https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Open Access Earth System Science Data Discussions

PlanetGSD 1.0: a cross-planetary grain-size distribution dataset from Earth, the Moon, and Mars

Jun Zhang[1,2], Yong Li[1*]

[1] Key Laboratory of Mountain Hazards and Engineering Resilience /Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610213, China

[2] State Key Laboratory of Hydroscience and Engineering; Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

*Correspondence to: Yong Li, ylie@imde.ac.cn

Abstract

Comparative studies of surface processes across planetary bodies are hindered by the lack of consistently parameterized, openly accessible soil data, especially the grain-size distribution (GSD) data. Here we present PlanetGSD 1.0, the first standardized and unified cross-planetary GSD database. It comprises 6,527 measurements from Earth (4,419 samples, 20 geomorphic settings), the Moon (379 samples, 8 missions), and Mars (1,729 rover-derived estimates, 4 landing areas), covering seven orders of magnitude in grain size (0.0001 – 600 mm). The traditional textural fractions have been transferred into a unique parameter set (μ, D_c, n) derived from the unified GSD (UGSD) function (median R^2 = 0.988), with quality metrics, georeferenced metadata, site-level Weibull statistics, and open-source analysis codes. Technical validation confirms high fitting quality across all samples (97.8% with R^2 > 0.95) and robust inter-operator reproducibility for Martian image-derived measurements (coefficient of variation < 8.3% for key parameters). The complete dataset is openly available on Figshare (Zhang, 2026) under CC BY 4.0 license. PlanetGSD 1.0

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

enables robust cross-planetary comparison of regolith properties, benchmarking of simulants, and data-driven landing site assessment, establishing a foundational resource for planetary science.

1. Introduction

Understanding the physical properties of surface materials is fundamental to comparative planetology, geotechnical engineering, and the planning of future exploration missions. Among these properties, grain-size distribution (GSD) is one of the most informative: it governs the mechanical strength, hydraulic conductivity, and geochemical reactivity of soils and regolith across planetary bodies (Carrier, 2003; Fedo et al., 2015; Iverson & Vallance, 2001). On Earth, GSD classification underpins soil taxonomy systems and is routinely used in geotechnical site characterization (Das, 2019). On the Moon, GSD controls regolith trafficability, dust mobilization, and the design of in-situ resource utilization (ISRU) systems (Colwell et al., 2007). On Mars, GSD records the transition from volcanic to sedimentary surface processes and provides constraints on past hydrological activity (Grotzinger et al., 2015).

Despite this shared importance, existing GSD datasets remain fragmented across disciplines and planetary targets, with no unified analytical framework for cross-body comparison. Terrestrial soil databases provide broad geographic coverage but lack the resolution needed for process-based comparison with extraterrestrial materials. Lunar and Martian GSD data are scattered across mission-specific archives and individual publications, with heterogeneous analytical methods that preclude direct integration.

Several major compilations illustrate these limitations. For terrestrial soils, ISRIC-WISE (Batjes, 2009) and HWSD (Shangguan et al., 2014) provide global coverage at ~1 km resolution but report only texture-class percentages (sand, silt, clay) rather than complete GSD curves.

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

POLARIS (Chaney et al., 2016) offers 30 m probabilistic texture estimates for the USA but similarly lacks full size distributions. These databases employ Earth-centric classification schemes (USDA, FAO) that are not transferable to extraterrestrial contexts. For lunar materials, the NASA Lunar Soils Grain Size Catalog (Graf, 1993) documents ~350 Apollo and Luna samples with cumulative mass percentages at 9–15 sieve fractions, but bin widths vary by mission protocol, and no unified parameterization is provided. The Lunar Sourcebook (Heiken et al., 1991) summarizes mean grain sizes and sorting coefficients for ~200 samples but does not provide a digital database or complete GSD curves. Recent Chang'e-5 and Chang'e-6 sample characterizations (Li et al., 2022, 2024; Zhang et al., 2022) add valuable data but remain isolated from the Apollo/Luna corpus. For Mars, GSD measurements are derived from rover imaging at individual landing sites: Gale Crater (Fedo et al., 2015; Weitz et al., 2018), Jezero Crater (Stack et al., 2020), Gusev Crater (Herkenhoff et al., 2004), and Meridiani Planum (Weitz et al., 2006). These studies employ site-specific image analysis protocols with varying resolution limits and binning schemes, and no cross-site compilation or standardized framework exists. Academic survey datasets in terrestrial geomorphology and sedimentology (e.g., Blott & Pye, 2001; Eshel et al., 2004) report individual GSD measurements in heterogeneous formats that resist systematic integration. Consequently, no unified framework exists for comparing GSD characteristics across planetary bodies. A systematic comparison of these existing resources with PlanetGSD 1.0 is provided in Table 1 (Section 2.1).

Compounding this data fragmentation, conventional GSD descriptors present additional limitations. Percentile indices (D_{10}, D_{50}, D_{90}) and moment statistics (Folk & Ward, 1957) reduce continuous distributions to summary values without direct links to physical processes. Standard curve-fitting models—including lognormal, Rosin–Rammler, and Fredlund functions—impose

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

unimodal forms that cannot adequately represent the polygenetic, multimodal distributions common in natural regolith (Bittelli et al., 1999). The Universal Grain-Size Distribution (UGSD) function (Yong et al., 2013, 2017) addresses these limitations through a flexible four-parameter formulation capable of capturing multimodal distributions, but has not previously been applied as a unified framework for cross-planetary comparison.

Here we present PlanetGSD 1.0, a standardized database of 6,527 GSD measurements compiled from Earth (4,419 samples spanning 20 geomorphic settings across 18 countries), the Moon (379 samples from Apollo, Luna, and Chang'e missions), and Mars (1,729 rover-derived estimates from Gale, Jezero, Gusev, and Meridiani sites), covering seven orders of magnitude in grain size (0.0001 – 600 mm). All samples have been harmonized using the UGSD function, achieving a median fitting R^2 of 0.988. Compared with existing resources, PlanetGSD 1.0 provides four key advances: (i) the first unified parameterization across three planetary bodies; (ii) complete raw cumulative curves retained alongside fitted parameters; (iii) site-level Weibull statistics enabling stochastic field generation; and (iv) fully open data, codes, and metadata under CC BY 4.0 license. With this database we aim to fill the dual gap of fragmented data and absence of a unified analytical framework, enabling robust cross-planetary comparison of regolith properties and supporting data-driven benchmarking of simulant materials and landing-site assessment.

This paper is organized as follows. Section 2 describes the data acquisition methods for each planetary body. Section 3 presents the data records, including dataset structure, content summaries, and key fields. Section 4 reports the technical validation results. Section 5 provides illustrative applications and identifies open questions. Section 6 and Section 7 are usage notes and conclusions, respectively.

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

2. Methods

2.1 Overview of the PlanetGSD dataset

The PlanetGSD dataset comprises 6,527 soil samples from three planetary bodies—Earth, the Moon, and Mars—integrated within a unified analytical framework based on the Universal Grain-Size Distribution (UGSD) function. This compilation enables systematic comparison of granular characteristics across terrestrial and extraterrestrial environments with contrasting surface processes, climatic regimes, and geological histories. The database is organized with multi-level metadata, including geographic coordinates, geomorphic context, mission information, and original data sources, all standardized to facilitate cross-planetary analysis.

Earth component. The terrestrial samples represent the full spectrum of Earth's surface environments, spanning five continents and encompassing 20 distinct geomorphic settings. These include gravity-driven deposits (debris flows, landslides, moraines), fluvial systems (river valleys, floodplains), aeolian environments (deserts, loess), pedogenic landscapes (vegetated slopes, grasslands), and coastal zones. This diversity captures grain-size signatures associated with major process regimes, providing a comprehensive terrestrial reference for comparison with extraterrestrial materials. The Earth data integrate field-collected samples and literature-derived measurements, all processed using standardized laboratory protocols.

Lunar component. The lunar samples originate from eight sample-return missions conducted between 1969 and 2024, including Apollo (USA), Luna (USSR), and Chang'e (China) programs. These missions provide geographic coverage of major lunar geological provinces, including mare basalts, highland terrains, and the farside South Pole-Aitken basin. All samples are regolith materials processed by impact gardening and space weathering, with grain-size distributions originally

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026


Open Access Earth System Science Data Discussions

measured by dry sieving. The lunar data are harmonized from multiple mission catalogs, enabling consistent cross-mission comparison.

Martian component. The Martian samples are derived from high-resolution rover imagery at four NASA landing sites: Gale Crater (Curiosity), Jezero Crater (Perseverance), Gusev Crater (Spirit), and Meridiani Planum (Opportunity). These sites represent diverse Martian geological environments, including lacustrine mudstones and deltaic deposits (Gale, Jezero), weakly altered olivine basalt plains (Gusev), and sulfate-rich evaporite sequences (Meridiani), spanning Noachian to Amazonian periods. Grain-size distributions were obtained through semi-automated image segmentation, capturing grain sizes from silt to gravel.

Database structure. All samples in PlanetGSD 1.0 are standardized to a consistent framework: raw cumulative GSD curves are preserved in their original bin resolutions, and each sample is fitted to the UGSD function to extract interpretable parameters (μ, D_c, n). Site-level statistics, including three-parameter Weibull distributions for μ populations, are provided to enable stochastic field generation and comparative analysis. Detailed sample-level metadata, including geographic coordinates, geomorphic setting, mission information, and original data citations, are provided in Supplementary Tables S1–S7.

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Open Access Earth System Science Data Discussions

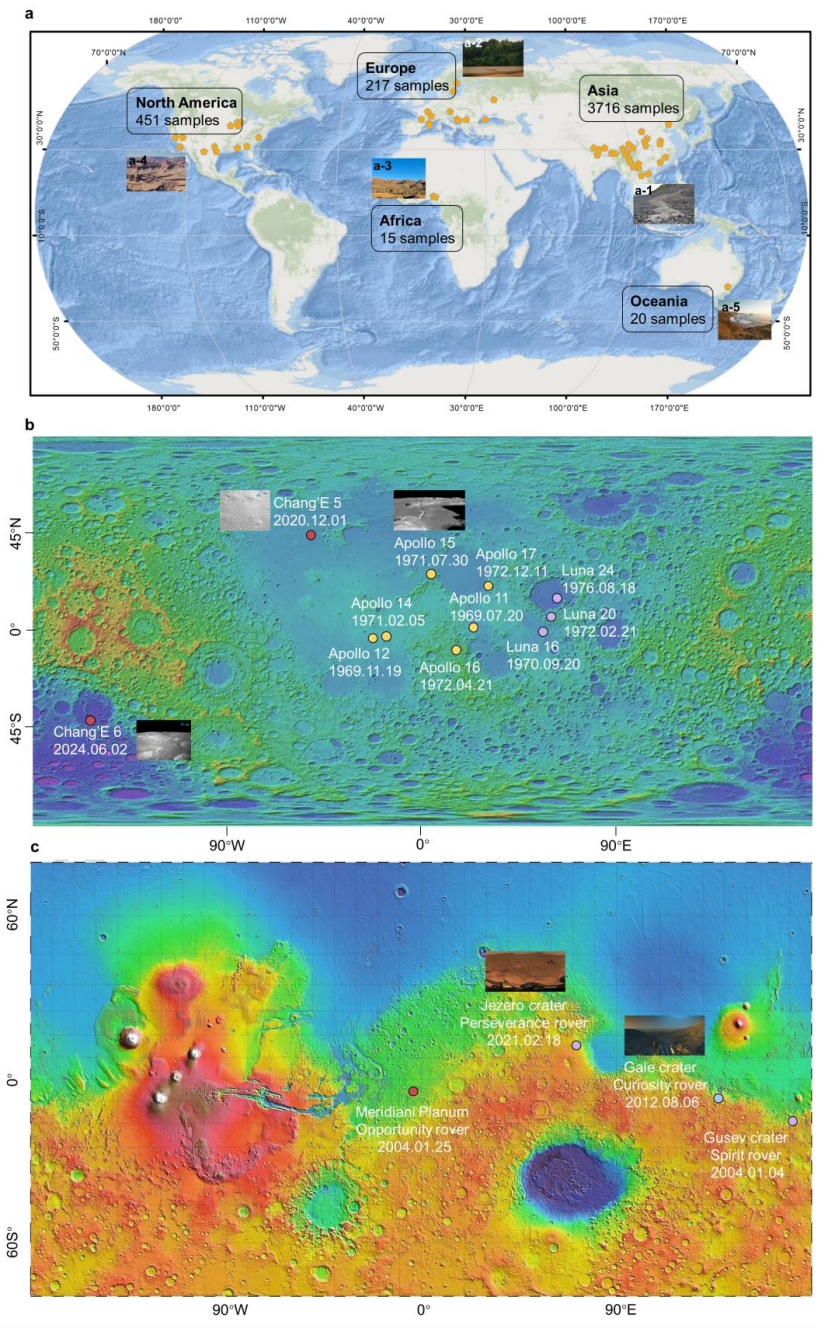


Figure 1. Global distribution of three terrestrial planetary soil sampling sites and representative photographic images. (a) Sampling locations across 18 countries on five continents: Asia (China, Singapore, South Korea, Thailand, Iran), Europe (Finland, Sweden, France, Bulgaria, Romania,

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Poland, Germany, Russia), Africa (Nigeria), North America (USA, Canada), and Australia. Representative soil environments are shown: (a-1) wide-graded debris-flow deposits (Jiangjia Gully, Southwest China); (a-2) arable land (Sweden); (a-3) gully terrain (Nigeria); (a-4) canyon (USA); and (a-5) coastal beach (Australia). (b) Sampling locations of lunar soils, predominantly collected near spacecraft landing sites. The insets show surface images from Apollo 15, CE-5, and CE-6, with mission dates annotated. The topographic basemap is the Lunar Reconnaissance Orbiter Laser Altimeter (LRO LOLA) color shaded relief (388 m resolution). Images courtesy of the China National Space Administration (CNSA). (c) Sampling locations of Martian soils across four distinct regions, with insets showing overviews of Gale and Jezero Craters. The basemap is the Mars Orbiter Laser Altimeter (MOLA) shaded relief topography. Images courtesy of NASA.

Table 1 summarizes the key characteristics of the three planetary datasets, including sample size, geographic coverage, analytical methods, grain-size ranges, and data sources. Table 2 provides detailed information for each sampling site, including sample counts, dominant soil types, and associated surface processes.

Table 1. Summary of the PlanetGSD dataset by planetary body

Component	Sample size	Geographic / landing-site coverage	Analytical method	Grain-size range (mm)	Primary data sources
Earth	4,419	18 countries on 5 continents; 20 geomorphic settings	Combined sieving (≥0.25 mm) + laser diffraction (<0.25 mm)	0.0001–20	Field campaigns (2015–2023); 43 peer-reviewed publications
Moon	379	8 sample-return missions (Apollo 11,12,14,15,16,17; Luna 16,20,24; Chang'e-5,-6)	Dry sieving	0.001–10	Graf (1993); Heiken et al. (1991); Zhang et al. (2022); Li et al. (2024)
Mars	1,729	4 rover landing sites (Gale, Jezero, Gusev, Meridiani)	Image-based grain segmentation (MAHLI, WATSON, MI, Mastcam, Mastcam-Z, RMI)	0.04–25	New analysis of PDS imagery (this study)

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

151 **Table 2.** Sampling sites, sample counts, dominant soil types, and associated surface processes

Planet	Site / Region	Sample count	Dominant soil types	Associated surface processes
Earth	Jiangjia Gully (China)	438	Debris-flow deposits	Gravity-driven mass movements
	Minjiang River (China)	825	Fluvial deposits	Fluvial transport and deposition
	Xichang (China)	960	Vegetated soils	Pedogenesis, root-soil interaction
	Horqin (China)	65	Grassland soils	Aeolian erosion, biotic processes
	Tibet (China)	107	Alpine meadow soils	Cryogenic processes
	Kunlun Mountains (China)	21	Permafrost and desert soils	Cryogenic, aeolian
	Malipo (China)	570	Red soils	Chemical weathering, pedogenesis
	Nevada (USA)	120	Desert soils	Aeolian erosion, arid weathering
	Michigan (USA)	31	Sandy soils	Glacial, coastal processes
	California (USA)	38	Coastal plain soils	Fluvial, coastal
	Canada	41	Fine loamy soils	Glacial, pedogenic
	Finland	108	Forest soils	Pedogenic, cryogenic
	Sweden	35	Arable soils	Agricultural, pedogenic
	Germany	40	Brown soils	Pedogenic
	France	45	Various soils	Fluvial, pedogenic
	Bulgaria	21	Agricultural soils	Pedogenic
	Poland	14	Forest soils	Pedogenic
	Romania	30	River plain soils	Fluvial
	Russia	32	Forest and plain soils	Pedogenic, fluvial
	Nigeria	15	Agricultural soils	Pedogenic, tropical weathering
	Australia	20	Coastal soils	Coastal, aeolian
	Singapore	32	Granite-derived soils	Tropical weathering
	South Korea	12	Red-yellow soils	Pedogenic
	Japan	42	Volcanic, forest soils	Volcanic, pedogenic
	Thailand	28	Siliceous soils	Tropical weathering
	Iran	35	Aeolian, marine sediments	Aeolian, coastal
Moon	Apollo 11	21	Mare basaltic regolith	Impact comminution, space weathering
	Apollo 12	20	Mare basaltic regolith	Impact comminution, space weathering
	Apollo 14	26	Highland ejecta	Impact comminution,

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

				mixing
	Apollo 15	63	Mare-highland boundary	Impact comminution
	Apollo 16	5	Highland regolith	Impact comminution, multi-source mixing
	Apollo 17	43	Mare-highland contact	Impact comminution
	Luna 16	-	Mare basaltic regolith	Impact comminution
	Luna 20	-	Highland regolith	Impact comminution
	Luna 24	-	Mare basaltic regolith	Impact comminution
	Chang'e-5	1	Mare basaltic regolith	Impact comminution, space weathering
	Chang'e-6	1	Farside SPA basin	Impact comminution
Mars	Gale Crater (11 sub-sites)	384	Lacustrine, aeolian, fluvial	Lake sedimentation, aeolian transport
	Jezero Crater (8 sub-sites)	722	Deltaic, lacustrine, volcanic	Delta deposition, lake environment
	Gusev Crater (10 sub-sites)	366	Volcanic, aeolian, hydrothermal	Flooding, volcanism, alteration
	Meridiani Planum (6 sub-sites)	257	Sulfate-rich, aeolian	Evaporite deposition, aeolian transport

2.2 Data acquisition methods

2.2.1 Field sampling (Earth)

Terrestrial field samples (2,847) were collected during campaigns conducted between 2015 and 2023 across 18 countries on five continents (Fig. 1a; Supplementary Tables S1–S5). Sampling targeted the upper 10 cm of the surface layer to ensure consistency with the actively weathering zone and to facilitate comparison with orbital remote sensing data. At vegetated sites, surface organic litter was removed prior to sampling.

The sampling strategy was tailored to geomorphic context: systematic grid-based transects for homogeneous terrain (plains, grasslands, deserts) and targeted sampling of distinct geomorphic units for heterogeneous environments (debris-flow deposits, landslides, moraines). Individual sample

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

mass ranged from 200 g to 2 kg (median 500 g) to ensure sufficient material for grain-size analysis and representative sampling of heterogeneous deposits. Samples were sealed in polyethylene bags with GPS coordinates recorded, transported to the laboratory, and stored at 4°C prior to analysis.

All field-collected samples were processed following standardized protocols (BS 1377-2, ASTM D422-63, ASTM D6913-04, ISO 11277:2009) using combined sieving and laser diffraction methods, as detailed in Section 2.3.1. Complete instrument specifications and quality control procedures are provided in Text S1. No additional replicate analyses or certified reference material measurements were performed specifically for this study.

2.2.2 Literature compilation (Earth and Moon)

Terrestrial literature data. In addition to field-collected samples, we compiled grain-size distribution data for 1,572 terrestrial samples from 43 peer-reviewed publications. These samples expand the geographic and environmental coverage of the terrestrial component, including soils from Europe (Finland, Sweden, France, Germany, Poland, Romania, Russia), North America (USA, Canada), Asia (China, Japan, South Korea, Thailand, Iran), Australia, and Africa (Nigeria). The compiled data represent diverse surface environments such as forested hillslopes, agricultural lands, grassland soils, desert regions, coastal zones, and loess-mantled terrains, complementing the field-collected samples and enhancing the representation of major process regimes (Text S2).

Lunar mission data. Lunar grain-size distributions were compiled from eight sample-return missions (Apollo, Luna, Chang'e) spanning 1969–2024. Data were digitized from the NASA Lunar Soils Grain Size Catalog (Graf, 1993) for Apollo and Luna samples, and from mission-specific publications for Chang'e-5 (Zhang et al., 2022) and Chang'e-6 (Li et al., 2024). All lunar samples were originally analyzed using standardized dry sieving protocols. Detailed mission information,

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

sample counts, and metadata are provided in Table 1 and Table S6.

2.2.3 Image-based analysis (Mars)

Martian GSD data were derived from high-resolution rover images for 1,729 sampling targets at four NASA landing sites: Gale Crater (Curiosity), Jezero Crater (Perseverance), Gusev Crater (Spirit), and Meridiani Planum (Opportunity). All images were obtained from NASA's Planetary Data System (PDS) Imaging Node (https://pds-imaging.jpl.nasa.gov/). Source image identifiers (PDS Product IDs) for all 1,729 samples are provided in Dataset S3.

Imaging systems. Grain-size measurements utilized both arm-mounted close-range imagers and mast-mounted cameras. Close-range systems include the Mars Hand Lens Imager (MAHLI) on Curiosity (Edgett et al., 2012), the Wide Angle Topographic Sensor for Operations and engineering (WATSON) on Perseverance, and the Microscopic Imager (MI) on Spirit and Opportunity (Herkenhoff et al., 2004). These instruments achieve pixel scales of 13–32 μm at working distances of 2–5 cm, enabling resolution of sand-sized grains down to approximately 40 μm. Mast-mounted systems—Mastcam on Curiosity (Bell et al., 2017), Mastcam-Z on Perseverance (Bell et al., 2021), and SuperCam Remote Micro-Imager (RMI) on Perseverance (Maurice et al., 2021; Wiens et al., 2021)—provide pixel scales of 74–450 μm depending on focal length and target distance, limiting grain-size resolution to medium sand and coarser fractions (>150 μm). Table 3 summarizes the key characteristics, resolution limits, and detection thresholds for all imaging systems used in this study.

Image selection and analysis. A total of 2,160 images meeting quality criteria (known target distance, viewing angle <30° from vertical, solar elevation >20°) were selected from the PDS archive. Grain-size distributions were measured using semi-automated image segmentation software developed by Shi et al. (2021) and Zhao et al. (2023), implemented in Mathematica. The

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

software identifies individual grains through edge detection and watershed segmentation, calculates equivalent circular diameter (ECD) for each grain, and constructs cumulative size distributions after binning into logarithmic size classes. For each image, 150–500 grains (median 280) were segmented. The geographic distribution of all four landing sites is shown in Figure 2. Complete image selection criteria, analysis protocols, resolution limits, and uncertainty estimates are provided in Text S3.

Table 3. Martian rover imaging systems and grain-size detection limits

Instrument	Rover(s)	Type	Pixel scale (µm/pixel)	Working distance	Min. detectable grain size (mm)	Reference
MI (Microscopic Imager)	Spirit, Opportunity	Arm-mounted	31 (fixed)	6.8 cm (fixed)	0.1	Herkenhoff et al., 2004
MAHLI (Mars Hand Lens Imager)	Curiosity	Arm-mounted	13.9–21.7	1.9–25 cm	0.04	Edgett et al., 2012
WATSON	Perseverance	Arm-mounted	~14	~2–20 cm	0.04	Beegle et al., 2015
Mastcam (M-100)	Curiosity	Mast-mounted	74–220	1–10 m	0.22	Bell et al., 2017
Mastcam-Z (left)	Perseverance	Mast-mounted	148 (at 2 m)	0.5–∞ m	0.44	Bell et al., 2021
RMI (SuperCam)	Perseverance	Mast-mounted	40 (at 2 m)	1.6–∞ m	0.12	Maurice et al., 2021

Note: Min. detectable grain size ≈ 3× pixel scale, the minimum threshold for reliable grain boundary segmentation. Pixel scales are from instrument calibration references cited in the last column.

2.3 Grain-size analysis and harmonization

2.3.1 Analytical methods

The PlanetGSD dataset integrates grain-size distributions obtained through three fundamentally different analytical methods: combined sieving–laser diffraction (terrestrial field samples and literature), dry sieving (lunar missions), and image-based segmentation (Mars rovers).

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

222 Table 4 summarizes the principles, grain-size ranges, advantages, and limitations of each method.

223

224 **Table 4.** Comparison of GSD analytical methods for grain-size determination across the three
225 planetary datasets

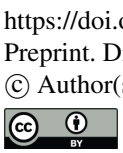
Methods	Principles	Grain-size range	Advantages	Limitations	Applied to
Combined sieving–laser diffraction	Sieving separates coarse fractions by physical passage through mesh; laser diffraction measures fine fractions by angular light scattering	0.0001–20 mm	Full size-range coverage (clay to gravel); high precision for fine fractions; standardized protocols (ASTM/ISO)	Time-intensive; requires laboratory access	Earth (field samples and literature)
Dry sieving	Mechanical shaking through nested sieve stacks	0.001–10 mm	Direct mass-based measurement; well-established for coarse sediments; large sample masses possible	Poor resolution for fines (<0.1 mm); clay fractions not resolved	Moon (all missions)
Image-based segmentation	Grain boundary detection from high-resolution images; equivalent circular diameter calculation	0.04–25 mm	In situ analysis; no sample return required; spatial context preserved	Truncated at fine end (≥40 μm); spherical grain assumption; 2D projection bias	Mars (rover images)

226

227 Figure 2 shows representative grain-size distribution curves from each planetary body,
228 illustrating the range of shapes captured in the dataset. The terrestrial examples (Fig. 2a) include
229 poorly sorted landslide deposits, bimodal debris-flow materials, and finer-grained vegetated soils.
230 Lunar samples (Fig. 2b) show the relatively narrow, fine-grained distributions characteristic of
231 impact-comminuted regolith. Martian samples (Fig. 2c) exhibit intermediate characteristics, with
232 some sites showing aeolian sorting and others preserving primary sedimentary textures.

233 The diversity of grain-size curves across and within planetary bodies—ranging from unimodal
234 to multimodal distributions, with varying degrees of sorting and skewness — highlights the
235 complexity of natural granular materials. Traditional descriptors such as percentile-based indices
236 (D_{10}, D_{50}, D_{90}) or unimodal distribution models (e.g., lognormal, Weibull) are insufficient to capture
237 this variability in a standardized manner, particularly for cross-planetary comparison where

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026



analytical methods and size ranges differ fundamentally. To address this limitation, we employ the Universal Grain-Size Distribution (UGSD) function (Section 2.3.2), which provides a flexible, four-parameter formulation capable of representing diverse curve shapes within a unified framework, enabling systematic comparison of granular characteristics across Earth, the Moon, and Mars.

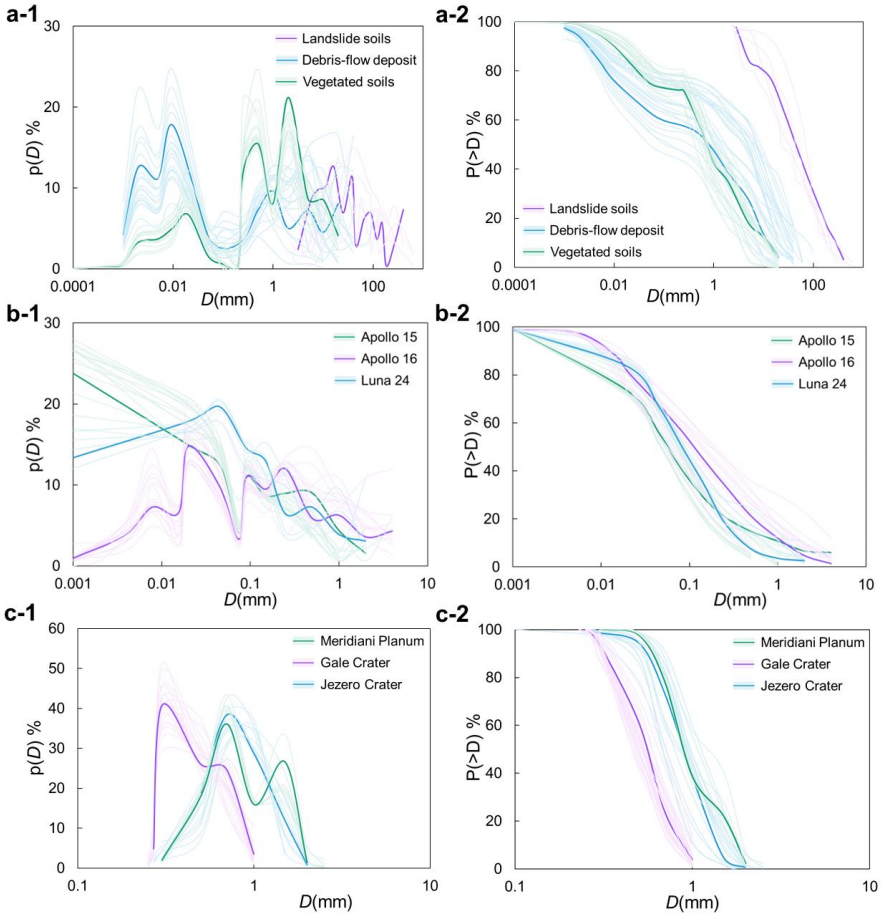


Figure 2. Representative grain-size distribution curves from each planetary body. Left column: frequency curves; right column: cumulative probability curves. (a) Terrestrial: landslide deposits, debris-flow deposits, and vegetated soils; (b) Lunar: Apollo 15, Apollo 16, and Luna 24; (c) Martian: Meridiani Planum, Gale Crater, and Jezero Crater.

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

248

249 **2.3.2 Data harmonization**

250 Cross-planetary comparison requires harmonization of data collected using different methods
251 with distinct size ranges and bin resolutions. Rather than applying interpolation or re-binning which
252 could introduce artifacts, we employ the UGSD function to the measured GSD data.

253 The Universal Grain-Size Distribution (UGSD) function, developed and validated in our
254 previous work (Yong et al., 2013, 2017; Zhang et al., 2023, 2025), describes the cumulative GSD
255 as:

256 $$P(D) = f(\mu)D^{-\mu}\exp(-(D/D_c)^n) \quad (3)$$

257 where μ is a power exponent positively correlated with fine-particle content, D_c is the characteristic
258 grain size, the normalized coefficient $f(\mu) = a\exp(-b\mu)$, with empirically determined constants a and
259 b); and n is a shape exponent controlling the coarse-tail steepness.

260 All 6,527 samples were fitted to the UGSD function (Eq. 3) using nonlinear least-squares
261 optimization. Figure 3 shows the rescaled UGSD curves for Earth (a), Moon (b), and Mars (c),
262 demonstrating the unified fitting performance across planetary bodies. Median R^2 values are 0.991
263 (Earth), 0.989 (Moon), and 0.983 (Mars), with 97.8% of all samples yielding $R^2 > 0.95$ (Table 9).
264 The terrestrial and lunar populations exhibit strongly left-skewed R^2 distributions, with >85% of
265 samples achieving $R^2 > 0.98$. The slightly lower Martian median is attributable to the truncation of
266 the fine tail below ~0.04–0.15 mm imposed by imaging resolution limits (Table 3), which reduces
267 the number of constraining data points available for curve fitting.

268

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

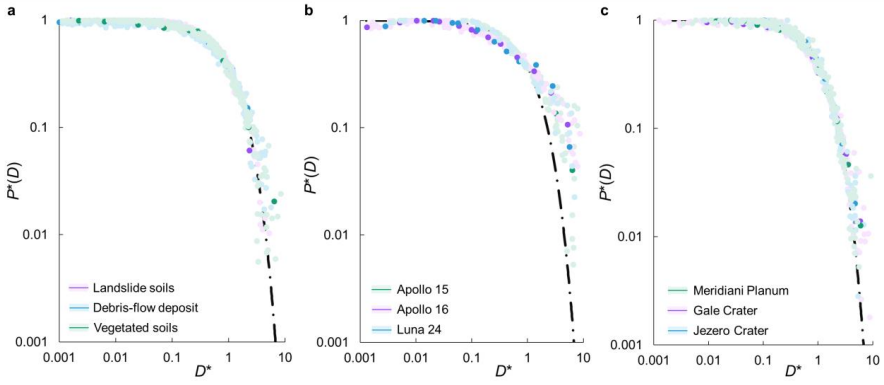


Figure 3. Rescaled UGSD fitting performance across the three planetary bodies. Each panel shows all measured cumulative GSD curves collapsed onto dimensionless axes by normalizing grain size D by the fitted characteristic size D_c and plotting the rescaled cumulative percentage $P(>D)/f(\mu)$ against $(D/D_c)^n$. (a) Earth; (b) Moon; (c) Mars.

All samples included in the dataset contain sufficient size fractions for reliable UGSD fitting (minimum five data points per sample). For curves truncated at coarse or fine ends (e.g., Martian samples missing fractions <40 μm), fitting is performed on the available portion, and truncation thresholds are documented in the metadata (Dataset S4). Complete quality control protocols and uncertainty estimate for all datasets are provided in Texts S1–S4. Parameter definitions and typical ranges for the dataset are summarized in Table 5.

Comparison with alternative parameterizations. To demonstrate the advantage of the UGSD function, we fitted all 6,527 samples to three alternative models (lognormal, Rosin-Rammler, and Fredlund three-parameter) and found that UGSD achieved higher median R^2 values than all three alternatives (Table S10).

Table 5. UGSD function parameters

Parameter	Symbol	Description	Typical range

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Power exponent	μ	Correlated with fine-particle content	0.01–3.5
Characteristic size	D_c	Representative grain diameter (mm)	0.01–20
Shape exponent	n	Controls coarse-tail steepness	0.1–2.0

2.4 Probability Distribution of UGSD Parameters

The UGSD parameters (μ, D_c, n) exhibit natural spatial variability across sampling sites, reflecting the inherent heterogeneity of granular materials under different surface processes. Among these, the exponent μ—which correlates positively with fine-particle content—serves as the most sensitive indicator of local textural characteristics and is therefore the primary focus of our probabilistic analysis. The characteristic size D_c also varies spatially and its distribution can be analyzed similarly, while the shape exponent n primarily reflects broad-scale differences between planetary bodies or major process regimes (e.g., $n \approx 1$ for Earth and Moon, $n > 1$ for Mars) and is not considered for site-scale probability distributions.

At the site level, the population of μ values was found to follow the three-parameter Weibull distribution:

$$f(\mu;\lambda,k,\mu_0) = \frac{k}{\lambda}\left(\frac{\mu-\mu_0}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{\mu-\mu_0}{\lambda}\right)^{k}\right) \quad (4)$$

where μ_0 is the location parameter (minimum μ value), λ is the scale parameter (characteristic spread), and k is the shape parameter (controlling distribution tail behavior). The three-parameter form was adopted because it provides greater flexibility than the two-parameter version, accommodating potential shifts in the minimum μ value that may arise from site-specific depositional or weathering histories. Fitting was performed using maximum likelihood estimation (scipy.stats.weibull_min.fit in Python).

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

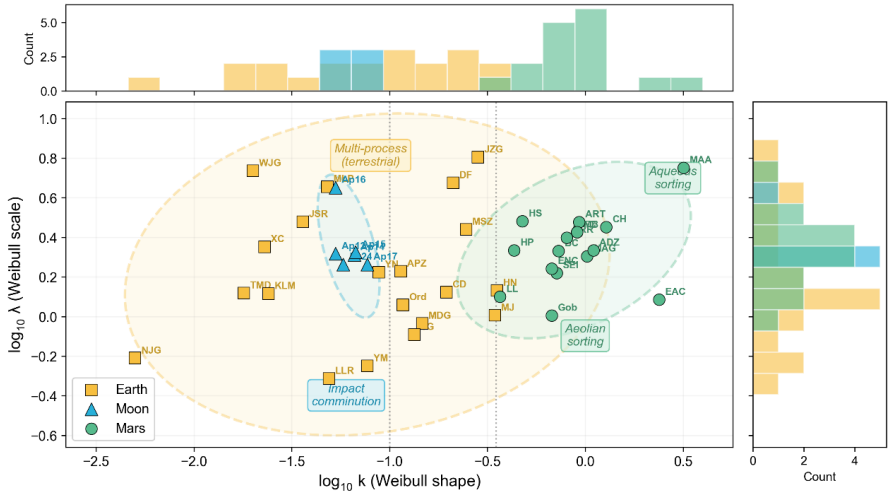
306 Table 6 summarizes the ranges of fitted Weibull parameters across the three planetary bodies,
307 and Figure 4 displays the site-level (k, λ) points for all 42 locations. One sees distinct point groups
308 for the three planetary bodies, supporting the utility of the k-λ parameter pair as a diagnostic
309 descriptor for cross-planetary comparison. Complete site-level Weibull parameters, including
310 geological context and dominant process annotations, are provided in Dataset S4 (Sheet 2).

311

312 **Table 6.** Weibull parameter ranges by planetary body

Planet	sites	k range	λ range	μ_0 range
Moon	6	0.053–0.077	1.83–4.48	0–0.028
Earth	20	0.005–0.354	0.49–6.38	0–0.294
Mars	16	0.367–3.180	1.01–5.64	0–3.142

313



314

315 **Figure 4.** Weibull k–λ parameter space for PlanetGSD 1.0 sites. Site-level data points are shown for
316 Earth (squares, 20), Moon (triangles, 6), and Mars (circles, 16). Marginal histograms show the
317 distribution of k (top) and λ (right) for each planetary body.

318

319 **2.5 From GSD measurements to random parameter fields**

320 The preceding sections established a two-step transformation of raw grain-size measurements

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

into a unified probabilistic framework (Fig. 5). First, the UGSD function converts the original GSD curves of all 6,527 samples into a compact set of interpretable parameters (μ, D_c, n), reducing high-dimensional textural data to physically meaningful indices (Section 2.3). Second, at each sampling site, the population of μ values is characterized by a three-parameter Weibull distribution (Section 2.4), capturing the local spatial variability of fine-particle content in a probabilistic form.

This transformation fundamentally changes the nature of the dataset: from a conventional collection of discrete observations to a probabilistic parameter library of grain-size distributions. In this library, each site is represented not by individual measurements but by the probability distribution of its granular characteristics (μ_0, k, λ), enabling stochastic simulation and extrapolation beyond the original sampling locations.

The probabilistic parameterization naturally leads to the concept of granular fields—spatially continuous representations of soil heterogeneity derived from the site-level Weibull parameters. Following the workflow of Zhang et al. (2022), μ-fields are generated through a three-step Monte Carlo procedure:

1. UGSD fitting of measured GSD data to obtain μ for each sample (as described in Section 2.3).
2. Weibull parameterization and Monte Carlo simulation – For each sampling site, the population of μ values is fitted to the three-parameter Weibull distribution (Section 2.4). An ensemble of N random μ values is then generated by sampling from this fitted distribution, with N specified by the user (default N = 1,000).
3. Spatial interpolation – The simulated μ values are assigned to a Digital Elevation Model (DEM)-registered grid via inverse-distance-weighted (IDW) interpolation (power = 2).

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Users can specify DEM resolution, interpolation radius, and the number of Monte Carlo realizations.

The resulting μ-fields provide spatially explicit estimates of grain-size characteristics that preserve local heterogeneity while enabling upscaling to regional scales. Because soil properties such as porosity, permeability, and shear strength can be expressed as functions of μ (Jiang et al., 2024), these granular fields can be further transformed into property fields for hydro-geophysical modeling, rover mobility simulation, or in-situ resource utilization assessment. The complete workflow, from raw measurements to μ-fields, is illustrated in Figure 5, and the implementation codes are provided in Code S1–S3.

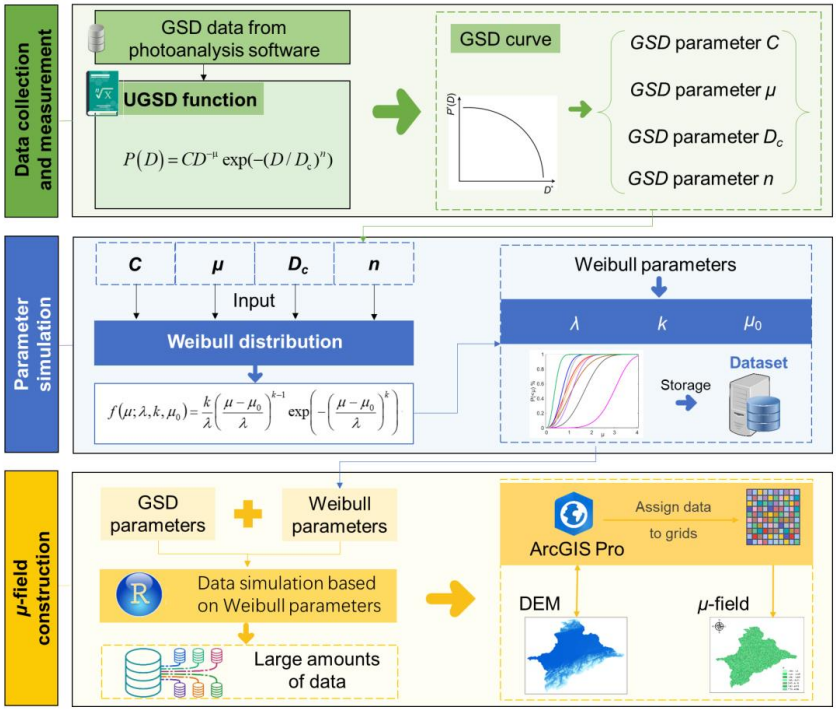


Figure 5. Workflow for generating spatially continuous μ-fields from raw GSD measurements. The process involves three stages: (1) UGSD fitting to extract μ for each sample; (2) Weibull

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

parameterization of site-level μ populations and Monte Carlo simulation of random μ values; (3) IDW interpolation onto DEM grids to produce continuous μ-fields. This transformation converts conventional observational data into a probabilistic parameter library, enabling stochastic simulation and cross-site comparison.

3. GSD Dataset Construction

The UGSD function has harmonized the textural data to a set of analytical parameters (μ, D_c, n), while preserving the integrity of the measurements. This allows to construct a new-type dataset of soil from the traditional measured data of texture or moisture of soils.

3.1 Database structure and file inventory

3.1.1 Folder organization

All datasets are provided in Microsoft Excel format (.xlsx) to accommodate multi-sheet structures that organize GSD curves, metadata, and quality flags within single files. The complete PlanetGSD 1.0 database is organized in a hierarchical folder structure with separate directories for data files, analysis codes, and documentation:

```
PlanetGSD_v1.0.0/
|-- data/
|-- codes/
|-- README.txt
```

3.1.2 File inventory

A complete inventory of all files in the PlanetGSD 1.0 repository is provided in Table 7, including file names, sizes, and brief descriptions.

Table 7. File inventory for PlanetGSD 1.0

File	Size	Description

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

S1_GSD_data_from_Earth.xlsx	145 KB	Raw cumulative GSD curves for 4,419 terrestrial samples, with geographic coordinates and metadata
S2_GSD_data_from_Moon.xlsx	36 KB	Raw cumulative GSD curves for 379 lunar samples, with mission identifiers and sample metadata
S3_GSD_data_from_Mars.xlsx	3,579 KB	Raw cumulative GSD curves for 1,729 Martian samples, with PDS image source IDs and geological context
S4_UGSD_parameters_and_Weibull.xlsx	107 KB	Two-sheet file: (Sheet 1) UGSD fitted parameters for all 6,527 samples; (Sheet 2) site-level Weibull statistics for 42 locations
S5_Mars_image_reference_numbers.xlsx	25 KB	Complete list of NASA PDS Product IDs for all 1,729 rover images used
S1_mu_field_generation.m	-	MATLAB code for stochastic μ-field generation
S2_UGSD_fitting.R	-	R code for data pre-processing and UGSD parameterization
S3_Weibull_fitting.py	-	Python code for Weibull distribution fitting and diagnostic plots
README.txt	-	Detailed dataset documentation, version history, and contact information

3.2 Database structure overview

The PlanetGSD 1.0 database is organized into five core datasets (S1–S5), each containing raw GSD measurements, fitted parameters, and associated metadata:

- Dataset S1 (Earth GSD) contains 4,419 terrestrial grain-size distributions with complete cumulative curves (17 size fractions from 0.0001 to 20 mm) and georeferenced metadata, integrating field-collected samples (2,847) and literature-derived data (1,572).
- Dataset S2 (Lunar GSD) comprises 379 lunar grain-size distributions from eight sample-return missions (Apollo, Luna, Chang'e), with 9–15 size fractions per sample depending on mission protocol.
- Dataset S3 (Mars GSD) provides 1,729 grain-size distributions derived from semi-

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

automated analysis of 2,160 rover images, with 8 logarithmic size bins (0.04–25 mm) and full PDS image traceability.

- Dataset S4 (UGSD and Weibull parameters) contains fitted UGSD parameters (C, μ, D_c, n) for all 6,527 samples (Sheet 1) and site-level three-parameter Weibull statistics (μ_0, k, λ) for 42 sampling locations (Sheet 2), with quality flags and geological context annotations.
- Dataset S5 (Mars image reference numbers) provides the complete list of NASA PDS Product IDs for all 1,729 rover images, ensuring full traceability to original mission data.

All datasets are provided in Microsoft Excel format (.xlsx) with multi-sheet structures to organize raw data, metadata, and quality flags within single files. Users are referred to the Figshare repository Zhang (2026) for direct data access.

3.3 Data dictionary

To facilitate correct interpretation and usage of the PlanetGSD 1.0 datasets, Table 8 provides complete field definitions for all variables appearing in Datasets S1–S5. The table includes field names, descriptions, data types, units, and example values, serving as a detailed reference to complement the database structure overview in Section 3.2.

Table 8. Data dictionary for PlanetGSD 1.0

Field	Description	Type	Unit	Example
sample_id	Unique sample identifier	String	—	Earth_DF_001
planet	Planetary body	String	—	Earth / Moon / Mars
site_name	Sampling site or mission name	String	—	Jiangjia Gully
latitude	Sampling latitude	Float (Decimal Degrees)	degrees	26.253
longitude	Sampling longitude	Float	degrees	103.117
elevation	Sampling elevation (Earth only)	Float	m	1,450
geomorphic_setting	Depositional or surface environment	String (Free text, see Table S9 for controlled vocabulary)	—	Debris flow

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

$D_{0.0001}$ to D_{20}	Cumulative mass percentage passing each size	Float (%)	%	2.3, 5.1, ...
$n_{fractions}$	Number of reported size fractions	Integer	—	17
data_source	Original data source	String	—	Field / Graf (1993)
μ	UGSD location parameter	Float	—	0.35
D_c	UGSD characteristic grain size	Float	mm	0.82
n	UGSD shape exponent	Float	—	0.48
R^2	Coefficient of determination for UGSD fit	Float	—	0.994
quality_flag	Data quality indicator	Integer (1 = good, 0 = flagged, 2 = excluded)	—	1 (good) / 0 (flagged)

4. Technical Validation

The validation of PlanetGSD 1.0 was conducted through a four-step procedure: (1) assessing UGSD fitting quality across all 6,527 samples; (2) evaluating Weibull parameterization quality at the site level; (3) verifying compiled data against original source publications; and (4) assessing inter-operator reproducibility for Martian image-derived measurements.

4.1 UGSD fitting quality

All 6,527 samples were fitted to the UGSD function (Eq. 3) using nonlinear least-squares optimization. Goodness-of-fit was quantified using the coefficient of determination (R^2). Figure 3 shows the rescaled UGSD curves for Earth (a), Moon (b), and Mars (c), demonstrating the unified fitting performance across planetary bodies.

Median R^2 values are 0.991 (Earth), 0.989 (Moon), and 0.983 (Mars), with 97.8% of all samples yielding $R^2 > 0.95$ (Table 9). The terrestrial and lunar populations exhibit strongly left-skewed R^2 distributions, with >85% of samples achieving $R^2 > 0.98$. The slightly lower Martian median is attributable to the truncation of the fine tail below ~0.04–0.15 mm imposed by imaging resolution limits, which reduces the number of constraining data points available for curve fitting.

Following the outlier screening protocol (Section 2.3.2), 23 samples (0.4% of total) with $R^2 <$

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

426 0.94 were excluded after manual review due to data-entry errors or highly irregular curve shapes
427 inconsistent with natural GSD. An additional 24 samples (0.4%) with $0.90 < R^2 < 0.94$ were retained
428 but flagged (quality_flag = 0) in Dataset S4 for user awareness. The remaining 6,480 samples
429 (99.3%) carry quality_flag = 1 and are recommended for analysis without restriction.

430 Comparison with alternative parameterizations. To demonstrate the advantage of the UGSD
431 function, we fitted all 6,527 samples to three alternative models (lognormal, Rosin-Rammler, and
432 Fredlund three-parameter) and found that UGSD achieved higher median R^2 values than all three
433 alternatives (Table S10).

434

435 **Table 9.** UGSD fitting quality statistics

Planet	samples	Median R^2	Mean R^2	% with $R^2 >$ 0.95	% with $R^2 <$ 0.90	K-S test pass rate
Earth	4,419	0.991	0.987	98.5	0.3	98.2%
Moon	379	0.989	0.985	97.9	0.5	97.1%
Mars	1,729	0.983	0.978	96.2	1.1	94.8%
Total	6,527	0.988	0.984	97.8	0.5	97.3%

436 Kolmogorov-Smirnov test at significance level $\alpha = 0.05$; pass rate indicates percentage of samples
437 where the UGSD fit is not significantly different from the measured distribution ($p > 0.05$).

438

439 **4.2 Weibull parameterization quality**

440 Three-parameter Weibull distributions were fitted to site-level μ populations for all 42 sites (20
441 Earth, 6 Moon, 16 Mars). Fit quality was assessed using two complementary approaches: the
442 Kolmogorov-Smirnov (K-S) goodness-of-fit test at significance level $\alpha = 0.05$, and visual inspection
443 of Q-Q (quantile-quantile) plots comparing observed μ quantiles against theoretical Weibull
444 quantiles.

445 Of the 42 sites, 39 (93%) passed the K–S test, indicating that the three-parameter Weibull

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

model adequately describes the μ population structure at most sampling locations. The three sites that did not pass are all terrestrial (MJ, DF, JZG) and exhibit bimodal μ distributions attributable to the mixing of geomorphically distinct lithological units within a single sampling area. Their Weibull parameters are retained in Dataset S4 (Sheet 2) with explanatory notes in the geological_description field; users may choose to exclude these sites or to fit separate Weibull distributions to the identified sub-populations.

4.3 Verification against published values

To verify the accuracy of compiled and processed data, we cross-checked subsets of samples from each planetary body against original source publications. This procedure tests for systematic errors introduced during data transcription, digitization, and processing.

(1) Lunar samples. Median grain sizes (D_{50}) for 50 Apollo samples were compared with values reported in the Lunar Sourcebook (Heiken et al., 1991). The mean absolute deviation between our compiled values and the Lunar Sourcebook was 4.2 μm (range: 0.3–11.8 μm), within the reported measurement uncertainty of ±5 μm for the original sieving analyses (Graf, 1993). No systematic bias was observed (mean signed deviation: +0.8 μm), confirming that digitization from the NASA Lunar Soils Grain Size Catalog did not introduce directional errors.

(2) Martian samples. D_{50} values for 30 Gale Crater samples were compared with those reported by Fedo et al. (2015) and Weitz et al. (2018) for overlapping locations. The mean absolute deviation was 0.08 mm (range: 0.01–0.22 mm), consistent with the inter-operator variability of ±15% documented for image-based grain-size measurements (see Section 4.4). No systematic offset was detected between our measurements and published values (mean signed deviation: −0.02 mm).

(3) Terrestrial samples. For 100 randomly selected literature-derived samples with data

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

available in the original publications, we verified cumulative grain-size percentages at the D_{10}, D_{50}, and D_{90} percentile levels. All values matched within rounding error (mean absolute difference: 0.3%; maximum: 0.8%). For the subset of 127 samples where original studies used hydrometer rather than laser diffraction methods, clay content (<2 μm) agreed within ±2.1% (mean absolute deviation) when cross-validated against laser diffraction measurements in overlapping datasets, consistent with known inter-method variability (Konert & Vandenberghe, 1997).

These checks confirm that data transcription and processing introduced no systematic errors into PlanetGSD 1.0.

4.4 Inter-operator reproducibility of Martian measurements

Because all Martian GSD data were derived from semi-automated image analysis, inter-operator variability represents an important component of measurement uncertainty. To quantify this variability, three independent analysts processed identical image sets (n = 30 images spanning all four landing sites) using the same grain-segmentation software (Shi et al., 2021; Zhao et al., 2023).

Results are summarized in Table 10. Inter-operator agreement was assessed using three metrics: (i) the coefficient of variation (CV) of D_{50} across the three analysts; (ii) the mean absolute percentage difference in cumulative grain-size fractions; and (iii) the resulting variation in fitted UGSD parameters.

Table 10. Inter-operator reproducibility for Martian image-based grain-size measurements (n = 30 images, 3 analysts)

Metric	Mean	Range	Notes
CV of D_{50}	8.30%	2.1–18.7%	Higher CV for images with more grain overlap

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Mean absolute % difference in cumulative fractions	6.20%	1.5–14.3%	Largest differences in finest resolvable fraction
CV of UGSD parameter μ	5.70%	1.2–12.4%	Propagated from grain boundary delineation differences
CV of UGSD parameter D_c	7.10%	2.3–15.8%	—
R^2 variation	±0.008	0.001–0.021	Fitting quality robust to operator differences

The primary source of inter-operator variability is grain boundary delineation for overlapping or partially obscured grains. Images with >30% grain overlap showed systematically higher CV (mean: 14.2%) compared to images with <10% overlap (mean: 3.8%). This operator-dependent uncertainty is propagated into the UGSD parameter confidence intervals reported in Dataset S4.

These inter-operator variability estimates (CV ≈ 5–8% for key parameters) are comparable to those reported in previous Mars grain-size studies using similar methods (Fedo et al., 2015; Weitz et al., 2018) and are substantially smaller than the inter-site variability observed across the four Martian landing sites, confirming that measurement uncertainty does not obscure geological signal in the dataset.

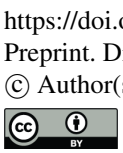
5. Illustrative applications and open questions

This section demonstrates three representative applications of PlanetGSD 1.0 and identifies open questions for future research.

5.1 Cross-planetary comparison of GSD signatures and μ-field generation

The unified UGSD parameterization enables not only direct comparison of grain-size characteristics across planetary bodies but also the generation of spatially continuous representations of soil heterogeneity. Using the site-level Weibull statistics derived in Section 2.4 and the Monte Carlo workflow described in Section 2.5, we generated μ-fields—spatially explicit

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026



maps of the UGSD parameter μ—for three representative sites, one from each planetary body (Fig. 6).

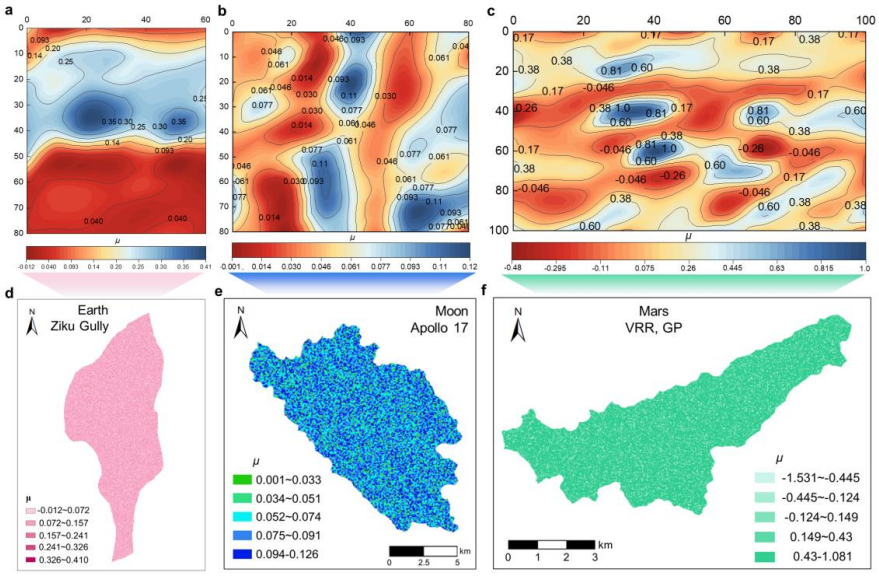


Figure 6. Spatial distribution of the UGSD parameter μ across three planetary bodies. The top row (a–c) shows μ-field maps for representative sites: (a) Ziku Gully, Earth; (b) Apollo 17 Taurus-Littrow valley, Moon; (c) Vera Rubin Ridge–Greenheugh Pediment area, Gale Crater, Mars. The bottom row (d–f) displays enlarged views of the corresponding areas at higher spatial resolution: (d) Ziku Gully; (e) Apollo 17 site; (f) Vera Rubin Ridge–Greenheugh Pediment area.

The Ziku Gully (Earth) exhibits a spatially homogeneous μ-field (range = 0.422), consistent with uniform pedogenic processes (Zhang et al., 2022). The Apollo 17 Taurus-Littrow valley (Moon) shows fine-scale heterogeneity (~100 m) without persistent gradients (range = 0.125), reflecting stochastic impact gardening as the dominant regolith-forming process (Zhang et al., 2025). The Vera Rubin Ridge–Greenheugh Pediment area in Gale Crater (Mars) exhibits a pronounced spatial

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

522 gradient (range = 1.612) that correlates with mapped lithological boundaries, suggesting process-
523 dependent GSD zonation (Sullivan et al., 2022; Weitz et al., 2018).

524 These contrasting patterns—homogeneous (Earth), randomly heterogeneous (Moon), and
525 strongly graded (Mars)—reflect the dominant surface processes on each planetary body and
526 illustrate how the probabilistic parameterization of GSD data (Section 2.4) can be transformed into
527 spatially explicit representations of grain-size heterogeneity. Such μ-fields provide valuable input
528 layers for geotechnical modeling, rover mobility simulation, and landscape evolution studies. Users
529 can generate analogous fields for other sites using the provided codes (Code S1) and their own
530 digital elevation model (DEM) data.

531 For a complementary view of the parameter space itself, the distribution of all 6,527 samples
532 in μ–D_c space is provided in Figure S2, where the three planetary populations occupy largely distinct
533 domains: Earth spans the μ range (−0.012 to +0.41), Moon occupies a narrow near-zero range (0.001
534 to 0.126), and Mars displays the broadest μ range (−1.531 to +1.081). The non-overlapping Weibull
535 shape parameter k ranges (Table 10) further confirm systematic differences among the three bodies,
536 supporting the utility of the k–λ parameter pair as a diagnostic descriptor for cross-planetary
537 discrimination.

538 **5.2 Regolith simulant benchmarking**

539 PlanetGSD 1.0 provides reference GSD parameters against which lunar and Martian simulant
540 materials can be quantitatively evaluated. Users can compute the UGSD and Weibull parameters of
541 candidate simulants using the provided analysis codes (Codes S1–S3) and compare them with the
542 site-level parameter ranges documented in Dataset S4. For example, the Weibull parameters for
543 Apollo mare sites (k = 0.053–0.077, λ = 1.83–4.48) define a target envelope for lunar simulants

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

intended to replicate mare regolith; simulants falling outside this envelope can be identified as texturally non-representative. Table 9 provides example analog matches linking target terrains to PlanetGSD reference sites to facilitate such comparisons.

5.3 Open questions and future directions

PlanetGSD 1.0 opens several avenues for future research, but also has limitations that should guide its use and development:

(1) Spatial coverage gaps. The terrestrial component is geographically biased toward Asia (73% of samples). Lunar samples are predominantly from nearside mare regions. Martian samples are limited to four low-latitude landing sites. Priority targets for future data collection include underrepresented terrestrial environments (tropical forests, polar deserts, South America), lunar farside and polar regions, and new Martian landing sites (e.g., Zhurong at Utopia Planitia).

(2) Analytical resolution limits. Martian GSD curves are right-censored below 40–150 μm depending on imaging instrument, truncating the fine tail of distributions. Validation against in-situ measurements from future sample-return missions will be essential to quantify systematic biases introduced by this truncation.

(3) Process attribution. While μ shows strong empirical correlations with depositional environments (e.g., $\underline{\mu} > 1.0$ for aeolian-dominated settings; $\mu < 0.5$ for lag deposits), the mechanistic links between UGSD parameters and specific surface processes remain incompletely understood. Controlled experiments and coupled process models are needed to establish causal relationships.

(4) Integration with remote sensing. Orbital datasets (thermal inertia, hyperspectral mineralogy, radar backscatter) provide indirect constraints on surface texture at global scales. Calibrating these proxies against UGSD parameters could enable planetary-scale mapping of grain-

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

size characteristics beyond the spatial reach of rover-based measurements.

(5) Climate change effects. Long-term terrestrial GSD records may capture the influence of changing precipitation patterns and temperature regimes on weathering and sediment transport. Coupling PlanetGSD with climate datasets could provide insights into how surface processes respond to environmental forcing—a question relevant to both Earth system science and interpreting Martian paleoclimate records.

6. Usage Notes

6.1 Recommended applications

PlanetGSD 1.0 is designed to support the following research applications:

(1) Comparative planetology. The unified UGSD parameterization enables direct comparison of grain-size signatures across Earth, Moon, and Mars. The $k - \lambda$ parameter domains (Figure 4) provide a quantitative framework for classifying samples by planetary provenance.

(2) Regolith simulant benchmarking. Lunar and Martian simulants can be quantitatively compared with natural planetary materials by computing their UGSD and Weibull parameters using Codes S2 and S3. Dataset S4 provides reference parameter values for all sampled sites to facilitate such comparisons.

(3) Landing site assessment. Pre-landing GSD estimates can be generated by identifying PlanetGSD entries with geological context similar to candidate landing sites. Table 11 provides example analog matches to illustrate this approach.

Table 11. Example analog matches for landing-site GSD estimation

Target terrain	PlanetGSD analog	Analog k range	Analog λ range
Phyllosilicate-bearing	Gale Crater Murray formation	1.8–2.5	2.0–2.5

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

plains	(MAG, VRR)		
Olivine-rich volcanic terrain	Gusev Columbia Hills (LL, HP)	0.4–0.6	1.3–2.2
Sulfate-rich evaporites	Meridiani Planum	0.5–0.9	1.5–3.0
Deltaic deposits	Jezero Crater delta	1.0–1.5	2.0–4.0

(4) Stochastic granular field generation. Using Code S1 and the fitted Weibull parameters (Dataset S4), users can generate spatially continuous μ-fields for geotechnical modeling, rover mobility simulation, or landscape evolution studies. The workflow accepts user-supplied DEM data and outputs μ-fields as GeoTIFF files compatible with standard Geographic Information System (GIS) software. Users can specify DEM resolution, interpolation radius, and the number of Monte Carlo realizations.

(5) Integration with remote sensing. The μ-field framework provides sub-pixel grain-size information that complements orbital datasets such as thermal inertia and radar backscatter. Coupling μ-fields with geophysical transfer functions enables derivation of spatially continuous porosity and cohesion estimates for engineering applications at future landing sites.

6.2 Methodological guidance for users

To ensure robust application of PlanetGSD 1.0, users should consider the following guidance:

Parameter interpretation. While the UGSD parameter μ correlates positively with fine-particle content and the Weibull shape parameter k reflects the width of the μ distribution, these are empirical descriptors derived from curve fitting. Direct translation of μ to physical quantities such as porosity, permeability, or shear strength should be undertaken with caution. Such applications are most robust when μ is coupled with complementary site-specific data or established geotechnical transfer functions (e.g., Zhang et al., 2022; Jiang et al., 2024). The parameters are primarily intended for comparative purposes—e.g., ranking sites by relative fineness or sorting—rather than for

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

607 absolute grain-size inference.

608 **Martian data handling.** All Martian GSD curves are right-censored below instrument-
609 specific resolution limits (Table 2). Users conducting analyses sensitive to the fine fraction (<0.1
610 mm) should treat cumulative percentages in the smallest size bin as lower bounds rather than
611 absolute values. For statistical applications, we recommend incorporating the truncation thresholds
612 (provided in Dataset S3 metadata) into uncertainty estimates. Sample-return measurements will be
613 essential for future validation of these image-derived GSDs.

614 **Lunar sample mass.** Lunar samples with mass <5 g are flagged (quality_flag = 0) in Dataset
615 S2 metadata due to potential sampling bias toward coarser fractions. Users requiring unbiased fine-
616 fraction statistics should apply mass-based filtering to retain only quality_flag = 1 samples.

617 **6.3 Limitations and caveats**

618 Users should be aware of the following limitations when using PlanetGSD 1.0:

619 **(1) Martian resolution limits.** All Martian GSD curves are right-censored below instrument-
620 specific detection limits (Table 2). Cumulative percentages in the fine tail represent lower bounds
621 rather than absolute values. This truncation contributes to lower UGSD fitting quality (median R^2 =
622 0.983) and may systematically affect Weibull parameter estimates.

623 **(2) Terrestrial sampling bias.** The terrestrial component is biased toward poorly sorted
624 deposits (debris flows, colluvium, moraines). Well-sorted aeolian dunes and littoral sands are
625 underrepresented. Users seeking to characterize these environments should supplement PlanetGSD
626 with dedicated datasets.

627 **(3) Lunar sample bias.** Lunar samples are predominantly from near-equatorial nearside mare
628 regions, with highland and farside materials underrepresented. This bias reflects the historical focus

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

of lunar exploration.

(4) Temporal mixing in Martian data. Martian samples span >3.5 Gyr of geological history. Users should consider depositional age when interpreting GSD characteristics; temporal context is provided in the metadata.

(5) Parameter interpretation. UGSD parameters are empirical descriptors derived from curve fitting. While they enable robust comparative analyses, direct physical interpretation requires caution and site-specific validation.

(6) Spatial extrapolation. The μ-fields generated using Code S3 rely on inverse-distance-weighted interpolation. In areas with sparse sampling or complex lithological boundaries, interpolated values may not accurately represent local grain-size characteristics.

6.4 Data updates and community contributions

PlanetGSD is conceived as a living database that will evolve with new missions and community input. Future versions (e.g., PlanetGSD 2.0) will incorporate:

- Additional terrestrial environments (aeolian dunes, glacial outwash, coastal sediments, loess sequences);
- Samples from ongoing and future missions (Mars Sample Return, Artemis lunar samples, OSIRIS-REx and Hayabusa2 asteroid samples);
- Extended grain-size ranges from improved analytical techniques;
- User-contributed datasets meeting quality standards.

Researchers wishing to contribute data should follow the standardized format:

- Complete cumulative grain-size percentages at standard size classes (or at original measured sizes with clear documentation);

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

- Geographic coordinates (WGS84 for Earth, Moon 2000/Mars 2000 for extraterrestrial bodies);
- Site metadata (geomorphic setting, sampling depth, lithology, depositional environment);
- Clear documentation of analytical methods and quality control procedures.

All contributed datasets will undergo the same validation and quality control checks applied to the original PlanetGSD compilation (see Technical Validation) before inclusion. Version history and a detailed changelog will be maintained in the repository README file.

7. Conclusions

The PlanetGSD 1.0 dataset establishes a unified framework for characterizing grain-size distributions across Earth, the Moon, and Mars, revealing that soils from these disparate planetary bodies conform to a single functional form—the Universal Grain-Size Distribution (UGSD). Key conclusions include:

1. **UGSD universality.** The UGSD function fits 6,527 samples from three planets with median $R^2 = 0.988$, demonstrating that granular configuration transcends genetic and environmental differences. The exponent n serves as a scaling factor classifying soils into categories that reflect textural organization inherited from formation processes.
2. ***μ* as process indicator.** The UGSD parameter μ effectively discriminates among depositional environments: $\mu > 1.0$ (aeolian-dominated), $\mu \approx 0.5$ (fluvial-aeolian transition), $\mu < 0.5$ (lag deposits, impact-influenced mixtures). Negative μ values record fine-fraction deficiencies from winnowing or analytical truncation.
3. **Probabilistic behavior.** Site-level μ populations follow three-parameter Weibull distributions, enabling stochastic generation of spatially continuous μ-fields that preserve

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

local heterogeneity while enabling upscaling to regional scales.

4. **Cross-planetary comparison.** The μ-D_c parameter space clusters soils by planetary body and process regime: Earth spans the widest range (diverse processes); Moon occupies a narrow, near-zero μ range (impact comminution); Mars displays the broadest μ range (complex aqueous-aeolian-volcanic interactions).

5. **Applications.** Granular fields support hydro-geophysical modeling, ISRU targeting, and mission planning by providing spatially explicit estimates of soil properties derived from limited point measurements.

The UGSD framework thus provides a quantitative, mechanistically grounded approach to soil characterization that unifies terrestrial and extraterrestrial regolith studies. By reducing high-dimensional texture data to interpretable indices and enabling probabilistic field generation, it bridges the gap between limited point measurements and spatially explicit process modeling, offering a new paradigm for planetary surface science.

Data Availability

The complete PlanetGSD 1.0 database is openly available at Figshare (https://doi.org/10.6084/m9.figshare.31569616, Zhang, 2026) under the Creative Commons Attribution 4.0 International License (CC BY 4.0). During peer review, the data can be accessed anonymously via the same URL. The repository contains five Excel files:

- 'S1_GSD_data_from_Earth.xlsx': 4,419 terrestrial samples with cumulative curves and metadata

- 'S2_GSD_data_from_Moon.xlsx': 379 lunar samples from eight sample-return missions

- 'S3_GSD_data_from_Mars.xlsx': 1,729 Martian samples derived from rover image analysis

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

- 'S4_UGSD_parameters_and_Weibull.xlsx': UGSD parameters for all samples and site-level Weibull statistics

- 'S5_Mars_image_reference_numbers.xlsx': PDS image Product IDs for all 1,729 Mars samples

Code Availability

All analysis codes are openly available in the PlanetGSD 1.0 repository at Figshare (https://doi.org/10.6084/m9.figshare.31569616). The repository contains three Python codes:

- Code S1 (MATLAB): Stochastic μ-field generation (requires MATLAB R2020a or later)

- Code S2 (R): Data pre-processing and UGSD parameterization (requires R ≥ 4.0)

- Code S3 (Python): Weibull distribution fitting and diagnostic plots (requires Python ≥ 3.9 with NumPy, SciPy, Matplotlib, and Pandas)

The codes require Python ≥ 3.9 with NumPy, SciPy, Matplotlib, and Pandas. Version history is maintained in the Figshare repository.

Acknowledgements

This research is supported by the National Youth Science Foundation (No. 42501091); the NSFC, China (42322703, 42271092); the Open Funds of the Key Laboratory of Mountain Hazards and Engineering Resilience, CAS (No. KLMHER-K22); the China National Postdoctoral Program for Innovation Talent (No. GZC20231347); the China Postdoctoral Science Foundation (No. 2023M741997).

Competing Interests

The authors declare no competing interests.

Author Contributions

J. Z.: Conceptualization, Data curation, Formal analysis, Software, Visualization, Writing –

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

718 Original Draft. **Y. L.:** Data curation, Validation, Writing – Review & Editing.

719

720 **Reference**

721 Batjes, N.H. (2009). Harmonized soil profile data for applications at global and continental scales:
722 updates to the WISE database. Soil Use and Management 5:124 – 127.
723 http://dx.doi.org/10.1111/j.1475-2743.2009.00202.x

724 Beegle, L., Bhartia, R., White, M., DeFlores, L., Abbey, W., Wu, Y. H., ... & Sobron, P. (2015,
725 March). SHERLOC: Scanning habitable environments with Raman & luminescence for
726 organics & chemicals. In 2015 IEEE aerospace conference (pp. 1-11). IEEE.
727 http://dx.doi.org/10.1109/AERO.2015.7119105

728 Bell III, J. F., Maki, J. N., Mehall, G. L., Ravine, M. A., Caplinger, M. A., Bailey, Z. J., ... & Wolff,
729 M. J. (2021). The Mars 2020 perseverance rover mast camera zoom (Mastcam-Z) multispectral,
730 stereoscopic imaging investigation. Space science reviews, 217(1), 24.
731 https://doi.org/10.1007/s11214-020-00755-x

732 Bell III, J. F., Godber, A., McNair, S., Caplinger, M. A., Maki, J. N., Lemmon, M. T., ... & Deen, R.
733 G. (2017). The Mars Science Laboratory Curiosity rover Mastcam instruments: Preflight and
734 in-flight calibration, validation, and data archiving. Earth and Space Science, 4(7), 396-452.
735 https://doi.org/10.1002/2016ea000219

736 Bittelli, M., Campbell, G. S., & Flury, M. (1999). Characterization of particle-size distribution in
737 soils with a fragmentation model. Soil Science Society of America Journal, 63(4), 782-788.
738 https://doi.org/10.2136/sssaj1999.634782x

739 Blott, S. J., & Pye, K. (2001). GRADISTAT: a grain size distribution and statistics package for the
740 analysis of unconsolidated sediments. Earth surface processes and Landforms, 26(11), 1237-
741 1248. https://doi.org/10.1002/esp.261

742 Bockheim, J. G., Hartemink, A. E. (2013). Classification and distribution of soils with lamellae in
743 the USA. Geoderma, 206, 92-100. https://doi.org/10.1016/j.geoderma.2013.04.014

744 Cabrol, N. A., Grin, E. A., Carr, M. H., Sutter, B., Moore, J. M., Farmer, J. D., ... & de Goursac, O.
745 (2003). Exploring Gusev Crater with Spirit: Review of science objectives and testable
746 hypotheses. Journal of Geophysical Research: Planets, 108(E12).

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

https://doi.org/10.1029/2002je002026

Carrier III, W. D. (2003). Particle size distribution of lunar soil. Journal of Geotechnical and Geoenvironmental Engineering, 129(10), 956-959.

Chaney, N. W., Wood, E. F., McBratney, A. B., Hempel, J. W., Nauman, T. W., Brungard, C. W., & Odgers, N. P. (2016). POLARIS: A 30-meter probabilistic soil series map of the contiguous United States. Geoderma, 274, 54-67. https://doi.org/10.1016/j.geoderma.2016.03.025

Christensen, P. R., Morris, R. V., Lane, M. D., Bandfield, J. L., & Malin, M. C. (2001). Global mapping of Martian hematite mineral deposits: Remnants of water-driven processes on early Mars. Journal of Geophysical Research: Planets, 106(E10), 23873-23885. https://doi.org/10.1029/2000je001415

Colwell, J. E., Batiste, S., Horányi, M., Robertson, S., & Sture, S. (2007). Lunar surface: Dust dynamics and regolith mechanics. Reviews of Geophysics, 45(2). https://doi.org/10.1029/2005rg000184

Das, B. M. (2019). Advanced soil mechanics. CRC press.

Demas, S. Y., Hall, A. M., Fanning, D. S., Rabenhorst, M. C., Dzantor, E. K. (2004). Acid sulfate soils in dredged materials from tidal Pocomoke Sound in Somerset County, MD, USA. Soil Research, 42(6), 537-545. https://doi.org/10.1071/SR03089

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

Eshel, G., Levy, G. J., Mingelgrin, U., & Singer, M. J. (2004). Critical evaluation of the use of laser diffraction for particle-size distribution analysis. Soil Science Society of America Journal, 68(3), 736-743. https://doi.org/10.2136/sssaj2004.7360

Fedo, C. M., McGlynn, I. O., & McSween Jr, H. Y. (2015). Grain size and hydrodynamic sorting controls on the composition of basaltic sediments: Implications for interpreting Martian soils. Earth and Planetary Science Letters, 423, 67-77. https://doi.org/10.1016/j.epsl.2015.03.052

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

Fredlund, M. D., Fredlund, D. G., & Wilson, G. W. (2000). An equation to represent grain-size

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

777 distribution. Canadian Geotechnical Journal, 37(4), 817-827. https://doi.org/10.1139/cgj-37-4-
778 817

779 Graf, J. C. (1993). Lunar Soils Grain Size Catalog (No. NASA-RP-1265).

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

783 Heiken, G., Vaniman, D., French, B.M., and Schmitt, J., 1991. Lunar Sourcebook: A User's Guide
784 to the Moon. Cambridge, UK.

785 Herkenhoff, K. E. (2004, December). High-Resolution Textures of Soils and Rocks at Gusev Crater
786 and Meridiani Planum From the Mars Exploration Rover Microscopic Imagers. In AGU Fall
787 Meeting Abstracts (Vol. 2004, pp. P13B-02).

788 Hwang, S. I., Lee, K. P., Lee, D. S., & Powers, S. E. (2002). Models for estimating soil particle-
789 size distributions. Soil Science Society of America Journal, 66(4), 1143-1150.
790 https://doi.org/10.2136/sssaj2002.1143

791 Hynek, B. M., & Di Achille, G. (2017). Geologic map of Meridiani Planum, Mars. US Geological
792 Survey Scientific Investigations Map, 3356. https://doi.org/10.3133/SIM3356

793 Isham, V., Cox, D. R., Rodriguez-Iturbe, I., Porporato, A., & Manfreda, S. (2005). Representation
794 of space–time variability of soil moisture. Proceedings of the Royal Society A: Mathematical,
795 Physical and Engineering Sciences, 461(2064), 4035-4055.
796 https://doi.org/10.1098/rspa.2005.1568

797 Iverson, R. M., & Vallance, J. W. (2001). New views of granular mass flows. Geology, 29(2), 115-
798 118. https://doi.org/10.1130/0091-7613(2001)029<0115:nvogmf>2.0.co;2

799 Jiang, H., Li, Y., Zou, Q., Zhang, J., Cui, J., Cheng, J., ... & Yao, H. (2024). A GSD-driven approach
800 to deriving stochastic soil strength parameters under hybrid machine learning models.
801 European Journal of Soil Science, 75(6), e70009. https://doi.org/10.1111/ejss.70009

802 Konert, M., & Vandenberghe, J. E. F. (1997). Comparison of laser grain size analysis with pipette
803 and sieve analysis: a solution for the underestimation of the clay fraction. Sedimentology, 44(3),
804 523-535. https://doi.org/10.1046/j.1365-3091.1997.d01-38.x

805 Lai, C. D., Murthy, D. N. P., & Xie, M. (2011). Weibull distributions. Wiley Interdisciplinary
806 Reviews: Computational Statistics, 3(3), 282-287. https://doi.org/10.1002/wics.157

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

807 Li, C., Hu, H., Yang, M. F., Liu, J., Zhou, Q., Ren, X., Ouyang, Z. (2024). Nature of the lunar far-
808 side samples returned by the Chang'E-6 mission. National Science Review, 11(11), nwae328.
809 https://doi.org/10.1093/nsr/nwae328
810 Li, C., Hu, H., Yang, M. F., Pei, Z. Y., Zhou, Q., Ren, X., ... & Ouyang, Z. (2022). Characteristics
811 of the lunar samples returned by the Chang'E-5 mission. National science review, 9(2),
812 nwab188. https://doi.org/10.1093/nsr/nwab188
813 Liu, T., Michael, G., Zuschneid, W., Wünnemann, K., & Oberst, J. (2021). Lunar megaregolith
814 mixing by impacts: Evaluation of the non-mare component of mare soils. Icarus, 358, 114206.
815 https://doi.org/10.5194/epsc2020-186
816 Mangold, N., Gupta, S., Gasnault, O., Dromart, G., Tarnas, J. D., Sholes, S. F., ... & Williford, K.
817 H. (2021). Perseverance rover reveals an ancient delta-lake system and flood deposits at Jezero
818 crater, Mars. Science, 374(6568), 711-717. https://doi.org/10.1126/science.abl4051
819 Maurice, S., Wiens, R. C., Bernardi, P., Cais, P., Robinson, S., Nelson, T., ... & Sobron, P. (2021).
820 The SuperCam instrument suite on the Mars 2020 rover: Science objectives and mast-unit
821 description. Space Science Reviews, 217(3), 47. https://doi.org/10.1007/s11214-021-00807-w
822 McLennan, S. M., Bell Iii, J. F., Calvin, W. M., Christensen, P. R., Clark, B. D., De Souza, P. A., ...
823 & Yen, A. (2005). Provenance and diagenesis of the evaporite-bearing Burns formation,
824 Meridiani Planum, Mars. Earth and Planetary Science Letters, 240(1), 95-121.
825 https://doi.org/10.1016/j.epsl.2005.09.041
826 Mitchell, J. K., Houston, W. N., Carrier III, W. D., & Costes, N. C. (1974). Apollo soil mechanics
827 experiment S-200 (No. NASA-CR-134306).
828 Rohatgi, A. (2020). Webplotdigitizer: Version 4.5. URL https://automeris. io/WebPlotDigitizer, 411,
829 860.
830 Ruff, S. W., Niles, P. B., Alfano, F., & Clarke, A. B. (2014). Evidence for a Noachian-aged
831 ephemeral lake in Gusev crater, Mars. Geology, 42(4), 359-362.
832 https://doi.org/10.1130/g35508.1
833 Shangguan, W., Gong, P., Liang, L., Dai, Y., & Zhang, K. (2014). Soil diversity as affected by land
834 use in China: consequences for soil protection. The Scientific World Journal, 2014(1), 913852.
835 https://doi.org/10.1155/2014/913852
836 Shi, Y., Zhao, S., Karunatillake, S., & Xiao, L. (2021, April). Semi-automated Image Segmenting

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Software for Martian Soil Granulometry. In EGU General Assembly Conference Abstracts (pp. EGU21-7872).

Squyres, S. W., Grotzinger, J. P., Arvidson, R. E., Bell III, J. F., Calvin, W., Christensen, P. R., ... & Soderblom, L. A. (2004). In situ evidence for an ancient aqueous environment at Meridiani Planum, Mars. science, 306(5702), 1709-1714. https://doi.org/10.1126/science.1104559

Stack, K. M., Williams, N. R., Calef III, F., Sun, V. Z., Williford, K. H., Farley, K. A., ... & Aileen Yingst, R. (2020). Photogeologic map of the perseverance rover field site in Jezero Crater constructed by the Mars 2020 Science Team. Space Science Reviews, 216(8), 127. https://doi.org/10.1007/s11214-020-00739-x

Sullivan, R., Baker, M., Newman, C., Turner, M., Schieber, J., Weitz, C., ... & Minitti, M. (2022). The aeolian environment in Glen Torridon, Gale crater, Mars. Journal of Geophysical Research: Planets, 127(8), e2021JE007174. https://doi.org/10.1029/2021JE007174

Weitz, C. M., Sullivan, R. J., Lapotre, M. G., Rowland, S. K., Grant, J. A., Baker, M., & Yingst, R. A. (2018). Sand grain sizes and shapes in eolian bedforms at Gale Crater, Mars. Geophysical Research Letters, 45(18), 9471-9479. https://doi.org/10.1029/2018GL078972

Weitz, C. M., Anderson, R. C., Bell III, J. F., Farrand, W. H., Herkenhoff, K. E., Johnson, J. R., ... & Sullivan, R. J. (2006). Soil grain analyses at Meridiani Planum, Mars. Journal of Geophysical Research: Planets, 111(E12). https://doi.org/10.1029/2005je002541

Wiens, R. C., Maurice, S., Robinson, S. H., Nelson, A. E., Cais, P., Bernardi, P., ... & Willis, P. (2021). The SuperCam instrument suite on the NASA Mars 2020 rover: Body unit and combined system tests. Space Science Reviews, 217(1), 4. https://doi.org/10.1007/s11214-020-00777-5

Yong, L., Chengmin, H., Baoliang, W., Xiafei, T., & Jingjing, L. (2017). A unified expression for grain size distribution of soils. Geoderma, 288, 105-119. https://doi.org/10.1016/j.geoderma.2016.11.011

Yong, L., Xiaojun, Z., Pengcheng, S., Yingde, K., & Jingjing, L. (2013). A scaling distribution for grain composition of debris flow. Geomorphology, 192, 30-42. https://doi.org/10.1016/j.geomorph.2013.03.015

Zhang, Jun (2026). GSD data from Earth, Moon, and Mars; Python code to simulater GSD dataset, calculate weibull parameters (for μ), and to create granular-field.. figshare. Dataset. https://doi.org/10.6084/m9.figshare.31569616.v3

Zhang, J., Li, Y., Cui, Y., Wu, Z., Xue, Y., Cheng, J., ... & Luo, A. (2025). Unity of terrestrial and

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

extraterrestrial soils in granular configuration. Earth and Planetary Science Letters, 654, 119239. https://doi.org/10.1016/j.epsl.2025.119239

Zhang, J., Li, Y., Yang, T., Liu, J., Guo, X., Yao, Y., 2023. A universal grain-size distribution of soil with scaling invariance. European Journal of Soil Science, 74(2), e13354. https://doi.org/10.1111/ejss.13354

Zhang, H., Zhang, X., Zhang, G., Dong, K., Deng, X., Gao, X., ... & Yang, M. (2022). Size, morphology, and composition of lunar samples returned by Chang'E-5 mission. Science China Physics, Mechanics & Astronomy, 65, 1-8. https://doi.org/10.1007/s11433-021-1818-1

Zhang, J., Li, Y., Liu, D., Jiang, N., Yang, T., Guo, X., & Yao, Y. (2022). The probability-based granular field of vegetated soils. Earth Surface Processes and Landforms, 47(13), 3100-3116. https://doi.org/10.1002/esp.5446

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

Zobeck, T. M., Gill, T. E., Popham, T. W. (1999). A two-parameter Weibull function to describe airborne dust particle size distributions. Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 24(10), 943-955. https://doi.org/10.1002/(sici)1096-9837(199909)24:10<943::aid-esp30>3.3.co;2-0

Reference from the Supplementary Material

Amelung, W., Zhang, X., Flach, K. W., & Zech, W. (1999). Amino sugars in native grassland soils along a climosequence in North America. Soil Science Society of America Journal, 63(1), 86-92. https://doi.org/10.2136/sssaj1999.03615995006300010014x

Arya, L. M., Bowman, D. C., Thapa, B. B., & Cassel, D. K. (2008). Scaling soil water characteristics of golf course and athletic field sands from particle-size distribution. Soil Science Society of America Journal, 72(1), 25-32. https://doi.org/10.2136/sssaj2006.0232

Barrett, L. R. (2001). A strand plain soil development sequence in Northern Michigan, USA. Catena, 44(3), 163-186. https://doi.org/10.1016/S0341-8162(00)00170-3

Ghosh, S., Deb, S., Ow, L. F., Deb, D., & Yusof, M. L. (2019). Soil characteristics in an exhumed cemetery land in Central Singapore. Environmental monitoring and assessment, 191(3), 174. https://doi.org/10.1007/s10661-019-7291-9

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

899 Gui, D., Lei, J., Zeng, F., Runge, M., Mu, G., Yang, F., & Zhu, J. (2010). Ordination as a tool to
900 characterize soil particle size distribution, applied to an elevation gradient at the north slope of
901 the Middle Kunlun Mountains. Geoderma, 158(3-4), 352–358.
902 https://doi.org/10.1016/j.geoderma.2010.06.002
903 Hartmann, W. K., & Neukum, G. (2001). Cratering chronology and the evolution of Mars. Space
904 Science Reviews, 96(1), 165-194. https://doi.org/10.1023/A:1011945222010
905 Hirmas, D. R., Allen, B. L. (2007). Degradation of pedogenic calcretes in west Texas. Soil Science
906 Society of America Journal, 71(6), 1878-1888. https://doi.org/10.2136/sssaj2006.0351
907 Jindaluang, W., Kheoruenromne, I., Suddhiprakarn, A., Singh, B. P., & Singh, B. (2013). Influence
908 of soil texture and mineralogy on organic matter content and composition in physically
909 separated fractions soils of Thailand. Geoderma, 195, 207-219.
910 https://doi.org/10.1016/j.geoderma.2012.12.003
911 Jin, Z., Dong, Y. S., Qi, Y. C., Liu, W. G., & An, Z. S. (2011). CHARACTERIZING VARIATIONS
912 IN SOIL PARTICLE-SIZE DISTRIBUTION ALONG A GRASS-DESERT SHRUB
913 TRANSITION IN THE ORDOS PLATEAU OF INNER MONGOLIA, CHINA. Land
914 Degradation & Development, 24(2), 141–146. https://doi.org/10.1002/ldr.1112
915 Liu, X., Zhang, G., Heathman, G. C., Wang, Y., Huang, C. H. (2009). Fractal features of soil particle-
916 size distribution as affected by plant communities in the forested region of Mountain Yimeng,
917 China. Geoderma, 154(1-2), 123-130. https://doi.org/10.1016/j.geoderma.2009.10.005
918 Milliken, R. E., Grotzinger, J. P., & Thomson, B. J. (2010). Paleoclimate of Mars as captured by the
919 stratigraphic record in Gale Crater. Geophysical Research Letters, 37(4).
920 https://doi.org/10.1029/2009gl041870
921 Minasny, B., McBratney, A. B., Field, D. J., Tranter, G., McKenzie, N. J., & Brough, D. M. (2007).
922 Relationships between field texture and particle-size distribution in Australia and their
923 implications. Soil Research, 45(6), 428-437. https://doi.org/10.1071/sr07051
924 Obara, H., Maejima, Y., Kohyama, K., Ohkura, T., & Takata, Y. (2015). Outline of the
925 comprehensive soil classification system of Japan–first approximation. Japan Agricultural
926 Research Quarterly: JARQ, 49(3), 217-226. https://doi.org/10.6090/jarq.49.217
927 Park, D. W. (1987). The loes-like red yellow soil of the south western coastal area of Korea in
928 comparison with the loess of China and Japan. Geojournal, 15, 197-200.

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

https://doi.org/10.1007/bf00157947

Rasmussen, C., Matsuyama, N., Dahlgren, R. A., Southard, R. J., Brauer, N. (2007). Soil genesis and mineral transformation across an environmental gradient on andesitic lahar. Soil Science Society of America Journal, 71(1), 225-237. https://doi.org/10.2136/sssaj2006.0100

Salako, F. K., Tian, G., Kirchhof, G., & Akinbola, G. E., 2006. Soil particles in agricultural landscapes of a derived savanna in southwestern Nigeria and implications for selected soil properties. Geoderma, 137(1-2), 90-99. https://doi.org/10.1016/j.geoderma.2006.06.017

Samiei-Fard, R., Heidari, A., Konyushkova, M., & Mahmoodi, S. (2021). Application of particle size distribution throughout the soil profile as a criterion for recognition of newly developed geoforms in the Southeastern Caspian coast. Catena, 203, 105362. https://doi.org/10.1016/j.catena.2021.105362

Schon, S. C., Head, J. W., & Fassett, C. I. (2012). An overfilled lacustrine system and progradational delta in Jezero crater, Mars: Implications for Noachian climate. Planetary and Space Science, 67(1), 28-45. https://doi.org/10.1016/j.pss.2012.02.003

Shaw, J. N., Odom, J. W., Hajek, B. F. (2003). Soils on quaternary terraces of the Tallapoosa River, Central Alabama. Soil science, 168(10), 707-717. https://doi.org/10.1097/01.ss.0000095144.68539.9b

Spriggs, G., Ray-Maitra, A. (2007). Particle-Size-Distribution of Nevada Test Site Soils (No. UCRL-TR-234937). Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States). https://doi.org/10.2172/922100

Torrent, J., Nettleton, W. D., Borst, G. (1980). Clay illuviation and lamella formation in a Psammentic Haploxeralf in southern California. Soil Science Society of America Journal, 44(2), 363-369. https://doi.org/10.2136/sssaj1980.03615995004400020031x

Wurman, E., Whiteside, E. P., & Mortland, M. M. (1959). Properties and genesis of finer textured subsoil bands in some sandy Michigan soils. Soil Science Society of America Journal, 23(2), 135-143. https://doi.org/10.2136/sssaj1959.03615995002300020017x

Supplementary Information

Supplementary Tables S1–S5 provide detailed site metadata for terrestrial samples organized

https://doi.org/10.5194/essd-2026-206
Preprint. Discussion started: 16 April 2026

Open Access Earth System Science Data Discussions

by geographic region. Supplementary Table S6 provides lunar sample metadata. Supplementary Tables S7–S8 provide Martian sub-site distribution and geological context.