long axis ratios as binned and placed into *CSDslice*, together with the best-fit shape output by *CSDslice* of 1:2.8:4, $R^2=0.8835$; c) Final CSD pattern as a screenshot from *CSDcorrections* v. 1.35 (Higgins, 2000 et seq.), showing kinked CSD derived from mixed populations, but with a reasonable crystal proportion (25.7%) compared to that observed in the thin section.

In practice, it is observed that R^2 values of over 0.8 give generally close agreement with the actual crystal shape, although care has to be taken when dealing with sample sets which involve less than about 200 crystal sections, as there are 25 bins in the fitting process, and at this level the statistics make the fitting process very susceptible to blips at high aspect ratios. Natural samples tend to display broader distributions than the idealised shapes of the database, being generally composed of either a family of similar crystal habits or even two or more very different crystal habits (Mock and Jerram, 2005).

Sequentially, populations of 10, 20, 50, 100, 200, 300, 500, and 1000 sections were fed into the *CSDslice* spreadsheet for each of the three analysed crystal habits 1:1:3, 1:2:3, and 1:10:10, and R^2 determined. Results for 50, 100 and 300 slices for the crystal habit 1:1:3 are shown in Fig. 4a–c, together with a plot of how quickly R^2 increases with the population size over the three considered shapes (Fig. 4d). Note that for acicular shapes, such as the crystal habit 1:1:3, R^2 increases more slowly. This is due to the low probability of sectioning the full length of an acicular shape and obtaining the short-long axis section. Conversely, the probability of sectioning tabular crystals in orientations that will give true short/intermediate and short/long axis sections is much higher, and so they produce a high R^2 for small populations. This is clear from both the graph in Fig. 4 and also the shapes of the curves of the database populations, which are focused for tabular populations and broadly scattered for acicular populations (see Fig. 1). This also leads to a higher level of variability in the case of the acicular shapes due to low probabilities of extreme sections, and therefore the need to look at larger numbers of sections. The curve for cubic populations is very similar to that for acicular populations due to the large proportion of triangular cuts, which do not carry as much information useful for determining shape as they sample only small portions of a shape's volume. As acicular aspect ratios become increasingly extreme, it becomes statistically harder to cut the longest axis. Even the database itself, compiled from 10,000 sections, shows some difference between different runs of the most extreme shape considered, 1:1:10.

As a test example, analysis is performed on a digitised image of the plagioclase phenocrysts in a porphyritic lava from Stromboli, Italy (Fig. 5a). The short axis/long axis data from *CSDslice* suggests a crystal habit of 1:2.8:4, with R^2 of 0.8835 (Fig. 5b). The resultant CSD plot is shown in Fig. 5c using *CSDcorrections* version 1.35 to calculate the 3D CSD from the 2D data. Clearly, the fit for natural data is less accurate than has been tested in this study using exact shapes. This highlights the fact

that natural populations often exhibit non-unique shapes; however, using an average estimate of shape markedly improves the calculated CSD (Mock and Jerram, 2005).

Further tests using sample sets of representative size (400 sections) but derived from mixtures of ideal shapes from the database were undertaken to examine the effect of mixed populations on the *CSDslice* spreadsheet. The sheet was able to accurately determine the dominant shape present in a mixed population (50%+ of the population) for both acicular and lath shapes.

6. Closing remarks

A user friendly database 'CSDslice' is introduced which allows the user to obtain objective estimates of crystal habit (long–intermediate–short axis) to be used when calculating 3D CSDs from 2D section data (e.g. for use with *CSDcorrections* (Higgins, 2000)). The dataset comprises 703 shapes with crystal habits in the range 1:10:10–1:1:10–1:1:1. When raw 2D data are placed into *CSDslice* the database returns the best 5 shape fits determined by linear regression automatically, with R^2 values ranging from 0 to a maximum of 1, indicating perfect fit. An R^2 value of >0.8 for over 200 considered crystal sections is recommended to give a good statistical fit. Acicular and cubic shapes have R^2 values of 0.4–0.6 when 200 2D crystal sections are considered, with close determinations of the actual shapes attained only by 300 considered crystal sections, suggesting accurate determination between 200–300 (e.g. at ~250 sections). Tabular shapes require less crystal sections to be considered as the longest dimension of the crystal is more frequently sectioned. This value is in close agreement with the findings of Mock and Jerram (2005) which suggested a sample size greater than ~200 to accurately measure a 3D CSD from 2D data.

Acknowledgements

The authors would like to thank William Kinman and Taber Hersum for constructive reviews of this manuscript and Alex Mock for kindly reviewing an earlier version. Darren Chertkoff is thanked for data concerning sample STR46. Comments from early users including James Day and Victoria Martin also improved the database. This work is supported by the EU 5th framework ERUPT program, grant code EVG1-CT2002-00058.

References

- Higgins, M., 1994. Determination of crystal morphology and size from bulk measurements on thin sections: numerical modelling. *American Mineralogist* 79, 113–119.

- Higgins, M.D., 2000. Measurement of crystal size distributions. *American Mineralogist* 85 (9), 1105–1116.
- Jerram, D.A., Cheadle, M.C., Philpotts, A.R., 2003. Quantifying the building blocks of igneous rocks: are clustered crystal frameworks the foundation? *Journal of Petrology* 44, 2033–2051.
- Marsh, B.D., 1988. Crystal size distribution (CSD) in rocks and the kinetics and dynamics of crystallization; I. Theory. *Contributions to Mineralogy and Petrology* 99, 277–291.
- Marsh, B.D., 1998. On the interpretation of crystal size distributions in magmatic systems. *Journal of Petrology* 39, 553–599.
- Mock, A., Jerram, D.A., 2005. Crystal size distributions (CSD) in three dimensions: insights from the 3D reconstruction of a highly porphyritic rhyolite. *Journal of Petrology*.