Citation: **Hirt, C**. (2018), Artefact detection in global digital elevation models (DEMs): The maximum slope approach and its application for complete screening of the SRTM v4.1 and MERIT DEMs, *Remote Sensing of Environment*, 207, 27-41, doi.org/10.1016/j.rse.2017.12.037.

Artefact detection in global digital elevation models (DEMs): The maximum slope approach and its application for complete screening of the SRTM v4.1 and MERIT DEMs

Christian Hirt

Institute for Astronomical and Physical Geodesy (IAPG) & Institute for Advanced Study (IAS),

Technical University Munich, Arcisstraße 21, 80333 Munich, Germany

c.hirt@tum.de

Orcid: http://orcid.org/0000-0002-3176-7939

Abstract

Despite post-processing efforts by space agencies and research institutions, contemporary global digital elevation models (DEMs) may contain artefacts, i.e., erroneous features that do not exist in the actual terrain, such as spikes, holes and line errors. The goal of the present paper is to illuminate the artefact issue of current global DEM data sets that might be an obstacle for any geoscience study using terrain information. We introduce the Maximum Slope Approach (MSA) as a technique that uses terrain slopes as indicator to detect and localize spurious artefacts. The MSA relies on the strong sensitivity of terrain slopes for sudden steps in the DEM that is a direct feature of larger artefacts. In a numerical case study, the MSA is applied for globally complete screening of two SRTM-based 3 arc-second DEMs, the SRTM v4.1 and the MERIT-DEM. Based on 0.1°x 0.1° sub-divisions and a 5 m/m slope threshold, 1,341 artefacts were detected in SRTM v4.1 vs. 108 in MERIT. Most artefacts spatially correlate with SRTM voids (and thus with the void-filling) and not with the SRTM-measured elevations. The strong contrast in artefact frequency (factor ~12) is attributed to the SRTM v4.1 hole filling. Our study shows that over parts of the Himalaya Mountains the SRTM v4.1 data set is contaminated by step artefacts where the use of this DEM cannot be recommended. Some caution should be exercised, e.g., over parts of the Andes and Rocky Mountains. The same holds true for derived global products that depend on SRTM v4.1, such as gravity maps. Primarily over the major mountain ranges, the MERIT model contains artefacts, too, but in smaller numbers. As a conclusion, globally complete artefact screening is recommended prior to the public release of any DEM data set. However, such a quality check should also be considered by users before using DEM data. MSA-based artefact screening is not only limited to DEMs, but can be applied as quality assurance measure to other gridded data sets such as digital bathymetric models or gridded physical quantities such as gravity or magnetics.

Key words: Digital elevation model, digital terrain model, artefact, outlier, data screening, void filling, maximum slope approach, quality, SRTM v4.1, MERIT-DEM, SRTM v2.1

1. Introduction

Since the beginning of the 21st century, remote sensing from dedicated space-borne platforms has revolutionized our knowledge of the Earth topography. Notably the (1) Shuttle Radar Topography

Mission (SRTM; Farr et al. 2007), the (2) Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER; Tachikawa et al. 2011), the (3) Advanced Land Observing Satellite with the Panchromatic Remote-sensing instrument for Stereo Mapping (ALOS/PRISM) and associated ALOS World 3D (AW3D) DEM, cf. Tadono et al. (2015) and the (4) TerraSAR-X Add-on for Digital Elevation Measurements (TanDEM-X; Wessel at al. 2016) have sampled the Earth surface geometry with unprecedented resolution and spatial coverage. As a result of these missions, digital elevation models (DEMs) as geometric representations of the surface relief have been produced with spatial resolutions of 1 to 3 arc-seconds (~30 to ~90 m in latitudinal direction) or better and near-global coverage (Hirt 2015). Today, DEM data sets form a critical backbone in several applications in engineering, geo- and environmental sciences. DEMs have become a common good, e.g., as base layer for personal navigation systems, OpenStreetMap and Google Maps.

In light of the widespread use, a realistic assessment of the DEM quality (that is, how closely the digital model represents the actual terrain surface) is important. DEMs may be subject to imperfections, such as vertical and horizontal errors, speckle-noise, voids (unobserved areas, also denoted as holes) and biases (offsets) that can vary regionally. Also, artificial features that misrepresent the actual terrain surface may be encountered in DEM data sets. Examples include artificial spikes, sinkholes, steps, pixel defects, line and masking (clipping) errors. Among all imperfections in DEM data sets, artefacts may be the most crucially problematic error source for DEM applications (Brown and Bara 1994, Lecours et al. 2017). In areas such as hydrology, hydrodynamics and flooding analysis (Wechsler 2007, Yamazaki et al. 2017), geostatistics, geomorphology, and geomorphometry (Pike et al. 2009, Reuter et al. 2009), geometrical and physical geodesy (Torge and Müller 2012, Hirt et al. 2014) these unwanted features may falsify the outcome of DEM-based analyses.

Over the years progress has been made with the production of improved DEM data sets, particularly based on SRTM and ASTER mission products. NASA's Jet Propulsion Laboratory (JPL) and the National Geospatial Intelligence Agency (NGA) have post-processed the initial ("unfinished") SRTM DEM release v1 (2003/2004) to remove some artificial error sources (pits, spikes) and fill minor voids in the "finished-grade" second version (Slater et al. 2006) that was released as v2.0 in 2004/2005 and v2.1 in 2009 (Crippen, pers. comm. 2017). A third model generation with improved void-filling based on ASTER elevation data has been released as Version v3 (2015), cf. NASA (2015) report. As far as ASTER is concerned, some of the spurious vertical artefacts (e.g., spikes representing clouds instead of the terrain) contained in the first "research-grade" release GDEM v1 (2009) have been removed in release GDEM v2 (2011) through inclusion of additional remote sensing data, and better delineation of water bodies has been achieved in GDEM v3 Abrams (2016) that is expected to be released in 2018 (Abrams, pers. comm. 2017). Further, research institutions have released post-processed SRTM versions that differ from the agency products in view of the void filling procedures and removal of additional errors. Examples include the SRTM v4.1 release by CGIAR-CSI (Reuter et al. 2007, Jarvis et al. 2008), the Altimetry Corrected Elevation (ACE2) DEM (Berry et al. 2010), the EarthEnv-DEM90 (Robinson et al. 2014), and the Multi-Error-Removed Improved-Terrain (MERIT) DEM by Yamazaki et al. (2017).

Unfortunately, despite these multiple year-long post-processing efforts by space agencies and research institutions, contemporary global DEMs are not free of artefacts, as the DEM user might be tempted to assume. A good statement was made by Wechsler (2007, p1482) that [modern digital data sets such as DEMs] may "*lure users into a false sense of security regarding the accuracy and*

precision of the data. Potential errors, and their effect on derived data and applications based on that data, are often far from users' consideration".

As we demonstrate in this paper, current global DEMs may contain spurious artefacts, such as pixel, line and edge defects, steps, pits and spikes as well as unfilled voids, and most of these unwanted features are a result of the DEM editing processes applied by the producers. The somewhat surprising presence of serious artefacts in current DEMs suggests that artefact testing procedures are not yet routinely applied prior to the release of DEMs.

1.1 Artefact detection in the literature

Methods for artefact detection and their removal have been discussed in several papers. Polidori et al. (1991) proposed to study the fractal geometry of a DEM to detect artefacts. Brown and Bara (1994) detected DEM systematic errors based on semivariance and fractal analysis. Oimoen (2000) investigated the detection and removal of production artefacts (mostly line errors) and emphasized the detrimental role of such features on derived DEM products "*such as slope, aspect and hydrology*". Albani and Klinkenberg (2003) describe a spatial filter for stripe removal in DEMs, while Arrell et al. (2008) tackle the problem of stripe removal with spectral techniques. Feng et al. (2004) presented a technique for removal of cloud-related spikes in ASTER elevation data. Lindsay and Creed (2005) study the removal of artificial depressions in DEMs in the context of hydrodynamic modelling, and Lindsay and Creed (2006) studied techniques to distinguish between artificial and natural depressions in DEM data sets. Reuter et al. (2009) state that because artefacts are "*distinct erratic features, most of them can be detected visually in 3D views, by using sun shading or simple GIS operations*" and note that [handling of artefacts is] "*especially important for land-surface parameters derived from second order derivatives (curvatures), aspect map and/or hydrological parameters*" (Reuter et al. 2009, p91). Villa Real et al. (2013) presents an algorithm for detection of vertical artefacts in DEMs that relies on comparisons against reference data. Polidori et al. (2014) noted the dependency of elevation derivatives on artefacts and tested the directional distribution of slopes that could possibly reveal artefacts in the data. Hirt et al. (2014) used extreme gravity values derived from topography to detect artificial depressions in SRTM elevation data. Merryman Boncori (2016) reports local shifts in SRTM DEM data that can be interpreted as artificial steps between DEM data tiles and Lecours et al. (2017) assess the influence of artefacts in digital bathymetry models on habitat maps. From the literature overview, geomorphometric quantities (Pike et al. 2009) such as horizontal gradients and slope (maximum inclination) are particularly sensitive for artefacts in DEM data sets: This dependency can be exploited, as shown in the present paper, to develop an approach for artefact detection in contemporary DEM data sets.

1.2 This study

The primary goal of the present paper is to detect and investigate spurious artefacts in current global DEM data sets that might be an obstacle in DEM applications requiring realistic terrain derivatives, such as hydrology and hydrodynamics, geomorphology, topographic mapping and gravity modelling. As secondary goal, the paper shall increase awareness in the producer and user community for the artefact problem that may even affect the most recently released products incorporating edited SRTM data.

We introduce the Maximum Slope Approach (MSA) as a technique that uses slopes as indicator to detect and localize spurious artefacts (Sect. 2). The MSA exploits the strong sensitivity of terrain

slopes for sudden steps in the DEM that are a direct feature of larger artefacts. We apply the MSA for globally complete screening (inspection) of two selected 3-arc second resolution SRTM releases (Sect. 3) in a numerical case study (Sect. 4). The chosen data sets are (i) the widely used SRTM v4.1 by CGIAR-CSI (Jarvis et al. 2008) and (ii) the new SRTM-based MERIT-DEM (Yamazaki et al. 2017), that can be considered a substantially edited elevation product where error sources have been reduced or removed. In the numerical case study, large terrain slopes will be automatically detected and localized in both SRTM data sets (Sect. 4.1), and semi-automatically classified into natural terrain features and artefacts (Sec. 4.2). The frequency of artefact occurrences in both products and their geographic distribution is analysed in Sect. 4.3 and discussed in Sect. 5.1. To exemplify artefacts and natural slopes we visualize selected DEM samples (see geographical location map in Fig. 1).

Our MSA-based artefact detection procedure is simple. It can in principle be applied with the DEM data itself; however, comparisons with a second DEM product increase the performance of the approach (Sect. 5.2). The MSA can easily be applied on all global DEMs, e.g., from the ASTER, ALOS/PRISM and TanDEM-X sensors, to ensure that spurious artefacts – if contained in the data – are detected before the public data release or application of the DEM. Application of the MSA on other gridded geodata products that may potentially contain artefacts, like planetary topography, Earth bathymetry, magnetics or gravity is possible as well (Sect. 6).

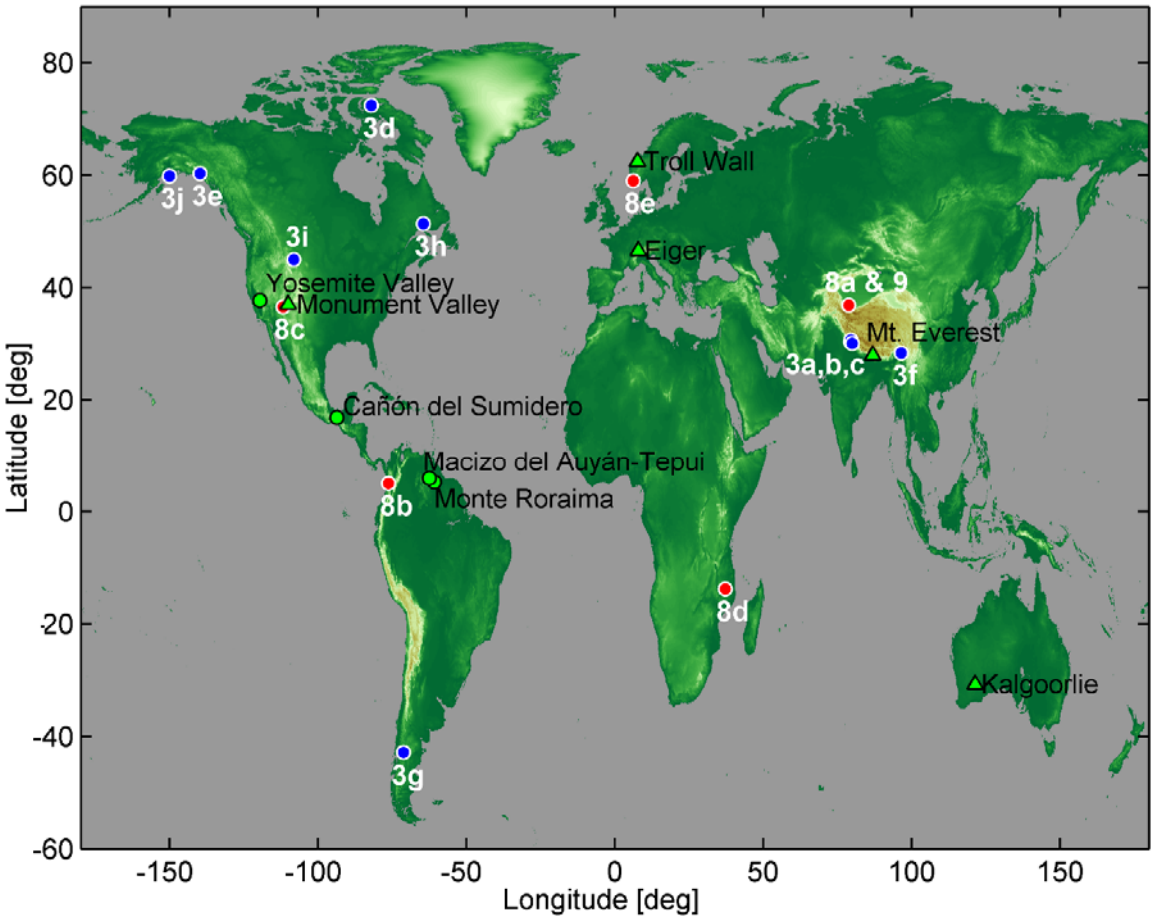


Fig. 1. World-wide location of selected DEM samples shown and discussed in this study. Green signatures show natural features (cf. Fig. 2 and 7), blue signatures are used for general examples of artefact types (cf. Fig. 3) and red signatures for artefacts types present in SRTM v4.1, but absent in MERIT (cf. Fig. 8 and Fig. 9).

2. Methods

2.1 Definition, cause and examples of artefacts

The term artefact, as aimed at and used in this study, denotes distinct step-like disruptions of the DEM-represented terrain surface that do not exist in the actual (real) terrain. From a statistical point of view, these step-like artefacts can be considered as gross errors (outlier) in the DEM heights. Where artefacts are present, the actual terrain is grossly misrepresented by the DEM. As characteristic feature of artefacts, large steps occur between neighbouring DEM cells (pixels). Consequently, terrain derivatives (gradients, slope) are directly influenced by artefacts, so can be used as indicators for their presence in the DEM. In our study, we only focus on step-like artefacts that result in large terrain derivatives.

Somewhat complementary to step-like artefacts, areas of excessive smoothness may be present in DEM data sets. These "smoothness artefacts" can occur, e.g., over mountainous terrain when a larger void area is filled through interpolation instead of auxiliary DEM data sources, or use of auxiliary DEMs too low in resolution (cf. Crippen et al. 2016, p127). As a result of the smoothness, terrain derivatives will be small, so may be difficult to distinguish from other naturally smooth topographic features, e.g., plateaus. Comparisons with independent DEM data (from other space missions) may be suitable to detect smoothness artefacts. Alternatively, fractal dimension measures might be deployed as indicator for these artefacts (Polidori et al. 1991). Smoothness artefacts are not dealt with in this study.

(Step-like) artefacts can occur in DEMs as single cell defects (e.g., unfilled voids), line errors (often along meridians or parallels, resulting e.g., from erroneous DEM merging), artificial spikes, sinkholes and steps (as result of inadequate void filling procedures), and steps at land-sea transitions (where coastlines or water body contours are inconsistent with the DEM heights). Visual examples for these artefacts are given in Sects. 2.2 and 4.3. In some cases, artefacts can also be the result of the actual DEM generation (from space observations to DEM heights). Well-known instances of these artefacts are the striping effect and spikes (clouds) in the ASTER GDEM data set (e.g., Feng et al. 2004). In other cases artefacts are the result of the DEM editing (post-processing), where unobserved areas (voids) of a "research-grade" DEM product are filled through interpolation or auxiliary data, data masks (e.g., coast lines) are applied, or different data sources are merged to extend the DEM coverage. Generally, the size of artefacts may vary between 1 (single cell defect) to 100s or more DEM cells, e.g., when line or void-filling artefacts are present in the DEM. As will be shown in Sect. 5, even regions covering several 10,000 DEM cells can be affected by artefacts.

In principle, artefacts may cause elevation changes between neighbouring DEM cells ranging between few 0.1 m/m and several 10 m/m. In this study, we aim to detect artefact-induced elevation changes that exceed a given threshold, e.g., 5 m/m (equivalent to a step of ~450 m for 90 m DEM resolution). Artefacts with associated lesser influence on the terrain slope are increasingly difficult to distinguish from the burgeoning number of natural terrain slopes of similar size, so are not in the focus of this study.

2.2 The maximum slope approach (MSA)

The maximum slope approach (MSA) relies on the (a) computation of terrain slopes from the DEM, and subsequent (b) analysis and classification of maximum slopes. The computation of slopes from

DEM data is a fundamental task of geomorphometry (Pike et al. 2009) and various algorithms have been proposed (e.g., Mark 1975, Skidmore 1989, Warren et al. 2004).

2.2.1 Slope computation

The gridded DEM data set is a discretisation of the bivariate height function $H(\varphi, \lambda)$ that depends on the geographic coordinates latitude φ and longitude λ. Slope is defined here as the maximum gradient in H (after Warren et al. 2004)

$$slope = \sqrt{H_\varphi^2 + H_\lambda^2} \tag{1}$$

with H_φ denoting the partial derivative of H in latitude, and H_λ on longitude direction. Our unit for $slope$ is meters/meters (e.g., 20 m elevation change over 90 m cell size); a conversion to slope angles is not required. For the approximation of the partial derivatives H_φ and H_λ, several methods are available, e.g., spline functions or polynomials (Warren et al. 2004) that use the surrounding cells around the computation point, e.g., a 3 x 3 or 5 x 5 cell neighbourhood. In principle, the more surrounding cells are involved in the determination of H_φ and H_λ, the smoother may be the estimates of the resulting slope, which can be useful, e.g., when the influence of outliers in H is to be mitigated (cf. Lindsay and Creed 2005, p3120, who state that the "*degree to which a surface derivative is affected by depression removal is related to the number of neighbouring cells used to calculate the derivative* […]") In view of the rationale of this study, however, it is desired that artefacts fully propagate into the partial derivatives H_φ and H_λ. This is achieved here by approximating the partials H_φ and H_λ with difference quotients computed in a 2 x 1 and 1 x 2 cell neighbourhood:

$$H_\varphi = \frac{\partial H}{\partial \varphi} \approx \frac{H_1 - H_2}{dS_{12}} \tag{2}$$

$$H_\lambda = \frac{\partial H}{\partial \lambda} \approx \frac{H_1 - H_3}{dS_{13}} \tag{3}$$

where H_2 is the cell located South of H_1 and H_3 is the cell located West of H_1. As will be shown in the numerical study, the chosen approximation for the partials H_φ and H_λ is suitable for artefact detection, but other choices could possibly be made for the computation of the derivatives. The distances between the cell centres are approximated as arc lengths on the surface of the ellipsoid

$$dS_{12} \approx M d\varphi \tag{4}$$

$$dS_{13} \approx N \cos\varphi \, d\lambda \tag{5}$$

where $d\varphi = d\lambda$ is the distance between the grid cells (in radian), M the meridian radius of curvature

$$M = \frac{a(1 - e^2)}{(1 - e^2 \sin^2 \varphi)^{3/2}} \tag{6}$$

and N the radius of curvature in the prime vertical

$$N = \frac{a}{(1 - e^2 \sin^2 \varphi)^{1/2}} \tag{7}$$

(Torge and Müller 2012, p95), that depend on latitude φ, and the GRS80 (Geodetic Reference System 1980) ellipsoid parameters $a = 6{,}378{,}137.0$ m (semi major-axis) and $e \approx 0.081\,819\,191\,0428$ (first numerical eccentricity), cf. Torge and Müller (2012, p109).

2.2.2 Global search for maximum slopes

Given the large number of elevation cells in a global 3 arc-second DEM (e.g., the MERIT data set comprises ~22 billion elevation cells), it is useful to apply a divide-and-conquer strategy for the detection of artefacts: Here, slopes (Eq. 1) are calculated over all cells of a given DEM subdivision (aka tile, e.g., 1° x 1°, with n rows and m columns) and maximum slope values

$$maxslope = max\left(slope_{i,j}\right), \quad i \in [1..n], j \in [1..m] \tag{8}$$

are computed as key indicator. The procedure is repeated for all tiles of the DEM under inspection, such that the maximum slope in each tile is determined.

- If for a DEM tile $maxslope$ falls below a certain threshold, e.g., 5 m/m, large artefacts (as defined in Sect. 2.1) are absent.
- Conversely, if $maxslope$ exceeds the threshold, the DEM tile contains suspiciously large slopes. Therefore, artefact(s) *could* be present in the tile. In that case further inspection is done in terms of smaller subdivisions, e.g., 0.1°x 0.1° tiles in order to narrow down the locations of the maximum slopes.

For all subdivisions (here 0.1°x 0.1°) with $maxslope$ exceeding the threshold too, the precise latitude φ and longitude λ of the maximum slope value are determined and local maps of the DEM heights H and their slope are generated for inspection and classification of the slopes.

2.2.3 Natural slopes in the terrain vs. threshold

In order to choose a suitable value for the threshold, it is useful to analyse how rugged topographic features such as major mountains and cliffs are represented in the 90-m resolution DEM and in the derived gradient and slope maps (Fig. 2). Naturally occurring slopes can reach maximum values of ~7-8 m/m over the Mount Everest area (panel a), ~5 m/m in the vicinity of the Eiger north face in Switzerland (panel b) and about ~8 m/m for Norway's Troll Wall (panel c) which features one of the steepest cliffs of Earth's land topography. Opposed to this, maximum slopes associated with the mesas of the U.S. Monument Valley are at the level of ~2-3 m/m (Fig. 2d).

As an example for major man-made terrain features, Fig. 2e shows the heights and geomorphometric parameters for the Kalgoorlie Superpit (Western Australia). It is seen that maximum gradients are below 2 m/m. Not displayed here, but this applies also for other man-made pits such as the Bingham Canyon Mine (U.S.). The chosen examples suggest that natural slopes are less than ~ 9 m/m and those of man-made features should not exceed a value of ~2 m/m in 3 arc-sec DEMs. Thus, a threshold of 10 m/m (equivalent to an elevation change among two adjoining 90m x 90 m cells of up to ~900 m) seems to be a safe choice to distinguish between slopes associated with artefacts and slopes of natural features (cf. Sect. 4).

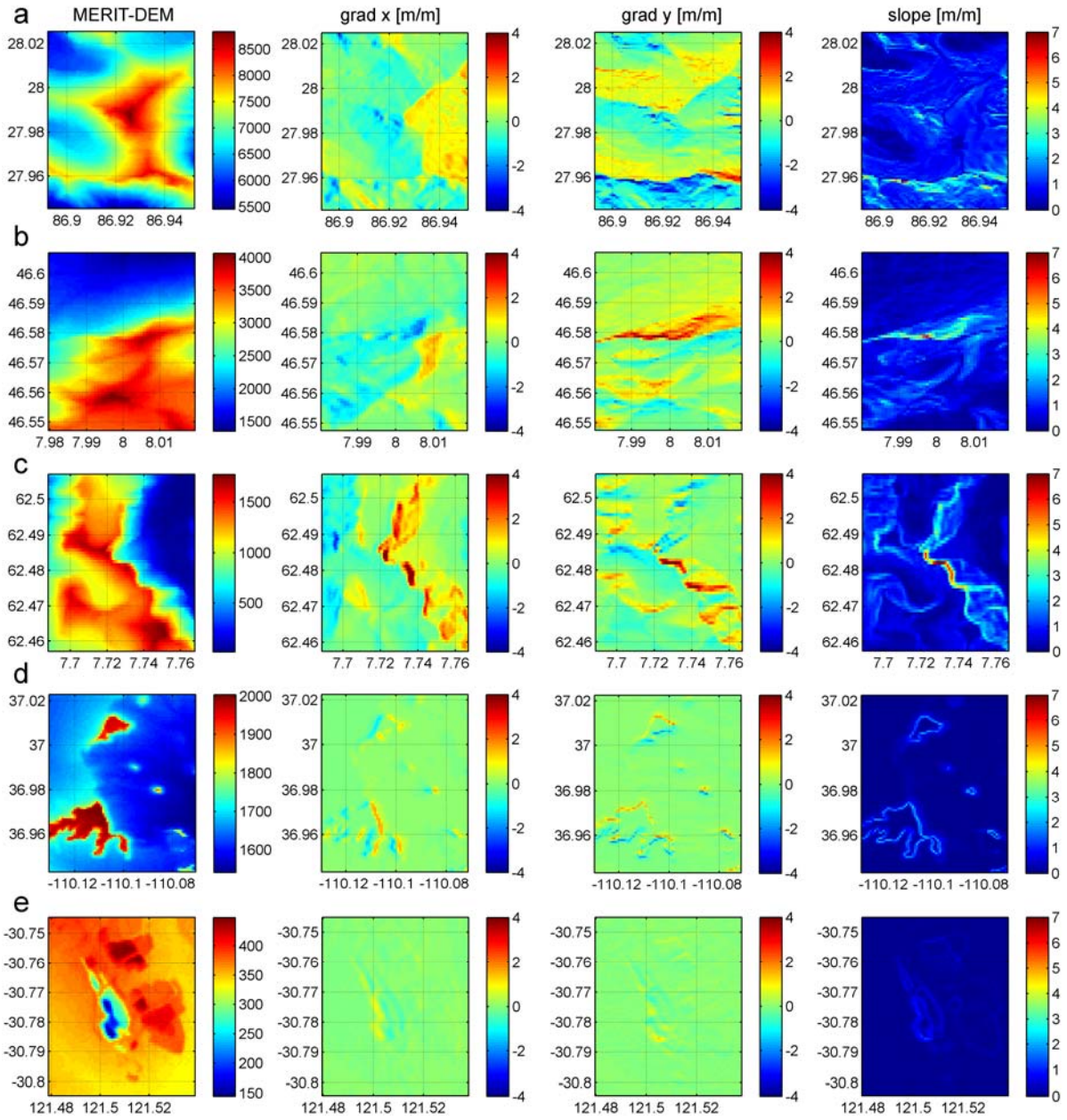


Fig.2. Elevations of selected topographic features (left), their horizontal gradients (middle columns) and their slope (right). Panels a) Mount Everest, b) Eiger (Switzerland), c) Troll Wall (Norway), d) Monument Valley (USA), e) Superpit in Kalgoorlie (Western Australia). Elevations in unit m, gradients and slopes in m/m, data from MERIT.

2.2.4 Classification artefact vs. natural feature

While a threshold of ~10 m/m separates the largest artefacts in 3 arc-sec DEMs from the natural terrain slopes, other artefacts with somewhat smaller associated maximum gradients (e.g., 5 to 10 m/m) can be detected in the DEM. In order to distinguish between natural terrain features and artefacts, however, some form of inspection and assessment of the candidate locations is required. Our criterion is geomorphological plausibility for natural features and, conversely, implausibility for artefacts. For instance, logical drainage structures visible in the height maps and slope maps around locations with large gradients are a good indicator for natural terrain features. Also, smoothly varying lines of large slopes may often be attributable to natural cliffs and walls (cf. Fig. 2a-2c).

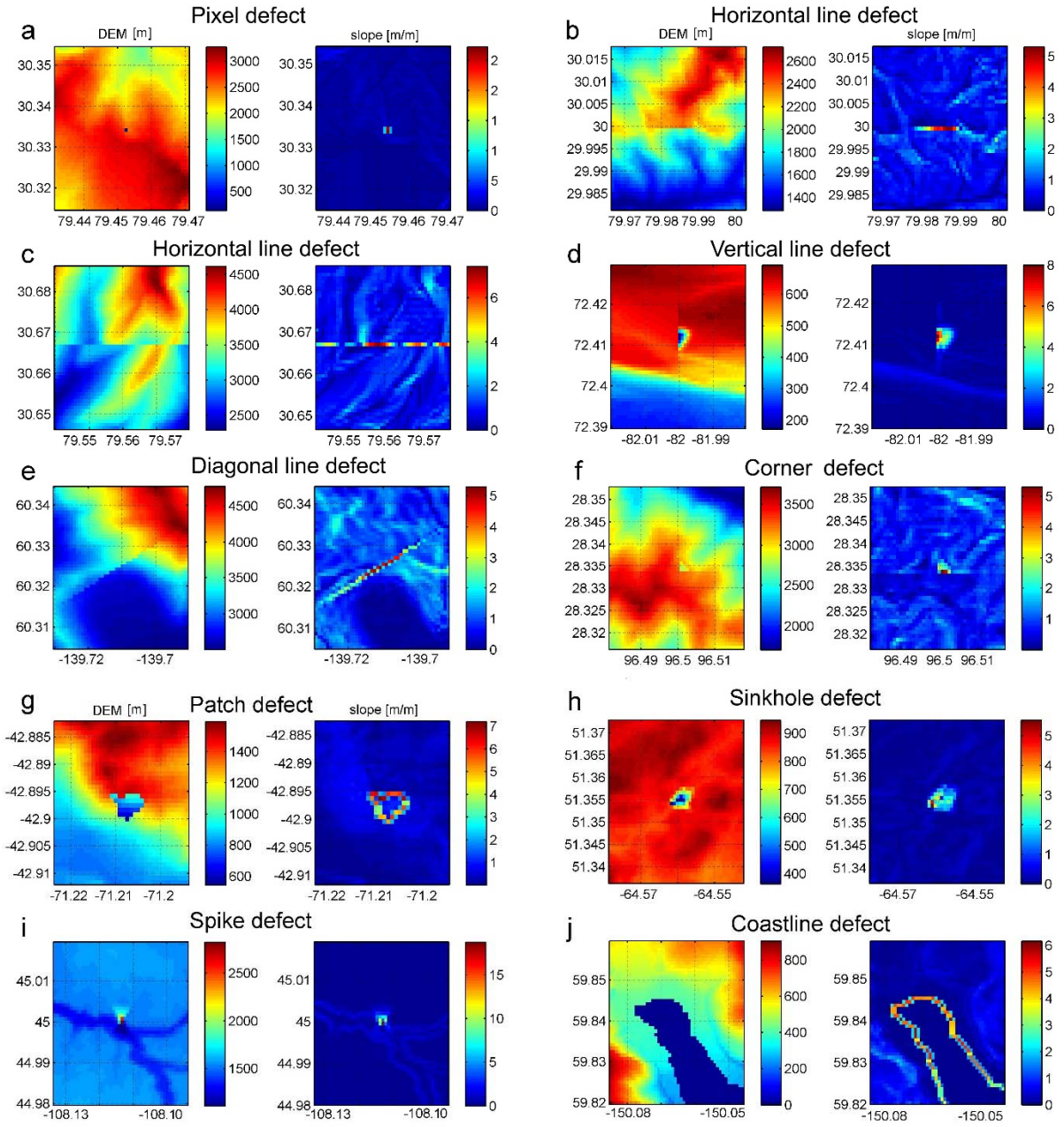


Fig. 3. Major types of artefacts present in digital elevation products such as a) pixel deflects, b-e) line defects in different directions, f) corner defects, g) patch defect, h) sinkhole defect, i: spike defect and j) coastline defect. Artefacts were localized in the MERIT-DEM. The image pairs show the DEM data (left, unit m) and the slopes (right, unit m/m). A detailed comparison with SRTM v2.1 suggests the void-filling (a, i, g) and the VFP-DEM auxillary data (b, c, f, h) as prime suspects for the artefacts.

In cases of artefacts (cf. Fig. 3), however, the affected terrain and slope maps may often exhibit

- unrealistic drainage patterns, e.g., one or more small or large-scale sinkholes with several 100 or 1000 m next to each other,
- unrealistic spikes, e.g., small-scale topographic peaks with several 100 or 1000 m amplitude away from mountain ridges where the presence of such features can be plausible,
- sharp steps along straight lines, often along φ – or λ –coordinate directions,
- unnaturally sharp steps along coast lines, e.g., of several 100 m elevation change over 90 m cells between land and sea level for a number of consecutive DEM cells,

so allows straightforward classification as artefact. In most cases, natural features (Fig. 2) and artefacts (Fig. 3) can easily be distinguished when their gradients are large. In ambiguous cases, digital imagery (e.g., from Google Maps) can be consulted and terrain textures inspected, or slopes computed from other, preferably independent DEM data to confirm or reject the presence of artefacts. If still doubtful, the feature will be classified in our work rather as natural feature than as artefact in order not to exaggerate the number of artefacts in the DEM models.

3. Data sets

To exemplify the global application of the MSA we have selected two post-processed 3 arc-second resolution SRTM releases, SRTM v4.1 by CGIAR-CSI (Jarvis et al. 2008) and the MERIT DEM (Yamazaki et al. 2017). Both data sets are briefly described below, and their choice is justified. In principle, any other global gridded DEM data set – from SRTM or other elevation mapping missions – can be chosen instead for a global MSA-based screening to assess the possible presence of artefacts.

3.1 SRTM v4.1

The Consortium for Geospatial Information (CGIAR-CSI) has used NASA's "finished-grade" SRTM v2.0 release as input data for the production of SRTM v4.1 (2008). The SRTM v2 model provided improved water body representation and coast lines, spikes and wells were removed and small void filling was applied via interpolation of surrounding elevations, while major voids were not filled (Slater et al. 2006, NASA 2015). CGIAR-CSI placed a main focus on void-filling with auxiliary data (DEMs from other sources) based on the procedures described in Reuter et al. (2007) and Jarvis et al. (2008). Among several auxiliary DEM data sources, SRTM v4.1 uses Viewfinder Panorama DEMs (http://viewfinderpanoramas.org) over Asia and South America available in or before ~2008. According to Jarvis et al. (2008), the following procedure was applied in the development of SRTM v4.1 to fill SRTM voids. First, contours or 3D points were produced from the SRTM v2 data. Second, elevations over the void regions were extracted from the auxiliary data. Third, for areas with high-resolution auxiliary data, the Hutchinson (1989) algorithm with drainage enforcement was applied, while over other areas the most appropriate interpolation technique (cf. Reuter et al 2007) was selected and applied. Fourth, the original SRTM was merged with the interpolated DEM data over the void areas, before the SRTM water body set was used as mask for clipping the merged DEMs along coastlines. The SRTM v4.1 data set has been released at resolution levels 3, 7.5 and 30 arc-seconds (http://data.cgiar-csi.org). It should be noted that the CGIAR-CSI SRTM version number 4.1, while post-dating NASA SRTM v2 release, pre-dates the official NASA SRTM v3 model from 2015.

Because of the measurement principle behind SRTM (radar interferometry), SRTM v4.1 is a mixture of a digital terrain model (representing the bare ground) and a digital surface model (representing the top of vegetation or buildings, were present), e.g., Hirt (2015).

From the literature, artefacts are indicated in v4.1: Hirt et al. (2014) have reported the presence of some unfilled or incorrectly filled SRTM voids and "*interpolation errors along the seams of 1-degree tiles over parts of Asia*" from analyses of gravity derived from 7.5 arc-second resolution SRTM v4.1 DEM data. A total of 37 locations with major artefacts was found in Hirt et al. (2014). Merryman Boncori (2016) has detected "*geolocation shifts and spatially correlated elevation differences*" in the SRTM v4.1 data set. The geolocation shifts, which can be interpreted as line-like artefacts along affected areas, can "*amount to height errors of tens or even hundreds of meters*" (Merryman Boncori 2016) in mountainous terrain. Yamazaki et al. (2017) noted "some *critical errors in the void filling*

procedures" in the SRTM v4.1 release, and gave one example of unrealistic interpolation over the Amazon basin (ibid, supplementary materials).

Motivated by these findings, and in light of the widespread application of the SRTM v4.1 data in the scientific community (evident from more than 1,100 Google Scholar citations as of 14 Sep 2017), the SRTM v4.1 DEM has been chosen for globally complete artefact screening. For the present study, the original SRTM v4.1 model was downloaded from http://data.cgiar-csi.org/srtm/tiles/GeoTIFF/ at 3-arc-second resolution in August 2017.

3.2 MERIT-DEM

The Multi-Error-Removed Improved-Terrain DEM (MERIT-DEM, short: MERIT) by the University of Tokyo (Yamazaki et al. 2017) is the most recent post-processed SRTM release. MERIT is based on ALOS/PRISM AW3D DEM (Tadono et al. 2015) North of 60° latitude and NASA's SRTM v2.1 (NASA 2015) South of 60° latitude. As auxiliary data source, DEM data collected and maintained by Viewfinder Panoramas (VFP-DEM) were used for void-filling both of SRTM v2.1 and AW3D. According to VFP (2017), their data is based on SRTM, ASTER GDEM, Russian and Nepal topographic maps and various other data sources over Asia, and SRTM, ASTER GDEM and local topographic maps over South America.

From Yamazaki (2017, pers. comm.), SRTM v2.1 information was generally honoured in MERIT, but sometimes merged with VFP-DEM, as follows. When SRTM and VFP agreed within 20 m, SRTM v2.1 was used, while in case of a disagreement of 100 m or more, VFP was used to define MERIT elevations. For elevation differences ranging between 20 and 100 m, MERIT elevations rely on a linear merge between SRTM v2.1 and VFP. Over SRTM v2.1 void areas, VFP elevations were smoothly merged with SRTM v2.1 using an average matching method that principally avoids "*elevation jumps along the gap boundaries*" (Yamazaki et al. 2017, supplementary materials, p5). The average matching method was applied both for inserting VFP-DEMs both into AW3D and SRTM v2.1 void areas (Yamazaki 2017, pers. comm.). From the MERIT data source mask (cf. Yamazaki et al. 2017, Fig. S1a ibid), about 0.5% of MERIT DEM cells (within the SRTM data area) rely mainly on VFP. North of 60° latitude, VFP was used as main source for an estimated ~30% of MERIT DEM cells.

For the development of MERIT, efforts were made to remove major error components from the DEMs, comprising locally varying absolute biases, stripe noises, speckle noise and tree height biases. Opposed to most other SRTM releases (such as SRTM v4.1), MERIT thus represents bare-ground elevations, so is technically a digital terrain model. This has been achieved by modelling and removing the tree height bias over vegetated areas (Yamazaki et al. 2017). Over built areas MERIT still contains a bias due to urban canopy, where it should be considered a mixture of a digital surface and terrain model. The removal of tree heights is a conceptual difference between MERIT and SRTM v4.1 that, however, is not critical for global artefact screening. This is because tree height biases, due to their relatively small amplitudes (about 20 m, cf. Yamazaki et al. 2017, Fig. 1 ibid) produce steps of below 1 m/m (at 90 m resolution) which are not targeted by our study.

The MERIT-DEM was selected as second SRTM-based model for global artefact screening because several error removal procedures were improved and applied by Yamazaki et al. (2017). As such, MERIT-DEM is probably an example for a modern "best-effort" reprocessed SRTM-based DEM. MERIT was downloaded in version V1.0.1 from http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/ in July 2017.

4. Global DEM screening – a case study

We have applied the MSA procedures described in Sect. 2 for a complete global inspection of the 3-arc second SRTM v4.1 (within -60° and 60° latitude) and MERIT-DEM (within -60° and +90° latitude).

4.1. Slopes in terms of 1 degree tiles

Fig. 4a shows the maximum slopes (Eq. 8) computed over 1° x 1° tiles for SRTM v4.1 and Fig. 4b for the MERIT-DEM, and selected statistics are reported in Table 1. The maximum slopes are shown as function of the geographical coordinates (latitude φ and longitude λ) of the tile centres. Both panels show that for the majority of tiles the maximum slopes are below ~2 m/m. Over these areas, the small maximum slopes indicate that significant step artefacts do not exist.

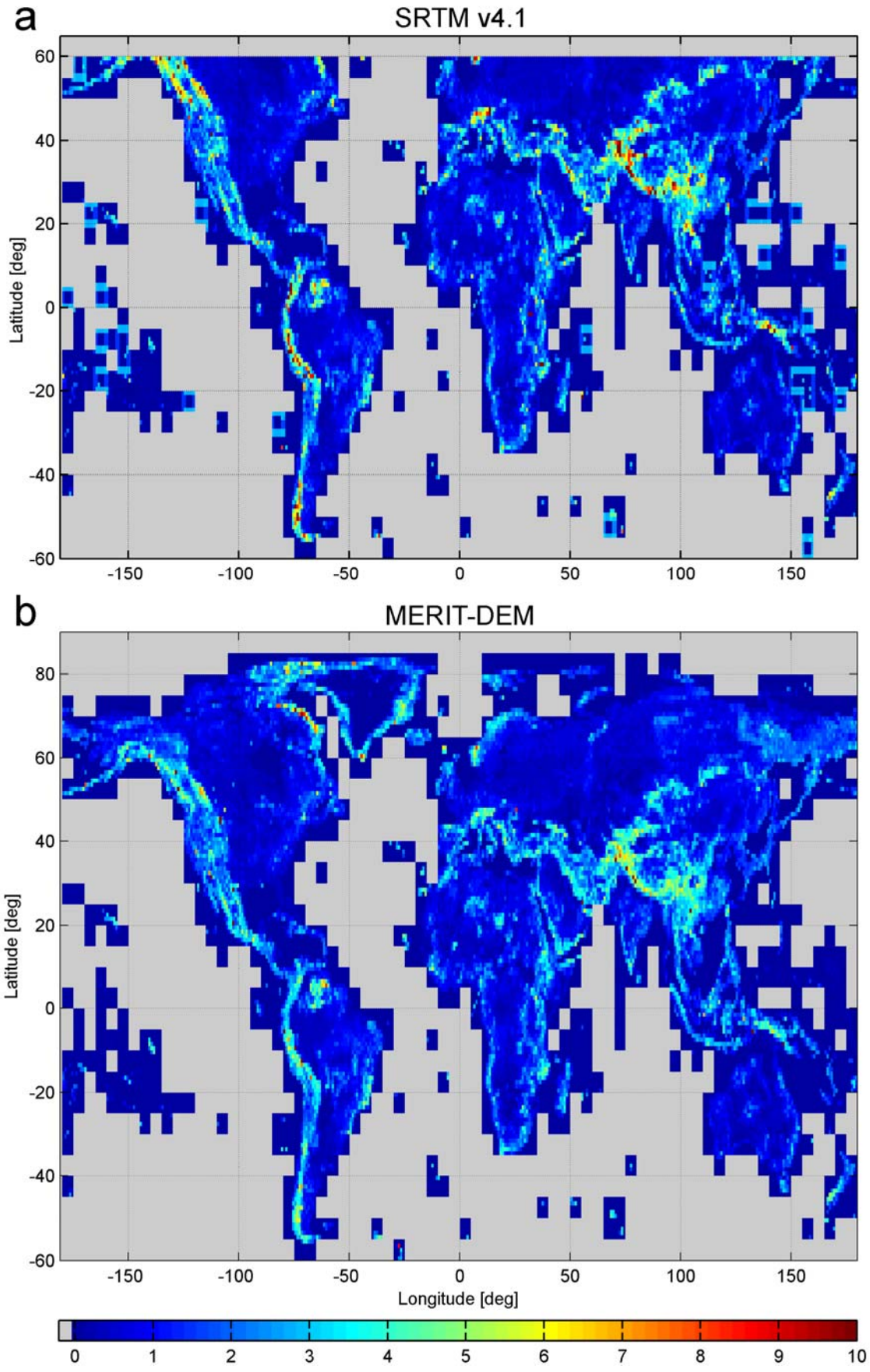
Larger maximum slopes naturally occur over Earth's major mountain ranges (Himalayas, Andes, Rocky Mountains), where (maximum) slopes in both models reach or exceed values of ~3-5 m/m (Fig. 4). A total of 589 (out of 14,889) tiles was found in the SRTM v4.1 model with slopes larger than 5 m/m, while in the MERIT model there are 232 (out of 19,223) tiles with slopes greater than 5 m/m within -60°$\leq\varphi\leq$90° and 177 tiles with slopes greater than 5 m/m within $\pm$60° latitude (Table 1). One-degree-tiles with very large maximum slopes (greater than 10 m/m) exist in the SRTM v4.1 model 79 times, but only 16 times in the MERIT model within $\pm$60° latitude. With few exceptions, tiles with large maximum slopes larger than 5 m/m are located where the topography is roughest (cf. Fig. 4). A cross-comparison between Fig. 4a and 4b shows that slopes larger than ~6 m/m not only occur more frequently in SRTM v4.1 over the major mountain ranges, but also over the European Alps, parts of Africa and New Zealand, while such large slopes are non-existent in the MERIT model over these regions.

Table 1. Statistics of tiles with large slopes detected in the SRTM v4.1 and MERIT DEM products

Category vs. model	**# occurrences in SRTMv4.1**	**# occurrences in MERIT DEM**	**# occurrences in MERIT DEM**
Latitude region	**-60° to +60°**	**-60° to +60°**	**-60° to +90°**
1° x 1° tiles with land topography, total	14899[a]	14339	19223
1° x 1° tiles with large slopes (≥ 5 m/m)	589	177	232
1° x 1° tiles with very large slopes (≥ 10 m/m)	79	16	18
0.1° x 0.1° tiles with large slopes (≥ 5 m/m)	1588	261	388
Classified as artefact	1341	108	123
Classified as natural feature	247	153	265
0.1° x 0.1° tiles with very large slopes (≥ 10 m/m)	83	18	21[b]
Classified as artefact	83	18	18
Classified as natural feature	0	0	3

[a]Note that SRTM V4.1 contains some erroneous 1° x 1° tiles with constant land heights over ocean areas devoid of islands

[b]Note that North of 60° latitude, at few locations exist on Baffin Island (Canada), where steep cliffs are aligned in East-West direction. Here, maximum slopes between ~10-14 m/m are found in the MERIT DEM, and these features are likely of natural character. It is important to note that, because meridians converge towards the poles, the metric DEM East-West cell dimension is much reduced in high latitudes. As a result, the slopes associated with Baffin-Island's topography tend to be larger than those of similar topography in mid-latitudes.



403

404 **Fig. 4.** Coarse screening of a) the SRTM v4.1 and b) MERIT elevation products. Shown are maximum slopes per 1° x 1° tile in
405 unit m/m.

406 **4.2 Slopes in terms of 0.1 degree tiles and classification**

407 For localisation, slopes were computed over all 0.1 degree sub-tiles within the "suspicious" 1 degree
408 tiles, and the φ and λ –coordinates of the maximum slopes were determined. To be able to detect
409 maximum slopes also at the tile boundaries, all sub tiles were extended by three DEM cells towards
410 either side, and all duplicate detections were eliminated as part of the screening exercise. Over some
411 0.1° sub-tiles, more than one location with large slopes could be detected, which either belonged to
412 the same topographic structure, or the same artefact, or indicated the presence of more than one
413 small-scale artefact (e.g., a local group of pixel defects). For our statistics (Table 1), these occurrences
414 were not further distinguished, and instead counted as one instance per 0.1° sub-tile. This statistical
415 discretization in terms of 0.1° tiles is a compromise in view of large-scale artefacts producing multiple
416 large slopes vs. small-scale groups of artefacts. With a finer discretisation (e.g., 0.02° sub-tiles), the
417 artefact count would likely further increase.

418 Table 1 reports the number of 0.1° sub-tiles with large maximum slopes, and Fig. 5 shows their
419 geographical distribution in both models. For the SRTM v4.1 model, a total of 1,588 occurrences was
420 found with slopes exceeding 5 m/m, out of which 83 sub-tiles even exceed a value of 10 m/m. This
421 compares to the 388 instances in the complete MERIT model with slopes larger than 5 m/m and 21
422 sub-tiles in excess of 10 m/m. In approximation, the SRTM v4.1 model contains 4 times more 0.1°
423 sub-tiles with large slopes than the MERIT model.

424 One part of these maximum slope occurrences reflects naturally large slopes of the terrain, while the
425 other part is caused by artefacts in the DEM model. In order to discriminate between both origins, all
426 DEM and slope data were plotted as color-coded maps (as in Fig. 2 and Fig. 3), and inspected visually.
427 In most cases, the maximum slopes could unambiguously classified either into a) natural terrain
428 features or b) artefacts based on the criterion of geomorphological im/plausibility (see Sect. 2.2.4).

429 Using the MERIT-DEM as an example, Fig. 2 shows selected natural or man-made terrain features
430 producing large slopes. It is seen that the slope maps exhibit smooth and not so sudden variations
431 along the cliffs or walls of the canyons or table mountains, suggesting they are geomorphologically
432 plausible. In contrast, Fig. 3 shows examples of various artefacts that may be encountered in the
433 MERIT DEM. The maximum slopes of what we classified as artefacts stand out from those of the
434 surrounding topography, either as straight or ringed lines or as spikes, so are geomorphologically
435 unrealistic.

436 **4.3 Results**

437 As key result of this work, Fig. 5 shows the geographical locations of all 0.1° tiles with slopes
438 exceeding 5 m/m in the SRTM v4.1 DEM (panel a) and MERIT-DEM (panel b), and their classification
439 as artefact or natural feature. Table 1 reports the number of occurrences of artefacts and natural
440 features in both models, and Fig. 6 shows the frequency distribution. To ensure comparability
441 between the areas behind the slope statistics of both DEMs, slopes from the MERIT model are
442 reported separately for ±60°, and -60° to +90° latitude.

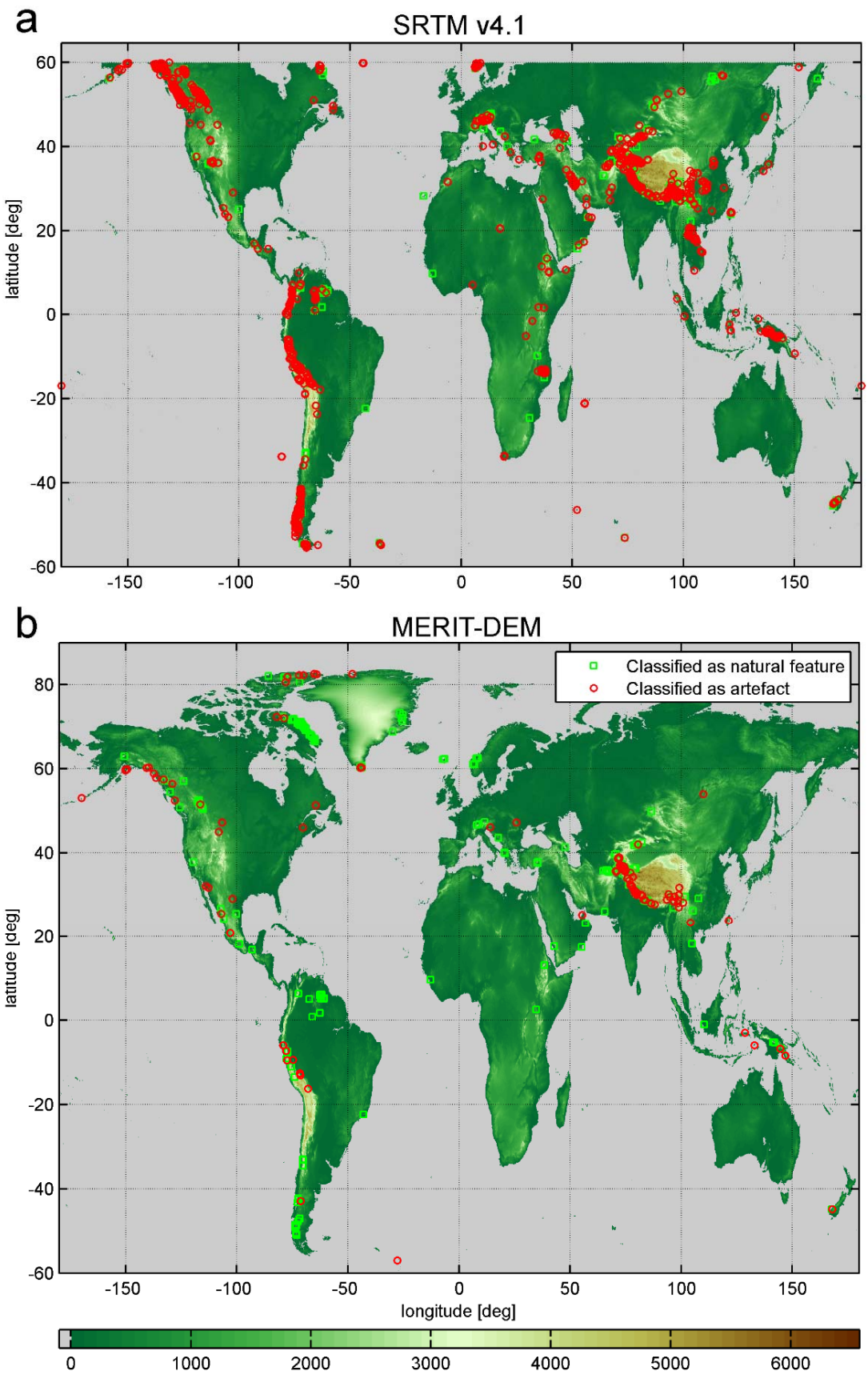


Fig. 5. Fine screening of the a) SRTM v4.1 and b) MERIT elevation products. Shown are all centres of 0.1° x 0.1° tiles with maximum slopes larger than 5 m/m, and their classification into artefacts (red signature) or natural feature (green signature).

From Table 1, SRTM v4.1 contains a somewhat larger number of large slopes classified as natural than MERIT (247 in SRTM v4.1 vs. 153 in MERIT within the latitude range ±60°). However, the number of artefacts differs substantially in both models: 1,341 occurrences in SRTM v4.1 vs. 108 in MERIT, suggesting that the SRTM v4.1 model contains about 12 times as many artefacts as the MERIT DEM.

Fig. 5 shows that artefacts tend to occur where the topography is roughest, that is, in the Rocky Mountains, along the Andes, and particularly along the Himalaya mountain range South and West to the Tibetan Plateau. A cross-comparison between Fig. 4a and 4b illustrates that the MERIT-DEM is generally much "cleaner" than the SRTM v4.1 data set, in the sense that artefacts are absent in MERIT where present in SRTM v4.1. Notwithstanding, also the MERIT DEM contains a number of 108 locations classified as artefacts; selected examples are given in Fig. 3.

Fig. 3 shows terrain and slope maps for various types of artefacts that may be encountered in DEM data sets. These are single pixel defects (Fig. 3a) and line defects along integer coordinate axes (Figs. 3b and 3d), line defects along random coordinate axes (Fig. 3c), diagonal line (Fig. 3e) and corner defects (Fig 3f). The line defects (Fig. 3b-3f) are obviously related to tiling during the DEM editing and hole filling. Fig. 3g shows a patch defect where two DEM data sources are in mismatch. Sinkholes (Fig. 3h) and spikes (Fig. 3i) can be the consequence when voids are filled through interpolation between an undetected pixel defect and surrounding elevations. Coastline defects (Fig. 3j) are the result of applying water masks that are inconsistent with the coastal topography.

To provide insight into the occurrence of natural vs. artificial slopes as function of the slope magnitude, Fig. 6 shows the frequency distribution of artefacts and natural features for SRTM v4.1 (panel a) and MERIT-DEM (panel b). Note that only slopes exceeding 5 m/m are shown, and slopes occurring in high latitudes in MERIT model (North of 60° latitude) are excluded from MERIT slope histogram (Fig 6b). From Fig. 6, none of the naturally occurring slopes exceeded a value of 9 m/m, which is why all very large slopes, in excess of 10 m/m, are confidently classified as artificial. The comparison of histograms shows that in any of the slope classes smaller than 15 m/m, the SRTM v4.1 model contains more or substantially more artefacts than the MERIT DEM, and this effect becomes particularly evident for classes between ~5 and ~8 m/m, each containing several 100 of SRTM v4.1 artefacts, while in none of the classes there are more than 30 MERIT artefacts.

Regarding the natural slopes in both DEMs, there is increasing frequency for smaller slopes visible, as would be expected. However, in any slope class considered, there are always more artefacts than natural features in SRTM v4.1 (Fig. 6a), and there are always more natural features in SRTM v4.1 than in MERIT (Fig. 6a vs. 6b). From a visual inspection of elevation and slope maps, the representation of extreme topographic features (e.g., walls or cliffs) appears to be somewhat smoother (i.e., smaller slopes) in MERIT than in SRTM v4.1. This observation may be a consequence of the Adaptive-Smoothing-Filtering used to smooth MERIT elevations at a pixel scale (see Yamazaki et al., 2017). In turn, this could thus explain the larger count of classified natural slopes in SRTM v4.1.

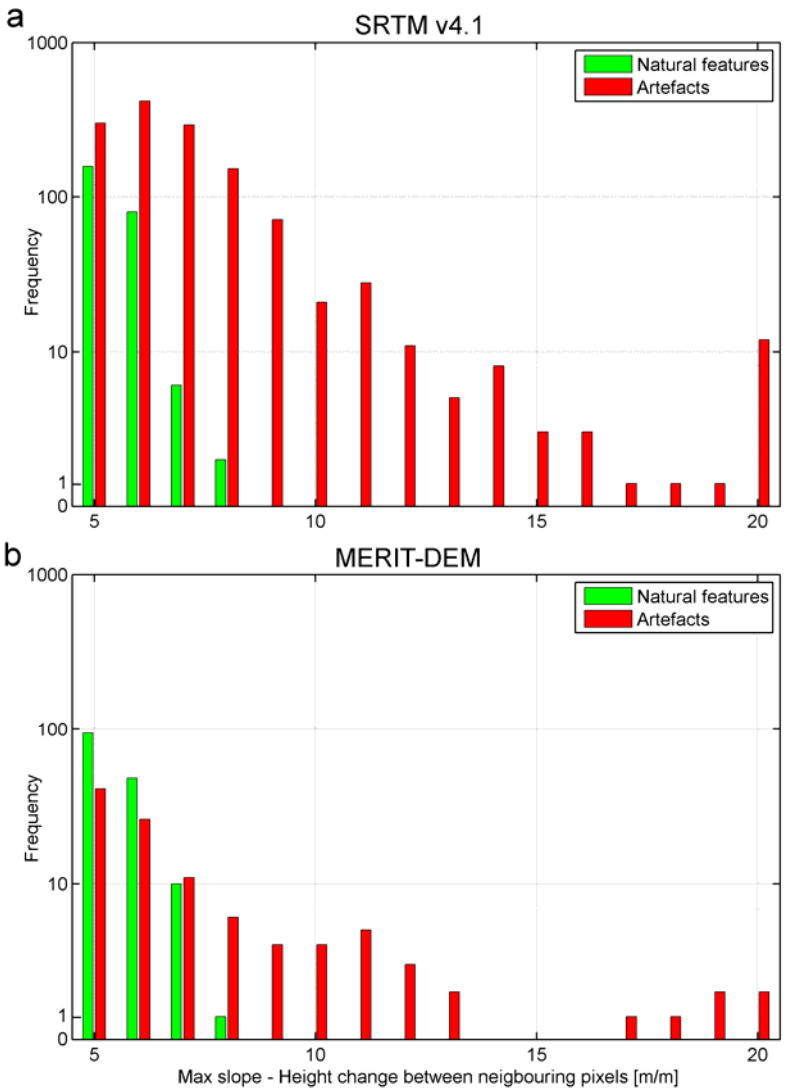


Fig. 6. Histogram of 0.1° x 0.1° tiles with maximum slopes larger than 5 m/m in a) the SRTM v4.1 and b) MERIT elevation products, classified into artefacts or natural feature. To ensure comparability, the regions in both models are ±60° latitude.

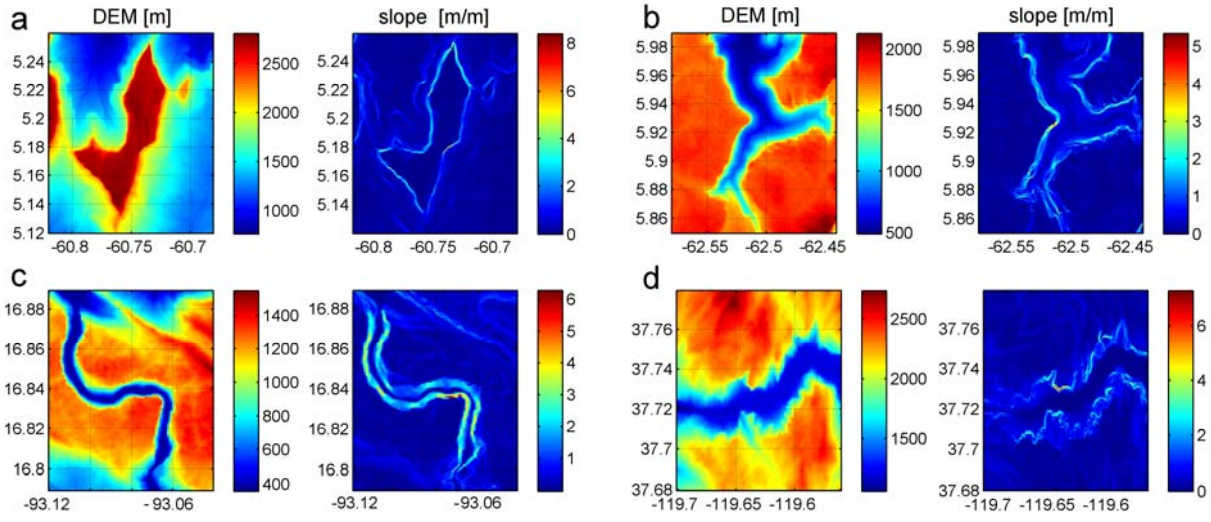


Fig. 7. Examples of natural topographic features with very large slopes discovered through the artefact screening, a) Monte Roraima Table Mountain (Venezuela/Brazil), b) Macizo del Auyán-Tepui (Venezuela), c) Cañón del Sumidero (Chiapas, Mexico), d) Yosemite Valley (USA). Data from MERIT, unit in m.

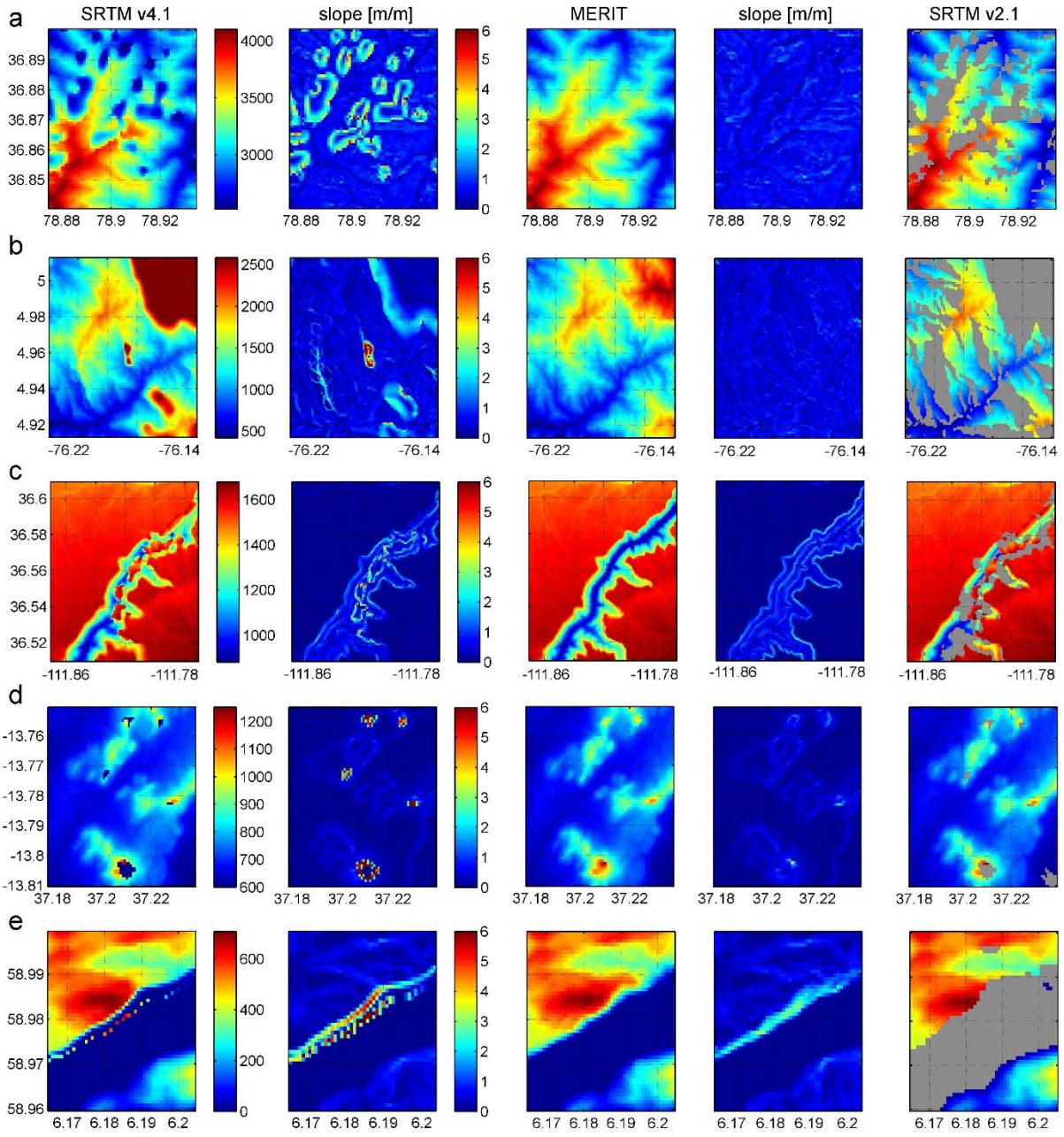


Fig. 8. Examples of major artefacts in the SRTMv4.1 (first and second column) and their absence in the MERIT product (third and forth column). For comparison purposes, the SRTM v2.1 non-void filled source data is shown in the fifth column. The image pairs show the DEM data (left) and the slopes (right). a) multiple sinkholes and misrepresentation of drainage structures, b) isolated spike, c) spikes and sinkholes inside a canyon, d) unfilled voids, e) coastline defects. While most of the visible artefacts are a consequence of the hole-filling, example c) shows one of the cases where SRTM v2.1 appears to misrepresent the terrain (inside the canyon).

Opposed to this, there are more naturally large slopes than artefacts smaller than ~6.5 m/m (Fig. 6b) in the MERIT model. From a detailed analysis of the MERIT-DEM, the largest natural slope within ±60° latitude reaches ~8.32 m/m, associated with Monte Roraima table mountain ($\varphi \approx 5.20°$, $\lambda \approx -60.74°$), see Fig. 7 for these and three other examples of extreme (natural) topographic slopes detected in the global screening. In SRTM v4.1, the largest slope not classified as artefact reaches ~8.25 m/m over the Catacombs Mountain (Rocky Mountains, $\varphi \approx 52.40°$, $\lambda \approx -117.74°$).

Fig. 8 finally shows five selected artefacts (out of 1,341) detected in the SRTM v4.1 model and the "clean" counterpart in the MERIT-DEM. Examples of artefacts that can be found in SRTM v4.1 include multiple sinkholes (Fig 8a), artificial spikes (Fig. 8b), unsound drainage patterns with spikes over a Canyon (Fig. 8c), unfilled voids (Fig 8d), and masking errors along coast lines (Fig. 8e). However, also most other artefact types shown in Fig. 3 for MERIT can be encountered in SRTM v4.1. From a cross-comparison with the (non-void filled) SRTM v2.1 release (Fig. 8, right column), the SRTM v4.1 artefacts are mostly correlated with the SRTM voids (see also Section 5.1).

5 Discussion

5.1 Model assessment

The numerical study has revealed a total of 1,341 sub-tiles of 0.1° x 0.1° size where artefacts with 5 m/m elevation change or more are present in the SRTM v4.1 data set, while 108 sub-tiles with artefacts were detected in the MERIT DEM, which is about ~8 % of the SRTM v4.1 artefact count. The vast majority of artefacts were found over highly-mountainous terrain of the Andes, Himalaya and Rocky Mountain chains. These areas have in common rough and steep topography where SRTM radar foreshortening, but also shadowing occurred, resulting in no-data areas (voids) in the SRTM data sets. One of the key differences between the two SRTM DEMs is the filling (cf. Sect. 3) of unobserved areas:

- SRTM v4.1 is based on NASA's SRTM v2.0 release that was essentially a non-void filled SRTM version, though small holes were filled through interpolation from surrounding interpolation (Slater et al. 2006, NASA 2015, p4). The majority of voids, however, was filled by CGIAR-CSI using interpolation procedures and auxiliary data described in Reuter et al. (2007) and Jarvis et al. (2008).
- MERIT-DEM is based on NASA's SRTM v2.1 release that - very similarly to SRTM v2.0 - contained voids. From USGS (2015, p1), all SRTM products up to version class 2 have voids "where SRTM was not successful in generating elevation measurements". For the generation of MERIT, VFP-DEMs were used as fill of these voids (Yamazaki et al. 2017, p5845).

From the references, we infer that the original SRTM v2 products behind CGIAR-CSI's SRTM v4.1 (relying on SRTM v2.0) and behind MERIT-DEM (relying on SRTM v2.1) are very similar, if not essentially the same. This is also corroborated by the statement that SRTM v2.1 "corrects some minor errors found in the original 3 arc-second version 2.0 product" (NASA 2015, p3). As such, the key differences between SRTM v4.1 and MERIT-DEM are a) different auxiliary data sets and b) different procedures applied for filling the voids with these auxiliary data. In order to further investigate the artefact issue, the non-void filled SRTM v2.1 data (available at https://dds.cr.usgs.gov/srtm/version2_1/) was inspected at all 1341 SRTM v4.1 and 108 MERIT artefact locations within ±60° latitude (Table 1):

- From 1341 artefacts in SRTM v4.1, 1337 instances (that is, ~99.7 %) were found to correlate with voids in SRTM v2.1 and thus with the void filling procedures or fillings. More specifically, 1091 (~81 %) of the artefacts are located at the edges of the SRTM v2.1 voids, and 246 (~18 %) inside the SRTM 2.1 voids.
- From the 108 artefacts in MERIT within ±60° latitude, 104 (that is, ~96.3 %) were found to correlate with voids in SRTM v2.1 and thus with the void filling procedures or fillings. About

94% of MERIT artefacts are located at the edges of the SRTM voids, and ~2 % inside the SRTM2.1 voids.

Because artificially large slopes were found to be present in SRTM v4.1 particularly at void edges, the large number of artefacts is primarily a result of the hole-filling by CGIAR-CSI. Also, in case of MERIT, most of the artefacts detected are a consequence of the hole-filling.

Conversely, our analysis demonstrates that the SRTM v2.1 data set makes a widely negligible contribution to the artefacts reported in Table 1. Thus, the SRTM v2.1 source data itself is generally fairly clean as far as distinct artefacts are concerned. If it were not, our MSA screening had detected more anomalously large gradients away from SRTM voids, which it hasn't.

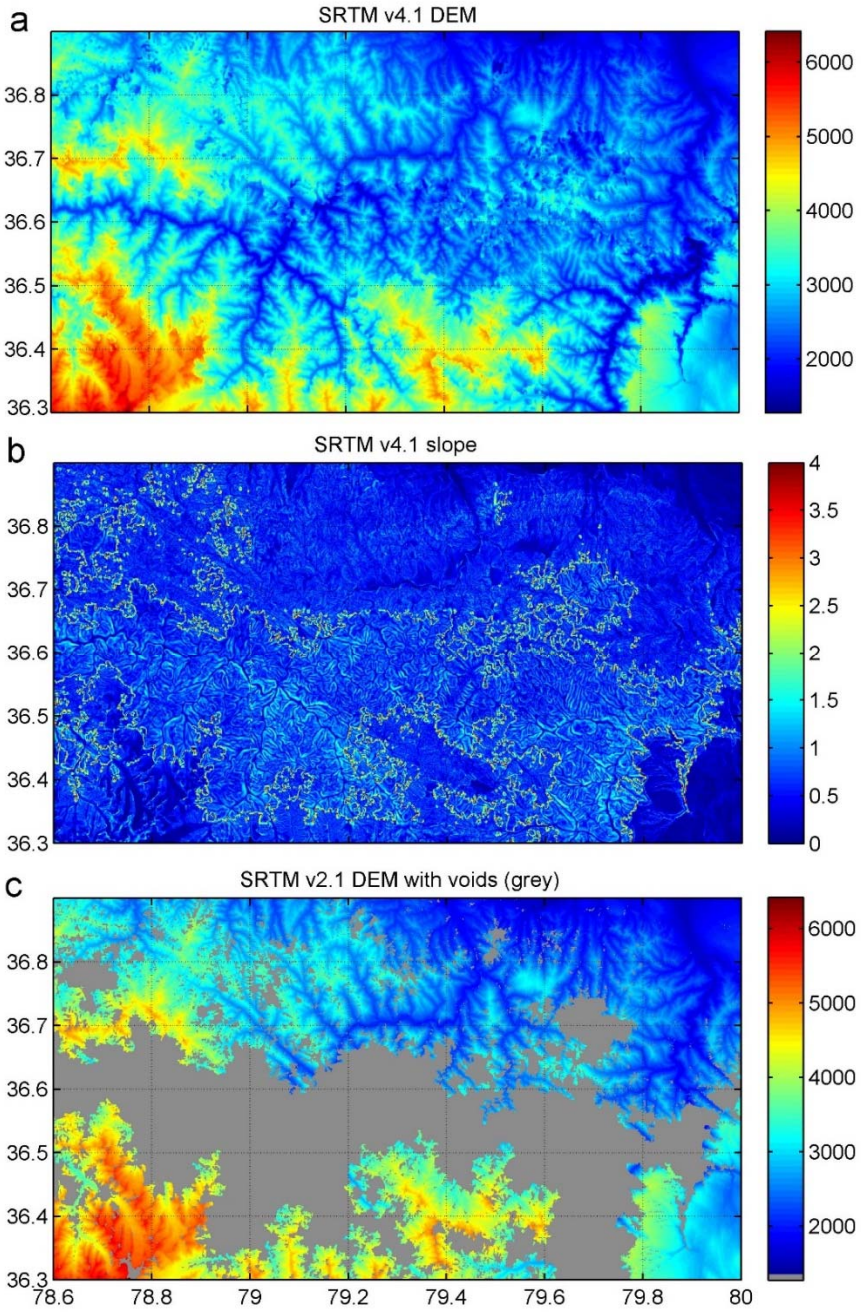
From the cross-comparison between MERIT and SRTM v4.1 artefacts at SRTM v2.1 voids it follows that the MERIT void-filling effort was generally more successful than that of SRTM v4.1. The different performance might also be related to improved auxiliary DEMs (e.g. VFP-DEM) available at a global scale in 2015, but only over some regions ~ten years ago, when SRTM v4.1 was developed. We finally note that several of the locations classified as natural features also spatially correlate with void areas, which is not surprising, given SRTM radar foreshortening and shadowing often occurred over steepest terrain.

Fig. 9 shows for a ~65 x ~125 km area, located West of the Tibetan Plateau, SRTM v4.1 DEM heights (panel a) where the derived slopes (panel b) unveil numerous locations with step-artefacts present. A detail inspection of Fig. 9 shows cliff-like errors in SRTM v4.1 everywhere the SRTM v2 and void fill data meet (panel c), suggesting offsets in the height component. These could be caused by the void fill data to be at another geodetic datum than SRTM (Crippen, pers. comm. 2017).

Notwithstanding the better performance in our global artefact screening, MERIT is by all means not free of artefacts. Various types of artefacts (Fig. 3), though in comparatively small numbers, have entered the model and can be occasionally encountered over the Himalaya mountain range, and rarely over the Andes and Rocky Mountains (cf. Fig. 5b). From Fig. 5b, the topography of Africa and Australia is represented in MERIT without spurious artefacts and Europe and New Zealand contain just a single artefact instance each. For all MERIT land topography North of 60° latitude, where several voids in ALOS/PRISM DEM data were filled with VFP-DEMs, a number of 15 artefacts were detected with the MSA. All in all, the number of artefacts contained in MERIT is rather small, and many of them should be removable through further editing, e.g., interpolation from valid DEM cells.

The focus of our study was on the detection and localisation of *distinct* artefacts in the global DEMs, with associated slopes reaching or exceeding an arbitrary threshold of 5 m/m. From Fig. 6, it is a safe prediction that an unknown number of artefacts below our threshold will be almost certainly present in both models. However, in view of the burgeoning number of natural terrain slopes for smaller thresholds (the global distribution is indicated in Fig. 6), it will become increasingly difficult to distinguish artefacts from the actual topography with the technique presented in this paper. A remedy will then be the cross-comparison with truly independent DEM data, e.g., from the TanDEM-X or ALOS/PRISM missions.

It is important to note that also in the SRTM itself larger errors may occasionally be present. These errors, primarily caused by interferometric unwrapping, have been detected in 160 one-degree-tiles (Bob Crippen, personal communication 2017). However, the amplitude of such errors is usually smaller than those of the most severe void-filling artefacts detected in this paper.



588

589 **Fig. 9.** Examples of regional contamination with artefacts, a) SRTM v4.1 elevations [m], b) SRTM v4.1 slopes [m/m], c) for
590 comparison purposes SRTM v2.1 [m] with voids. Along the SRTM v2.1 void edges, SRTM v4.1 slopes often reach values of
591 ~8 to 10 m/m.

592 **5.2 MSA-algorithm**

593 The Maximum Slope Approach was deliberately designed as a simple, yet effective technique for
594 detection of large artefacts, allowing its straightforward application for global screening of these and
595 other DEMs. It may be argued that to a certain degree the classification of maximum slopes based
596 on geomorphological plausibility (Sect. 2.2.4) is subjective (analyst-dependent) and not fully
597 automated (the detection of candidate locations for artefacts is done fully automatically, but the
598 classification depends on decisions made by the analyst).

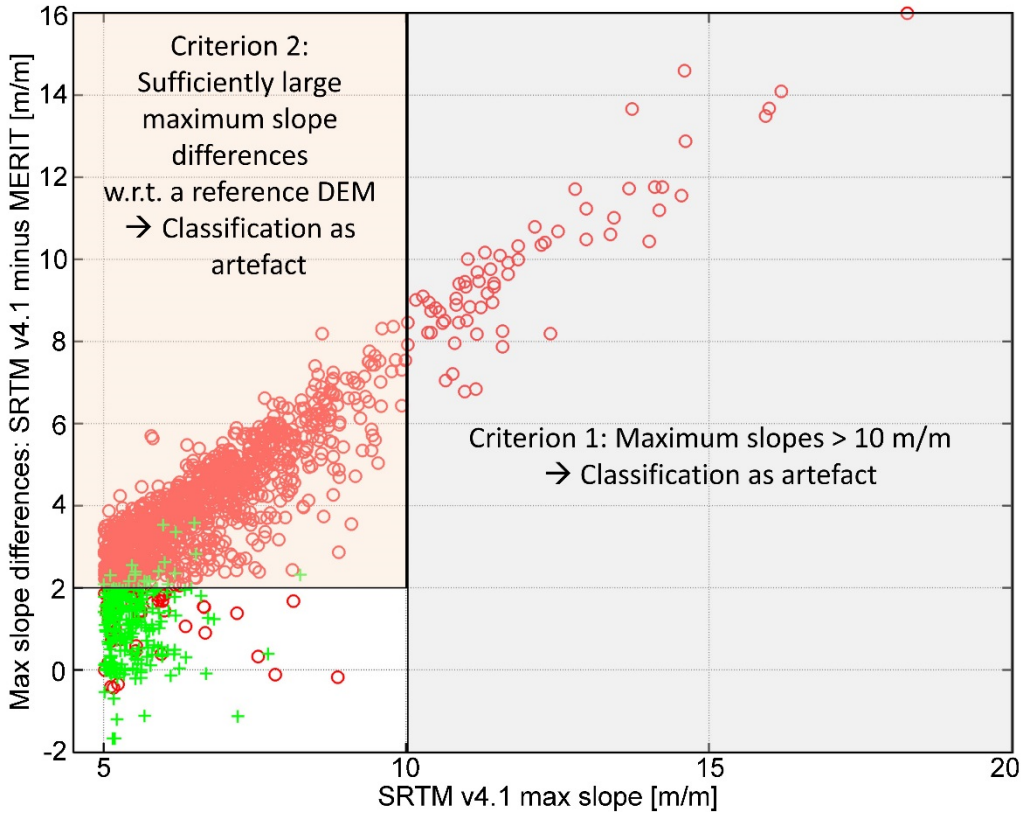


Fig. 10. Comparison of maximum slopes from the SRTM V4.1 model with the counterpart from the MERIT-DEM (that serves as a reference model here), and their classification as natural features (green) and artefacts (red). Shown are differences in maximum slopes (SRTM V4.1 max slope minus MERIT max slope) as function of the SRTM V4.1 slope value. The figure shows two domains (criteria) that are useful for automated classification. The domain of criterion #1 includes very large slopes that always reflect artefacts. Criterion #2 includes large differences in maximum slopes that separates most artefacts in one vs. natural features in the other model.

A remedy to both issues is the comparison of maximum slopes computed from different DEMs at all locations where artefacts are indicated. Fig. 10 shows the result of this experiment for the 1,588 large slopes detected in SRTM v4.1, where corresponding maximum slopes are computed from the MERIT-DEM (within a rectangular area of 0.1° length and width centred to the SRTM v4.1 maximum slope). It is seen that differences between maximum slopes from SRTM v4.1 and MERIT can be a useful tool for a widely automated classification of slopes into a) artefacts and b) natural features (Fig. 10). If maximum slopes from both DEMs are sufficiently distinct from each other (here a value of 2 m/m is used as threshold), artefacts in one of the models are indicated, while more similar values of maximum slopes around the same location are a sign for natural features. In few cases (red circles in the white area in Fig. 10), similar values of maximum slopes may reflect nearby artefacts present in both DEMs, that were manually re-classified here. Fig. 10 also shows that exceptionally large maximum slopes above 10 m/m can be automatically classified as artefact in 3 arc-sec DEMs. Notwithstanding it is seen that in the majority (>95 % of cases) the two criteria shown in Fig. 10 allow for a widely automated classification, and only some corrections are done through the analyst.

The focus of our work was on artefact detection and analysis, rather than on setting up a largely automated processing chain for DEM data cleaning that removes artefacts through application of suitable algorithms. For the subsequent data cleaning, different techniques might be needed to remove different kinds of artefacts. In all cases, the maximum slopes provide good a-priori information on the (φ ,λ) location of the artefact within the DEM. Then, single pixels defects (Fig. 3a)

may be easily extracted via outlier search techniques, and regions with areal defects, spikes and sinkholes (Figs. 3g - 3i) may be extracted through image segmentation (Zaitoun and Aqel 2016), and interpolated (repaired) with established algorithms such as the Delta-Surface-Fill method (Grohman et al. 2006). More difficult however will be the removal of step artefacts along coastlines (Fig 3j, 8e), or extended line artefacts such as in Fig. 3b and 3c requiring the use of auxiliary or original DEM data.

6 Conclusions and recommendations

This study has introduced the Maximum Slope Approach (MSA) as a method to detect and investigate artefacts in gridded digital elevation models (DEMs). The MSA relies on the fact that artefacts (e.g., pixel, line and coastline defects, spikes and sinkholes) often introduce sharp steps in the terrain surface, so can be detected by searching for unnaturally large slopes. While slopes larger than 10 m/m were shown to always reflect artefacts in 3 arc-second resolution DEM data within ±60° latitude, slopes smaller than ~9 m/m may be produced by natural terrain features, too. The threshold value, tested here for 3 arc-second DEMs, may be different for other resolution classes.

In order to discriminate between artefacts and natural features, elevation and slope maps were analysed for geomorphological plausibility on the one hand. Comparison against slopes from other DEM data sets (Fig. 10) on the other hand may increase the objectivity of the classification and degree of automation. The general outcome of this study – artefacts are still an issue in contemporary DEMs – shall be well reproducible, independent of the chosen method, their variation, or perception of geomorphological realism through the analyst.

In our numerical case study the MSA was exemplarily applied for globally complete screening of two SRTM-based global DEMs, the SRTM v4.1 and the MERIT-DEM releases. Based on 0.1°x 0.1° sub-divisions and a 5 m/m slope threshold, 1,341 artefacts were detected in SRTM v4.1 vs. a comparatively small number of 108 in MERIT. Importantly, the vast majority of detected artefacts in both models spatially correlate with areas not observed by SRTM, showing the artefacts to be mainly a problem of the void-filling, and not of the SRTM-measured elevations (Section 5.1). The strong contrast in artefact frequency (factor ~12.4) was attributed to different procedures and data sets used for the SRTM hole filling.

- The case study revealed that the SRTM v4.1 data set is contaminated by step artefacts over parts of the Himalayas (Fig. 9). Over this region, the use of SRTM v4.1 is not recommended, while caution should be exercised when using SRTM v4.1 over parts of the Andes and Rocky Mountains, but also other areas as indicated in Fig. 5. The same holds true for derived products that rely on the SRTM v4.1 data set such as gravity maps (e.g., Hirt et al. 2013, Hirt et al. 2014).
- It was shown that the MERIT model as an example for a recent “best-effort” error-reduced SRTM release contains artefacts, too. These were found primarily over the three major mountain ranges, but in significantly smaller numbers than the v4.1 data set (Fig. 5). In a relative sense, our study provide evidence of substantial improvements of the MERIT model over SRTM v4.1, as far as gross errors are concerned. Nonetheless, care should be exercised over the detected locations, and the artefacts should be removed in a next release of the MERIT model in order to improve the model quality.

We note that many contemporary DEM quality assessment studies rely on comparisons between DEM and independent ground control points (GCP, e.g., from ground surveying), which are usually

available locally or regionally. Such comparisons deliver vertical accuracy estimates for DEM heights, often confirming or refining accuracy estimates of DEM producers. However, due to their coverage and point density, GCPs are little or not suited to detect artefacts that may be "hidden" anywhere in the DEM, as such compromising the quality of the DEM in global applications. The observation that "*the global (average) error is small*", but "*local error values can be large, and also spatially correlated*" (Holmes et al. 2000, p154) is immediately applicable for DEM artefacts.

As a general conclusion, globally complete artefact screening of DEM data sets, that involves every single DEM cell, is recommended prior to any public release. Specifically, the screening techniques may be important in the context of the ongoing production of new global DEM data sets (e.g., NASADEM as a fully reprocessed SRTM with revised void filling, Buckley et al. 2015, Crippen et al. 2016; World-DEM from TanDEM-X, Wessel et al. 2016). However, MSA-based quality checks should also be considered by users before deploying DEM data in any application requiring geomorphometrically sound terrain information. The presence of pronounced artefacts in recent DEMs suggests that such global model screening techniques are not yet routinely applied in practice.

While the statements that "*No dataset is perfect. Each has its own limitations. Nevertheless, the temptation is to assume that digital datasets are perfect*." (Giles et al. 2010, p141) are true, artefact screening as proposed in this paper can certainly reduce the limitations of DEM data, and thus contributing to improving their quality. It is further recommended that artefact screening techniques be considered for DEM quality assurance procedures, as outlined e.g., in USGS (1997).

The Maximum Slope Approach is a simple technique capable of finding spurious artefacts in DEM data sets that disrupt the terrain surface and terrain slopes. Application of the MSA is not only limited to DEMs, but could be useful for other gridded data sets of geometric or physical quantities such as digital bathymetric models (e.g., Becker et al. 2009), global relief models (e.g., Hirt and Rexer 2015), planetary topography models (e.g., Jin 2015), gravity anomaly grids (Hirt et al. 2013) or magnetic anomaly grids (Meyer et al. 2017).

Acknowledgement

The author would like to thank the German National Research Foundation (DFG) for providing funding under grant agreement Hi 1760/01. Thanks are extended to Robert Crippen (NASA/JPL) and an anonymous reviewer for their thoughtful comments and for clarifying the versions of data sources behind the DEM data used in this study, and to Dai Yamazaki (U Tokyo) for providing a MERIT data source mask.

References

Abrams, M. (2016), ASTER Global DEM Version 3, and new ASTER Water Body Dataset. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 107-110.

Albani M, and B. Klinkenberg (2003), a Spatial Filter for the Removal of Striping Artifacts in Digital Elevation Models. Photogrammetric Engineering & Remote Sensing Vol. 69, No. 7, July 2003, pp. 755–765.

Arrell K., S. Wise, J. Wood and D. Donoghue (2008), Spectral filtering as a method of visualising and removing striped artefacts in digital elevation data. Earth Surface, Processes and Landforms, 33(6), pp. 943-961. doi: 10.1002/esp.1597.

Becker, J. J., Sandwell, D. T., Smith, W. H. F., Braud, J., Binder, B., Depner, J., et al. (2009). Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30_PLUS. Marine Geodesy, 32(4), 355-371.

Berry, P. A. M., R. G. Smith, and J. Benveniste (2010), ACE2: The New Global Digital Elevation Model, In: Gravity, Geoid and Earth Observation, 231–237, Springer, Berlin, doi:10.1007/978-3-642-10634-7_30.

Brown, D.G. and T.J. Bara (1994), Recognition and reduction of systematic error in elevation and derivative surfaces from 7 1/2-minute DEMs, Photogrammetric Engineering & Remote Sensing, 60(2), 189-194.

Buckley, S. (2015), NASADEM Overview and First Results: Shuttle Radar Topography Mission (SRTM) Reprocessing and Improvements, AGU Fall Meeting, Paper IN21A-1678, San Francisco.

Crippen, R., S. Buckley, P. Agram, E. Belz, E. Gurrola, S. Hensley, M. Kobrick, M. Lavalle, J. Martin, M. Neumann, Q. Nguyen, P. Rosen, J. Shimada, M. Simard, W. Tung (2016), NASADEM Global Elevation Model: Methods and Progress The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B4, 2016, XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.

Farr, T.G., Rosen P.A., Caro E., Crippen R., Duren R., Hensley S., Kobrick M., Paller M., Rodriguez E., Roth L., Seal D., Shaffer S., Shimada J., Umland J., Werner M., Oskin M., Burbank D. & Alsdorf D. (2007), The Shuttle Radar Topography Mission. Rev. Geophys. 45, RG2004, doi:10.1029/2005RG000183.

Feng C., Ma J.-W., Dai Q., Chen X. (2004), An Improved Method for Cloud Removal in ASTER Data Change Detection, International Geoscience and Remote Sensing Symposium (5), 3387 – 3389.

Giles J. R. A., S. H. Marsh, and B. Napier (2010), Dataset acquisition to support geoscience, In: Elevation Models for Geoscience (Ed. Fleming C. et al.), Geological Society, London, Special Publications, 345, 135–143, DOI: 10.1144/SP345.14

Grohman, G., G. Kroenung, and J. Strebeck (2006), Filling SRTM voids: The Delta Surface Fill method. Photogrammetric Engineering and Remote Sensing, v. 72, no. 3, p. 213-216.

Hirt, C. (2015), Digital Terrain Models. In: Encyclopedia of Geodesy (ed. E. Grafarend), Springer, Berlin, New York. ISBN 978-3-319-01868-3, doi: 10.1007/978-3-319-02370-0_31-1.

Hirt, C., and M. Rexer (2015), Earth2014: 1 arc-min shape, topography, bedrock and ice-sheet models - available as gridded data and degree-10,800 spherical harmonics, International Journal of Applied Earth Observation and Geoinformation 39, 103-112, doi:10.1016/j.jag.2015.03.001.

Hirt, C., S.J. Claessens, T. Fecher, M. Kuhn, R. Pail, M. Rexer (2013), New ultra-high resolution picture of Earth's gravity field, Geophysical Research Letters, Vol 40, doi: 10.1002/grl.50838.

Hirt C., M. Kuhn, S.J. Claessens, R. Pail, K. Seitz, T. Gruber (2014), Study of the Earth's short-scale gravity field using the ERTM2160 gravity model, Computers & Geosciences, 73, 71-80. doi:10.1016/j.cageo.2014.09.00.

Holmes, K.W., O.A. Chadwick, P.C. Kyriakidis (2000), Error in a USGS 30-meter digital elevation model and its impact on terrain modeling, Journal of Hydrology 233 (2000) 154-173.

Hutchinson, MF (1989), A new procedure for gridding elevation and stream line data with automatic removal of spurious pits, Journal of Hydrology, 106 (1989) 211-232.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E. (2008), Hole-filled SRTM for the globe Version 4. Available from the CGIAR-CGI SRTM 90m database: http://srtm.csi.cgiar.org.

Jin, S. (2015), Planetary Geodesy and Remote Sensing, CRC Press, Taylor and Francis.

Lecours, V., R. Devillers, E.N. Edinger, C.J. Brown & V.L. Lucieer (2017), Influence of artefacts in marine digital terrain models on habitat maps and species distribution models: a multiscale assessment. Remote Sensing in Ecology and Conservation, doi: 10.1002/rse2.49.

Lindsay, J.B. and I.F. Creed (2005), Removal of artifact depressions from digital elevation models: towards a minimum impact approach, Hydrol. Process. 19, 3113–3126, DOI: 10.1002/hyp.5835.

Lindsay, J.B.and I.F. Creed (2006), Distinguishing actual and artefact depressions in digital elevation data, Computers & Geosciences 32 (2006) 1192–1204.

Mark, D.M. (1975), Geomorphometric parameters: A review and evaluation, Geografiska Annaler. Series A, Physical Geography Vol. 57, No. 3/4, pp. 165-177.

Merryman Boncori, J. P. (2016), Caveats Concerning the Use of SRTM DEM Version 4.1 (CGIAR-CSI). Remote Sensing, 8(10), 793. doi:10.3390/rs8100793.

Meyer B., R. Saltus and A. Chulliat (2017), EMAG2: Earth Magnetic Anomaly Grid (2-arc-minute resolution) Version 3. National Centers for Environmental Information, NOAA. Model. doi:10.7289/V5H70CVX.

NASA (2015), The Shuttle Radar Topography Mission (SRTM) Collection User Guide, available at: https://lpdaac.usgs.gov/sites/default/files/public/measures/docs/NASA_SRTM_V3.pdf.

759 Oimoen, M.J. (2000), An effective filter for removal of production artifacts in u.s. geological survey 7.5-minute
760 digital elevation models, Paper presented at Presented at the Fourteenth International Conference on Applied
761 Geologic Remote Sensing, Las Vegas, Nevada, 6-8 November 2000.
762 Pike, R.J., I.S. Evans and T. Hengl (2009), Geomorphometry: A brief guide, In: Developments in Soil Science,
763 Volume 33, p 3-30, DOI: 10.1016/S0166-2481(08)00001-9.
764 Polidori L., J. Chorowicz, and R, Guillande, (1991), Description of terrain as a fractal surface, and application to
765 digital elevation model quality assessment, Photogrammetric Engineering and Remote Sensing, 57 (10), p. 1329-
766 1332, 1991.
767 Polidori L., M. El Hage and M. De Morisson Valeriano (2014), Digital elevation model validation with no ground
768 control: application to the topodata dem in brazil. Bol. Ciênc. Geod., sec. Artigos, Curitiba, v. 20, no 2, p.467-479,
769 abr-jun, 2014.
770 Reuter H., Nelson A. and Jarvis A. (2007), An evaluation of void filling interpolation methods for SRTM data.
771 International Journal of Geographic Information Science 21:9, 983–1008.
772 Reuter, H.I. T. Hengl, P. Gessler and P. Soille (2009), Preparation of DEMs for Geomorphometric Analysis, In:
773 Developments in Soil Science, p. 87-120, doi:10.1016/S0166-2481(08)00004-4.
774 Robinson, N., J. Regetz, and R. P. Guralnick (2014), EarthEnv-DEM90: A nearly-global, void-free, multi-scale
775 smoothed, 90m digital elevation model from fused ASTER and SRTM data, ISPRS J. Photogramm. Remote Sens.,
776 87, 57–67, doi:10.1016/j.isprsjprs.2013.11.002.
777 Skidmore, AW (2007), A comparison of techniques for calculating gradient and aspect from a gridded digital
778 elevation model, International Journal of Geographical Information Systems, 3:4, 323-334, DOI:
779 10.1080/02693798908941519.
780 Slater, J.A., G. Garvey, C. Johnston, J. Haase, B. Heady, G. Kroenung, and J. Little (2006), The SRTM Data
781 "Finishing" Process and Products, Photogrammetric Engineering & Remote Sensing 72(3), 237–247
782 Tachikawa, T, Masami Hato, Manabu Kaku and Akira Iwasaki (2011), Characteristics of ASTER GDEM Version 2,
783 IGARSS, July 2011. Available on http://www.jspacesystems.or.jp/ersdac/GDEM/
784 ver2Validation/IGARSS2011_Proceedings_GDEM2.pdf.
785 Tadono, T., Takaku, J., Tsutsui, K., Oda, F., & Nagai, H. (2015), Status of "ALOS World 3D (AW3D)" global DSM
786 generation. In 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3822-3825).
787 IEEE.
788 Torge W, Müller J (2012) Geodesy. 4th Edition. Berlin, New York: W. de Gruyter.
789 USGS (1997) U.S. Geological Survey: Standards For Digital Elevation Models, Part 3: QualityControl, available via
790 https://nationalmap.gov/standards/pdf/3DEM0897.pdf.
791 USGS (2015), Shuttle Radar Topography Mission (SRTM) Quick Guide, available at
792 https://lpdaac.usgs.gov/sites/default/files/public/measures/docs/SRTM%20Quick%20Guide.pdf
793 Villa Real, L.C., J.E.L. Aban and S.A. Husain (2013), A novel noise removal algorithm for vertical artifacts in digital
794 elevation models, Proceedings of the Asian Association on Remote Sensing, Bali, Indonesia, 2013.
795 VFP (2017), Viewfinder Panorama DEM collection, available via http://viewfinderpanoramas.org/dem3.html.
796 Warren, S.D., M.G. Hohmann, K. Auerswald and H. Mitasova (2004), An evaluation of methods to determine
797 slope using digital elevation data, CATENA Volume 58, Issue 3, 10 December 2004, Pages 215-233,
798 doi:10.1016/j.catena.2004.05.001
799 Wechsler, S.P. (2007), Uncertainties associated with digital elevation models for hydrologic applications: a
800 review, Hydrol. Earth Syst. Sci., 11, 1481–1500, 2007
801 Wessel, B., Breunig, M., Bachmann, M., Huber, M., Martone, M., Lachaise, M. et al. (2016), Concept and First
802 Example of TanDEM-X High-resolution DEM. In EUSAR 2016: 11th European Conference on Synthetic Aperture
803 Radar, Proceedings of (pp. 1-4). VDE VERLAG GmbH.
804 Yamazaki, D., D. Ikeshima, R. Tawatari, T. Yamaguchi, F. O'Loughlin, J.C. Neal, C.C. Sampson, S. Kanae, P.D. Bates
805 (2017), A high accuracy map of global terrain elevations, Geophysical Research Letters, Doi:
806 10.1002/2017GL072874.
807 Zaitoun N.M. and MJ. Aqel (2016), Survey on Image Segmentation Techniques, Procedia Computer Science 65,
808 797 – 806, doi:10.1016/j.procs.2015.09.027.