Flood-driven sediment dynamics on gravel bars mapped using UAV photogrammetry and machine learning: Sense River, Switzerland.

Nahin Rezwan[1], David Mair[2], Alexander C. Whittaker[1], Fritz Schlunegger[2], Ariel Henrique do Prado[2], Sadia Hossain Setu[3]

Abstract

Gravel-bed rivers are shaped by complex interactions between hydrological forcing, sediment sorting, and channel morphology, yet fine-scale, spatially continuous observations of these processes remain rare. We combine UAV structure-from-motion photogrammetry with machine-learning grain segmentation to quantify flood-driven sediment redistribution in a minimally disturbed gravel-bed river (Sense River, Switzerland). Two surveys of four gravel bars (2021 and 2024) mapped individual clasts in images at centimetre resolution, allowing spatial and temporal analysis of grain-size patterns. We show clear intra-bar fining from crests to tails and a reach-scale morphology control on sorting: bend-associated bars are moderately to well sorted, while straight reaches are more poorly sorted. Grain-size distributions converge to self-similar forms across all bars, with analysis of ca. 1.86 million grains providing unprecedented empirical validation of scale-invariant sorting, an improvement by orders of magnitude over conventional pebble counts. To understand the detailed hydraulic controls during the moderately large flood captured between surveys (ca. 180 m³/s; ca. 2-10 years recurrence), we performed detailed hydraulic modelling for one bar, estimating spatial fields of shear stress, Shields parameter, and stability conditions during the flood. We also differentiated the topography between the two surveys to map the relative elevation change. The crest and margin armour remained largely stable, whereas the tail part was extensively reworked. A hydraulically driven mobilisation model reproduced observed mobility with ca. 65% overall accuracy (up to 82% in tails) but under-predicted movement on crests. We also show that where floods were large enough to mobilise the grains, coarse patches were rapidly buried or completely replaced, demonstrating that local hydraulic geometry can override patch stability. Overall, bar adjustment was deposition-dominated for that bar, consistent with the waning stage of the flood, during which reduced shear stresses promoted net deposition. Our data indicates that flows <200 m³/s can remobilise large bar areas, and analysis of gauging data for the Sense River since 1928 shows that such events are becoming more frequent. Our results highlight the important geomorphic role of moderate to moderately large floods in such rivers and demonstrate high-resolution, hydraulically informed grain mapping as a robust tool for predicting gravel-bed river response under changing flood regimes.

Response to Reviewer 3: General and Specific Comments

Reviewer comments are indicated in regular and black text, while our responses are formatted as blue text

We sincerely thank Reviewer 3 for the detailed and thoughtful evaluation of our manuscript. We appreciate the reviewer's careful reading and the recognition of the value of the dataset, particularly the integration of UAV photogrammetry with automated grain segmentation and the unusually dense grain-size dataset. The reviewer raises several important conceptual and methodological issues, particularly regarding the need for a clearly formulated working hypothesis, more precise use of geomorphological terminology, stronger methodological justification, and improved figure clarity.

We agree that addressing these points will significantly strengthen the manuscript. In the revised version, we therefore clarify the conceptual framing of the study, standardise terminology, expand methodological explanations, and improve the presentation of figures.

Below we respond to the specific comments in detail.

1. Lack of a clearly defined working hypothesis
The manuscript does not clearly formulate a working hypothesis regarding sediment transport or morphological changes. Although the dataset and method are innovative, the manuscript does not fully exploit the potential of such high-resolution data to test specific process-based hypotheses.

We agree that we want readers to understand the conceptual framing of the study as clearly as possible. In the revised manuscript we therefore restructure the end of the Introduction to state the aim, objectives, and working hypothesis of the study explicitly. The aim of the paper is to understand how a moderately large flood reorganises the surface texture and morphology of gravel bars using spatially continuous grain-size mapping and high-resolution topographic data. To address this aim, we map grain-size distributions across gravel bars, quantify topographic change associated with the flood, and evaluate how spatial variations in hydraulic competence relate to observed patterns of sediment adjustment. In doing so, we test the hypothesis that the response of gravel bars to a high-flow event is spatially heterogeneous and influenced by the pre-existing bar geometry, which controls the distribution of hydraulic forcing across the bar surface. While we did articulate some of these ideas in the existing version of the manuscript, we very much see the value in spelling this out together for readers and we thank the reviewer for their comment.

2. Definitions of key geomorphological and sediment-transport concepts

(a) Confusion between spatial scales

There is confusion between grain-scale observations and tile-scale analysis. Although grains are segmented individually, the analysis is aggregated into 25 m² tiles, meaning that interpretation occurs at the patch scale rather than at grain scale.

We agree with the reviewer that this distinction needs to be clarified more explicitly. In the revised manuscript we will introduce clear definitions of spatial scales used in the analysis, including grain, tile (5 m × 5 m = 25 m²), bar unit (crest, middle, tail), bar, and reach.

We emphasise in section 3.3 of our revised manuscript that although individual grains are segmented from imagery, the statistical analysis is conducted at the tile scale. Consequently, the results represent patch-scale textural variability rather than grain-scale transport processes. The revised text clarifies this distinction between the resolution of grain-size measurement and the scale at which geomorphic inference is made.

(b) Use of the term "mobility"

The manuscript uses terms such as mobility, remobilisation, and clast-level mobility, although the methodology does not track individual grains or measure transport directly.

We thank the reviewer for this important comment. We wish to avoid any suggestion that implies direct observation of individual grain-scale transport, which is beyond the scope of the present dataset. We do not track individual grains or reconstruct clast trajectories between surveys, and we will make this limitation clearer in the revised manuscript.

At the same time, we note that the terminology of sediment mobility is commonly used in fluvial geomorphology when interpreting the potential for entrainment or reworking from hydraulic proxies such as dimensionless shear stress and critical threshold exceedance, as well as from observed surface and topographic change (e.g., Buffington and Montgomery, 1997; Wilcock and Crowe, 2003; do Prado et al., 2025). Such studies also do not track individual grains but they do freely talk about sediment mobility.

In that sense, our analysis does not measure individual grain motion directly, but it does evaluate spatial patterns of likely sediment disturbance using a combination of grain-size change, DoD-derived morphological adjustment, and modelled hydraulic competence. In addition, repeat imagery allows us to identify areas where the bar surface remained stable and others where substantial reworking occurred, even if individual clasts are not tracked.

To clarify this distinction, we revise the text in the manuscript to avoid terms that imply direct clast-by-clast tracking, while retaining the broader geomorphic interpretation of mobility where it refers to inferred sediment reworking potential.

3. Methods

3.1 Photogrammetry and topographic data

The absence of ground control points (GCPs) is not necessarily problematic but should be discussed, along with point-cloud alignment and uncertainty estimation.

We are happy to provide more information and a clear discussion of our UAV-SfM results. The UAV surveys were conducted using an UAV equipped with an on-board GNSS system that is capable of in-flight RTK corrections, which significantly improves georeferencing accuracy compared with standard GPS positioning. As a result, the photogrammetric models were of high precision and accuracy (see related response to RC1 above). Similar setups have been demonstrated to result in high-quality topographic point clouds without requiring ground control points (e.g., James et al., 2017).

In the revised manuscript we describe our image acquisition and photogrammetric processing in more detail in section 3.3, we provide SfM model accuracy and uncertainty results in the supplementary material, and now we present result-relevant information of these in discussion section 4.4.

The use of the 75th percentile method for DEM normalisation is insufficiently explained.

We employed a simple subtraction of the 75th percentile value here to normalise the DEM elevation to calculate the relative difference between the DSMs subsequently. The references were intended to provide further information on more complex workflows, and our approach is not intended as a probabilistic uncertainty model. In the revised manuscript, we clarify that the 75th-percentile approach was used solely to remove systematic vertical offsets and between independently reconstructed DSMs. We have revised 3.3 section to communicate this clearly.

DSM uncertainties and the justification of the ±0.06 m threshold are not clearly explained.

We consider the survey noise (and referencing errors) by estimating the uncertainties in the sparse point cloud from the SfM model (details added in the supplementary material). Due to the online and in-flight RTK corrections, we could achieve high positioning accuracies during the surveys. This allowed us to obtain precise SfM models with average positioning errors < 2 cm for cameras in z direction (total 3D RMSE errors of < 3.5 cm), and an average sparse point cloud precision of < 3 cm (see supplementary material and response to the next comment below). These sparse cloud errors can be considered the main source of DSM noise (e.g., James et al., 2020), aside from external referencing errors. We add this information and discussion in sections 3.1.2 and 4.4, and we include references to more sophisticated uncertainty estimation methods, i.e., Wheaton et al. (2010) and Bailey et al. (2020). Therefore, the ±0.06 m threshold corresponds to approximately twice the estimated uncertainty and was adopted as a conservative threshold to identify significant elevation changes. Only elevation differences exceeding this threshold are interpreted as significant morphological change.

3.2 Grain detection

How are grains located on tile borders handled?
We agree that this requires clarification. The image segmentation algorithm filters out grains for which the centre-points fall in the outermost 10% of each image tile. Thereby, grains with intersecting tile boundaries ¬are in all likelihood excluded by default in our setup (only boulders larger than 1m could be still cut, theoretically). Statistical descriptors (D16, D50, D84) are derived only from grains that fall into the central 64% of each image tile, which are therefore very likely to be fully segmented grains. This clarification is being added to the Methods section. In addition we have updated the text in section 3.2 and figure 3e to communicate our approach clearly.

3.3 Hydraulic calculations
The presentation of hydraulic metrics is confusing, particularly regarding how stability is evaluated and which grain-size percentile is used.
We agree that the presentation of the hydraulic metrics could be clearer and we are very happy to clarify our hydraulic metrics. In the revised manuscript we clarify the calculation of hydraulic parameters and the criteria used to evaluate sediment stability.

Sediment stability is assessed using a threshold-based interpretation derived from the Shields criterion. Specifically, the dimensionless shear stress (τ^*) calculated for each grid cell is compared with a critical entrainment threshold ($\tau c = 0.06$) representing the onset of gravel motion (e.g., Shields, 1936; Buffington and Montgomery, 1997). Where τ exceeds this threshold, sediment is interpreted as potentially mobile, whereas values below this threshold indicate conditions where the surface material is expected to remain stable under the modelled hydraulic forcing.

In addition, we estimate the hydraulically required grain size for stability by rearranging the Shields relation. This represents the grain size that would be required for sediment to remain stable under the modelled hydraulic conditions. Stability calculations are based on the median grain size (D50). See also our detailed response of how we handle the related questions to this posed by reviewers 1 and 2.

4. Spatial variability of sorting
Sorting is not clearly defined, and mapping sorting parameters could provide additional insight into spatial sediment variability.

We appreciate this suggestion and agree that sorting metrics could provide additional information about spatial sediment organisation. In the current analysis we focus primarily on spatial patterns of median and coarse grain sizes (D50 and D84) and their relation to hydraulic forcing. In sedimentology, sorting refers to the dispersion or spread of grain sizes within a grain-size distribution, typically quantified using statistical measures derived from the full grain-size distribution (e.g., Folk and Ward, 1957). In our analysis, sorting is interpreted qualitatively from the shape and steepness of the cumulative grain-size distribution curves, where steeper curves indicate narrower grain-size distributions and broader curves indicate greater dispersion of grain sizes. This definition and clarification have been added in Section 4.2 of the revised manuscript, where we also distinguish between sedimentological sorting and spatial variability of grain-size statistics across the bar surface.

5. Representativeness of Bar 3

Bar 3 shows the least change among the four bars, raising questions about its representativeness.
We appreciate this observation. Bar 3 was selected for detailed hydraulic analysis because it retained a well-preserved morphology and grain-size structure following the flood, allowing a reasonable comparison between pre- and post-event conditions. In contrast, several of the other bars experienced extensive morphological reworking during the flood, with parts of the bar surface being substantially reshaped or removed, which we mentioned in section 4.3 of the original manuscript and Figure 6. This made it difficult to perform a consistent hydraulic comparison between surveys for those locations.

Bar 3 therefore provides a suitable surface for evaluating how spatial variations in bar geometry and slope relate to hydraulic competence and sediment stability during the flood event. The aim of the analysis is not to present Bar 3 as the most dynamically active bar in the reach, but rather as a location where the relationship between pre-existing bar morphology, hydraulic forcing, and observed sediment adjustment can be evaluated in a consistent way. In the revised manuscript we further clarify this rationale in Section 4.3.

6. Lack of information on bar and reach configuration

Important information about channel and reach morphology is missing, including channel width, slopes, and characteristic discharges.
We agree that additional context is necessary. In the section 2 of the revised manuscript, we are including the further information describing the channel and reach configuration, including characteristic discharge metrics.

7. Representativeness of 25 m² tiles and use of the term “fine scale”

The term “fine scale” may be misleading depending on river size, and the choice of tile size should be discussed.

We appreciate this comment and agree that the terminology should be used more carefully. In the revised manuscript we moderate the use of the term “fine scale” and clarify that the spatial resolution of analysis is defined by the tile size (5 m × 5 m = 25 m²).

The tile size was selected to balance two considerations: (i) ensuring that each tile contains a sufficiently large number of grains for reliable statistical analysis, and (ii) maintaining spatial resolution capable of capturing variability across bar units. Given that the largest clasts in the study reach approach ca. 1 m in diameter, a 5 m grid was considered appropriate to capture local grain-size variability while maintaining robust sample sizes. We are clarifying this in the section 3.2 of the revised manuscript.

In section 3.2 of the revised manuscript, we also discuss how tile size could be adjusted in future studies depending on river scale, sediment size, and computational constraints. In this work, we have produced a nearly continuous grain-size distribution map which demonstrates that even changing tile size will not significantly change the results. It means that the magnitude could have been changed a bit but the grain-size distribution pattern and the fundamental results should be the same.

8. Figure clarity and design
Several issues with figure readability, including small text, inconsistent scale bars, colour reuse, and missing flow-direction indicators.

In the revised manuscript we increase text sizes, standardise colour schemes across figures, and ensure consistent scale bars. We also revise panel layouts and legends to improve readability and reduce visual clutter.

We thank Reviewer 3 again for the detailed and constructive feedback. Addressing these comments will significantly strengthen the conceptual clarity, methodological transparency, and presentation of the manuscript.

We thank the reviewer for the additional minor and line-by-line comments. These have been carefully addressed throughout the revised manuscript and the text has been edited accordingly to improve clarity and consistency.

References:

Buffington, J. M., & Montgomery, D. R. (1997). A systematic analysis of eight decades of incipient motion studies, with special reference to gravel-bedded rivers. Water Resources Research, 33(8), 1993–2029. https://doi.org/10.1029/96WR03190

do Prado, A., Galeazzi, C. P., Mair, D., Garefalakis, P., de Almeida, R. P., Whittaker, A. C., & Schlunegger, F. (2025). Gravel bar shielding: A mechanism responsible for bar stability in gravel- and cobble-bedded streams. Earth Surface Processes and Landforms, 50(7), e70107. https://doi.org/10.1002/esp.70107

Folk, R. L., & Ward, W. C. (1957). Brazos River bar [Texas]; a study in the significance of grain size parameters. Journal of Sedimentary Research, 27(1), 3–26. https://doi.org/10.1306/74D70646-2B21-11D7-8648000102C1865D

James, M. R., Robson, S., and Smith, M. W.: 3-D uncertainty-based topographic change detection with structure-from-motion photogrammetry: precision maps for ground control and directly georeferenced surveys, Earth Surf. Proc. Land., 42, 1769–1788, https://doi.org/10.1002/esp.4125, 2017a

Phlip Bailey, Joseph Wheaton, Matt Reimer, & James Brasington. (2020). Geomorphic Change Detection Software (7.5.0). Zenodo. https://doi.org/10.5281/zenodo.7248344

Wheaton, J.M., Brasington, J., Darby, S.E. and Sear, D.A. (2010), Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. Earth Surf. Process. Landforms, 35: 136-156. https://doi.org/10.1002/esp.1886