Towards ice thickness inversion: an evaluation of global DEMs in the glacierized Tibetan Plateau

Wenfeng Chen[1,2], Tandong Yao[1], Guoqing Zhang[1], Fei Li[1,2], Guoxiong Zheng[2,3], Yushan Zhou[1], Fenglin Xu[1,2]

[1] State Key Laboratory of Tibetan Plateau Earth System, Resources and Environment, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China

Correspondence: Wenfeng Chen (chenwf@itpcas.ac.cn) and Guoqing Zhang (guoqing.zhang@itpcas.ac.cn)

Abstract. Accurate estimates of regional ice thickness, which are generally produced by ice-thickness inversion models, are crucial for assessments of available freshwater resources and sea level rise. Digital elevation model (DEM) derived from surface topography of glaciers is a primary data source for such models. However, the scarce in situ measurements of glacier surface elevation limit the evaluation of DEM uncertainty. Hence the influence of DEM uncertainty on ice-thickness modelling remains unclear over the glacierized area of the Tibetan Plateau (TP). Here, we examine the performance of six widely used and mainly global-scale DEMs: AW3D30 (30 m), SRTM-GL1 (30 m), NASADEM (30 m), TanDEM-X (90 m), SRTM v4.1 (90 m) and MERIT (90 m) over the glacierized TP by comparing with ICESat-2 laser altimetry data while considering the effects of glacier dynamics, terrain factors, and DEM misregistration. The results reveal NASADEM as the best performer in vertical accuracy, with a small mean error (ME) of 0.9 m and a root mean squared error (RMSE) of 12.6 m, followed by AW3D30 (2.6 m ME and 11.3 m RMSE). TanDEM-X also performs well (0.1 m ME and 15.1 m RMSE), but suffers from serious errors and outliers on steep slopes. SRTM-based DEMs (SRTM-GL1, SRTM v4.1, and MERIT) (13.5−17.0 m RMSE) had an inferior performance to NASADEM. Errors in the six DEMs increase from the south-facing to the north-facing aspect and become larger with increasing slope. Misregistration of the six DEMs relative to ICESat-2 footprint in most glacier areas is small (less than one grid spacing). In a next step, the influence of six DEMs on four ice-thickness models: GlabTop2, Open Global Glacier Model (OGGM), Huss-Farinotti (HF), Ice Thickness Inversion Based on Velocity (ITIBOV) is intercompared. The results show that GlabTop2 is sensitive to the accuracy of both elevation and slope, while OGGM and HF are less sensitive to DEM quality and resolution, and ITIBOV is the most sensitive to slope accuracy. NASADEM is the best choice for ice-thickness estimates over the whole TP.

1 Introduction

The Tibetan Plateau (TP), which includes the Pamir, Hindu Kush, Karakoram, Himalaya, and Tibet regions, covers an area of ~3 million km² and has a mean elevation of more than 4000 m a.s.l. (Fig. 1). It accounts for more than 82% of the Earth's land surface area above 4000 m a.s.l. (Fielding et al., 1994), and is often referred to as the Third Pole of Earth or the Asian Water Tower (Yao et al., 2012) due to its high elevation and abundant water resources in the form of glaciers, snow, permafrost, lakes, and rivers. The TP has a glacierized area of ~8.3×10^4 km² (RGI Consortium, 2017) with an ice volume of ~6.2×10^3 km³ (Farinotti et al., 2019), mainly distributed in the Karakoram and Himalaya regions.

Ice thickness is a crucial parameter for assessing the contribution of glaciers to global sea level rise (Kraaijenbrink et al., 2017), quantifying regional water availability (Huss and Hock, 2018; Immerzeel et al., 2020), and evaluating cryosphere-related hazards (Linsbauer et al., 2016; Zheng et al., 2021). In the TP, owing to the lack of in-situ ice thickness measurements (Welty et al., 2020), regional glacier thickness is mainly estimated by ice-thickness inversion models (ITIMs) using open-access digital elevation models (DEMs) (Farinotti et al., 2009; Farinotti et al., 2019; Frey et al., 2014). The DEM is a fundamental part of most regional ITIMs (Farinotti et al., 2017), and is often used to determine center flow lines (Maussion et al., 2019), shear stress (Frey et al., 2014; Wu et al., 2020), apparent mass balance (Farinotti et al., 2009), and for ice-thickness interpolation (Huss and Farinotti, 2012). In addition to its use in ITIMs, the DEM has been an essential input for a wide range of TP glaciology studies, such as glacier inventory (Bhambri et al., 2011; Frey et al., 2012; Ke et al., 2016; Mölg et al., 2018), glacier mass change (Brun et al., 2017; Shean et al., 2020; Zhou et al., 2018), glacier related disasters (Allen et al., 2019; Kääb et al., 2018; Zhang et al., 2019) and projections of glacier or glacial lake evolution (Kaser et al., 2010; Kraaijenbrink et al., 2017; Zheng et al., 2021). The uncertainty in the DEMs can lead to different ITIM outcomes (Frey and Paul, 2012; Fujita et al., 2017; Furian et al., 2021; Kääb, 2005), especially for those ITIMs in which the DEM is a crucial input. For example, the sensitivity of the glacier bed topography (GlabTop) model to slope increases for shallower slopes (Paul and Linsbauer, 2012), and an overestimation of slope by ~10% would result in an underestimation of ice thickness of ~32% (Linsbauer et al., 2012). Localized elevation errors and data gaps could affect the estimated ice thickness by 5−25% using Huss-Farinotti model (Huss and Farinotti, 2012). Therefore, it is imperative to choose a suitable DEM source for regional glacier thickness modelling (Koldtoft et al., 2021). Farinotti et al. (2017 and 2021) intercompare the performance of several ITIMs and suggest that consideration of the uncertainty in the input data could improve the model output. However, to our knowledge, the uncertainty in different open access DEMs and their influence on various ITIM outputs over the TP have not been evaluated.

Currently, open-access DEMs covering the whole TP are mainly created by stereo mapping sensors such as ALOS AW3D30 (Tadono et al., 2015), C- or X-band interferometry synthetic aperture radar (InSAR) such as TanDEM-X, and SRTM-C based products such as NASADEM (Crippen et al., 2016). Shadows and the layover effect of InSAR technology (González and Fernández, 2011), along with the deficient orientation of photogrammetrically stereo images (Mukherjee et al., 2013) or low stereo-correlation (Hugonnet et al., 2021) propagated during DEM production, may introduce errors and voids. Filling these voids with other data could result in increased uncertainty (Liu et al., 2019). Additionally, the rugged terrain of glaciers and

the low contrast of snow cover can often lead to geometric distortion and missing data (Reuter et al., 2007; Takaku et al., 2020). Estimates of the accuracy of DEMs in different terrains and landforms, and for different vegetation coverage and land use have been conducted outside the TP using Global Navigation Satellite Systems (GNSS) measurements or high-resolution DEMs (González-Moradas and Viveen, 2020; Grohmann, 2018; Hawker et al., 2019; Uuemaa et al., 2020). The performance of specific DEMs varied in these studies, indicating that the local terrain and land cover influenced the DEM accuracy. In the TP, glaciers are distributed across different climatic zones and have a wide range of elevations with rugged terrain (Fielding et al., 1994; Thompson et al., 2018). GNSS measurements are not accessible for most glaciers, and publicly available high resolution DEMs are of lower quality due to its long temporal coverage (Shean, 2017). The assessment of DEM accuracy in specific regions with limited GNSS measurements or high-resolution DEMs is insufficient to determine the performance of global DEMs across the whole glacierized TP.

Liu et al. (2019) evaluated the performance of seven open-accessed DEMs over the TP with sparse the Ice, Cloud, and land Elevation Satellite (ICESat) altimetry data and suggested that AW3D30 has a high degree of accuracy. However, ICESat data with a footprint of 70 m (larger than the resolution of the estimated DEMs) could result in intra-pixel errors in steep slopes (Uuemaa et al., 2020). Besides, glacial regions were not considered in Liu et al. (2019), due to the variations of glaciers over time. Misregistration among DEMs, which may lead to evaluation bias (Han et al., 2021; Hugonnet et al., 2021; Van Niel et al., 2008), was also neglected. Bearing these issues in mind, and considering the limitations of optics sensors in rugged terrain and the glacier accumulation area (Chen et al., 2021), it is clear that a further assessment of the performance of AW3D30 over glacier area is required. Recently, TanDEM-X (released in 2017) and NASADEM (released in 2020) have been reported to have large improvements in accuracy relative to previous DEM products for various land-cover types (Wessel et al., 2018), floodplain sites (Hawker et al., 2019), slightly undulating terrain (Altunel, 2019), and mountain environments (Gdulová et al., 2020). Nonetheless, their performance over the rugged and glacierized TP remains unclear.

The purpose of this study is to evaluate the optimal DEM to use for regional ice thickness estimation over the TP. We first evaluated the performance of six widely used DEMs: AW3D30, SRTM-GL1, NASADEM, TanDEM-X, SRTM v4.1, and MERIT which are derived from different sensors and have different resolutions, against ICESat-2 data which has been proven to have high vertical accuracy and resolution (Brunt et al., 2019, 2021; Li et al., 2021), but with sparse tracks (Fig. 1). The elevation differences between these DEMs and the ICESat-2 are systematically analyzed concerning aspect, slope, elevation, and glacier zones. The influence on the accuracy assessment from the glacier elevation changes, terrain and misregistration among DEMs is then quantified. Finally, we compare the performance of ice thickness modelled by using the six DEMs against in situ measurements of ice thickness by ground penetrating radar (GPR).

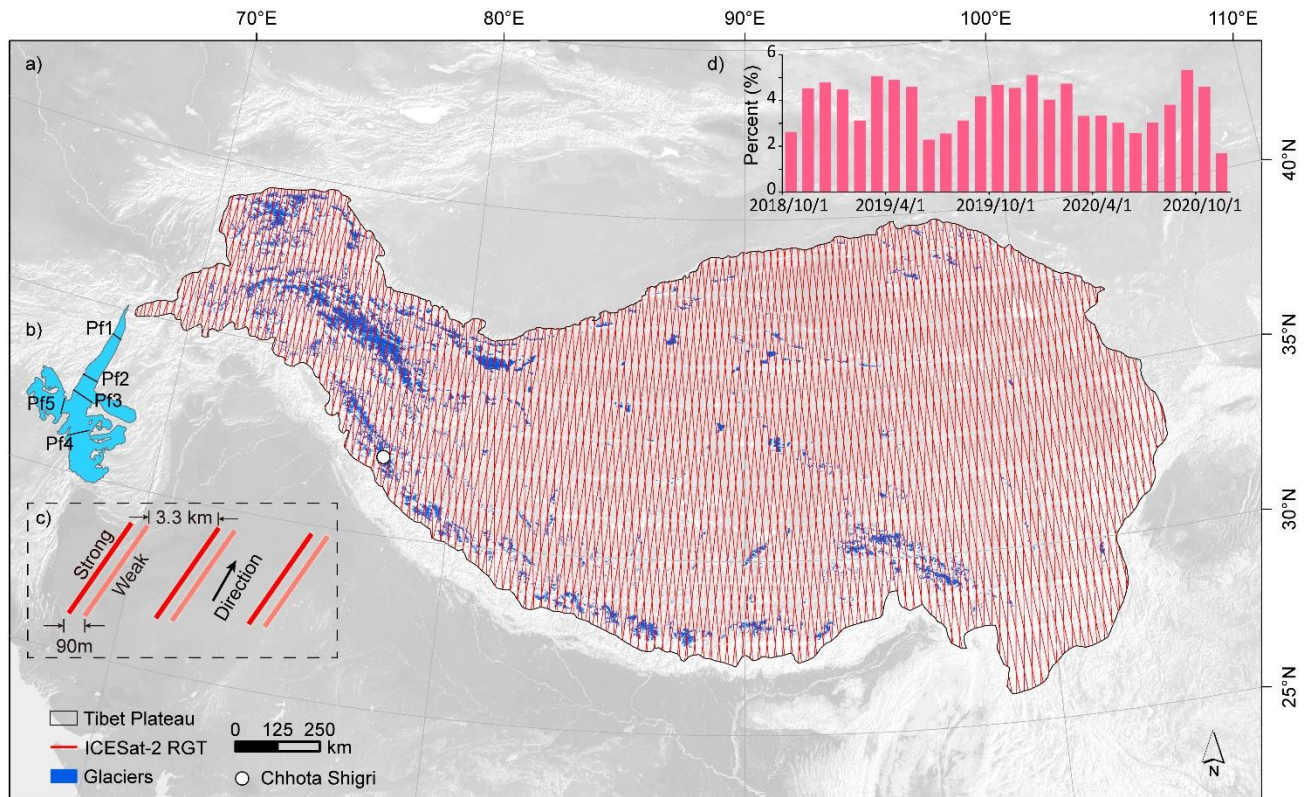



Figure 1. Location of the TP and its ICESat-2 reference ground tracks (RGTs). a) ICESat-2 tracks over the TP and glaciers. b) Location of ground penetrating radar (GPR) profiles over the Chhota Shigri Glacier which is used as an example (for location, see a). c) Relative location of six beams when Advanced Topographic Laser Altimeter System (ATLAS) has backward orientation. The distance between RGTs is 28.8 km. d) Percentage of ICESat-2 data in different months from October 2018 to November 2020. The boundary of the TP is derived from SRTM above 2500 m a.s.l. (Zhang et al., 2013).

2 Data and Methods

2.1 ICESat-2 elevation data

ICESat-2, a follow-on mission to the ICESat, was launched on 15 September 2018, with the goal of acquiring Earth's geolocated surface elevation is referenced to the WGS84 ellipsoid at the photon level. ICESat-2 ATLAS (Advanced Topographic Laser Altimeter System) emits a pulse every 0.7 m along the track covering a horizontal circular area with 0.5 m in vertical extent and ~17 m in diameter. This design diameter value varied due to the photo-counting lidar technology and potentially the atmospheric conditions (Magruder et al., 2020). ICESat-2's ATL03 and ATL06 products can both be used as elevation reference. The ATL03 product has a spacing of ~0.7 m and can provide more terrain details than the ATL06 product.

In this study, based on the resolution of global DEM and computational cost, we select the ICESat-2 Level-3A land-ice ATL06 product as an elevation reference. ATL06 heights are median-based heights derived from a linear-fit model over each segment corrected for first-photon bias and transmit-pulse shape. The segment has a length of 40 m centered on reference points at 20-

110 m intervals along the track. The ATL06 product has better than 5 cm height accuracy and better than 20 cm surface measurement precision in the Antarctic (Brunt et al., 2019,2020; Li et al., 2021) and Qilian Shan (Zhang et al., 2020). The product also contains land background points. The RGI6.0 glacier inventory (RGI Consortium, 2017) was used to extract points falling on glaciers (Fig. 2).

ICESat-2 ATL06 data covering the TP from October 2018 to November 2020 was downloaded from

115 https://earthdata.nasa.gov/ (Fig. 1). There are 2436 files containing about 100 GB of data in total. The fields: Location (latitude, longitude), surface elevation (h_li), elevation uncertainty (h_li_sigma) and quality (atl06_quality_summary) were used. By combining the quality field (atl06_quality_summary=0) (Smith et al., 2019) with the glacier inventory, a total of 3.5 million points out of 0.16 billion records over the TP were selected (Fig. 1). The slope, aspect and elevation value of the cell center of the DEMs were extracted for the ICESat-2 footprints.

120

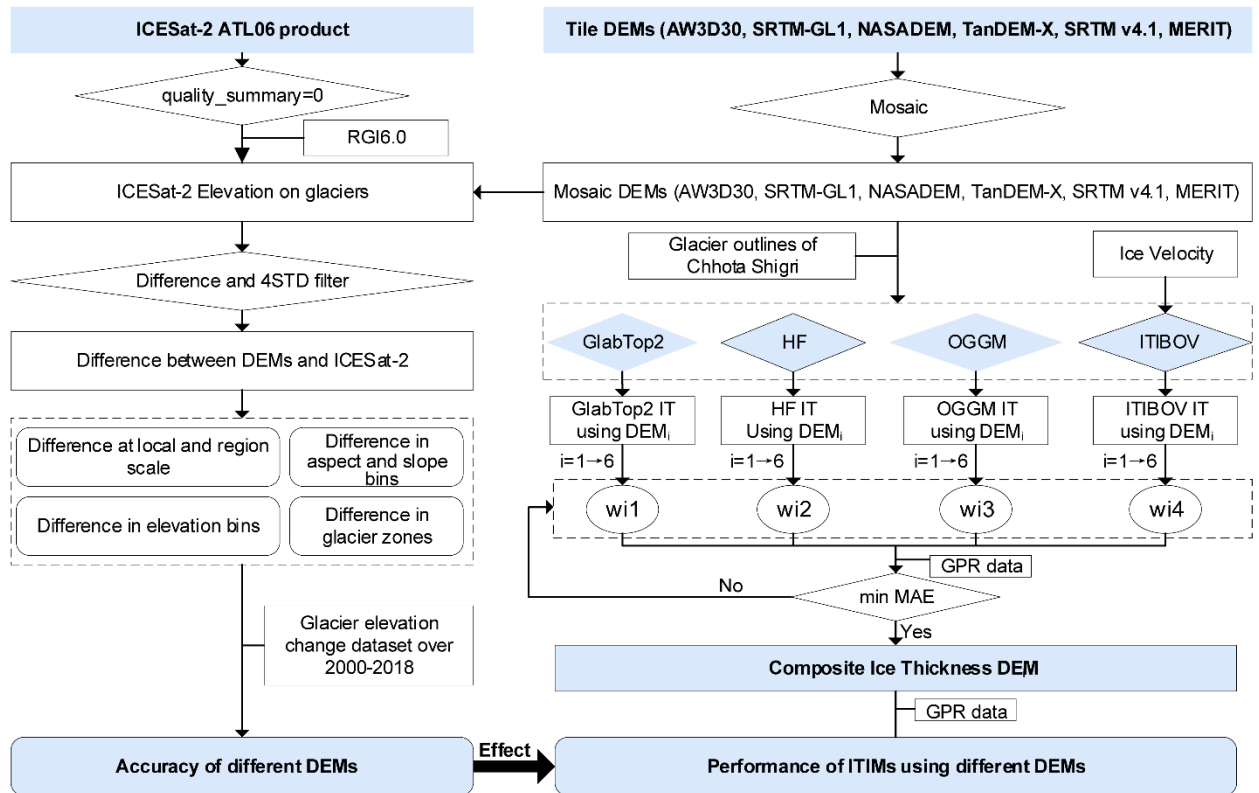



Figure 2. Flow chart showing the targets and methods used in this study including accuracy evaluation of DEMs and their effects on ice-thickness inversion models. The wi1, wi2, wi3 and wi4 denote the weight of each modelled ice thickness, i from 1 to 6 denotes six different DEMs, and the number 1−4 denotes the four ice thickness inversion models.

2.2 Global DEMs

Six global-scale DEMs were selected for evaluation of ITIM sensitivity, based on popularity, data source, resolution and sensor type (Table 1).

1) ALOS World 3D - 30 m (AW3D30) is acquired by the optics stereo sensor loaded on the Advanced Land Observing Satellite (ALOS) which operated from 2006 to 2011 with a horizontal resolution of 30 m. Approximate 10 % of global land area, mainly in tropical rainforest areas and the polar areas, has voids mostly due to cloud or snow/ice covers constatation in source imageries. Data gaps are filled with SRTM, ASTER GDEM v3, ArcticDEM v3, and TanDEM-X 90 (Takaku et al., 2020). After filling gaps, the accuracies in void-filled and void-free areas are nearly consistent (Takaku et al., 2020). Data was available at https://www.eorc.jaxa.jp/ALOS/en/aw3d30/index.htm after user registration.

2) TanDEM-X 90 m DEM (hereafter TanDEM-X) is a product derived from the first bistatic X band SAR mission of the world which took place from 2014 to 2016 (Bachmann et al., 2021). It is a pixel-reduced product of the global TanDEM-X DEM with a grid spacing of 0.4 arcseconds (12 m). The official reported absolute vertical and horizontal accuracy is better than 10 m at the 90% confidence level. It is noted that the current release is a non-edited version: areas with outliers, noise and voids remain. The original data was collected during different seasons and years, inducing errors due to (seasonal and long-term) accumulation/ablation on glaciers. Data was acquired from https://download.geoservice.dlr.de /TDM90/.

3) NASADEM is a new product released in 2020, which is derived by reprocessing the original SRTM signal data using updated interferometric unwrapping algorithms and auxiliary data, such as ICESat, to reduce voids and improve vertical accuracy (Crippen et al., 2016). Remnant voids are filled mainly by Global Digital Elevation Model (GDEM) v3 data. This data was downloaded from https://search.earthdata.nasa.gov/.

4) SRTM-GL1 (30 m) is an extensively used DEM in ITIMs. The first open-access ice-thickness database of global glaciers also adopted SRTM-GL1 as its DEM source (Farinotti et al., 2019). Voids were primarily filled by ASTER GDEM2.

5) SRTM v4.1, with a spatial resolution of 90 m, is produced by the method proposed by Reuter et al. (2007), including merging tiles, filling small holes iteratively and interpolating across the holes using a range of methods, according to the size of the hole, and the land type surrounding it (https://cgiarcsi.community/data/srtm-90m-digital-elevation-database-v4-1/). SRTM v4.1 was also used to compare against the performance of SRTM-GL1 to estimate the influence of resolution.

6) MERIT is also widely used with a spatial resolution of 90 m. It was developed by removing absolute bias, stripe noise, speckle noise, and tree height bias from the existing spaceborne DEMs (SRTM3 v2.1 and AW3D30 v1) using multiple

satellite datasets and filtering techniques (Yamazaki et al., 2017). Its accuracy was significantly improved, especially in
155 flat regions (Yamazaki et al., 2017). The overall accuracy is similar to TanDEM-X in floodplain sites (Hawker et al. 2019), but lower in short vegetation. The dataset was downloaded from http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/. SRTM v4.1 and MERIT were selected to compare with TanDEM-X, and simultaneously estimate the influence of DEM resolution on ITIMs.

Elevations of ICESat-2 data, NASADEM_SHHPv001 and TanDEM-X are based on WGS84 ellipsoid reference, and
160 elevations of the other four DEMs are based on EGM96 geoid (Table 1). The function of geoidheight provided by MATLAB was used to calculate geoid height to unify their references.

162 **Table 1.** Details of the selected DEMs.

Item	Version	Acquisition time	Reference	Release time	Resolution (m)	Sensor type	Source
AW3D30	v2.2	2006–2011	EGM96 geoid	Apr. 2019	30	Optical	Accessed 2021-10-19 from https://www.eorc.jaxa.jp/ALOS/en/aw3d30/index.htm
SRTM-GL1	v003	Feb. 2000	EGM96 geoid	Sep. 2015	30	SAR C- band	Accessed 2021-10-19 from https://earthexplorer.usgs.gov/
NASADEM	SHHPv001	Feb. 2000	WGS84 ellipsoid	Feb. 2020	30	SAR C- band	Accessed 2021-10-19 from https://search.earthdata.nasa.gov/
TanDEM-X	v1.0 (Non-edited version)	2010–2015	WGS84 ellipsoid	Feb. 2019	90	SAR X- band	Accessed 2021-10-19 from https://download.geoservice.dlr.de/TDM90/
SRTM	v4.1	Feb. 2000	EGM96 geoid	Nov. 2018	90	SAR C- band	Accessed 2021-10-19 from https://drive.google.com/drive/folders/0B_J08t5spvd8RWRmYmtFa2puZEE
MERIT	v1	Feb. 2000	EGM96 geoid	Oct. 2018	90	SAR C- band	Accessed 2021-10-19 from http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/

163

2.3 Ice thickness inversion methods

165 Tiles of six DEMs (AW3D30, TanDEM-X, NASADEM, SRTM-GL1, SRTM v4.1, and MERIT) were used to form a mosaic of terrain data covering the whole TP. Four ice-thickness inversion models (GlabTop2, HF, OGGM, ITBIOV) were used to estimate the glacier thickness. The Chhota Shigri Glacier located in western Himalaya with available GPR data (Fig.1) was selected as an example to evaluate the influence of DEM uncertainty on the ITIMs. The GPR data were measured based on a pulse radar system in October 2009 (Azam et al., 2017) and is available at Farinotti et al. (2021). Full details of the ITIMs are

170 given below.

GlabTop (Glacier bed topography) is based on the theory that glacier thickness is mainly determined by the slope of the terrain (Linsbauer et al., 2009, 2012; Paul and Linsbauer, 2012). It is assumed that the glacier is an ideal plastic fluid, with basal slip being ignored. Based on the empirical relationship between mean shear stress along the centrelines and the range of glacier elevation (Haeberli and Hoelzle, 1995) (Eq. 1), the actual basal shear stress τ can be determined.

$$\tau = 0.005 + 1.598\Delta H - 0.435\Delta H^2$$

$$\tau = 150 kPa, if\ \Delta H > 1600 \tag{1}$$

175 where ΔH is the elevation range of the glacier. The ice thickness h can then be determined from Eq. (2)

$$h = \frac{\tau}{f\rho g \sin\alpha} \tag{2}$$

where f is the shape factor, ρ is glacier density (850 ± 60 kg/m^3) (Huss, 2013), g is the acceleration due to gravity (9.8 m/s^2) and α is the slope. Glabtop2 is an automated method for calculating ice thickness, similar to GlabTop, but avoids digitizing the branch lines. For details refer to Frey et al. (2014).

The HF (Huss- Farinotti) model is based on the mass balance principle which relates the surface mass balance of the glacier

180 (b) to the ice flux and variation in the glacier thickness. Given the ice flux, ice thickness can be calculated according to Glen's ice flow law (Farinotti et al., 2009a; Huss and Farinotti, 2012).

$$h = \sqrt[n+2]{\frac{q(1-fsl)}{2A}\frac{n+2}{\left(f\rho g\sin\alpha\right)^n}} \tag{3}$$

where h is the mean elevation of band thickness, q is the ice flux, $fsl=0.8$ is the basal slip correction factor, $n=3$ is the exponent of flow law, ρ is glacier density (850 ± 60 kg/m^3), g is the acceleration due to gravity (9.8 m/s^2), $f=0.8$ is the valley shape factor (Cuffey and Paterson, 2010) and A is the Glen flow rate factor (3.24×10^{24} Pa^{-3} s^{-1}) (Cuffey and Paterson, 2010;

185 Gantayat et al., 2014).

This method defines a new variable $\tilde{b} = \dot{b} - \rho g\frac{\partial h}{\partial t}$, where $\tilde{b}$ is the apparent mass balance, $\dot{b}$ is the glacier surface mass balance, and $\frac{\partial h}{\partial t}$ is the glacier surface elevation change. $\tilde{b}$ is linearly related to the elevation. In the absence of mass balance data and

thickness change data on the surface of a glacier, the ice flux q can be obtained by estimating $\tilde{b}$, which is determined from Huss and Farinotti et al. (2012). Ice thickness in each elevation band can then be determined by substituting into Equation (3). Finally, h is extrapolated, in combination with the slope, to obtain the distributed ice thickness, according to the parameters in Huss and Farinotti (2012).

The Open Global Glacier Model (OGGM) is based on the same concept as HF, but has two main differences (Maussion et al., 2019). Firstly, the method described in Kienholz et al. (2014) is used to automatically obtain the middle streamlines and watershed division. Secondly, the apparent mass balance data are reconstructed from the local climatic dataset from variables such as precipitation and temperature.

The ice thickness inversion based on velocity (ITIBOV) model is inverted from the shallow ice approximation, it obtains the ice thickness by combining the surface velocity field with the Glen ice flow law (Gantayat et al., 2014; Glen, 1955):

$$h = \frac{n+1}{2A} \frac{(1-k)u_s}{(f \rho g \sin \alpha)^n} \tag{4}$$

where h is ice thickness, u_s is glacier surface velocity, and k is the contribution ratio of basal slip velocity relative to the u_s. We used the mean velocity over 1985−2019 from ITSLIVE dataset (Gardner, 2019) as the u_s input. We assumed that basal slip only occurred during the warm seasons, and k was calculated by dividing the annual glacier velocity by winter glacier velocity (Wu et al., 2020). Data from the Global Land Ice Velocity Extraction from Landsat 8 (GoLIVE) dataset with a date separation length of fewer than 96 days are used to estimate the monthly velocity (Fahnestock et al., 2016; Scambos, 2016), allowing the winter velocity (December, January and February) and annual mean velocity to be calculated. Basal factor k was calculated as 0.80 (Fig. S1). The shared parameters, such as creep factor, shape factor and basal creep factor are same in all four models.

An ensemble of the output from different models can improve the modelled thickness (Farinotti et al., 2017; Farinotti et al., 2021). Therefore, after calculating the ice thickness from four models using different DEMs, we calculated an ensemble ice thickness using the same DEM but with different models. First, the ensemble ice thickness was the sum of the four models with weights w1, w2, w3, and w4, respectively. The sum of four weights equals 1.70% of the GPR data are adopted as calibration data. 30% of the GPR results are adopted as validation data. Then, the four weights are iteratively changed to achieve the minimal mean absolute error between calibration data and model result. Finally, the mean absolute error (MAE) between ensemble ice thickness and validation data are calculated.

2.4 Accuracy assessment method

The error in the DEMs is considered to be the difference between the DEM elevation and the ICESat-2 measurements. To remove the influence of outliers, elevation differences greater than four standard deviations were removed. Mean error (ME), MAE, median error, root mean square error (RMSE), standard deviation (STD), and normalized median absolute deviation (NMAD) were calculated for the error analysis. NMAD and ME were used to assess the disturbance from extreme errors

(Höhle and Höhle, 2009; Gdulová et al., 2020). When calculating the ME, the positive and negative biases cancel each other, making the error smaller; therefore, the STD together with ME could be a complementary indicator for assessment.

$$ME = \frac{1}{n}\sum_{1}^{n}(H_{DEM} - H_{ICESat-2}) \tag{5}$$

$$MAE = \frac{1}{n}\sum_{1}^{n}abs(H_{ICESat-2} - H_{DEM}) \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}(H_{ICESat-2} - H_{DEM})^2} \tag{7}$$

$$STD = \sqrt{\frac{1}{n\text{-}1}\sum_{1}^{n}(H_{ICESat-2} - H_{DEM} - ME)^2} \tag{8}$$

$$NMAD = 1.4826 * median(abs(H_{ICESat-2} - H_{DEM})) \tag{9}$$

220   Glacier surface elevation change ranged between -21 and 17 m/yr over the TP during 2000−2018 (Shean et al., 2020). Therefore, the disparity of acquisition date between ICESat-2 and six DEMs (Table 1) could introduce large errors due to the changing glacier geometry. This applies in particular to TanDEM-X and AW3D30, which are collected in different months and years (Table 1). However, the other four DEMs are produced from NASA's Shuttle Radar Topography Mission during the 11-day mission in February 2000. We selected ICESat-2 data acquired in February 2019 and 2020. Then the glacier elevation
225   dynamic magnitude during February 2000 and February 2019/2020 (Shean et al., 2020), are subtracted from the selected ICESat-2 elevation. By comparing the elevation differences from the adjusted ICESat-2 and the four DEMs, we could partly estimate the impacts of glacier dynamic on accuracy assessment.

3 Results

3.1 Accuracy of DEMs

230   The four standard deviations (that is 4 sigma) was chosen to not only filter the differences between ICESat-2 and DEMs to exclude extreme outliers, but also keep most records in the further accuracy analysis. The ratio of excluded outliers relative to the record of each DEM is less than 1%. Overall, there is no irregular deviation existed among these DEMs and ICESat-2 elevation after filtering (Fig. 3). The ICESat-2 vs DEMs values are distributed tightly around the fit line with a slope coefficient of 1, with no obvious differences among the R^2. NASADEM and TanDEM-X performed the best in terms of intercept and fit
235   RMSE, with very little difference to the ICESat-2 data.

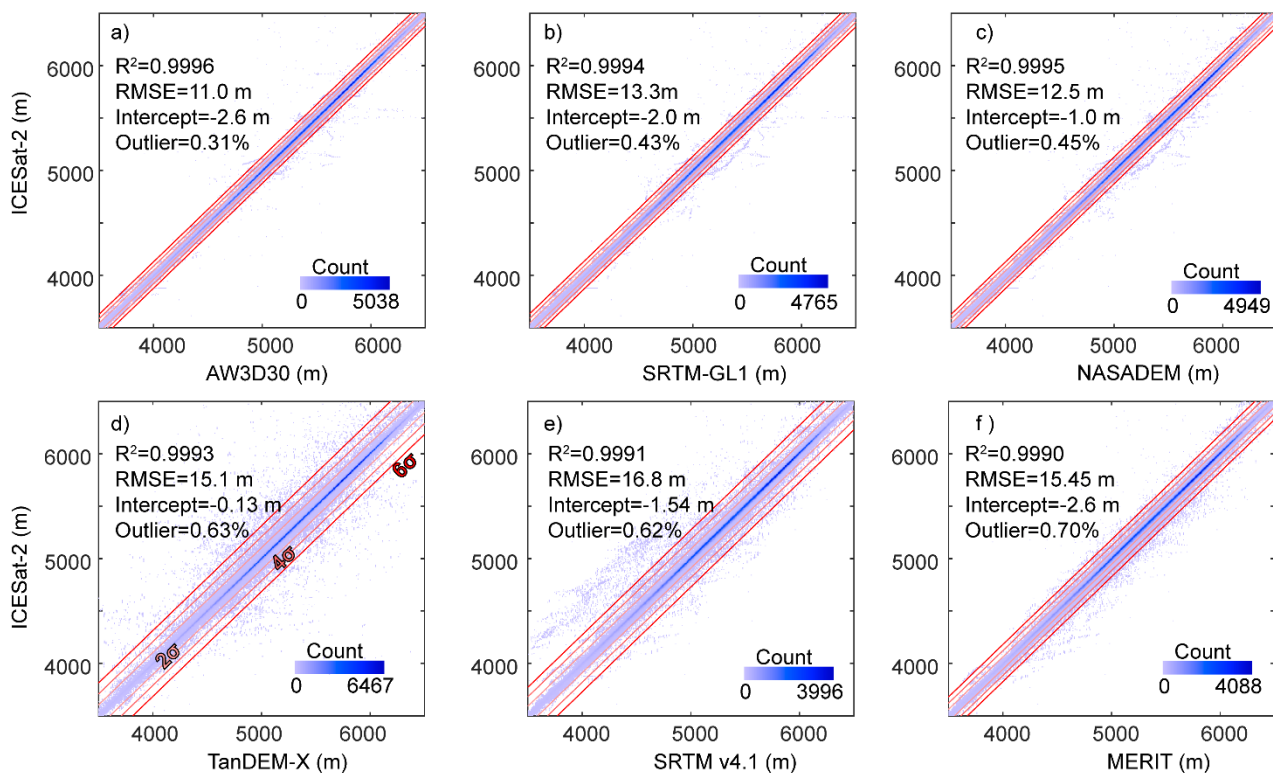



Figure 3. Differences between six DEMs and ICESat-2 elevation. a) AW3D30, b) SRTM-GL1, c) NASADEM, d) TanDEM-X, e) SRTM v4.1, and f) MERIT. The gradually lighter red lines in panel d denote the range within 6, 4 and 2 sigma of the mean. The text at the top left of each panel gives the fit results for data within 4 sigma of the mean. 'Outlier' denotes the proportion of outliers relative to the total records. 'R²', 'RMSE', and 'Intercept' are fit results when the slope coefficient is set to 1. The elevation range was cut to 3500−6500 m a.s.l., the range in which most elevations values are located, to show clearly the effect of using different multiples of the std from the mean.

Statistically, Median and ME did not differ much, which indicated that extreme values did not influence the ME much after the 4 std filter was applied (Fig. 4). STD was slightly larger than NMAD, especially for TanDEM-X, indicating larger discrepancies due to the DEM errors and noise (Höhle and Höhle, 2009). NASADEM performed better than the other two 30-m resolution DEMs in ME. AW3D30 behaved best in RMSE (11.3 m), MAE (8.2 m). SRTM-GL1 and NASADEM are both produced from the same original SAR data, but differ in RMSE (13.5 vs 12.6 m), MAE (10.0 vs 9.4 m), and ME (2.0 vs 0.9 m). The new algorithm and auxiliary data applied in NASADEM do indeed improve the absolute accuracy of the product over glacierized terrain. The quality of TanDEM-X was the best out of the 90-m resolution DEMs with the smallest RMSE (15.1 m), MAE (8.9 m), ME (0.1 m), and STD (15.1 m). SRTM v4.1 and MERIT are both error-reduced products from SRTM3 v2 (Reuter et al., 2007; Yamazaki et al., 2017), and they have similar ME (–1.5 vs 2.6 m) and RMSE (17.0 vs 15.6 m).

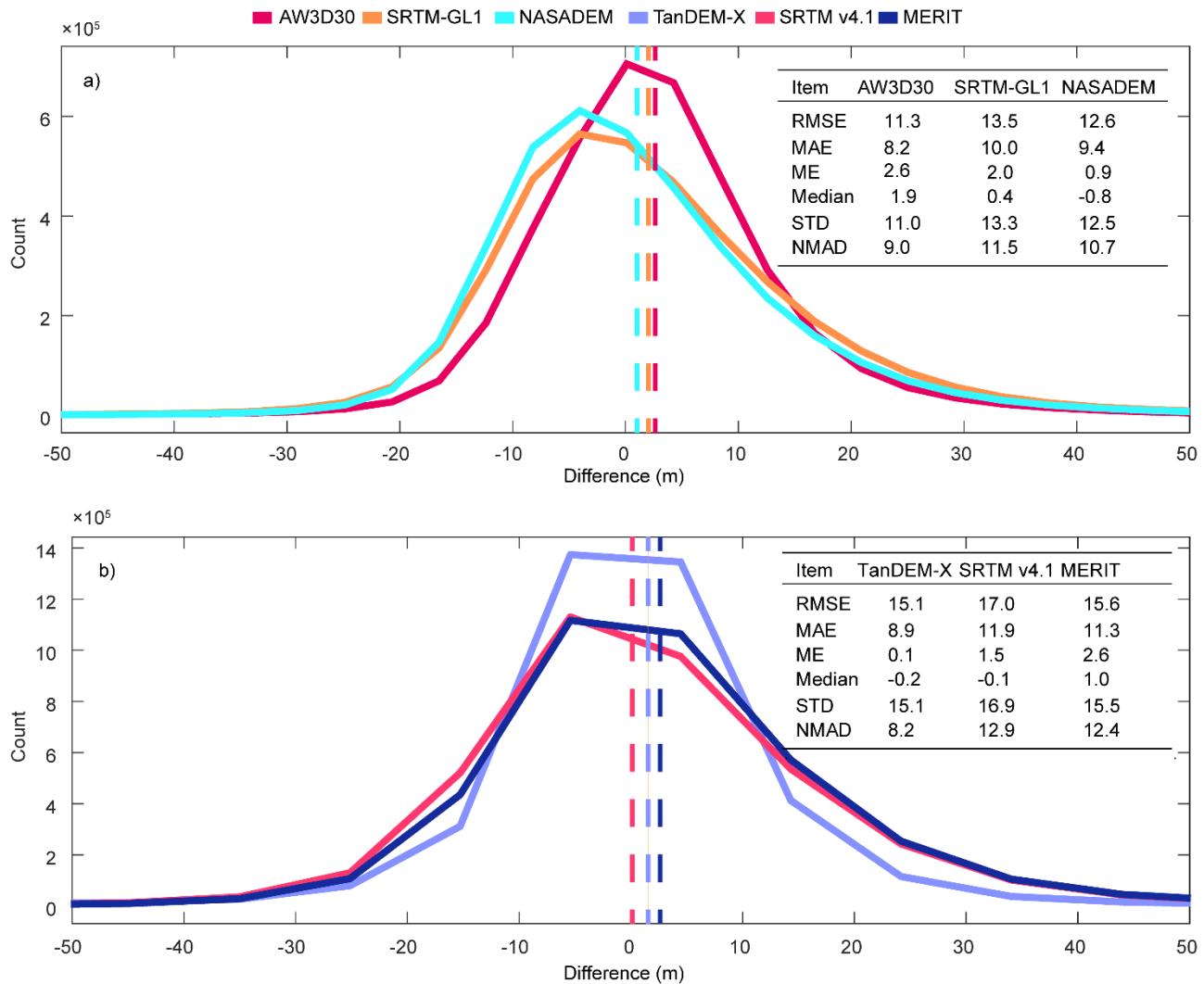



Figure 4. Overall difference (m) statistics between six DMEs and ICESat-2 elevation. a) 30-m-resolution DEMs, AW3D30, SRTM-GL1 and NASADEM. b) 90-m-resolution DEMs, TanDEM-X, SRTM v4.1 and MERIT. The vertical dashed lines denote the mean elevation difference of each DEM between ICESat-2. ME is Mean error, MAE is mean absolute error, Median is median error, RMSE is root mean square error, STD is standard deviation, and NMAD is normalized median absolute deviation.

Spatially, the ME in southeast Tibet is more positive than that in the Himalayas, and it is slightly negative in western Kunlun and the Karakoram mountains (Fig. 5a). It is worth noting that in the Himalayas and southeast Tibet, the ME of the other four DEMs is more positive than that of TanDEM-X and AW3D30. ME of TanDEM-X are mainly at ±5 m, but with some large values in several regions. SRTMGL-1, NASADEM, SRTM v4.1 and MERIT have nearly the same distribution of ME, and show negative ME values in the West Kunlun and Karakoram. ME of NASADEM is smaller than that of SRTM-GL1 in most

regions of TP, but is bigger in West Kunlun and Karakoram. Overall, the STD of 30-m resolution DEMs is much better than that of 90-m resolution DEMs (Fig. 5b). STD along the Hindu Kush-Himalaya and southeast Tibet was larger than that in other

265 regions. STD in southeast Tibet was relatively larger (>12 m). Specifically, the STD of AW3D30 and NASADEM was smallest and share similar spatial distribution features. The STD of NASADEM was improved over the most part of TP, compared with that of SRTM-GL1 (Fig. 5b). This indicates that some disturbances from noise and errors may exist in the SRTM-GL1, SRTM v4.1 and MERIT in the West Kunlun and Karakoram. TanDEM-X performs well in overall statistics (Fig. 4b) and ME (Fig.5a), but didn't show large improvements in STD and was even worse in some areas. The STD and ME of SRTM v4.1 and MERIT

270 have the same spatial distribution (Fig. 5b), and have similar overall STD (both ~15 m) and ME (both ~2 m) values (Fig. 4b).

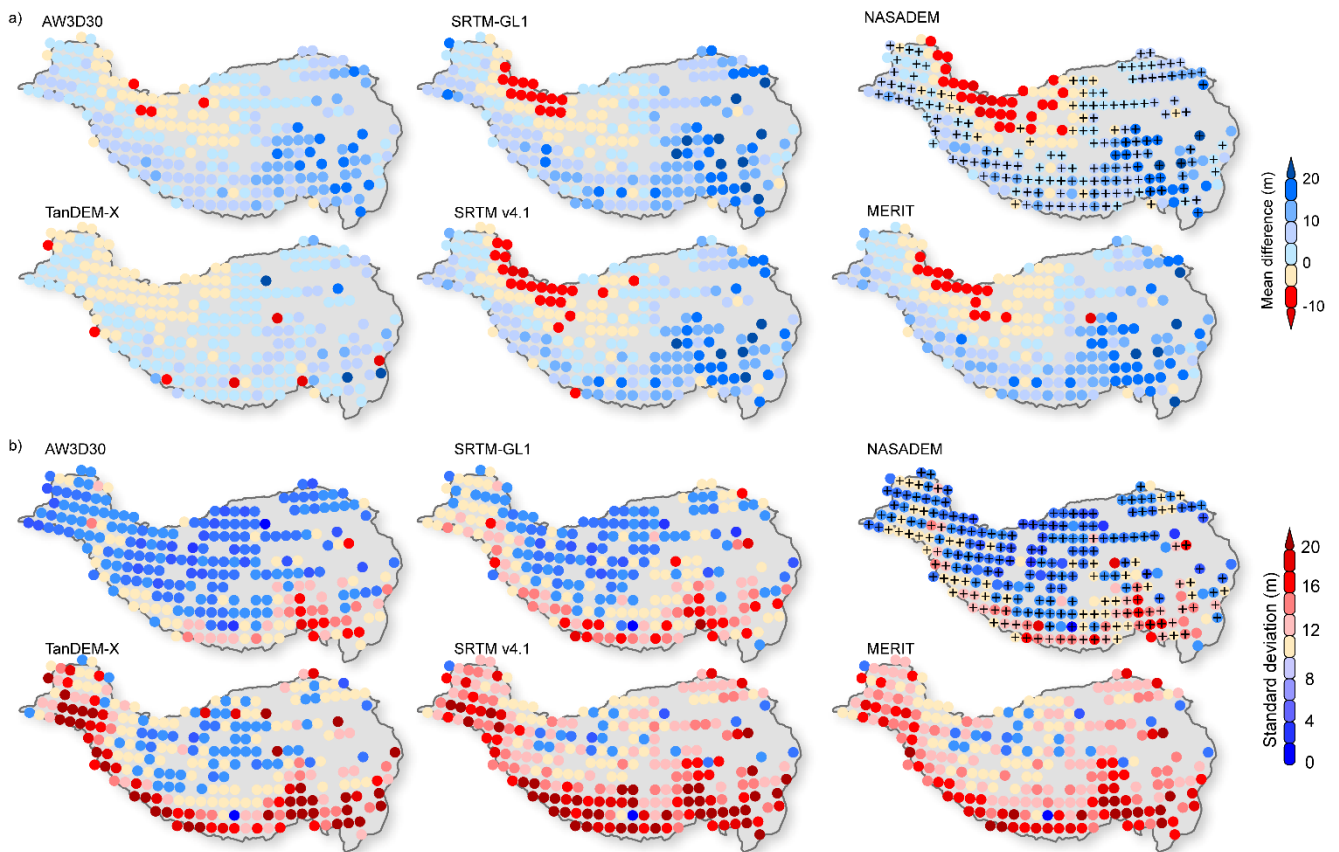



Figure 5. Aggregated spatial mean error (ME) (a) and standard deviation (STD) (b) between six DEMs and ICESat-2 elevation for 1°×1° cells across the TP. The cross symbol denotes that NASADEM performs better than SRTM-GL1 in ME or STD.

3.2 Differences between DEMs and ICESat-2 in aspect, slope and elevation

The influence of aspect is most apparent for SRTM-GL1, with a median value of about –5 m in the south aspect which increased in magnitude gradually towards to the north aspect (~ 5 m). A similar pattern, but with a smaller amplitude is found for the NASADEM, TanDEM-X, MERIT (-1 to 1 m) and AW3D30 (0 to 2 m) (Fig. 6a).

The median differences of the 30-m DEMs generally increased along the slope. However, for the 90-m DEMs, the difference increased with slope at first, but then decreased on steep slopes. NASADEM and TanDEM-X had minimum mean median values of about 0.9 and 1.2 m, respectively (Fig. 6b). For all DEMs, the spreads of differences become larger as the slope becomes steeper. This increase is most obvious for TanDEM-X and SRTM v4.1, with rates of 1.29 m/degree (r=0.97, p<0.01) and 1.11 m/degree (r=0.89, p<0.01). This indicated that errors of both DEMs suffered from steep slope effects. AW3D30 and NASADEM have a similar mean spread (19.2 vs 20.8 m). On slopes of less than 20º, TanDEM-X has the best quality with a mean median value of -0.2 m and mean spread of 11.7 m, respectively. MERIT shows a slight advantage over SRTM v4.1 with a reduced spread for steep slopes. Overall, relative to the other DEMs, AW3D30 and NASADEM perform best for steep slopes in terms of spread and median value.

The differences for all DEMs generally decreased with elevation, with fluctuations around zero at very high elevations (Fig. 6c). AW3D30 has a smaller difference at low elevation relative to NASADEM and SRTM-GL1. For NASADEM and SRTM-GL1, the differences along the elevations show a similar distribution and varied from –10 to 10 m over the range 4500−6500 m a.s.l., where measurements are concentrated (Fig. 6d); However, NASADEM behave best out of six DEMs in the high elevation. The difference of TanDEM-X varied from around –5 to 5 m between 4500 and 6500 m a.s.l. The SRTM v4.1 and MERIT differences changed almost similarly from –20 to 40 m, but show differences at the highest elevations.

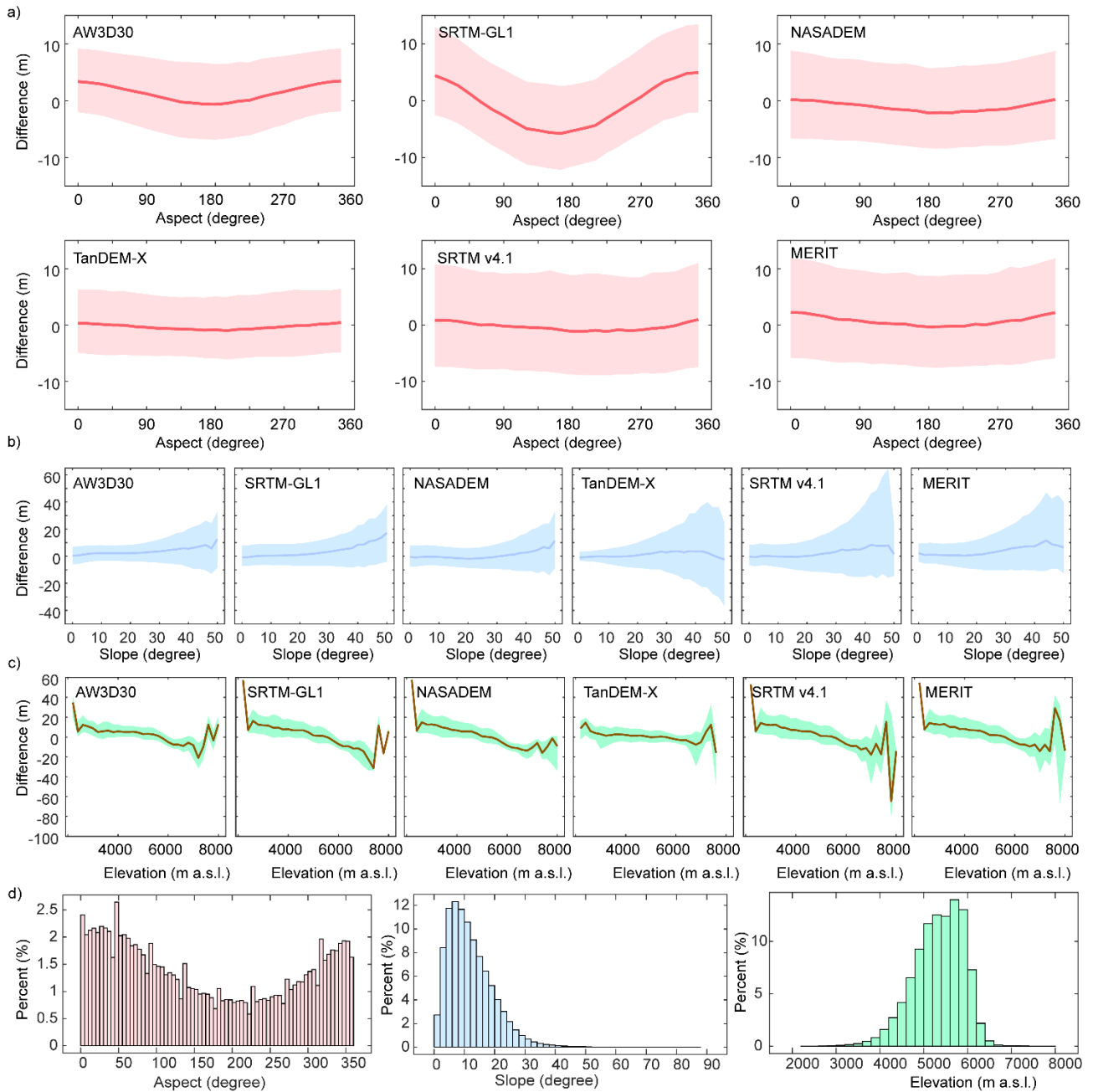



Figure 6. Differences between six DEMs and ICESat-2 with terrain factors. (a) 5° aspect bin. (b) 2° slope bin. (c) 200 m elevation bin. (d) Percentage (%) of data in each aspect, slope and elevation bin.

3.3 Differences between DEMs and ICESat-2 in different glacier zones

Elevations differences in different glacier zones were also estimated and are shown in Fig. 7a−d. We divided it into four sub-zones using the maximal, median and minimum elevation from the RGI glacier inventory (Fig. 7e). Here we consider Zone 1 and Zone 4 to be the ablation area and accumulation area, respectively. Zone 2 and Zone 3 are transition areas. Crests of the probability distribution of differences located in the positive axis range in Fig. 7a move to the left in Fig. 7b−d. Correspondingly, ME, MAE and RMSE all decrease from Zone 1 (ablation area) to Zone 2 (transition area) (Fig.7 and Table S1). Spatially, areas in the glacier terminus are subject to more melting (Brun et al., 2017) leading to this decrease. The ME of the SRTM based products SRTM-GL1, SRTM v4.1, NASADEM and MERIT are all around 10 m in Zone1 and decreased similarly by 8.1, 7.6, 7.5 and 7.2 m towards Zone2, respectively (Table S1). Temporally, the ME of the DEM acquired in earlier periods is bigger. The ME is 8.1m for AW3D30, which was acquired in 2006−2011, bigger than that of TanDEM-X (3.9 m), which was acquired in 2010−2015.

ME, MAE and RMSE in Zone 3 and Zone 4, near or in the accumulation area, are almost all smaller than the corresponding values in Zone 1 (Fig. 7 and Table S1). ME of all DEMs decreased to negative values in Zone 3 and Zone 4. Usually, in the accumulation area, glaciers have a positive or less negative elevation change (Li and Lin, 2017; Maurer et al.,, 2019; Rankl and Braun, 2016), therefore, accumulation may be concerned with changes in Zone3 and Zone 4. The observed shift in the ME from zone 1 to zone 4 is a sign of influence from thinning or accumulation between the time of collection of the six DEMs and the ICESat-2 data.

In terms of STD, NASADEM performed best in Zone 3 and Zone 4, with values ranging from 8.8 to 11.1 m (Table S1). AW3D30 had the best performance of all DEMs in Zone 1 and Zone 2, with an STD varying from 10.0 to 12.3 m. The STD of TanDEM-X was better than that of SRTM v4.1 and MERIT in Zone 1 and Zone 2, but was worse in Zone 3 and Zone 4.

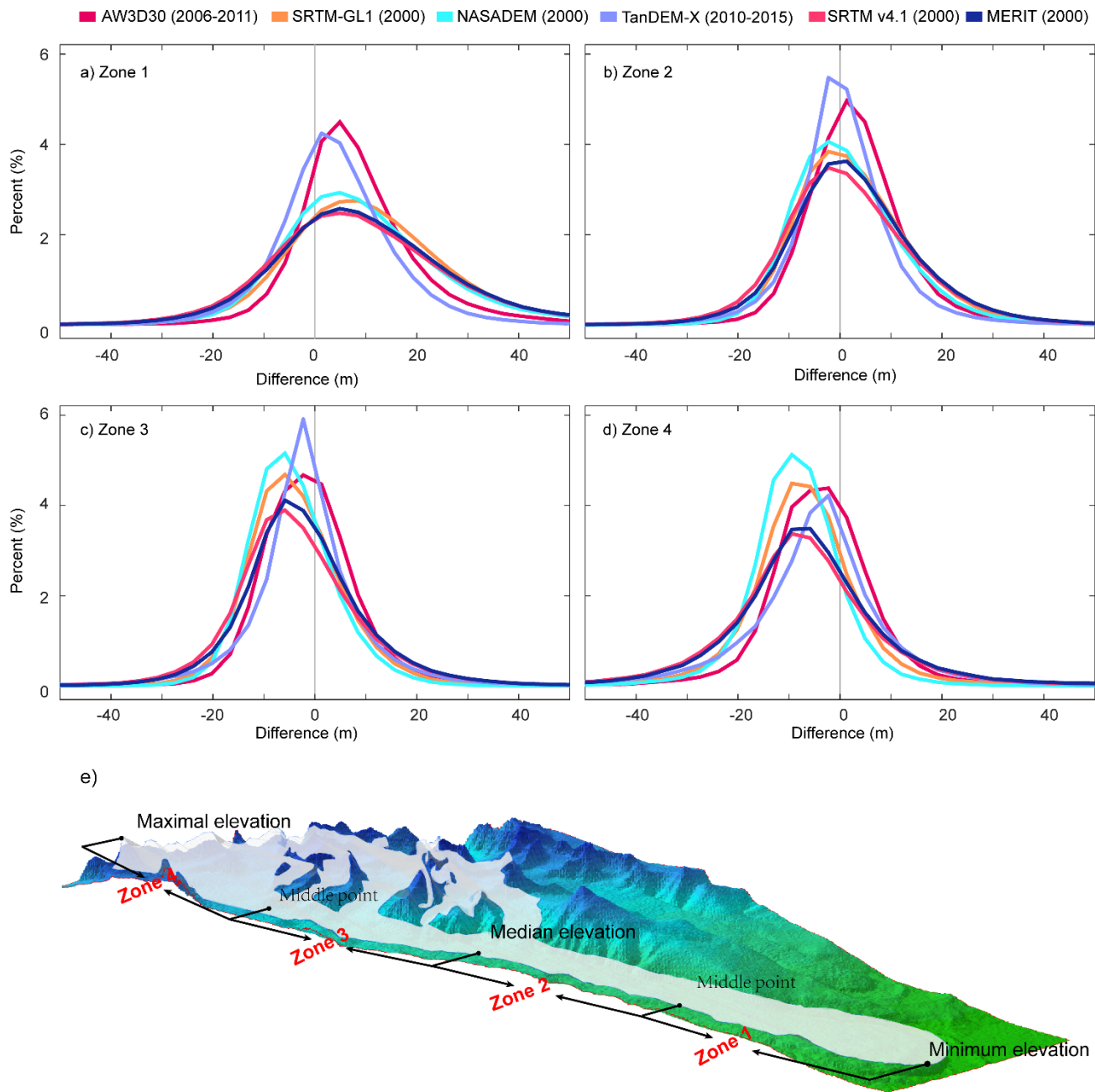



Figure 7. Probability distribution of the difference between six DEMs elevation and ICESat-2 elevation in different glacier zones (a–d) as defined in panel (e). The acquision date of each DEM is labeled.

3.4 Sensitivity of modelled ice thickness to DEMs

The models are not adjusted independently according to the difference between the output and GPR results. Therefore, the results are not indicators of the performance of the models but rather references for examining the influence of different DEMs on ice thickness estimated using different ITIMs. The effects of the DEMs on the model outcomes are presented in Figure 8 and are quite obvious. Mean ice thickness differs, according to the DEM used, by up to 134%, 6%, 47% and 19% for GlabTop2, HF, ITIBOV and OGGM, respectively. The deepest ice thickness differs by up to 53%, 25%, 13% and 13% for GlabTop2, HF, ITIBOV and OGGM, respectively.

The mean ice thicknesses from GlabTop2 and ITIBOV using the 90-m DEMs (they are TanDEM-X, SRTM v4.1 and MERIT) are ~30 m less than those obtained when using 30-m DEMs (that is AW3D, SRTM-GL1 and NASADEM) (Fig. 8). GlabTop2, HF and OGGM using AW3D30, and ITIBOV using NASADEM output the largest mean thickness. GlabTop2 and ITIBOV using TanDEM-X, OGGM and HF using SRTM-GL1 output the smallest mean thickness.

The influence of different DEMs on ITIMs can also be identified when making a comparison with the GPR results (Fig. 8 and Table 2). If the median error is used as the criterion, GlabTop2 and ITIBOV using NASADEM, HF using AW3D30, and OGGM using SRTM v4.1 achieved the relatively best simulation (Fig. 9). If RMSE was used, GlabTop2 using NASADEM, HF using SRTM-GL1, ITIBOV using AW3D30 and OGGM using TanDEM-X performed best (Table 2).

In different glacier zones, each DEM-model combination has its merits and weaknesses (Table 2). NASADEM indicates better performance (number of smaller RMSE in five profiles) relative to other DEMs by using GlabTop2 and ITBIOV models. AW3D30 conducts better by using ITBIOV and OGGM models. TanDEM-X is better by using OGGM model. While five models are composited, NASADEM behaves better. Overall, NASADEM and AW3D30 performs better in different glacier zones in all models.

Thickness (m)

0 40 80 120 160 200 240 280 320 360 400

1 km

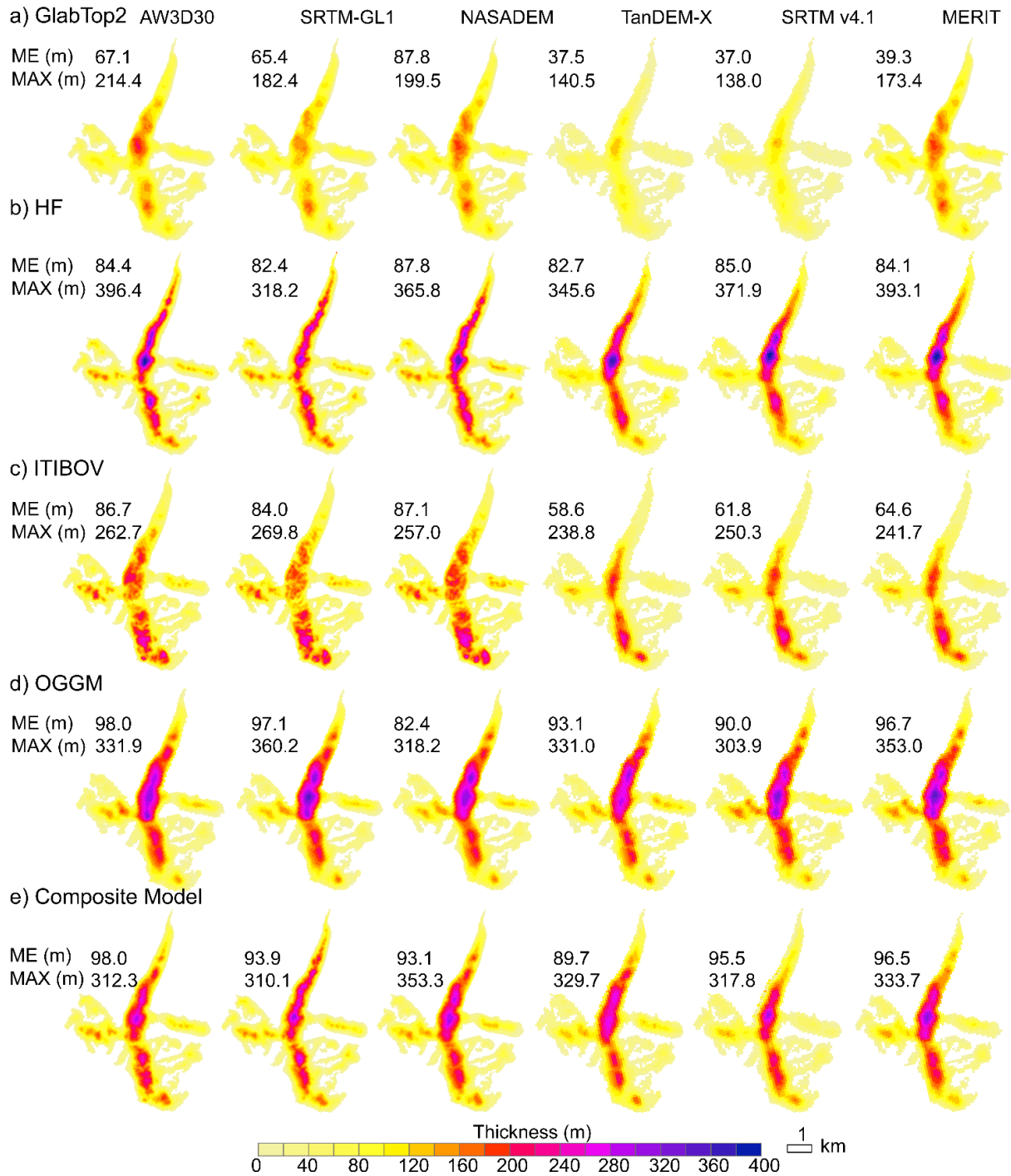



Figure 8. Distribution of modelled ice thickness of Chhota Shigri Glacier (location shown in Fig.1) using AW3D30, SRTM-GL1, TanDEM-X, SRTM v4.1, NASADEM and MERIT. (a) Glabtop2; (b)HF; (c) ITIBOV; (d) OGGM; (e) composite result. Mean (ME) and maximum (MAX) modelled ice thickness are given in each panel.

As similar to the procedure of Farinotti et al. (2017), results from the four models are further composed to achieve the minimum MAE between the modelled and GPR thicknesses (Fig. 8e). The weights for each model in ten experiments are shown in Table S2. After composition, the mean thickness using different DEMs ranged from 90 (acquired based on TanDEM-X) to 98 m (acquired based on AW3D30). NASADEM and AW3D30 achieved minimum MAE, which are 36.7 and 44.1 m, respectively. RMSEs of combined ice thickness modelled from different DEMs are reduced by ~21 m (~25%), when compared to the RMSE for one ITIM. The mean errors and median errors of all DEMs are at the range of ±10 m, except for that of AW3D30 and TanDEM-X which is at a level of around 20m. The spreads of error of 30-m DEMs are 33% smaller than those of 90-m DEM. Error spread from NASAM was minimum (75.1 m), followed by AW3D30 (77.3 m).

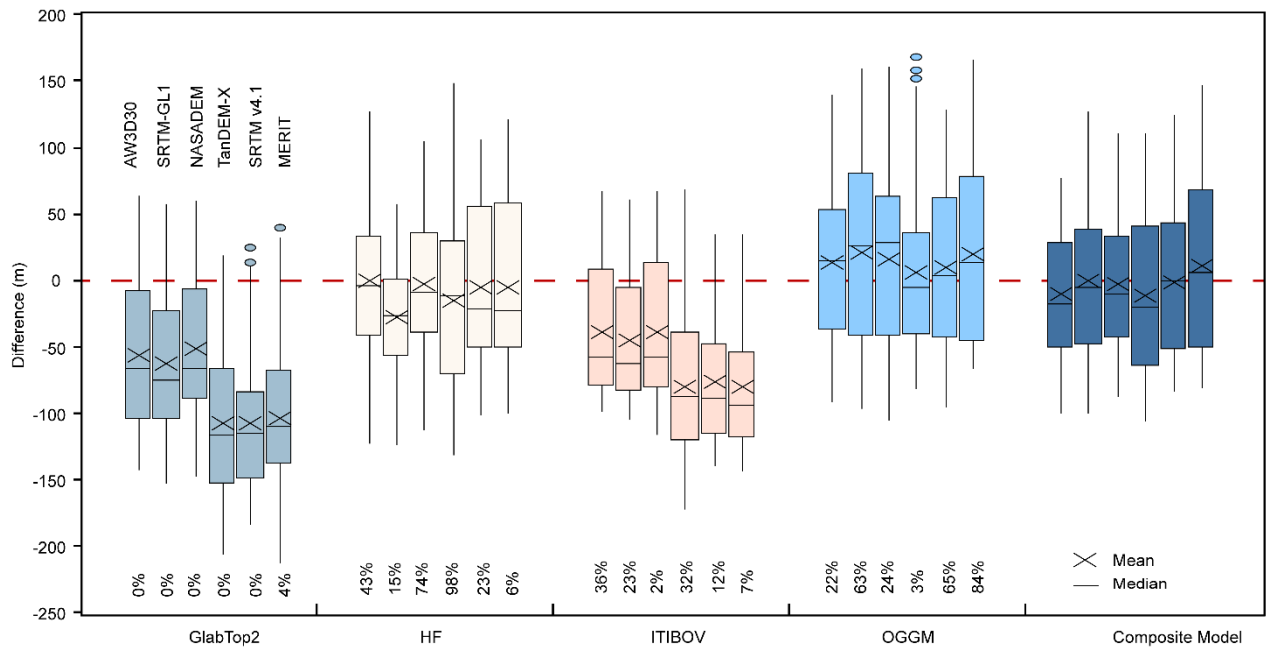



Figure 9. Point-by-point deviation comparison between the modelled and measured ice thickness from GlabTop2, HF, ITIBOV, OGGM and the composite result. In each group, the boxes are plotted in the order: AW3D30, SRTM-GL1, NASADEM, TanDEM-X, SRTM v4.1 and MERIT. Different models using the same DEM are aggregated by weights (labelled at the bottom) to achieve minimum mean absolute error.

Table 2. Root mean squared error (RMSE, m) of modelled ice thickness compared with ground penetrating radar (GPR) measurements on each profile (pf1 to pf5). The locations of profiles are shown in Figure 1. Bold numbers denote the best model performance on each profile using different DEMs.

Model	DEM	No. of ground penetrating radar profiles					
		pf1	pf2	pf3	pf4	pf5	All
GlabTop2	AW3D30	56.6	84.5	**45.4**	**102.4**	103.3	79.7
	SRTMG-GL1	**54.8**	94.6	62.4	104.3	90.6	83.3
	NASADEM	53.8	**75.8**	48.5	104.6	**84.4**	**75.3**
	TanDEM-X	103.5	143.3	104.8	137.0	115.0	122.0
	SRTM v4.1	101.9	133.6	93.2	132.3	135.1	118.7
	MERIT	110.0	120.7	70.9	154.7	132.2	118.1
HF	AW3D30	**27.2**	29.8	83.7	30.7	92.5	61.6
	SRTM-GL1	28.4	58.4	**33.3**	30.6	88.1	**50.0**
	NASADEM	37.2	**26.5**	62.7	**30.1**	82.2	51.8
	TanDEM-X	50.7	63.4	60.9	66.5	**83.9**	65.4
	SRTM v4.1	49.3	36.4	74.9	44.8	86.4	61.3
	MERIT	51.4	33.6	87.6	42.4	86.2	65.8
ITBIOV	AW3D30	69.3	61.9	**45.6**	**66.4**	70.8	**61.4**
	SRTMG-GL1	70.7	71.7	52.6	70.6	61.4	64.8
	NASADEM	**67.3**	**61.3**	58.3	80.3	**57.1**	65.3
	TanDEM-X	98.3	116.4	61.1	91.2	109.7	94.0
	SRTM v4.1	102.9	114.5	51.8	73.4	111.3	88.3
	MERIT	108.3	117.6	54.8	80.4	112.7	91.9
OGGM	AW3D30	**32.0**	47.2	**80.5**	62.5	**34.6**	58.9
	SRTMG-GL1	33.5	56.9	101.6	65.1	41.1	70.2
	NASADEM	32.2	55.3	92.3	68.8	43.6	67.1
	TanDEM-X	42.7	**35.6**	86.1	**52.2**	35.4	**58.4**
	SRTM v4.1	37.4	46.1	81.4	68.7	39.9	61.6
	MERIT	38.3	52.1	104.1	53.8	49.7	69.7
Composite	AW3D30	37.2	27.9	**44.4**	45.8	68.3	**45.8**
	SRTMG-GL1	39.1	29.8	64.3	63.3	**47.5**	52.7
	NASADEM	**34.9**	**22.1**	65.4	**41.1**	70.6	51.3
	TanDEM-X	50.9	46.3	60.5	57.9	69.0	57.5
	SRTM v4.1	41.8	41.4	59.3	59.5	55.1	53.1
	MERIT	42.0	49.1	85.3	53.8	55.3	62.5

365 ## 4 Discussion

4.1 Influence of glacier elevation change on the assessment of DEMs

The identified extreme outliers (Fig. 3) are mostly located near the glacier terminus, high elevation and high slope regions (Fig. 10a−b). Extreme glacier melt, such as in south-eastern Tibet, and surges, as observed in the Karakoram, can also lead to

dramatic elevation changes, resulting in large differences (Fig. 10c). This glacier elevation change effect is also reflected in
370 the spatial distribution of difference (Fig. 5), elevation bins (Fig. 6c) and glacier zones (Fig. 7). The differences at lower
elevations are positive, and generally decrease with elevation, consistent with the fact that glaciers melt at lower elevations
and accumulate at higher elevations (Cuffey and Paterson, 2010). The differences in all DEMs with elevation and glacier zones
comply with these features (Fig. 6c and Fig. 7). NASADEM was acquired in 2000 and TanDEM-X was acquired in 2010−2015,
and the value of NASADEM is more positive than TanDEM-X in the ablation zone. The relatively more positive and larger
375 values of ME and STD along the Hindu Kush-Himalaya, southern Tibet (Fig. 5) and negative ME values in the West Kunlun
and Karakoram (Fig. 5) are also related to glacier elevation change (Hugonnet et al., 2021).

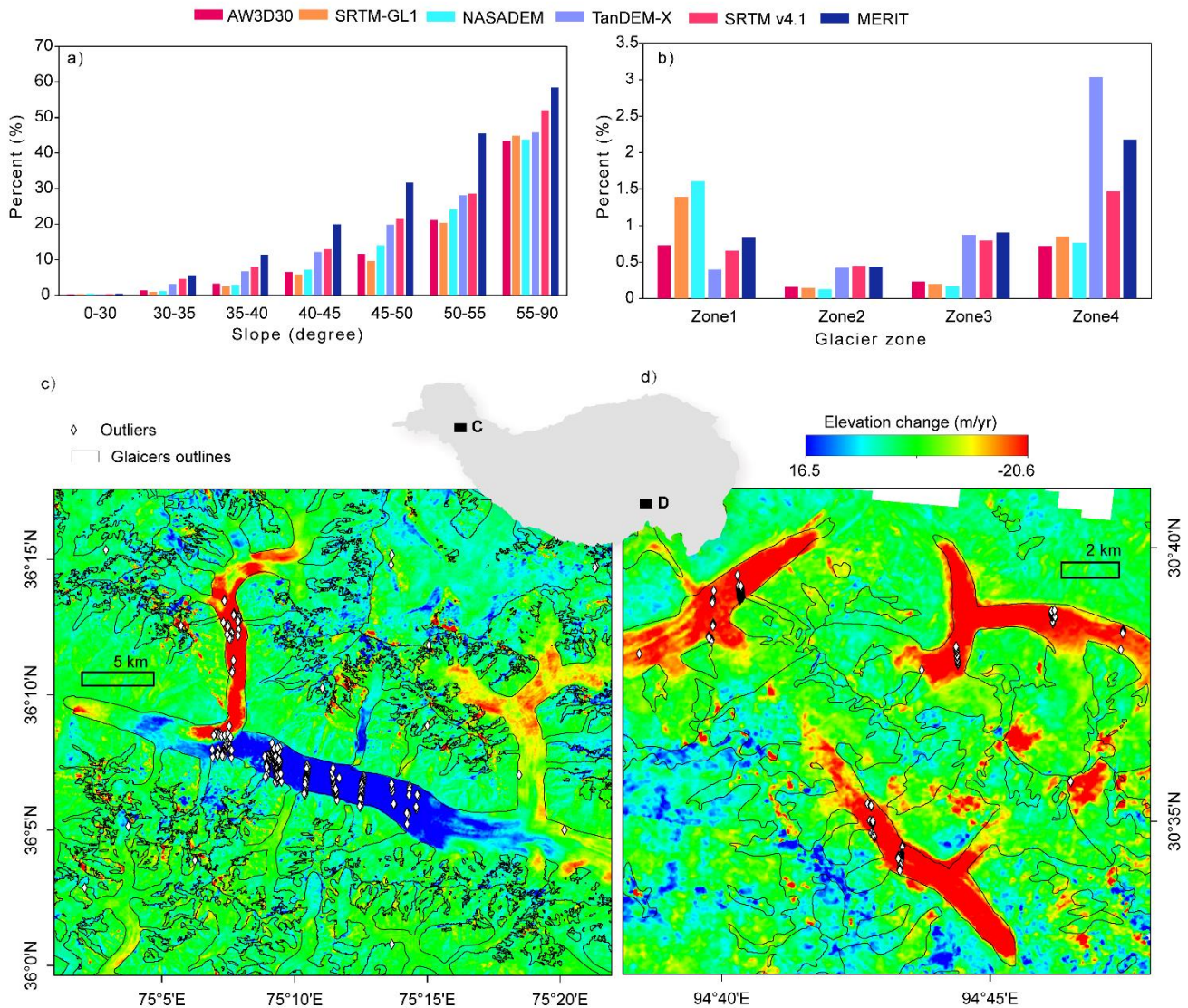



Figure 10. Distribution of excluded extreme outliers. The proportion of outliers accounting for the total number in slope bins (a) and each glacier Zone (b). Examples of locations of excluded points overlaid with glacier surface elevation change in the Karakoram (c) and southern TP (d). Locations of these two examples in panels c and d are labelled C and D in the central insert, respectively. Glacier elevation change data covering 2000−2019 is from Shean et al. (2020).

Table 3. Comparisons of differences between four SRTM based DEMs and ICESat-2 elevation over glacier zones before and after adjustment. ICESat-2 data acquired in February are used to calculate the differences. Glacier zones are defined according to Fig. 7e.

Item	Zone	Before (m)				After (m)			
		SRTM-GL1	NASADEM	SRTM v4.1	MERIT	SRTM-GL1	NASADEM	SRTM v4.1	MERIT
Mean error	1	27.8	26.6	23.9	27.4	13.4	12.1	9.7	12.8
	2	10.0	9.1	8.3	10.5	2.9	2.0	1.7	3.6
	3	-0.2	-0.8	0.7	1.7	-3.5	-3.9	-2.2	-1.4
	4	-12.4	-12.0	-8.5	-6.0	-9.3	-10.1	-7.5	-5.3
Absolute mean error	1	33.8	33.0	36.3	34.3	17.8	16.8	20.5	18.5
	2	16.4	16.1	19.0	18.0	9.0	8.6	11.9	10.9
	3	11.2	10.8	14.4	13.8	8.3	7.7	11.6	10.7
	4	18.0	16.0	23.4	20.6	13.0	12.1	19.6	16.9
Standard deviation	1	37.7	37.9	53.5	39.4	20.3	20.3	40.1	22.8
	2	20.6	20.8	36.1	26.8	12.0	11.6	29.2	21.8
	3	14.6	14.5	23.7	22.7	10.7	9.8	20.7	19.3
	4	29.6	27.4	39.5	29.4	20.9	20.8	36.9	26.0
RMSE	1	46.9	46.3	58.6	48.0	24.3	23.7	41.2	26.2
	2	22.9	22.7	37.0	28.8	12.4	11.8	29.2	22.1
	3	14.6	14.5	23.7	22.7	11.3	10.6	20.8	19.4
	4	32.1	29.9	40.4	30.0	22.9	23.2	37.6	26.6

After removing the glacier elevation change using the glacier elevation change dataset covering 2000−2018 (Shean et al., 2020), the mean difference in Zone 1 and Zone 2 decreased sharply by ~14 m and ~7 m for the SRTM based DEMs, respectively (Table 3). However, similar improvements are not obvious in Zone 3 and Zone 4. This may be related to the slight elevation change in the accumulation region (Brun et al., 2017; Shean et al., 2020), and high uncertainty due to steeper slopes and higher elevations (Fig. 6b−c). Apart from these factors, penetration depth of C-band should be related to the remaining errors in accumulation area. Mean errors in SRTM based DEMs are 5.3−10.1 m are at the same order with the C-band penetration (Rignot et al. 2001). MAE, STD and RMSE all improved a lot in four regions after this adjustment.

ICESat-2 data covering the period from October 2018 to October 2020 repeat every 91 days. Therefore, variations of ICESat-2 elevation data caused by glacier fluctuations have influenced the error statistics (Fig. 11a). Precipitation on the TP mainly occurs in June−August (Maussion et al., 2014). Hence, after precipitation accumulation on glaciers in spring and summer, the

elevation increased, and the mean difference decreased. With little accumulation, glaciers experience more melt and sublimation in autumn and winter (Li et al., 2018). As a result, the glacier surface elevation decreases, and the mean difference increases. However, the magnitude of these changes is much smaller, at a level of less than 3 m (Fig. 11a), compared with the large ME, MAE and RMSE magnitude of most of the DEMs (with the exceptions of TanDEM-X and NASADEM) (Fig. 4). When taking all points from different seasons into consideration, the ICESat-2 dataset gives average elevation over the 2018-2020 period, the seasonal effects could also partly cancel each other out. If only the ICESat-2 data from February was used (Table 3), NASADEM and TanDEM-X still perform better than others. Therefore, we conclude that the seasonal fluctuations of ICESat-2 data have little influence on the assessments of the DEMs under the above conditions.

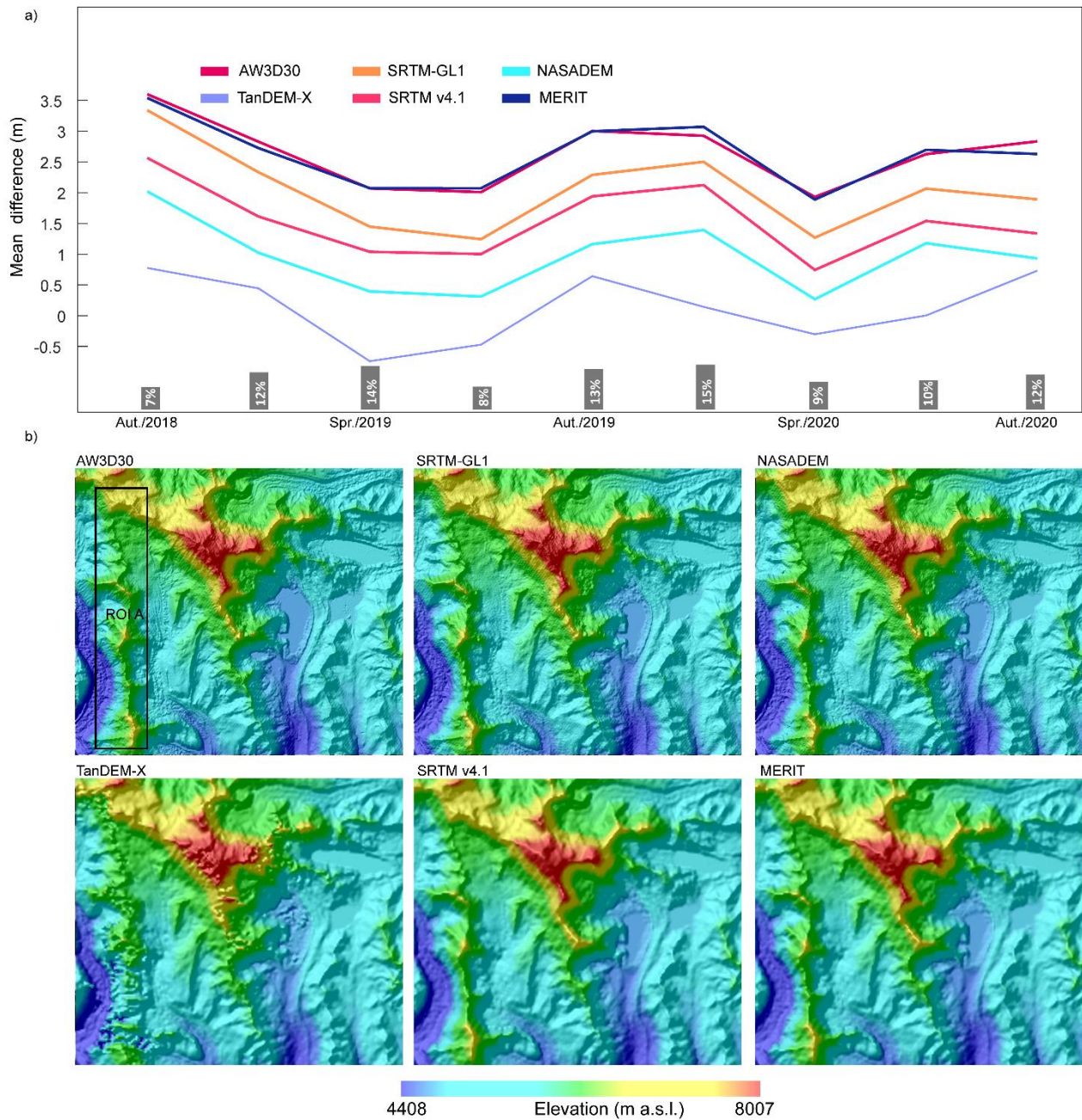
400

405

Figure 11. Influence on elevation differences between ICESat-2 and six DEMs from glacier elevation change and terrain factors. (a) Mean absolute difference between six DEMs and ICESat-2 in different seasons during 2018−2020. The spring, summer, autumn, and winter are defined as March−May, June−August, September−November and December−February,

respectively. The histogram at the bottom shows the percentage of the total number of points in each season. (b) Examples of
410 elevation and shaded relief of six DEMs in the Shisha Pangma region. The rectangle denotes the area of interest.

4.2 Influence of terrain on the assessment of DEMs

The elevation differences depend strongly on terrain factors (Fig. 6a−b). The largest errors are concentrated in the north aspect,
as was also reported in previous studies (Gorokhovich and Voustianiouk, 2006; Shortridge and Messina, 2011), in which they
415 were attributed to the orientation of the sensor (Gdulová et al., 2020; Shortridge and Messina, 2011). However, here, the data
from different sensors all show this aspect dependence, and we infer that it may be related to the predominance of slopes for
certain aspects. There are many more measurements with steeper slopes in the north aspect, and fewer measurements with
flatter slopes in the south aspect (Fig. 12). The error and spread become larger with steeper slopes (Fig. 7b), as also reported
by Liu et al. (2019) and Uuemaa et al. (2020), which may be due to geometric deformation or shadow (Liu et al. 2019).
420 Therefore, the error variation with aspect tends to be related to steeper slopes (Gdulová et al., 2020; Gorokhovich and
Voustianiouk, 2006).

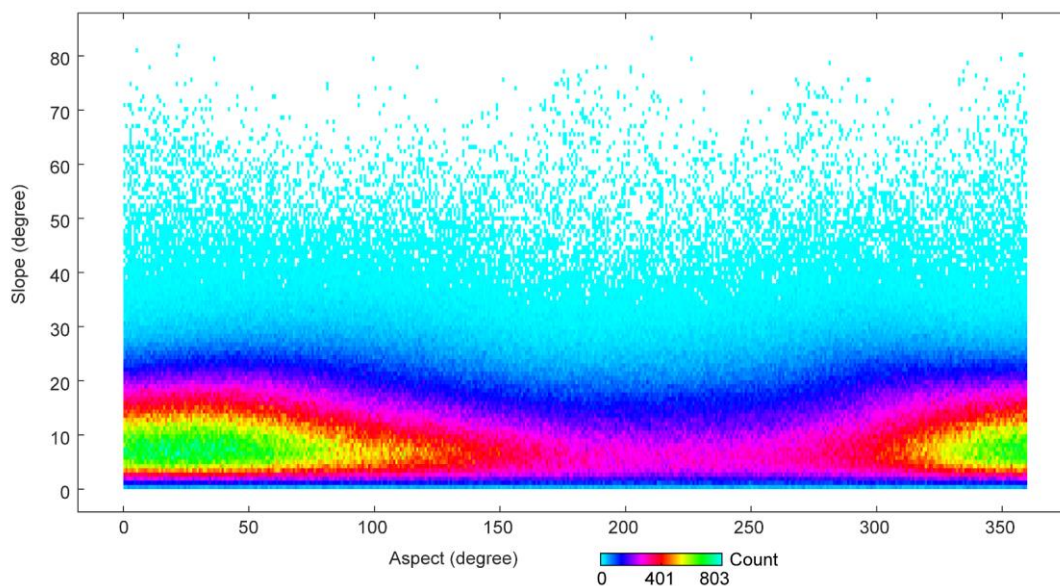



Figure 12. Distribution of measurements in different aspects against the slope.

425 Though points in the 55°−90° slope region account for a small fraction (Fig. 6d), almost half the points in the 55°−90° slope
region are identified as extreme outliers (Fig. 10a). Differences also show large discrepancies for all DEMs in the steeply
sloping regions where voids and large errors are frequent (Falorni, 2005). Steep slopes combined with low resolution led to
variations in the spread of differences in Fig. 6b. Spreads of differences were larger on steep slopes for the 90-m DEMs than

those of the 30-m DEMs. Intra-pixel variation aggravates this effect in steeply sloping regions (Uuemaa et al., 2020), lower
430   resolution or reduced pixel DEMs smooth the terrain details and lead to inaccurate elevation compared with the 20-m footprint
of ICESat-2 points. The spread and the number of outliers gradually increased with the slope, especially for the TanDEM-X
case (Fig. 7b). Using the terrain in the rugged Shishapangma region (Fig. 12b) as an example, we can see that the elevation
from TanDEM-X suffers from substantial errors along the ridge at high elevations and the output appears almost blurred. Even
so, TanDEM-X still has overall accuracy advantages over SRTM v4.1 and MERIT, indicating the high quality of TanDEM-X
435   in low relief regions (Fig. 7b).

4.3   Influence of misregistration on the assessment of DEMs

Six DEMs are produced from different sensors or by different methods. Pixels in DEMs do not represent exactly the same
location. This misregistration among DEMs, which has been ignored in previous research (González-Moradas and Viveen,
440   2020; Liu et al., 2019), is an important error source when looking at DEM differences (Hugonnet et al., 2021; Van Niel et al.,
2008). This study intends to give direct insights into the quality of uncorrected DEM products, so the misregistration problem
was not tackled before the evaluations were carried out. However, the influence of misregistration was evaluated. According
to the sinusoidal relationship between aspect and error differences between two DEMs (Van Niel et al., 2008), using the co-
registration method in Nuth and Kääb (2011) and ICESat-2 points outside the glaciers, offset pixels relative to ICESat-2 in x-
445   and y- direction at the 1°×1° grid scale were estimated by nonlinear least squares fitting method across the TP. Misregistration
was found to be less than one grid spacing (Figure 13). The misregistration of SRTM-GL1 relative to ICESat-2 is the largest;
the misregistration of other DEMs is always less than 0.2 grid spacing. Considering that only the cell centre value was used,
the sub-grid spacing shift may have little influence (Van Niel et al., 2008).

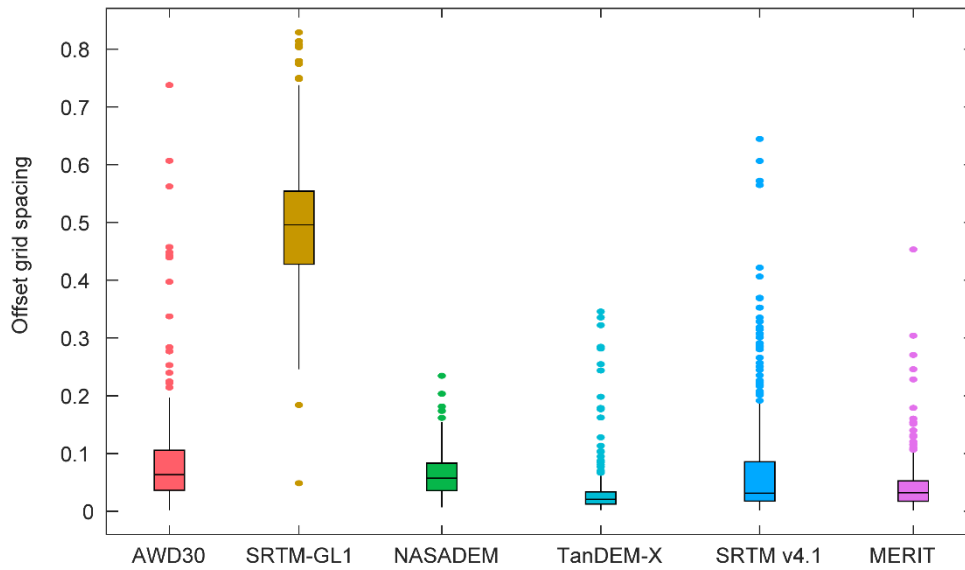



Figure 13. Distribution of offset grid spacings of DEMs relative to ICESat-2 on a 1°×1° grid. Only the grid squares with R^2 greater than 0.5 and the number of record greater than 1000 are considered.

4.4 Influence of DEMs on ice thickness estimated by ITIMs

Even with the same parameters, the same ITIM using different DEMs will yield different thickness patterns (Figs. 8, 9). The quality of a DEM indeed influences the performance of the ITIMs. However, the different models have various levels of robustness to the quality of the input DEMs. Different DEMs resulted in differences in largest and smallest mean ITIM ice thickness at a range of 3.6−32 m (Fig. 8).

Generally, the outcome with GlabTop2 and ITIBOV using 30-m DEMs is 51% and 43% better than with the 90-m DEMs in mean error, respectively. With GlabTop2, elevation data was used to determine not only the slope, but also the shear stress (Frey et al., 2014). A sensitivity test based on the Equations 1−4 was executed and the modelled ice thickness differences before and after adding perturbations to the input parameters such as slope and elevation were compared. An error of +5° in slope caused more than a –34.1 % difference in the output for slopes of less than 20°. Additionally, relative elevation errors had an enormous impact (Fig. 14 b). For glaciers with an elevation range of less than 400 m, which accounted for 41% of the total number and 5% of the total area over the TP, +10, +30, and +50 m errors in elevation range caused more than +2%, +6% and +10% differences in output. Such errors in elevation range had greater influence (Fig. 14b), especially for small glaciers, which have smaller elevation ranges. These two errors propagate and lead to a much larger overall error (Table 3). Thus, GlabTop2 using NASADEM and AW3D30 with best quantity as input achieved the best RMSE in comparison with GPR measurements. In contrast to the other ITIMs, the ITIBOV model directly estimated the ice thickness at each grid cell according

470   to cell velocity information without interpolation. The slope sensitivity of ITIBOV is higher than that of GlabTop2, with an

error of +5º in slope causing more than a –71.4% difference in the ice thickness for slopes of less than 20º (Fig. 14a). The

flatter the slope, the more sensitive ITIBOV is to the slope error (Fig. 14a). Although along- and across-track slope data are

provided in the ICESat-2 ATL06 product, they are incompatible with the slope estimated from DEMs due to their different

data formats and algorithms used (Burrough and McDonell, 1998; Smith et al., 2019). Moreover, the surface terrain of glaciers

475   changes with time due to accumulation, melting and transient dynamics (Dehecq et al., 2018; Shean et al., 2020). Nevertheless,

the accuracy of the DEMs estimated here could also provide some information about slope accuracy. With the higher-accuracy

NASADEM and AW3D30 as input of ITIBOV most accurate ice thicknesses were obtained (Table 2).

For HF and OGGM, the modelled ice thicknesses did not show large differences when using 30 or 90 m DEMs as input.

(Fig. 9) , thereby suggesting a minor impact of DEM resolution on ice thickness reconstruction (Pelto et al., 2020). In the HF

480   model, elevation data was used for convergence calculation of apparent mass balance and mean slope in elevation bins

(Farinotti et al. 2009; Farinotti et al. 2019), whereas, in OGGM, it is used to extract flowlines, shear stress at flowlines and

mass balance at an elevation (Maussion et al. 2019). These two models show roughness to the input DEM (Fig. 14a). Although

NASADEM and TanDEM-X were more accurate, the output of HF and OGGM using these two DEMs did not yield better

results compared to when using the other DEMs (Fig. 9). The STD of RMSE values for HF and OGGM using six DEMs are

485   6.2 and 4.9 m, respectively (Table 2). STD of mean ice thickness by HF and OGGM using six DEMs are 1.1 and 6.0 m (Fig.

8).

When the results from different ITIMs are ensembled, the influences of uncertainty and resolution in the input DEMs on the

modelled ice thickness still exist (Fig. 9 and Table 2). The RMSE of ITIMs from 30-m DEMs was 16.8% less than that from

90-m DEMs. Models using AW3D30 and NASADEM, which have higher resolution and better accuracy, yielded most

490   accurate thickness estimates. However, glacier surface elevation changes with climate, AW3D30 acquired in different years

and seasons represents glacier terrain in different periods. This could result in the spatial inconsistencies of the output of ITIMs

in large-scale ice-thickness inversion. Above, we suggest NASADEM as the best input of ITIMs for ice-thickness estimates

over the TP. This conclusion is of significance for ice thickness inversion models using DEMs in TP. However, it should be

noted that the result may be not suitable for studies in other glacierized mountainous regions. Because various errors

495   exist in DEMs, such as speckle noise, stripe noise and absolute bias; they behave differently across the Earth (Yamazaki

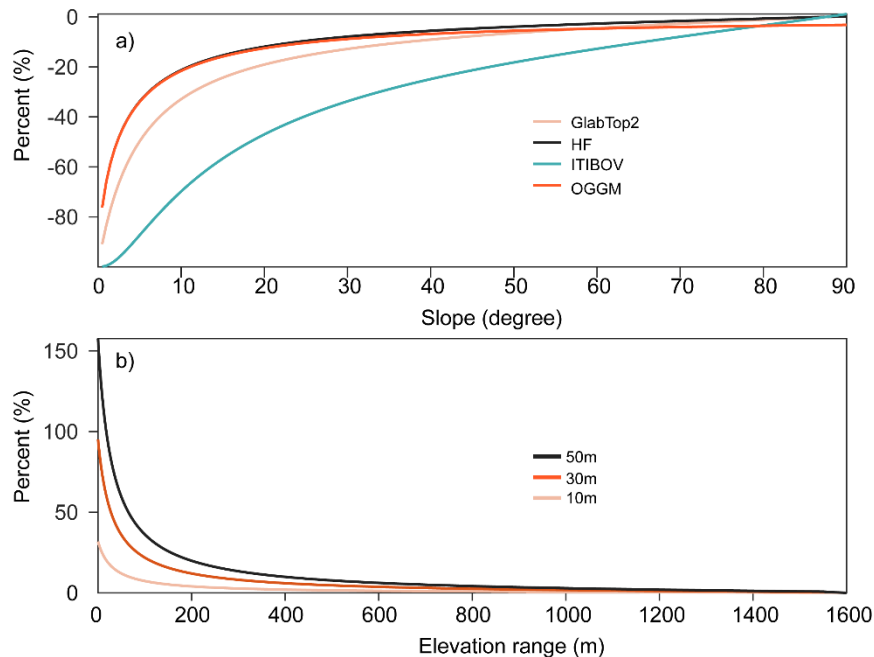
et al., 2017; Takaku et al., 2020). Our method to assess the accuracy of DEMs is repeatable in different regions, combined

with the recently released glacier elevation change data on Earth (Hugonnet et al., 2021). Furthermore, benefiting from

the high accuracy and dense coverage of ICESat-2 data, the quality of DEMs can also be improved, similar to what has

been done in the production of MERIT (Yamazaki et al., 2017). For example, the misregistration in DEMs could be

500   corrected and terrain-related errors could be reduced by unitizing the relation of difference against slope, aspect and

elevation in Fig. 6.

 Figure 14. Sensitivity test of slope and elevation on ice-thickness inversion models. (a) Percentage difference of modelled ice thickness from GlabTop2, HF, ITIBOV and OGGM when there is +5° slope error; (b) Percentage difference of modelled ice thickness from GlabTop2 when the elevation range error is +10, +30 and +50 m for different elevation ranges.

5 Conclusions

In the present study, six DEMs (i.e. AW3D30, SRTM-GL1, NASADEM, TanDEM-X, SRTM-GL1 and MERIT) from different sensors with different spatial resolutions were evaluated using ICESat-2 data. The influence of glacier dynamics, terrain and misregistration on the DEM accuracy was analysed. Out of the three 30-m DEMs, NASADEM was the best performer in vertical accuracy with an ME of 0.9 m and an RMSE of 12.6 m. Out of the three 90-m DEMs, TanDEM-X performed best with an ME of 0.1 and an RMSE of 15.1 m. The quality of TanDEM-X was stable and unprecedented on shallow slopes, but suffered from serious problems on steep slopes, especially along the steep ridges. AW3D30 has similar accuracy to NASADEM, and is even better in STD, MAE and RMSE when not considering the effect of glacier dynamics. SRTM-based DEMs (i.e. SRTM-GL1, SRTM v4.1 and MERIT) (~15 m RMSE) were inferior to NASADEM. MERIT shows little improvement over SRTM v4.1 in glacierized terrain. The influence of glacier elevation change on the elevation difference is larger for DEMs acquired in earlier period, at low elevations and in the ablation region. However, this does not influence the conclusion that NASADEM performed the best, followed by TanDEM-X but with serious outliers in the high elevation region. For all the DEMs, the errors increased from the south-aspect slope to the north-aspect slope, controlled by the increasing

error with slope. Misregistration errors in the glacier region are within one grid spacing, and have little influence on the evaluation benefiting from the 20 m footprint of ICESat-2 relative to the 30- or 90-m resolution DEMs.

The influence of DEM accuracy on ice-thickness inversion models depends on the model properties. Generally, a higher resolution DEM was helpful for better model outcomes. The widely used GlabTop2 model is very sensitive to the accuracy of

525 both elevation and slope; using NASADEM as the input, this model facilitated the best outcome. Although the OGGM and HF models are less sensitive to the quality of DEM, the use of NASADEM or AW3D30 was still beneficial. Among the four ice-thickness inversion models, ITIBOV was the most sensitive to slope accuracy. Ice-thickness inversion models using AW3D30 or NASADEM as input gave the most accurate thickness estimates. These two DEMs also perform the best when four ice-thickness inversion results were aggregated by the minimum MAE optimization method.

530 Considering the influence of inconsistency in data acquisition time on generating glacier terrain, we suggest that NASADEM is the best choice for ice-thickness inversion models over the whole TP. AW3D30 could be a good substitute but with limitations from its mixed acquiring dates. TanDEM-X is an appropriate alternative for glaciological research focusing on the flat glacier terminus, but it requires further improvement for use in steep terrain or ice-thickness inversion.

535 **Code availability.** The codes for processing the ICESat-2 data are available at https://github.com/cnugis/ProcessICESat-2.

Data availability. ICESat-2 ATL06 is available at https://nsidc.org/data/atl06 (Accessed 2021-10-19); AW3D30 is available at https://www.eorc.jaxa.jp/ALOS/en /aw3d30/index.htm (Accessed 2021-10-19); SRTM-Gl1 is available at https://earthexplorer.usgs.gov/ (Accessed 2021-10-19); NASADEM is available at https://search. earthdata.nasa.gov/

540 (Accessed 2021-10-19) ; TanDEM-X 90m is available at https://download.geoservice.dlr.de/TDM90/ (Accessed 2021-10-19); SRTM v4.1 is available at https://drive.google.com/drive/folders/0B_J08t5spvd8RWRmYmtFa2puZEE (Accessed 2021-10-19); MERIT is available at http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/ (Accessed 2021-10-19); Glacier elevation change data is available at https://zenodo.org/record/3600624 (Accessed 2021-10-19). The database containing elevation, slope and aspect and some other fields in MATLAB mat format is available at https://zenodo.org/record/5267309.

545

Supplement. The supplement related to this article is available online at:

Author contribution. W.F Chen, T.D Yao and GQ. Zhang designed the outline of this study. W.F Chen processed the data and made all the figures. Fei Li estimated the ice thickness using OGGM. Guoxiong Zheng, Yushan Zhou and Fenglin Xu

550 reviewed and edited the manuscript. All authors contributed to the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the two anonymous reviewers, David E. Shean, and the editor Arjen Stroeven for the constructive comments that improved the manuscript.

Financial support. This study was supported by grants from the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (2019QZKK0201), the Strategic Priority Research Program (A) of the Chinese academy of Sciences (XDA20060201) and the Natural Science Foundation of China (41871056).

Review statement. This paper was edited by Arjen Stroeven and reviewed by two anonymous referees.

References

Allen, S.K., Zhang, G., Wang, W., Yao, T., & Bolch, T.: Potentially dangerous glacial lakes across the Tibetan Plateau revealed using a large-scale automated assessment approach. Sci. Bull., 64, 435-445, https://doi.org/10.1016/j.scib.2019.03.011, 2019

Altunel, A.O.: Evaluation of TanDEM-X 90 m Digital Elevation Model. Int. J. Remote Sens., 40, 2841-2854, https://doi.org/10.1080/01431161.2019.1585593, 2019

Azam, M.F., Wagnon, P., Ramanathan, A., Vincent, C., Sharma, P., Arnaud, Y., Linda, A., Pottakkal, J.G., Chevallier, P., Singh, V.B., & Berthier, E: From balance to imbalance: a shift in the dynamic behaviour of Chhota Shigri glacier, western Himalaya, India. J GLACIOL, 58, 315-324, https://doi.org/10.3189/2012JoG11J123,2012

Bachmann, M., Kraus, T., Bojarski, A., Schandri, M., Boer, J., Busche, T., Bueso Bello, J.L., Grigorov, C., Steinbrecher, U., Buckreuss, S., Krieger, G., & Zink, M.: The TanDEM-X Mission Phases - Ten Years of Bistatic Acquisition and Formation Planning. IEEE J-STARS, 1-1, https://doi.org/10.1109/JSTARS.2021.3065446, 2021

Bhambri, R., Bolch, T., & Chaujar, R.K.: Mapping of debris-covered glaciers in the Garhwal Himalayas using ASTER DEMs and thermal data. Int. J. Remote Sens., 32, 8095-8119, https://doi.org/10.1080/01431161.2010.532821, 2011

Brun, F., Berthier, E., Wagnon, P., Kaab, A., & Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances, 2000-2016. Nat. Geosci., 10, 668-673, https://doi.org/10.1038/ngeo2999, 2017

Brunt, K.M., Neumann, T.A., & Smith, B.E.: Assessment of ICESat-2 Ice Sheet Surface Heights, Based on Comparisons Over the Interior of the Antarctic Ice Sheet. Geophys. Res. Lett., 46, 13072-13078, https://doi.org/10.1029/2019GL084886, 2019

Brunt, K. M., Smith, B. E., Sutterley, T. C., Kurtz, N. T., and Neumann, T. A.: Comparisons of Satellite and Airborne Altimetry with Ground-Based Data From the Interior of the Antarctic Ice Sheet, Geophys Res Lett, 48, e2020GL090572, https://doi.org/10.1029/2020GL090572, 2021.

Burrough, P.A., & McDonell, R.A.: Principles of Geographical Information Systems. Oxford University Press, New York, 190 pp., 1988

585   Chen, W., Yao, T., Zhang, G., Li, S., & Zheng, G.: Accelerated glacier mass loss in the largest river and lake source regions of the Tibetan Plateau and its links with local water balance over 1976–2017. J. Glaciol., 1-15, https://doi.org/10.1017/jog.2021.9, 2021

Crippen, R., Buckley, S., Belz, E., Gurrola, E., Hensley, S., Kobrick, M., Lavalle, M., Martin, J., Neumann, M., Nguyen, Q., Rosen, P., Shimada, J., Simard, M., & Tung, W.: NASADEM global elevation model: methods and progress. Int. Arch.

590   Photogramm. Remote Sens. Spatial Inf. Sci., XLI-B4, 125–128, https://doi.org/10.5194/isprs-archives-XLI-B4-125-2016, 2016.

Cuffey, K., & Paterson, W.S.B.: The Physics of Glaciers 4th Edition. Academic Press, 2010

Dehecq, A., Gourmelen, N., Gardner, A.S., Brun, F., Goldberg, D., Nienow, P.W., Berthier, E., Vincent, C., Wagnon, P., & Trouvé, E.: Twenty-first century glacier slowdown driven by mass loss in High Mountain Asia. Nat. Geosci., 12, 22–27,

595   https://doi.org/10.1038/s41561-018-0271-9, 2019

Fahnestock, M., Scambos, T., Moon, T., Gardner, A., Haran, T., & Klinger, M.: Rapid large-area mapping of ice flow using Landsat 8. Remote Sens. Environ., 185, 84-94, https://doi.org/10.1016/j.rse.2015.11.023, 2016

Falorni, G.: Analysis and characterization of the vertical accuracy of digital elevation models from the Shuttle Radar Topography Mission. J. Geophys. Res.-Earth, 110, https://doi.org/10.1029/2003JF000113, 2005

600   Farinotti, D., Brinkerhoff, D.J., Clarke, G.K.C., Fürst, J.J., Frey, H., Gantayat, P., Gillet-Chaulet, F., Girard, C., Huss, M., Leclercq, P.W., Linsbauer, A., Machguth, H., Martin, C., Maussion, F., Morlighem, M., Mosbeux, C., Pandit, A., Portmann, A., Rabatel, A., Ramsankaran, R., Reerink, T.J., Sanchez, O., Stentoft, P.A., Singh Kumari, S., van Pelt, W.J.J., Anderson, B., Benham, T., Binder, D., Dowdeswell, J.A., Fischer, A., Helfricht, K., Kutuzov, S., Lavrentiev, I., McNabb, R., Gudmundsson, G.H., Li, H., & Andreassen, L.M.: How accurate are estimates of glacier ice thickness? Results from ITMIX,

605   the Ice Thickness Models Intercomparison eXperiment. The Cryosphere, 11, 949-970, https://doi.org/10.5194/tc-11-949-2017, 2017

Farinotti, D., Brinkerhoff, D.J., Fürst, J.J., Gantayat, P., Gillet-Chaulet, F., Huss, M., Leclercq, P.W., Maurer, H., Morlighem, M., Pandit, A., Rabatel, A., Ramsankaran, R., Reerink, T.J., Robo, E., Rouges, E., Tamre, E., van Pelt, W.J.J., Werder, M.A., Azam, M.F., Li, H., & Andreassen, L.M.: Results from the Ice Thickness Models Intercomparison eXperiment Phase

610   2 (ITMIX2). Front. Earth. Sci., 8, https:// doi: 10.3389/feart.2020.571923, 2021

Farinotti, D., Huss, M., Bauder, A., Funk, M., & Truffer, M.: A method to estimate the ice volume and ice-thickness distribution of alpine glaciers. J. Glaciol., 55, 422-430, https://doi.org/10.3189/002214309788816759, 2009

Farinotti, D., Huss, M., Furst, J.J., Landmann, J., Machguth, H., Maussion, F., & Pandit, A.: A consensus estimate for the ice thickness distribution of all glaciers on Earth. Nat. Geosci., 12, 168-173, https://doi.org/10.1038/s41561-019-0300-3 ,2019

615   Fielding, E., Isacks, B., Barazangi, M., & Duncan, C.: How flat is Tibet? Geology, 22, 163-167, https://doi.org/10.1130/0091-7613(1994)022<0163:HFIT>2.3.CO;2, 1994

Frey, H., Machguth, H., Huss, M., Huggel, C., Bajracharya, S., Bolch, T., Kulkarni, A., Linsbauer, A., Salzmann, N., & Stoffel, M.: Estimating the volume of glaciers in the Himalayan–Karakoram region using different methods. The Cryosphere, 8, 2313-2333, https://doi.org/10.5194/tc-8-2313-2014, 2014

620   Frey, H., & Paul, F.: On the suitability of the SRTM DEM and ASTER GDEM for the compilation of topographic parameters in glacier inventories. Int. J. Appl. Earth. Obs., 18, 480-490, https://doi.org/10.1016/j.jag.2011.09.020, 2012

Frey, H., Paul, F., & Strozzi, T.: Compilation of a glacier inventory for the western Himalayas from satellite data: methods, challenges, and results. Remote Sens. Environ., 124, 832-843, https://doi.org/10.1016/j.rse.2012.06.020, 2012

Fujita, K., Suzuki, R., Nuimura, T., & Sakai, A.: Performance of ASTER and SRTM DEMs, and their potential for assessing
625   glacial lakes in the Lunana region, Bhutan Himalaya. J. Glaciol., 54, 220-228, https://doi.org/10.3189/002214308784886162, 2017

Furian, W., Loibl, D., & Schneider, C.: Future glacial lakes in High Mountain Asia: an inventory and assessment of hazard potential from surrounding slopes. J. Glaciol., 1-18, https://doi.org/10.1017/jog.2021.18, 2021

Gantayat, P., Kulkarni, A.V., & Srinivasan, J.: Estimation of ice thickness using surface velocities and slope: case study at
630   Gangotri Glacier, India. J. Glaciol., 60, 277-282, https://doi.org/10.3189/2014JoG13J078, 2014

Gardner, A.S., M. A. Fahnestock, and T. A. Scambos.: ITS_LIVE Regional Glacier and Ice Sheet Surface Velocities. Data archived at National Snow and Ice Data Center, 2019

Gdulová, K., Marešová, J., & Moudrý, V.: Accuracy assessment of the global TanDEM-X digital elevation model in a mountain environment. Remote Sens. Environ., 241, https://doi.org/10.1016/j.rse.2020.111724, 2020

635   Glen, J.W.: The Creep of Polycrystalline Ice. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 228, 519-538, https://doi.org/10.1098/rspa.1955.0066, 1955

González-Moradas, M.d.R., & Viveen, W.: Evaluation of ASTER GDEM2, SRTMv3.0, ALOS AW3D30 and TanDEM-X DEMs for the Peruvian Andes against highly accurate GNSS ground control points and geomorphological-hydrological metrics. Remote Sens. Environ., 237, https://doi.org/10.1016/j.rse.2019.111509, 2020

640   González, P.J., & Fernández, J.: Error estimation in multitemporal InSAR deformation time series, with application to Lanzarote, Canary Islands. J. Geophys. Res.-Sol. Ea., 116, https://doi.org/10.1029/2011JB008412, 2011

Gorokhovich, Y., & Voustianiouk, A.:. Accuracy assessment of the processed SRTM-based elevation data by CGIAR using field data from USA and Thailand and its relation to the terrain characteristics. Remote Sens. Environ., 104, 409-415, 2006

Grohmann, C.H.: Evaluation of TanDEM-X DEMs on selected Brazilian sites: Comparison with SRTM, ASTER GDEM and
645   ALOS AW3D30. Remote Sens. Environ., 212, 121-133, https://doi.org/10.1016/j.rse.2006.05.012, 2018

Han, H., Zeng, Q., & Jiao, J.: Quality Assessment of TanDEM-X DEMs, SRTM and ASTER GDEM on Selected Chinese Sites. Remote Sen.-Basel, 13, https://doi.org/10.3390/rs13071304, 2021

Hawker, L., Neal, J., & Bates, P.: Accuracy assessment of the TanDEM-X 90 Digital Elevation Model for selected floodplain sites. Remote Sens. Environ., 232, 111724-111738, https://doi.org/10.1016/j.rse.2019.111319, 2019

650 Höhle, J., & Höhle, M.: Accuracy assessment of digital elevation models by means of robust statistical methods. ISPRS J. Photogramm., 64, 398-406, https://doi.org/10.1016/j.isprsjprs.2009.02.003, 2009

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., & Kääb, A.: Accelerated global glacier mass loss in the early twenty-first century. Nature, 592, 726-731, https://doi.org/10.1038/s41586-021-03436-z, 2021

655 Huss, M.: Density assumptions for converting geodetic glacier volume change to mass change. The Cryosphere, 7, 877-887, Density assumptions for converting geodetic glacier volume change to mass change, 2013

Huss, M., & Farinotti, D.: Distributed ice thickness and volume of all glaciers around the globe. J. Geophys. Res.-Sol. Ea., 117, 1-10, https://doi.org/10.1029/2012JF002523, 2012

Huss, M., & Hock, R.: Global-scale hydrological response to future glacier mass loss. Nat. Clim. Change., 8, 135-140, 660 https://doi.org/10.1038/s41558-017-0049-x, 2018

Immerzeel, W.W., Lutz, A.F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., Hyde, S., Brumby, S., Davies, B.J., Elmore, A.C., Emmer, A., Feng, M., Fernandez, A., Haritashya, U., Kargel, J.S., Koppes, M., Kraaijenbrink, P.D.A., Kulkarni, A.V., Mayewski, P.A., Nepal, S., Pacheco, P., Painter, T.H., Pellicciotti, F., Rajaram, H., Rupper, S., Sinisalo, A., Shrestha, A.B., Viviroli, D., Wada, Y., Xiao, C., Yao, T., & Baillie, J.E.M.: Importance and vulnerability of the world's water towers. Nature, 665 577, 364-369, https://doi.org/10.1038/s41586-019-1822-y, 2020

Kääb, A.: Combination of SRTM3 and repeat ASTER data for deriving alpine glacier flow velocities in the Bhutan Himalaya. Remote Sens. Environ., 94, 463-474, https://doi.org/10.1016/j.rse.2004.11.003, 2005

Kääb, A., Leinss, S., Gilbert, A., Bühler, Y., Gascoin, S., Evans, S.G., Bartelt, P., Berthier, E., Brun, F., Chao, W.-A., Farinotti, D., Gimbert, F., Guo, W., Huggel, C., Kargel, J.S., Leonard, G.J., Tian, L., Treichler, D., & Yao, T.: Massive collapse of two 670 glaciers in western Tibet in 2016 after surge-like instability. Nat. Geosci., 11, 114-120, https://doi.org/10.1038/s41561-017-0039-7, 2018

Kaser, G., Grosshauser, M., & Marzeion, B.: Contribution potential of glaciers to water availability in different climate regimes. Proc. Natl. Acad. Sci. USA, 107, 20223-20227, https://doi.org/10.1073/pnas.1008162107, 2010

Ke, L., Ding, X., Zhang, L.E.I., Hu, J.U.N., Shum, C.K., & Lu, Z.: Compiling a new glacier inventory for southeastern 675 Qinghai–Tibet Plateau from Landsat and PALSAR data. J. Glaciol., 62, 579-592, https://doi.org/10.1017/jog.2016.58, 2016

Kienholz, C., Rich, J.L., Arendt, A.A., & Hock, R.: A new method for deriving glacier centerlines applied to glaciers in Alaska and northwest Canada. The Cryosphere, 8, 503-519, https://doi.org/10.5194/tc-8-503-2014, 2014

Koldtoft, I., Grinsted, A., Vinther, B., & Hvidberg, C.: Ice thickness and volume of the Renland Ice Cap, East Greenland. J. Glaciol., 67(264), 714-726. doi:10.1017/jog.2021.11, 2021

680 Kraaijenbrink, P.D.A., Bierkens, M.F.P., Lutz, A.F., & Immerzeel, W.W.: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. Nature, 549, 257-260, https://doi.org/10.1038/nature23878, 2017

Li, G., & Lin, H.: Recent decadal glacier mass balances over the Western Nyainqentanglha Mountains and the increase in their melting contribution to Nam Co Lake measured by differential bistatic SAR interferometry. Global Planet. Change, 149, 177-190, https://doi.org/10.1016/j.gloplacha.2016.12.018, 2017

685 Li, R., Li, H., Hao, T., Qiao, G., Cui, H., He, Y., Hai, G., Xie, H., Cheng, Y., and Li, B.: Assessment of ICESat-2 ice surface elevations over the Chinese Antarctic Research Expedition (CHINARE) route, East Antarctica, based on coordinated multi-sensor observations, The Cryosphere, 15, 3083–3099, https://doi.org/10.5194/tc-15-3083-2021, 2021.

Li, S., Yao, T., Yang, W., Yu, W., & Zhu, M.: Glacier Energy and Mass Balance in the Inland Tibetan Plateau: Seasonal and Interannual Variability in Relation to Atmospheric Changes. J. Geophys. Res.-Atmos., 123, 6390-6409,
690 https://doi.org/10.1029/2017JD028120, 2018

Linsbauer, A., Frey, H., Haeberli, W., Machguth, H., Azam, M.F., & Allen, S.: Modelling glacier-bed overdeepenings and possible future lakes for the glaciers in the Himalaya—Karakoram region. Ann. Glaciol., 57, 119-130, https://doi.org/10.3189/2016AoG71A627, 2016

Linsbauer, A., Paul, F., & Haeberli, W.: Modeling glacier thickness distribution and bed topography over entire mountain
695 ranges with GlabTop: Application of a fast and robust approach. J. Geophys. Res.-Sol. Ea., 117, 1-17, https://doi.org/10.1029/2011JF002313, https://doi.org/10.1016/j.geomorph.2019.04.012, 2012

Liu, K., Song, C., Ke, L., Jiang, L., Pan, Y., & Ma, R.: Global open-access DEM performances in Earth's most rugged region High Mountain Asia: A multi-level assessment. Geomorphology, 338, 16-26, https://doi.org/10.1016/j.geomorph.2019.04.012, 2019

700 Magruder, L. A., Brunt, K. M., and Alonzo, M.: Early ICESat-2 on-orbit geolocation validation using ground-based corner cube retroreflectors, Remote Sensing, 12, 3653, https://doi.org/10.3390/rs12213653, 2020.

Maurer, J.M., Schaefer, J.M., Rupper, S., & Corley, A. Acceleration of ice loss across the Himalayas over the past 40 years. Sci. Adv., 5, eaav7266, https://doi.org/10.1126/sciadv.aav7266, 2019

Maussion, F., Butenko, A., Champollion, N., Dusch, M., Eis, J., Fourteau, K., Gregor, P., Jarosch, A.H., Landmann, J.,
705 Oesterle, F., Recinos, B., Rothenpieler, T., Vlug, A., Wild, C.T., & Marzeion, B.: The Open Global Glacier Model (OGGM) v1.1. Geosci. Model. Dev., 12, 909-931, https://doi.org/10.5194/gmd-12-909-2019, 2019

Maussion, F., Scherer, D., Mölg, T., Collier, E., Curio, J., & Finkelnburg, R. Precipitation Seasonality and Variability over the Tibetan Plateau as Resolved by the High Asia Reanalysis. J. Climate, 27, 1910-1927, https://doi.org/10.1175/JCLI-D-13-00282.1, 2014

710 Mölg, N., Bolch, T., Rastner, P., Strozzi, T., & Paul, F.: A consistent glacier inventory for Karakoram and Pamir derived from Landsat data: distribution of debris cover and mapping challenges. Earth Syst. Sci. Data, 10, 1807-1827, https://doi.org/10.5194/essd-10-1807-2018, 2018

Mukherjee, S., Joshi, P.K., Mukherjee, S., Ghosh, A., Garg, R.D., & Mukhopadhyay, A. Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). Int. J. Appl. Earth. Obs., 21, 205-217,
715 https://doi.org/10.1016/j.jag.2012.09.004, 2013

Nuth, C., & Kääb, A.: Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. The Cryosphere, 5, 271-290, https://doi.org/10.5194/tc-5-271-2011, 2011

Paul, F., & Linsbauer, A.: Modeling of glacier bed topography from glacier outlines, central branch lines, and a DEM. Int. J. Geogr. Inf. Sci., 26, 1173-1190, https://doi.org/10.1080/13658816.2011.627859, 2012

720 Pelto, B.M., Maussion, F., Menounos, B., Radić, V., & Zeuner, M.: Bias-corrected estimates of glacier thickness in the Columbia River Basin, Canada. J. Glaciol., 66, 1051-1063, https://doi.org/10.1017/jog.2020.75, 2020

Ramsankaran, R., Pandit, A., & Azam, M.: Spatially distributed ice-thickness modelling for Chhota Shigri Glacier in western Himalayas, India. Int. J. Remote Sens., 39, 3320-3343, https://doi.org/10.1080/01431161.2018.1441563, 2018

Rankl, M., & Braun, M.: Glacier elevation and mass changes over the central Karakoram region estimated from TanDEM-X
725 and SRTM/X-SAR digital elevation models. Ann. Glaciol., 57, 273-281, https://doi.org/10.3189/2016AoG71A024, 2016

Reuter, H.I., Nelson, A., & Jarvis, A.: An evaluation of void-filling interpolation methods for SRTM data. Int. J. Geogr. Inf. Sci., 21, 983-1008, https://doi.org/10.1080/13658810601169899, 2007

Rignot, E., Echelmeyer, K., & Krabill, W.: Penetration depth of interferometric synthetic-aperture radar signals in snow and ice. Geophys Res Lett, 28, 3501-3504, https://doi.org/10.1029/2000gl012484

730 Scambos, T., M. Fahnestock, T. Moon, A. Gardner, and M. Klinger.: Global Land Ice Velocity Extraction from Landsat 8 (GoLIVE), Version 1. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center, 2016

Shean, D.E.: High Mountain Asia 8-meter DEM Mosaics Derived from Optical Imagery, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, 2017

Shean, D.E., Bhushan, S., Montesano, P., Rounce, D.R., Arendt, A., & Osmanoglu, B.: A Systematic, Regional Assessment
735 of High Mountain Asia Glacier Mass Balance. Front. Earth. Sci., 7, https://doi.org/10.3389/feart.2019.00363, 2020

Shortridge, A., & Messina, J.: Spatial structure and landscape associations of SRTM error. Remote Sens. Environ., 115, 1576-1587, https://doi.org/10.1016/j.rse.2011.02.017, 2011

Smith, B., Fricker, H.A., Holschuh, N., Gardner, A.S., Adusumilli, S., Brunt, K.M., Csatho, B., Harbeck, K., Huth, A., Neumann, T., Nilsson, J., & Siegfried, M.R.: Land ice height-retrieval algorithm for NASA's ICESat-2 photon-counting
740 laser altimeter. Remote Sens. Environ., 233, https://doi.org/10.1016/j.rse.2019.111352, 2019

Takaku, J., Tadono, T., Doutsu, M., Ohgushi, F., & Kai, H.: Updates of 'Aw3d30' Alos Global Digital Surface Model with Other Open Access Datasets. Int. Arch. Photogramm., XLIII-B4-2020, 183-189, https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-183-2020, 2020

Thompson, L.G., Yao, T., Davis, M.E., Mosley-Thompson, E., Wu, G., Porter, S.E., Xu, B., Lin, P.-N., Wang, N., Beaudon,
745 E., Duan, K., Sierra-Hernández, M.R., & Kenny, D.V.: Ice core records of climate variability on the Third Pole with emphasis on the Guliya ice cap, western Kunlun Mountains. Quaternary. Sci. Rev., 188, 1-14, 2018

Uuemaa, E., Ahi, S., Montibeller, B., Muru, M., & Kmoch, A.: Vertical Accuracy of Freely Available Global Digital Elevation Models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM). Remote Sens.-Basel, 12, https://doi.org/10.1016/j.quascirev.2018.03.003, https://doi.org/10.1016/j.rse.2007.11.003, 2020

750 Van Niel, T.G., McVicar, T.R., Li, L., Gallant, J.C., & Yang, Q.: The impact of misregistration on SRTM and DEM image differences. Remote Sens. Environ., 112, 2430-2442, https://doi.org/10.1016/j.rse.2007.11.003, 2008

Welty, E., Zemp, M., Navarro, F., Huss, M., Fürst, J.J., Gärtner-Roer, I., Landmann, J., Machguth, H., Naegeli, K., Andreassen, L.M., Farinotti, D., & Li, H.: Worldwide version-controlled database of glacier thickness observations. Earth System Science Data, 12, 3039-3055, https://doi.org/10.5194/essd-12-3039-2020, 2020

755 Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., & Roth, A.: Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data. ISPRS J. Photogramm., 139, 171-182, https://doi.org/10.1016/j.isprsjprs.2018.02.017, 2018

Wu, K., Liu, S., Zhu, Y., Liu, Q., & Jiang, Z.: Dynamics of glacier surface velocity and ice thickness for maritime glaciers in the southeastern Tibetan Plateau. J. Hydrol., 590, https://doi.org/10.1016/j.jhydrol.2020.125527, 2020

760 Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., & Bates, P.D.: A high-accuracy map of global terrain elevations. Geophys. Res. Lett., 44, 5844-5853, https://doi.org/10.1002/2017GL072874, 2017

Yao, T., Thompson, L.G., Mosbrugger, V., Zhang, F., Ma, Y., Luo, T., Xu, B., Yang, X., Joswiak, D.R., Wang, W., Joswiak, M.E., Devkota, L.P., Tayal, S., Jilani, R., & Fayziev, R.: Third Pole Environment (TPE). Environ. Dev., 3, 52-64, 2012

765 Zhang, G., Bolch, T., Allen, S., Linsbauer, A., Chen, W., & Wang, W.: Glacial lake evolution and glacier–lake interactions in the Poiqu River basin, central Himalaya, 1964–2017. J. Glaciol., 65, 347-365, https://doi.org/10.1016/j.envdev.2012.04.002, 2019

Zhang, G., Yao, T., Xie, H., Kang, S., & Lei, Y.: Increased mass over the Tibetan Plateau: From lakes or glaciers? Geophys. Res. Lett., 40, 2125-2130, https://doi.org/10.1002/grl.50462, 2013

770 Zhang, Y., Pang, Y., Cui, D., Ma, Y., & Chen, L.: Accuracy Assessment of the ICESat-2/ATL06 Product in the Qilian Mountains Based on CORS and UAV Data. IEEE J-STARS, 14,1558-1571, https://doi.org/10.1109/JSTARS.2020.3044463, 2020

Zheng, G., Allen, S.K., Bao A. M., Juan A. B., Matthias H., Zhang G. Q., Li J. L., Ye Y., Jiang L. L., Yu T., Chen W.F., & Stoffel, M.: Increasing risk of glacial lake outburst floods from future Third Pole deglaciation. Nat. Clim. Change., 11, 411–
775 417, https://doi.org/10.1038/s41558-021-01028-3, 2021

Zhou, Y., Li, Z., Li, J., Zhao, R., & Ding, X.: Glacier mass balance in the Qinghai–Tibet Plateau and its surroundings from the mid-1970s to 2000 based on Hexagon KH-9 and SRTM DEMs. Remote Sens. Environ., 210, 96-112, https://doi.org/10.1016/j.rse.2018.03.020, 2018