The manuscript describes a method of assessing grain size distribution and parameterizing this characteristic, and presents a database of terrestrial, lunar and Martian regolith/soil samples using this method. The database itself is a unique and useful collation of potentially cross-referenceable data, and the method itself is well-described and has the potential to provide information that could be used to better understand regolith characteristics. I suggest the following to improve the usefulness of the work to the planetary community in assessing and comparing regolith across regions and bodies. Note that my expertise is in assessing grain sizes using images, so I have left a review of the mathematical aspects of this manuscript to others more capable of discussing them.

Reply: Thank you for your positive evaluation.

General

- Define specialized terms being used earlier in the manuscript:
- Line 46-47, 54: Specify what curves are being referred to; specify what is meant by "30 m".

Reply: In the revised manuscript, we have explicitly defined that the "curves" mentioned in the text refer specifically to grain size frequency distribution curves (showing the relative abundance of each particle size fraction). This definition has been added at the first occurrence of the term (Lines 53-54 in the revised manuscript), where we now write: "complete GSD curves (hereafter GSD 'curves' showing the relative abundance of each particle size fraction)." Furthermore, "30 m" refers to the spatial resolution of the POLARIS dataset (30 meters per grid cell).

- Lines 80-83: Lots of jargon here without definition. To make the manuscript more accessible across disciplines, consider defining these terms or at least pointing to a later Section and defining them there.

Reply: We agree that terms such as "cumulative curves", "Weibull statistics", "stochastic field generation", and "unified parameterization" may be unclear to readers unfamiliar with soil mechanics or geostatistics.

In the revised manuscript, we have made the following changes:

At the first occurrence of these terms (Lines 80-83), we have added brief clarifications and explicitly pointed readers to the later sections where these concepts are defined in detail. Specifically:

“complete raw cumulative curves” → revised to “complete raw cumulative GSD curves (mass percentage passing)” (Line 106 in the revised manuscript)

“Weibull statistics” → revised to “Weibull distribution statistics (see Section 2.4 for full definition, formula, and application to grain size analysis)” (Lines 107-108 in the revised manuscript)

“stochastic field generation” → revised to “stochastic field generation (see Section 2.5 for detailed methodology)” (Lines 108-109 in the revised manuscript)

- More details are needed regarding what and why specific assumptions were made in assessing actual geologic materials. For example, I see a “spherical grain” is assumed by the program. What is the logic behind this assumption other than it simplifies things for the program? If that’s the logic, provide the rationale as to why this is a reasonable assumption to make for this work. The limitations listed are not complete (e.g., as on the Moon, clay fractions aren’t resolved by any martian imager at this time, but this is not noted); nor do they appear to be the most salient limitations to “harmonization” (e.g., if clay fractions aren’t resolvable then why not just cut off the grain size distribution there?).

Reply: We thank the reviewer for raising these important methodological points.

Regarding the spherical grain assumption: We agree that the original statement “it simplifies things for the program” is insufficient justification. In fact, the spherical grain approximation is a standard practice in 2D image-based granulometry, supported by extensive literature. For instance, Pizzati et al. (2023) demonstrated that approximating particles as perfect spheres with a shape correction factor yields 2D-to-3D conversions that match laser diffraction measurements. This assumption has also been adopted in previous Martian grain-size studies (e.g., Lapotre et al., 2017). In the revised manuscript, we have added these citations (see Line 270 in Table 4) to justify the assumption. We also clarify that we apply the equivalent circular diameter (ECD) method, which is standard for sand-sized and coarser grains where shape deviation does not substantially bias the distribution statistics.

Regarding the unresolved fine fractions: The reviewer correctly points out that Martian imagers cannot resolve clay-sized particles (<2 μm) or even silt-sized grains in some cases. Following the reviewer's suggestion, we have revised our approach (see Lines 357-364) to explicitly truncate the grain-size distributions at the instrument-specific detection limits (Table 3). For each

Martian sample, only grains larger than the minimum detectable size (~3× pixel scale, i.e., 0.04–0.15 mm depending on the instrument) are retained for UGSD fitting. The truncated portion is not extrapolated or artificially filled; instead, we treat the cumulative percentages at the smallest size bin as lower bounds. This truncation is now clearly noted in Section 2.3.2 and the limitations section (Section 6.3). We agree that this is a more conservative and transparent approach than fitting across unresolved size ranges.

Regarding completeness of limitations: We have expanded Section 6.3 (Lines 751-812) to explicitly list the unresolved clay/silt fractions as a key limitation for Martian data. We have also added a discussion of why truncation is preferable to extrapolation. Specifically, we now state that because no Martian imager can resolve grains below ~40 μm, any fitting across that range would be unconstrained; we therefore truncate rather than extrapolate.

Pizzati, M., Mantovani, L., Lisotti, A., Storti, F., & Balsamo, F. (2023). Particle size distributions in Earth Sciences: a review of techniques and a new procedure to match 2D and 3D analyses. EGUsphere, 2023, 1-71. https://doi.org/10.5194/egusphere-2023-2636, 2023.

Lapotre, M. G., Ehlmann, B. L., & Minson, S. E. (2017). A probabilistic approach to remote compositional analysis of planetary surfaces. Journal of Geophysical Research: Planets, 122(5), 983-1009. https://doi.org/10.1002/2016JE005248.

- There is insufficient detail on the methodology for assessing grain size. What does semi-autonomous mean? What parameters were used to determine grain size? How were grains discriminated from composite grains (i.e. grains-within-grains)? How did the authors adjust methodology among the different types of images from Mars, particularly RMI (a fisheye-lens imager that on MSL is greyscale) and MAHLI/WATSON? MAHLI/WATSON in particular acquire a range of resolutions; which were used? How was assessment of two-dimensional images regularized with lunar dry sieving, and with the large number of laboratory techniques that may be brought to bear for Earth-based samples? Real-world limitations of the data with respect to resolution also should be discussed in more depth. MAHLI and WATSON have acquired images <1 cm from the surface, yielding better than 20 µm/pixel. However, a reasonable estimate of resolution is typically at least across 3 pixels (as noted in Table 3), meaning no better than ~60 μm/pixel. This assumes a nearly pristine target; with even a small covering of ubiquitous martian dust, that resolution degrades. The manuscript acknowledges the problem without discussing how it was addressed.

Reply: We thank the reviewer for these important technical questions.

Regarding "semi-automated": Our analysis workflow follows established photoanalytical methods (Shi et al., 2024; Zhao et al., 2023). "Semi-automated" means that while grain boundary detection and size/shape calculations are performed algorithmically (implemented in Mathematica/Python), several steps require manual intervention: (1) image quality filtering (blurred or out-of-focus regions are masked); (2) parameter tuning for foreground-background separation (thresholds adjusted per image due to variable illumination); (3) manual correction of erroneous grain merging. This hybrid approach balances reproducibility with the flexibility needed for heterogeneous Mars images. Detailed operator decisions are documented in Supplementary Text S3 (Lines 1958-1968).

Regarding parameters used to determine grain size: we fit each sample to the UGSD function, which characterizes the full GSD using three interpretable parameters: μ (power exponent correlated with fine-particle content), D_c (characteristic grain size, in mm), and n (shape exponent controlling coarse-tail steepness). These three parameters collectively capture the entire grain-size distribution curve in a compact form, enabling quantitative comparison across samples and planetary bodies. The UGSD function and parameter definitions are provided in Section 2.3.2 (Lines 350-396) and summarized in Table 5 (Line 396).

Regarding discrimination of composite grains (grains-within-grains): We have already considered this issue in our image analysis workflow. Specifically, our two-step protocol addresses composite grains as follows. First, the watershed segmentation algorithm separates touching grains based on local intensity minima. Second, each segmented image undergoes manual inspection. According to our established criteria: (1) small grains sitting on larger grains are counted separately if ≥ 50 % of their boundary is visible; (2) surface textures, cements, or alteration rinds are not counted as separate grains; (3) images with foreground coverage > 5 % (e.g., granules overlying fine matrix) are excluded following Shi et al. (2024). In the revised manuscript, we have elaborated this protocol in Section 2.2.3 and Supplementary Text S3 (Lines 1243-1252) to provide a more complete and transparent methodological description.

Regarding instrument-specific adjustments: MAHLI/WATSON (color, arm-mounted, 14-77 μm/pixel) and RMI (grayscale, mast-mounted, ≈ 40 μm/pixel at 2 m, with fisheye distortion) require different processing. For MAHLI/WATSON, we use the highest resolution (14-21 μm/pixel), manually mask out-of-focus areas, and apply edge detection with adaptive thresholds. For RMI, we analyze only the central 40 % of the field of view to minimize distortion (< 2 %) and apply histogram equalization for contrast enhancement. Cross-instrument validation on 30 co-located targets shows D_{50} agreement within ±12% (Supplementary Material Text S5.2, Lines 1979-1991). These instrument specifications have been added to Table 3 (Line 258).

Regarding resolution selection: We exclusively use the highest-resolution images available for each target. For MAHLI, this means working distances of 2-10 cm, yielding pixel scales of 14-32 μm/pixel. The specific focus motor count (which inversely correlates with working distance) is recorded for each image to enable scale calculation ($p = 6.9 + 3.52 \times w$, where $p = \mu$m/pixel, w = working distance in cm; Edgett et al., 2012; Minitti et al., 2013). For WATSON and MI, fixed scales (14 μm/pixel and 31 μm/pixel, respectively) are used when available. The pixel scale of each image is reported in Table 3 (Line 258 in the revised manuscript).

Regarding 2D-to-3D harmonization: The fundamental incompatibility between 2D image analysis and 3D mass-based methods (dry sieving, laser diffraction) is acknowledged. Our approach is comparative rather than absolute: (1) Martian GSDs are processed exclusively through the same image-based method, enabling internal comparisons; (2) Lunar and terrestrial data are mass-based; (3) Cross-method comparisons are performed only at the parameter level (UGSD μ, Dc, n) rather than raw percentages. We also apply stereological conversion where appropriate (Pizzati et al., 2023): ECD-based distributions are multiplied by a shape factor of 1.3-1.5 when directly compared to sieve data, as justified in Section 2.3.2 (Lines 427-431 in the revised manuscript).

Regarding resolution limits and dust effects: The reviewer raises a critical point. MAHLI achieves 14 μm/pixel at minimum working distance (~2.1 cm), theoretically resolving ~40 μm grains at the 3-pixel threshold (Table 3). However, three factors degrade effective resolution: (1) The point spread function (PSF) of the optics requires ⩾3 pixels for reliable grain boundary detection (Eibl et al., 2016); (2) Foreground-background contrast is reduced by ubiquitous Martian dust (Chen-Chen et al., 2023); (3) Dust aggregates (<100 μm) are often indistinguishable from true grains (Friday et al., 2013).

Following the methodology of Karunatillake et al. (2014) and Shi et al. (2024), we adopt a conservative minimum detectable grain size of 5-7 pixels (not 3 pixels) for Martian images. This corresponds to ~70-100 μm for MAHLI at 14 μm/pixel. For the VFS class (<125 μm), individual grains are often not resolvable; such samples are classified based on their homogeneous appearance and dust-like chemistry rather than explicit grain counts (Shi et al., 2024). To address dust contamination: (1) We preferentially select images acquired after dust removal tool (DRT) brushing when available; (2) For unbrushed targets, we manually mask regions with obvious dust aggregates; (3) The fine tail (< min detectable size) is truncated, not extrapolated (Lines 259-266).

We have revised Table 3 accordingly.

Chen-Chen, H., Pérez-Hoyos, S., Sánchez-Lavega, A., & Peralta, J. (2023). Characterisation of deposited dust particles on Mars insight lander Instrument Context Camera (ICC) lens. Icarus, 392, 115393. https://doi.org/10.1016/j.icarus.2022.115393.

Edgett, K. S., Yingst, R. A., Ravine, M. A., Caplinger, M. A., Maki, J. N., Ghaemi, F. T., Goetz, W. (2012). Curiosity's Mars hand lens imager (MAHLI) investigation. Space science reviews, 170, 259-317. https://doi.org/10.1007/s11214-012-9910-4.

Eibl, M. A., & Fedo, C. M. (2016, March). A Mars Analog Study of 2D Textural Image Analysis: Effects of Shadows, Image Resolution, and Comparisons to Actual Sediment Textures from Aeolian Dune Sand, Moses Lake, WA. In 47th Annual Lunar and Planetary Science Conference (No. 1903, p. 2321).

Friday, M. E., Fedo, C. M., McGlynn, I. O., & McSween, H. Y. (2013, March). The accuracy of 2D assessment of sediment textures, and application to Mars. In 44th Annual Lunar and Planetary Science Conference (No. 1719, p. 2361).

Karunatillake, S., Wray, J. J., Gasnault, O., McLennan, S. M., Rogers, A. D., Squyres, S. W., ... & Olsen, N. (2014). Sulfates hydrating bulk soil in the Martian low and middle latitudes. Geophysical Research Letters, 41(22), 7987-7996. https://doi.org/10.1002/2014GL061136.

Minitti, M. E., Kah, L. C., Yingst, R. A., Edgett, K. S., Anderson, R. C., Beegle, L. W., ... & Van Beek, T. (2013). MAHLI at the Rocknest sand shadow: Science and science-enabling activities. Journal of Geophysical Research: Planets, 118(11), 2338-2360.

https://doi.org/10.1002/2013JE004426.

Pizzati, M., Mantovani, L., Lisotti, A., Storti, F., & Balsamo, F. (2023). Particle size distributions in Earth Sciences: a review of techniques and a new procedure to match 2D and 3D analyses. EGUsphere, 2023, 1-71. https://doi.org/10.5194/egusphere-2023-2636, 2023.

Shi, Y., Zhao, S., Karunatillake, S., Cousin, A., Zhao, J., & Xiao, L. (2024). Sorting and eathering trends of soil at Gale Crater, Mars: Implications for regional pedological processes. Journal of Geophysical Research: Planets, 129(12), e2024JE008598. https://doi.org/10.1029/2024JE008598.

Zhao, S., Karunatillake, S., & Shi, Y. (2023). Semi-automated granulometry software for Martian soil analysis (beta). Zenodo. https://doi.org/10.5281/zenodo.7998487.

- Grain size distribution is a fundamental characteristic of rocks and fines, and can in some cases be a discriminator of geologic history or provenance. However, the geologic variations in each site are not clearly explained, nor more importantly, how these differences may influence resulting grain size distribution. In general, the manuscript makes a good case that the UGSD equation can mathematically separate out generalized soil/regolith populations (Earth, Moon, Mars), thus allowing comparison. However, without a deeper understanding of the geologic provenance and history of the samples shown in Figure 4 (a key figure), their predictive use is unclear. For example, it is to be expected that impact gardening would strongly influence the grain size distribution of any lunar sample, so that these samples should be tightly clustered. But why should this tight cluster then plot so firmly in the "multi-process terrestrial" region? There are terrestrial squares that land right in this lunar oval (I'm not clear on the acronyms used for each site); why is this? Without being able to at least hypothesize about the results in Figure 4, the predictive use of the method remains theoretical rather than practical. I agree with the authors' assessment in lines 559-561; process attribution is currently not reliable with this method, lessening its value.

Reply: Thanks for the comments and they offer an opportunity to explain our ideas of the UGSD paradigm. Figure 4 exhibits the Weibull-distribution parameters (k, λ) of the UGSD exponent μ (Eq.4 in the text), which is used to illustrate the variability of the UGSD in earth, moon and Mars;

and actually the clusters fall into three areas, although with overlaps. In particular, the two points highlighted in the comment are worth noting:

Why do lunar samples plot within the terrestrial field? and **Why do terrestrial squares plot inside the lunar oval?** The lunar samples are tightly clustered because soils on moon are relatively generated in simple ways, e.g., dominated by impact comminution and alteration by physical and mechanical surface processes. These produce a narrow range of GSD characteristics. Consequently, the (k, λ) cluster fall within the range of terrestrial soils. On the other hand, the overlap points of terrestrial soils are those from mass movements (e.g., landslides and debris-flows), periglacial colluvium (Kunlun Mountains), and glacial till—are dominated by physical fragmentation with minimal chemical weathering. These processes are quantitatively similar to the surface processes on the moon and produce similar GSD characteristics. The overlap therefore does not indicate identical processes, but rather demonstrates that the $k-\lambda$ space captures textural similarity across different genetic environments.

For clericity we have now revised Figure 4 caption (Lines 467-473) and added a complete acronym table (Table S5, https://doi.org/10.6084/m9.figshare.32362083) to aid interpretation.

What is the predictive value if process attribution is not reliable? First, we agree with the reviewer that the UGSD is insufficient for making predictions of mechanical processes and identifying geological environments. Limited by the scope of this data-focused journal, we do not specifically address the issues related to prediction and general application. In fact, the primary intention of proposing the dataset holds twofold significance, one is soil-data parameterization — compressing high-dimensional textural data into interpretable, comparable low-dimensional UGSD parameters, as discussed in the present article; and the other is the data base for dynamical analysis of surface processes, which we have discussed elsewhere. For example, while the (k, λ)-plot distinguish soils in different planets on large scale and in total statistic sense, the UGSD parameters (μ, D_c)-plot, distinguishes soils from different sites related to local processes. As shown in the following Figure A, the (μ, D_c)-clusters in the four study areas in Mars exhibit great variety in soil types and are indicative for surface processes. Moreover, from the (μ, D_c)-fields one may derive random fields of mechanical and physical parameters of soils (e.g., the cohesion, friction angle, porosity, and hydraulic conductivity), which are the fundamental inputs for dynamical analysis.

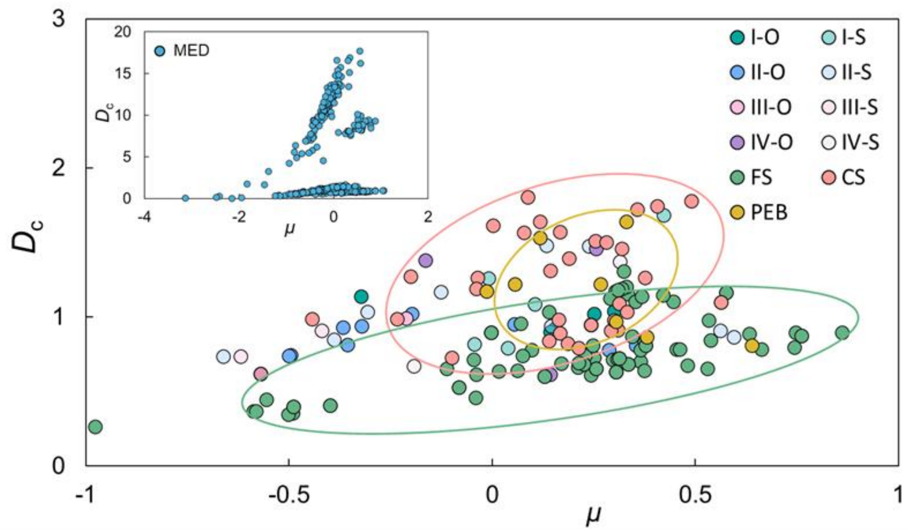

Figure A. (μ, D_c)-clusters for soil types in Mars, indicating local conditions of soil genesis and surface processes

Revisions made:

Expanded Figure 4 caption to explain the terrestrial-lunar overlap (Lines 467-473)

Added acronym definitions and geological descriptions (Table S5, https://doi.org/10.6084/m9.figshare.31569616)

Expanded Section 6.3 to clarify the limitations of process attribution (Lines 842-920)

Added discussion of predictive *vs*. theoretical value in Section 5.1 (Lines 712-723)

Specific

- Lines 34-35: Probably appropriate to reference Carrier's works here, too.

Reply: We have added this citation.

- Lines 35-36: This is probably too broad – grain size distribution is one factor in assessing both volcanic and sedimentary regoliths, but I think the second clause (past hydrological activity) is fully accurate.

Reply: We agree that the original phrasing "records the transition" may be overly broad. We have revised the sentence to: "On Mars, GSD provides important constraints on the transition from volcanic to sedimentary surface processes and on past hydrological activity (Grotzinger et al., 2015; Rivera-Hernández et al., 2020)." This wording acknowledges that GSD is one factor among several (e.g., sedimentary structures, mineralogy, geochemistry) used to interpret Martian surface processes, while still highlighting its value for constraining hydrological activity.

Grotzinger, J. P., Gupta, S., Malin, M. C., Rubin, D. M., Schieber, J., Siebach, K., ... & Wilson, S. A. (2015). Deposition, exhumation, and paleoclimate of an ancient lake deposit, Gale crater, Mars. Science, 350(6257), aac7575. https://doi.org/10.1126/science.aac7575

Rivera-Hernández, F., Sumner, D. Y., Mangold, N., Banham, S. G., Edgett, K. S., Fedo, C. M., ... & Wiens, R. C. (2020). Grain size variations in the Murray formation: Stratigraphic evidence for changing depositional environments in Gale crater, Mars. Journal of Geophysical Research: Planets, 125(2), e2019JE006230. https://doi.org/10.1029/2019JE006230

- Line 85: Not sure what you mean by data-driven. As opposed to what?

Reply: By "data-driven" we simply mean an approach based directly on measured data rather than on a pre-defined theoretical model.

- Line 95: Please define the UGSD earlier than Line 255.

Reply: Thank you for the comment. We have checked Line 65 (in the revised manuscript) and confirm that UGSD is already defined at its first appearance as "Universal Grain-Size Distribution (UGSD)." To avoid any confusion and redundancy, we have now removed the redundant definition at Line 374 and replaced it with: "The UGSD function, developed and..." We believe this addresses the reviewer's concern.

- Line 181: What process was used to digitize lunar data?

Reply: We digitized the lunar data from Graf (1993) by taking screenshots of each PDF page, using GPT to convert the screenshots into editable text, pasting the extracted text into Excel, and then manually verifying all data against the original tables.

- Lines 206-211: I'm not familiar with this software, but some treatment of its ability to identify grains using known analog targets would be very helpful.

Reply: The software is a Mathematica-based semi-automated segmentation tool (Zhao et al.,

2021) developed from the algorithm of Karunatillake et al. (2013, 2014). Its ability to identify grains has been validated in three ways:

(1) Comparison with manual segmentation: Using 57 MAHLI and MI images, the software was qualitatively compared against BASEGRAIN, ImageJ Trainable WEKA Segmentation, and ENVI classification tools, showing superior speed and accuracy.

(2) Known analog targets: Karunatillake et al. (2014, Icarus, Part 2, Section 3.4) tested the algorithm on terrestrial basaltic sand from Costa Rica (0.1–1.0 mm) with manually placed foreground pebbles (5–10 mm). The algorithm successfully segmented the foreground pebbles while correctly excluding background sand grains, demonstrating its ability to handle grain size contrasts in a known analog setting.

(3) Internal consistency: The same study (Part 2, Fig. 3) demonstrated that manual segmentation has substantial internal inconsistency (areal mismatch of 35–50 %), whereas the algorithm provides consistent results across repeated runs.

We have added a brief summary of these validation approaches in the revised manuscript at Lines 252-260, citing Karunatillake et al. (2014) and Zhao et al. (2021). The terrestrial analog test is now explicitly referenced.

Karunatillake, S., McLennan, S. M., Herkenhoff, K. E., Husch, J. M., Hardgrove, C., & Skok, J. R. (2014). A martian case study of segmenting images automatically for granulometry and sedimentology, part 1: algorithm. Icarus, 229, 400-407. https://doi.org/10.1016/j.icarus.2013.10.001

Zhao, S., Shi, Y., Karunatillake, S., & Xiao, L. (2021, March). Computational photoanalysis software for Martian soil Granulometry. In 52nd Lunar and Planetary Science Conference (No. 2548, p. 1959).

- Figure 2: Please define the y-axis for both columns. Also, the sweeping comment of Lines 231-232 is not supported sufficiently.

Reply: We have revised the caption of Figure 2 to clearly define the Y-axes for both columns. The left column Y-axis is now labeled as “percentage of grain mass per size class (%)” and the right column as “cumulative percentage passing (%)”. (Lines 373-377 in the revised manuscript)

We agree that the original statement requires appropriate literature support. We have revised the text (originally Lines 231-232, now Lines 354-358 in the revised manuscript) to include specific citations. The revised sentence now reads:

"Martian samples thus exhibit intermediate characteristics: some sites (e.g., Meridiani Planum) show well-sorted distributions indicative of aeolian sorting (Kozakiewicz et al., 2025), while others (e.g., Gale Crater) preserve multi-modal or poorly sorted signatures reflecting primary sedimentary textures (Milliken et al., 2014; Kapui et al., 2018)."

The supporting evidence is as follows:

Aeolian sorting: Kozakiewicz et al. (2025) documented active aeolian processes and well-sorted sand grains in Meridiani Planum.

Preserved primary textures: Milliken et al. (2014) described preserved dune bedforms in Gale Crater; Kapui et al. (2018) demonstrated using Earth analogs that mature basaltic sands can retain original grain-size signatures distinguishing fluvial from aeolian transport.

Kapui, Z., Kereszturi, A., Kiss, K., Szalai, Z., Újvári, G., Hickman-Lewis, K., ... & Westall, F. (2018). Fluvial or aeolian grains? Separation of transport agents on Mars using earth analogue observations. Planetary and Space Science, 163, 56-76. https://doi.org/10.1016/j.pss.2018.06.007

Kozakiewicz, J., Maj, D., Mol, S., Sobucki, M., Michaels, T., & Frodyma, N. (2025). Seasonality of aeolian landforms on Meridiani Planum, Mars. Icarus, 425, 116325. https://doi.org/10.1016/j.icarus.2024.116325

McSween, H. Y., McGlynn, I., & Fedo, C. (2010, December). Soils in Gusev Crater, Mars: What We Can And Cannot Learn From Surface Sediments. In AGU Fall Meeting Abstracts (Vol. 2010, pp. P52B-02).

Milliken, R. E., Ewing, R. C., Fischer, W. W., & Hurowitz, J. (2014). Wind-blown sandstones cemented by sulfate and clay minerals in Gale Crater, Mars. Geophysical Research Letters, 41(4), 1149-1154. https://doi.org/10.1002/2013GL059097

- Lines 494-498: The fact that the authors see less error across martian sites may be due to a number of reasons other than "geological signal", including the fact that nearly every image captures some fraction of the global fines mobilized by storms. Other possibilities should be discussed.

Reply: Thank you for this insightful comment. Indeed, attributing the observed inter-site variability solely to "geological signal" is an oversimplification. Two factors should be considered:

(1) Resolution limits of image-based measurements. As shown in Table 3, all Martian GSD curves are right-censored below ~0.04–0.15 mm (depending on the instrument). This truncation of the fine tail reduces the number of constraining size fractions available for curve fitting, which may artificially reduce apparent inter-site variability in UGSD parameters. In other words, some of the observed similarity across sites could be an artifact of what the instruments cannot see, rather than a true geological signal.

(2) Limited geographic and environmental coverage. The four landing sites (Gale, Jezero, Gusev, Meridiani) are all located in low-latitude regions with broadly similar aeolian regimes. This sampling does not capture the full range of Martian surface textural diversity (e.g., high-latitude mantling terrains, dust-dominated regions, or polar layered deposits).

We note that while dust storms are known to mobilize and transport fine particles globally (Kahre et al., 2017; Senel et al., 2021), the grain-size fractions most affected by this process (silt and clay, <62 μm) are below the detection limit of the rover imagers used in this study (Table 3). Therefore, the potential homogenizing effect of dust storms cannot be directly evaluated with our data and is not invoked as an explanation for the observed inter-site variability.

We have revised the discussion in the manuscript accordingly (Lines 694-711).

Kahre, M. A., Murphy, J. R., Newman, C. E., Wilson, R. J., Cantor, B. A., Lemmon, M. T., & Wolff, M. J. (2017). The Mars dust cycle. The atmosphere and climate of Mars, 18, 295.

Senel, C. B., Temel, O., Lee, C., Newman, C. E., Mischna, M. A., Muñoz-Esparza, D., ... & Karatekin, Ö. (2021). Interannual, seasonal and regional variations in the Martian convective boundary layer derived from GCM simulations with a semi-interactive dust transport model. Journal of Geophysical Research: Planets, 126(10), e2021JE006965. https://doi.org/10.1029/2021JE006965

- Figure 6: This figure is not effectively demonstrating the statements made in lines 517-525.

How does the top row relate to the bottom row? How do either of these rows relate to actual geologic maps as suggested in the text? I can't discern how the martian region shows "strongly graded" grain size distribution, for instance. It looks homogeneous to me – in fact they all do.

Reply: We agree that direct comparison with actual geologic maps would strengthen the interpretation. However, for the Apollo 17 and Gale Crater sites, in-situ grain-size transects at the scale of our simulations (kilometers) are not available. For Laowa Gully, we do not possess detailed geologic maps or measured grain-size spatial transects for validation.

Therefore, we have revised the text to present these simulations as demonstrations of methodological capability rather than validated predictions (Lines 721-750). Specifically:

1. We no longer claim that the patterns "reflect" or "are consistent with" actual geologic processes; instead, we state they "represent" or "emulate" the expected patterns based on first principles.

2. We explicitly acknowledge the lack of validation data for planetary sites and frame the simulations as generating testable hypotheses.

3. We clarify that the terrestrial case (Laowa Gully) serves as a baseline (homogeneous) simulation, with field validation deferred to future work.

- Lines 550-554: Gaps in geographic coverage likely have less meaning with respect to the predictive power of this method than gaps in representation of geologic provenance/environment. Asia has little scientific value as a discriminator, for example, while "playa lake" has far more.

Reply: We fully agree that gaps in geologic provenance/environment representation are more consequential for the predictive power of our method than geographic coverage per se. "Asia" is indeed a poor discriminator, whereas "playa lake," "debris-flow fan," or "aeolian dune" carry direct process-based meaning.

We have revised the text (Lines 550-554) to reframe the discussion from geographic to geologic/environmental terminology. Specifically, we now discuss underrepresented geomorphic settings (e.g., playa lakes, aeolian dune fields, glacial outwash plains, coastal sediments, tropical weathering profiles, and cold-desert soils) rather than geographic regions (e.g., South America). We have also explicitly noted that the current terrestrial dataset is dominated by gravity-driven and fluvial deposits (debris flows, landslides, alluvial fans), while other process regimes remain underrepresented. This revision directly addresses the reviewer's concern that environmental context is more scientifically meaningful than geographic labels.

The revised text is provided below and highlighted in the revised manuscript. Thank you for guiding us toward a more process-based framing.

"**(1) Gaps in geologic and environmental representation.** The terrestrial component is dominated by gravity-driven and fluvial deposits (debris flows, landslides, alluvial fans), while other geomorphic settings remain underrepresented. These include playa lakes, aeolian dune fields, glacial outwash plains, coastal sediments, tropical weathering profiles, and cold-desert soils (e.g., Antarctic dry valleys). The current geographic bias toward Asia (73% of samples) is less critical than the underrepresention of these process-based endmembers. Lunar samples are predominantly from nearside mare regions, with highland, farside, and polar materials. Martian samples are limited to four low-latitude landing sites, lacking high-latitude and dust-dominated terrains. Priority targets for future data collection include underrepresented geomorphic settings on Earth (as listed above), lunar farside and polar regions, and new Martian landing sites (e.g., Zhurong at Utopia Planitia, and high-latitudes sites such as Arcadia Planitia)."

- Lines 582-585: Being able to demonstrate grain size distribution of a site would provide important ancillary data, but note that there are other, more well-tested ways of determining geological context for landing site assessment, such as geologic maps.

Reply: We agree that geologic maps remain the most well-tested and authoritative means of determining geological context for landing site assessment. We did not intend to suggest that GSD estimates could replace established methods such as geologic mapping, spectroscopic analysis, or thermal inertia measurements.

We have revised the text to clarify that pre-landing GSD estimates derived from PlanetGSD serve as ancillary data that complement, rather than substitute for, traditional approaches. The revised text now reads (Lines 854-860):

"(3) Landing site assessment. While geologic mapping and remote sensing analyses (e.g., spectroscopy, thermal inertia) remain the primary methods for determining geological context, pre-landing GSD estimates can be generated as ancillary data by identifying PlanetGSD entries with geological context similar to candidate landing sites. Such estimates provide additional constraints on surface mechanical properties (e.g., trafficability, drilling resistance) that are not directly accessible from orbital data alone. Table 11 provides example analog matches to illustrate this complementary approach."

This revision positions GSD-based estimation as a supplementary tool that adds value to existing well-established methods, rather than as a standalone alternative. We thank the reviewer for guiding us toward a more accurate framing.