Studies in Knowledge and Innovation
Vol. 2 No. 4A (2026)

Averaging Elevations First Changes Lunar Slope Values Without Reordering Six Sites

Jordan Wright[1] · Luke Hayes[2] · Benjamin Lewis[3]

[1] Spain · [2] Portugal · [3] Czechia

Published: April 30, 2026 | Article 837

Abstract

A regional slope ranking can survive a processing choice that changes the values beneath it. We applied two orders of calculation to six released 5 m lunar terrain models: derive native slopes and average them, or average elevations and derive new slopes. Matched 10, 20 and 40 m cells covered one fixed interior per site. Elevation-first processing lowered the equal-site mean by 0.085°, 0.167° and 0.270°, respectively, while paired-cell absolute differences averaged 0.425°, 0.632° and 0.688°. All six site rankings remained unchanged. At 40 m, 35.6–41.1% of cells nevertheless had higher elevation-first values. Numerical agreement, local direction and regional ordering therefore gave different answers. These are representation differences within interpolated terrain models, with no independent terrain-truth or vehicle-safety validation.

Averaging Elevations First Changes Lunar Slope Values Without Reordering Six Sites

Abstract

A regional slope ranking can survive a processing choice that changes the values beneath it. We applied two orders of calculation to six released 5 m lunar terrain models: derive native slopes and average them, or average elevations and derive new slopes. Matched 10, 20 and 40 m cells covered one fixed interior per site. Elevation-first processing lowered the equal-site mean by 0.085°, 0.167° and 0.270°, respectively, while paired-cell absolute differences averaged 0.425°, 0.632° and 0.688°. All six site rankings remained unchanged. At 40 m, 35.6-41.1% of cells nevertheless had higher elevation-first values. Numerical agreement, local direction and regional ordering therefore gave different answers. These are representation differences within interpolated terrain models, with no independent terrain-truth or vehicle-safety validation.

When a terrain summary travels to a coarser map

Suppose a terrain product supplies a fine elevation grid but an application needs a coarser slope map. One workflow calculates slope on the original elevation grid and averages those angles into the destination cells. Another first averages elevations and then calculates slope on the coarser surface. Both can produce a map in degrees with exactly the same dimensions. They do not necessarily produce the same numerical quantity. An analyst who retains the output grid spacing while omitting the calculation order leaves an important part of the variable undefined.

This distinction already has a terrestrial research history. Grohmann directly compared derivatives of resampled elevation with resampled terrain derivatives and showed that the order matters in regional analyses (Grohmann, 2015). Related work on DEM grid size connects landscape representation to changes in derived slope distributions (Zhang and Montgomery, 1994). Our contribution is not a new resampling method or the discovery of scale dependence. We quantify the distinction for six particular lunar south-polar terrain products, keeping their physical reporting support, source alignment and derivative operator fixed. The question is whether changed local slope values also change the ordering of these named sites.

The elevation source makes that question concrete. Barker and colleagues developed improved LOLA-based terrain models for four south-polar regions and analyzed their height uncertainties and consequences for illumination calculations (Barker et al., 2021). The associated public archive also supplies two additional legacy site products. These arrays support a controlled processing comparison because both workflows begin with the same released elevations. A surface-gradient estimator is still required: the familiar weighted neighborhood associated with Horn provides one explicit choice (Horn, 1981). We retain that choice rather than searching for whichever estimator gives the largest contrast.

The first analytical proposition, P1, is that calculation order changes the reported slope distribution and can lower its regional mean. Averaging height can suppress short spatial variations before a derivative is formed; averaging already-derived angles operates on another field. The second proposition, P2, is that a numerical change need not reorder sites. These propositions distinguish magnitude from ranking, rather than making two interchangeable definitions of success. Measurement theory provides a general reason to preserve how a numerical variable is constructed as well as its units (Stevens, 1946).

Existing terrain-resolution studies show that grid size influences first, second and compound derivatives (Kienzle, 2004). Lunar research also characterizes slopes and roughness over stated measurement

baselines (Rosenburg et al., 2011). Accordingly, our scale comparison is conditioned on its estimator and source model. It is not evidence that any grid resolution is universally best. Alternative local surface formulations exist, including the terrain-analysis approach of Zevenbergen and Thorne (Zevenbergen and Thorne, 1987); choosing among them would introduce another comparison beyond the fixed operation-order question addressed here.

Six elevation products and one reporting interior

We acquired all six surface-elevation GeoTIFFs directly linked by the dedicated NASA Goddard PGDA index at https://pgda.gsfc.nasa.gov/data/LOLA_5mpp/ on 16 November 2026. They are Site01, Connecting ridge; Site04, Shackleton rim; Site07, a peak near Shackleton; Site11, de Gerlache rim; Site20, Leibnitz beta plateau; and Site23, Malapert massif. Their filenames end in _final_adj_5mpp_surf.tif. Provider timestamps are 5 or 12 June 2020; filenames, byte counts and actual SHA256 checksums identify the versions, because no semantic release number is declared. The six files total 275,490,836 bytes.

Five tiles contain 3,200 by 3,200 Float32 cells, covering 16 by 16 projected kilometres. Site23 contains 4,200 by 4,200 cells over 21 by 21 kilometres. The native source total is 68,840,000 cells. Each tile has 5 m projected spacing and surface height Z in metres. The provider explicitly describes interpolation filling empty terrain cells. A dense output grid therefore does not mean that elevations are independently measured every five metres or that all features at that scale are resolved.

GeoTIFF tags specify a south-polar stereographic projection, central latitude -90°, central longitude 0°, unit scale at the pole and a 1,737,400 m sphere. The provider README identifies the MOON_ME frame of DE421; generic embedded datum text instead says unknown, a discrepancy we retain in the source audit. The upper-left tiepoint locates a cell edge because the rasters use PixelIsArea registration. At zero-based column c and row r, native centers are $x = x0 + 5(c + 0.5)$ and $y = y0 - 5(r + 0.5)$. Projection scale and the distinction between projected and surface distance follow the specified map geometry (Snyder, 1987).

Every source value was inspected for range and fill status before derived slopes were calculated. All six files declare NaN as the missing value; actual NaN, infinity and exact-zero counts are zero. Finite negative elevations occur in four tiles and are retained. There is no elevation-dependent selection. The provider's four-core-site accuracy results are not automatically assigned to Site20 and Site23. Data quality depends on the proposed use, including what metadata and validation are available (Wang and Strong, 1996); readable arrays alone do not establish equal uncertainty across these products.

We anchor every coarsening to its own source origin. Factors two, four and eight give 10, 20 and 40 m cells, each an arithmetic mean of a nonoverlapping group of native elevations. All source dimensions are divisible by eight. The common reporting mask selects the 40 m parent cells after removing their outermost row and column, equivalent to a fixed 40 m margin at each source edge. It contains 1,065,549 parents across the six tiles and 68,195,136 native cells. Finer reporting cells are precisely the children of those same parents. No outcome determines the crop.

That margin supplies a source buffer for every required derivative neighborhood. No padding, extrapolation or edge substitution enters. Site01 and Site04 overlap by 66 projected square kilometres in the original files; their cropped reporting domains still overlap by 64.6464 square kilometres. Other pairs do not overlap. Consequently, the summed counts describe processed tile support and include repeated geography. They are not counts of distinct terrain observations or evidence of six independent regional draws.

Two operators applied to identical cells

Tang and colleagues compared multiple raster slope algorithms at different resolutions, demonstrating why an algorithm name is part of a reproducible terrain measure (Tang et al., 2013). Here Horn's 3 by 3 estimator is fixed throughout. For spacing h, its eastward derivative is (NE + 2E + SE - NW - 2W - SW)/(8h), and its northward derivative is (NW + 2N + NE - SW - 2S - SE)/(8h). Letters name adjacent elevations. The magnitude is unchanged if the sign convention for the row direction is reversed.

We convert projected derivatives to a local reference-sphere horizontal scale. At each cell's actual center, $\rho = \sqrt{x^2 + y^2}$ and $k = 1 + \rho^2/(4R^2)$. The gradient magnitude used for slope is k times the projected gradient magnitude, because a projected distance is k times the corresponding reference-sphere distance. The reported angle is atan of that magnitude, converted to degrees. We use native centers for native slopes and the centers of the averaged elevation cells for coarse slopes. This does not calculate distance along the sloping ground, adjust for local terrain height, or validate a mechanical model.

Estimator precision and source uncertainty are distinct concerns. Florinsky analyzes the accuracy of local topographic variables and the role of derivative formulations (FLORINSKY, 1998). We do not infer that Horn supplies a truth standard. It supplies an unchanged operator whose physical neighborhood expands as spacing increases. Scale-space work likewise treats scale as an explicit dimension of a representation (Lindeberg, 1994). Our arithmetic block averaging is not claimed to implement a Gaussian scale-space construction; the useful connection is the requirement to state what changed before comparing derived quantities.

Workflow A derives native Horn angles first and takes their equal-weight mean within each destination cell. Workflow B averages native elevations first, derives Horn angles using the destination spacing and then selects the same cells. At every site and spacing, D = B - A is a paired difference in degrees. Both workflows report on identical coarse cells and the same physical interior. An agreement question consequently concerns these paired values, not just their correlation or a visual resemblance of maps (Bland and Altman, 1986).

All means use equal projected pixels. At a given grid spacing this is equivalent to averaging over projected planar area; it is not weighting by actual lunar surface area. We preserve that definition when comparing workflow means. Coarse cells in a site have equal native support, so the sitewide A mean is invariant across the three destination spacings, apart from rounding. This internal identity checks the common mask. It does not imply that A's cellwise distribution is unchanged.

For each of the 18 site-spacing combinations we calculate means, medians and 95th percentiles of A and B, paired mean difference, absolute difference, squared difference and the fraction with D positive. Quantiles use linear interpolation between adjacent ordered values. Quantile definitions can differ across statistical packages, making that convention part of the result (Hyndman and Fan, 1996). Source origin, masks and derivative buffers are retained with the outputs rather than reconstructed from a plotted outline.

The primary summary averages the six site mean differences equally at each spacing. It gives a named tile the same importance whether its side is 16 or 21 km. Weighting must follow the estimand (Solon et al., 2015); this choice describes a typical value among these six cases, rather than a unique-area average of the lunar surface. We also average the six paired-cell mean absolute differences equally. Deleting each site in turn yields six five-site summaries whose observed minimum and maximum quantify deterministic sensitivity. No confidence coverage is attached to this range.

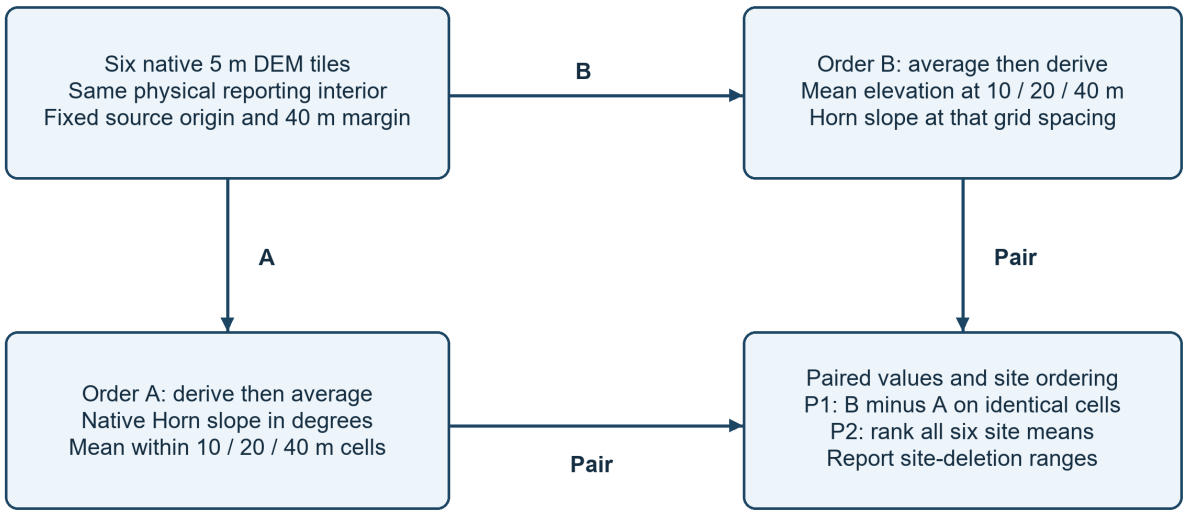


Figure 1. The two processing paths share original elevations and a fixed reporting mask. P1 concerns the paired angle contrasts; P2 concerns the ordering of all six site means. The diagram specifies operations rather than a causal intervention on actual terrain.

Mean attenuation coexists with local increases

The equal-site A mean is 14.775° at every destination spacing. Workflow B yields 14.690° at 10 m, 14.608° at 20 m and 14.505° at 40 m. The corresponding B - A contrasts are -0.085°, -0.167° and -0.270°. Every individual site's contrast is negative at all three spacings. Table 1 reports these actual effects together with their finite support and deletion sensitivity.

Table 1. Equal-site effects for six named DEM tiles. A derives native slope before averaging; B averages elevation before deriving slope. Each MAE is computed within sites and then averaged equally across six sites. The site-deletion range contains six remaining-five-site means; it is not a confidence interval. Summed paired-cell counts include overlap and do not denote independent replications.

Spacing (m)	Paired cells, summed	Mean B-A (°)	Mean paired MAE (°)	Site-deletion range for mean B-A (°)
10	17,048,784	-0.085	0.425	-0.088 to -0.082
20	4,262,196	-0.167	0.632	-0.174 to -0.159
40	1,065,549	-0.270	0.688	-0.281 to -0.252

These mean decreases are consistent with P1, but they do not say that every cell becomes less steep. At 40 m, B exceeds A in 35.6-41.1% of a site's cells. At 10 m, the corresponding proportions are 43.8-45.9%. A smaller regional mean therefore coexists with a substantial area of local positive differences. The changing mixture is compatible with smoothing and a wider derivative neighborhood redistributing local gradients; this experiment does not isolate either mechanism's contribution.

Classic image-scale theory describes changes in represented structure as resolution changes (Koenderink, 1984). The present local increases should not be read as evidence that coarsening creates actual steeper terrain. They occur in a different numerical representation of the same elevation source. Because the slope transformation includes gradient magnitude and an angle, averaging it is not the same operation as first

averaging a signed height field. The finite positive fractions are reported rather than replaced by a universal claim about smoothing direction.

Paired-cell absolute differences average 0.425°, 0.632° and 0.688° across the sites. Each is larger than its corresponding signed mean shift. The increase from 10 to 40 m therefore affects local agreement more than a single regional bias number suggests. Lin's distinction between concordance and simple association is relevant to this interpretation (Lin, 1989): a preserved regional relationship does not establish equality of the values that an application would read.

Table 2. All six site mean slopes under the two workflows. Common physical support makes A's regional mean invariant across destination spacings. These are finite projected-pixel summaries; the table makes no site-selection recommendation.

Site	Mean A, all spacings (°)	Mean B, 10 m (°)	Mean B, 20 m (°)	Mean B, 40 m (°)
Site01	12.577	12.478	12.369	12.213
Site04	20.692	20.622	20.558	20.476
Site07	10.741	10.651	10.552	10.437
Site11	14.401	14.328	14.261	14.174
Site20	12.831	12.753	12.679	12.587
Site23	17.407	17.307	17.227	17.140

The ascending site order is Site07, Site01, Site20, Site11, Site23 and Site04. It is identical for A and B at every spacing. None of the 15 site pairs reverses in any of the three comparisons, and no exact ties occur. Thus P2 is realized in these data: the workflow contrast changes values without changing this particular ranking. Studies of DEM resolution and terrain-based modeling show why both variable characterization and the downstream use deserve separate scrutiny (Thompson et al., 2001). Our ranking result supplies no guarantee for an additional site or a more local ordering within one tile.

Distributions, sensitivity and the limits of the comparison

Table 3 reports both workflows' medians and upper percentiles for every site and spacing. The upper 95th percentile is lower for B in all 18 comparisons. At Site04 and 40 m, for example, A's 95th percentile is 33.43° and B's is 32.96°. At Site07 the corresponding values are 17.98° and 17.50°. These are positions in finite paired-field distributions, not estimates of a permissible operating slope.

Table 3. Complete site-spacing distribution summaries on matched coarse cells. Percentiles use the stated linear interpolation rule. Positive-difference percentages refer to cells with B > A, not to missingness or classification accuracy.

Site	Spacing (m)	A median (°)	B median (°)	A 95th (°)	B 95th (°)	Cells with B>A (%)
Site01	10	12.29	12.23	22.64	22.44	43.9
Site01	20	12.31	12.16	22.49	22.19	40.8
Site01	40	12.32	12.01	22.32	21.90	35.6
Site04	10	20.65	20.56	34.26	33.82	45.9
Site04	20	20.63	20.45	33.87	33.31	44.3
Site04	40	20.55	20.31	33.43	32.96	41.1
Site07	10	10.63	10.57	18.45	18.19	43.8
Site07	20	10.64	10.52	18.26	17.83	41.1
Site07	40	10.67	10.42	17.98	17.50	37.0
Site11	10	14.42	14.40	23.82	23.62	45.6
Site11	20	14.46	14.37	23.67	23.41	43.6
Site11	40	14.48	14.32	23.52	23.24	39.9
Site20	10	11.62	11.54	23.99	23.76	44.2
Site20	20	11.61	11.44	23.82	23.47	42.3
Site20	40	11.57	11.30	23.58	23.23	38.4
Site23	10	17.92	17.89	27.72	27.46	44.9
Site23	20	17.97	17.86	27.51	27.21	43.2
Site23	40	17.99	17.78	27.31	27.07	39.7

We retain absolute and squared discrepancies because they assign different emphasis to local differences. RMSE weights large departures more heavily than MAE (Chai and Draxler, 2014). Neither measure can become an accuracy score without an independently justified reference. Here A and B are alternative source transformations. MAE expresses their average separation in degrees (Willmott and Matsuura, 2005), while the signed contrast captures the net directional change after local positive and negative differences have canceled.

At 40 m the six site mean differences range from -0.363° to -0.216°. The equal-site mean remains between -0.281° and -0.252° when any one tile is removed. The analogous deletion ranges are -0.088° to -0.082° at 10 m and -0.174° to -0.159° at 20 m. These calculations show that the mean direction is not determined by a single named tile. They do not quantify source-height uncertainty or the probability that another terrain region behaves similarly. Separating model uncertainty from deterministic sensitivity is necessary for interpreting such intervals (Smith, 2013).

The six tiles are a provider-defined legacy cohort, not a probability sample of candidate lunar sites. Sampling theory cannot turn complete processing of a purposive domain into population representativeness (Neyman, 1934). Site overlap further prevents an interpretation based on six

independent terrains. We leave the overlap intact because the estimand gives equal weight to the released named products, and we report it explicitly rather than describing summed cell counts as unique surface coverage.

More fundamentally, interpolation, altimeter coverage and residual source errors precede both workflows. An operation-order contrast is not an estimate of those errors. Work on measurement error emphasizes that a variable's uncertain construction can affect later quantitative relationships (Carroll et al., 2006). Neither a fine grid nor an unchanged ranking removes that issue. The original four-site source paper's uncertainty analysis addresses another question with additional assumptions and evidence; it cannot be replaced by the small differences between our two workflows.

The broader lesson is that a geometric measurement carries a scale definition. Mandelbrot's classic coastline example made dependence on measurement scale explicit (Mandelbrot, 1967), but we fit no fractal model and extrapolate no scaling law from three destination spacings. Our claim is confined to the observed slope differences at 10, 20 and 40 m. Terrain physics, interpolation and estimator footprint would all need attention before applying another scale.

What a numeric slope input should retain

The complete comparison answers two separate uses of a terrain summary. If the use is this six-site ranking, both workflows return the same order. If the use is a local or regional numerical angle, the processing order remains consequential and must be specified. Keeping the original elevations, the gradient definition, grid origin, scale correction and averaging measure makes that value reproducible. It does not establish which value a vehicle can tolerate.

Lunar-R1's qualitative terrain case discusses checking slope and obstacle parameters against mechanical limits (Xiao et al., 2026). That use makes the provenance of a supplied angle relevant: a slope input needs a calculation scale and support, even when its subsequent reasoning is textual. We do not reconstruct Lunar-R1's example values, execute its models, assess its answers or draw conclusions about its training procedure. Our evidence concerns the environmental measurement supplied before any such reasoning.

Evaluation research argues that a benchmark's observations must support its intended construct (Bowman and Dahl, 2021). The same discipline applies to the boundary between a map-processing result and a capability claim. A representation contrast cannot establish planning competence, and a preserved site ranking cannot establish terrain safety. Broad benchmark critiques warn against allowing one convenient measurement to stand for many abilities (Raji et al., 2021). No ability score is computed here.

The protocol was written after source inspection and before calculating the slope contrasts. This is a documented secondary analysis rather than prospective registration. The archive retains exact image identities, common masks, 18 summary rows and per-parent spatial moments. An independent scalar calculation reconstructs native and coarse stencils from original heights at fixed positions, and a second aggregation checks the reported effects over all parent summaries. Such recoverable links between inputs and numbers are central to computational reproducibility (Peng, 2011).

The retained records also state access conditions, coordinate conventions and transformations needed to reuse the computation, consistent with FAIR's emphasis on reusable research objects (Wilkinson et al., 2016). Accountable human source verification and final approval remain pending. Within the six acquired terrain products, elevation-first processing lowers regional mean slope while leaving the complete site order intact; local differences in both directions show why those two findings should travel together.

References

1. Mandelbrot, Benoit. (1967). How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. Science. https://doi.org/10.1126/science.156.3775.636
2. Raji, Inioluwa Deborah; Bender, Emily M.; Paullada, Amandalynne; Denton, Emily; Hanna, Alex. (2021). AI and the Everything in the Whole Wide World Benchmark. arXiv preprint. https://arxiv.org/abs/2111.15366
3. Rosenburg, M. A.; Aharonson, O.; Head, J. W.; Kreslavsky, M. A.; Mazarico, E.; Neumann, G. A.; Smith, D. E.; Torrence, M. H.; Zuber, M. T.. (2011). Global surface slopes and roughness of the Moon from the Lunar Orbiter Laser Altimeter. Journal of Geophysical Research. https://doi.org/10.1029/2010je003716
4. Lin, Lawrence I-Kuei. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics. https://doi.org/10.2307/2532051.
5. Zevenbergen, Lyle W.; Thorne, Colin R.. (1987). Quantitative analysis of land surface topography. Earth Surface Processes and Landforms. https://doi.org/10.1002/esp.3290120107
6. Kienzle, Stefan. (2004). The Effect of DEM Raster Resolution on First Order, Second Order and Compound Terrain Derivatives. Transactions in GIS. https://doi.org/10.1111/j.1467-9671.2004.00169.x
7. Thompson, James A.; Bell, Jay C.; Butler, Charles A.. (2001). Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma. https://doi.org/10.1016/s0016-7061(00)00081-1
8. Neyman, Jerzy. (1934). On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. Journal of the Royal Statistical Society,97(4),558-606. https://doi.org/10.1111/j.2397-2335.1934.tb04184.x.
9. Willmott, CJ; Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research. https://doi.org/10.3354/cr030079.
10. Lindeberg, Tony. (1994). Scale-space theory: a basic tool for analyzing structures at different scales. Journal of Applied Statistics. https://doi.org/10.1080/757582976
11. Barker, Michael K.; Mazarico, Erwan; Neumann, Gregory A.; Smith, David E.; Zuber, Maria T.; Head, James W.. (2021). Improved LOLA elevation maps for south pole landing sites: Error estimates and their impact on illumination conditions. Planetary and Space Science. https://doi.org/10.1016/j.pss.2020.105119
12. Zhang, Weihua; Montgomery, David R.. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. Water Resources Research. https://doi.org/10.1029/93wr03553
13. Solon, Gary; Haider, Steven J.; Wooldridge, Jeffrey M.. (2015). What Are We Weighting For?. Journal of Human Resources. https://doi.org/10.3368/jhr.50.2.301.
14. Grohmann, Carlos H.. (2015). Effects of spatial resolution on slope and aspect derivation for regional-scale analysis. Computers & Geosciences. https://doi.org/10.1016/j.cageo.2015.02.003
15. Stevens, S. S.. (1946). On the Theory of Scales of Measurement. Science. https://doi.org/10.1126/science.103.2684.677.
16. Wang, Richard Y.; Strong, Diane M.. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems. https://doi.org/10.1080/07421222.1996.11518099
17. Smith, Ralph C.. (2013). Uncertainty Quantification: Theory, Implementation, and Applications. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611973228
18. Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. “Measurement Error in Nonlinear Models.” 2006. https://doi.org/10.1201/9781420010138.
19. Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; ’t Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://doi.org/10.1038/sdata.2016.18.
20. Chai, T.; Draxler, R. R.. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. Geoscientific Model Development. https://doi.org/10.5194/gmd-7-1247-2014.

21. Snyder, John P.. (1987). Map projections: A working manual. Professional Paper. https://doi.org/10.3133/pp1395

22. FLORINSKY, IGOR V.. (1998). Accuracy of local topographic variables derived from digital elevation models. International Journal of Geographical Information Science. https://doi.org/10.1080/136588198242003

23. Peng, Roger D.. (2011). Reproducible Research in Computational Science. Science. https://doi.org/10.1126/science.1213847

24. Hyndman, Rob J.; Fan, Yanan. (1996). Sample Quantiles in Statistical Packages. The American Statistician. https://doi.org/10.1080/00031305.1996.10473566

25. Koenderink, Jan J.. (1984). The structure of images. Biological Cybernetics. https://doi.org/10.1007/bf00336961

26. Tang, Jing; Pilesjö, Petter; Persson, Andreas. (2013). ESTIMATING SLOPE FROM RASTER DATA - A TEST OF EIGHT ALGORITHMS AT DIFFERENT RESOLUTIONS IN FLAT AND STEEP TERRAIN. Geodesy and Cartography. https://doi.org/10.3846/20296991.2013.806702

27. Bland, J. Martin, and Douglas G. Altman. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet*, vol. 327, no. 8476, 1986, pp. 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8.

28. Horn, B.K.P.. (1981). Hill shading and the reflectance map. Proceedings of the IEEE. https://doi.org/10.1109/proc.1981.11918

29. Xiao, Xin-Yu; Tian, Ye; He, Zhixian; Chen, Qilin; Wang, Shiqi; Xia, Qianchen. (2026). Lunar-R1: Towards Efficient Reasoning of Large Language Models for Lunar Exploration via Reinforcement Learning. Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing: Industry Track. https://aclanthology.org/2026.emnlp-industry.5/.

30. Bowman, Samuel R.; Dahl, George. (2021). What Will it Take to Fix Benchmarking in Natural Language Understanding?. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. https://doi.org/10.18653/v1/2021.naacl-main.385