Supplement — Introducing PebbleCounts: A grain-sizing tool for photo surveys of dynamic gravel-bed rivers

Benjamin Purinton[1] and Bodo Bookhagen[1]

[1]Institute of Earth and Environmental Science, Universität Potsdam, Potsdam, Germany

Correspondence: Ben Purinton (purinton@uni-potsdam.de)

S1. Command-line Variables for *PebbleCounts*

Table S1 shows the command-line variables for *PebbleCounts* (KMS approach) and Table S2 shows the command-line variables for *PebbleCountsAuto* (AIF approach).

Table S1. Command-line variable flags in *PebbleCounts* and their meaning. The default values are effective for most images.

Variable Flag	**Meaning (units)**	**Default Value(s) and Suggested Range**
im	Image to run, including path to folder	No default
ortho	Georeferenced orthoimagery flag	No default, 'y' for orthoimagery, 'n' for top-down
input_resolution	Input resolution if not orthoimage (mm)	No default, calculate from eq. (3)
subset	Interactively subset image	Default no ('n')
*sand_mask**	Name, including path, to a sand mask if one already exists	No default
*otsu_threshold**	Percentage of Otsu value to threshold shadows by (percentage of 100)	No default, suggested value of 50
*maxGS**	Expected maximum a-axis grain size (m)	Default 0.3
*cutoff**	Minimum b-axis length to be counted (pixels)	Default 20, can be raised
*min_sz_factors**	Factors to multiply *cutoff* at each scale, used to cleanup masks for easier clicking	Default [50, 5, 1] for three scales (large to small) for $\sim$1 mm/pixel imagery, double for < 0.8 mm/pixel
*win_sz_factors**	Factors to multiply *maxGS* by at each scale	Default [10, 3, 2] for three scales (large to small), can be changed $\pm$0.5–1.5 to get more or less windows
*improvement_ths**	Improvement threshold values that tell k-means when to halt (fraction of 1)	Default [0.01, 0.1, 0.1] for three scales (large to small), can be varied from 0.01–0.2
*coordinate_scales**	Fraction to scale *x,y* coordinates (fraction of 1)	Default [0.5, 0.5, 0.5] for three scales (large to small), can be varied from 0.3–0.7
*overlaps**	Fraction of overlap between windows (fraction of 1)	Default [0.5, 0.3, 0.1] for three scales (large to small), can be varied from 0–0.5 at each scale
*first_nl_denoise**	Strength of first non-local means denoising	Default 5, can be varied $\pm$1
*nl_means_chroma_filts**	Strength of windowed non-local means denoising	Default [3, 2, 1] for three scales (large to small), can be varied $\pm$1
*bilat_filt_szs**	Size of bilateral filtering windows (pixels)	Default [9, 5, 3] for three scales (large to small), can be varied from 3–9
*tophat_th**	Upper percentile threshold to take from top-hat filter for edge detection (fraction of 1)	Default 0.9, can be varied from 0.8–0.95
*sobel_th**	Upper percentile threshold to take from sobel filter for edge detection (fraction of 1)	Default 0.9, can be varied from 0.8–0.95
*canny_sig**	Canny filtering sigma value for edge detection	Default 2, can be varied from 1–2
resize	Value to resize windows by (fraction of 1)	Default 0.8, can be varied from 0.5–0.99 if you want a smaller (0.5) or larger (0.99) pop-up window

*Influence on results

Table S2. Command-line variable flags in *PebbleCountsAuto* and their meaning. The default values are effective for most images.

Variable Flag	**Meaning (units)**	**Default Value(s) and Suggested Range**
im	Image to run, including path to folder	No default
ortho	Georeferenced orthoimagery flag	No default, 'y' for orthoimagery, 'n' for top-down
input_resolution	Input resolution if not orthoimage (mm)	No default, calculate from eq. (3)
subset	Interactively subset image	Default no ('n')
*sand_mask**	Name, including path, to a sand mask if one already exists	No default
*otsu_threshold**	Percentage of Otsu value to threshold shadows by (percentage of 100)	No default, suggested value of 50
*cutoff**	Minimum b-axis length to be counted (pixels)	Default 20, can be raised
*percent_overlap**	Maximum allowable overalp between neighboring ellipses for filtering suspect grains (percentage of 100)	Default 15, can be varied from 5–30
*misfit_threshold**	Maximum allowable misfit between ellipse and grain mask for filtering suspect grains (percentage of 100)	Default 30, can be varied from 10–50
*min_size_threshold**	Minimum area of grain, used to clean the mask (pixels)	Default 10 for $\sim$1 mm/pixel imagery, 40 for $<$ 0.8 mm/pixel
*first_nl_denoise**	Strength of first non-local means denoising	Default 5, can be varied $\pm$1
*tophat_th**	Upper percentile threshold to take from top-hat filter for edge detection (fraction of 1)	Default 0.9, can be varied from 0.8–0.95
*sobel_th**	Upper percentile threshold to take from sobel filter for edge detection (fraction of 1)	Default 0.9, can be varied from 0.8–0.95
*canny_sig**	Canny filtering sigma value for edge detection	Default 2, can be varied from 1–2
resize	Value to resize windows by (fraction of 1)	Default 0.8, can be varied from 0.5–0.99 if you want a smaller (0.5) or larger (0.99) pop-up window

*Influence on results

S2. GSDs Separated by Site

Figure S1 and S2 show the results of the KMS approach to grain sizing on a site-by-site basis before and following a 20-pixel lower truncation of the distribution. On the left columns are the initial runs of KMS *PebbleCounts* without any lower truncation, and on the right columns are the re-running with a 20-pixel truncation. Aside from three sites (S09, S10, and S24) the p-values for the KS-tests without truncation were < 0.01, indicating a robust rejection of the null hypothesis. Following truncation, all 12 sites have p-values > 0.1, except S16, which has p=0.09.

Despite a high p-value, S24 demonstrates a stronger bias in the GSD towards coarser grains (up to 0.5 ψ discrepancy), as indicated by the high A_{diff} value of -1.36. Here, the KS-test pass is likely caused by the small sample size remaining after truncation (n=24), the least of any site. The poor performance of S24 was expected given the large size range with many sub-cm pebbles and a few large boulders, strong cast shadows from the large grains, and intra-granular edges on angular boulders with quartz veins (see Figure 9b in the main manuscript). Importantly, S24 is the only site not from a major river stem, but rather from a debris-flow fan draining a small tributary catchment in the Quebrada del Toro. The site S34 also had a high A_{diff}=-2.11. In this case, poor performance is due to significant blurriness of this image, and again a small sample size (n=47).

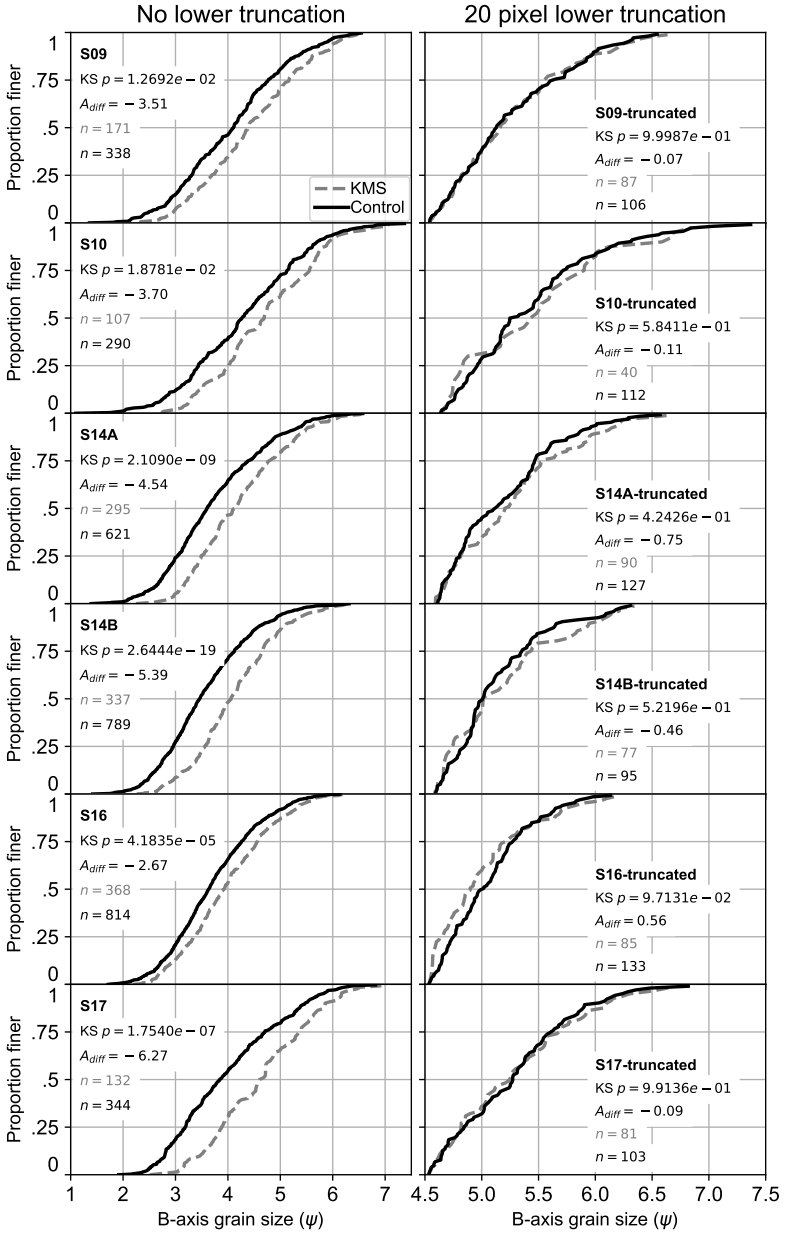

Figure S1. Results from KMS *PebbleCounts* on a site-by-site basis with the initial run in the left columns (no truncation) and the truncated (at 20-pixels) run in the right columns. The control data are given as a solid black line with the number of pebbles (n) shown in black. KMS *PebbleCounts* results are in gray and dashed. The p-value results of a KS-test are also shown. A_{diff} is the approximate integral between the curves. Note the reduction in x-axis scale between the columns, where the right, truncated distributions are plotted on a narrower range to emphasize the remaining discrepancies.

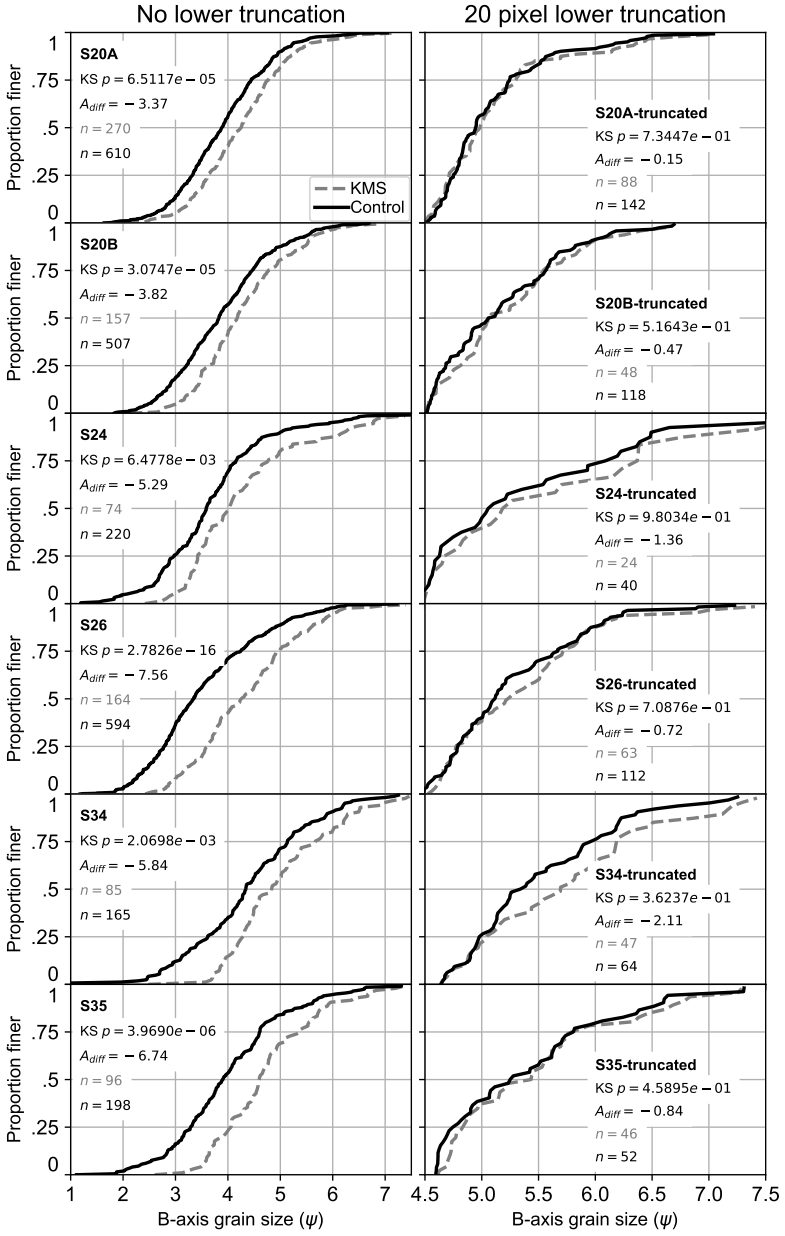

Figure S2. Results from KMS *PebbleCounts* on a site-by-site basis with the initial run in the left columns (no truncation) and the truncated (at 20-pixels) run in the right columns. The control data are given as a solid black line with the number of pebbles (n) shown in black. KMS *PebbleCounts* results are in gray and dashed. The p-value results of a KS-test are also shown. A_{diff} is the approximate integral between the curves. Note the reduction in x-axis scale between the columns, where the right, truncated distributions are plotted on a narrower range to emphasize the remaining discrepancies.

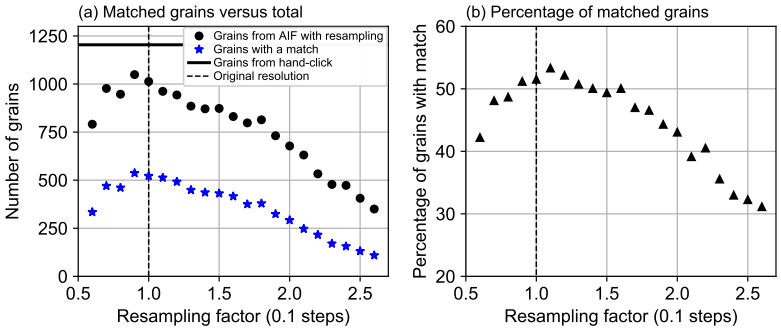

Figure S3. Matching grains found in each filtered mask versus the resampling factor (where 1 is the original image) for the ∼1.16 mm/pixel resolution images. Matches are defined as an AIF grain within 5 pixels of the hand-clicked line or the KMS grain centroid and with a 1 cm maximum b-axis difference between the AIF grain and the match.

S3. Resampling and Parameter Selection in AIF Approach

Figure S3 demonstrates the percentage of grains with a match found in the AIF approach (matches are defined as in main manuscript Figure 18) when increasing resampling from a factor of 0.6–2.6 by 0.1 steps using Lanczos resampling (Lanczos, 1950). As the resampling factor increases, there is progressive reduction in the number of found grains after filtering, therefore we selected the original resolution (resampling factor of one). Figure S4 and Figure S5 demonstrate two cases where the resampling slightly improved the resulting GSD. Both images were of relatively low quality with significant blurring and the presence of many weak edges between grains of similar color.

We selected a maximum percent misfit between the ellipse and grain of 30% as the 90^{th} percentile of misfits for the KMS approach was 30%. Furthermore, we allowed a maximum overlap between neighboring ellipses of 15%, visually selected to minimize overlapping grain measurement and over-segmentation of discrete grains. For the higher resolution imagery it was necessary to use a lower sobel and top-hat threshold (0.85), since we consider all the edges at once in the AIF approach, rather than in a windowed subset as in the KMS approach, and many edges are not found when using the 0.9 threshold given the increased number of pixels under consideration.

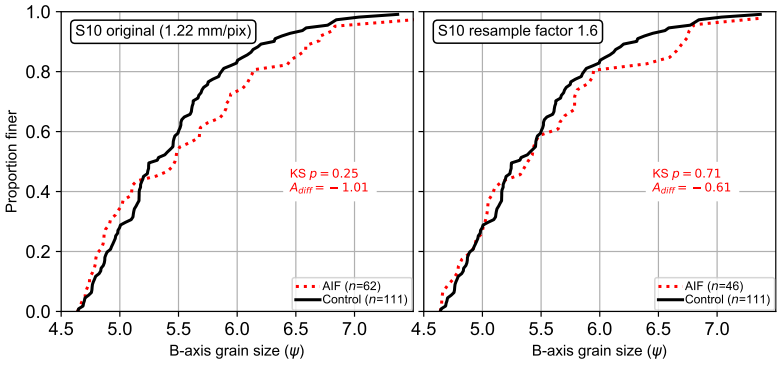

Figure S4. Slight improvement (increase in p and decrease in A_{diff}) in result using a 1.6-times resampling factor prior to running the AIF algorithm for the difficult (somewhat blurry, weak edges) S10 orthoimage.

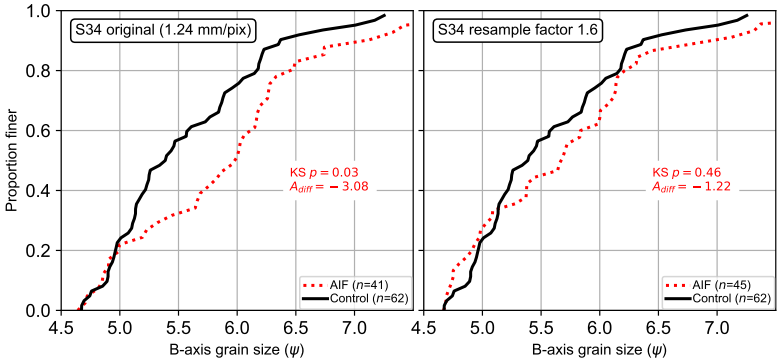

Figure S5. Slight improvement (increase in p and decrease in A_{diff}) in result using a 1.6-times resampling factor prior to running the AIF algorithm for the difficult (very blurry, weak edges) S34 orthoimage.

References

Lanczos, C.: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Natl. Bur. Stand. B, 45, 255–282, 1950.