Solid Earth Discuss., 6, C1614–C1621, 2015
www.solid-earth-discuss.net/6/C1614/2015/

Solid Earth
Discussions

Interactive comment on "GrainSizeTools: a Python script for estimating the dynamically recrystallized grain size from grain sectional areas" *by* M. A. Lopez-Sanchez and S. Llana-Fúnez

M. A. Lopez-Sanchez and S. Llana-Fúnez

malopez@geol.uniovi.es

Received and published: 27 February 2015

First we would like to thank Petr Jeřábek for his review. Based on his comments we have prepared the following reply. Original comments by the reviewer appear appear between quotes.

"I do not see the justification of publishing/advertising the description of the script in a scientific journal. I could understand a presentation of more robust scientific software packages covering many aspects/techniques of data processing but the presented script is just not enough."

First, we are aware that as a researchers our goal is to produce knowledge nor code

or a script. Having said this, we disagree with the view that the reviewer has on our work. The manuscript is not just a description of a script. In fact, only two sections of eight are focused in the script: 1) the section that deals with the description of the script; 2) the section that offers a comparison with other applications/scripts to estimate the frequency peak grain size. The main goal of the study is first to test which single measure of grain size, applying 2D methods, give the best estimate of the differential stress (or rate of mechanical work).

In our analysis we tested the mean, the median (included in the new version of the manuscript), the area-weighted mean and the frequency peak of a population of apparent grain sizes. With respect to the latter, it was estimated by considering two non-parametric approaches, the middle (or central) value of the modal interval using histograms and, for the first time, the peak of the Gaussian KDE. We finally suggest the use of the peak of the Gaussian KDE as best estimator.

The second goal is to provide a tool. We introduce a free, open-source and easy to handle script with the aim to facilitate the adoption of this measure of grain size (i.e. the peak of the Gaussian KDE) in the context of paleopizometry or –wattometry studies. To our knowledge there are no previous scripts implemented this method to find the frequency peak grain size via the Gaussian KDE. Of course, anyone with programing skills in Python, R, Julia, MATLAB, etc., can use the Gaussian kernel-density estimator implemented in their scientific packages. Unfortunately, not everybody has programming skills so we decided to release a script and a brief tutorial (https://sourceforge.net/p/grainsizetools/wiki/Home/) with this in mind (i.e. there is no need of prior knowledge of Python language to use the script and get the results).

The following reasons reinforce the publication of our analysis:

1) To the best of our knowledge, there are no previous studies that address the question of which measure of grain size give the best estimate of the differential stress (or rate of mechanical work) in dynamically recrystallized mylonites comparing the mean, the

median and the frequency peak (usually poorly referred as the mode) in grain size distributions. Ranalli (1984) compared the mean and the median grain size with the same goal. However, although the conclusions he arrived are valuable, his study is not by far as complete as ours since his study relied on probabilistic considerations, but not deal with real mylonites.

2) It is demonstrated that the peak of the Gausssian KDE is far superior to the use of modal intervals in histograms to find the frequency peak grain size. This is also one of the reasons portrayed in the manuscript to use our script (or the method described) over the StripStar script or the CSDCorrections in case you wanted to use the frequency peak as a grain size estimator.

3) It is also recommended the use of the peak of the Gaussian KDE over other measures of grain size (mean, median and area-weighted mean) in paleopiezometric studies. Although the simulations performed on a natural sample seem to indicate that the mean, the peak of Gaussian KDE, the median and the area-weighted grain size are equally good as descriptors of grain size, we recommend the Gaussian KDE peak for the following reasons: i) this parameter is potentially less prone to be shifted by the presence of outliers and the image/optical resolution limitations; and ii) regardless you consider a population of apparent 2D grain sizes or the actual 3D grain population, it yields comparable results. This could be useful in the future for comparative purposes if a method capable of measuring directly the actual (3D) grain size distribution becomes mainstream.

4) It is addressed how many grains are needed for this type of studies with a robust theoretical base, not just relying on experience. This is obtained performing a Monte Carlo simulation using bootstrapping over a real sample. In fairness to the reviewer, we admit that the title of the article was particular misleading because it focused on the script (the tool) and not in the comparative between the different measures of grain size (the knowledge). For this reason, we changed the title to "An evaluation of different measures of grain size for paleopiezometry or paleowattometry studies". Furthermore, the

organization of the new manuscript was also improved to make the aims and findings more clear and concise, relegating the role of the script to the background.

"I see two main troubles with scientific justification of the present manuscript 1) The belief that we need to know the single number to characterize and describe sometimes rather complex populations of grain sizes is to me obscure and definitely not step forward. 2) The mentioned simplicity of the procedure and advantages of the presented grain size determination technique in comparison with the other existing techniques (StripStar and CSDcorrections) are not so obvious."

Regarding the first point, we are aware that to fully quantify a continuous population of grains several parameters may be needed that describe numerically the type of continuous distribution of grain sizes. For example, assuming that sizes in dynamically recrystallized grains follow a log-normal distribution we will need the scale and the shape parameters at their original scale or the mean and the standard deviation of the log(grain size) to fully characterize it (see Limpert et al. 2001). We also aware that the future in microstructural analysis is to incorporate grain size distributions in the rheological interpretations. An example of such approach is the interaction of grain size sensitive and insensitive deformation mechanisms during deformation of a polycrystalline aggregate (e.g. Heilbronner and Bruhn 1998; Herwegh et al. 2005; Herwegh et al. 2014; Czaplińska et al. in press). We are indeed currently working on this issue in order to introduce such capabilities within the GrainSizeTools script in the near future.

"a question emerges as whether the simplification used in the contemporary statistical treatment of the grain size populations is not counterproductive. Do we want/need the simplification if the complexity and more numbers characterizing the population may correspond to various mechanisms of the recrystallization process? The single number approach has been used since seventies so isn't already a time to move on?"

With respect to the first question, at least for paleopiezomety or –wattometry studies a single value is needed for comparative purposes. So it seems that we need that simplification. The key question is whether this simplification is valid or not for our purposes and we think the answer is presumably positive. To clarify, imagine just for simplicity that the grain size population in dynamically recrystallized mylonites follows a normal distribution of grain sizes. If different values of stress (or rate of mechanical work) yields different and unique values of mean, median or frequency peak grain size, then you can use a single value for comparative purposes. On the other hand, if different values of stress (or rate of mechanical work) can yield grain size populations with the same mean but different standard deviations you cannot use this single value for comparative purposes. We know that a normal distribution of grain size population is unrealistic in dynamically recrystallized mylonites, but in case you consider a log-normal distribution of grain size, the same general principle applies. Because some authors in the past were capable of established that the grain size population show no changes in the mean grain size during the ongoing deformation (e.g. Means, 1983; Barnhoorn et al. 2004; Stipp et al. 2006) or found an empirical relation between the stress applied and the grain size in mylonites using a single measure of grain size, we consider to be highly probable that for a single value of stress (or rate of mechanical work) only a unique mean (or median or frequency peak) grain size is obtained. We admit that to date this is just a hypothesis that awaits further testing. However, if this assumption were not true no paleopiezometric estimate done so far would be reliable, which seems rather unlikely given the amount of available experimental studies in alloys, ceramics and rocks that established an empirical relation between different values of mean grain size and stress.

Consequently, the answer to the second question about if it is time to move on is partly a YES, in the sense that we need put more effort in the methods to describe better the full grain size populations of grains because they have a broader scale of applications. But also a NO, because for paleopiezometry (-wattometry) studies we still need a single measure of grain size. Moreover, even if it is demonstrated in the future that the mean grain size (or similar) is not a valid measure, we will still need a single value for this type of studies and the big question will be: which one?

"The mentioned simple procedure of obtaining the grain size number via the presented technique is apparently not that simple because it is actually the list of numbers corresponding to areas of individual grains that needs to be imported into the script. However, it is usually the production of these numbers that is time consuming and demands several steps in the procedure. The subsequent statistical treatment of the grain areas in many cases is not demanding at all (matlab, python etc.). The development of this new technique of grain size determination is also not a step forward when compared to the existing techniques (such as StripStar and CSDcorrections) as demonstrated by the table 3. The table 3 presents the best estimation on grain size numbers obtained by the three techniques but basically the numbers range between 33.5 to 34.3 microns. Moreover the theory used to explain the technique is already well known and described elsewhere."

We think the reviewer is mixing three different issues here.

In the first place we agree with the reviewer that the generation of grain boundary outlines is the most time consuming step and probably the step more prone to introduce errors. We did not address this issue at all in our manuscript, partly because there are very nice textbooks dealing with this issue, such as Heilbronner and Barret (2014).

Secondly, the reviewer claims that the subsequent treatment of data, which is what the GrainSizeTools scrip was written for, can be done in others environments and is not demanding at all. As previously indicated, it is true that the treatment of data can be done in others environments but the questions about whether this data treatment is demanding or not is relative. Not all the people have programming skills or know how to implement a kernel density estimator or just and area-weighted plot in Excel, to give an example with a widely used computer program. In fact, in the first place we created our own code because we need to create number and area-weighted plots with Gaussian kernel density estimator and algorithms to set the optimal bin sizes or bandwidth for the study and, to be honest, we did not find scripts that met our requirements, despite the reviewer claims that they already exist. In the end, we decided that it would be useful to release the script implemented to perform the study to the public, in the way that it can be used by people without programming skills just following a brief tutorial or to advanced users with the aim of reuse, modify or extend the code for their own purposes. Finally, the reasons why we choose to implement the new script in Python and not to continue the development of previous scripts, such as StripStar, are clearly portrayed in the new version of the manuscript.

Finally, the reviewer claims that the GrainSizeScript script is not a step forward when compared to the StripStar script of the CSDCorrections application and that they produced similar results as demonstrated in our table 3 within the manuscript. In the first place, they have different goals, StripStar and CSDCorrections were implemented to derive the actual grain size from the population of apparent grains, while GrainSizeTools was implemented obtain a single measure of grain size for paleopiezometric studies and is, therefore, more focused on reproducibility. This is explicitly mentioned within the manuscript. Regarding the comparison between them and the results showed in table 3, this comparative was carried out to demonstrate that the peak of the Gaussian KDE produced comparable results, within the error, to the frequency peak of the actual (3D) distribution of grains. This has implications since it indicates that if you use the frequency peak as a measure of grain size, 2D and 3D grain size populations yields comparable results. On the other hand, the results showed in table 3 do not indicate that the three scripts/applications are equally suitable in case you want to use the frequency peak as estimator for paleopiezometric studies. As demonstrated in the manuscript, the use of the Gaussian KDE over histograms to find the frequency peak grain size increases reproducibility. Only the GrainSizeTools script implemented this approach. Both, StripStar and CSDCorrections, ask the user to define the number of classes and this user-defined step will penalizes reproducibility across studies. In fact, by using different number of classes with StripStar and CSDCorrections it was found that the estimations of the frequency peak grain size varied between 7 and 10 microns (see ranges in table 3), which implies errors up to 18 % respect to the average of the mean grain sizes used as reference.

Interactive comment on Solid Earth Discuss., 6, 3141, 2014.