***INPUT AND OUTPUT FILES FOR BINARY_TRAVERSER***

These instructions are also available at http://www.geo.umass.edu/climate/lewis/analysis/

**Note:**

$X_p$ refer to 'parameters'. They are chosen in the parameter control panel (see detailed instructions file, Fig. B2).

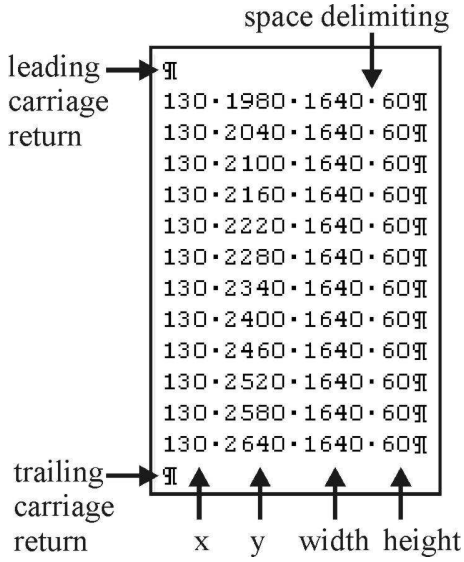
$X_m$ refer to measurements. They are chosen in the measurement control panel (see detailed instructions file, Fig. B3).

**1– ROI coordinate input file:**

EXPLANATION: The coordinates, in pixels, of the ROI.

Tips:

- this file can be created in, and exported from, a spreadsheet program.
- be sure there are no leading or trailing blank lines (Fig. C1).
- be sure that no ROI have no particles in them.
- To test your ROI coordinate input file, it would be a good idea to visualize your ROI before running the macro. This will help the user determine whether any ROI have no particles, and to be sure all ROI are properly positioned. A tool called '*ROI_imager*' is included to help you do this.




o

**Figure C1.** Coordinate file structure. The file should have no leading or trailing carriage returns or newlines.

FILE STRUCTURE:

Column 1: left edge of the measuring box x-coordinate

Column 2: top edge of the measuring box y-coordinate

Column 3: measuring box width

Column 4: measuring box height

**2 - CORENAME_PARAMETER_STATISTICS.txt:**

*(e.g 3eg-1_AREA_STATISTICS.txt)*

EXPLANATION: A comma separated file that is output by the macro after it runs. It contains the chosen statistics for each ROI. If applicable, units are in microns, or microns squared for area, as set in $B_p$.

FILE STRUCTURE:

*X:* left edge of the measuring box x-coordinate

*Y:* top edge of the measuring box y-coordinate

*Width:* measuring box width

*Height:* measuring box height

*Area:* measuring box area

*Mid_Y:* y-coordinate of the midpoint of the measuring box

*Number_Parts:* the number of particles within the ROI that satisfy the boundary criteria

*Cumulative:* the sum of all measurements of a parameter for an ROI. This only really makes sense for area, where you want to calculate % black pixels.

*Mean:* mean

*Median:* median (if the number of measurements is odd, then this is $m_{50}$. If the number of measurements is even, then the median is the average of the two middlemost measurements).

*Mode*: the most frequently occurring measurement. Numbers are rounded to the nearest integer for all parameters except aspect ratio and circularity. For aspect ratio and circularity, mode may not be useful, since numbers are reported to 4 decimal places.

*Max*: maximum number

*Min:* minimum number

*StDev:* standard deviation, where:

$$s = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n-1}}$$

*Skewness:* where:

$$s = \frac{\sum (X_i - \overline{X})^3}{ns^3}$$

*Kurtosis:* where:

$$s = \frac{\sum (X_i - \overline{X})^4}{ns^4}$$

*Perc#Value*: The value of the percentile specified in $E_m$ (# is 1 to 5 corresponding to the percentile number.

*Perc#Closest:* The closest possible percentile to the percentile specified in $E_m$. (Perc#Closest will be very close to the percentile specified in $E_m$ when n is larger than a few hundred; # is 1 to 5 corresponding to the percentile number.).

*Histogram:* All bin counts less than the lower limit of the first bin value are counted in the first histogram column (directly under the 'histogram' column header). A number of unlabeled columns follow after the 'histogram' column header: these contain the counts for each histogram bin. A particle is counted in a bin if it is greater than the lower bin limit and is less than or equal to the upper bin limit as defined in $F_m$, and described in CORENAME_PARAMETER_HISTOGRAM_LIMITS.txt (part 3). On the first bin (the column immediately to the right of the column labeled 'histogram'), a particle is counted if it is greater than *or equal to* the lower bin limit. The number of columns is the same as the number of histogram bins

*plus one*. *Plus one* because the last column is the number of particles greater than the upper bin limit set in $F_m$. Bin boundaries can increase linearly, can increase on the phi scale, or $\log_{10}$ scale. This is explained in the Measurement Dialog section that describes $F_m$.

**3 - CORENAME_PARAMETER_HISTOGRAM_LIMITS.txt**

*(3eg-1_AREA_HISTOGRAM_LIMITS.txt)*

EXPLANATION: The upper and lower limits of each histogram bin. The number of rows equals the number of bins specified in $F_m$.

**4- CORENAME_PARAMETER_RAWDATA_#.txt**

These data files contain the 'raw data' for each particle within a ROI. The data are one column wide, and are for the parameter written in the file name (i.e. area, circularity, etc.). The files are numbered with 1 being measurements from the topmost ROI. Data are scale-transformed where appropriate (i.e. for area, major axis and minor axis, but not for circularity or angle).

**5- CORENAME_CRACKAREA.txt**

This comma delimited file is produced if $D_p$ is chosen. The file lists the total area of disturbances within a binary crack image for each ROI.

Units are in microns, except for 'Area' and 'Crack Area', which are in microns squared.

FILE STRUCTURE:

*X:* left edge of the measuring box x-coordinate

*Y:* top edge of the measuring box y-coordinate

*Width:* measuring box width

*Height:* measuring box height

*Area:* measuring box area

*Mid_Y:* y-coordinate of the midpoint of the measuring box

*Crack_Area:* the black area for each measuring box, representing cracks or other disturbances.

The percentage of cracks per ROI can be calculated using *Area* and *Crack_Area*.