DendroScan

User manual

v.1.0

22/09/2020

**DendroScan v1.0**

This program is free software: you can redistribute them and/or modify them under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

This software package is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with REFIR. If not, see <https://www.gnu.org/licenses/>.

If you wish to contribute to the development of DendroScan or to report bugs or other problems with the software, please contact me per email (tobi@hi.is).

Inhalt

[1. Introduction 4](#_Toc33648404)

[1.1 General information 4](#_Toc33648405)

[1.2 Content 4](#_Toc33648406)

[1.3 System requirements 5](#_Toc33648407)

[1.4 Installation and start 5](#_Toc33648408)

[2 Dendrogram mode 6](#_Toc33648409)

[2.1 When to use 6](#_Toc33648410)

[2.2 Theoretical background 6](#_Toc33648411)

[2.3 Steps to compute a dendrogram 7](#_Toc33648412)

[2.3.1 Initiation and file selection 7](#_Toc33648413)

[2.3.2 Selection of shape parameters 8](#_Toc33648414)

[2.3.3 Constructing dendrogram 9](#_Toc33648415)

[2.3.4 Change alpha 9](#_Toc33648416)

[2.3.5 The SPI (statistical power index) indicator 9](#_Toc33648417)

[2.3.6 Dendrogram display 10](#_Toc33648418)

[2.3.7 Export results 10](#_Toc33648419)

[2.3.8 Clear window and restart program 10](#_Toc33648420)

[3 “T-test” mode 11](#_Toc33648421)

[3.1 When to use 11](#_Toc33648422)

[3.2 Theoretical background of a two-tailed t-test 11](#_Toc33648423)

[3.3 Steps to compute two-tailed t-tests 11](#_Toc33648424)

[3.3.1 Initiation and file selection 11](#_Toc33648425)

[3.3.2 Selection of shape parameters 12](#_Toc33648426)

[3.3.3 Performing two-tailed t-tests 13](#_Toc33648427)

[3.3.4 Change alpha 13](#_Toc33648428)

[3.3.5 Interpretation of Levene-test plots 13](#_Toc33648429)

[3.3.6 Interpretation of t-test results 13](#_Toc33648430)

[3.3.7 Export t-test results 14](#_Toc33648431)

[3.3.8 Clear window and restart program 14](#_Toc33648432)

[4 “E-test” mode 14](#_Toc33648433)

[4.1 When to use 14](#_Toc33648434)

[4.2 Theoretical background of equivalence tests 14](#_Toc33648435)

[4.3 Steps to conduct e-tests 15](#_Toc33648436)

[4.3.1 Initiation and file selection 15](#_Toc33648437)

[4.3.2 Selection of shape parameters 16](#_Toc33648438)

[4.3.3 Conducting e-tests 16](#_Toc33648439)

[4.3.4 Change alpha 17](#_Toc33648440)

[4.3.5 Interpretation of Levene-test plots 17](#_Toc33648441)

[4.3.6 Interpretation of e-test results 17](#_Toc33648442)

[4.3.7 Export e-test output 18](#_Toc33648443)

[4.3.8 Clear window and restart program 18](#_Toc33648444)

[Appendix 19](#_Toc33648445)

[References 21](#_Toc33648446)

# 1. Introduction

## 1.1 General information

In Volcanology, studies on tephra often compare the morphology of ash particles between different eruptions or eruptive stages to reconstruct eruptive styles or the responsible fragmentation mechanisms (e.g., Heiken 1972, 1974; Heiken and Wohletz 1985). For this purpose, a multitude of morphometric systems have been developed which quantify the particle shapes in 2D via nondimensional shape parameters. Five of these morphometric systems are used in the free particle shape analysis software PARTISAN[[1]](#footnote-2) (Dürig et al. 2018), which processes binarized images of particles (e.g., particle silhouettes by scanning electron microscope images) and computes 23 non-dimensional shape parameters (see Table 1). The results of each run are then stored in a .csv file.

**DendroScan** is tailored for post-processing the PARTISAN output files in order to assist researchers with statistical analyses of multiple multi-dimensional morphometric data sets.

**DendroScan** provides all statistical tools needed for conducting a dendrogrammatic analysis of particle morphometry (DAPM), as introduced by Dürig et al. (Dürig et al. 2020b). In particular, the program allows the user to:

* construct and display dendrograms, which sort multi-dimensional data sets by morphometric similarities based on one-way analysis of variances (ANOVA).
* test two data sets for significant differences by conducting two-tailed t-tests (Student 1908; Welch 1947).
* verify whether two data sets are statistically equivalent by means of e(quivalence)-tests (Dürig et al. 2012, 2020b)

## 1.2 Content

DendroScan is a GNU licensed Matlab program, which consists of a main script (“main.m”) and a number of auxillary scripts in a folder labelled “functions”. These scripts include:

* “cdfTukey.m”
* “dendrogram\_funct.m”
* “etest\_main.m”
* “games\_howell.m”
* “linkageold.m”
* “ttest\_main.m”
* “vartestn\_new.m”

Peter Nagy and Pierre Mégevand are acknowledged for providing “cdfTukey.m” (Nagy 2020) and “games\_howell.m” (Mégevand 2020).

## 1.3 System requirements

**DendroScan** was developed and tested with Matlab R2019b on both Linux (debian) and Windows operation systems. **DendroScan** can be run with Matlab in default installation and does not require any additional Matlab libraries.

## 1.4 Installation and start

To run **DendroScan**, simply unzip the software package in a folder of your choice. The program is started by opening “main.m” with Matlab and executing it.

The GUI opens and the user is offered a test selection field in the upper left corner of the GUI (see Fig.1).

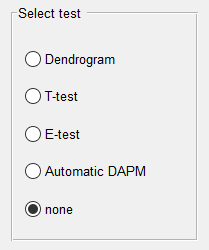


Fig.1 Test selection field

By default, the radio button “none” is active. The user can choose between four statistical test options. Three of them offer manual tests, whereas the fourth option is to conduct an automatic DAPM (dendrogrammatic analysis of particle morphometry).

# 2 Dendrogram mode

## 2.1 When to use

Select this test mode to statistically compare more than two data sets and sort them by conducting a dendrogram analysis, as introduced by Dürig et al. (Dürig et al. 2020b).

## 2.2 Theoretical background

The dendrogram analysis is conducted by following a sequence of steps:

1. For each shape parameter, Levene-tests (Levene 1960) are performed in order to check the homogeneity of variances of the data sets.
2. Then a one-way analysis of variances (ANOVA) is conducted for all shape parameters.  
   Depending on the outcome of the Levene-tests, subsequently either the Tukey’s range test are applied (also known as Tukey honestly significant difference HSD) as post-hoc correction for assumed homogeneous variances (Tukey 1949). Alternatively, a Games-Howell post-hoc adjustment (Games et al. 1979) is applied for heterogeneous variances.

The ANOVA provides the error likelihoods (“*p*-values”) of the null hypothesis, according to which the tested data sets are from the same population. If the error likelihood *p* is below the level of significance *α*, the null hypothesis can be rejected: the data sets are then verified to be “significantly different” in the tested hypothesis.

1. The distance matrix *X* is constructed for all selected *z* shape parameters and *N* tested samples, with *pijk* being the *p*-value of data set *i* tested with the one from data set *j* in the *k*-th shape parameter:

(1)

with *Yijk* being defined as:

(2)

1. In the last step dendrograms are plotted, using *X* as distance matrices. These plots group the tested data sets according to their morphometric differences, with the y-axis reflecting their “dissimilarity”.

Based on the dendrogram, data sets of high morphometric similarities are grouped in clusters.

Importantly, there is a negative correlation between the number of data sets analysed *N*, and the statistical power. The larger *N*, the lower the capability of the dendrogram analysis to detect genuine differences between their shape parameters. In these cases it might be necessary to repeat a dendrogram analysis with using fewer data sets, e.g. by focusing separately on individual clusters of supposedly “similar” data sets (Dürig et al. 2020b).

In order to provide quality guidance, **DendroScan** estimates a statistical power index, labelled “SPI”. The SPI values range between 0 and 100, and are computed by:

(3)

where *f* is the probability density function:

(4)

with *N* being the number of data sets.   
The adjustment parameters *µ* and *s* are set to 2.4 and 4.5. These values were selected based on empirical findings, which have shown that for N > 7 the reliability of the dendrogram significantly decreased. The relationship between *N* and the SPI is illustrated in Fig. 2.

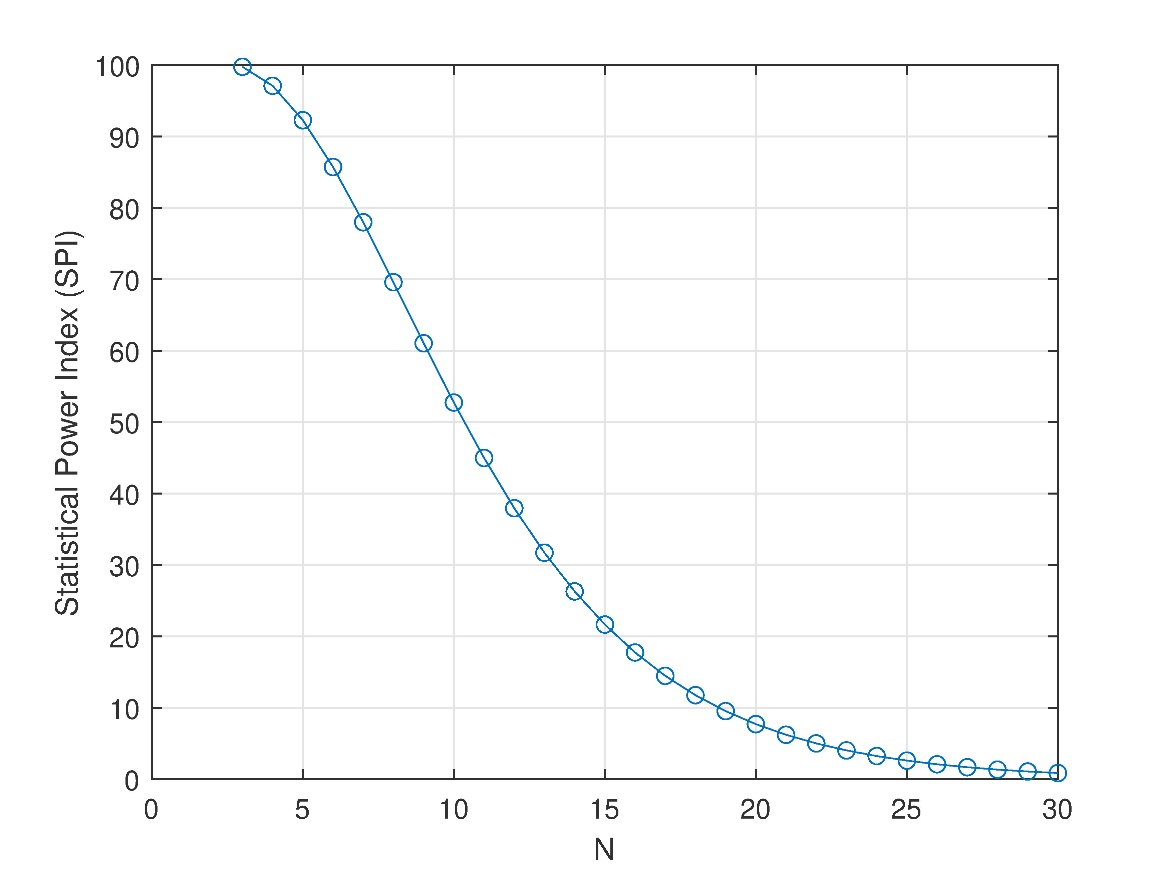


Fig. 2 Relationship between SPI and number of data sets. A high SPI reflects a high reliability of the ANOVA-based dendrogram results.

## 2.3 Steps to compute a dendrogram

### 2.3.1 Initiation and file selection

The dendrogram mode is activated by clicking on the radio button “dendrogram” (see Fig.1). Under the test selection field, a file selection field appears with two buttons (see Fig.3).

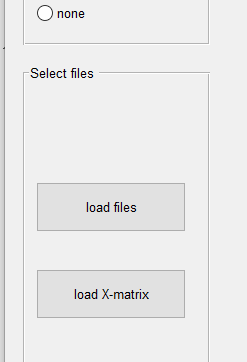
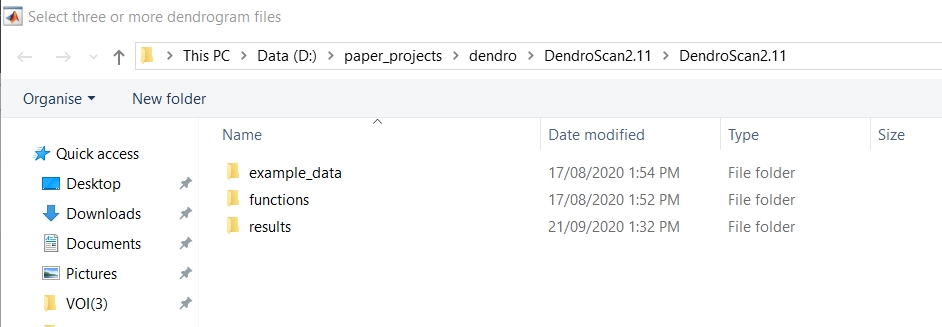
 

Fig.3 File selection field in “dendrogram” mode (left). Clicking on either button opens a file browser window (right).

* **“load files”**: click on this button to pick all files which you want to be included in the dendrogram. A file browser window is opened. Select the correct path of the .csv input data files you want to be analysed by clicking on each file and pressing the control-key.
* **“load X-matrix”**: click on this button, if you already have stored an X-matrix file from a previous run. In this mode, no ANOVA is conducted (i.e. steps 1 to 3 described in section 2.2 are skipped). Instead, the dendrogram is directly displayed.

### 2.3.2 Selection of shape parameters

In the next step a selection field is unlocked (see Fig.4), which lists all 23 shape parameters (see Table 1 in Appendix) provided by the PARTISAN output.

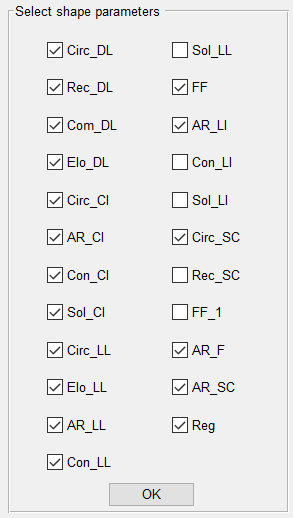


Fig.4 Shape parameter selection field.

By checking or unchecking the according check boxes, shape parameters can be included or excluded from computing the dendrogram.

By default, 18 shape parameters are checked, following the considerations by Dürig et al. (Dürig et al. 2020b). Note that the shape parameters which are by default deselected are redundant, and their selection would not contribute with any more morphometric information. Selecting these additional parameters would lead to a biased X-matrix and is therefore not recommended. For more details about the shape parameters and their mathematical definitions, the reader is referred to Dürig et al. (2018).

### 2.3.3 Constructing dendrogram

After the shape parameter selection process is finished, click “OK” to proceed.

A window pops up with a progress bar, while the X-matrix is computed. When the processing is finished, the SPI is displayed and the dendrogram is displayed together with a number of additional fields and buttons (see Fig.5).

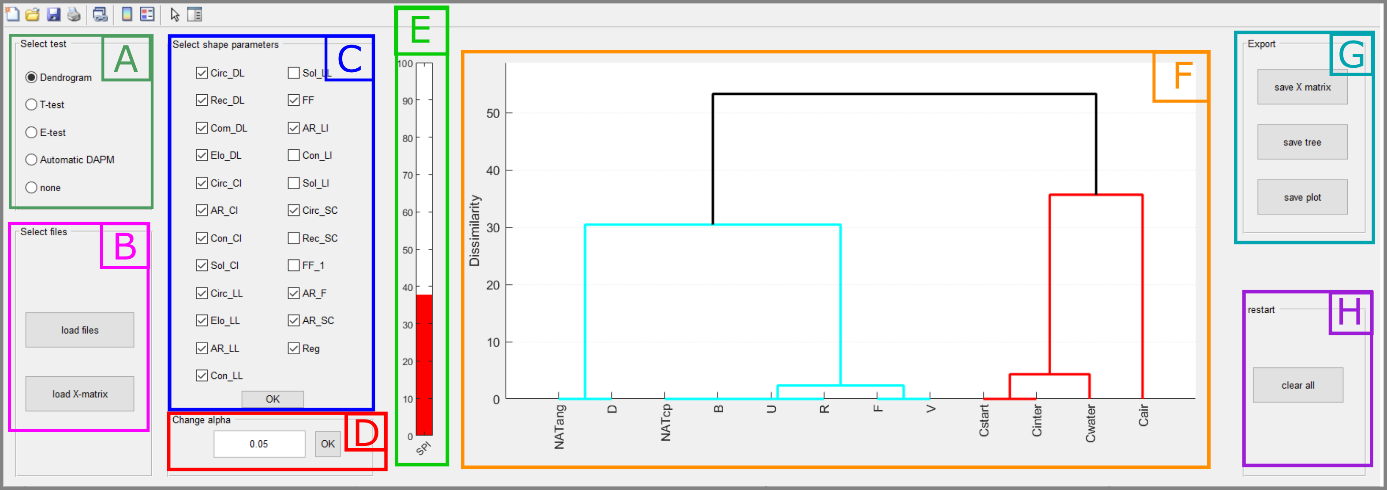


Fig.5 Screenshot of the DendroScan GUI in dendrogram mode after successful computation. A) test selection field; B) file selection field; C) shape parameter selection field; D) significance level field; E) SPI (Statistical Power Index) indicator; F) dendrogram output area; G) export field; H) reset field.

### 2.3.4 Change alpha

By default, the X-matrix (and with it the dendrogram) is computed with the recommended significance level α of 5%.

However, it is also possible to repeat the computation of the dendrogram with a different α value. For this purpose, change α in the according field (field D in Fig. 5) and press the “OK” button.

### 2.3.5 The SPI (statistical power index) indicator

The SPI (statistical power index) ranges between a value of 0 to 100 and is visualized by a vertical bar (see Fig.5 E). A high value reflects a high reliability of the ANOVA results, which underly the dendrogram.

The traffic lights colour scheme of the SPI indicator gives the user some “first glance” guidance:

* **red/empty bar**: SPI < 50%; dendrogram is of low reliability and high probability of type II errors. This means data sets which are grouped together in the dendrogram, could still be characterized by significant, yet unrevealed differences. It is therefore highly recommended to repeat the dendrogrammatic analysis with fewer data sets, for example by focussing separately on the individual clusters.
* **yellow bar**: 50% < SPI < 75%; dendrogram is of moderate reliability and medium probability of type II errors. Data sets grouped together in the dendrogram might still be characterized by significant, yet unrevealed differences. The user is therefore advised to repeat the dendrogrammatic analysis with fewer data sets, for example by focussing separately on the individual clusters.
* **green bar**: SPI > 75%; dendrogram is of (relatively) high reliability and low probability of type II errors.

### 2.3.6 Dendrogram display

Fig.5 (F) illustrates a typical dendrogram. Data sets grouped by similarities form “clusters”. The main clusters are characterized by different colours.

### 2.3.7 Export results

In the file export area (see Fig.5 G) the user is provided three options to export results from the dendrogrammatic analysis, by clicking on the following buttons:

* **“save X-matrix”**: clicking on this button results in opening a file browser window, where the user can choose export destination path and filename. **DendroScan** then creates a .csv file, which contain the entries of the computed X-matrix.
* **“save tree”**: clicking on this button opens a a file browser window, where export destination path and filename can be specified. **DendroScan** then creates two .csv files: i) a “tree file” with the specified name (<filename.csv>), and   
  ii) a “label file” named <filename\_label.mat>.   
  The tree data in these files represent the dendrogram in a Matlab readable format, the dendrogram can simply be reproduced. This can be done, for example, by using the subsequent Matlab command lines :  
   >> t = csvread('filename.csv');

>> load filename\_labels.mat

>> dendrogram(t,'labels',labels)

* **“save plot”:** clicking on this button opens a a file browser window, where the user can specify destination path and filename, under which an image of the dendrogram is stored as “.png” file.

### 2.3.8 Clear window and restart program

In the reset field of the GUI (lower right corner, see Fig.5 H), a “clear all” button is located. When clicking on this button, the screen is cleared and the program is reset.

Note:  
If any of the settings are changed (e.g., if shape parameters are deactivated, or different files are selected), DendroScan updates the dendrogram, as soon as any of the “OK” buttons is clicked. It is therefore not strictly necessary to click on “clear all” before a new round of analysis.

# 3 “T-test” mode

## 3.1 When to use

Select this test mode to statistically verify whether two data sets are significantly different in any of their shape parameters.   
In the DAPM approach, t-tests are conducted after construction of a dendrogram with sufficient statistical power.

## 3.2 Theoretical background of a two-tailed t-test

A t-test is a statistical method based on the Student’s t-distribution (Student 1908) that has previously been applied in particle morphometry (e.g., Dellino et al. 2001; Dürig et al. 2012). It provides the error likelihood (“p-value”) of the null hypothesis, which states that the two tested data sets are identical. If the error likelihood *p* is below the level of significance *α*, the null hypothesis can be rejected: the data sets are then verified to be “significantly different” in the tested hypothesis (Student 1908; Zabell 2008).

Before t-tests are applied, to assess whether the variances of the data sets are homogeneous. **DendroScan** uses Levene-tests (Levene 1960) for this purpose. In cases where the variances of the compared data sets are verified to be homogeneous, “pooled variance t-tests” are computed (Student 1908). If this precondition is not met a “separated variance t-test”, developed by Welch (1947) is conducted.

While a t-test is a very robust and reliable method to test two sets of randomly selected samples, its reliability decreases when the same data sets are repeatedly used (Bender and Lange 2001). In such cases the likelihood of a type I error (i.e., the test indicates a significant difference where there is none) increases. For this reason, it is not recommended to conduct a DAPM to more than one sample known to be from the same population.

Statistically, the t-test can verify a significant difference (when *α*< *p*) between two data sets. The t-test is not suitable, however, to prove mathematically a statistical “similarity” when *α*> *p*. For this purpose, the user is advised to use e-tests, instead.

## 3.3 Steps to compute two-tailed t-tests

### 3.3.1 Initiation and file selection

The t-test mode is activated by clicking on the radio button “t-test” (see Fig.1). Under the test selection field, a file selection field appears with the button “load files” (see Fig.6).

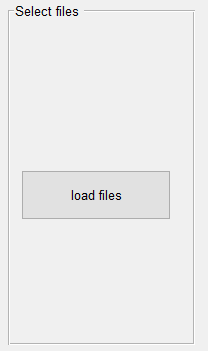
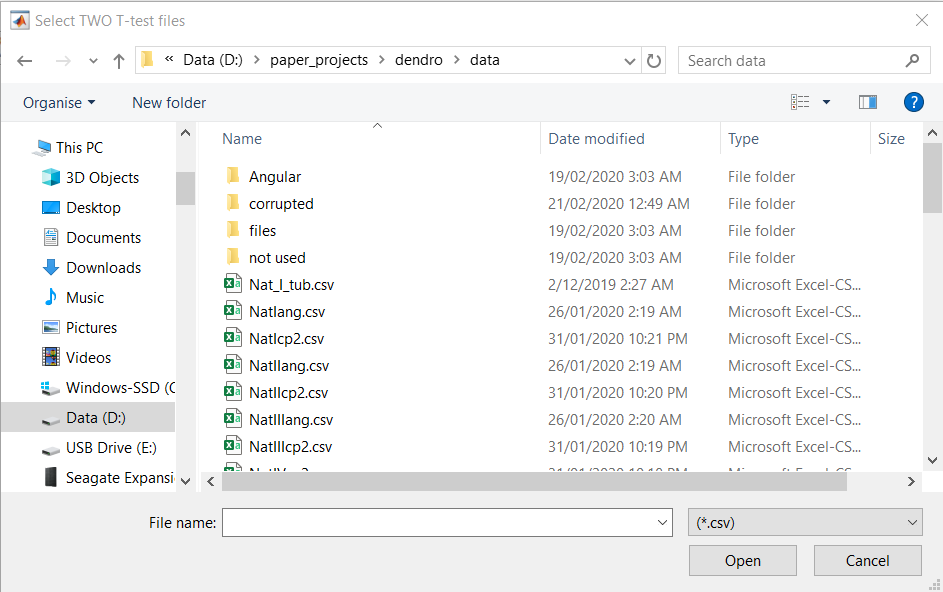
 

Fig.6 File selection field in “t-test” mode (left). Clicking on either button opens a file browser window (right).

* **“load files”**: click on this button to pick two files which you want to compare via a two-tailed t-test.

### 3.3.2 Selection of shape parameters

In the next step a selection field is unlocked (see Fig.7), which lists all 23 shape parameters (see Table 1 in Appendix) provided by the PARTISAN output. By default, 18 shape parameters are selected according to Dürig et al. (Dürig et al. 2020b).

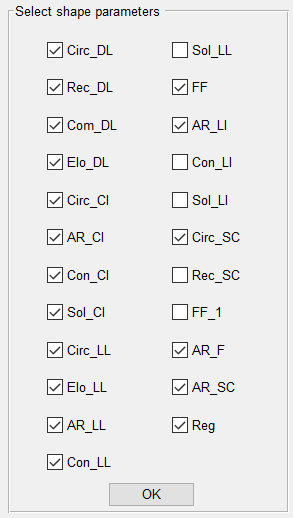


Fig.7 Shape parameter selection field.

By checking or unchecking the according check boxes, shape parameters can be included or excluded from the t-test.

### 3.3.3 Performing two-tailed t-tests

After the shape parameter selection process is finished, click “OK” to proceed.

When the computation of t-tests is finished, the test results for all selected shape parameters are displayed (see Fig.8).

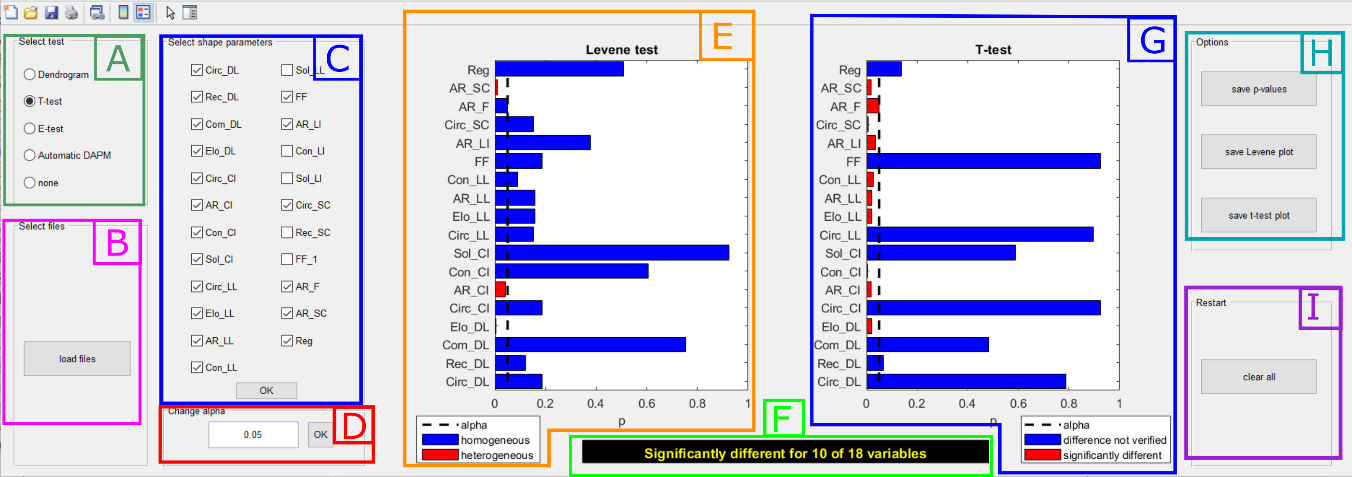


Fig.8 Screenshot of the DendroScan GUI after application of t-tests. A) test selection field; B) file selection field; C) shape parameter selection field; D) significance level field; E) Levene-test result plot; F) results summary field; G) t-tests result plot; H) export field; I) reset field

### 3.3.4 Change alpha

By default, a significance level α of 5% is used. If you want to repeat the t-test with a different α value, change α in the according field (field D in Fig.8) and press the “OK” button. The t-test is then repeated with the modified α-value.

### 3.3.5 Interpretation of Levene-test plots

Levene-test results for each of the selected shape parameter are illustrated by a horizontal bar plot (see Fig.8 E). Each bar represents the according p-value. The vertical dotted line reflects the significance level.

* **blue bar**: Levene-test’s p-value is above 5%; the tested data sets are characterized by homogeneous variances. Pooled variance t-test is applied.
* **red bar**: Levene-test’s p-value is below 5%; the tested data sets are characterized by heterogeneous variances. Separate variance t-test is applied.
* **no bar**: Levene-test’s p-value is considerably lower than 5%; the tested data sets are characterized by extremely heterogeneous variances. Separate variance t-test is applied.

### 3.3.6 Interpretation of t-test results

T-test results for each of the selected shape parameter are illustrated by a horizontal bar plot (see Fig.8 G). Each bar represents the according p-value. The vertical dotted line reflects the significance level.

* **blue bar**: t-test’s p-value is above 5%; significant difference cannot be verified
* **red bar**: t-test’s p-value is below 5%; data sets are characterised by significantly different in the tested shape parameter.
* **no bar**: t-test’s p-value is considerably lower than 5%; difference in the tested shape parameter is highly significant.

The t-test results are summarized in the results summary field (Fig.8 F).

### 3.3.7 Export t-test results

In the file export area (see Fig.8 H) the user is provided three options to export the t-test results, by clicking the following buttons:

* **“save p-values”**: clicking this button results in opening a file browser window, where the user can choose export destination path and filename. **DendroScan** then creates a .csv file containing three columns:   
  - column 1: list of all 23 shape parameters  
  - column 2: p-values of Levene-tests (selected shape parameters)  
  - column 3: p-values of t-tests (selected shape parameters)

Note:  
p-values for non-selected shape parameters are not computed and listed as “NaN” in the .csv files.

* **“save Levene plot”**: clicking this button opens a a file browser window, where export destination path and filename can be specified. **DendroScan** then saves the Levene plot as .png file.
* **“save t-test plot”:** clicking this button opens a a file browser window, where export destination path and filename can be specified. **DendroScan** then saves the t-test plot as .png file.

### 3.3.8 Clear window and restart program

In the reset field of the GUI (lower right corner, see Fig.8 I), a “clear all” button is located. When clicking on this button, the screen is cleared, and the program is reset.

# 4 “E-test” mode

## 4.1 When to use

Select the “e-test” (equivalence test) mode to statistically verify whether two data sets are statistically equivalent in their shape parameters.   
When following the DAPM approach, e-tests are applied as a last step for data sets which did show in t-tests no significant differences in any of their shape parameters.

## 4.2 Theoretical background of equivalence tests

While both ANOVA and t-tests are useful for proving significant differences, failing these tests alone can - from a mathematical point of view - not be used as proof for equivalence (Walker and Nowacki 2011). This gap is covered by e-tests. For image particle analysis, this method was introduced by Dürig et al. (2012). For each shape parameter, this method computes whether the confidence interval *Δ* (with level of significance being *α*) from one sample is within a given acceptable range *Dmax*, denoted “equivalence margin” (Rasch and Guiard 2004; Wellek 2010), which specifies a lower and an upper boundary *ΔL* and *ΔU*. To verify equivalence, two one-sided t-tests are conducted, one of either side of the equivalence margin, testing the composed null hypotheses *H01:* *Δ < ΔL* and *H02: Δ >* *ΔU*. T-test results leading to the rejection of both hypotheses imply that *ΔL* < *Δ < ΔU* and serve as proof for statistical equivalence (Rasch and Guiard 2004; Wellek 2010).

In contrast to earlier methods (Dürig et al. 2012, 2020b), t-tests are applied for both cases, for data sets with homogeneous variances as well as for cases where the variances of the two compared data sets are heterogeneous. The application of t-tests followed the strategy described in section 3.2.

Equivalence tests provide only useful information in conjunction with their “*Dmax\_cal*” values. In other words, a statement if two data sets are “statistically equivalent” in a certain shape parameter is only of use when the according *Dmax\_cal* value is provided, too.

**DendroScan** provides two options:

* To apply e-tests with pre-defined *Dmax\_cal* values (saved in a “calibration file”). This option can be chosen if one wants to use literature values (e.g., from Dürig et al. (2020a)).
* To compute calibrated *Dmax\_cal* values based on selected files (“standards”). As standards serve files which contain data sets that are known (or at least assumed) to be statistically equivalent, and hence reflect the “acceptable natural variation” of shape parameters. It is recommended to use files with a sample size of at least 20 and larger. **DendroScan** uses the standards for calibration, by conducting e-tests with a *Dmax\_cal* value of 0.01 and increasing it stepwise by 0.01, until the tests indicate statistical equivalence between all used standards.

## 4.3 Steps to conduct e-tests

### 4.3.1 Initiation and file selection

The e-test mode is activated by clicking on the radio button “e-test” (see Fig.1). Under the selection field, three radio buttons show up (see Fig.9), which offer the user to choose between:

* **“Import cal file”**: click on this option to import a calibration file which contains previously stored calibrated *Dmax\_cal* values.
* **“run calibration”**: if this option is chosen a file selection window opens, which allows the user to pick two or more .csv data files (“standards”). Based on these standards, **DendroScan** computes the *Dmax\_cal* values for subsequent e-tests. To select, click on the files by holding down the control-key.

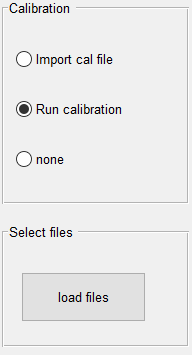


Fig.9 Calibration type selection field in “e-test” mode. “Run calibration” is selected, in which case DendroScan conducts a calibration based on standards. The standards can be selected, by clicking the “load files” button.

After importing or computing the *Dmax\_cal* values, a second file selection window opens, which asks the user to pick the two .csv data files for comparing them via e-tests. This step can be repeated at any time by clicking on the button “load files” in the lower left corner of the GUI.

### 4.3.2 Selection of shape parameters

In the next step a selection field occurs (see Fig.10), which lists all 23 shape parameters (see Table 1 in Appendix) provided by the PARTISAN output. By default, 18 shape parameters are selected according to Dürig et al. (Dürig et al. 2020b).

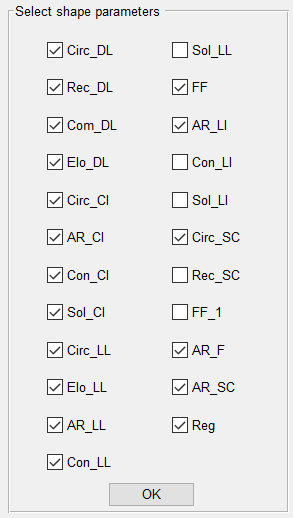


Fig.10 Shape parameter selection field.

By checking or unchecking the according check boxes, shape parameters can be included or excluded from the e-tests.

### 4.3.3 Conducting e-tests

After the shape parameter selection process is finished, click “OK” to proceed.

When the computation of e-tests is finished, the test results for all selected shape parameters are displayed (see Fig.11).

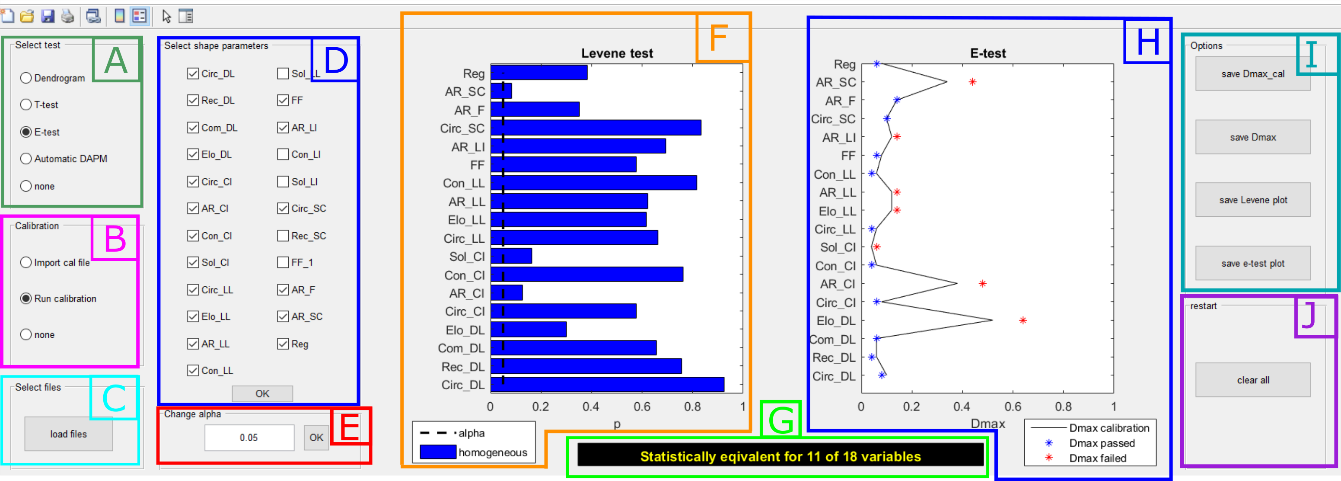


Fig.11 Screenshot of the DendroScan GUI after computation of e-tests. A) test selection field; B) calibration type field; C) file selection field; D) shape parameter selection field; E) significance level field; F) Levene-test result plot; G) results summary field; H) e-tests result plot; I) export field; J) reset field

### 4.3.4 Change alpha

By default, a significance level α of 5% is used. If you want to repeat the t-test with a different α value, change α in the according field (field E in Fig.11) and press the “OK” button. The e-test is then repeated with the modified α-value.

### 4.3.5 Interpretation of Levene-test plots

Levene-test results for each of the selected shape parameter are illustrated by a horizontal bar plot (see Fig.11 F). Each bar represents the according p-value. The vertical dotted line reflects the significance level.

* **blue bar**: Levene-test’s p-value is above 5%; the tested data sets are characterized by homogeneous variances. For the according shape parameter’s e-test pooled variance t-tests are applied.
* **red bar**: Levene-test’s p-value is below 5%; the tested data sets are characterized by heterogeneous variances. Separate variance t-tests are applied for this shape parameter.
* **no bar**: Levene-test’s p-value is considerably lower than 5%; the tested data sets are characterized by extremely heterogeneous variances. Separate variance t-tests are applied for this shape parameter.

### 4.3.6 Interpretation of e-test results

The outcome is illustrated by a plot featuring the *Dmax* values for each of the tested shape parameters (see Fig.11 H). Hereby, *Dmax* (indicated by an asterisk marker) is a test parameter, which represents the minimum *Dmax\_cal* value for which the tested data sets would be statistically equivalent, according to the e-test.   
In addition, the equivalence margin (*Dmax\_cal*) is illustrated by a black line.

Statistical equivalence is verified if *Dmax* ≤ *Dmax\_cal*

The colours of the markers indicate the outcome of the e-tests:

* **blue marker**: *Dmax* value is equal to or smaller than the equivalence margin. Statistical equivalence is verified for this shape parameter.
* **red marker**: *Dmax* value is larger than equivalence margin. Statistical equivalence is not verified.

A summary of the e-test results is presented in the results summary field (Fig.11 G).

Note:  
According to DAPM two data sets are only verified to be “statistically equivalent”, when statistical equivalence is proven for all tested shape parameters (Dürig et al. 2020b).

### 4.3.7 Export e-test output

In the file export area (see Fig.8 I) the user is provided three options to export the e-test results, by clicking the following buttons:

* **“save cal. file”**: clicking this button opens a file browser window, where the user can choose export destination path and filename. **DendroScan** then creates a .csv calibration file which contains the list of *Dmax\_cal* values which define the equivalence margin. Such a file can be imported into Dendroscan via the “load cal.” function in the calibration field (Fig. 11 B). For non-selected shape parameters, “NaN” is displayed.
* **“save results”**: clicking this button opens a a file browser window, where export destination path and filename can be specified. **DendroScan** then creates a .csv file, which contains four columns:   
  - column 1: list of all 23 shape parameters  
  - column 2: p-values of Levene-tests   
  - column 3: computed *Dmax* values, (hypothetical minimum *Dmax\_cal* values for which data sets would be statistically equivalent)   
  - column 4: equivalence margins (calibrated *Dmax\_cal* values)   
  For each shape parameter, statistical equivalence is verified if *Dmax* ≤ *Dmax\_cal*.

Note:  
*Dmax* values for non-selected shape parameters are not computed and listed as “NaN” in the .csv files.

* **“save Levene plot”**: clicking this button opens a a file browser window, where export destination path and filename can be specified. **DendroScan** then saves the Levene plot as .png file.
* **“save e-test plot”:** clicking this button opens a a file browser window, where export destination path and filename can be specified. **DendroScan** then saves the e-test plot as .png file.

### 4.3.8 Clear window and restart program

In the reset field of the GUI (lower right corner, see Fig.8 J), a “clear all” button is located. When clicking on this button, the screen is cleared, and the program is reset.

# 5 “Automatic DAPM” mode

## 5.1 When to use

DAPM stands for “dendrogrammatic analysis of particle morphology”. Select the “Automatic DAPM” mode if you want to automatically sort multiple data sets according to their morphometrical dissimilarities and to identify data sets of identical shapes.

## 5.2 Theoretical background of equivalence tests

The DAPM approach, as suggested by Dürig et al. (Dürig et al. 2020b) foresees the application of ANOVA-based dendrogram construction in order to group the data into clusters. This step is reiterated until no further grouping can be done.

In the next step t-tests are conducted for all samples still grouped together.

In a final step, e-tests are performed for those samples, which are not found to be significantly different in any of the selected shape parameters.

Under some circumstances, however, grouping of data sets in a dendrogram can be ambiguous: if samples are separated by 0 distance in some of the shape parameters, it might occur that the grouping into the one or the other subcluster becomes arbitrary, depending on the linkage method and, e.g., the order of files processed. This might lead to an erroneous sorting in an early step of the dendrogrammatic analysis.

In order to counter that potential issue, **DendroScan** conducts t-tests for all samples which show 0 distance towards each other (even if they might have been grouped in a different cluster at an earlier step.

The advantage of this extra step is that an erroneous sorting in an early step dendrogram would be corrected. However, if a multitude of very similar samples are processed (resulting in many distances of 0 in the X-matrices), this might result in a multitude of additional t-tests, which - as described in section 3.2- increases the likelihood of a type I errors. It is therefore recommended to avoid the processing of “redundant” data sets (e.g., several samples which are from the same population and therefore known to be of the same shape). It is very much advised to use such “redundant” samples as standards for calibration purposes, instead.

If this advise is considered, the accuracy of an automatic DAPM is higher than manually performed.

The automatic DAPM conducts the following steps:

1. Computation of ANOVA-based matrix X and dendrogram linkages.
2. Constructing agglomerative clusters from the linkage with a distance threshold of 70% of the maximum distance (dissimilarity value).
3. Repeat steps 1 and 2 for each (sub)cluster, until no further (smaller) subclusters can be generated. During each iteration, the X matrix is updated with the new distances.
4. Determine which samples have 0 distance with each other. Conduct t-tests for these samples and modify the X-matrix by adding a value of 1 for each shape parameter which was found to be significantly different. Distance values unequal to 0 are increased by the maximum distance for all additional entries.
5. Display a preliminary dendrogram, which reflects the updated X-matrix.
6. Conduct e-tests for those samples, which still have a distance of 0 in the updated X-matrix. Run calibration by using standards or use calibration files provided by the user. Modify the X-matrix accordingly, by adding a distance value of 1 for each shape parameter, which was not verified to be statistically equivalent. Again, distance values unequal to 0 are increased by the maximum distance for all additional entries.
7. Display the final dendrogram, which sorts the data sets according to the updated X-matrix. Samples which are grouped together with a dissimilarity of 0 are verified to be identical.

In order to provide quality guidance, **DendroScan** computes a statistical reliability index (“SRI”) along with the DAPM. Similar to the SPI, its values range between 0 and 100, and are computed by:

(5)

where *f* is the probability density function:

(6)

with *M* being the maximum number of t-tests conducted for the same data set.   
The adjustment parameters *µ* and *s* are set to 1 and 2.5. The relationship between *M* and the SPI is illustrated in Fig. 2. Note that in contrast to the SPI, a low SRI indicates a high likelihood of type I errors, which means that the significant differences might be indicated, where there are none.

Fig. 12 Relationship between SRI and maximum number of t-tests conducted for the same data set.

## 5.3 Steps to conduct automatic DAPM

### 5.3.1 Initiation and file selection

The automatic DAPM mode is activated by clicking on the respective radio button “e-test” (see Fig.13).

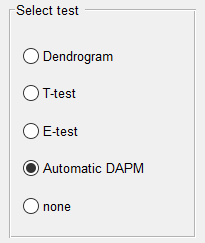


Fig. 13 An automatic DAPM is started by selecting the respective radio button in the test selection field.

Under the test selection field, a button shows up

* “**load files**”: Clicking on this button opens a file manager. Select all files, for which an automatic DAPM should be conducted.

### 5.3.2 Selection of shape parameters

In the next step a selection field occurs (see Fig.14), which lists all 23 shape parameters (see Table 1 in Appendix) provided by the PARTISAN output. By default, 18 shape parameters are selected according to Dürig et al. (Dürig et al. 2020b).

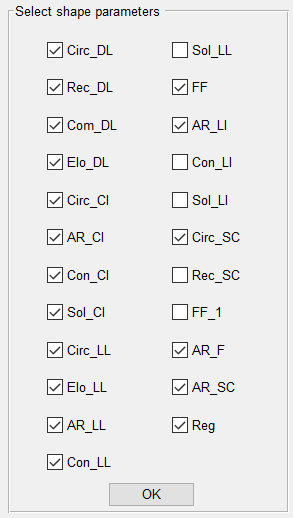


Fig.14 Shape parameter selection field.

By checking or unchecking the according check boxes, shape parameters can be included or excluded from the e-tests.

### 5.3.3 Preliminary DAPM diagram and Statistical Reliability Index (SRI)

After finishing the shape parameter selection, click the “OK” button to proceed. **DendroScan** computes a sequence of dendrogrammatic analysis. When finishing the t-tests (step 4 in section 5.2), a preliminary dendrogram is displayed (see Fig.15).

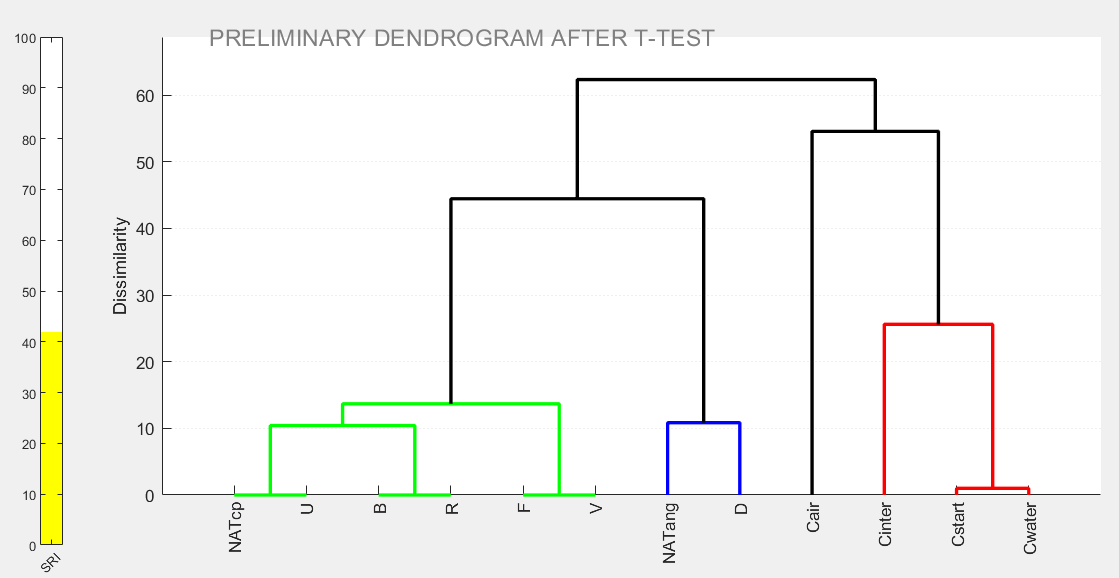


Fig.15 Screenshot showing preliminary dendrogram, computed in automatic DAPM mode.

The diagram displayed at this stage is labelled by a water mark stating “PRELIMINARY DENDROGRAM AFTER T-TEST”.

A coloured bar on the left side of the dendrogram indicates the SRI (Statistically Reliability Index), which is based on the maximum number of t-tests *M* conducted with the same data set during step 4 of the DAPM (see 5.2) and ranges between 0 and 100.

The SRI is visualized by a vertical bar. A high value reflects a high reliability of the t-test results, which underly the dendrogram. The traffic lights colour scheme of the SRI indicator gives the user some “first glance” guidance:

* **red/empty bar**: SRI < 30%; dendrogram is of low reliability and high probability of type I errors. This means data sets might have wrongly grouped apart in the dendrogram. Depending on the objective, it might therefore be better to repeat the DAPM with fewer (or less similar) data sets.
* **yellow bar**: 30% < SRI < 75%; dendrogram is of good reliability and medium probability of type I errors. Data sets grouped apart in the dendrogram might still be more similar, than the DAPM suggests. It might therefore be a good idea to repeat the automatic DAPM with fewer (or less similar) data sets.
* **green bar**: SRI > 75%; dendrogram is of high reliability and low probability of type I errors.

### 5.3.4 Selecting of calibration files or standards for e-tests

The preliminary dendrogram serves mainly to assist the user with the decision on the standards to be used for the succeeding e-tests.

Example:

In the example presented in Fig.14, three sample pairs are grouped together and will be analysed by e-test in the following step: “NATcp” vs “U”, “B” vs ”R” and “F” vs “V”. Assume, the user has two calibration files available: one for “NATcp”, and one for “NATang”. In all three e-tests, using the calibration file for “NATcp” samples would be preferable, since the samples in question are closer grouped to that data set, than to “NATang”.

The color of the bar follows a traffic light scheme, indicating the reliability of the ANOVA

With displaying the preliminary dendrogram an e-test selection field appears in the upper right side of the window, offering a number of radio buttons:

* **Import cal file:** chose this option to import a calibration file which contains previously stored calibrated *Dmax\_cal* values. For each e-test, DendroScan will open a separate file browser and request a calibration file.
* **Run calibration:** chose this option to pick two or more .csv data files (“standards”). Based on these standards, **DendroScan** computes the *Dmax\_cal* values for subsequent e-tests. For each e-test, DendroScan will open a separate file browser, requesting the user to select standards.
* **Import cal file (all):** click on this option to import a calibration file which contains previously stored calibrated *Dmax\_cal* values. This file will be used for all e-tests conducted during the automatic DAPM.
* **Run calibration (all):** chose this option to pick two or more .csv data files (“standards”). Based on these standards, **DendroScan** computes the *Dmax\_cal* values for subsequent e-tests. These *Dmax\_cal* values will be used for all e-tests conducted during the automatic DAPM.
* **none**: this is the default selection. As long as the radio button is on “none”, the program is paused.

Depending on the user’s choice, **DendroScan** proceeds with computing or using the according *Dmax\_cal* values for applying the e-tests. As a result, the X-matrix is updated, and visually presented via a final dendrogram (Fig.16).

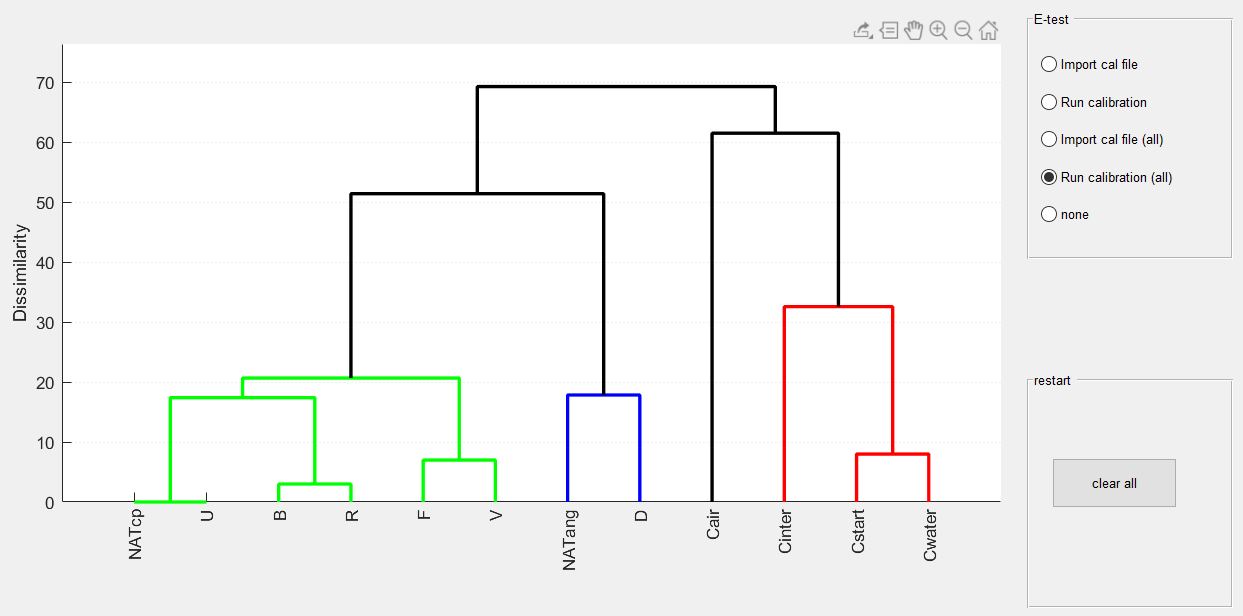


Fig.16 Screenshot showing final dendrogram (i.e., after e-tests), computed in automatic DAPM mode.

### 5.3.5 Output generated by automatic DAPM

Besides the dendrogram displayed in the GUI window, DendroScan also creates output files, into a subfolder located in the “results” folder. The name of the subfolder is “DAPM\_” followed by date and time of generation. For example: “DAPM\_20200417\_2200”.

Eight files are generated:

* 3 files, named “DAPM\_dendrogram\_prelim” in .eps, .png and .tif format. These image files contain the preliminary dendrograms.
* 3 files, named “DAPM\_dendrogram\_final” in .eps, .png and .tif format, containing the resulting dendrograms.
* a file denoted “X-matrix.csv”, which contains the final X-matrix (including the updates after t-tests and e-tests).
* a logfile, denoted “DAPM\_LogFile.txt”.

The entries of the log files (see Fig.17) are grouped into three sections:

* “INFO”: lists time, number and names of analyzed files and selected shape parameters.
* “WARNINGS”: lists all data sets, which are tested via t-tests three times or more.
* “RESULTS”: lists all e-test results of compared data pairs. The results include a statement about the number of shape parameters, which were found to be statistically equivalent.

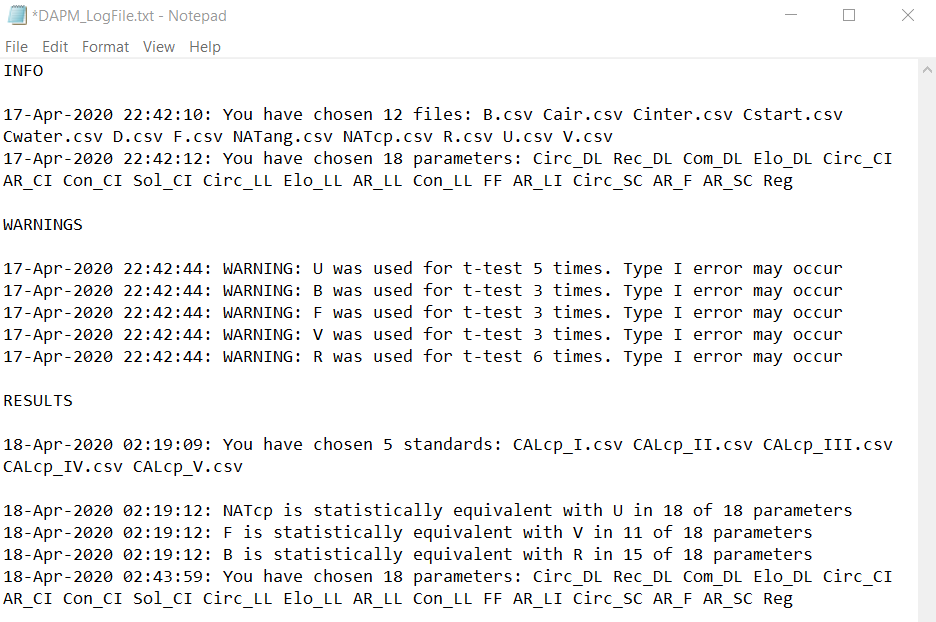


Fig.16 Screenshot showing entries of a logfile created with automatic DAPM.

### 5.3.6 Clear window and restart program

In the reset field of the GUI (lower right corner, see Fig.16), a “clear all” button is located. When clicking this button, the screen is cleared, and the program is reset.

# References

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

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

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

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 (2012) 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, White JDL, Zimanowski B, et al (2020a) Boosting deep-sea eruptions with induced fuel-coolant explosions. Nat Geosci in review:

Dürig T, White JDLL, Zimanowski B, et al (2020b) Dendrogrammatic analysis of particle morphometry links experimental particles with Havre submarine ash and reveals a deep-sea fragmentation style. Bull Volcanol in prep.:

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

Heiken G (1972) Morphology and petrography of volcanic ashes. Geol Soc Am Bull 83:1961–1988

Heiken G (1974) An atlas of volcanic ash. Smithson Contrib to Earth Sci 1–101. https://doi.org/10.5479/si.00810274.12.1

Heiken G, Wohletz K (1985) Volcanic ash. University of California Press, Berkeley

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. Contrib to Probab Stat Essays …

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

Mégevand P (2020) pierremegevand/games\_howell

Nagy P (2020) Cumulative distribution function of the studentized range (for Tukey’s HSD test)

Rasch D, Guiard V (2004) The robustness of parametric statistical methods. Psychol Sci

Schmith J, Höskuldsson Á, Holm PM (2017) Grain shape of basaltic ash populations: implications for fragmentation. Bull Volcanol 79:1–16. https://doi.org/10.1007/s00445-016-1093-5

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

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

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

Zabell SL (2008) On student’s 1908 article “the probable error of a mean.” J Am Stat Assoc. https://doi.org/10.1198/016214508000000030

# Special kudos

Peter Nagy (2020). Cumulative distribution function of the studentized range (for Tukey's HSD test) (https://www.mathworks.com/matlabcentral/fileexchange/37450-cumulative-distribution-function-of-the-studentized-range-for-tukey-s-hsd-test), MATLAB Central File Exchange. Retrieved February 20, 2020.

Pierre Mégevand (2020). pierremegevand/games\_howell (https://www.github.com/pierremegevand/games\_howell), GitHub. Retrieved February 20, 2020.

# Appendix

Table 1: Non-dimensional shape parameters provided by the open source software PARTISAN in .csv files, which serve DendroScan as data input.

|  |  |  |
| --- | --- | --- |
| shape parameter | notation | morphometric system by |
| **Circ\_DL** | **circularity** | **Dellino and La Volpe (1996)** |
| **Rec\_DL** | **rectangularity** |
| **Com\_DL** | **compactness** |
| **Elo\_DL** | **elongation** |
| **Circ\_CI** | **circularity** | **Cioni et al (2014)** |
| **AR\_CI** | **aspect ratio** |
| **Con\_CI** | **convexity** |
| **Sol\_CI** | **solidity** |
| **Circ\_LL** | **circularity** | **Leibrandt and Le Pennec (2015)** |
| **Elo\_LL** | **elongation** |
| **AR\_LL** | **aspect ratio** |
| **Con\_LL** | **convexity** |
| **Sol\_LL** | **solidity** | **Liu et al (2015)** |
| **FF** | **form factor** |
| **Con\_LI** | **convexity** |
| **Sol\_LI** | **solidity** |
| **AR\_LI** | **aspect ratio** |
| **Circ\_SC** | **circularity** | **Schmith et al (2017)** |
| **AR\_F** | **Feret aspect ratio** |
| **FFa** | **form factor** |
| **AR\_SC** | **reciprocal aspect ratio** |
| **Reg** | **regularity** |

Table 2: Notation of variables

|  |  |
| --- | --- |
| parameter | description |
| ***p*** | p-value, error likelihood of the null hypothesis, according to which the tested data sets are from the same population |
| ***pijk*** | p-value of data set *i* tested with the one from data set *j* in the *k*-th shape parameter |
| ***α*** | significance level |
| ***Xij*** | X-matrix |
| ***N*** | number of data sets |
| ***f*** | probability function |
| ***µ*** | adjustment parameter |
| ***s*** | adjustment parameter |
| ***SPI*** | statistical power index |
| ***Dmax\_cal*** | equivalence margin |
| ***dlower*** | lower equivalence bound |
| ***dupper*** | upper equivalence bound |
| ***δ*** | range determined by the confidence level (1- ***α***) |
| ***Dmax*** | test parameter, minimum ***Dmax\_cal*** value for which e-test would be passed |

1. PARTISAN can be downloaded here: <https://doi.org/10.4401/ag-7865> [↑](#footnote-ref-2)