This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

A review of statistical tools for morphometric analysis of juvenile pyroclasts

Tobias Dürig[1], Pierre-Simon Ross[2], Pierfrancesco Dellino[3], James D. L. White[4], Daniela Mele[3], Pier Paolo Comida[2]

1. Institute of Earth Sciences, University of Iceland, Sturlugata 7, 101 Reykjavík, Iceland tobi@hi.is

2. Institut national de la recherche scientifique, 490 rue de la Couronne, Québec (Qc), G1K 9A9, Canada, rossps@ete.inrs.ca, pier_paolo.comida@ete.inrs.ca

3. Dipartimento di Scienze della Terra e Geoambientali, University of Bari, Via Edoardo Orabona, 4, 70125 Bari, Italy, pierfrancesco.dellino@uniba.it, daniela.mele@uniba.it

4. Department of Geology, University of Otago, 360 Leith Street, Dunedin 9016, New Zealand, james.white@otago.ac.nz

* Corresponding author, tobi@hi.is

Abstract

Morphometric analyses are based on multiparametric datasets that describe quantitatively the shapes of objects. The stochastic nature of fracture-formation processes that break up magma during explosive eruptions yields mixtures of particles that have highly varied shapes. In volcanology morphometric analysis is applied to these mixtures of particles with diverse shapes for two purposes: (1) to fingerprint tephra from individual eruptions and use the fingerprints to distinguish among tephra layers and determine their extents, and (2) to reconstruct eruption processes, by linking particles formed by known fragmentation processes in experiments with particles from natural pyroclastic deposits. Here we review the most commonly adopted statistical techniques for morphometric analysis of pyroclasts. We provide sets of objects with different shapes, along with their morphometric data, in order to demonstrate and illustrate the methods. They can be used not only for addressing the processes of fragmentation during explosive eruptions, but also for the characterization of other types of solid particles with complex morphologies.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Introduction

The use of shape descriptors in the analysis of juvenile pyroclasts offers many options for quantitative analyses and interpretation of magma-fragmentation processes. The shapes of juvenile pyroclasts are highly varied, reflecting varied pre-fragmentation magma textures and the stochastic nature of magma fragmentation and fracture processes (Lawn 1993; Dürig and Zimanowski 2012; Dürig et al. 2012b; Taddeucci et al. 2021). Data analysis and interpretation thus requires the use of sophisticated statistical techniques. This article offers a review of such methods, as developed in the last 25 years, and intends to provide a morphometric "toolbox" that includes the most commonly used analytical and statistical techniques. To help readers fully exploit their morphometric data, we describe how to apply morphometry, and discuss the mathematical pre-conditions and caveats associated with the statistical techniques presented. Along with the recommendations on data acquisition provided by Ross et al. (2021) and Comida et al. (2021), this paper intends to serve the volcanological community as a basis for a discussion on standardized protocols for the analysis of juvenile pyroclasts. The techniques presented here are, however, not limited to use with pyroclasts – they can be applied to any particles, and indeed to any set of morphometric data.

The morphological (and possibly also textural) information for one grain is specified by a set of M variables (i.e., shape parameters), which might or might not be statistically independent from one another. Two-dimensional shape parameters can range from descriptors of basic geometrical characteristics (e.g., Dellino and La Volpe 1996; Cioni et al. 2014; Leibrandt and Le Pennec 2015; Liu et al. 2015) to more complex parameters that are the result of fractal analysis (Dellino and Liotino 2002; Maria and Carey 2002) and curvature plots (Tunwal et al. 2020) or extracted from Fourier shape analysis (Barrett 1980; Suzuki et al. 2015; Chávez et al. 2020). In recent years, novel scanning techniques have allowed the retrieval of 3-dimensional shape parameters, such as, e.g., fractal dimension (Rausch et al. 2015; Vonlanthen et al. 2015; Dioguardi et al. 2017), aspect ratio of the best fit ellipsoid (Vonlanthen et al. 2015) or 3D-sphericity (Mele et al. 2011; Vonlanthen et al. 2015; Dioguardi et al. 2017). An N x M matrix of M variables describing N particles is defined as a "morphometric data set" (Dürig et al. 2020c). An example of a morphometric data set is a table with shape parameters for randomly collected tephra grains at a certain location, or for juvenile particles extracted from a specific size fraction of a certain pyroclastic bulk sample. To compare two morphometric data sets, each set should ideally be of equal size, allowing comparison of N x M values. A typical analysis for 50 grains per sample or size fraction (e.g., Dürig et al. 2018; Comida et al. 2021; Ross et al. 2021) and four shape parameters per grain (Dellino and La Volpe 1996; Dürig et al. 2012a) requires cross-comparisons among 200 individual values in the data set. This is what multivariate statistical methods are designed to accomplish.

The typical goals of morphometric data analysis are:

- to provide a quantitative summary of particle shape descriptors that complements other data (such as grain size, componentry or stratigraphic information);
- to compare two or more data sets with one another, in order to investigate whether they are statistically equivalent or instead show significant differences;
- to determine the fragmentation mechanism(s) that generated the pyroclasts and help reconstruct eruption processes.

The data-analysis representations and the statistical techniques to be used for reaching the three aforementioned goals are described in the following sections. They are intended as step-by-step guides and recommendations for the statistical tests and techniques to be used. To demonstrate them we provide a set of artificial 2D silhouettes (see Fig. 1 and Fig. 2). Supplementary data (Online Resource 1) includes all the binary images and their morphometric descriptors. Furthermore, we use silhouettes of 88-63 µm (narrow +4ϕ) sized ash particles sampled from the 1959 Kīlauea Iki eruption (Hawaii), and the 2012 Havre eruption (Kermadec arc) to illustrate the use of discriminatory diagrams. These silhouettes and the obtained shape parameters can be retrieved from Online Resource 2.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

We note that all statistical analysis described below can be applied to any shape parameters. For demonstration we use the parameters suggested by Dellino and La Volpe (1996), consisting of circularity *Circ_DL*, elongation *Elo_DL*, rectangularity *Rec_DL* and compactness *Com_DL*, defined by:

$$Circ\_DL = \frac{2\sqrt{A\pi}}{p} \quad (1)$$

where *A* is the projected area of the particle's silhouette, and *p* its perimeter.

$$Elo\_DL = \frac{a}{m} \quad (2)$$

with *a* being the longest segment inside the particle parallel to the long side of the minimum area bounding rectangle, and *m* being the mean intercept perpendicular to *a*.

$$Rec\_DL = \frac{p}{2b+2w} \quad (3)$$

where *b* and *w* are the long and short side of the minimum area bounding rectangle, respectively. The compactness is defined by:

$$Com\_DL = \frac{A}{b \cdot w} \quad (4)$$

Descriptive morphometry

Summarising statistical reports

The underlying data for morphometric analyses consist of morphometric data sets and images of particles or particle silhouettes. It is best practice to append these raw data to a publication (Ross et al. 2021), or to lodge them in an open-access data archive, linked from the publication. For each data set, at least the *sample size* (*N*) and both *mean* and *standard deviation* for each of the measured parameters should also be presented in a data table. In addition, providing the minimum and maximum value, or the median (50% percentile) can support additional interpretations. For example, a median that differs considerably from the

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

mean indicates the presence of outliers that might deserve further exploration. Since some of the multivariate statistical tests require that samples are normally distributed, it is also useful to calculate the kurtosis and skewness for each parameter (indicating the distribution's "tailedness" and its asymmetry, compared to its mean). A "perfectly" normal distribution is characterized by a value of 0 for both kurtosis and skewness (Davis 2002).

Table 1 shows how such basic statistics can be presented.

Binary diagrams

The distribution of each parameter can be presented in frequency plots (i.e., histograms, see Fig. 3a). These can be arranged as matrices, sorted by parameter (columns) and stratigraphic sampling location (rows) (see Dellino and La Volpe 1996; Coltelli et al. 2008). A more compact presentation of data uses range plots, which display the total span of the respective parameters as horizontal bars along a stratigraphic axis. Outliers may be overemphasized in range plots, and this can be overcome by using "boxplots", which indicate the distribution's quartiles (Fig. 3b). The morphometric range plots or boxplots can be arranged together with other parameters of interest, such as grain size or chemical composition (see e.g., Verolino et al. 2019).

A quick way to visualize data sets is to prepare binary diagrams, using one parameter per axis (e.g., Fig. 3c and 3d). For *M* parameters, this approach results in $\sum_{i=1}^{M-1} i$ diagrams. For example, if using four shape parameters, six unique pairs need to be prepared. Depending on one's objective this may work well initially, with a relatively small database. Some of the examples listed in Table 2 are binary diagrams that have been used in morphological studies for interpretation of underlying ash generation processes. We anticipate, however, that binary plots will become confusing when data from different volcanoes and different eruptive styles are brought together. More sophisticated methods of data analysis and presentation are thus needed.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Comparison of morphometric data sets

Morphometric data analysis is useful for comparing tephra from different eruptions or eruptive phases (e.g., Dellino and La Volpe 1996; Taddeucci et al. 2002; Cioni et al. 2008; Iverson et al. 2014; Verolino et al. 2019) and for linking characteristics of particles from experiments with those coming directly from a pyroclastic deposit (e.g., Büttner et al. 2002; Dürig et al. 2012a, 2020b; Schipper et al. 2013; Jordan et al. 2014).

Testing for equality of variances

Before comparing two or more data sets with one another one must test their levels of variance, because the best approach to subsequent analysis depends on whether the variances of the data sets are equal within an acceptable range. Thus, when planning a comparative morphometric analysis, the first step is to test the equality of variances (Mele et al. 2011). A common test tailored for such a task is the F-test, named after the Fisher-Snedecor probability or “F” -distribution. The F-test evaluates two data sets against the null hypothesis H_0 that their variances are equal, by comparing the ratio of their variances (“F-scores”) with a critical threshold that is specified by the selected level of significance α (Davis 2002). Typically, 5% is selected for α. Using the F-distribution and a lookup table, the F-score is translated to a “*p*-value”, which gives the error likelihood of incorrectly rejecting the null hypothesis. If the *p*-value is smaller than α, H_0 can be rejected. In this case the F-test has shown that the variances of the two tested data sets are heterogeneous.

The “Levene test” expands the F-statistic to allow also the comparison of variances of more than two data sets (Levene 1960). Other than that, this type of test is equivalent to the F-test and serves the same purpose (Dürig et al. 2012a).

The Levene test was explicitly designed to be robust against violation of normality, whereas F-tests assume data sets with normal distributions. In practice, however, F-tests have also been shown to be very robust even when used with non-normally distributed data (Donaldson 1968).

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

For reporting the outcome of F-tests or Levene-tests (Fig. 4a and Table 3) we recommend that the analyst provides the *p*-values along with α. Note that choosing any level of significance, α, other than 5% requires explicit justification.

Two-tailed t-tests

A common task in morphometry is to verify that differences in the means of two data sets are statistically significant, given the sizes of the data sets and their standard deviations. A two-tailed t-test is commonly used in such cases (e.g., Dellino et al. 2001; Büttner et al. 2002; Mele et al. 2011; Dürig et al. 2012a, 2020b, a; Schipper et al. 2013; Jordan et al. 2014; Schmith et al. 2017). The test begins with the null hypothesis H_0 stating that both groups (the two pyroclastic samples being compared) are extracted from the same population.

Two different types of t-tests exist, depending on the data sets' homogeneity of variances. When the variances of the two data sets can be inferred to be equal, a pooled variance "Student's t-test" (Student 1908; Davis 2002) is used. With $\bar{X}_1$, $\bar{X}_2$ being the means, $s_1 = s_2$ the standard deviations, and N_1, N_2 being the sample sizes of the two data sets, the t-value is computed according to:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p}{N_1} - \frac{s_p}{N_2}}} \quad (5)$$

using the pooled standard deviation s_p, defined as:

$$\sigma_p{}^2 = \frac{(N_1 - 1)\cdot s_1{}^2 + (N_2 - 1)\cdot s_2{}^2}{N_1 + N_2 - 2} \quad (6)$$

If, instead, the variances of the two data sets are different (e.g., as indicated by a Levene test), it is better to apply a separate variance or "Welch's t-test" (Welch 1947), in which the t-value is computed by:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1}{N_1} - \frac{s_2}{N_2}}} \quad (7)$$

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

In both cases the t-values follow the Student's t-distribution curve, with which they can be translated into a "*p*-value". This parameter expresses the probability that test results under the assumption of H_0 are at least as extreme as the observed outcomes. A very low p-value therefore represents a very low probability that H_0 is true. If p is below a pre-defined level of significance α, the null hypothesis can be rejected: the data sets are therefore "significantly different" under the tested hypothesis (Davis 2002). Usually, a level of significance of 5% is used, although sometimes 10% has been chosen for morphometric studies (Büttner et al. 2002). We recommend that researchers report the results of each t-test by providing the type of t-test conducted (or, alternatively the *p*-value of the previously conducted test for equality of variance), the *p*-value, the sample sizes N_1, N_2 and the selected level of significance (see Table 3 and Fig. 4b).

We note that t-tests are parametric; they assume random sampling and that the tested data sets are normally distributed. These conditions may not be met, but t-tests are popular because they show a certain degree of robustness against violations of the assumption of normality. For example when using t-tests, Type I errors (i.e. indicating a significant difference, when in reality there is none) are relatively low, when:

- comparing data sets with samples from two different shape of distributions and unequal sizes, but equal variances (Havlicek and Peterson 1974);
- comparing data sets with samples from non-normal distributions and unequal variances, but comparable sizes (Ahad and Yahaya 2014).

Type I errors are, however, significantly increased when multiple types of inhomogeneities coincide, e.g. unequal variances, non-normal distributions and unequal sample sizes (Ahad and Yahaya 2014).

To avoid these Type I errors, analysts should apply t-tests to data sets with sample sizes N that are not too different from one another. Testing the normality of the data sets with Shapiro-Wilk or Kruskal-Wallis tests (Davis 2002) and listing their possible kurtosis could further help demonstrate the validity of t-test results, but it is not strictly mandatory.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

When the condition of random sample selection is not fulfilled (which is common in geological investigations), the reliability of a t-test is considerably reduced, especially if the same data set is repeatedly used for different comparisons (Bender and Lange 2001). As an example, think of three data sets (“A”, “B”, “C") which should be compared with each other. After having tested “A” with “B”, a subsequent t-test comparing “A” with “C” would use the sample “A” for a second time and therefore violate the “random selection” condition. This leads to an increase in the likelihood of a Type I error. To counter this effect, a so called “post-hoc correction” has to be applied, for example using a Bonferroni correction (Bonferroni 1936). Such corrections, however, increase the likelihood of Type II errors (genuine differences are no longer detected) and reduces the t-test’s statistical power (Perneger 1998; Bender and Lange 2001).

As a general guideline, when planning to repeatedly apply t-tests, the use of different randomly selected subsets of data sets is advised. If this is not feasible, a post-hoc correction should be applied. Alternatively, applying a one-way analysis of variances (ANOVA) or Dendrogrammatic Analysis of Particle Morphometry (DAPM) might be a better option.

One-way Analysis of Variances (ANOVA)

The term “ANOVA” refers to statistical procedures that serve to verify the differences of means across multiple data sets, based on tests that follow the F-distribution. In contrast to t-tests, ANOVA is designed to simultaneously test more than two data sets for significant differences (Davis 2002). The ANOVA’s null hypothesis is: H_0: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_n$ with μ_i being the mean of the *i-th* compared data set out of *n*. The alternative hypothesis H_1 is that at least one of the means is significantly different. As with F-tests, the F-values are computed and translated into an error likelihood *p* of improperly rejecting H_0. H_1 is verified if $p < \alpha$ (the level of significance). When reporting results of ANOVA, we recommend reporting both α and resulting *p* -values (see Table 4).

ANOVA assumes that data sets (Davis 2002):

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

a) are composed of randomly selected samples;
b) contain normally distributed samples;
c) have homogeneous variances.

These tests have been shown to be robust against violations of condition b) and c), particularly in cases where sample sizes are not too dissimilar (Ersoy et al. 2006; Blanca et al. 2017). We recommend to always use ANOVA analyses with similar-sized data sets.

In analogy to t-tests, Type I errors increase when using a data set for several repeated tests with ANOVA. To reduce Type I errors, post-hoc corrections can be applied. The most appropriate correction method depends on the validity of condition c). Examples of post-hoc corrections are:

- the Tukey's range test (also known as Tukey honestly significant difference) for data sets of homogeneous variances (Tukey 1949)
- the Games-Howell post-hoc adjustment (Games et al. 1979) which is a good option for testing data sets of heterogeneous variances.

Instead of just reporting whether H_1 is verified, post-hoc correction methods provide n x n matrices with adjusted p-values for all n tested data sets. However, like post-hoc corrections for t-tests, these adjustment procedures come with the cost of decreased statistical power, which becomes evident when comparing large numbers of data sets (Dürig et al. 2020c).

Equivalence tests (e-tests)

The failure of a two-tailed t-test or ANOVA to demonstrate a difference between data sets is not sufficient to mathematically prove similarity of two data sets (Walker and Nowacki 2011). For example, let us assume we are comparing the mean circularity of +1 phi juvenile clasts extracted from two pyroclastic samples, using a two-tailed t-test as described above. Thus, the null hypothesis H_0 is that both groups are extracted from the same population, because with the two-tailed t-test, we are hoping to reject H_0 by getting a p-value below α. That would demonstrate statistically – with a certain confidence level (1-α) – that there is a significant difference in the means. If instead the *p*-value is greater than α, we **fail to reject** H_0, but that does not imply that we can automatically **accept** H_0. To actually conclude that our +1 phi juvenile clasts are all likely to be derived from the same population, a different statistical test is needed, with different hypotheses.

A statistical method introduced to verify the equivalence of morphometric data sets is the equivalence test, or "e-tests" (Dürig et al. 2012a). An e-test checks whether the mean μ and the confidence interval $\Delta = [-C;C]$ of a data set lie within an acceptable range, specified by an equivalence margin D_{max}, so that (Rasch and Guiard 2004; Wellek 2010):

$$\mu - D_{max} < \mu - C < \mu < \mu + C < \mu + D_{max} \qquad (8)$$

For verification, an e-test uses one-tailed t-tests for each side of the equivalence margin, testing the composed null hypotheses H_{01}: $\mu - C < \mu - D_{max}$ and H_{02}: $\mu + C > \mu + D_{max}$. If the one-tailed t-test results lead to a rejection of both null hypotheses, relationship (8) is valid and statistical equivalence is verified (Rasch and Guiard 2004; Wellek 2010; Dürig et al. 2012a).

The 'classic' e-tests used exclusively Student's t-tests and could therefore only provide reliable results for data sets with homogeneous variances (e.g., Dürig et al. 2012a). With the recently published free and open software *DendroScan* (Dürig et al. 2020a), the range of application has been extended for cases of inhomogeneous variances by also including Welch's t-tests to the e-test procedures. Since e-tests are based on t-tests, the same conditions and assumptions apply.

The validity of e-tests depends on the quality of the pre-defined equivalence margin. An underestimation of D_{max} would result in a corridor that is too small, and therefore in Type II errors, where equivalences remain undetected. Conversely, overly large values would lead to Type I errors. It is therefore crucial to provide, along with *p*-values and α, *also* the D_{max} values when reporting the results of e-tests.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Since equivalence margins are specific to each shape parameter and each case (i.e., eruption or eruptive phase) (Dürig et al. 2012a, 2020b), a common strategy to find the appropriate D_{max} values is to use tailored calibration tests on so called "standards", i.e. subsamples of grains coming from the same population as the one to be tested. For calibration, e-tests are reiteratively computed for each shape parameter, by increasing the D_{max} values stepwise (e.g., by 0.01, with an initial value of 0.01), until the e-tests indicate statistical equivalence in all shape parameters (Dürig et al. 2020c, a).

For demonstration, the morphometric data set "d" (Fig. 2) is statistically tested for equivalency with data set "a" (Fig. 1). With this aim, e-tests are computed with *DendroScan* by using the data sets "AA", "AB" and "AC" (see Online Resource 1) as standards. According to the results (see Fig. 5), "a" and "d" can be treated as statistically equivalent.

Principal component analysis (PCA)

The statistical tests described thus far must be applied separately for each shape parameter. A morphometric comparison of *M* parameters requires the execution of at least 2 x *M* tests (for example, in the combination of F-tests and two-tailed t-tests). The number of results quickly becomes large and difficult to present, in the same way as does the presentation of multiple binary diagrams (see previous section).

In mathematical terms, each pyroclast can be represented by a data point in an *M*-dimensional vector space. Principal component analysis ("PCA") is a multivariate method that can be used to reduce the dimensions of this vector space (Maria and Carey 2002; Scasso and Carey 2005; Cioni et al. 2008; Suzuki et al. 2015; Schmith et al. 2017; Nurfiani and de Maisonneuve 2018; Pardo et al. 2020). In other words, PCA can reduce the number of variables in a way that retains as much of the original information as possible. It can also be used to explore the relationships among the original variables. PCA initially extracts *M* factors (denoted "principal components") by finding linear combinations of the original variables in the M-dimensional space. The principal components are constructed to be orthogonal to one another (so as to be statistically independent), and their length is

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

proportional to the total variance of the original data set. Next, the *M* principal components are sorted by their total variances (e.g., see Table 5). These total variances quantify the variance that can be explained by the principal components alone and are also denoted "Eigenvalues". The number of components extracted is based on a compromise between analytical tractability and loss of information. A typical decision criterion is the Kaiser normalization criterion (Kaiser 1958; Davis 2002), which suggests that only principal components with an Eigenvalue of 1 or larger be considered. PCA is particularly useful in complex multivariate analysis, when dealing with a multitude of different parameters, e.g., from different morphometric systems, to reveal redundancies (i.e., variables that actually do not add additional information) and to help find the most meaningful parameters.

In a simple example shown by Table 5, we applied PCA to four shape parameters. This approach would lead to the use of principal components 1 and 2 and therefore a dimensional reduction from four to two.

Table 6 (left) shows the Pearson correlation coefficients for each variable and component, denoted "factor loadings". Often, it is useful to redistribute the factor loadings in a way that facilitates interpretation of a component's meaning. A typical approach for achieving this goal uses the "varimax rotation" (Davis 2002), which rotates the components (and with them the coordinate axes), but keeps the components orthogonal. Table 6 (right) provides an example of the resulting component matrix after such a rotation: now component 1 can be seen as a measure of Circ_DL and Rec_DL, while component 2 is mainly measuring Com_DL and Elo_DL.

For each of our demonstration data sets (Fig. 1, Fig. 2) the resulting component scores are listed (see Online Resource 3). For four data sets, Figure 3e shows the (unrotated) principal components. Data points of four objects were individually tagged in Figures 3c through 3f to 'track' them. For example, Figures 3c and 3d show that "b50" (black triangle within a black circle) is characterized by high circularity and rectangularity, medium compactness and low elongation. Although the (unrotated) components contain this information, it is difficult to

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

reconstruct it from Figure 3e. Only after the "varimax" rotation (Fig. 3f) does it become apparent that "b50" is characterized by a large value for rotated component 1 and a low value for rotated component 2. According to the resulting rotated-component matrices (Table 6), we know that the first rotated component is strongly correlated with circularity and rectangularity, whereas the second one shows a positive correlation mostly with elongation and is negatively correlated with compactness. It is hence possible to infer the original shape parameters from Figure 3f. Although this example only includes four original variables, PCA would become even more useful if 10 or 20 morphometric parameters were involved.

When presenting PCA results, it is mandatory to specify the decision criterion used to choose the level of dimensional reduction (in our example: "Kaiser normalization criterion") and type of rotation (here: "varimax") applied. Along with the total variances (Table 5) and the resulting component scores for each sample, it is also recommended that resulting component score coefficients be reported. Such a table is also known as a "Component Score Coefficient Matrix" (e.g., Table 7).

The resulting principal components are statistically *independent* variables. PCA can therefore be used as a first step for the subsequent application of statistical methods that require independent variables, such as discriminant function analysis (see next section).

Factor analysis

Being closely related to PCA, factor analysis methods, such as the R-mode type factor analysis (Dellino and La Volpe 1996; Davis 2002; Dellino and Liotino 2002) are used to reduce the number of explanatory variables (i.e., morphometric parameters) without losing relevant information. As for PCA, the original variables are linearly combined to construct the equivalent to principal components, which are (unsurprisingly) named "factors". In contrast to PCA, however, factors are not orthogonal, and therefore not independent variables. Instead, the new axes are orientated in a way that optimally describes the original data variances. Factor analysis can be used (1) as a data reduction method, although PCA may be better suited for this, (2) as an "exploratory" tool, in order to find hidden and not directly measurable ("latent") dependencies between variables that explain the distribution of factor scores, or (3) to test the validity of an *a priori* model ("confirmatory factor analysis"). In the context of morphometry, to our knowledge, only the first usage has been applied so far (Dellino and La Volpe 1996; Dellino and Liotino 2002). We strongly recommend that researchers report factor scores along with the factor loadings, eigenvectors and score weights, since all of these parameters are required for a full analysis and interpretation of the data (Dellino and La Volpe 1996; Davis 2002; Dellino and Liotino 2002).

Cluster analysis

Cluster analysis is the collective term for a suite of exploratory statistical techniques to sort observations (here: particles) according to their relatedness and assign them to relatively homogeneous groups ("clusters"). When applied to morphometric data sets, the members of such clusters are characterized by sharing a set of features, while simultaneously being distinct from members of the other clusters (Dellino et al. 2001; Davis 2002). Analysing these groupings and investigating common links that connect the members of a cluster play an important role in morphometric analysis and are used to infer the influences of eruptive processes on particle formation (Dürig et al. 2020c). The implementations of cluster analysis are many. In the following we concentrate only on those most commonly used in morphometry: hierarchical cluster analysis and the k-means procedure.

Hierarchical cluster analysis

Hierarchical cluster analysis can be further differentiated into "agglomerative" and "divisive" analyses.

An agglomerative hierarchical cluster analysis starts with a single observation (i.e., pyroclast), treating it as a preliminary cluster. From the remaining particles, the one identified as "most similar" to the first one is joined. This procedure is repeated until all pyroclasts are included.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

The “divisive” hierarchical cluster analysis works in reverse (“top-down”): starting with all particles as one cluster, the algorithm partitions it into two least-similar sub-groups. This procedure is reiterated for each sub-group, then repeated to the next cluster level and so on.

The algorithm’s decision on which particle to add (or, in the divisive class of cluster analysis, where to divide the original cluster) depends on:

- what is used as a measure of dissimilarity;
- which points within the clusters are used as references for measuring the group’s distance (known as linkage in the context of cluster analysis).

In morphometric cluster analysis, common measures of dissimilarity are the normalized Euclidian distance (Dellino et al. 2001; Maria and Carey 2002; Cioni et al. 2008) or the squared Euclidian distance (Rausch et al. 2015). A specially defined distance is used in DAPM (see below).

Linkages commonly used in morphometric cluster analyses are: “single linkage” (Dellino et al. 2001), where the clustering algorithm computes the distances between the nearest neighbours, or “complete linkage“ (Maria and Carey 2002; Dürig et al. 2020c), which uses the farthest neighbours of each group. Other examples include “average” or “median” linkage (Davis 2002). With the plethora of implementations of this method, it is critical that researchers provide sub-type, measure of dissimilarity and linkage method used when publishing results obtained by a hierarchical cluster analysis.

In morphometric studies, hierarchical cluster analyses has often been applied at the level of individual pyroclasts (Maria and Carey 2002; Rausch et al. 2015), in order to group clasts of similar origin. Occasionally the means of measured shape parameters for groups of pyroclasts have been used instead (e.g., Dellino et al. 2001), and in this case it is critical to clearly explain the contents of the groups, and how particles were assigned to each group.

Any in-built correlation between shape parameters would result in a bias in the actual groupings based on the Euclidean distances. In order to reduce this effect, hierarchical cluster analyses are often combined with principal component analyses, and applied to the statistically independent principal components found (Maria and Carey 2002; Cioni et al. 2008).

The output of a hierarchical cluster analysis is a tree diagram, or “dendrogram”, that displays the dissimilarities among the tested shapes. An example of a dendrogram is shown in Fig. 6. It is the result of an agglomerative hierarchical cluster analysis, constructed on data sets “a”, “b”, “c” and “d”, using the squared Euclidean distance with complete linkage. From the example, it is evident that interpretation of the groupings illustrated in the dendrogram is not trivial; this is because differences among individual particles within data sets causes the particles to be grouped into different clusters. For example, from Figure 6 it is not immediately clear that data sets “a” and “d” are, in fact, statistically equivalent.

k-means procedure

The k-means procedure, introduced by MacQueen (1967), is a special type of cluster analysis. In contrast to a hierarchical cluster analysis, where the number of clusters k is provided as output, the k-means procedure works with a user-defined fixed value for k. It classifies the data by assigning it to the k clusters and computes their centroids. The algorithm begins by randomly selecting k data points as initial seeds (Davis 2002). It then assigns the N observations to the “most similar” seeding points, by using the minimum increase of variance as decision criteria. Using the centroids of each of the k clusters as the next seed, this procedure is reiterated, until stable centroids of the clusters (k-means) are obtained (Davis 2002).

The k-means procedure can be used when the user can guess the number of clusters into which the data will/should cluster. A possible field of application is data reduction, by replacing the individual data points with data from the k centroids. Another use of the k-means procedure is to explore similarities among morphometric data sets with varying k. For example, using $k = 3$ and $k = 2$ in a comparison of volcanic ash samples from three different volcanoes, researchers provided information about the degree to which each volcano has tephra that can be distinguished from those of other volcanoes (Avery et al. 2017).

For demonstration, let us assume our aim is to find out which of the three sets of objects “a”, “d” and “e” (see Figs. 1-2) show the highest morphometric similarity, based on the four shape parameters by Dellino and La Volpe (1996). We start by applying the k-means procedure using k = 3 and compare the outcome with the results for k = 2 (see Fig. 7 and Online Resource 4). For illustration purposes, after application of the k-means procedure, we applied PCA with varimax rotation and plotted the two principal components to illustrate the clusters.

For k = 3 (Fig. 7a), all objects of group “e” were assigned to cluster 1 (red), whereas the bulk of “a” and “d” objects were grouped into cluster 2 (blue). Cluster 3 (green) comprises only two objects (one from sample “a” and one from “d”). For k = 2 (Fig. 7b), all objects of “e” are members of cluster 1 (blue), whereas most members of “a” and “d” were assigned to cluster 2 (red). We can infer from these results that the shapes of “a” and “d” objects are overall more similar to each other than to those from sample “e”.

Although data clustering is a useful method for data exploration, interpretation of the cluster assignments may become complex and is somewhat user-dependent. Also, the k-means procedure requires normality of input data (implying large sample sizes) and is less robust than, e.g., the e-test. When publishing results of the *k*-means procedure, the initial conditions, along with the coordinates of the k centroids, need to be reported to facilitate interpretation of groupings.

Dendrogrammatic Analysis of Particle Morphology (DAPM)

DAPM is a recently published technique designed for comparative analysis of multiple data sets (Dürig et al. 2020c). Technically, it can be seen as a special variant of hierarchical cluster analysis which combines all the aforementioned statistical tests (F-tests, ANOVA, two-tailed t-tests and e-tests) in order to produce dendrograms displaying degrees of dissimilarity among data sets. In contrast to the other types of cluster analyses, which are usually applied to individual particles (Maria and Carey 2002; Cioni et al. 2008; Rausch et al. 2015), DAPM is tailored for analysis of dissimilarities and similarities among different data sets, each representing many particles, by means of their variances.

When analysing *Q* data sets, the DAPM's initial step is to compare all data sets with *M* shape parameters by F-tests, followed by ANOVA with the appropriate post-hoc correction (Tukey's range test or Games-Howell post-hoc adjustment). Starting from these results, the elements of a distance matrix *X* are computed by:

$$X_{ij} = \sum_{k=1}^{M} Y_{ijk} \tag{9}$$

where Y_{ijk} is calculated according to:

$$Y_{ijk} = \begin{cases} log\left(1 + \frac{1}{p_{ijk}}\right) & if\ p_{ijk} < 0.05 \\ 0 & if\ p_{ijk} \geq 0.05 \end{cases} \tag{10}$$

and p_{ijk} is the ANOVA's *p*-value of data set *i* tested with the one from data set *j* in the k-th shape parameter.

The use of the entries of X as measures of dissimilarity, along with the complete linkage method, allows the construction of a "level 1" dendrogram, which groups the *Q* data sets according to their relative morphometric differences (Dürig et al. 2020c).

If the number of data sets is relatively large ($Q > 7$ (Dürig et al. 2020a)), this initial sorting is to be treated as preliminary because, according to the considerations above, the statistical power of ANOVA is expected to be low. Still, the level 1 diagram can be used to identify the main morphometric clusters and to split the *Q* data sets into sub-sets, for which the computation procedure of *X* is repeated, resulting in several "level 2" dendrograms. By increasing stepwise the "levels", this procedure is reiterated for each of the new sub-clusters, until no further cluster separation is possible.

Data sets grouped with a dissimilarity of 0 represent the highest level dendrograms. They are tested, by using the *M* shape parameters, with two-tailed t-tests.

The data sets that “fail” the t-tests (i.e., no significant differences indicated for any of the M shape parameters), are submitted to e-tests, using pre-defined equivalent margins.

With this procedure, it is not only possible to sort multiple data sets according to their ANOVA-verified dissimilarity, but also to identify those data sets among them that are statistically equivalent according to a clear set of rules. This is useful, for example, in linking pyroclasts obtained from experiments with natural ones.

Recently released freeware, which allows the analyst to perform a DAPM automatically, is *DendroScan* (Dürig et al. 2020a). An example of a DAPM-based dendrogram is provided in Figure 8.

The use of eq. (9) in the construction of the distances X could lead to an overestimation of some morphologic features, especially when there is an underlying correlation between some of the shape parameters. The decision about which parameters to select for DAPM should therefore be based on the final aim of the analysis. If the aim is to identify the most significantly different data sets, and separate them from those which are statistically equivalent, it is recommended to use all the shape parameters. This approach ensures completeness of the morphological features in DAPM. If the aim is, instead, to interpret the degree of dissimilarities between significantly different data sets, it is advisable to use a reduced set of statistically independent parameters (Dürig et al. 2020a), possibly by using binary diagrams to ensure the absence of correlation.

When publishing results from DAPM, researchers should include the shape parameters, D_{max} values and level of significance α used, along with the dendrograms produced at the highest data levels. In addition, the distance matrices X might be provided. Researchers must state the linkage used for computing the dendrograms.

Determination of fragmentation mechanism and eruptive style

Discriminative and interpretive diagrams

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

An important application of morphometric analysis is in the classification of tephras according to their genesis, for example the distinction of ash produced by phreatomagmatic versus magmatic fragmentation processes. Qualitative schemes of tephra classification by particle morphology have a decades-long history (Heiken 1974; Wohletz 1983; Büttner et al. 1999; Taddeucci et al. 2002; White and Valentine 2016; Németh and Kósik 2020). Some attempts have also been made to develop user-independent methods for quantitative and reproducible classification (e.g., Büttner et al. 2002; Murtagh and White 2013; Schmith et al. 2017). The most common method for discriminating among eruption mechanisms is to plot samples in classification diagrams (for examples, see Table 2) that distinguish fields of different eruptive conditions. Plots used for discriminative interpretation range from simple binary diagrams (e.g., Cioni et al. 2014; Leibrandt and Le Pennec 2015; Liu et al. 2015), to diagrams plotting combinations of shape parameters (Büttner et al. 2002; Murtagh and White 2013; Iverson et al. 2014; Alvarado et al. 2016), to more complex approaches, in which interim parameters are derived from linear interpolation based on binary plots (Schmith et al. 2017).

To demonstrate the use of two of these classification diagrams, we compare two morphometric data sets obtained from silhouettes of ash particles from the narrow +4ϕ (88-63 μm) grain size fraction (see Online Resource 2). The first data set, denoted “Iki”, describes the shape of grains produced during continuous lava fountaining episodes of the 1959 Kīlauea Iki eruption (Richter et al. 1970; Mueller et al. 2018, 2019). The second data set was obtained from ash particles produced in significant amounts during the 2012 eruption of Havre, a silicic deep-sea volcano (Carey et al. 2018). Based on morphometric comparisons with samples from lab experiments, it was found that a phreatomagmatic mechanism played a key role in the ash generating episode(s) of this eruption (Dürig et al. 2020b, c). Table 8 presents an overview of the resulting shape parameters *Circ_DL*, *Elo_DL*, *Rec_DL* and *Com_DL.*

Figure 9a shows the classification diagram by Büttner et al. (2002) plotted with data from the two demonstration sets. This plot has been designed to distinguish grains that were the product of brittle fragmentation from those generated under ductile fragmentation conditions.

While the authors originally identified a threshold of approximately 0.88 on the y-axis for shoshonite clasts (Büttner et al. 2002), a revised value of 0.71 was suggested in a later study for Havre ash (Dürig et al. 2018). Both thresholds are displayed in Figure 9a as dashed horizontal lines. When using the threshold suggested by Dürig et al. (2018), the majority of the Iki samples fall into the ductile field, while the bulk of Havre samples plot in the brittle field. There are, however, also outliers in both data sets, which demonstrate the necessity of using sufficiently large numbers of pyroclasts to extract useful information from these types of diagrams. Note also that the morphometric variance among the Iki grains is larger than that of Havre particles, reflecting the considerably larger standard deviations of the underlying shape parameters (see Table 8).

Figure 9b presents an alternative classification diagram, following the suggestion of Murtagh and White (2013). Here, the suggested boundary is an ascending line (illustrated as a dashed line in Fig. 9b), that separates morphometric data points from particles of phreatomagmatic origin (left side) from those of magmatic origin (right side). For our test samples, the diagram is fairly successful in sorting the two populations by their eruptive mechanism, especially when focussing on the mean values of both data sets. However, rather large minorities of 14 (27.5%) and 12 (25%) grains from the Iki and Havre samples, respectively, are sorted into the 'wrong' sector. We also note that there is a substantial overlap between the two samples when taking their standard deviations into account. In our demonstration this does not affect the overall outcome, because the samples studied here are representing end members on the scale of eruptive styles. For other samples, the results might be far less clear, rendering the method unreliable (Schmith et al. 2017). The same applies to the discrimination diagram by Büttner et al. (2002). Although it has been widely used for distinguishing between phreatomagmatic and magmatic grains (e.g., Németh and Cronin 2011; Murtagh and White 2013; Iverson et al. 2014; Alvarado et al. 2016), a number of studies found that it is difficult to identify a universal threshold that defines a clear distinction between fields. This probably reflects the roles of magma chemical composition and physical magma components (melts, bubbles and crystals) on the mechanical behaviour of magma under stress, which is the direct control on shapes of pyroclasts (Murtagh and White 2013; Schmith et al. 2017; Dürig et al. 2018). Morphometric analyses can therefore only provide a piece of the puzzle, to be considered along with whole-deposit componentry and granulometry (Mele et al. 2020), and analysis of particles' microtextures and surface features (White and Valentine 2016; Ross et al. 2021).

Discriminant function analysis (DFA)

As a multivariate statistical method that classifies data sets and provides error likelihoods of classification, discriminant function analysis (DFA) has potential to establish relationships between particle morphology and fragmentation processes or eruption styles. DFA requires data sets with known group membership (e.g., data sets from known purely single-process magmatic or single-process phreatomagmatic fragmentation processes).

Similar to PCA and factor analysis, as a first step, the discriminant function analysis seeks to reduce the number of variables by combining the original variables in a way that maximizes the differences between groups and minimizes the variance within each group (Davis 2002). Next, the algorithm finds a discriminant function, which is tailored to separate the data into the previously defined groups, i.e. to discriminate between them. The predictive quality of this function can be tested by computing the percentage of known data sets that are correctly classified, and this success percentage is then listed in classification matrices. Based on the discriminant function, the DFA is subsequently able to predict the group-membership of unclassified data sets. Furthermore, the structure of the separation function provides insights into which of the variables (i.e. parameters) has the most discriminatory power (Avery et al. 2017).

A caveat, additional to starting with data sets of known origin, is that DFA requires normally distributed data and independent variables for both the initial data sets and those subsequently investigated. A way to obtain independent variables is to first apply a PCA then use the resulting principal components as input. When reporting results from a DFA, researchers should provide comprehensive information, including the definition of grouping variables, normality of data, equality of co-variance matrix, missing data and outliers, variables used, correlation matrices, software and version used, classification matrices, discriminant function and classification function weights (for details, see e.g., Huberty and Hussein 2003).

Supervised machine learning

Supervised machine learning methods can be somewhat similar in aim to what has been described for DFA in which the algorithms are trained to recognize membership in a group. Machine learning approaches that have been used for classification of particle shapes are, e.g., decision-trees, random forest (Tunwal et al. 2018) and convolutional neural networks (Shoji et al. 2018). In the first method, a Classification And Regression Trees (CART) algorithm uses the training data to build a decision tree, which then is applied as a predictive model to classify the unknown data set (Loh 2011). A fixed decision tree might fit too exactly to the noise-affected training data and not take stochastic variations of the test data into account (“overfitting”). To counter this effect, a random forest algorithm can be used, which builds and combines large numbers of decision trees based on random selection of shape parameters and sampling of training data (Breiman 2001). Convolutional neural networks (CNN) take an alternative path and are specifically designed for image recognition. CNN algorithms process pixel intensities in several layers, where the early layers focus on simple features and later layers recognize patterns of increased complexities. By using large numbers of particle images as training data, CNN can classify new particles according to the acquired model. Potentially we could teach an algorithm to distinguish magmatic from phreatomagmatic particles (Shoji et al. 2018). Future applications might combine machine learning with some of the previously discussed statistical techniques. For example, PCA and the k-means procedure might be used as a first step to obtain training data, before applying CNN. Similar approaches have already been successfully applied in other fields of research (e.g., Tang et al. 2017; Rustam et al. 2020). A disadvantage of machine learning approaches is, however, that the algorithms are somewhat like “black boxes” and as such might lead to misinterpretation by the user.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Conclusions and Outlook

We have provided an overview of the statistical methods commonly used to analyse morphometric data sets. Table 9 summarizes the purpose, mathematical pre-conditions and output of each of the previously discussed tests and algorithms. Our aim is to explain these techniques in a way accessible to geologists, and we have illustrated the methods using simple particle shapes. With this statistical toolkit at hand, morphometric data sets can be explored while simultaneously understanding the mathematical limitations that attach to each of the methods applied.

Although we present a broad overview for volcanology, the presented analytical tools represent only a small selection of all techniques available. With ever-increasing computational capabilities, machine learning techniques may become more and more important as complementary analytical tools, leading to more-complex routines for shape analysis. Together with the ongoing development of 3D scanning technologies, the near future promises new advances in the quest to decode the volcanological information ingrained in the shapes of volcanic particles.

Tables

Table 1 Example for basic descriptive statistics of the morphometric data sets, based on *N* samples. The particles are shown in Figure 1 and 2. Note that this data serves only for demonstration purposes. For representativity of “real” volcanic ash analyses, more particles (i.e., larger sample sizes *N*) would be required.

data set		N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
							Statistic	Std. Error	Statistic	Std. Error
a	Circ_DL	12	1.07	1.35	1.13	0.07	2.71	0.64	8.21	1.23
	Rec_DL	12	0.83	1.03	0.86	0.06	3.07	0.64	9.77	1.23
	Com_DL	12	0.51	0.99	0.74	0.11	0.21	0.64	3.76	1.23
	Elo_DL	12	1.06	2.41	1.50	0.32	2.18	0.64	6.90	1.23
b	Circ_DL	4	1.63	3.19	2.39	0.67	0.12	1.01	-1.20	2.62
	Rec_DL	4	1.18	2.30	1.73	0.48	0.14	1.01	-1.24	2.62
	Com_DL	4	0.66	0.67	0.66	0.00	1.83	1.01	3.43	2.62
	Elo_DL	4	1.31	1.44	1.39	0.06	-1.23	1.01	0.77	2.62
c	Circ_DL	8	1.05	1.81	1.37	0.29	0.41	0.75	-1.57	1.48
	Rec_DL	8	0.83	1.26	1.02	0.17	0.32	0.75	-1.70	1.48
	Com_DL	8	0.62	0.79	0.72	0.07	-0.33	0.75	-1.74	1.48
	Elo_DL	8	1.29	1.50	1.40	0.07	-0.04	0.75	-0.58	1.48
d	Circ_DL	12	1.07	1.35	1.13	0.08	2.35	0.64	5.91	1.23
	Rec_DL	12	0.83	1.07	0.86	0.07	3.38	0.64	11.59	1.23
	Com_DL	12	0.50	1.00	0.74	0.11	0.20	0.64	4.56	1.23
	Elo_DL	12	1.00	2.29	1.48	0.30	1.78	0.64	6.07	1.23
e	Circ_DL	8	1.10	1.55	1.32	0.15	0.06	0.75	-0.78	1.48
	Rec_DL	8	0.84	1.12	0.98	0.09	-0.01	0.75	-0.79	1.48
	Com_DL	8	0.70	0.79	0.75	0.03	-0.02	0.75	-0.83	1.48
	Elo_DL	8	2.05	2.26	2.15	0.07	0.27	0.75	-1.09	1.48
h	Circ_DL	12	1.33	4.61	2.43	0.98	1.14	0.64	0.98	1.23
	Rec_DL	12	0.84	2.64	1.43	0.53	1.27	0.64	1.30	1.23
	Com_DL	12	0.41	0.73	0.46	0.09	2.94	0.64	9.10	1.23
	Elo_DL	12	1.08	2.46	1.92	0.37	-0.89	0.64	1.60	1.23
r	Circ_DL	7	1.14	1.73	1.44	0.21	-0.10	0.79	-1.19	1.59
	Rec_DL	7	1.00	1.41	1.21	0.15	-0.18	0.79	-1.20	1.59
	Com_DL	7	0.87	1.00	0.93	0.05	0.15	0.79	-1.02	1.59
	Elo_DL	7	1.31	1.40	1.35	0.03	-0.18	0.79	0.72	1.59
s	Circ_DL	8	1.27	2.98	2.12	0.67	-0.13	0.75	-1.58	1.48
	Rec_DL	8	0.86	1.23	1.05	0.15	-0.33	0.75	-1.79	1.48
	Com_DL	8	0.20	0.58	0.36	0.15	0.67	0.75	-0.94	1.48
	Elo_DL	8	1.61	2.86	2.15	0.43	0.18	0.75	-0.32	1.48

Table 2: List with examples of interpretative diagrams and diagrams used to discriminate between particles formed by phreatomagmatic vs. magmatic fragmentation. Morphometric parameters are those used in the presenting publication (fourth column). In addition, the terms applied to the shape parameters used in the open freeware PARTISAN (Dürig et al. 2018) are provided in the fifth column.

Introduced by	Purpose	Plot (y-axis; x-axis)	Morphometric system	PARTISAN output variables
Büttner et al. (2002)	Brittle vs ductile (phreatomagmatic vs magmatic)	Rectangularity x Compactness; Elongation x Circularity	Dellino and La Volpe (1996)	Rec_DL x Com_DL; Elo_DL x Circ_DL
Murtagh and White (2013)	Phreatomagmatic vs magmatic	Elongation x Compactness; Rectangularity x Circularity	Dellino and La Volpe (1996)	Elo_DL x Com_DL; Rec_DL x Circ_DL
Cioni et al. (2014)	Interpretative binary diagram	Ellipse Aspect Ratio; Solidity	Cioni et al. (2014)	AR_CI; Sol_CI
Leibrandt and Le Pennec (2015)	Interpretative binary diagram	Convexity; Circularity	Leibrandt and Le Pennec (2015)	Con_LL;Circ_LL
Liu et al. (2015)	Interpretative binary diagram	Convexity; Solidity	Liu et al. (2015)	Con_LI; Sol_LI
Schmith et al. (2017)	Elongated vs non-elongated grains; slope of linear interpolation defines “regularity index”	Regularity; Feret Aspect Ratio	Schmith et al. (2017)	Reg_SC; AR_SC
Schmith et al. (2017)	Phreatomagmatic vs magmatic	regularity index; percentage of elongated grains	Schmith et al. (2017)	

Table 3: Results for Levene tests and t-tests used to compare the morphometric data sets "a" with "c". Numbers are *p*-values in percent. When the *p*-value is below the level of significance α (often 5%), the tested variances (in case of Levene-test) or means (in case of a t-test) can be inferred to be significantly different. A graphic form of displaying these results is presented on Figure 4.

α = 5%	Circ_DL	Rec_DL	Com_DL	Elo_DL
Levene test	0.13	0.23	54.41	6.88
t-test	6.03	4.13	46.16	71.50

Table 4: Example for ANOVA results. Resulting F- and *p*-values are displayed. A *p*-value below the level of significance (here 5%) indicates a significant difference. For example, when comparing the morphometric data sets “a”, “b”, “c”, and “e” (two centre columns), ANOVA reveals significant differences in circularity (Circ_DL), rectangularity (Rec_DL) and elongation (Elo_DL). This result implies that at least two of the four samples are different in these parameters.

morphometric data sets	a,b,c		a, b, c, e		a, b, c, e, h, r, s	
α = 5%	F	p (%)	F	p (%)	F	p (%)
Circ_DL	25.151	<0.05	20.954	<0.05	8.72	<0.05
Rec_DL	26.039	<0.05	22.021	<0.05	7.626	<0.05
Com_DL	1.144	33.8	1.196	32.9	37.514	<0.05
Elo_DL	0.545	58.8	22.988	<0.05	13.574	<0.05

Table 5: Dimensional reduction by principal component analysis (PCA) with “varimax” rotation: In this example, PCA was conducted via software SPSS®, based on 8 data sets (“a, b, c, d, e, h, r, s”) and using the four shape parameters employed by Dellino and La Volpe (1996) as variables. Thus, PCA initially extracted four principal components from the four original shape parameters Circ_DL, Rec_DL, Com_DL and Elo_DL. The total variances of the principal components are denoted “Eigenvalues”. In order to reduce the number of explaining parameters one could follow the Kaiser normalization criterion (Kaiser 1958; Davis 2002), which suggests consideration of only principal components with eigenvalue 1 or larger. Using the first two components instead of the original four shape parameters would reduce the dimension by two, but still be sufficient to explain ~93.6% of the total variance.

	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.471	61.77	61.77
2	1.273	31.81	93.59
3	0.240	6.009	99.59
4	0.017	0.41	100.00

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Table 6: Components and rotated-component matrices: values represent the Pearson correlation between variables and components, denoted “factor loadings”. The original components show cross-correlations, which complicate interpretations. On the right, the redistributed factor loadings after a so called “varimax” rotation (Davis 2002) are shown. Now Component 1 is dominantly measuring circularity (Circ_DL) and rectangularity (Rec_DL), while Component 2 can be seen as mainly a measure of compactness (Com_DL) and elongation (Elo_DL).

	Components		Rotated components	
	1	2	1	2
Circ_DL	0.940	0.328	0.932	0.348
Rec_DL	0.740	0.649	0.984	-0.026
Com_DL	-0.842	0.420	-0.379	-0.861
Elo_DL	0.576	-0.754	-0.039	0.948

Table 7: Component score coefficient matrix resulting from PCA, using four shape parameters as input variables and the “varimax” rotation (see also Table 5). The principal component score of a sample is calculated by linearly combining the sample-specific shape parameter values, weighted with the according component score coefficients.

Component Score Coefficient Matrix		
	Component	
	1	2
Circ_DL	0.457	0.045
Rec_DL	0.556	-0.201
Com_DL	-0.051	-0.472
Elo_DL	-0.199	0.604

Table 8: Basic descriptive statistics of the morphometric data sets used to plot the discrimination diagrams shown in Figure 9. The shape parameters were obtained from silhouettes of ash particles sampled from the 1959 Kīlauea Iki (“Iki”) and the 2012 Havre eruption.

		N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
Data set							Statistic	Std. Error	Statistic	Std. Error
Iki	Circ_DL	51	1.05	2.69	1.74	0.39	0.26	0.33	0.06	0.66
	Rec_DL	51	0.83	1.31	1.03	0.12	0.43	0.33	-0.43	0.66
	Com_DL	51	0.33	0.83	0.57	0.14	0.38	0.33	-0.82	0.66
	Elo_DL	51	1.14	18.39	3.69	3.32	2.83	0.33	8.83	0.66
Havre	Circ_DL	48	1.13	1.70	1.33	0.14	0.90	0.34	0.33	0.67
	Rec_DL	48	0.90	1.23	0.98	0.07	1.58	0.34	3.00	0.67
	Com_DL	48	0.55	0.86	0.76	0.06	-1.33	0.34	2.86	0.67
	Elo_DL	48	1.24	5.13	2.19	0.84	1.75	0.34	3.15	0.67

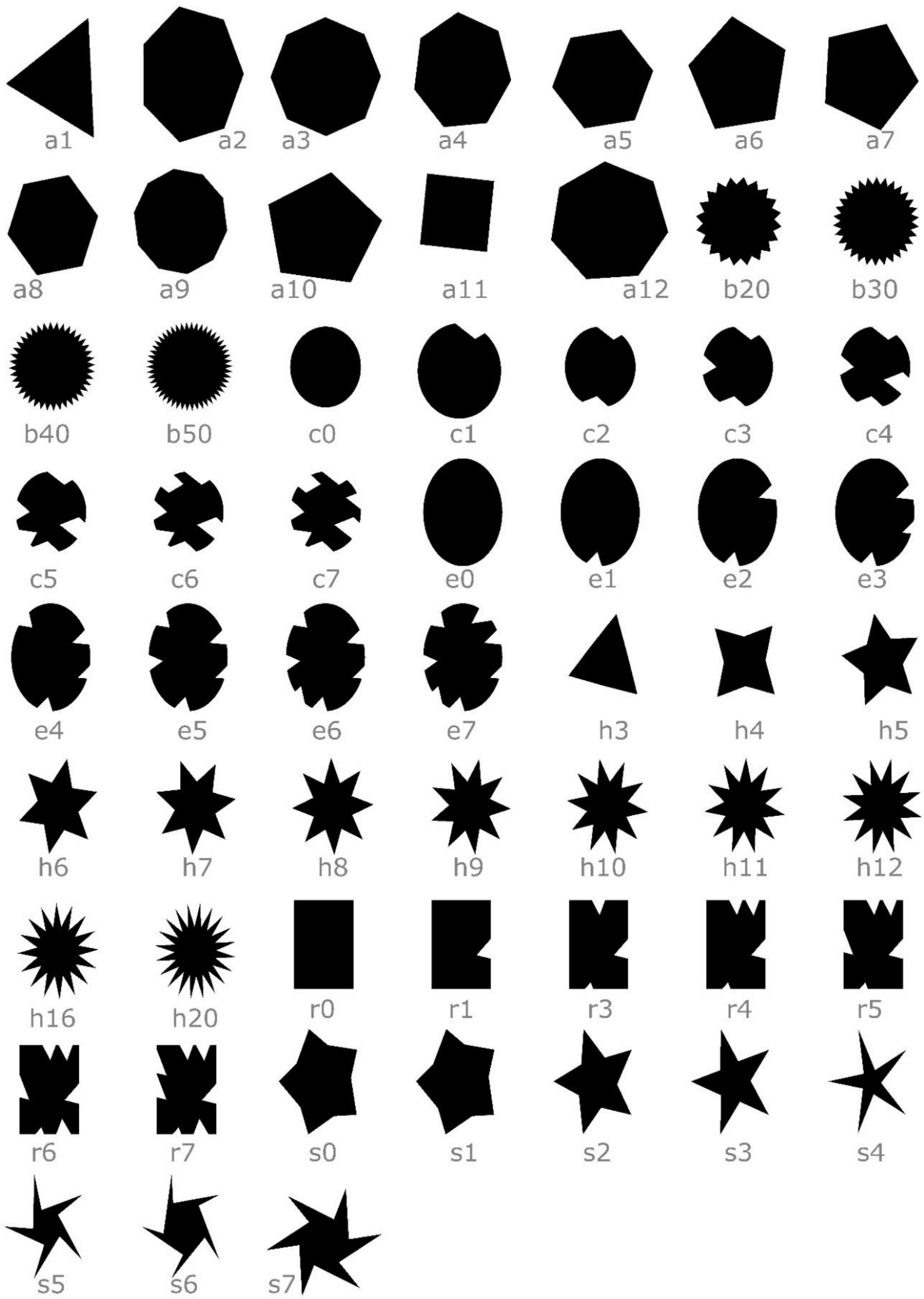
Table 9: List of methods, fields of application, necessary condition(s) and main output. Brackets indicate that a method remains fairly robust if conditions are violated.

Method	Purpose	Condition	Output	Additional information to be disclosed
F-test	Testing if variances of two data sets are equal	(Normal distribution)	*p*-value	α, sample sizes (*N*)
Levene test	Testing if variances of multiple data sets are equal		*p*-value	α, *N*
Student's t-test	Testing 2 data sets of equal variances for significant differences	(Normal distribution)	*p*-value	α, *N*
Welch's t-test	Testing 2 data sets of unequal variances for significant differences	(Normal distribution)	*p*-value	α, *N*
ANOVA	Testing 3 or more data sets for significant differences	(Normal distribution)	*p*-value	α, *N*, post-hoc correction (if applied)
Equivalence test	Testing 2 data sets for statistical similarities	(Normal distribution)	Yes/no, D-value	α, D_{max} values, standards used for calibration
Factor analysis	Dimensionality reduction, revealing underlying "latent" variables	Normal distribution; otherwise correction for non-normality needed, e. g., adjustment by Satorra-Bentler (1994)	Factor scores	Type of factor analysis, factor loadings, eigenvectors, score weights
Principal component analysis	Dimensionality reduction	(Normal distribution)	Principal components PC1, PC2	Decision criterion, type of rotation, Eigenvalues, component score coefficients
Hierarchical cluster analysis	Sorting of individual grains based on their Euclidian distance		Dendrogram (dissimilarity axis)	Type of cluster analysis, measure of dissimilarity, linkage
K-means procedure	Sorting of individual grains, based on their distance from *k* centroids, where number *k* is pre-defined; data reduction: resulting cluster centroids can be used instead of individual data points.	Normal distribution	Sorting of grains into the *k* clusters; coordinates of the *k* cluster centroids	Number of clusters *k*; coordinates of the initial seeding points
DAPM	Sorting of data sets, based on the outcome of sequential ANOVA, t-tests and equivalence tests	(Normal distribution)	Dendrogram (dissimilarity axis), distance matrices X	Used shape parameters, α, D_{max} values, standards used for calibration, type of linkage
Discriminant analysis	Discriminating data sets	Normal distribution, independent variables	Classification matrix	Definition of grouping variables, variables used, co-variance matrix, missing data, outliers, correlation matrices, software, discriminant

				function, classification function weights

Figures

Figure 1



This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Fig. 1 Binary images (silhouettes) of “artificial” objects used for demonstration. The names of particles are indicated. The compiled morphometric data sets are labelled “a”, “b”, ”c”, ”e”, ”h”, ”r” and “s”.

Figure 2

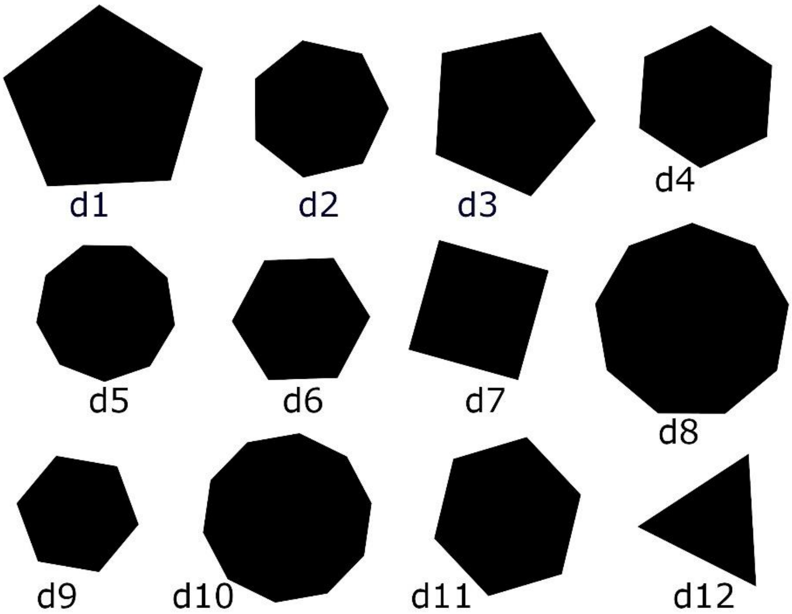


Fig. 2 Binary images of data set “d”. Using e-tests and DAPM we test whether “d” is statistically equivalent with “a” from Figure 1.

Figure 3

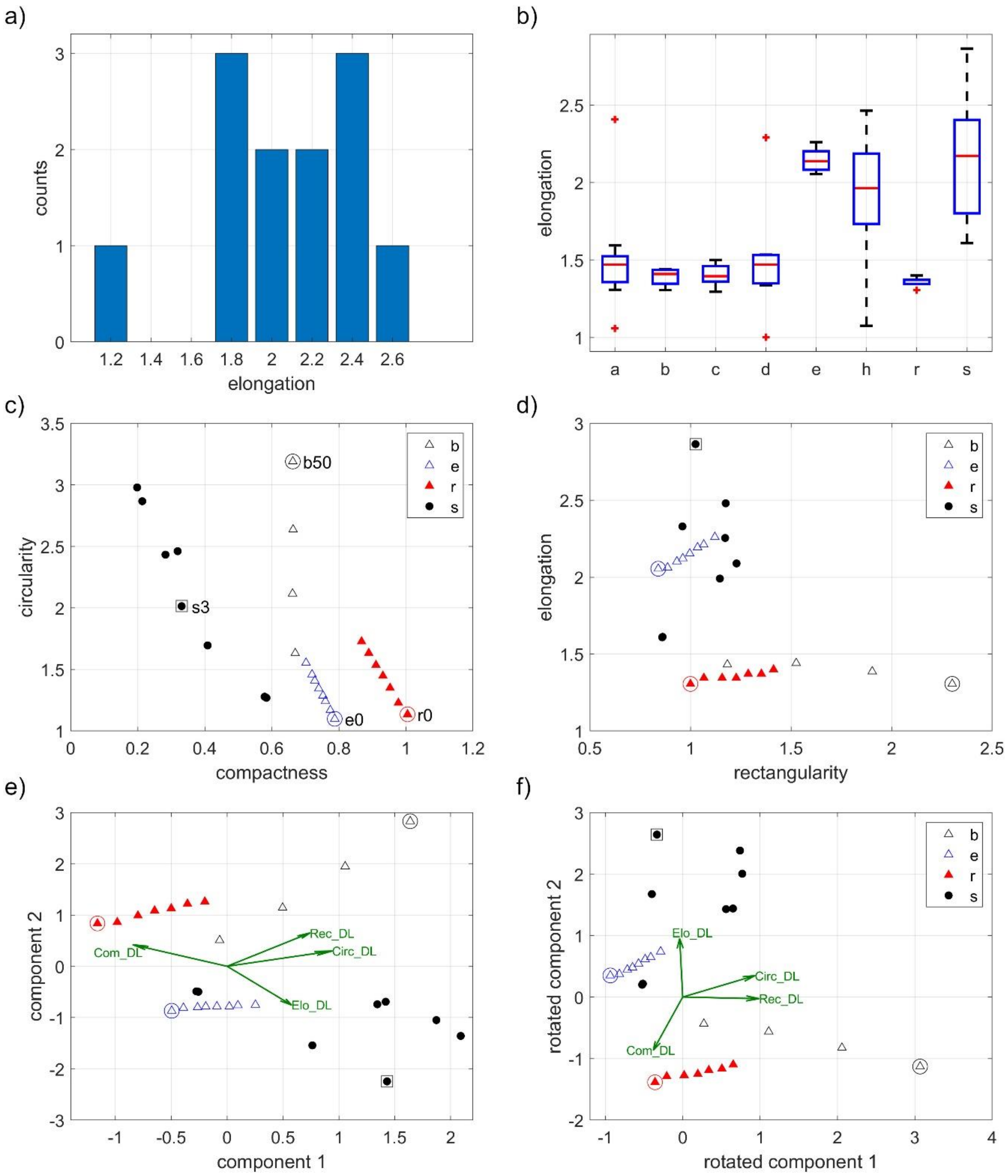


Fig. 3 Typical plots for presenting statistical results. a) Histogram showing the frequency distribution of elongation for data set "h". b) The data set-specific ranges of elongation are displayed as boxplots. The median is marked as a red central bar, and the 25th and 75th percentiles are indicated by the bottom and top edges of the blue boxes, respectively. The whiskers show the overall range covering all data points not considered outliers, while the latter are presented by red cross symbols. In c) and d) two examples of binary plots are shown. For four shape parameters, there are six combinations of binary plots. Data points for four objects are marked in the diagrams c) through f). e) With principal component analysis, the information of four shape parameters can be visualized in condensed form. The factor

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

loadings are illustrated by green vectors. They illustrate how the original variables influence the principal components. f) After “varimax” rotation, the data can be easier interpreted as products of the original shape parameters. While rotated component 1 is almost entirely depending on Circ_DL and Elo_DL, rotated component 2 is dominantly influenced by Com_DL and Rec_DL.

Figure 4

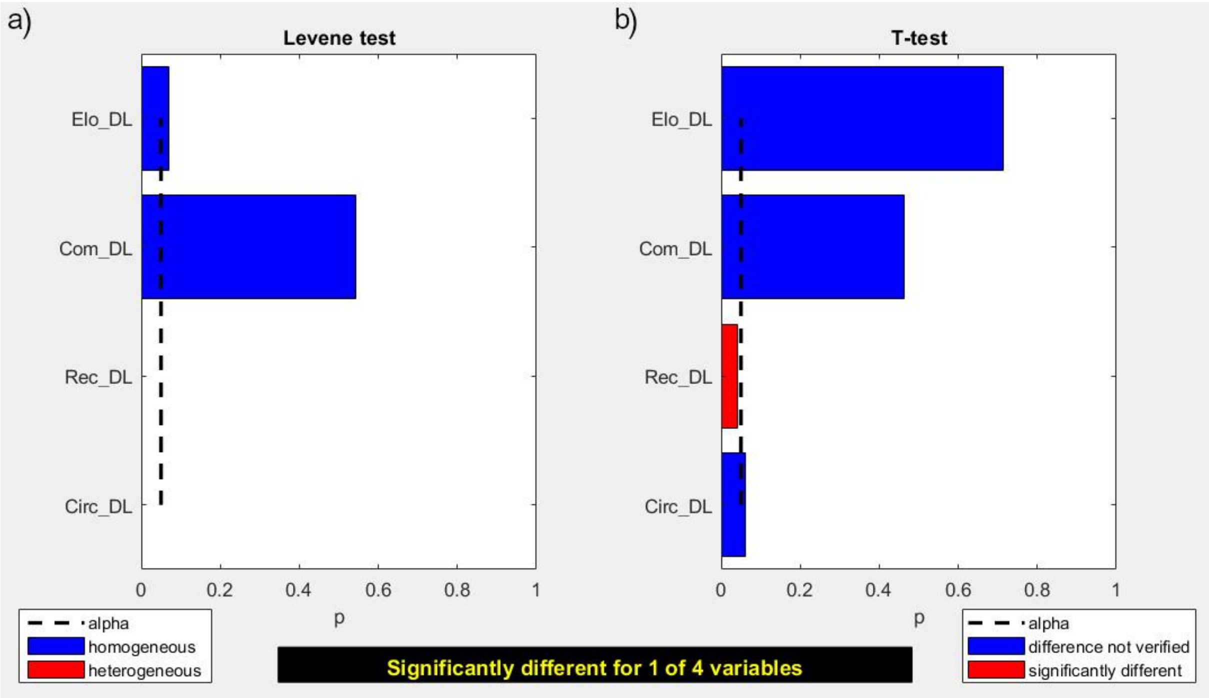


Fig. 4 Levene test (a) and two-tailed t-test (b) results plotted for data sets “a” and “c” in four shape parameters and a significance level α of 5%. The figure is a screenshot from DendroScan. a) For Levene tests, the null hypothesis H_0 is that the variances of the tested data sets are equal. If the *p*-value is smaller than α, then H_0 can be rejected. Here, the variances for the tested morphometric data sets are homogeneous in elongation (Elo_DL) and compactness (Com_DL), but heterogeneous for rectangularity (Rec_DL) and circularity (Circ_DL). b) The two-tailed t-tests works with the null hypothesis H_0 that the tested data are from the same population. Differences between the data sets are proven to be significant, if the *p*-value is smaller than α and H_0 is rejected. Here, the data sets “a” and “c” are significantly different in Rec_DL.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Figure 5

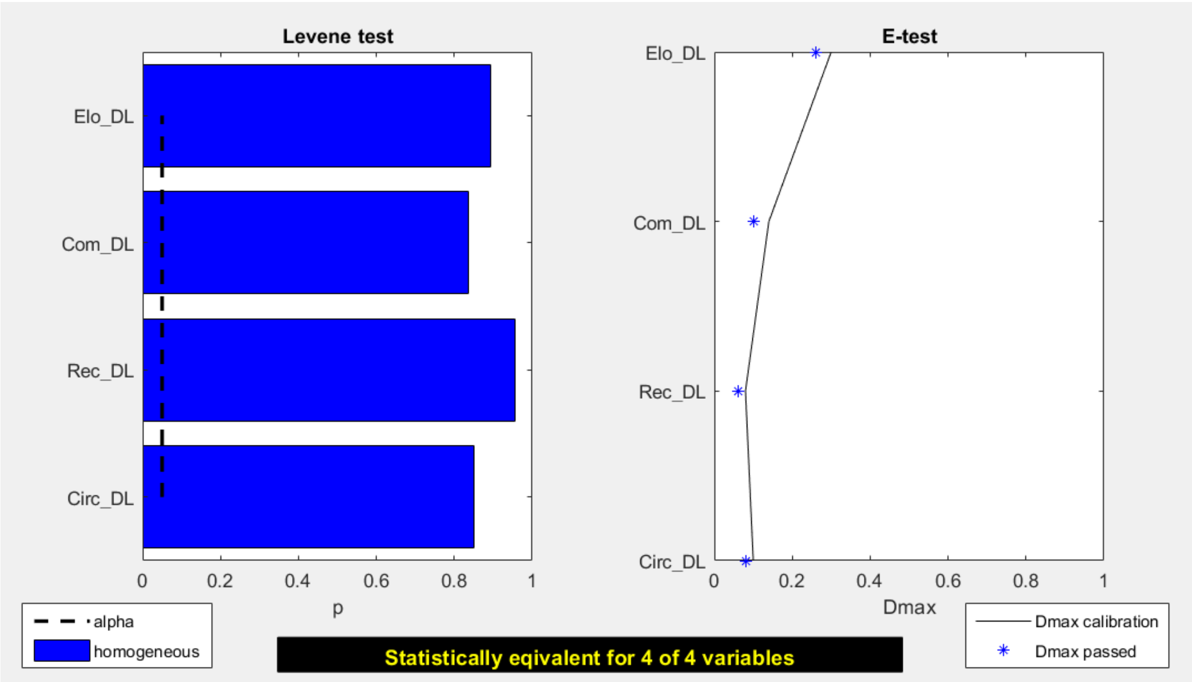


Fig. 5 *DendroScan* screenshot presenting results of e-tests. Here, the morphometric data sets “AA”, “AB” and “AC” (see Online Resource 1) are used as standards. The computed D_{max} values lie within the equivalence margins (indicated in the right diagram by a black line). Therefore, the e-tests show that the data sets “d” (Fig. 2) and “a” (Fig. 1) are statistically equivalent in all four tested shape parameters.

Figure 6

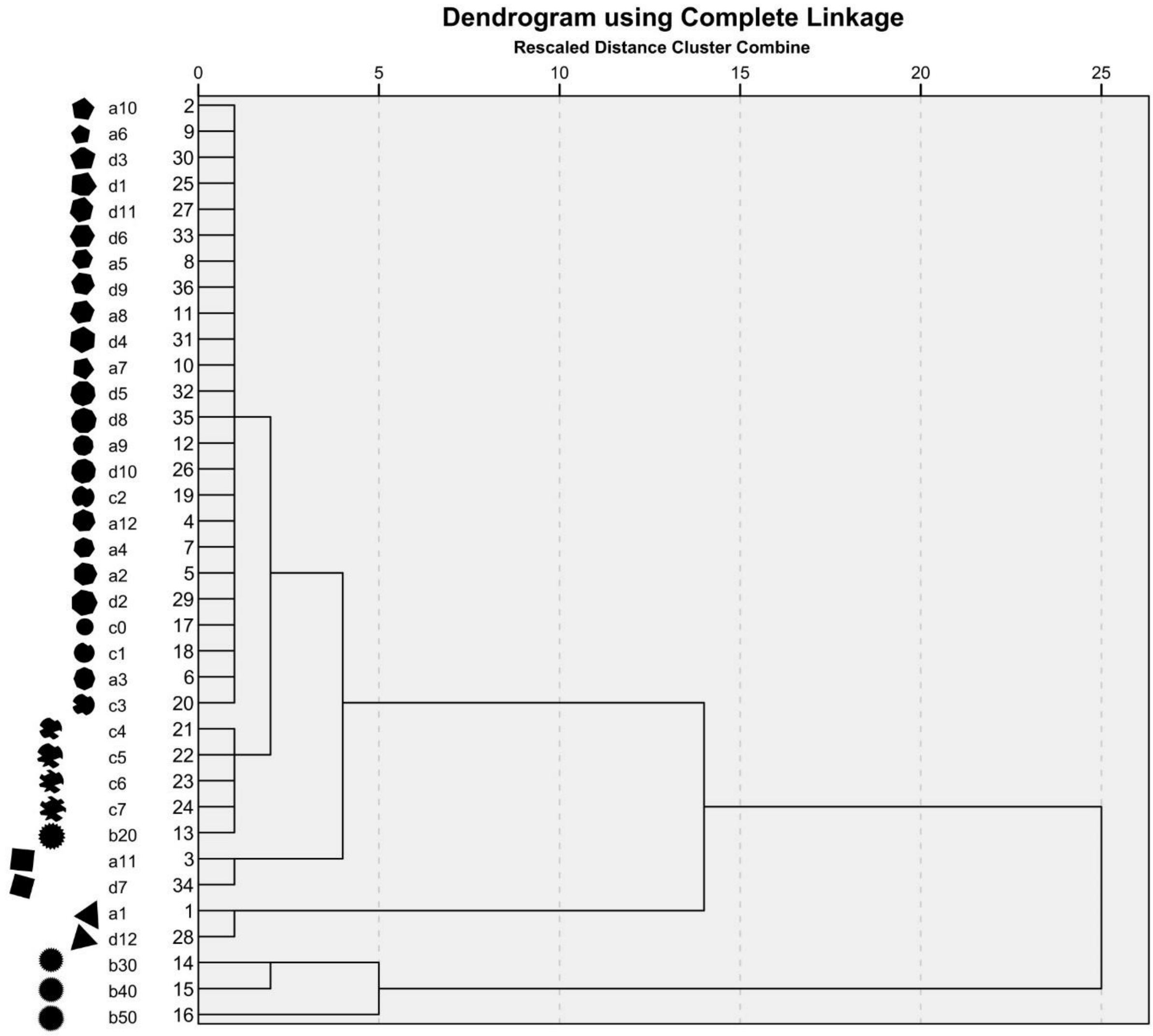


Fig. 6 Dendrogram illustrating the results of an agglomerative cluster analysis on the data sets “a, b, c, d”, conducted with the statistical program SPSS®. Squared Euclidean distance with complete linkage is used as measure of distance. Note that in this case the statistical equivalence of data sets “a” and “d” (which can be verified by e-tests), is not easy to identify by hierarchical cluster analysis.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Figure 7

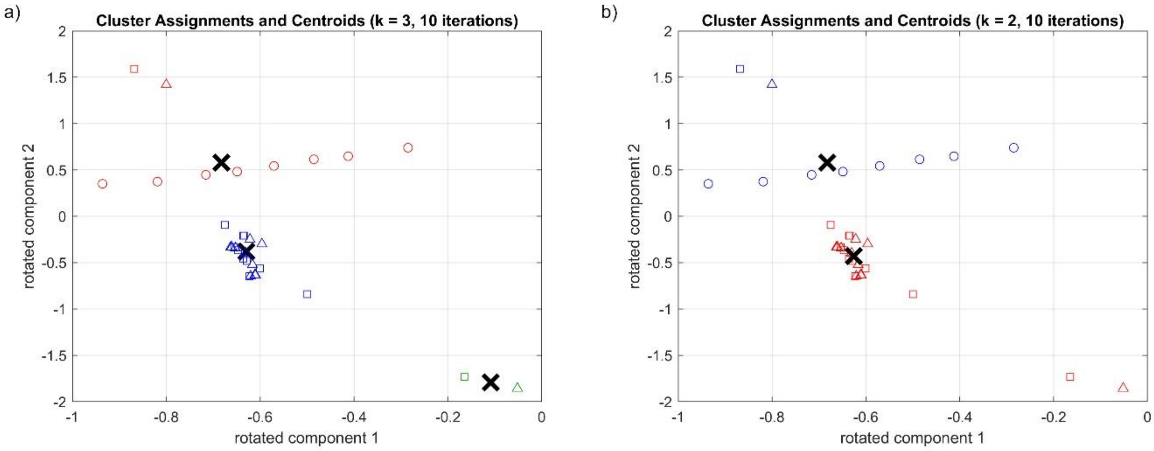


Fig. 7 Demonstration of k-means procedure, applied to data sets “a” (squares), “d” (triangles) and “e” (circles). Colour indicates membership of a certain cluster. The crosses represent the clusters’ centroids. a) For k = 3, most data points from “e” were assigned to cluster 1 (red), while the bulk of “a” and “d” was classified as cluster 2 (blue). Only two data points were assigned to cluster 3 (green). b) For k=2 , the bulk of data grouped together is the objects from “e” as cluster 1 (blue), whereas cluster 2 is exclusively composed of “a” and “d” objects.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Figure 8

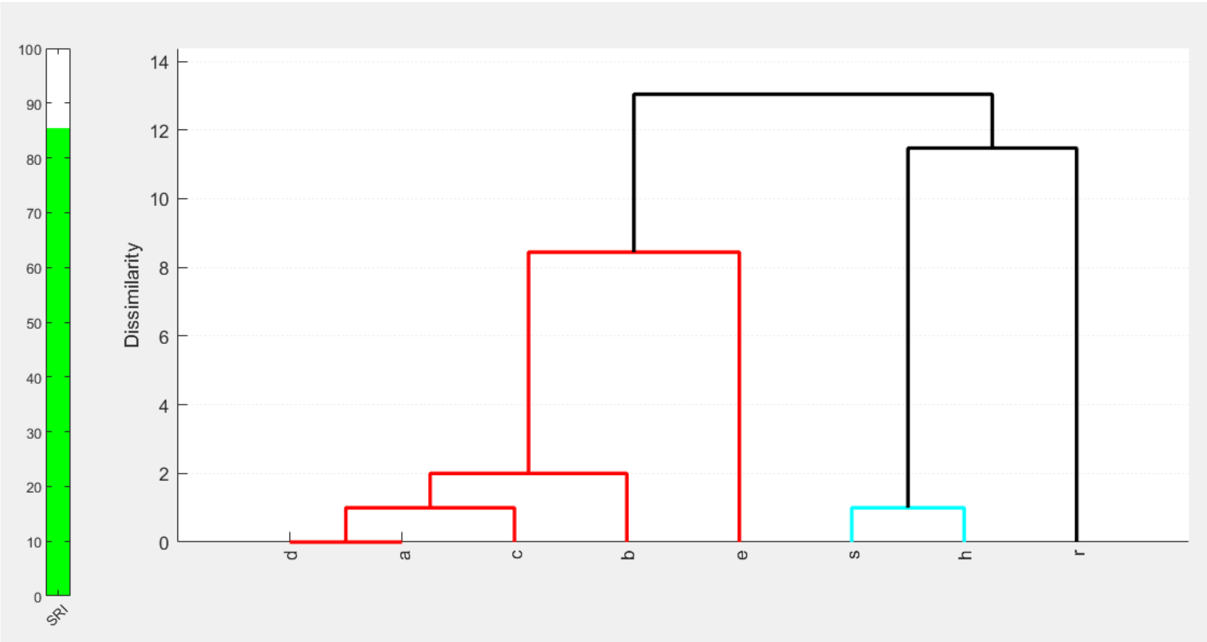


Fig. 8 Results of DAPM for the morphometric data sets “a”, “b” , “c”, “d”, “e”, “h”, “r”, “s” (see also Fig. 1 and Fig. 2), when using the four shape parameters Circ_DL, Rec_DL, Com_DL and Elo_DL and a level of significance of 5%. For computation of D_{max}, the data sets “AA”, “AB” and “AC” were used. The DAPM-based dendrogram shows data sets “a” and “d” to be equivalent and identifies 4 main clusters. Output was produced by the freeware *DendroScan*. The green bar on the left side indicates the “statistical reliability index” (SRI). With 85, this DAPM output can be seen as very reliable (Dürig et al. 2020a). The corresponding files with X-matrix and D_{max} values can be found in Online Resource 1.

Figure 9

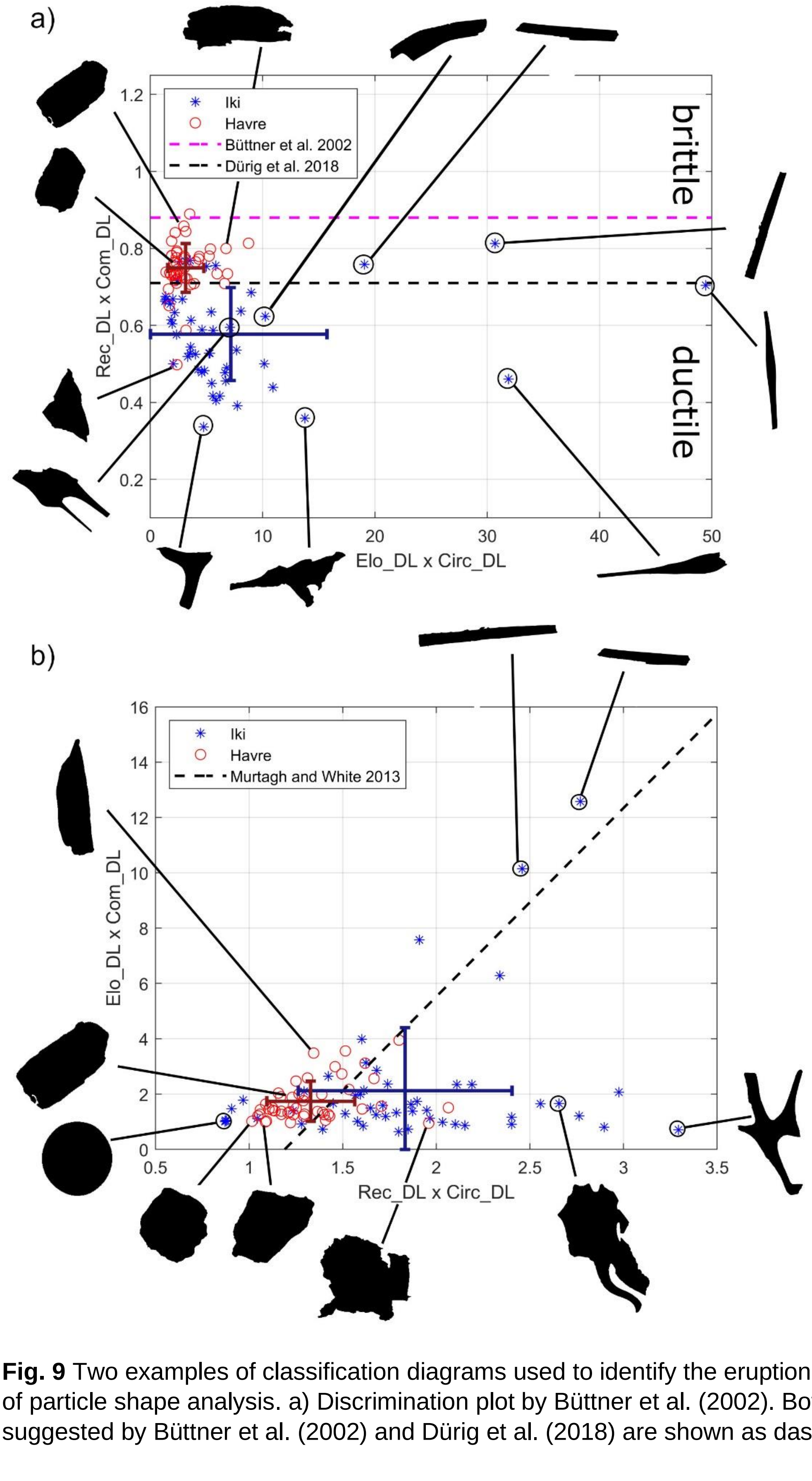


Fig. 9 Two examples of classification diagrams used to identify the eruption style by means of particle shape analysis. a) Discrimination plot by Büttner et al. (2002). Both thresholds suggested by Büttner et al. (2002) and Dürig et al. (2018) are shown as dashed horizontal

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

lines, dividing the plot in an upper and a lower sector, respectively. Data points in the upper sector indicate that particles have been generated by brittle fragmentation, while particles produced by ductile fragmentation mechanisms are characterized by shape parameters that fall in the lower sector. b) The diagram suggested by Murtagh and White (2013) uses the dashed line as threshold to discriminate between phreatomagmatic (left side) and magmatic grain shapes (right side). In both diagrams, the standard deviations for both populations are indicated by error bars, with centres indicating their mean values.

This is a post-peer-review, pre-copyedit version of an article published in Bulletin of Volcanology. The final authenticated version is available online at: https://doi.org/10.1007/s00445-021-01500-0

Acknowledgements

Louise Steffensen Schmidt is thanked for reading an earlier version of the manuscript. We thank Adrian Hornby and an anonymous reviewer, as well as associate editor Benjamin Andrews, for their constructive comments that helped improving the manuscript. Furthermore, we gratefully acknowledge Sebastian Mueller, Bruce Houghton and Wendy Cockshell for providing the Kīlauea Iki 1959 samples. TD is supported by the Icelandic Research Fund (Rannís), grant Nr. 206527-051.

Online Resources

Online Resource 1 (.rar):

Silhouettes of “artificial” objects and data sets with morphometric parameters used for demonstration.

Online Resource 2 (.rar):

Silhouettes of “natural” ash silhouettes from the 1959 Kīlauea Iki and the 2012 Havre eruption, as well as data sets with their morphometric parameters.

Online Resource 3 (.xlsx):

Table with principal components for all demonstration silhouettes, before and after “varimax” rotation.

Online Resource 4 (.xlsx):

Tables showing the results from k-means procedure, applied to data sets “a”, “d” and “e”, using 10 iterations with k=3 and k=2.

Declarations

Funding

TD was supported by the IRF (Rannís) PostDoctoral project grant 206527-051

Conflicts of interest/Competing interests

There are no competing interests.

Availability of data and material

All example data presented are provided in Online Resource 1.

Code availability

Not applicable.

References

Ahad NA, Yahaya SSS (2014) Sensitivity analysis of Welch's t-test. In: AIP Conference Proceedings. pp 888–893

Alvarado GE, Mele D, Dellino P, et al (2016) Are the ashes from the latest eruptions (2010–2016) at Turrialba volcano (Costa Rica) related to phreatic or phreatomagmatic events? J Volcanol Geotherm Res 327:407–415. https://doi.org/10.1016/j.jvolgeores.2016.09.003

Avery MR, Panter KS, Gorsevski P V, et al (2017) Distinguishing styles of explosive eruptions at Erebus, Redoubt and Taupo volcanoes using multivariate analysis of ash morphometrics. J Volcanol Geotherm Res 332:1–13. https://doi.org/10.1016/j.jvolgeores.2017.01.010

Barrett PJ (1980) The shape of rock particles, a critical review. Sedimentology. https://doi.org/10.1111/j.1365-3091.1980.tb01179.x

Bender R, Lange S (2001) Adjusting for multiple testing—when and how? J Clin Epidemiol 54:343–349. https://doi.org/10.1016/S0895-4356(00)00314-0

Blanca MJ, Alarcón R, Arnau J, et al (2017) Non-normal data: Is ANOVA still a valid option? Psicothema 29:552–557. https://doi.org/10.7334/psicothema2016.383

Bonferroni CE (1936) Teoria statistica delle classi e calcolo delle probabilità. Pubbl del R Ist Super di Sci Econ e Commer di Firenze 8:3–62

Breiman L (2001) Random Forrests. Mach Learn

Büttner R, Dellino P, La Volpe L, et al (2002) Thermohydraulic explosions in phreatomagmatic eruptions as evidenced by the comparison between pyroclasts and products from Molten Fuel Coolant Interaction experiments. J Geophys Res Solid Earth 107:2277. https://doi.org/10.1029/2001JB000511

Büttner R, Dellino P, Zimanowski B (1999) Identifying magma-water interaction from the surface features of ash particles. Nature; London 401:688–690. https://doi.org/http://dx.doi.org.ezproxy.otago.ac.nz/10.1038/44364

Carey R, Soule SA, Manga M, et al (2018) The largest deep-ocean silicic volcanic eruption of the past century. Sci Adv 4:e1701121. https://doi.org/10.1126/sciadv.1701121

Chávez GM, Castillo-Rivera F, Montenegro-Ríos JA, et al (2020) Fourier Shape Analysis, FSA: Freeware for quantitative study of particle morphology. J Volcanol Geotherm Res 404:107008. https://doi.org/10.1016/j.jvolgeores.2020.107008

Cioni R, D'Oriano C, Bertagnini A (2008) Fingerprinting ash deposits of small scale eruptions by their physical and textural features. J Volcanol Geotherm Res 177:277–287. https://doi.org/10.1016/j.jvolgeores.2008.06.003

Cioni R, Pistolesi M, Bertagnini A, et al (2014) Insights into the dynamics and evolution of the 2010 Eyjafjallajökull summit eruption (Iceland) provided by volcanic ash textures. Earth Planet Sci Lett 394:111–123. https://doi.org/10.1016/j.epsl.2014.02.051

Coltelli M, Miraglia L, Scollo S (2008) Characterization of shape and terminal velocity of tephra particles erupted during the 2002 eruption of Etna volcano, Italy. Bull Volcanol 70:1103–1112. https://doi.org/10.1007/s00445-007-0192-8

Comida PP, Ross P-S, Dürig T, et al (2021) Standardized data acquisition for morphological and textural studies of juvenile pyroclasts in primary fragmentation studies; 2. Choice of size fraction and method optimization for particle cross-sections. Bull Volcanol submitted:

Davis JC (2002) Statistics and Data Analysis in Geology, 3rd edition, 3rd edn. John Wiley & Sons, New York; Chichester; Brisbane

Dellino P, La Volpe L (1996) Image processing analysis in reconstructing fragmentation and transportation mechanisms of pyroclastic deposits. The case of Monte Pilato-Rocche Rosse eruptions, Lipari (Aeolian islands, Italy). J Volcanol Geotherm Res 71:13–29. https://doi.org/10.1016/0377-0273(95)00062-3

Dellino P, La Volpe L, Isaia R, Orsi G (2001) Statistical analysis of textural data from complex pyroclastic sequences: implications for fragmentation processes of the Agnano-Monte Spina Tephra (4.1 ka), Phlegraean Fields, southern Italy. Bull Volcanol 63:443–461. https://doi.org/10.1007/s004450100163

Dellino P, Liotino G (2002) The fractal and multifractal dimension of volcanic ash particles contour: a test study on the utility and volcanological relevance. J Volcanol Geotherm Res 113:1–18

Dioguardi F, Mele D, Dellino P, Dürig T (2017) The terminal velocity of volcanic particles with shape obtained from 3D X-ray microtomography. J Volcanol Geotherm Res 329:. https://doi.org/10.1016/j.jvolgeores.2016.11.013

Donaldson TS (1968) Robustness of the F-Test to Errors of Both Kinds and the Correlation Between the Numerator and Denominator of the F-Ratio. J Am Stat Assoc. https://doi.org/10.2307/2284037

Dürig T, Bowman M, White J, et al (2018) PARTIcle Shape ANalyzer PARTISAN –

an open source tool for multi-standard two-dimensional particle morphometry analysis. Ann Geophys 61:VO671. https://doi.org/10.4401/ag-7865

Dürig T, Mele D, Dellino P, Zimanowski B (2012a) Comparative analyses of glass fragments from brittle fracture experiments and volcanic ash particles. Bull Volcanol 74:691–704. https://doi.org/10.1007/s00445-011-0562-0

Dürig T, Schmidt LS, White JDL, Bowman MH (2020a) DendroScan: an open source tool to conduct comparative statistical tests and dendrogrammatic analyses on particle morphometry. Sci Rep 10:21682. https://doi.org/10.1038/s41598-020-78698-0

Dürig T, Sonder I, Zimanowski B, et al (2012b) Generation of volcanic ash by basaltic volcanism. J Geophys Res Solid Earth 117:B01204. https://doi.org/10.1029/2011JB008628

Dürig T, White JDL, Murch AP, et al (2020b) Deep-sea eruptions boosted by induced fuel–coolant explosions. Nat Geosci 13:498–503. https://doi.org/10.1038/s41561-020-0603-4

Dürig T, White JDL, Zimanowski B, et al (2020c) Deep-sea fragmentation style of Havre revealed by dendrogrammatic analyses of particle morphometry. Bull Volcanol 82:67. https://doi.org/10.1007/s00445-020-01408-1

Dürig T, Zimanowski B (2012) “Breaking news” on the formation of volcanic ash: Fracture dynamics in silicate glass. Earth Planet Sci Lett 335:1–8. https://doi.org/10.1016/j.epsl.2012.05.001

Ersoy O, Chinga G, Aydar E, et al (2006) Texture discrimination of volcanic ashes from different fragmentation mechanisms: A case study, Mount Nemrut stratovolcano, eastern Turkey. Comput Geosci 32:936–946. https://doi.org/10.1016/j.cageo.2005.10.013

Games PA, Keselman HJ, Clinch JJ (1979) Tests for homogeneity of variance in factorial designs. Psychol Bull 86:978–984. https://doi.org/10.1037/0033-2909.86.5.978

Havlicek LL, Peterson NL (1974) Robustness of the T Test: A Guide for Researchers on Effect of Violations of Assumptions. Psychol Rep. https://doi.org/10.2466/pr0.1974.34.3c.1095

Heiken G (1974) Atlas of Volcanic Ash. Smithson Contrib to Earth Sci 1–101. https://doi.org/10.5479/si.00810274.12.1

Huberty CJ, Hussein MH (2003) Some Problems in Reporting Use of Discriminant Analyses. J Exp Educ 71:177–192. https://doi.org/10.1080/00220970309602062

Iverson NA, Kyle PR, Dunbar NW, et al (2014) Eruptive history and magmatic stability of Erebus volcano, Antarctica: Insights from englacial tephra. Geochemistry, Geophys Geosystems 15:4180–4202. https://doi.org/10.1002/2014GC005435

Jordan SC, Dürig T, Cas RAF, Zimanowski B (2014) Processes controlling the shape of ash particles: Results of statistical IPA. J Volcanol Geotherm Res 288:19–27. https://doi.org/10.1016/j.jvolgeores.2014.09.012

Kaiser HF (1958) The varimax criterion for analytic rotation in factor analysis.

Psychometrika 23:187–200. https://doi.org/10.1007/BF02289233

Lawn B (1993) Fracture of Brittle Solids. Cambridge University Press

Leibrandt S, Le Pennec J-L (2015) Towards fast and routine analyses of volcanic ash morphometry for eruption surveillance applications. J Volcanol Geotherm Res 297:11–27. https://doi.org/10.1016/j.jvolgeores.2015.03.014

Levene H (1960) Robust tests for equality of variances. In: Olkin I, Ghurye SG, Hoeffding W, et al. (eds) Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press, Menlo Park, CA, pp 278–292

Liu EJ, Cashman K V, Rust AC (2015) Optimising shape analysis to quantify volcanic ash morphology. GeoResJ 8:14–30. https://doi.org/10.1016/j.grj.2015.09.001

Loh WY (2011) Classification and regression trees. Wiley Interdiscip Rev Data Min Knowl Discov 1:14–23. https://doi.org/10.1002/widm.8

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability. pp 281–297

Maria A, Carey S (2002) Using fractal analysis to quantitatively characterize the shapes of volcanic particles. J Geophys Res Solid Earth 107:ECV 7-1-ECV 7-17. https://doi.org/10.1029/2001JB000822

Mele D, Costa A, Dellino P, et al (2020) Total grain size distribution of components of fallout deposits and implications for magma fragmentation mechanisms: examples from Campi Flegrei caldera (Italy). Bull Volcanol 82:31. https://doi.org/10.1007/s00445-020-1368-8

Mele D, Dellino P, Sulpizio R, Braia G (2011) A systematic investigation on the aerodynamics of ash particles. J Volcanol Geotherm Res 203:1–11. https://doi.org/10.1016/j.jvolgeores.2011.04.004

Mueller SB, Houghton BF, Swanson DA, et al (2018) Intricate episodic growth of a Hawaiian tephra deposit: case study of the 1959 Kīlauea Iki eruption. Bull Volcanol 80:73. https://doi.org/10.1007/s00445-018-1249-6

Mueller SB, Houghton BF, Swanson DA, et al (2019) Total grain size distribution of an intense Hawaiian fountaining event: case study of the 1959 Kīlauea Iki eruption. Bull Volcanol 81:43. https://doi.org/10.1007/s00445-019-1304-y

Murtagh RM, White JDL (2013) Pyroclast characteristics of a subaqueous to emergent Surtseyan eruption, Black Point volcano, California. J Volcanol Geotherm Res 267:75–91. https://doi.org/10.1016/j.jvolgeores.2013.08.015

Németh K, Cronin SJ (2011) Drivers of explosivity and elevated hazard in basaltic fissure eruptions: the 1913 eruption of Ambrym Volcano, Vanuatu (SW-Pacific). J Volcanol Geotherm Res 201:194–209

Németh K, Kósik S (2020) The role of hydrovolcanism in the formation of the Cenozoic monogenetic volcanic fields of Zealandia. New Zeal J Geol Geophys 1–26. https://doi.org/10.1080/00288306.2020.1770304

Nurfiani D, de Maisonneuve CB (2018) Furthering the investigation of eruption styles through quantitative shape analyses of volcanic ash particles. J Volcanol

Geotherm Res 354:102–114

Pardo N, Avellaneda JD, Rausch J, et al (2020) Decrypting silicic magma/plug fragmentation at Azufral crater lake, Northern Andes: insights from fine to extremely fine ash morpho-chemistry. Bull Volcanol 82:79. https://doi.org/10.1007/s00445-020-01418-z

Perneger T V (1998) What's wrong with Bonferroni adjustments. BMJ 316:1236–1238. https://doi.org/10.1136/bmj.316.7139.1236

Rasch D, Guiard V (2004) The robustness of parametric statistical methods. Psychol Sci 46:175–208

Rausch J, Grobéty B, Vonlanthen P (2015) Eifel maars: Quantitative shape characterization of juvenile ash particles (Eifel Volcanic Field, Germany). J Volcanol Geotherm Res 291:86–100. https://doi.org/10.1016/j.jvolgeores.2014.11.008

Richter DH, Eaton JP, Murata KJ, et al (1970) Chronological narrative of the 1959-60 eruption of Kilauea Volcano, Hawaii. US Geol Surv Prof Pap 537-E E:1–73

Ross P-S, Dürig T, Comida PP, et al (2021) Standardized data acquisition for morphological and textural studies of juvenile pyroclasts in primary fragmentation studies; 1. Overview and workflow. Bull Volcanol submiitted:

Rustam Z, Hartini S, Pratama RY, et al (2020) Analysis of architecture combining Convolutional Neural Network (CNN) and kernel K-means clustering for lung cancer diagnosis. Int J Adv Sci Eng Inf Technol. https://doi.org/10.18517/ijaseit.10.3.12113

Satorra A, Bentler PM (1994) Corrections to test statistics and standard errors in covariance structure analysis. In: Eye A, Clogg CC (eds) Latent variables analysis: Applications for developmental research. Sage Publications Inc., Thousand Oaks, CA.

Scasso RA, Carey S (2005) Morphology and formation of glassy volcanic ash from the August 12-15, 1991 eruption of Hudson volcano, Chile. Lat Am J Sedimentol Basin Anal 12:3–21

Schipper CI, Sonder I, Schmid A, et al (2013) Vapour dynamics during magma–water interaction experiments: hydromagmatic origins of submarine volcaniclastic particles (limu o Pele). Geophys J Int 192:1109–1115. https://doi.org/10.1093/gji/ggs099

Schmith J, Höskuldsson Á, Holm PM (2017) Grain shape of basaltic ash populations: implications for fragmentation. Bull Volcanol 79:14. https://doi.org/10.1007/s00445-016-1093-5

Shoji D, Noguchi R, Otsuki S, Hino H (2018) Classification of volcanic ash particles using a convolutional neural network and probability. Sci Rep. https://doi.org/10.1038/s41598-018-26200-2

Student (1908) The Probable Error of a Mean. Biometrika 6:1–25. https://doi.org/10.2307/2331554

Suzuki K, Fujiwara H, Ohta T (2015) The evaluation of macroscopic and microscopic textures of sand grains using elliptic Fourier and principal component analysis:

Implications for the discrimination of sedimentary environments. Sedimentology 62:1184–1197. https://doi.org/10.1111/sed.12183

Taddeucci J, Cimarelli C, Alatorre-Ibargüengoitia MA, et al (2021) Fracturing and healing of basaltic magmas during explosive volcanic eruptions. Nat Geosci 14:248–254. https://doi.org/10.1038/s41561-021-00708-1

Taddeucci J, Pompilio M, Scarlato P (2002) Monitoring the explosive activity of the July-August 2001 eruption of Mt. Etna (Italy) by ash characterization. Geophys Res Lett 29:71–74. https://doi.org/10.1029/2001GL014372

Tang JL, Wang D, Zhang ZG, et al (2017) Weed identification based on K-means feature learning combined with convolutional neural network. Comput Electron Agric 135:63–70. https://doi.org/10.1016/j.compag.2017.01.001

Tukey JW (1949) Comparing individual means in the analysis of variance. Biometrics 5:99–114

Tunwal M, Mulchrone KF, Meere PA (2020) A new approach to particle shape quantification using the curvature plot. Powder Technol 374:377–388. https://doi.org/10.1016/j.powtec.2020.07.045

Tunwal M, Mulchrone KF, Meere PA (2018) Quantitative characterization of grain shape: Implications for textural maturity analysis and discrimination between depositional environments. Sedimentology 65:1761–1776. https://doi.org/10.1111/sed.12445

Verolino A, White JDL, Dürig T, Cappuccio F (2019) Black Point – Pyroclasts of a Surtseyan eruption show no change during edifice growth to the surface from 100 m water depth. J Volcanol Geotherm Res 384:85–102. https://doi.org/10.1016/j.jvolgeores.2019.07.013

Vonlanthen P, Rausch J, Ketcham RA, et al (2015) High-resolution 3D analyses of the shape and internal constituents of small volcanic ash particles: The contribution of SEM micro-computed tomography (SEM micro-CT). J Volcanol Geotherm Res 293:1–12. https://doi.org/10.1016/j.jvolgeores.2014.11.016

Walker E, Nowacki AS (2011) Understanding equivalence and noninferiority testing. J Gen Intern Med 26:192–196. https://doi.org/10.1007/s11606-010-1513-8

Welch BL (1947) The Generalization of `Student’s’ Problem when Several Different Population Variances are Involved. Biometrika 34:28–35. https://doi.org/10.2307/2332510

Wellek S (2010) Testing Statistical Hypotheses of Equivalence and Noninferiority. Chapman and Hall/CRC

White JDL, Valentine GA (2016) Magmatic versus phreatomagmatic fragmentation: Absence of evidence is not evidence of absence. Geosphere 12:1478–1488. https://doi.org/10.1130/GES01337.1

Wohletz KH (1983) Mechanisms of hydrovolcanic pyroclast formation: Grain-size, scanning electron microscopy, and experimental studies. J Volcanol Geotherm Res 17:31–63. https://doi.org/10.1016/0377-0273(83)90061-6