

Supporting Information for “From regular to random: a unifying framework for step-pool spacing”

Christian M. Erikson¹ and Jens M. Turowski¹

¹*GFZ Helmholtz Centre for Geosciences, 14473 Potsdam, Germany*

Minimum spacing adjustment routine

To generate the “regular reference line” we added Gaussian noise to a initially regular sequence. In the case where adding Gaussian noise caused the minimum spacing constraint to be violated, we modified the sequence by first identifying the step immediately upstream of the step in violation. We then calculated the distance between that upstream step and the immediately downstream step of the step in violation. If the distance between these steps lacked sufficient space to reposition the step in violation with a valid minimum spacing, we repeated the search process by moving one step farther downstream until the distance between the original upstream step and the current downstream step was greater than or equal to twice the minimum spacing times the number of steps in violation of the minimum spacing constraint. That is, we found the next downstream step with enough space to reposition all steps between it and the upstream step with a valid minimum spacing between all steps.

When a downstream step to reposition the steps in violation existed, we divided the available space, accounting for one minimum spacing on either end, by the number of steps needing to be repositioned to define an average allowed interval length. We then calculated the position in the middle of the allowable region. Finally, we repositioned the original step in violation by placing it at this middle point and then adding Gaussian noise. We repeated this

process for each downstream step until all steps met the minimum spacing condition, ensuring that all violating steps were validly repositioned.

In the case that there was no downstream step with sufficient space for all steps in violation to be redistributed, we removed the final step from the sequence and attempted to use the newly created available space to reposition the remaining steps in a manner satisfying the minimum spacing requirement. Removal of the final step from the sequence did not always provide sufficient space for the minimum spacing constraint to be satisfied. When this occurred, we propagated spacing from the downstream point in a regular manner until all step spacings met the minimum spacing requirement. In some cases, this process resulted in a fully regular final spacing.

Alternative regular reference lines

We created an alternative algorithm for generating a regular reference line in order to compare it to our primary algorithm. In the alternative algorithm, we placed a step within a given interval where the location within the interval was determined by randomly sampling from a Gaussian distribution. For each placed step, we then subdivided the initial interval into two subintervals. We assigned a probability of placing the next step in the new subintervals proportionally to the length of each interval. We then repeated this process, continually creating new subintervals and updating the probability. Because the resulting step spacing was a result of two random processes, the final step spacing was always approximately normally distributed, consistent with the central limit theorem. We again repeated this process as part of a Monte Carlo simulation to estimate bounds on the alternative line (Fig. S2, gold line). We also used this

alternative regular reference line for calculation of the range normalized coefficient of variation (CV_N) (Fig. S3).

In addition to the alternative algorithm, we also considered the effect of Gaussian distributions with different standard deviations on our primary algorithm. We show two possibilities using standard deviation equal to one-third (dark blue) and one-sixth (light blue) of the mean (Fig. S2). These fractions of the mean were chosen so that the occurrence of steps shifting beyond the original position of the adjacent step after adding noise was rare, $\sim 0.3\%$ probability for one-third of the mean (3σ for a Gaussian distribution) and $\sim 2 * 10^{-7}\%$ probability for one-sixth (6σ). Nonetheless, we accounted for the cases when extreme shifts did occur using the same procedure outlined in the Methods. The line generated with the smaller standard deviation has the effect of making sequences seem less regular than the primary regular reference line (Fig. 2, dark blue line).

Observation sensitivity

Because the number of observations in a sequence ranged from three steps to more than 200 steps, we used subsampling to determine the effect of sample size on plotting position within the diagnostic diagram. We subsampled the “Erlenbach 2013b” dataset (Erikson, Turowski, et al., 2026), which consisted of 120 spacings, in two ways. First, we randomly sampled individual spacings. Second, we randomly selected a continuous block within the full reach. In both cases, we changed the number of sampled steps over a range from three to 100 and repeated each sample number 200 times to estimate the mean and variance. We compared the results of both sampling routines to the diagnostic diagram colored by observation number (Fig. S1).

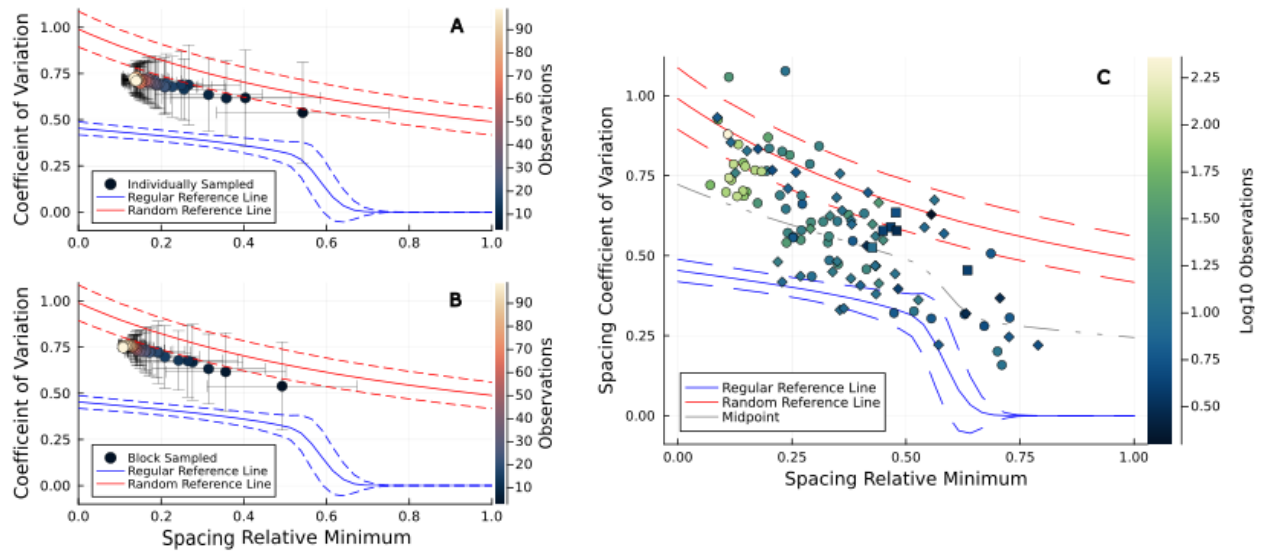


Fig S1: Results of a subsampling analysis on the “Erlenbach 2013b” survey from Erikson, Turowski, et al. (2026) using individual random sampling (A) and randomly sampling continuous blocks (B). The number of observations for each sequence in the full dataset are indicated by color in the same plotting space as Fig. 2 (C).

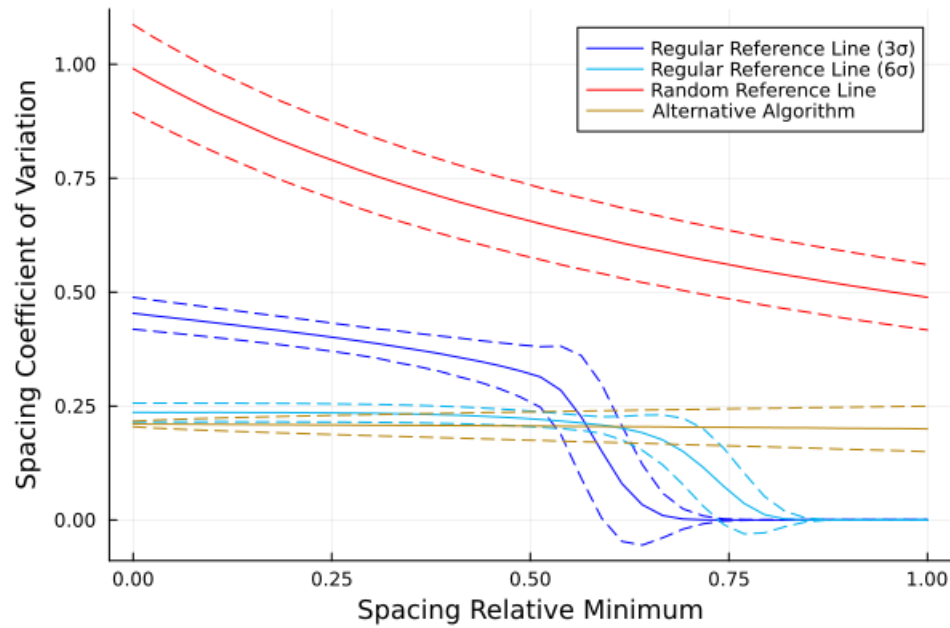


Fig S2: Reference lines using various algorithms. The red and dark blue lines are the random and regular reference lines as in Fig. 2. The light blue line uses the same algorithm as the dark blue regular reference line but with a smaller standard deviation in the Gaussian distribution from which noise was added. The gold line is an alternative algorithm in which steps are sequentially added to an initially empty sequence with a probability of placement proportional to the size of the empty interval.

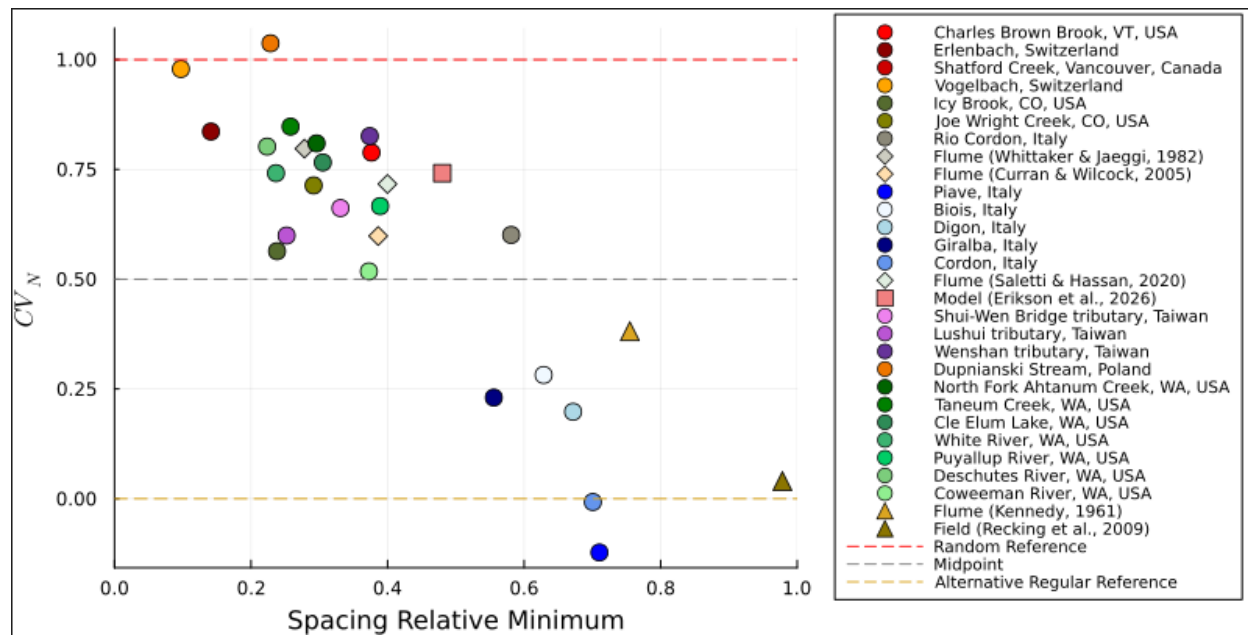


Fig S3: Source-averaged coefficients of variation relative to the space between reference lines (CV_N). The upper and lower dashed lines correspond to the random and regular reference lines and the dashed gray line indicates the middle point between them. The regular reference line is generated with the alternative algorithm where steps are added to open intervals with a probability proportional to the length of the interval.