Uncertainty analysis of digital elevation models by spatial inference from stable terrain
Romain Hugonnet, Fanny Brun, Etienne Berthier, Amaury Dehecq, Erik Schytt Mannerfelt, Nicolas Eckert, Daniel Farinotti

To cite this version:
Romain Hugonnet, Fanny Brun, Etienne Berthier, Amaury Dehecq, Erik Schytt Mannerfelt, et al.. Uncertainty analysis of digital elevation models by spatial inference from stable terrain. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE, 2022, pp.1-17. 10.1109/JSTARS.2022.3188922. hal-03719624

HAL Id: hal-03719624
https://hal.archives-ouvertes.fr/hal-03719624
Submitted on 11 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Uncertainty analysis of digital elevation models by spatial inference from stable terrain

Romain Hugonnet, Fanny Brun, Etienne Berthier, Amaury Dehecq, Erik Schytt Mannerfelt, Nicolas Eckert and Daniel Farinotti

Abstract—The monitoring of Earth’s and planetary surface elevations at larger and finer scales is rapidly progressing through the increasing availability and resolution of digital elevation models (DEMs). Surface elevation observations are being used across an expanding range of fields to study topographical attributes and their changes over time, notably in glaciology, hydrology, volcanology, seismology, forestry and geomorphology. However, DEMs frequently contain large-scale instrument noise and varying vertical precision that lead to complex patterns of errors. Here, we present a validated statistical workflow to estimate, model, and propagate uncertainties in DEMs. We review the state-of-the-art of DEM accuracy and precision analyses, and define a conceptual framework to consistently address those. We show how to characterize DEM precision by quantifying the heteroscedasticity of elevation measurements, i.e. varying vertical precision with terrain- or sensor-dependent variables, and the spatial correlation of errors that can occur across multiple spatial scales. With the increasing availability of high-precision observations, our workflow based on independent elevation data acquired on stable terrain can be applied almost anywhere on Earth. We illustrate how to propagate uncertainties for both pixel-scale and spatial elevation derivatives, using terrain slope and glacier volume changes as examples. We find that uncertainties in DEMs are largely underestimated in the literature, and advocate that new metrics of DEM precision are essential to ensure the reliability of future Earth and planetary surface elevation assessments.

Index Terms—Geostatistics, error propagation, remote sensing, variogram, spatial correlation, surface height.

I. INTRODUCTION

DIGITAL elevation models (DEMs) are gridded, numerical representations of surface elevation. DEMs have a long history of interpolation from point measurements and digitized historical maps [1], [2]. Nowadays, DEMs are mostly generated from radar interferometry [3], [4], optical stereophotogrammetry [5], [6] or laser scanning [7], [8] of a planetary surface. When produced from these remote sensing techniques, DEM grid cells essentially represent surface elevations timestamped to the date of instrument acquisition. With the ever-improving coverage and precision of satellite and airborne sensors [9], land surface assessments based on DEMs are advancing towards estimates that are both more spatially and more temporally resolved [10], [11]. Additionally, the recent unlocking of historical optical archives has created unprecedented potential for studying half a century of Earth’s surface elevation [12]–[14].

Studies that harness elevation observations can generally be divided into two groups. The first group relies on single-acquisition and often gap-filled DEMs to extract essential topographic characteristics, e.g., in river discharge and flood modelling [15]–[17], geomorphological terrain analysis [18]–[21], tectonic monitoring [22]–[25], avalanche risk prediction [26], land classification [27], [28], onshore inundation and sea-level rise forecasting [29]–[31] and planetary surface characterization [32], [33]. The second group requires multiple acquisitions to study surface elevation changes over time, e.g., for landslide and rock avalanche detection [34]–[36], seasonal snow depth assessment [37]–[39], lava flow volume quantification [40], [41], canopy height evolution [42]–[44] and glacier, ice sheet and ice shelf mass balance estimation [45]–[47].

In both groups, and for all applications, the interpretation of results and its robustness are intrinsically intertwined with the accuracy and precision of the underlying DEMs.

Accuracy and precision are related to systematic and random errors. In the case of DEMs, they have been the focus of specific research [48]–[51], software development [52] and questioning [53]–[56] since the beginning of the numerical era. Yet, these efforts are dwarfed by the tremendous increase of studies that rely on DEMs [57] and the processing of ever larger data volumes [58]–[60]. Most critically, the analysis of many modern studies is still confined to simplified metrics for accuracy and precision (e.g., [61]–[63]) that mix systematic and random errors and fail to describe the strong spatial variations and correlations in errors observed in DEMs (e.g., [64]–[66]).

Here, we present a statistical workflow to robustly estimate and propagate uncertainties in DEMs; most specifically, we:

- perform a literature review of analyses dealing with DEM accuracy and precision;
- propose a framework based on spatial statistics to consistently address DEM accuracy and precision;
- present robust inferential methods to estimate elevation heteroscedasticity and spatial correlation of errors;
- analyze the impact on the uncertainty of elevation derivatives, using terrain slope and glacier volume changes as examples;
- provide access to our methods through the open, tested and documented Python package xDEM.
II. LITERATURE REVIEW

A. Mitigating poor DEM accuracy before studying precision

The term accuracy has been used to describe either systematic errors or, in some instances, both systematic and random errors, leading to some confusion. In the present article, we define accuracy as the description of systematic errors only, also known as "trueess" [67], which is related to elevation biases. Poor accuracy is common in DEMs and has been a major source of error in elevation assessments, particularly during the advent of space-borne DEMs. Limitations in instrument positioning, orientation, or post-processing often lead to erroneous horizontal referencing [69], vertical shifts [69], [70] or tilts [71], [72] that propagate into elevation biases (Fig. 1a). By utilizing terrain with elevation assumed stable over time, methods performing 3-dimensional alignment of DEMs have flourished, relying on either generic registration methods [73]–[75], least squares approaches [71], [76] or specifically-developed DEM registration based on terrain constraints [77], [78]. These methods proved robust for aligning a DEM either to an external reference DEM, or to accurate geolocated point elevation data such as space-borne laser altimetry [79], [80]. The above registration methods are only successful at correcting elevation biases common to the entire DEM grid, however. Other biases remain present once 3-dimensional alignment is attained and can arise from resolution [81], [82], specific image deformations and instrument biases [13], [66] or physical properties of the observed terrain such as radar penetration into snow and ice [83], [84] or into forest canopy [43]. Most of these biases are instrument- or application-dependent and, therefore, require specific considerations. Notwithstanding those, poor DEM accuracy has been largely addressed by the robustness of registration methods that have become increasingly widespread, thereby shifting the focus towards the next limiting factor: better quantifying DEM precision.

B. The inherent variability of vertical precision

Precision describes random errors [67] and is related to elevation variance. One aspect of DEM precision consists of the pixel-scale dispersion of elevations that we refer to as "vertical precision". DEMs are generated from acquisitions that possess intrinsic, random measurement errors. At the pixel scale, instrument resolution, spectral range, and encoding depth of optical sensors directly affect the quality of stereocorrelation [5], [6], [85], radar slant angle and height of ambiguity play an important role in interferometric coherence [86], [87] while laser wavelength, sunlight background radiation, target reflectivity, and backscattering properties modulate laser signal-to-noise ratio [88], [89]. Many instrument- or processing-related metrics constitute quality indicators of the estimated elevations. These indicators have been almost exclusively used for the filtering of observations of lesser quality, however, and only occasionally as a tool towards improved modelling of sensor-specific variability in vertical precision (e.g., [90]). Besides, the geometry of instrument acquisition can exacerbate random errors depending on the relief of the observed landforms (Fig. 1b). Vertical precision has indeed been long shown to decrease with terrain slope [48], [91]–[94]. Several assessments account for this variability by partitioning the elevation variance into categories of flat and steep terrain (e.g., [59]). Most studies use a single metric to describe vertical precision, however, often reporting a standard deviation (e.g., ±2 meters). Such simple metrics are insufficient in describing the heteroscedasticity of elevation measurements, i.e. the variability in vertical precision. Although some studies quantified and modelled this heteroscedasticity [64], [95], [96], this modelling was generally performed without validation of the underlying methodology and, most critically, without considering the effect of spatial correlations.

C. The correlated noises that plague DEMs

Another aspect of DEM precision concerns the inter-pixel spatial dependency of random errors, here referred to as "spatial correlations". Spatial correlations describe structures of noise that show a location-dependent pattern, which can often be traced back to limitations during acquisition or post-processing. Along-track undulations have been observed in many DEMs generated from air- and space-borne sensors (Fig. 1b), including the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [47], [66], the Satellite Pour l’Observation de la Terre (SPOT) [97], [98], Pléiades [39], [99] and the Shuttle Radar Topography Mission (SRTM) [88], [100], [101] (Fig. 1a). Processing noise is common in DEMs requiring image digitization including aerial photographs [102], [103] or historical satellite imagery such as Corona and Hexagon KeyHole-9 (KH-9) [13], [104] (Fig. 1c). To mitigate these correlated noises, DEM correction methods have emerged [82], [105] but are still burgeoning for specific
types of errors [69], [106], [107], and their performance is highly dependent on the type of terrain. Furthermore, nearly all DEMs contain structural short-range correlation of errors. The degree to which a DEM grid spacing represents its native resolution [108], [109] and how that resolution has possibly been degraded through interpolation [110] determine the severity of these short-range correlations. When upsamped to a larger grid spacing, vertical precision improves directly as a function of the underlying spatial correlations [111]. Spatial correlations are generally quantified using an empirical variogram [112], [113] estimated either on the basis of differences with independent elevation observations [2], [114] or those with simulated elevation surfaces [115], [116]. Many studies have used variograms, but have almost exclusively used short range models (i.e. 5 to 20 times the pixel size). Few studies modelled longer-range correlations, that is, correlations that persist over distances several orders of magnitude larger than the pixel size [13], [47], [65]. The widespread occurrence of long-range noise in DEMs thus constitutes a critical limitation in the analysis of DEM precision, and one that directly affects uncertainty propagation.

D. Uncertainty propagation to elevation derivatives

To propagate elevation variance into uncertainties of elevation derivatives (i.e. variables that are derived from elevations), a large set of methods has been applied that generally relies on spatial statistics. Spatial statistics, also known as geostatistics [112], [113], [117], provide a large body of theories and methods that, among others, can address spatial uncertainty analyses [118]–[120] by characterizing spatial correlations that depend only on the distance between observations. These uncertainty propagation methods can be subdivided into two groups: (i) Monte Carlo techniques that simulate multiple random realizations of correlated error fields [121]–[123], notably including Sequential Gaussian simulation [117] and Fourier randomization [124]; and (ii) gradient techniques that analytically approximate the variance of a derivative through simplified equations, that can be either based on Taylor series expansion [121] for any derivative of elevation, or approximations of variogram integration [65] for spatial derivatives. The first group has been widely used for topographic variables, notably in hydrology [16], [57], [125], [126] and occasionally for spatial derivatives in glaciology [127]. The second group is used less frequently, both for Taylor series expansions developed in few applications [128], [131], and for variogram integration implemented mainly in glaciology and geomorphology [65], [132]. Although both groups are expected to perform similarly, Monte Carlo techniques are computationally expensive, especially at fine resolution. Analytical approximations, instead, require a theoretical description of variance propagation that can reach a high degree of complexity for some derivatives [133]. To our knowledge, few studies [132] constrained these propagation methods with estimates of heteroscedasticity and spatial correlation of errors into a single framework for DEMs, and none tested the underlying assumptions of spatial statistics. In the following, we propose such a framework, and later describe methods to robustly estimate its key components.

III. Problem formulation

A. Elevation bias and variance at each location

We consider the elevation observation \( \hat{h}(x, y, t) \) located at \((x, y)\) in space and \(t\) in time, and pertaining to the DEM \( D \). Annotating the true unknown elevation at the same location \( h(x, y, t) \), we can state that the elevation observation has a bias \( \delta h(x, y, t) \) if, over a large number of repeated measurements \( i \in I \) of elevation \( \hat{h}(x, y, t)_i \) at \((x, y, t)\), we have:

\[
\hat{h}(x, y, t)_i - h(x, y, t) = \delta h(x, y, t).
\]

The repeat elevation measurements around the bias \( \delta h(x, y, t) \) are subject to random measurement errors \( \epsilon(x, y, t) \) with variance \( \sigma^2_h(x, y, t) \), whose distribution is not necessarily normal and might depend on time and location:

\[
\hat{h}(x, y, t)_i = h(x, y, t) + \delta h(x, y, t) + \epsilon_i(x, y, t).
\]

In practice, acquiring a large number of repeat measurements at both the same location and time is not feasible, and we therefore turn towards inferential methods to estimate these biases and variance.

B. Inference from stable terrain

DEM benefit from a great asset, largely uncommon to other remote sensing data, which is that large proportions of planetary surface elevations remain virtually unchanged through time. In fact, elevation changes caused by erosion, short vegetation growth, or continental drift are typically small compared to the precision of the measurement. Terrains such as bare rock or grasslands – later referred to as "stable terrain" – thus provide the means of analyzing multiple elevation measurements acquired at different points in time as if they were acquired from simultaneous measurements \( \hat{h}(x, y, t)_i \):

\[
\frac{dh(x, y, t)}{dt} \approx 0 \text{ for } (x, y) \in \text{stable terrain}.
\]

While this temporal consistency unlocks the potential to analyze elevation acquisitions independently of time \( t \), it is impeded by the number of required DEMs. For each location \((x, y)\), the number of samples to perform the statistical analysis would always be at best equal to the total number of independent acquisitions, requiring a large number of DEMs. Therefore, we investigate the spatial properties of elevation biases and variance.

C. Spatial homogeneity after affine alignment

Elevation biases and variance are inherent to instrumental limitations, to the physical properties of the observed terrain, as well as its topography (see previous Sections II-A and II-B). Among many types of location-specific biases, a general exception is that of grid misalignment to the true elevations \( h(x, y, t) \) that follows specific geometric distributions linked to the gridded nature of DEMs (Fig. 1a). In our framework, we therefore split elevation biases into two categories: affine biases \( \delta h_A \) that are common to the entire DEM (e.g., translation, rotation, scaling), and non-affine "specific" biases \( \delta h_S \)
that occur at the grid cell level and vary with instrumental and topographical effects (Fig. 2):

\[ \delta h(x, y, t) = \delta h_A(x, y, t) + \delta h_S(x, y, t). \quad (4) \]

Once an alignment is attained by the affine transformation \( A \) giving \( A(x, y, t) = \delta h_A(x, y, t) \), we assume that, for a single DEM \( D \), specific elevation biases \( \delta h_S \) and elevation variance \( \sigma_h^2 \) have a spatial distribution that is homogeneous with the properties of the instrument and the observed terrain \( P \). We use this spatial homogeneity to substitute space for time. For example, we consider that elevations \( h(x_1, y_1, t) \) and \( h(x_2, y_2, t) \) of \( D \) acquired on the same surface type (e.g., bare rock), and under the same topographical attributes (e.g., flat) will have similar specific biases and variance:

\[
\delta h_s(x_1, y_1) \approx \delta h_s(x_2, y_2) \quad \text{for } P(x_1, y_1) = P(x_2, y_2). \quad (5)
\]

Combining the assumptions of Eqs. (3) and (5) and provided that we describe all the properties \( P \) of spatial homogeneity, a large sample size can be used to infer \( \delta h_s \) and \( \sigma_h \) at each location \((x, y)\) from a single difference between a DEM and an independent source of elevation data. The properties of spatial homogeneity \( P \) could differ between biases and variance. In the following, we assume that specific elevation biases, if they exist, are independently corrected and focus on characterizing the elevation variance \( \sigma_h^2 \).

### D. Elevation difference with an independent source

After performing affine alignment of elevations \( \hat{h}_1(x, y, t_1) \) from a first source \( D_1 \) and elevation \( \hat{h}_2(x, y, t_2) \) of a second source \( D_2 \), we subtract them to derive elevation differences \( \hat{d}h_{1-2}(x, y) \). Assuming independence between the error of each elevation source, the variance of the difference is:

\[
\sigma^2_{\hat{d}h_{1-2}}(x, y) = \sigma^2_{\hat{h}_1}(x, y) + \sigma^2_{\hat{h}_2}(x, y). \quad (6)
\]

By selecting a second source to observe \( \hat{h}_2 \) that is of higher precision than the first source that observes \( \hat{h}_1 \), the analysis of the differences \( \hat{h}_2 - \hat{h}_1 \) will largely capture the variance of the first source. For example, if the second source is three times more precise than the first, Eq. (6) implies that about 95% of the variance of the elevation difference will originate from the first source, yielding:

\[
\sigma_{\hat{h}_1}(x, y) \approx \sigma_{\hat{d}h_{1-2}}(x, y). \quad (7)
\]

Alternatively, if \( h_1 \) and \( h_2 \) originate from independent acquisitions of the same instrument and processing, we have:

\[
\sigma_{\hat{h}_1}(x, y) = \frac{\sigma_{\hat{d}h_{1-2}}(x, y)}{\sqrt{2}}. \quad (8)
\]

Thus, we use elevation differences to infer on \( \sigma_h \), which can be converted from either Eqs. (7) or (8).
E. Discriminating elevation bias from variance in spatial statistics

To further analyze elevation variance, we need to discriminate bias from variance. When analyzing elevation differences, what appears as a bias at the local scale could also be a bias from variance. When analyzing elevation differences, second-order stationarity implies that the following assumptions should be fulfilled (see Supplementary Section II-A):

1) a first assumption of stationary mean, i.e. that the average of elevation differences \( dh(x, y) \) is constant over large areas;
2) a second assumption of stationary variance, i.e. that the variance of elevation differences \( \sigma_{dh}(x, y) \) is constant over large areas; and
3) a third assumption of spatially consistent covariance, i.e. that the correlation between random errors of elevation differences only depends on the distance between observations.

Large areas here refer to areas slightly smaller than the size of the study domain, typically within an order of magnitude. As such, a correlated error with a correlation range that is orders of magnitude larger than the size of the study domain might be considered a vertical bias common to the entire DEM grid (Fig. 2). And, inversely, such a bias placed in the context of a larger study domain might be considered as a correlated error, if the elevation differences fulfill the above assumptions.

Thanks to the affine alignment of our elevation differences, we verify the first assumption of stationary mean. However, the heteroscedasticity of elevations (see Section II-B) invalidates the second and third assumptions, and therefore a nonstationary framework needs to be defined.

F. A non-stationary spatial framework for DEM analysis

To perform spatial statistics with a non-stationary variance, transformation of the data towards a stationary variance is necessary. The transformation depends on the nature of the spatial variability and correlations. In DEMs, we identify two types of correlation: short-range ones related to resolution, and long-range ones related to correlated noise or digitization artefacts. While the latter appear unrelated to the heteroscedasticity of elevation, the former are similarly linked to local instrument- and terrain-dependent variables (see Sections II-B and II-C). We thus subdivide elevation variance into elevation heteroscedasticity and spatial correlation of errors (Fig. 2) assuming that longer-range correlations are independent of elevation heteroscedasticity, which yields:

\[
\sigma^2_{dh}(x, y) = \sigma^2_{sr}(x, y) + \sigma^2_{lr},
\]

where \( \sigma^2_{sr}(x, y) \) is the variable short-range variance at \( (x, y) \), and \( \sigma^2_{lr} \) is the constant long-range variance.

Using the variable spread \( \sigma_{dh}(x, y) \), the elevation differences can be standardized into a standard score \( z_{dh}(x, y) \) with unit variance, which fulfills the second assumption of second-order stationarity:

\[
z_{dh}(x, y) = \frac{dh(x, y)}{\sigma_{dh}(x, y)}.
\]

Additionally, the spatial covariance \( C_{zd}(d) \) of \( z_{dh} \), related to the variogram \( \gamma_{zd}(d) = 1 - C_{zd}(d) \), is also free of the influence of heteroscedasticity and now fulfills the third assumption of second-order stationarity:

\[
\gamma^2_{zd}(d) = \left( \frac{\sigma_{dh,sr}}{\sigma_{dh}} \right)^2 \gamma^2_{sr}(d) + \left( \frac{\sigma_{dh,lr}}{\sigma_{dh}} \right)^2 \gamma^2_{lr}(d),
\]

where \( d \) is the spatial lag, i.e. the distance between two given observations, \( \sigma_{dh,sr} \) is the average of \( \sigma_{dh,sr} \) in the DEM \( D \), and \( \gamma_{sr} \) and \( \gamma_{lr} \) are the short- and long-range variogram functions.

With all the assumptions in our framework fulfilled, we can now reliably use spatial statistics for uncertainty propagation. To this end, we require an estimate of the elevation dispersion \( \sigma_{dh}(x, y) \) and of the variogram of the standard score \( \gamma_{zd}(d) \), which describe the heteroscedasticity and the spatial correlation of errors, respectively. We also need to ensure that our assumption of spatial homogeneity remains valid when using stable terrain as an error proxy to infer heteroscedasticity and spatial correlations on moving terrain. In the following, we address these aspects by utilizing near-simultaneous data and implementing robust methods.

IV. DATA

A. Mont-Blanc case study: simultaneous DEMs

To demonstrate the methods associated with our proposed framework, we present a case study of two DEMs generated one day apart in the Mont-Blanc massif, French Alps (Fig. 3b, Table I). These DEMs were produced with a spatial posting of 5 m from SPOT-6 and Pléiades stereo images using the Ames Stereo Pipeline [134]. We utilize the temporal closeness of the two acquisitions to assess if stable terrain can be used as a proxy for moving terrain, considering a negligible elevation change on moving terrain.

We present an additional case study in the Northern Patagonian Icefield to illustrate the influence of the quality of stereo-correlation, a sensor-dependent variable, on elevation heteroscedasticity (Supplementary Section I-A with additional refs. [66], [135], [136]). This case study is based on simultaneously acquired ASTER [47] and SPOT-5 images. Furthermore, the DEMs used to illustrate noise patterns (Figs. 1 and S1) are described in the Supplementary Section I-B with additional refs. [137]–[139].

| Table I: Nearly-simultaneous Pléiades and SPOT-6 DEMs used for the Mont-Blanc case study. |
|-----------------------------------------------|
| Instrument | Acquisition time | Resolution of stereo-pair |
| Pléiades   | 24/10/2017, 12:00 CET | 1.5 m  |
| SPOT-6     | 25/10/2017, 12:30 CET | 0.7 m  |

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
B. Inventory and land cover products

We define moving terrain as glacierized, forested and seasonally snow-covered terrain, and exclude water bodies from our analysis. The remaining terrain is assumed to be stable. We mask glaciers using the Randolph Glacier Inventory 6.0 (RGI 6.0) outlines \[140\], which are delineated from images with a typical resolution of 15–30 m. We mask forests and water bodies using the ESA Climate Change Initiative Land Cover version 2.0.7 \[141\] which has a resolution of 300 m. Forested terrain corresponds to either broad-leaved, needle-leaved, evergreen, or deciduous tree cover classes. We mask glaciers using the Randolph Glacier Inventory 6.0 outlines \[140\], which are delineated from images with a typical resolution of 15–30 m. We mask forests and water bodies using the ESA Climate Change Initiative Land Cover version 2.0.7 \[141\] which has a resolution of 300 m. Forested terrain corresponds to either broad-leaved, needle-leaved, evergreen, or deciduous tree cover classes.

We identify specific elevation biases over forested terrain between the SPOT-6 and Pleiades DEMs – likely owing to different native resolution, orientation and spectral bands (Fig. S3) – and thus exclude this terrain from our analysis. Our end-of-summer acquisitions contain little snow outside of glacierized surfaces. Therefore, we did not mask off-ice snow cover. Ultimately, in our analysis, moving terrain corresponds to glacierized terrain.

![Glacierized, Forested, Example glaciers](image)

**Fig. 3.** Mont-Blanc case study. Hillshade of Pleiades DEM and land cover for the Mont-Blanc case study. Example glaciers serve to illustrate Section VI-C.

V. Methods

A. Robust statistics and alignment

We use the median instead of the mean as a robust estimator of central tendency, and the normalized median absolute deviation (NMAD) instead of the standard deviation as a measure of statistical dispersion. Both choices are to mitigate the effects of frequent outliers in DEMs \[142\]. Combining these estimators with the dense sampling of stable terrain also ensures robustness to elevation changes of potentially unmasked moving terrain. This includes rare events such as landslides or ground subsidence, or events that can occur over a small portion of the analyzed terrain such as volcanic uplift or sediment transport. We coregister DEMs on stable terrain for horizontal and vertical shifts following the aspect-slope relation described in \[77\] and we correct for possible tilts through least squares optimization of a plane \[71\].

B. Heteroscedasticity

We estimate elevation heteroscedasticity by sampling an empirical dispersion of elevation differences \(\hat{\sigma}_{dh}\) using the NMAD of binned categories along the terrain slope \(\alpha\) \[143\] and the terrain maximum absolute curvature \(c\) (Figs. 4a-c and S4). Maximum absolute curvature is defined as the maximum of the absolute profile curvature and the absolute planform curvature at each location \[144\]. All terrain attributes are estimated from the Pleiades DEM that contains the least data gaps. When available, the binning can also include an instrument quality factor \(q\), such as the quality of stereo-correlation (Supplementary Section I-A, Figs. S5-S7) or interferometric coherence.

We numerically model the empirical dispersion \(\hat{\sigma}_{dh}\) as a function \(\sigma_{dh}\) of the terrain- and sensor-dependent variables \((\alpha, c, q)\) by multidimensional linear interpolation of the binned data (Fig. S8). The modelling of this variability can also be performed by fitting parametric models, for example an exponential model with the slope or a linear model with the maximum curvature (Fig. S9). These are more robust in the case of small sample sizes of elevation differences.

We standardize the elevation differences \(dh\) following Eq. 10 using the modelled dispersion \(\sigma_{dh}(\alpha, c, q)\):

\[
z_{dh} = \frac{dh}{\sigma_{dh}(\alpha, c, q)}. \tag{12}\]

After standardization, we verify that the standard score of the elevation differences matches a normal distribution by quantile-quantile plotting, and by comparison to a normal distribution fit \[142, 145\] (Fig. S10). The substantial improvement validates our choice of terrain slope and maximum curvature as key variables to describe elevation heteroscedasticity, as those largely explain the departure of random elevation errors from normality.

C. Spatial correlations

We estimate spatial correlations by sampling an empirical variogram \(\hat{\gamma}\) on the standard score \(z_{dh}\) using Dowd’s estimator \[146, 147\] (Fig. 5):

\[
2\hat{\gamma}_{z_{dh}}(d) = 2.198 \cdot \text{median}(z_{dh}(x, y) - z_{dh}(x', y'))^2, \tag{13}\]

where \(z_{dh}\) is the standard score of elevation differences, and locations \((x, y)\) and \((x', y')\) are separated by a spatial lag \(d\). Dowd’s estimator is based on median absolute deviations, and consequently more robust than the Matheron \[148\] or Cressie-Hawkins \[149\] estimators classically used (see Supplementary Section II-B based on additional ref. \[150\]). We
verify the increased robustness of Dowd’s estimator for the Mont-Blanc case study (Figs. S11 and S12).

To improve the variogram estimation, we introduce a pairwise subsampling method based on iterative subsetting of pairwise combinations between a disk and multiple rings centered on a random point (see Supplementary Section II-C). As variograms were historically sampled from point measurements [112], traditional sampling methods are less computationally efficient on large grids. Most critically, they are inefficient at sampling pairwise distances evenly across spatial scales, which is substantially improved by our method to estimate more reliably both short-range and long-range correlations (Fig. S13). Finally, we derive empirical variograms for 100 independent realizations with the same binning. We estimate our final empirical variogram by the mean of all realizations at sampling pairwise distances evenly across spatial scales, for which we can study the distribution of errors.

For the Mont-Blanc study, we find no significant improvement in least-squares residuals when fitting more than three models, which are capable of capturing one short-range and two long-range correlations (Fig. S14 Table S2). The two long-range correlations match the along- and cross-track lengths of low-amplitude undulations in the elevation differences (Fig. 5b). We thus use three models to avoid the possible overfitting of a larger number of summed models. Generally, \( k \) should be chosen to reflect the number of distinct ranges in the patterns of DEM noise. For instance, ASTER undulations are characterized by two wavelengths of 1–2 km and 5–10 km in the along-track direction, and a cross-track distance of 60 km (Fig. S11b), which better fits three distinct long-range models [47] for a total of four ranges.

We also identify a low sensitivity to different variogram model types (Fig. S15 Tables S3 and S4), which shows that adequately modelling the multi-range nature of the spatial correlations is more important than refining that of their spatial form (Fig. 5b). For the Mont-Blanc case study, we reached the smallest least-squares residuals using a gaussian model \( G(s, r, x \rightarrow x') \) at short ranges, and spherical models \( S(s, r, x \rightarrow x') \) at long ranges [151] and used those henceforth:

\[
G(s, r, d) = s \left( 1 - e^{-\left( \frac{d}{r} \right)^2} \right)^2, \quad \text{and} \quad S(s, r, d) = \begin{cases} s \left( \frac{d^3}{5} - \frac{1}{3} \left( \frac{d}{r} \right)^3 \right) & \text{if } 0 < d < r, \\ s & \text{if } d \geq r. \end{cases}
\]

**D. Uncertainty propagation**

1) Simulation methods for elevation derivatives: For derivatives of elevation with a complex spatial gradient, such as terrain slope and aspect later analyzed, we use simulation methods. We find similar results using Fourier randomization [124, 152] and unconditional Gaussian simulation [117, 153], and thus only use the former in the following. For 1,000 realizations, we simulate a random correlated error field of the standard score \( z_{dh} \) based on the modelled spatial correlation \( \gamma_{z_{dh}} \) in Eq. 14. We then de-standardize \( z_{dh} \) using Eq. 12 and add the resulting elevation error field to the studied DEM. For each of these DEM realizations with an added error field, we then compute the terrain attribute of interest (e.g., terrain slope or aspect), for which we can study the distribution of errors.
2) Theoretical approximation methods for spatial derivatives: For spatial derivatives such as the average $\bar{dh}$ of elevation changes $dh$ in an area $A$, we derive an exact analytical solution of the uncertainty in the spatial average $\sigma_{\bar{dh}}$:

$$\sigma_{\bar{dh}}^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} \sigma_{dh_i} \sigma_{dh_j}, \quad (17)$$

where $N$ denotes the number of samples $i$ falling in the area $A$, $\sigma_{dh_i}$ is the vertical precision of pixel $i$, and $\rho_{ij} = (1 - \gamma_{\bar{dh}}(d))$ is the spatial correlation between pixel $i$ and pixel $j$ based on their distance $d$.

In practice, Eq. (17) raises the issue of scaling exponentially with the number of samples, possibly resulting in trillions of calculations. To remedy this, we propose an approximation for spatially contiguous areas, inspired by the approach of [65] that computes a single aerial integral by approximating the area $A$ by a disk of the same area. Here, for each pixel $k$ of a random subset of $K$ pixels within the $N$ pixels, we compute the single aerial integral of the variogram numerically. We then approximate the variogram integral by the average of these subset aerial integrals (see Supplementary Section II-D):

$$\frac{\sigma_{\bar{dh}}^2}{A} \approx \frac{\sigma_{\bar{dh}}^2}{A} \frac{1}{K} \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N} (1 - \gamma_{\bar{dh}}(x_k - x_i)), \quad (18)$$

where $\frac{\sigma_{\bar{dh}}^2}{A}$ is the average variance of the elevation differences of pixels $i$ in the area $A$:

$$\sigma_{\bar{dh}i}^2 = \frac{1}{N} \sum_{j=1}^{N} (1 - \gamma_{\bar{dh}}(x_i - x_j)), \quad (19)$$

We show that our method improves the accuracy of the theoretical approximation of [65] by accounting for more complex area shapes than disks while maintaining computational efficiency (Fig. S16). Additionally, these formulations can be linked to a number of effective samples, which describes the number of samples among the $N$ pixels in area $A$ that are statistically independent based on the spatial correlations modelled by $\gamma_{\bar{dh}}$ (see Supplementary Section II-D).

Once uncertainties have been integrated from a spatial support (e.g., pixels) to a larger spatially contiguous ensemble (e.g., glaciers), they can be propagated again to a larger ensemble (e.g., all glaciers in a region) following Krige’s relation of transitivity [112], [154]. For this, Eq. (17) can be applied for each pair of spatially contiguous ensembles $i$ and $j$ of area $A$, with the same variogram $\gamma_{\bar{dh}}$ composed of the $k$ summed models $V_k(s_k, r_k, d)$:

$$\sigma_{\bar{dh}i}^2 = \frac{1}{(\sum_i A_i)^2} \sum_k \sum_i \sum_j \left( \sigma_{\bar{dh}i,k} \sigma_{\bar{dh}j,k} \right) A_i A_j,$$

where $d_{i-j}$ is the distance between the centroids of ensemble $i$ and $j$, and $\sigma_{\bar{dh}i,k}$ is the spatially integrated uncertainty.
of ensemble \( i \) associated to the variogram model \( \hat{V}_k \), partial sill \( s_k \) and range \( r_k \) with pixel pairs \( n \) and \( m \) (see Eq. 17):

\[
\sigma_{dh_{k,i}} = \frac{1}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} (s_k - \hat{V}_k(s_k, r_k, x_n - x_m)) \sigma_{dh_n} \sigma_{dh_m}.
\]

Furthermore, we use a Monte Carlo spatial sampling method to validate our uncertainties of spatially averaged elevations, thus indirectly verifying the robustness of our modelled spatial correlations (Fig. 5c). We randomly sample up to 10,000 circular patches of area \( A \) without replacement. We compute the mean \( \bar{dh} \) inside circular patches, keeping only those with more than 80% valid elevation differences \( dh \) to mitigate the effects of missing data. We use the NMAD of 10,000 realizations to empirically estimate the uncertainty of the spatially averaged \( dh \) of area \( A \), and repeat this procedure for varying area sizes \( A \) (Fig. 5c). This method substitutes repeated correlated simulation of Fourier randomization or Gaussian simulation by a repeated spatial sampling, relying on the assumption of spatial homogeneity of variance on stable terrain (Section III-C). As it requires a large number of independent patches to produce a robust estimate, the area size \( A \) for which it can estimate an uncertainty is limited to sizes much smaller than that of the spatial domain. It is also highly dependent on the availability of stable terrain. Therefore, we use it only for validation purposes.

VI. RESULTS AND DISCUSSION

In Section VI-A below, we discuss the use of stable terrain as an error proxy based on the methods applied to the Mont-Blanc case study. In Sections VI-B and VI-C we then analyze the impacts of heteroscedasticity and spatial correlations when propagating elevation variance into uncertainties of pixel-scale elevation derivatives such as terrain slope, or spatial derivatives such as glacier volume changes. In those two sections, we provide examples based on the Mont-Blanc case study and determine the impact of our methods for a set of assumptions on the variance properties during uncertainty propagation:

- either homoscedastic elevation (constant variance, shortened "homosc.") or heteroscedastic elevation (variable variance, "heterosc."); and
- either no spatial correlation (shortened "no corr.") or only short-range correlations ("short-range"), or both short- and long-range correlations ("long-range").

In this exercise, the most realistic case refers to the one that accounts for potential elevation heteroscedasticity and potential short- and long-range correlations. Uncertainties are reported as a symmetric confidence interval of 1\( \sigma \) (68% confidence level) or 2\( \sigma \) (95%), specified in each case.

A. Validation of stable terrain as an error proxy

We test the validity of using stable terrain as a proxy of elevation errors for moving terrain on the nearly simultaneous DEMs of the Mont-Blanc case study. We find that elevations on moving terrain exhibit the same heteroscedasticity with slope and curvature than those on stable terrain, with less than 1% of binned samples that differ by more than 30% (Fig. 4-c). We additionally verify that this elevation heteroscedasticity is continuous between neighbouring bins when using robust estimators, thereby consolidating our assumption of spatial homogeneity (Section III-C Eq. 5). By extending this assumption to the case of moving terrain, we infer a complete map of vertical precision (Fig. 5).

We find similar spatial correlations of errors between stable and moving terrain (Fig. 5). Values of partial sills and ranges of the variogram models that describe these correlations are within the same orders of magnitude (Table II), despite greater differences at long ranges due to the limited pairwise samples available on moving terrain. Using our Monte Carlo sampling method, we validate the increased robustness of using multiple correlation ranges to estimate uncertainties across spatial scales (Fig. 5). Our results indicate that using a short-range model alone underestimates elevation uncertainties by several orders of magnitude for areas larger than 0.1 km².

\begin{table}[h]
\centering
\caption{Estimated variogram model parameters for the spatial correlations of elevation errors in the Mont-Blanc case study. Stable and moving terrain is distinguished (columns). Gaussian components are listed for the short-range model and spherical components for the long-range ones, as in Fig. 5. Partial sills are expressed as a percentage of the total variance.}
\begin{tabular}{|l|c|c|}
\hline
Model parameters & Stable & Moving \\
\hline
Sill of short-range model & 93\% & 95\% \\
Range of short-range model & 30 m & 38 m \\
Sill of 1st long-range model & 2\% & 4\% \\
Range of 1st long-range model & 3,900 m & 2,400 m \\
Sill of 2nd long-range model & 5\% & 1\% \\
Range of 2nd long-range model & 11,200 m & 10,800 m \\
\hline
\end{tabular}
\end{table}

For elevation heteroscedasticity, our results highlight the importance of elevation standardization to ensure an adequate scaling when inferring on another type of terrain (e.g., from steep, stable terrain to flat, moving terrain). Yet, our analysis only exemplifies snow- and ice-covered terrain with high-resolution stereophotogrammetric DEMs. The physical properties of the observed terrain in relation to the utilized sensor might in some cases invalidate our assumption of spatial homogeneity. For instance, we found that our standardization did not mitigate the larger errors of elevation over forested areas (Fig. 5). In such a case, an upfront investigation of specific elevation biases is required. After these biases are corrected, a refined modelling of elevation heteroscedasticity based on sensor-dependent variables can help to reach a good description of the properties of spatial homogeneity. We indeed found a strong relationship with the quality of stereo-correlation for the case study of the Northern Patagonian Icefield (Figs. 56 and 57). The rougher resolution (15 m) and spectral range (8 bits) of the ASTER stereo images, compared to those of SPOT-6 and Pléiades (metric resolution and 12-bits), leads to a significant variability in elevation errors with terrain texture.
For spatial correlations, we highlight the value of standardization to reduce variability for empirical variogram estimation (Figs. S3b-d and S12). It is especially useful to deconvolve the long-range correlations with small magnitude to the short-range ones. Heteroscedasticity may indeed explain the short-range variogram anisotropy found by previous studies [155]. We nevertheless identify a slight difference in the well-constrained short correlation range between stable and moving terrain (30 m vs 38 m, respectively; Table II). This difference might be due to the rougher interpolation of stereophotogrammetric block-matching algorithms over bright, lower-texture glacierized terrain. In some cases, sensor properties or processing schemes influence not only the magnitude of spatial variability but also the scale of correlations. Developing a statistical framework that continuously includes these effects might be overly complex for most analyses that, instead, could adjust estimates of short-range correlation depending on the type of observed terrain.

We conclude that stable terrain is a valid proxy for error analysis, provided that elevation heteroscedasticity is taken into account. However, the quality of statistical inference from this error proxy depends directly on the number of stable terrain samples available. For some DEMs, these samples might be scarce in the proximity of continuous expanses of moving terrain (e.g., at the margins of ice sheets or large forests) and thus insufficient to perform robust inference. To address this, the stable terrain of independent DEMs, possibly located elsewhere, could be utilized if they are generated from the same instrument and processing chain. Many DEMs indeed have consistent error properties between segments acquired under similar conditions around the world (e.g., [47], [59], [156]). For instruments with correlated noise of varying amplitude, such as Pléiades or ASTER, long-range correlations can be more robustly inferred from a multiple-acquisition average of variograms.

B. Impact on pixel-scale derivatives of elevation: example with terrain slope and aspect

We illustrate the propagation of elevation uncertainty to the slope and aspect in a 4 km² area around the Mont-Blanc summit (Fig. 6a). We select this area due to its wide range of slopes and aspects, and its small extent facilitating computationally expensive simulations. To avoid the circularity of the aspect variable when assessing uncertainty, we divide it into northness (i.e. cosine of the aspect) and eastness (i.e. sine of the aspect) which denote, respectively, the north-south and east-west tilt of the slope.

We propagate uncertainties in the Pléiades DEM by simulating random elevation error fields (see Section V-D) for every set of assumptions (Fig. S17). For this example, we assume that SPOT-6 and Pléiades have random errors of similar amplitude, and estimate the random errors of the Pléiades DEM following Eq. 8. We generally note a strong deviation from normality and asymmetry in the simulated uncertainty distribution of terrain attributes (Fig. S18). While this asymmetry requires specific considerations for in-depth terrain analysis, we here provide a simplified picture by estimating a symmetric 1-σ uncertainty derived from the half-difference between the 16th and 84th percentile of the simulated slope, northness or eastness of each pixel.

Our analysis reveals that elevation heteroscedasticity plays a major role in the spatial distribution of uncertainties in slope and aspect. In particular, it exacerbates errors in steep and rough terrain. Spatial correlations moderately affect uncertainties by slightly reducing their amplitude (Fig. 6b-c). We interpret the latter to be due to an increase in the spatial coherence of terrain derivatives when the elevation errors are spatially correlated. Since topographical attributes are derived over a 3x3 pixel window, the closer the short-range spatial correlations are to a 3-pixel length, the larger the impact on the amplitude change (Fig. S19).

By aggregating uncertainties into slope categories, we show that uncertainties in flat terrain are overestimated when assuming homoscedasticity and no spatial correlation, while those in steep terrain are underestimated by up to a factor...
of 10 (Fig. 6c). Slope uncertainties decrease near slopes of 90 degrees, likely because elevation errors tilt the terrain in different orientations while generally maintaining a steep slope, which translates into aspect uncertainties. We reach similar conclusions when aggregating uncertainties by maximum absolute curvature categories, our second variable that describes elevation heteroscedasticity (Fig. S20).

C. Impact on spatial derivatives of elevation: example with glacier volume changes

Fig. 7. Uncertainty propagation to glacier mean elevation changes at glacier volume changes

We consider 84 glaciers in the Mont-Blanc massif that have at least 85% of their area covered by valid elevation differences. We analyze the mean elevation changes within the outline of each glacier, which can be converted to volume changes after multiplication by the glacier area, and propagate uncertainties for each set of assumptions.

We find that spatial correlations strongly hamper the decrease in uncertainty with increasing glacier area (Fig. 7c). Long-range correlations are the main contributor to uncertainty for large areas, mirroring the validation of Fig. 3. While our case study has long-range correlations of only 7% of the variance, uncertainties of mean elevation changes for glaciers larger than 10 km² are underestimated by a factor of about 25 when based solely on short-range correlation. This is striking, and even more so when realizing that the underestimation is nearly by a factor of 150 when totally omitting spatial correlations. This dramatic increase is explained by the fact that long-range correlations essentially correspond to local biases.

Heteroscedasticity has a moderate influence on the uncertainty of each glacier, impacting its amplitude by a factor of 1 to 3. The uncertainty of glaciers located in flat areas is overestimated when using a homoscedastic assumption due to the larger average variance over rougher, stable terrain. On the contrary, the uncertainty of the steepest glaciers is underestimated (Fig. 7b). Using the empirical comparison provided by the nearly simultaneous volume changes (Fig. 7b), we show that the uncertainties for the mean elevation change are most realistic when accounting for long-range spatial correlation. In such a case, 89% of the ranges intersect zero (the true volume change) at the 2σ level (i.e. 95% confidence), in contrast to only 30% for short ranges and 7% for no correlation. Yet, our uncertainties are slightly too low.

We identify the cause of this underestimation as the omission of a longer-range correlation close to the size of the DEM and thereby difficult to constrain. This longer-range correlation arises from the fact that along-track undulations are fully correlated in the cross-track direction with 20 km swath. Directional variography could help characterize such correlations, but would lead to a more difficult uncertainty propagation, with exacerbated complexity when combining several DEMs. Instead, we maintain an omnidirectional variogram to describe correlations, but assess a conservative estimate based on artificial undulations (Fig. 21). This results in the replacement of the 11.2 km correlation with a 20 km one (DEM swath width) and a partial sill twice larger. We then find that 93% of the uncertainties for glacier larger than 0.2 km² intersect zero at the 95% confidence level, confirming the increased robustness with these considerations. Only 87% do so for smaller glaciers, however. This discrepancy might be explained by unaccounted heteroscedasticity from landform-projected shadows that particularly affects small glaciers in steep and north-facing slopes.

| Uncertainty of spatial average | Griaz and Bourgeat | Bossons and Taconnaz | All glaciers |
|-------------------------------|---------------------|----------------------|-------------|
| σ_{M}no corr. (m)              | 0.027               | 0.006                | 0.002       |
| σ_{M}short-range corr. (m)    | 0.027               | 0.006                | 0.002       |
| σ_{M}long-range corr. (m)     | 0.061               | 0.049                | 0.024       |
| Glacier area (km²)            | 0.9                 | 16.3                 | 131.1       |

When uncertainties of volume change of several glaciers are propagated into that of the massif, correlations also come into play. We illustrate the propagation at different spatial scales.
by considering several glacier groups: one group with the two small neighbouring glaciers of Griaz and Bourgeat, another group with the two large neighbouring glaciers of Bossons and Taconnaz, and a third group including all 84 glaciers (Fig. 3). We find identical uncertainties when considering no correlation, or only short-range correlations (Table III). This reflects the fact that all glaciers are separated by at least 30 m, i.e. a distance larger than that of our short-range correlation (Table II). Long-range spatial correlations have a large impact on the total uncertainty, however, with a tenfold understimation of the uncertainty for all glaciers in the massif when omitting them. Increased uncertainties from long-range correlations mostly affect large neighbouring glaciers, as shown for Bossons and Taconnaz, but also affect smaller, disconnected glaciers such as Griaz and Bourgeat. The latter is true as long as the glaciers are within the correlation range of 11.2 km.

VII. CONCLUSIONS

In this study, we reviewed the literature on the accuracy and precision of DEMs. On the basis of the raised considerations regarding variable vertical precision and correlated noises, we proposed a non-stationary spatial framework for DEM uncertainty analysis. This framework allows to perform inference on a single difference between a DEM and independent elevation data on stable terrain, and to distinguish elevation biases from elevation variance. We developed robust methods to estimate and model both elevation heteroscedasticity and spatial correlation of elevation errors. We then validated that stable terrain is a reliable error proxy for other terrain types using pairs of DEMs derived from nearly simultaneous acquisitions for the Mont-Blanc massif and the Northern Patagonian Icefield.

We illustrated the impact of our methods when propagating uncertainties to pixel-scale and spatial derivatives of elevation. For the pixel-scale terrain slope, uncertainties are underestimated by up to a factor 10 in rough and steep topography when omitting elevation heteroscedasticity. For glacier volume changes, the uncertainty of the volume change of a glacier of 10 km³ is underestimated by a factor of 25 when omitting correlations with ranges of 3.9 and 11.2 km, despite their small cumulative magnitude of only 7% of the variance. This understimation of long-range spatial correlation affects many studies relying on instruments plagued by noise, such as the widely used DEMs from SRTM and ASTER.

We provide an implementation of our methods in the Python package xDEM [157], which includes, in particular, DEM alignment, correction, and uncertainty analysis. Spatial statistics have long been used for uncertainty analysis, yet often suffered from a lack of accessibility [126]. The wider application of such analysis was still deemed as "unrealized" a decade ago [133], possibly also due to the scarcity of open-source and documented tools for spatial statistics. By providing our methodological tools within the frame of a package embedded in high-level programming languages that efficiently pairs with remote sensing analysis, we hope to foster a consistent, reproducible and accessible uncertainty analysis of DEMs.

We highlight the genericity of our spatial framework for uncertainty analysis and of our estimation methods for dense and outlier-prone grid data. Our framework holds the potential to be extended to other geospatial data. Gridded surface displacement, for instance, profit from the same error proxy of stable terrain and are increasingly used in a variety of applications. To describe the precision of such spatially structured data, we advocate for the use of additional metrics. These metrics should describe potential heteroscedasticity and spatial correlation of errors, reported, for example, in a tabular manner — parameters of variogram models; discrete categories of heteroscedasticity. Ultimately, the adoption of such new metrics is critical to progress towards a realistic description of error structure in geospatial data, and a robust propagation of uncertainties in Earth system science assessments.

DATA AND CODE AVAILABILITY STATEMENT

The code developed for the processing and analysis of all data, and the generation of figures and tables in this article is available at https://github.com/rhugonnet/dem_error_study with routines implemented in the DEM analysis Python package xDEM available at https://github.com/GlacioHack/xdem with supporting documentation at https://xdem.readthedocs.io and the spatial statistics Python package SciKit-GStat available at https://github.com/mmnaellicke/scikit-gstat with supporting documentation at https://scikit-gstat.readthedocs.io.

AUTHOR CONTRIBUTIONS

R.H., F.B. and E.B. designed the study. F.B. performed an early analysis of spatial correlation for the Mont-Blanc case study. R.H. developed and tested non-stationary spatial statistics methods for DEMs with inputs from N.E., A.D., F.B. and E.B. R.H. implemented the methods in the Python package xDEM with main inputs from A.D. and E.M. All authors interpreted the results. R.H. performed the literature review and led the writing of the paper, and all other co-authors contributed.

ACKNOWLEDGMENTS

We thank Dorde Masovic for realizing Fig. 2 with vector graphics. We thank Alexandre Wadou for insightful exchanges on spatial statistics and Simon Gascoin for discussions on along-track undulations in Pléiades DEMs. We also thank Eiyjofur Magnússon and Joaquín Belart for providing insights into Gaussian simulations for DEMs. We thank Jonathan Bamber, César Deschamps-Berger and Matthias Huss for comments on early versions of the manuscript. SPOT6/7 data were obtained from GEOSUD (ANR-10-EQPX-20, programme ‘Investissements d’Avenir’). INRAE is a member of Labex OSUG. R.H. acknowledges a fellowship from the University of Toulouse. R.H., A.D., E.M. and D.F. acknowledge funding from the Swiss National Science Foundation, grant number 184634. E.B. acknowledges support from the French Space Agency (CNES) through ISIS and TOSCA programmes.

REFERENCES

[1] H. Taud, J.-F. Parrot, and R. Alvarez, “DEM generation by contour line dilation,” Comput. Geosci., vol. 25, no. 7, pp. 775–783, Aug. 1999.
[2] Q. Weng, “Quantifying uncertainty of digital elevation models derived from topographic maps,” in Advances in Spatial Data Handling, Springer Berlin Heidelberg, 2002, pp. 403–418.
[3] P. A. Rosen, S. Hensley, I. R. Joughin, F. K. Li, S. N. Madsen, E. Rodriquez, and R. M. Goldstein, “Synthetic aperture radar interferometry,” Proc. IEEE, vol. 88, no. 3, pp. 333–382, Mar. 2000.
[4] R. Bürgmann, P. A. Rosen, and E. J. Fielding, “Synthetic aperture radar interferometry to measure earth’s surface topography and its deformation,” Annu. Rev. Earth Planet. Sci., vol. 28, no. 1, pp. 169–209, May 2000.
Y. Bühler, S. Kumar, J. Veitinger, M. Christen, A. Stoffel, and Others,

T. Bolch, U. Kamp, J. Olsenholler, and M. Olubić, “Using ASTER planet SkySat triplet stereo and video imagery,” ISPRS J. Photogramm. Remote Sens., vol. 6, p. 233, 2018.

J. Oksanen and T. Sarjakoski, “Error propagation analysis of DEM-model inputs towards identifying the seismic fault,” Front Earth Sci., vol. 3, p. 32, 2015.

S. Kulp and B. H. Strauss, “Global DEM errors underpredict coastal vulnerability to sea level rise and flooding,” Front Earth Sci., vol. 4, p. 36, 2016.

A. Hooijer and R. Vernimmen, “Global LiDAR elevation data reveal greatest sea-level rise vulnerability in the tropics,” Nat. Commun., vol. 12, no. 1, p. 3592, Jun. 2021.

E. K. Barker, E. Mazurkovic, G. A. Neumann, M. T. Zuber, J. Haruyama, and D. E. Smith, “A new lunar digital elevation model from the lunar orbiter laser altimeter and SELENE terrain camera,” Icarus, vol. 273, pp. 346–355, Jul. 2016.

G. LeFavour and D. Alsdorf, “Water slope and discharge in the amanzon river estimated using the shuttle radar topography mission digital elevation model,” Geophys. Res. Lett., vol. 32, no. 17, Sep. 2005.

E. C. Geyman, W. J. van Pelt, A. C. Maloof, H. F. Aas, and J. Kohler, “Historical glacier change on svalbard predicts doubling of mass loss by 2100,” Nature, vol. 601, no. 7893, pp. 374–379, Jan. 2022.

M. Nolan, C. Larsen, and M. Sturm, “Mapping snow depth from manned aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry,” The Cryosphere, vol. 9, no. 4, pp. 1445–1463, Aug. 2015.

R. Marti, S. Gascoin, E. Berthier, M. d. Pinel, T. Houet, and D. Laffly, “Mapping snow depth in open alpine terrain from stereo satellite images,” The Cryosphere, vol. 12, no. 1, p. 3592, Jun. 2021.

C. Deschamps-Berger, S. Gascoin, E. Berthier, J. Deems, E. Gutmann, A. Dehecq, D. Shean, and M. Dumont, “Snow depth mapping from stereo satellite imagery in mountainous terrain: evaluation using airborne laser-scanning data,” The Cryosphere, vol. 14, no. 9, pp. 2925–2940, Sep. 2020.

M. Bagnardi, P. J. González, and A. Hooper, “High-resolution digital elevation model from tri-stereo pi-1 satellitie imagery for lava flow volume estimates at fogo volcano,” Geophys. Res. Lett., vol. 43, no. 12, pp. 6267–6275, Jun. 2016.

G. B. M. Pedersen, J. M. C. Belart, E. Magnússon, O. K. Vilumundaddróttir, F. Kielz, F. S. Sigurmundsson, G. Gisladóttir, and J. A. Benediktsson, “Hekla volcano, iceland, in the 20th century: Lava volumes, production rates, and effusion rates,” Geophys. Res. Lett., vol. 45, no. 4, pp. 1805–1813, Feb. 2018.

C. Ticehurst, S. Phinn, and A. Held, “Using multitemporal digital elevation model data for detecting canopy gaps in tropical forests due to cyclone damage: An initial assessment,” Austral Ecol., vol. 32, no. 1, pp. 59–69, Feb. 2007.

L. Magruder, H. Leigh, and A. Neuenschwander, “Evaluation of terrain and canopy height products in central african tropical forests,” Int. J. Remote Sens., vol. 37, no. 22, pp. 5365–5387, Nov. 2016.

K. Gudalović, J. Marešová, V. Barták, M. Szostak, J. Červenka, and V. Moudrý, “Use of TanDEM-X and SRTM-C data for detection of
deforestation caused by bark beetle in central european mountains,” Remote Sensing, vol. 13, no. 15, p. 3042, Aug. 2021.

[45] D. Felikson, T. C. Bartholomau, G. A. Catania, N. J. Korsgaard, K. H. Kjar, M. Mörhlig, B. Noël, M. van den Broeke, L. A. Stearns, E. L. Shroyer, D. A. Sutherland, and J. D. Nash, “Inland thinning on the greenland ice sheet controlled by outlet glacier geometry,” Nat. Geosci., vol. 10, no. 5, pp. 366–369, May 2017.

[46] D. E. Shean, L. K. Joughin, P. Dutrieux, B. E. Smith, and E. Berthier, “Ice shelf basal melt rates from a high-resolution digital elevation model (DEM) record for pine island glacier, antarctica,” The Cryosphere, vol. 13, no. 10, pp. 2633–2656, 2019.

[47] R. Hugonnet, R. McNabb, E. Berthier, B. Menounos, C. Nuth, L. Girod, R. Hugonnet, R. McNabb, E. Berthier, B. Menounos, C. Nuth, L. Girod, “Validation of vertical accuracy of freely available digital elevation model (DEM),” Int. J. Appl. Earth Obs. Geoinf., vol. 21, pp. 205–217, Apr. 2013.

[48] S. A. Kulp and B. H. Strauss, “New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding,” Nat. Commun., vol. 10, no. 1 p. 4844, Oct. 2019.

[49] A. Gruber, B. Wessel, M. Huber, and A. Roth, “Operational TanDEM-X DEM calibration and first validation results,” ISPRS J. Photogramm. Remote Sens., vol. 73, pp. 39–49, Sep. 2012.

[50] A. Dehecq, R. Millan, E. Berthier, N. Gourmelen, E. Trouvé, and V. Vionnet, “Elevation changes inferred from TanDEM-X data over the Mont-Blanc area: Impact of the X-Band interferometric bias,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 8, pp. 3870–3882, Aug. 2016.

[51] J. P. Best and N. D. McKay, “A method for registration of 3-D shapes,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 14, no. 2, pp. 239–256, Feb. 1992.

[52] A. Myronenko and X. Song, “Point set registration: coherent point drift,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 12, pp. 2268–2280, Dec. 2010.

[53] Y.-Q. Zhou, J. Park, and V. Koltun, “Fast global registration,” in Computer Vision – ECCV 2016. Springer International Publishing, 2016, pp. 766–782.

[54] T. Zhang and M. Cen, “Robust DEM co-registration method for terrain changes assessment using least trimmed squares estimator,” Adv. Space Res., vol. 41, no. 11, pp. 1827–1835, Jan. 2008.

[55] M. Huber, B. Wessel, R. McNabb, E. Berthier, N. Gourmelen, E. Trouvè, and A. Roth, “Operational TanDEM-X DEM calibration and first validation results,” ISPRS J. Photogramm. Remote Sens., vol. 73, pp. 39–49, Sep. 2012.

[56] J. Oksanen, “Digital elevation model error in terrain analysis,” Ph.D. dissertation, University of Helsinki, Faculty of Science, Nov. 2006.

[57] C. Papasaika-Hansuch, “Symposium on fusion of digital elevation models,” Ph.D. dissertation, ETH Zurich, 2012.

[58] D. F. Maune, Digital elevation model technologies and applications: the DEM users manual. Asprs Publications, 2007.

[59] P. F. Fisher and N. J. Tate, “Causes and consequences of error in digital elevation models,” Progress in Physical Geography: Earth and Environment, vol. 30, no. 4, pp. 467–489, Aug. 2006.

[60] S. P. Wechsler, “Uncertainties associated with digital elevation models for hydrologic applications: a review,” Hydrol. Earth Syst. Sci., vol. 11, no. 4, pp. 1481–1500, Aug. 2007.

[61] S. E. Lakshmi and K. Yarrakula, “Review and critical analysis on digital elevation models,” Geofizika, vol. 35, no. 2, pp. 129–157, 2018.

[62] L. Polidori and M. E. Hage, “Digital elevation model quality assessment methods: A critical review,” Remote Sensing, vol. 12, no. 21, p. 3522, Oct. 2020.

[63] F. Hebler and R. S. Purves, “The influence of elevation uncertainty on derivation of topographic indices,” Geomorphology, vol. 111, no. 1, pp. 4–16, Oct. 2009.

[64] T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seiber, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and M. Kobrick, “The shuttle radar topography mission,” Rev. Geophys., vol. 41, no. 1, pp. 1–16, Jan. 2003.

[65] D. Alsdorf, “The shuttle radar topography mission,” Rev. Geophys., vol. 39, no. 2, pp. 223–232, Feb. 2001.

[66] J. T.ward, M. Lefsky, T. Markus, A. Marchant, C. Naujokaitis, P. Morcrette, and A. Schepers, “The ICESat mission: On-orbit measurement performance,” Proc. IEEE, vol. 98, no. 5, pp. 735–750, May 2010.

[67] Q.-Y. Zhou, J. Park, and V. Koltun, “Fast global registration,” in IEEE Conference on Computer Vision – ECCV 2016. Springer International Publishing, 2016, pp. 766–782.

[68] M.-J. Noh and I. M. Howat, “Automated coregistration of repeat digital elevation models for surface elevation change measurement using geometric constraints,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 4, pp. 2247–2260, Apr. 2014.

[69] J. B. Abshire, X. Sun, H. Riris, M. J. Sirota, J. F. McCrary, S. Palm, D. Yu, and P. Liiva, “Geoscience laser altimeter system (GLAS) on the ICESat mission: Orbit measurement performance,” Geophys. Res. Lett., vol. 32, no. 21, 2005.

[70] W. Abdalati, H. J. Zwally, R. Bindschadler, B. Csató, S. L. Farrell, H. A. Fricker, D. Harding, R. Kwok, M. Lefsky, T. Markus, A. Marakhov, T. Neumann, S. Palm, B. Schutz, B. Smith, J. Sinninghe, and C. Webb, “The ICESat-2 laser altimetry mission,” Proc. IEEE, vol. 98, no. 5, pp. 735–750, May 2010.

[71] J. Vaze, J. Teng, and G. Spencer, “Impact of DEM accuracy and resolution on topographic indices,” Environmental Modelling & Software, vol. 25, no. 10, pp. 1086–1098, Oct. 2010.

[72] J. Gardelle, E. Berthier, and Y. Arnaud, “Impact of resolution and radar penetration on glacier elevation changes computed from DEM differencing,” J. Glaciol., vol. 58, no. 208, pp. 419–422, 2012.

[73] J. Vaze, J. Teng, and G. Spencer, “Impact of DEM accuracy and resolution on topographic indices,” Environmental Modelling & Software, vol. 25, no. 10, pp. 1086–1098, Oct. 2010.

[74] J. Gardelle, E. Berthier, and Y. Arnaud, “Impact of resolution and radar penetration on glacier elevation changes computed from DEM differencing,” J. Glaciol., vol. 58, no. 208, pp. 419–422, 2012.

[75] J. Vaze, J. Teng, and G. Spencer, “Impact of DEM accuracy and resolution on topographic indices,” Environmental Modelling & Software, vol. 25, no. 10, pp. 1086–1098, Oct. 2010.

[76] J. Gardelle, E. Berthier, and Y. Arnaud, “Impact of resolution and radar penetration on glacier elevation changes computed from DEM differencing,” J. Glaciol., vol. 58, no. 208, pp. 419–422, 2012.

[77] J. Vaze, J. Teng, and G. Spencer, “Impact of DEM accuracy and resolution on topographic indices,” Environmental Modelling & Software, vol. 25, no. 10, pp. 1086–1098, Oct. 2010.
[133] G. B. M. Heuvelink, “Analysing uncertainty propagation in GIS: Why is it not that simple?” in Uncertainty in Remote Sensing and GIS, Chichester, UK: John Wiley & Sons, Ltd. Jul. 2006, pp. 155–165.

[134] R. A. Beyer, O. Alexandrov, and S. McMichael, “The anes stereo pipeline: NASA’s open source software for deriving and processing terrain data,” Earth and Space Science, vol. 5, no. 9, pp. 537–548, Sep. 2018.

[135] NASA/METI/AIST/Japan Space systems, and U.S./Japan ASTER Science Team, “ASTER level 1A data set - reconstructed, unprocessed instrument data,” 2001.

[136] E. Rupnik, M. Daikir, and M. Pierrat Deseilligny, “MicMac – a free, open-source solution for photogrammetry,” Open Geospatial Data, Software and Standards, vol. 2, no. 1, p. 14, 2017.

[137] NASA JPL, “NASADEM SRTM-only height and height precision mosaic global 1 arc second;” 2020.

[138] M.-J. Noh and I. M. Howat, “The surface extraction from TIN based search-space minimization (SETSM) algorithm,” ISPRS J. Photogramm. Remote Sens., vol. 129, pp. 55–76, Jul. 2017.

[139] C. Porter, P. Morin, I. Howat, M.-J. Noh, B. Bates, K. Peterman, S. Keesey, M. Schlenk, J. Gardiner, K. Tomko, M. Willis, C. Kellher, M. Cloutier, E. Husby, S. Foga, H. Nakamura, M. Platson, M. Wethington, Jr, C. Williamson, G. Bauer, J. Enos, G. Arnold, W. Kramer, P. Becker, A. D’Souza, P. Cummins, F. Laurier, and M. Bojesen, “ArcticDEM;” Nov. 2018.

[140] RGI Consortium, “Randolph glacier inventory – a dataset of global glacier outlines: Version 6.0: Technical report, global land ice measurements from space, colorado, USA;” 2017.

[141] ESA, “Land cover CCI product user guide version 2. tech. rep;” 2017.

[142] J. Höhle and M. Höhle, “Accuracy assessment of digital elevation models by means of robust statistical methods;” ISPRS J. Photogramm. Remote Sens., vol. 64, no. 4, pp. 398–406, Jul. 2009.

[143] B. K. P. Horn, “Hill shading and the reflectance map,” Proc. IEEE, vol. 69, no. 1, pp. 14–47, Jan. 1981.

[144] L. W. Zevenbergen and C. R. Thorne, “Quantitative analysis of land surface topography;” Earth Surf. Processes Landforms, vol. 12, no. 1, pp. 47–56, Jan. 1987.

[145] M. B. Wilk and R. Gronadeskian, “Probability plotting methods for the analysis of data;” Biometrika, vol. 55, no. 1, pp. 1–17, Mar. 1968.

[146] P. A. Dowd, “The variogram and kriging: Robust and resistant estimators;” in Geostatistics for Natural Resources Characterization, G. Verly, M. David, A. G. Journel, and A. Marechal, Eds. Dordrecht: Springer Netherlands, 1984, pp. 91–106.

[147] M. Mälén and H. D. Schneider, “Scikit-GIS:Scikit-GIS 0.2.6: A scipy flavored geostatistical analysis toolbox written in python;” Nov. 2019.

[148] G. Matheron, Les variables régionalisées et leur estimation: une application de la théorie de fonctions aléatoires aux sciences de la nature, Masson et CIE, 1965, vol. 4957.

[149] N. Cressie and D. M. Hawkins, “Robust estimation of the variogram: I;” Journal of the International Association for Mathematical Geology, vol. 12, no. 2, pp. 115–125, Apr. 1980.

[150] M. G. Genton, “Highly robust variogram estimation;” Math. Geol., vol. 30, no. 2, pp. 213–221, Feb. 1998.

[151] T. M. Burgess and R. Webster, “Optimal interpolation and isarithmic mapping of soil properties;” J. Soil Sci., vol. 31, no. 2, pp. 315–331, Jan. 1980.

[152] S. Müller and L. Schüler, “GeoStat-Framework/GSTools v1.3.3 ‘pure pink’;” 2021.

[153] E. J. Pebesma and C. G. Wesseling, “gstof : A program for geostatistical modelling, prediction and simulation;” Comput. Geosci., vol. 24, no. 1, pp. 17–31, 1998.

[154] R. Webster and M. A. Oliver, Geostatistics for Environmental Sciences, 2nd ed. Wiley, Nov. 2007.

[155] H. Liu and K. C. JezeK, “Investigating DEM error patterns by directional variograms and fourier analysis;” Geogr. Anal., vol. 31, no. 3, pp. 249–266, Sep. 2010.

[156] E. Rodríguez, C. S. Morris, and J. E. Belz, “A global assessment of the SRTM performance;” Photogrammetric Engineering & Remote Sensing, vol. 72, no. 3, pp. 249–260, 2006.

[157] xdem contributors, “xdem;” May 2021.

Romain Hugonnet received his MSc. degree in 2017 from Ecole Centrale de Lille, France. Since 2018, he has been working as a Ph.D. student for the Laboratory for Space Studies in Geophysics and Oceanography, Toulouse, France and the Laboratory of Hydraulics, Hydrology and Glaciology, ETH Zürich, Switzerland. His research aims at monitoring the Earth system by remote sensing, with a focus on glaciology and digital elevation models. His fields of expertise are satellite imagery and spatiotemporal analyses, including uncertainty quantification.

Fanny Brun received her Ph.D. degree in 2018 from the University Grenoble Alpes, France. Since 2020, she has been working for IRD in the Institute of Environmental Geosciences, Grenoble, France. Her research aims at observing glacier changes from remote sensing and field techniques, with a focus in Asia. She currently serves as Scientific Editor for the Journal of Glaciology.

Etienne Berthier received his Ph.D. degree in 2005 from the University of Toulouse, France. Since 2007, he has been working for CNRS in the Laboratory for Space Studies in Geophysics and Oceanography, Toulouse, France. His research aims at observing and understanding the effect of climate change on mountain glaciers and ice caps and quantifying their contribution to sea level rise. His field of expertise is satellite remote sensing, including the retrieval of ice topography from optical images to measure ice elevation changes. He currently serves as the Scientific Editor for The Cryosphere and La Météorologie (in French). He is member of the European Geophysical Union and the International Glaciological Society.

Amaury Dehecq received his Ph.D. degree in remote sensing in 2015 from the University Grenoble Alpes, France. From 2019 to 2021, he worked as a postdoctoral researcher at ETH Zurich’s Laboratory of Hydraulics, Hydrology and Glaciology (VAW). Since 2021, he has been working for the French National Research Institute for Sustainable Development (IRD) and based at the Institute for Geosciences and Environmental research (IGE), Grenoble, France. His research focuses on estimating 20th century mountain glacier changes from remote sensing observations, using a combination of modern satellite, declassified spy satellite, and historical airborne/terrestrial images.

Erik Schytt Mannerfelt received his MSc degree in 2020 from UiT — The Arctic University of Norway. Between 2020 and 2022, he has worked at the Laboratory of Hydraulics, Hydrology and Glaciology, ETH Zurich, Switzerland, as a research assistant. He currently is working at the Laboratory of Hydraulics, Hydrology and Glaciology, ETZH, Switzerland, as a research assistant. His research focuses on fieldwork-, modelling-, and photogrammetry-based methods to constrain late Little Ice Age to contemporary glacier changes and dynamics.
Nicolas Eckert received his Ph.D. degree in 2007 from AgroparisTech. Since 2008 he is researcher at INRAE Grenoble, co-head of the mountain risk team since 2018. He is also task officer in charges of Environmental risks for the AQUA department at INRAE, and for the ALLENV1 research federation. He serves as Associate Chief Editor for Journal of Glaciology and as Scientific Editor for Cold Regions Science and Technology. His research is at the crossroads between geosciences and statistical modelling, with applications to mountain risks, mountain climatology and glaciers. Additional interests in the socio-historical component of risk makes him involved in interdisciplinary research addressing all dimensions of mountain risks.

Daniel Farinotti received his doctoral degree in 2010 from the Swiss Federal Institute of Technology in Zurich (ETH Zurich). Since 2016 he leads the Professorship of Glaciology at ETH Zurich’s Laboratory of Hydraulics, Hydrology and Glaciology (VAW), a position jointly affiliated to the Swiss Federal Institute for Forest, Snow and Landscape Research WSL. His research focuses on the evolution of glaciers and the implications for water resources, notably including the estimation of glacier ice thickness from surface characteristics, the long-term modelling of glacier mass budgets, the estimation of the runoff contributions from glacierized catchments, or the implications for water resource management in high-mountain environments.