Finding Earth's Best Testbeds for Moon and Mars Rovers: A Quantitative Terrain Comparison

Chaitanya Sanganeria[1]

Received September 28, 2025
Accepted April 12, 2026
Electronic access April 30, 2026

Terrestrial analog site identification for planetary rover testing often relies on logistical convenience, qualitative geomorphological similarity, and the availability of field infrastructure. While these approaches have produced widely used analog sites, it remains difficult to scale or systematize analog selection across large geographic regions using consistent, quantitative criteria. In this study, I present a low-cost, image-based screening framework that ranks terrestrial candidate sites by their visual-textural similarity to representative lunar and Martian terrains. Using grayscale orbital/satellite imagery, I compute two image-derived texture descriptors: (i) a brightness variability index capturing spatial intensity heterogeneity, and (ii) a local-gradient roughness proxy capturing small-scale intensity transitions. These descriptors are not treated as physical elevation or mechanical roughness; instead, they are standardized image-texture measures extracted under a consistent preprocessing pipeline, with images reprojected and resampled to a common spatial scale to reduce resolution-driven bias. Similarity between Earth sites and extraterrestrial reference regions is quantified in a multi-dimensional feature space that incorporates both central tendency and spatial variability of the descriptors. Applied to a curated multi-body image set and evaluated with bootstrap resampling, the framework yields stable rankings and frequently places established analog regions (including Devon Island) among the closest matches within this dataset. I emphasize that this method is intended as an initial screening tool rather than a substitute for physics-based terrain assessment, and I discuss key limitations related to illumination geometry, albedo/sensor differences, and the need for validation against elevation-derived roughness measures.

Keywords: planetary analogs; rover testing; terrain similarity; image texture; Moon; Mars; Devon Island; remote sensing

Introduction

Planetary rovers are routinely tested on Earth before deployment in extraterrestrial environments, yet the selection of appropriate terrestrial test sites remains largely heuristic [1,2]. Sites such as Devon Island in the Canadian High Arctic, volcanic terrains in Iceland, and hyper-arid regions of the Atacama Desert are frequently chosen based on qualitative similarity, logistical accessibility, or historical precedent [2–5]. While these locations have proven valuable, the absence of a standardized, scalable method for comparing terrestrial terrains to planetary surfaces complicates the identification of new or complementary analog sites across large regions [1,2].

Remote sensing imagery offers a potential avenue for systematic, global-scale screening of candidate terrains [1,6,7]. Large archives of Earth, lunar, and Martian imagery enable comparative analysis without immediate in situ measurements [7]. However, the use of image-derived quantities must be approached cautiously. Pixel intensity patterns are influenced by illumination geometry (incidence/emission/phase angles), surface albedo and composition, atmospheric effects (Earth), sensor bandpasses, onboard processing, and spatial resolution [8,9]. Consequently, intensity-based descriptors should not be naively interpreted as physical elevation or mechanical roughness, especially across heterogeneous sources [8].

The objective of this study is therefore limited but practical: to test whether simple, consistently extracted image-texture descriptors can be used to rank terrestrial sites by their visual-textural similarity to selected lunar and Martian terrains [10,11]. Rather than attempting to infer rover mobility directly, I operationalize "resemblance" as proximity in a defined feature space derived from image-texture statistics [10]. The resulting rankings are intended to support early-stage site screening and prioritization, to be followed by more detailed analyses using photometric correction, digital elevation models (DEMs) [12,13], and field-based validation in established analog programs [2,6,8].

Scope and Definitions

Throughout this paper, we use the term "similarity" to denote proximity in a defined image-feature space, not equivalence in physical, mechanical, or geological properties [8]. Two image-derived quantities are employed:

[1] *Institut Le Rosey*

- Brightness variability index (BVI): a normalized measure of spatial variation in grayscale intensity within an image.

- Roughness proxy (RP): a texture descriptor based on local intensity gradients given by:

$$R = (1/N) \sum_{i=1}^{N} \sum_{\text{neighbors}} |p_i - p_j|$$

Where p_i is the pixel and p_j is the average value of the pixels around it

Importantly, neither BVI nor RP is interpreted as physical elevation, slope, rock abundance, or mechanical roughness[8]. Because illumination, albedo, and sensor processing can dominate intensity patterns, these descriptors are best viewed as standardized image-texture measures that may correlate with terrain appearance under constrained conditions but are not direct physical measurements[8,9]. The method does not model gravity, regolith mechanics, or rover–terrain interaction and must not be used to predict mobility or traction performance[6]. Its role is limited to comparative visual screening under a standardized processing pipeline.

Dataset

A curated dataset of grayscale or grayscale-converted orbital/satellite images was assembled for Earth, the Moon, and Mars from publicly available scientific image archives[6]. For each planetary body, ten images were used in the primary analysis. Images were selected to represent relatively homogeneous terrains; Earth scenes were chosen to minimize cloud cover and strong atmospheric obscuration.

Reproducibility note: The present manuscript reports region names and computed descriptor summaries for each scene. Full per-image provenance (sensor/mission, acquisition date/time, resolution in m/pixel, projection/CRS, and a retrievable identifier/link) is necessary for strict replication and is identified as a required improvement in future revisions and follow-on work[6,14].

Table 1 summarizes the dataset and computed descriptors used in this study. To avoid physical overinterpretation, the former "altitude" columns are explicitly treated as normalized intensity statistics (scaled for readability) and reported in arbitrary units (a.u.), not meters[8].

Image Preprocessing

All images were converted to grayscale prior to analysis. Natural-color Earth images were converted using a standard luminance transformation:

$$I = 0.299R + 0.587G + 0.114B,$$

where R, G, and B denote the red, green, and blue channels, respectively. Lunar and Martian images, which were already grayscale, were processed directly[7].

To reduce the influence of absolute illumination differences, grayscale intensities were normalized on a per-image basis using min–max normalization:

$$I_{\text{norm}} = (I - I_{\min})/(I_{\max} - I_{\min}),$$

where $I_{\min}$ and $I_{\max}$ are the minimum and maximum pixel intensities within the image. This normalization preserves relative spatial variability while discarding absolute brightness offsets. However, normalization alone does not correct for differing illumination geometries or photometric functions; therefore, intensity-based descriptors may still be confounded by phase angle, incidence/emission angles, shadowing, and surface reflectance properties[8,9]. In the Discussion, I treat these factors as primary limitations and motivate photometric correction as future work[8,9].

Image-Derived Metrics

Brightness Variability Index

The brightness variability index (BVI) quantifies spatial heterogeneity in normalized intensity. For each image, the mean and standard deviation of I_{norm} were computed over all valid pixels. The standard deviation serves as the primary descriptor of brightness variability, while the mean provides a complementary measure of overall tone. Because intensity is influenced by photometric and sensor factors, BVI is treated strictly as a texture/appearance descriptor rather than a physical topography measure[8,9].

Roughness Proxy

RP is computed from local intensity gradients. For each pixel p_i, the absolute difference between its intensity and the mean intensity of its 8-connected neighborhood is computed[15], and the image-level RP is defined as the mean of these local differences over all valid pixels. Similar local-gradient descriptors are widely used in image texture analysis to quantify edge density and high-frequency spatial variation[10,11]. RP therefore captures small-scale intensity transitions, but it does not directly measure rover-scale mechanical roughness or hazards without validation against physical ground truth (e.g., DEM-derived slope/roughness or field measurements)[8,14].

To assess scale sensitivity, RP was recalculated for a subset of images after downsampling/resampling experiments, confirming that absolute values vary with resolution while relative rankings can remain stable within a narrow, controlled

Table 1 Surface roughness and normalized density of terrestrial, lunar, and Martian analog sites

Planet	Location	Country/Region	Mean normalized intensity (scaled; a.u.)	Std. normalized roughness proxy (a.u.)	Mean roughness proxy (a.u.)
Earth	Askja Caldera	Iceland	86.03	27.39	6.07
Earth	North Atacama Desert	Chile	152.17	55.32	15.8
Earth	South Atacama Desert	Chile	138.71	51.14	15.22
Earth	Devon Island	Canada	129.95	27.64	11.11
Earth	Mauna Kea	United States-Hawaii	95.74	25.74	6.14
Earth	Meteor Crater	United States-Arizona	152.49	42.21	13.54
Earth	Nevada Desert	United States-Nevada	175.73	24.49	5.51
Earth	Nevada Test Site	United States-Nevada	156.19	26.88	6.78
Earth	Utah Desert	United States-Utah	137.96	34.8	18.01
Earth	Wadi Rum	Jordan	159.23	32.52	13.17
Moon	Mare Serenitatis	Moon1	152.49	50.32	6.82
Moon	Mare Imbrium	Moon2	195.14	49.55	7.76
Moon	Lacus Somniorum	Moon3	120.57	32.29	5.72
Moon	Oceanus Procellarum	Moon4	156.03	29.37	10.73
Moon	Mare Crisium	Moon5	120.61	61.43	11.7
Moon	Mare Nectaris	Moon6	111.08	32.95	4.2
Moon	Mare Nubium	Moon7	133.33	29.67	5.22
Moon	Mare Frigoris	Moon8	113.71	67.05	10.73
Moon	Mare Vaporum	Moon9	98.94	21.98	2.65
Moon	Mare Cognitum	Moon10	105.8	32.04	2.46
Mars	Elysium Planitia	Mars1	88.43	23.4	13.27
Mars	Isidis Planitia	Mars2	57.46	27.91	5.53
Mars	Noctis Labyrinthus	Mars3	155.91	22.98	8.47
Mars	Gale Crater	Mars4	113.09	38.34	6.55
Mars	Valles Marineris	Mars5	108.93	31.57	7.18
Mars	Hellas Basin	Mars6	133.88	43.53	6.28
Mars	Arabia Terra	Mars7	108.8	38.06	10.68
Mars	Utopia Planitia	Mars8	133.47	39.95	4.09
Mars	Terra Cimmeria	Mars9	134.55	27.98	5.13
Mars	Argyre Planitia	Mars10	111.14	34.09	8.47

resolution band. This sensitivity motivates reporting resolution metadata and enforcing a common analysis scale [14].

Similarity Metric

Similarity between terrestrial and extraterrestrial terrains was evaluated in a four-dimensional feature space consisting of the mean and standard deviation of normalized intensity and the mean and standard deviation of the roughness proxy [10]:

$$F = [\text{mean}(I_{\text{norm}}), \text{std}(I_{\text{norm}}), \text{mean}(RP), \text{std}(RP)].$$

For each Earth site, a feature vector F_{Earth} was constructed and compared to the centroid of the Moon and Mars reference feature sets, computed as the mean feature vector across the lunar images and across the Martian images, respectively.

Euclidean distance in this feature space was used as a dissimilarity measure. For clarity, lower distances correspond to higher similarity. Distances were computed between each Earth site and reference feature centroids for the Moon and Mars computed from the corresponding image sets [10].

Table 2 reports the computed distances (labeled as "Similarity" in the original draft but interpreted here as distances). To avoid ambiguity, these values should be interpreted as distances d, where smaller d indicates greater similarity.

$$\text{Similarity Score} = \frac{1}{1 + \sqrt{(E_{\text{avg}} - P_{\text{avg}})^2 + (R_{\text{avg}} - Q_{\text{avg}})^2}}$$

In this formula:

$E_{avg} \rightarrow$ The average elevation (altitude) of the Earth image being compared.

$P_{avg} \rightarrow$ The average elevation (altitude) of the reference planetary body (Moon or Mars).

$R_{avg} \rightarrow$ The average roughness of the Earth image being compared.

$Q_{avg} \rightarrow$ The average roughness of the reference planetary body (Moon or Mars).

Sensitivity Analysis

To assess robustness, a bootstrap resampling analysis was performed. Reference feature centroids for the Moon and Mars were recomputed multiple times with one image removed at each iteration. Earth sites were re-ranked for each resample, and the frequency with which each site achieved the top rank was recorded. This analysis provides a measure of ranking stability and guards against overinterpretation based on small sample sizes [16].

Table 2 Dissimilarity (Euclidean distance; lower is more similar).

Earth Image Name	Similarity to Moon	Similarity to Mars
Askja_Caldera.jpg	0.022	0.034
Atacama_Desert.jpg	0.041	0.025
Atacama_Desert2.jpg	0.080	0.038
Devon_Island.jpg	0.186	0.060
Mauna_Kea.jpg	0.028	0.050
Meteor_Crater_Arizona.jpg	0.042	0.025
Nevada_Desert.jpg	0.022	0.016
Nevada_Test_Site.jpg	0.038	0.023
Utah_Desert.jpg	0.070	0.038
Wadi_Rum_Jordan.jpg	0.033	0.022

Results

Using the defined feature space and Euclidean distance, the framework produces a ranked list of terrestrial sites relative to the lunar and Martian reference centroids. Based on Table 2, Nevada_Desert and Askja_Caldera are among the closest matches to the lunar reference set (distance 0.022), while Nevada_Desert provides the closest match to the Martian reference set (distance 0.016). These rankings should be interpreted strictly as within-dataset ordering of image-texture similarity under the preprocessing pipeline and selected scenes, not as a statement about rover mobility, geology, or mechanical terrain equivalence[7,8].

Devon Island is an established operational analog site with extensive prior field use[17], but in this specific image-texture dataset it does not rank among the closest matches by the current distance metric[2]. This discrepancy underscores a key limitation of intensity-based screening: image appearance can be strongly affected by illumination geometry, albedo/composition, and sensor processing choices, and any single curated image set may not represent the full diversity of a region[8,9]. A more rigorous evaluation would require (i) more comprehensive sampling per region, (ii) standardized acquisition constraints and metadata reporting, and (iii) validation against physics-based or elevation-derived terrain measures[8,14].

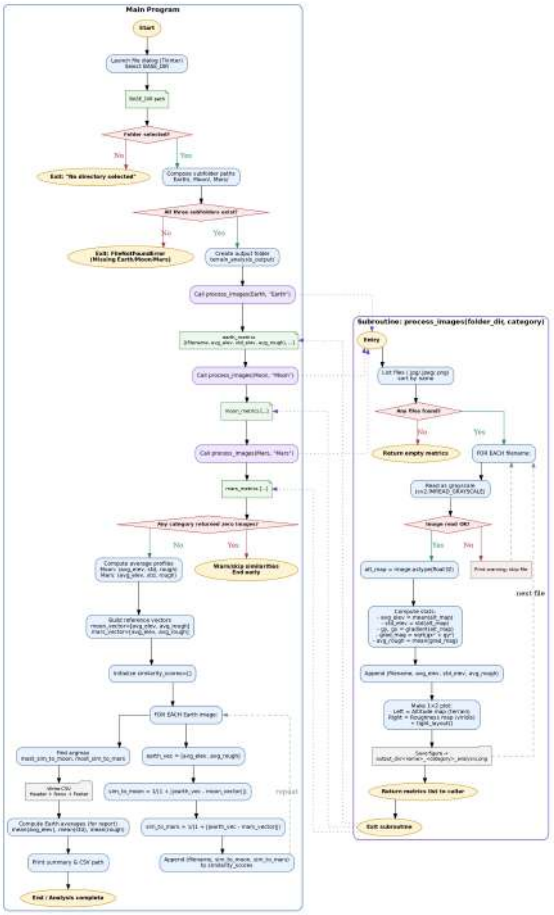

Fig. 1 Detailed flowchart illustrating the terrain analysis program workflow. The diagram depicts the complete process from folder selection to similarity computation, including the process_images subroutine. It outlines key decision points (e.g., folder validation, image read success), iterative loops for processing multiple images, and error-handling branches. The flow clearly distinguishes between the main program logic and the subroutine, showing data exchanges and the generation of analysis outputs such as visual maps, statistical summaries, similarity scores, and CSV reports.

Discussion

The primary contribution of this study is a lightweight screening baseline that can be applied broadly using widely available imagery. The method is deliberately simple and therefore limited: it reduces each scene to a small set of image-texture summary statistics, which may support early-stage prioritization but cannot replace physics-based terrain assessment and rover engineering evaluation[18].

Illumination geometry and photometric response are major confounders. Although per-image min–max normalization reduces absolute brightness offsets, it does not enforce comparable incidence/emission angles or account for bidirectional reflectance effects[19]. Consequently, differences in I_{norm} statistics and gradient-derived RP may reflect imaging conditions as much as surface structure. In future work, photometric correction and explicit reporting of illumination metadata would be necessary to strengthen cross-body comparisons.

Resolution dependence is another central limitation. Gradient-based descriptors can change with spatial resolution (m/pixel), and strict comparability requires a controlled spatial

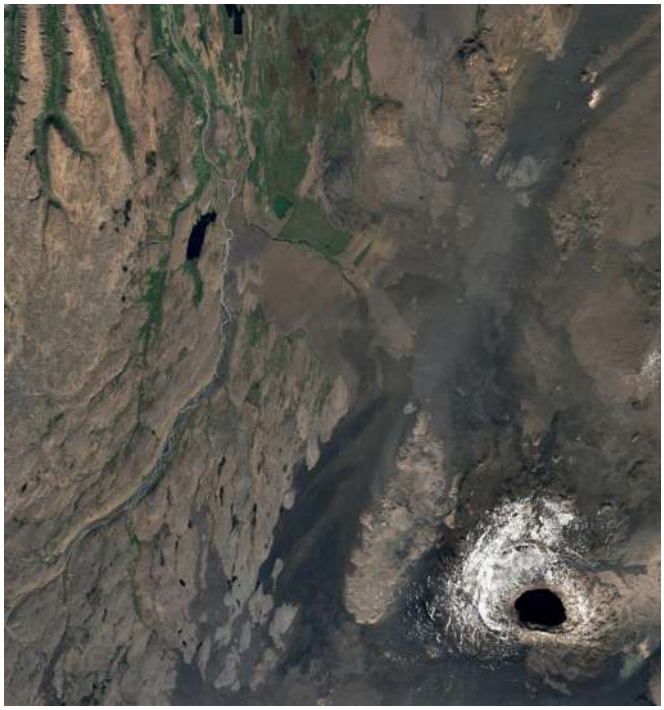

Fig. 2 Natural-color satellite image of Askja Caldera, Iceland, showing volcanic terrain with surrounding ridges and snow-covered areas

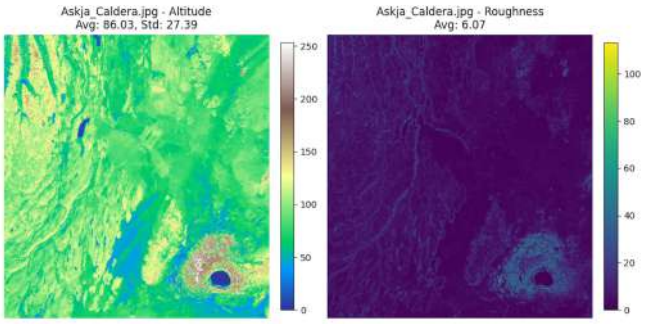

Fig. 3 Terrain analysis of Askja Caldera, Iceland. The left panel is the altitude map (average altitude = 86.03 m, standard deviation = 27.39 m). The right panel shows the roughness index map (average roughness = 6.07).

scale and explicit resampling strategy[12]. Additionally, collapsing scenes to mean and standard deviation discards distributional and spatial-correlation information that can matter for rover traversability (e.g., clustered hazards versus diffuse roughness). Future work can incorporate distribution-sensitive texture descriptors (e.g., histograms, percentiles, lacunarity, or multi-scale features) and compare results against elevation-derived slope/roughness from DEM products[13,20].

Finally, this framework does not model gravity, regolith mechanics, or wheel–soil interactions, which are critical for rover mobility outcomes and cannot be inferred reliably from intensity texture alone[21,22]. Therefore, the method should be used only as an initial filter and followed by detailed physi-

Fig. 4 Grayscale satellite image of Mare Imbrium, Moon, highlighting impact craters and rugged lunar highlands.

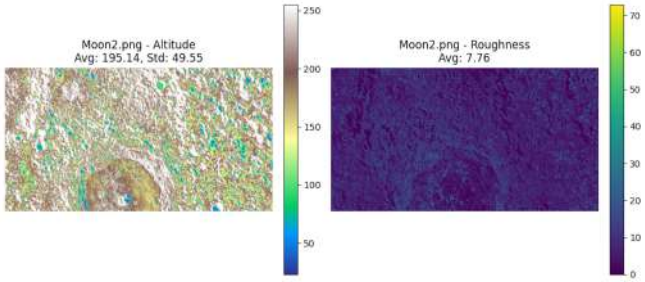

Fig. 5 Terrain analysis of Mare Imbrium, Moon. The left panel is the altitude map (average altitude = 195.14 m, standard deviation = 49.55 m). The right panel shows the roughness index map (average roughness = 7.76)

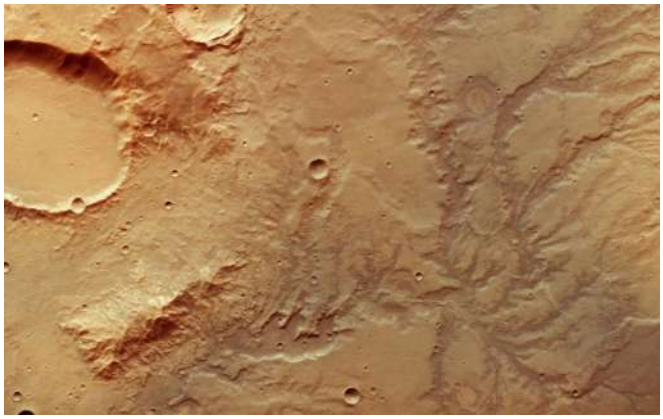

Fig. 6 Natural-color satellite image of Terra Cimmeria, Mars, showing a heavily cratered and eroded ancient highland region.

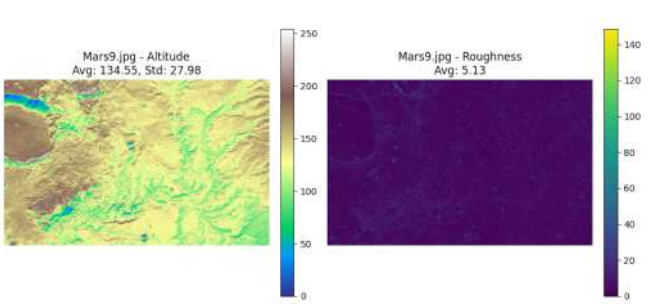

Fig. 7 Terrain analysis of Terra Cimmeria, Mars. The left panel is the altitude map (average altitude = 134.55 m, standard deviation = 27.98 m). The right panel shows the roughness index map (average roughness = 5.13).

Fig. 8 Natural-color satellite image of Devon Island, Canada, Earth's largest uninhabited island

cal characterization and/or field testing in established analog programs[3].

Conclusion

This study presents an image-based framework for preliminary screening of terrestrial rover analog sites using standardized image-texture descriptors computed under a consistent preprocessing pipeline. By operationalizing resemblance as proximity in a defined feature space and assessing ranking stability through resampling, the method provides a transparent, low-cost way to prioritize candidate sites for deeper investigation. However, because illumination, sensor differences, and spatial resolution can strongly influence intensity-derived descriptors, and because key physical factors for mobility (e.g., gravity and wheel–soil mechanics) are out of scope, the framework should be treated as an initial screening tool rather than a standalone decision method. Future work should add full per-image metadata for reproducibility, apply photometric correction where possible, and validate similarity rankings against DEM-derived roughness/slope measures and established analog-site findings.

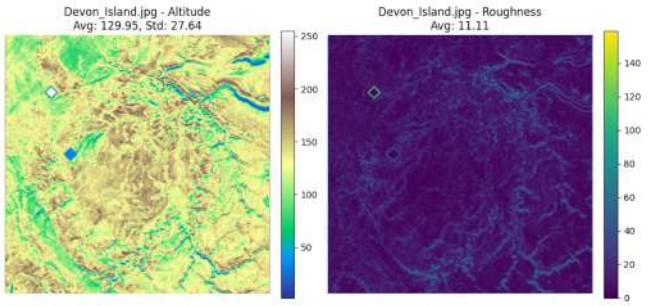

Fig. 9 Terrain analysis of Devon Island, Canada. The left panel is the altitude map (average altitude = 129.95 m, standard deviation = 27.64 m). The right panel shows the roughness index map (average roughness = 11.11).

References

1 B. Foing, C. Stoker, J. Zavaleta, P. Ehrenfreund, C. Thiel and P. Sarrazin, Field astrobiology research in Moon–Mars analogue environments: instruments and methods. Int J Astrobiol, 10(3), 141–160 (2011). https://doi.org/10.1017/S1473550411000036.

2 P. Lee and G. Osinski, The Haughton-Mars Project: overview of science investigations at the Haughton impact structure and surrounding terrains, and relevance to planetary studies. Meteorit Planet Sci, 40(12), 1755–1758 (2005). https://doi.org/10.1111/j.1945-5100.2005.tb00144.x.

3 T. Barfoot, P. Furgale, G. Osinski, N. Ghafoor and K. Williams, Devon Island as a proving ground for planetary rovers. In: Body and Machine. Advances in Intelligent and Soft Computing, vol. 83, Springer, Berlin, 277–288 (2010). https://doi.org/10.1007/978-3-642-16259-6_21.

4 G. e. a. Neumann, Lunar impact basins revealed by Gravity Recovery and Interior Laboratory measurements. Sci Adv, 1(9) (2015). https://doi.org/10.1126/sciadv.1500852.

5 A. e. a. Azua-Bustos, Hypolithic cyanobacteria supported mainly by fog in the coastal Atacama Desert of Chile. Microb Ecol, 61(3), 568–581 (2011). https://doi.org/10.1007/s00248-010-9784-z.

6 M. Malin and K. Edgett, Mars Global Surveyor Mars Orbiter Camera: Interplanetary cruise through primary mission. J Geophys Res, 106(E10), 23429–23570 (2001). https://doi.org/10.1029/2000JE001455.

7 R. Yingst, B. Cohen, B. Hynek, M. Schmidt, C. Schrader and A. Rodriguez, Testing Mars Exploration Rover-inspired operational strategies for semi-autonomous rovers on the moon II: The GeoHeuristic Oper-

ational Strategies Test in Alaska. Acta Astronaut, 99, 24–36 (2014). https://doi.org/10.1016/j.actaastro.2014.01.019.

8 A. McEwen, Photometric functions for photoclinometry and other applications. Icarus, 92(2), 298–311 (1991). https://doi.org/10.1016/0019-1035(91)90053-V.

9 C. Chen, Q. Qin, L. Chen, H. Zheng, W. Fa, A. Ghulam and C. Zhang, Photometric correction and reflectance calculation for lunar images from the Chang'E-1 CCD stereo camera. J Opt Soc Am A, 32(12), 2409–2422 (2015). https://doi.org/10.1364/JOSAA.32.002409.

10 R. Haralick, K. Shanmugam and I. Dinstein, Textural features for image classification. IEEE Trans Syst Man Cybern, SMC-3(6), 610–621 (1973). https://doi.org/10.1109/TSMC.1973.4309314.

11 R. Gonzalez and R. Woods, Digital Image Processing, 4th ed., Pearson, New York, NY (2018).

12 M. Kreslavsky and J. Head, Kilometer-scale roughness of Mars: Results from MOLA data analysis. J Geophys Res Planets, 105(E11), 26695–26711 (2000). https://doi.org/10.1029/2000JE001259.

13 A. Cord, D. Baratoux, N. Mangold, P. Martin, P. Pinet, R. Greeley, F. Costard, P. Masson, B. Foing and G. Neukum, Surface roughness and geological mapping at subhectometer scale from the High Resolution Stereo Camera onboard Mars Express. Icarus, 187(2), 348–373 (2007). https://doi.org/10.1016/j.icarus.2006.10.004.

14 D. Smith, M. Zuber, G. Neumann, F. Lemoine, E. Mazarico and M. Torrence, Initial observations from the Lunar Orbiter Laser Altimeter (LOLA). Geophysical Research Letters, 37(18) (2010). https://doi.org/10.1029/2010GL043751.

15 M. Rosenburg, O. Aharonson, J. Head, M. Kreslavsky, E. Mazarico, G. Neumann, D. Smith, M. Torrence and M. Zuber, Global surface slopes and roughness of the Moon from the Lunar Orbiter Laser Altimeter. J Geophys Res Planets, 116(E2) (2011). https://doi.org/10.1029/2010JE003716.

16 B. Efron and R. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall/CRC, New York, NY (1993).

17 R. e. a. Arvidson, Ancient aqueous environments at Endeavour Crater, Mars. Science, 343(6169) (2014). https://doi.org/10.1126/science.1248097.

18 G. Osinski, J. Spray and P. Lee, Impactites of the Haughton impact structure, Devon Island, Canadian High Arctic. Meteorit Planet Sci, 40(12), 1789–1812 (2005). https://doi.org/10.1111/j.1945-5100.2005.tb00147.x.

19 M. e. a. Shepard, The roughness of natural terrain: A planetary and remote sensing perspective. J Geophys Res Planets, 106(E12), 32777–32795 (2001). https://doi.org/10.1029/2000JE001429.

20 M. Kreslavsky, J. Head, G. Neumann, M. Rosenburg, O. Aharonson, D. Smith and M. Zuber, Lunar topographic roughness maps from Lunar Orbiter Laser Altimeter (LOLA) data: Scale dependence and correlation with geologic features and units. Icarus, 226(1), 52–66 (2013). https://doi.org/10.1016/j.icarus.2013.01.034.

21 K. Iagnemma, S. Kang, H. Shibly and S. Dubowsky, Online terrain parameter estimation for wheeled mobile robots with application to planetary rovers. IEEE Trans Robot, 20(5), 921–927 (2004). https://doi.org/10.1109/TRO.2004.829462.

22 J. e. a. Zakrajsek, Exploration rover concepts and development challenges. AIAA Space (2005). https://doi.org/10.2514/6.2005-2525.