

LETTER • **OPEN ACCESS**

Validation of FABDEM, a global bare-earth elevation model, against UAV-lidar derived elevation in a complex forested mountain catchment

To cite this article: Christopher B Marsh *et al* 2023 *Environ. Res. Commun.* **5** 031009

View the [article online](#) for updates and enhancements.

You may also like

- [Predicting Autism Spectrum Disorder \(ASD\) for Toddlers and Children Using Data Mining Techniques](#)
Rasool Azeem Musa, Mehdi Ebady Manaa and Ghassan Abdul-Majeed
- [Intra-molecular chaperone: the role of the N-terminal in conformational selection and kinetic control](#)
Chung-Jung Tsai, Buyong Ma and Ruth Nussinov
- [On the spectra of quenched random perturbations of partially expanding maps on the torus](#)
Yushi Nakano and Jens Wittsten

Environmental Research Communications



LETTER

Validation of FABDEM, a global bare-earth elevation model, against UAV-lidar derived elevation in a complex forested mountain catchment

OPEN ACCESS

RECEIVED
2 November 2022REVISED
28 February 2023ACCEPTED FOR PUBLICATION
17 March 2023PUBLISHED
29 March 2023Christopher B Marsh , Phillip Harder and John W Pomeroy

Centre for Hydrology, University of Saskatchewan, Canada

E-mail: chris.marsh@usask.ca**Keywords:** vegetation-free digital elevation model, UAV-lidar, Copernicus-30, FABDEM, bare earth validation

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Abstract

Space-based, global-extent digital elevation models (DEMs) are key inputs to many Earth sciences applications. However, many of these applications require the use of a 'bare-Earth' DEM versus a digital surface model (DSM), the latter of which may include systematic positive biases due to tree canopies in forested areas. Critical topographic features may be obscured by these biases. Vegetation-free datasets have been created by using statistical relationships and machine learning to train on local-scale datasets (e.g., lidar) to de-bias the global-extent datasets. Recent advances in satellite platforms coupled with increased availability of computational resources and lidar reference products has allowed for a new generation of vegetation- and urban-canopy removals. One of these is the Forest And Buildings removed Copernicus DEM (FABDEM), based on the most recent and most accurate global DSM Copernicus-30. Among the more challenging landscapes to quantify surface elevations are densely forested mountain catchments, where even airborne lidar applications struggle to capture surface returns. The increasing affordability and availability of UAV-based lidar platforms have resulted in new capacity to fly modest spatial extents with unrivalled point densities. These data allow an unprecedented ability to validate global sub-canopy DEMs against representative UAV-based lidar data. In this work, the FABDEM is validated against up-scaled lidar data in a steep and forested mountain catchment considering elevation, slope, and Terrain Position Index (TPI) metrics. Comparisons of FABDEM with SRTM, MERIT, and the Copernicus-30 dataset are made. It was found that the FABDEM had a 24% reduction in elevation RMSE and a 135% reduction in bias compared to the Copernicus-30 dataset. Overall, the FABDEM provides a clear improvement over existing deforested DEM products in complex mountain topography such as the MERIT DEM. This study supports the use of FABDEM in forested mountain catchments as the current best-in-class data product.

1. Introduction

Space-based, global-extent digital elevation models (DEMs) are key inputs to many Earth sciences including flood hazard mapping (Sampson *et al* 2016, Garrote 2022), hydrological modelling (O'Callaghan and Mark 1984, Beven and Freer 2001, Gharari *et al* 2020), shortwave radiation estimates (Garnier and Ohmura 1968, Dozier and Frew 1990, Marsh *et al* 2012), landslide hazard assessment (Fenton *et al* 2013), ecology (Illés *et al* 2011, Moudrý *et al* 2018), geomorphology (Liu *et al* 2009), and landform classification (Reu *et al* 2013). However, many of these applications require the use of a 'bare-Earth' digital elevation model (DEM) versus a digital surface model (DSM), the latter of which may include systematic positive biases due to tree canopies in forested areas (O'Loughlin *et al* 2016). These biases occur as vegetation obscures and/or obstructs the Earth's surface from most space-based remote sensing techniques (Magruder *et al* 2021) resulting in the

largest errors over heavily vegetated surfaces (Carabajal and Harding 2005). Critical topographic features such as slopes, river networks, and ecological patterns can be obscured by these biases (Yamazaki *et al* 2017). As many researchers and countries continue to rely on space-based elevation datasets in either totality or to complement local high-resolution datasets (Yamazaki *et al* 2017), there is large motivation to use DEMs that are free of vegetation and urban canopy biases.

The Shuttle Radar Topography Mission (SRTM, Farr *et al* (2007)) DSM is one of the most widely used global DSM datasets (Uemaa *et al* 2020) and has generated many derivative products. Vegetation-free DEM datasets have been created by using statistical relationships and machine learning to train on local-scale datasets (e.g., lidar) to de-bias the global-scale datasets by removing the impact of vegetation. Such approaches have resulted in many SRTM-derived products such as O’Loughlin *et al* (2016), Zhao *et al* (2018), Magruder *et al* (2021), CoastalDEM (Kulp and Strauss 2018), and MERIT (Yamazaki *et al* 2017). Despite their widespread use and global extent, these datasets can be difficult to apply to environmental studies due to artefacts being present and are often related to void processing despite significant post-processing (Hirt 2018, Garrote 2022). These errors can then be magnified in steep, forested mountain catchments with complex topography (Hirt 2018). Source dataset errors can also propagate to derivative datasets, e.g., Hirt (2018). Recent advances in new satellite platforms have led to improved base DSM datasets. The recently released Copernicus-30 (COP30) TanDEM-X X-band global radar 1’ (≈ 30 m) DSM (AIRBUS 2020) is currently considered one of the most accurate global DSM (Gesch 2018, Grohmann 2018, Guth and Geoffroy 2021, Hawker *et al* 2022, Garrote 2022). Coupled with improved geospatial reference frames (Goyal *et al* 2022), increased availability of computational resources, and lidar reference products, it has allowed for a new generation of vegetation- and urban-canopy removals. The COP30 derived The Forest And Buildings removed Copernicus DEM (FABDEM) (Hawker *et al* 2022) is an example of this and represents a new generation of DEM that may be of substantial interest to global environmental studies.

The increasing affordability and availability of Unoccupied Aerial Vehicle (UAV)-based lidar platforms have resulted in new capacity to fly modest spatial extents with an unprecedented lidar point density (Kucharczyk *et al* 2018). This density greatly increases the number of sub-canopy laser returns, allowing for better constraining the sub-canopy topography (Harder *et al* 2020). Therefore, these data allow a novel opportunity to validate global sub-canopy DEMs against representative UAV-based lidar data. Although there is no expectation of matching a high-resolution UAV-lidar dataset, upscaling the lidar data to match the global datasets a ‘true’ dataset with which to validate them is possible.

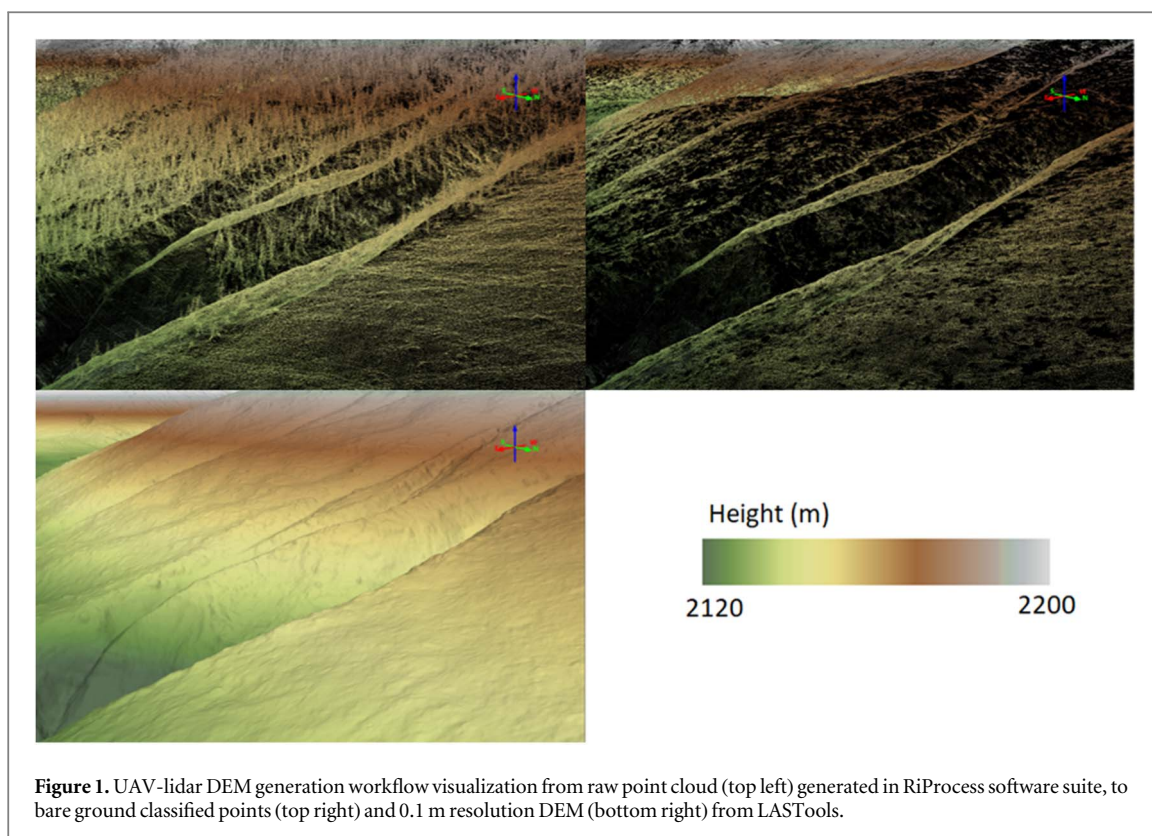
Globally, mountain headwaters provide fresh water for many billions of lives (Viviroli *et al* 2020) and, as a result, substantial effort is invested in quantifying hydrological, ecological, and geomorphological change in these regions. However, the complex topography present in mountain headwater catchments represents a mix of steep slopes and forested surfaces that could be problematic with deforestation algorithms. In this work, the FABDEM is validated against a UAV-lidar dataset in a steep, forested mountain catchment. Comparisons with SRTM, MERIT, and the Copernicus-30 datasets are also made due to their historical use, to place the comparison in the context of existing datasets, and align with the accuracy assessment of the original FABDEM manuscript (Hawker *et al* 2022). End-users and practitioners have tremendous interest in knowing how accurate vegetation-removed datasets are, especially in the context of other widely used datasets, e.g., SRTM. Specifically, this study investigates in a steep, forested mountain basin: a) how much error is present in the elevation data of the FABDEM, b) how much error is present in key landscape attributes of the FABDEM, and c) are the distributions of these landscape attributes well represented in the FABDEM?

2. Methodology

Digital datasets that describe the elevation of the Earth’s surface may either include the natural and built surface features (e.g., vegetation, buildings) or may only include the Earth’s surface. The former of these are referred to as Digital Surface Models (DSM) and the latter are Digital Elevation Models (DEM) that are ‘bare Earth’ representations of the surface. Herein, these terms are used to distinguish between the datasets. When both DEM and DSM are referred to collectively, the term ‘dataset’ is used.

2.1. Data

To evaluate the accuracy of the FABDEM DEM, a reference DEM is needed. This reference DEM was acquired from a UAV-based lidar platform. Due to the slow and low flight of a UAV platform, compared to other remote sensing platforms, a significantly higher number of lidar returns are available (> 100 pts m^{-2}) and therefore it is possible to more accurately characterize the sub-canopy topography. FABDEM was also compared to the hydrologically conditioned, vegetation-removed, MERIT DEM. The MERIT DEM was included due to its widespread use, as well as to provide a benchmark sub-canopy dataset to compare FABDEM. As MERIT and



FABDEM are derivative from the SRTM and Copernicus-30 DSMs respectively, both SRTM and Copernicus-30 datasets were included in the comparison. Other SRTM-derivative datasets, such as NASADEM, are not considered in this comparison. This is firstly due to focusing on vegetation-removed DEMs (i.e., further comparisons of DSMs against a lidar sub-canopy DEM are not particularly instructive). Secondly, there is a strong motivation to consider near-global datasets. Many datasets, e.g., NASADEM, exclude much of northern Canada, Alaska, Norway, Sweden, Finland, and Russia. These are rugged areas with significant boreal forest canopy coverage and are areas with users who would be primarily interested in DEMs such as FABDEM and MERIT.

2.1.1. UAV-Lidar

The Fortress Mountain Research Basin lidar data was acquired in late summer (September 2018) with a Riegl miniVUX1-UAV lidar with an integrated Applanix APX-15 inertial measurement unit (IMU) mounted on a DJI M600 Pro UAV platform. Complete information on the data collection, processing, and accuracy assessment, including data and code, for the UAV-lidar DEM generation, can be found in Harder *et al* (2020). The area of interest was mapped in a sequence of terrain following flights at $7\text{ m} \cdot \text{s}^{-1}$, 110 m above the ground surface, and with 80 m parallel line spacing. The highest elevations of the basin area >2500 m.a.s.l. were not surveyed due to visual line-of-sight limitations and logistical challenges. Absolute positioning of the platform was facilitated with the logging of Global Navigation Satellite System (GNSS) data over the course of the campaign with a Leica GS16 base station located within the flight area and precise point positioning with post-processing with the Natural Resources Canada PPP system (Natural Resources Canada 2023). Post-processing (see Harder *et al* (2020) for more details) integrated IMU and GNSS base station data within the Applanix POSPAC UAV software to generate trajectory data of the UAV-lidar system. The RiProcess software suite, propriety Riegl software, integrated and optimized the trajectory and laser scan data to generate a raw lidar point cloud (figure 1, top left). The raw point cloud was classified into surface and vegetation points (figure 1, top right), with the `lasground_new` function of LASTools (Isenburg 2023) which differentiated ground surface points from vegetation with an adaptive TIN (Triangular Irregular Network) algorithm adapted from Axelsson (2000). A 0.1 m resolution DEM (figure 1, bottom left) and canopy height model (CHM) were generated with the `blast2dem` and `lascanopy` functions, respectively, in LASTools. Uncertainty of the DEM was assessed with respect to 119 surface elevations observed with a Leica GS16 RTK rover-base GNSS system. Vertical differences between the manual survey and the UAV-lidar DEM were found to have Root Mean Squared Error (RMSE) of 0.11 m, 0.13 m, and 0.14 m for open, shrub, and forested surfaces respectively (Harder *et al* 2020).

2.1.2. FABDEM

The Forest And Buildings removed Copernicus DEM (FABDEM) version 1.0 (Hawker *et al* 2022) is a 1' (≈ 30 m) DEM free of buildings and forests (Hawker *et al* 2022). It is derived from the Copernicus GLO-30 (herein COP30) (AIRBUS 2020) DSM by using a random forest machine learning to remove vegetation and urban buildings from the COP30 DSM. Global forest canopy height training data from the Global Ecosystem Dynamics Investigation (GEDI) (Dubayah *et al* 2020) derived Global Forest Canopy Height 2019 data-set (Potapov *et al* 2021) was used for latitudes between 52° N and 52° S. For latitudes outside the GEDI extent, ICESat-2 (Ice, Cloud, and land Elevation Satellite) (Neuenschwander *et al* 2021) laser altimetry canopy heights were used. In both cases, the Copernicus Global Land Service collection 3 epoch 2015 canopy coverage was used to constrain the forest canopy heights. Forests were removed where both the forest height is greater than 3 m and tree cover was greater than 10% (Hawker *et al* 2022).

2.1.3. MERIT

The MERIT (Yamazaki *et al* 2017) is a global 3' (≈ 90 m) DEM derived from the SRTM 3' DSM and has corrected the speckle noise, stripe noise, and tree height bias in the SRTM dataset. Vegetation biases were removed as a function of tree density and tree height (Yamazaki *et al* 2017). Tree height was determined from the space-borne laser altimetry ICESat (Ice, Cloud, and land Elevation Satellite) (Harding and Carabajal 2005) with a global tree density map (Hansen *et al* 2013) and tree height map (Simard *et al* 2011). Although this vegetation-removed DEM improved on the flat and lowlands, there is likely low accuracy in mountainous areas due to steep slopes (Yamazaki *et al* 2017).

2.1.4. SRTM

The Making Earth System Data Records for Use in Research Environments (MEaSUREs) Shuttle Radar Topography Mission (SRTM) is a global C-band radar 1' (≈ 30 m) DSM (Farr *et al* 2007). It was acquired in 2000. Specifically, the version 3, void-filled SRTMGL1NV003 (<https://lpdaac.usgs.gov/products/srtmgl1nv003/>) product was used in this study.

2.1.5. Copernicus-30

The Copernicus-30 (COP30) TanDEM-X X-band global radar 1' (≈ 30 m) DSM (AIRBUS 2020). It is currently considered the most recent and most accurate global DSM (Gesch 2018, Grohmann 2018, Guth and Geoffroy 2021, Hawker *et al* 2022). Not only does COP30 include more recent topography, but it also uses the Earth Gravitational Model 2008 (EGM2008) geoid model (Pavlis *et al* 2012) as compared to EGM1996 as used by the other datasets. EGM2008 is substantially improved over EGM1996 and Goyal *et al* (2021) reported differences between the two up to 12 m between the two models. The COP30 dataset is an ESA/Airbus product acquired between 2010 and 2015.

2.2. Study region

The study region is a 4.74 km^2 domain, shown in red in figure 2. It is situated almost entirely within the Fortress Mountain Research Basin (FMRB), located in the Kananaskis River Valley, Alberta, Canada. In figure 2 the black outline is the FMRB proper, the red outline is the lidar acquisition study region, and the blue line is Fortress Creek. Shown in green is the vegetation coverage greater than 3 m, derived from Potapov *et al* (2021) for visualization purposes. Topography from the Copernicus-30 DSM dataset is shown in greyscale. The FMRB is representative of the heavily forested, subalpine, rugged mountain terrain of the Canadian Rockies. Each dataset is shown in figure 3. The basin ranges in elevation from 2000 m to 2900 m (study area: 1936 m to 2348 m) and comprises forested slopes transitioning into above-tree-line alpine (Harder *et al* 2020). The FMRB forests are primarily composed of fir and spruce ranging from 4 m in height in the younger stands to ≈ 9.5 m in the mature stands (Langs *et al* 2021) transitioning to larches near the treeline, this is shown in figure 4. The distribution of canopy height is shown in figure 5. The distribution is split for vegetation ≤ 3 m and > 3 m, corresponding to the threshold for the vegetation-removal algorithm in FABDEM (see data section for more details).

2.3. Processing

The lidar dataset was reprojected to geographic WGS84 from UTM 11N and a bounding box of the irregular lidar domain was computed. The COP30, MERIT, SRTM, and FABDEM datasets were then extracted to this rectangular subdomain. As each dataset is aligned to a different grid, each dataset was regridded to a common alignment (defined by the lidar dataset) and resolution (30 m) using bilinear interpolation (2×2 window) via the QGIS Align Rasters functionality. This was the only non-automated step in this workflow. These aligned datasets were then clipped to match the irregular domain of the lidar dataset (red outline in figure 2) using the gdalwarp tool from the Geospatial Data Abstraction Library (GDAL; GDAL/OGR contributors 2022). Once

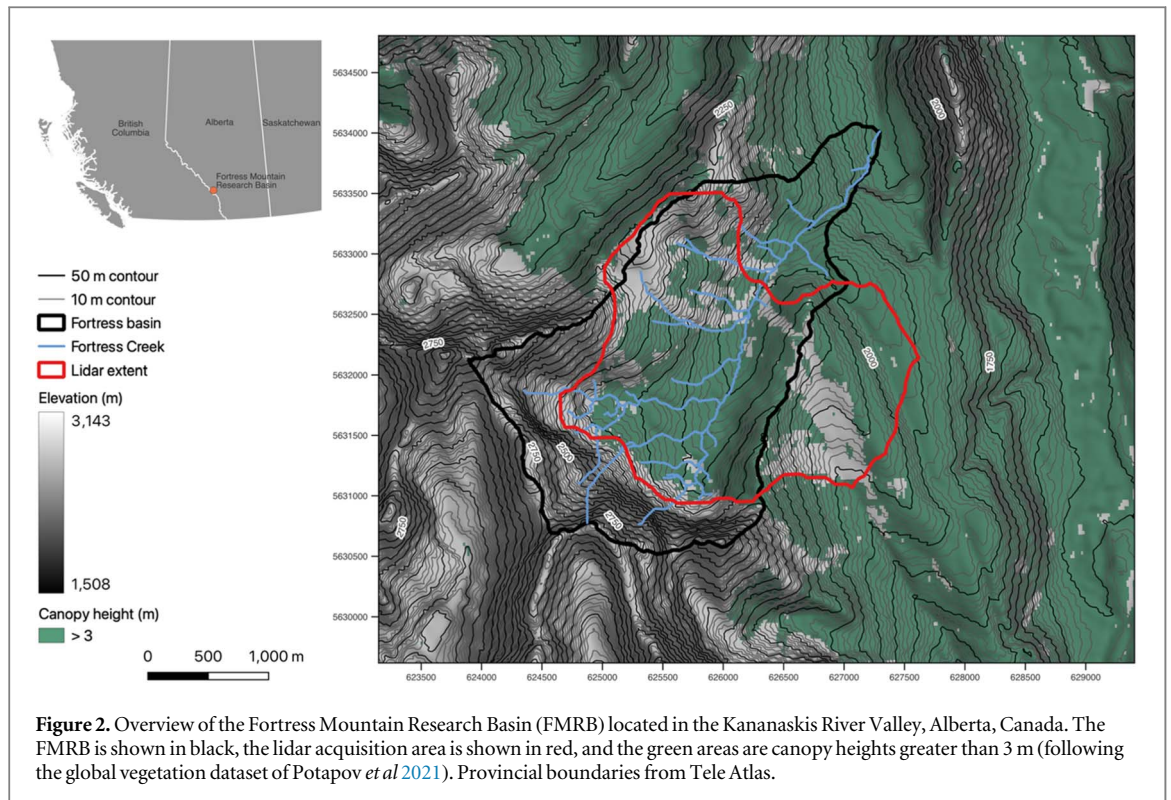
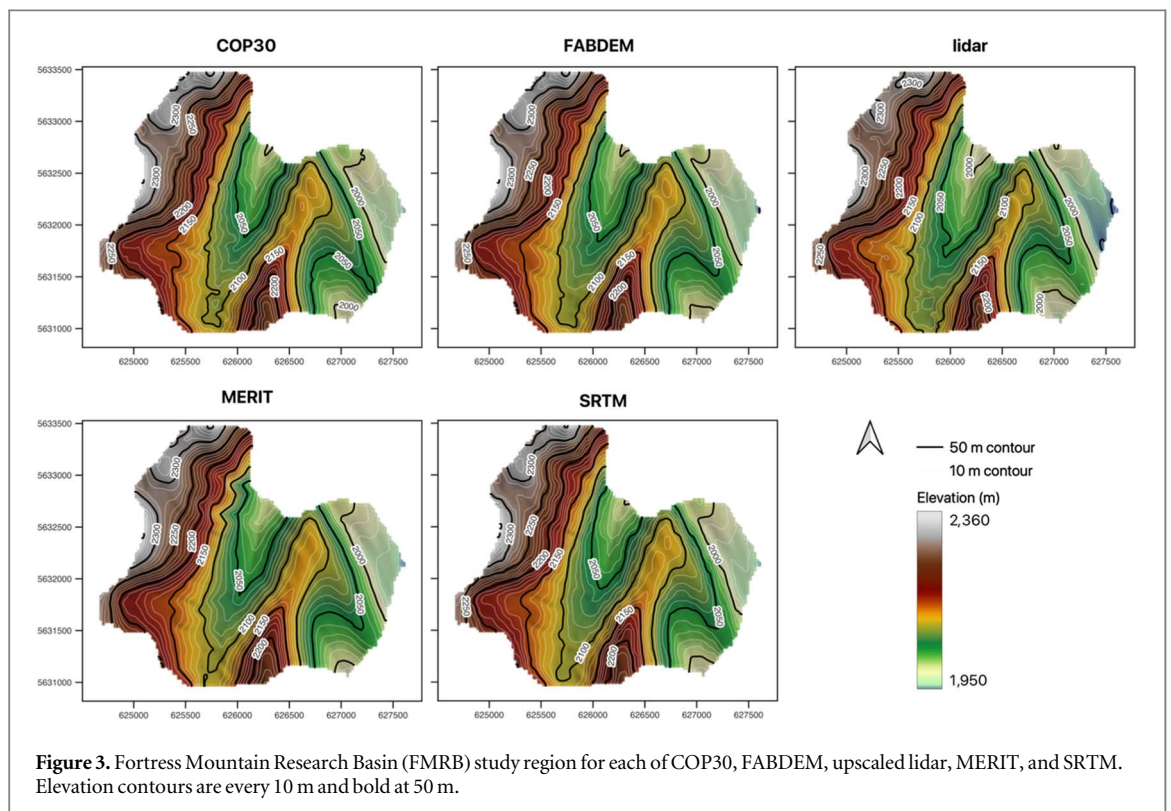


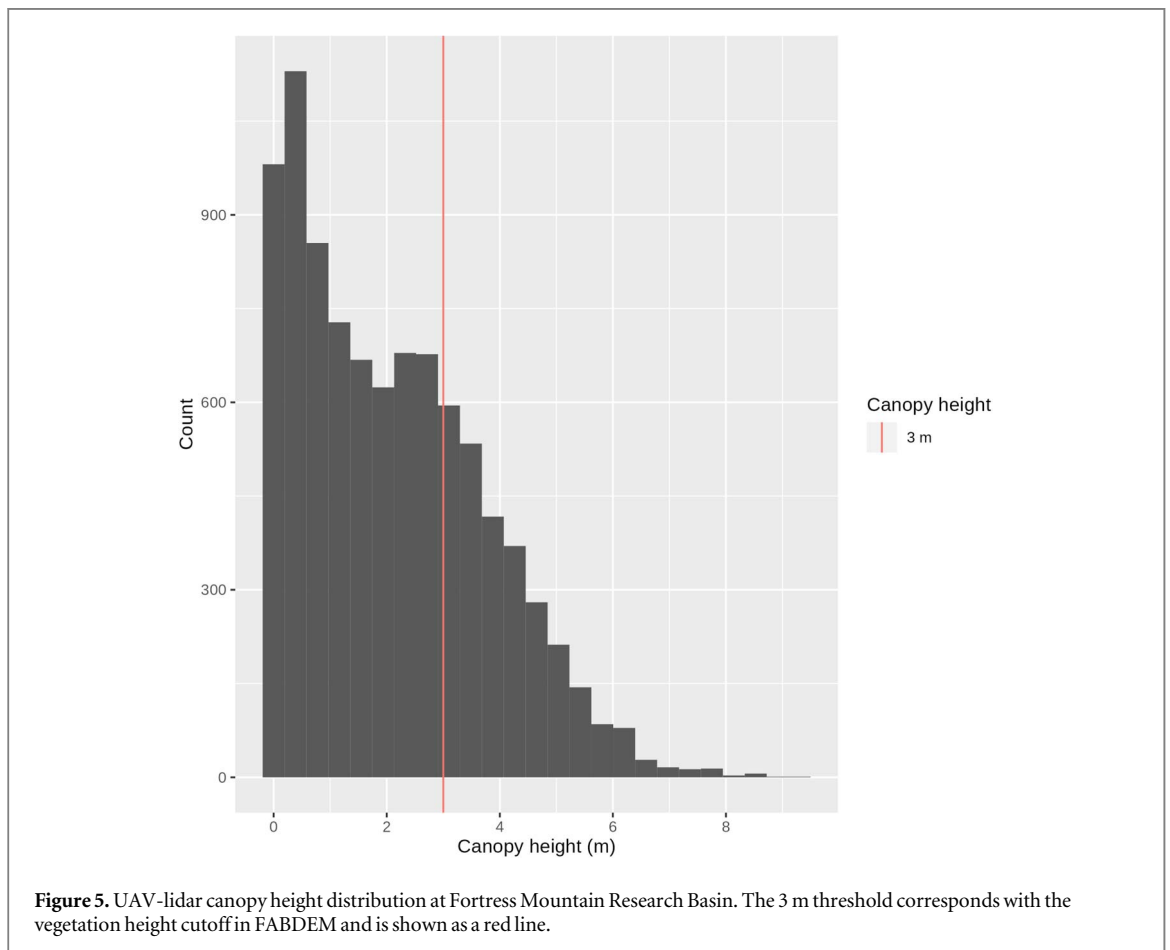
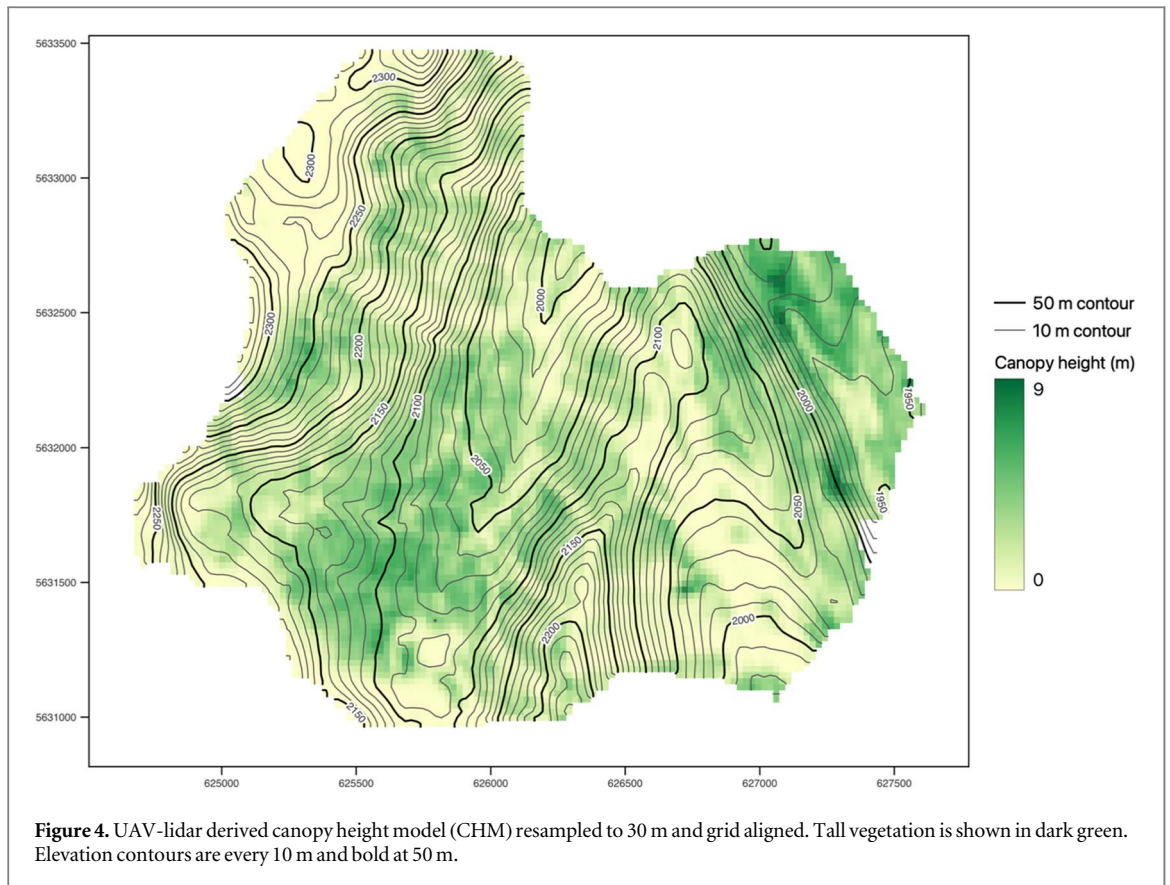
Figure 2. Overview of the Fortress Mountain Research Basin (FMRB) located in the Kananaskis River Valley, Alberta, Canada. The FMRB is shown in black, the lidar acquisition area is shown in red, and the green areas are canopy heights greater than 3 m (following the global vegetation dataset of Potapov *et al* 2021). Provincial boundaries from Tele Atlas.



aligned and clipped, all the aligned datasets were projected to EPSG 32611 (UTM zone 11N) to be able to compute secondary landscape characteristics. These characteristics were Terrain Position Index (TPI) and slope, both generated using the gdaldem GDAL tool.

2.4. Landscape attributes

The complex topography present in mountain headwater catchments represents a mix of steep slopes and forested surfaces that could prove to be problematic with deforestation algorithms. Topographic indices derived



from remote sense elevation data are commonly used as surrogates for field measurements (Moudrý *et al* 2018). In this study, two landscape attributes were computed for each input elevation dataset and compared to those computed from the lidar. Although an almost infinite set of possible elevation-derived landscape characteristics could be compared, slope and TPI were selected due to their common usage in the hydrological and environmental sciences.

Slope was selected as it is the first derivative of the elevation field. The use of slope is widespread in many Earth science investigations such as in the: simulation of subsurface flow (Beven 1981; Beven and Freer 2001), impact of topography on near-surface wind fields (Ryan 1977, Wagenbrenner *et al* 2019, Dujardin and Lehning 2022), estimation of avalanches (Bernhardt and Schulz 2010), impact of steep terrain on incident shortwave radiation (Garnier and Ohmura 1968, Dozier and Frew 1990, Marsh *et al* 2012), use in hydrological modelling (Kienzle 2011, Gharari *et al* 2020), simulation of blowing snow (Marsh *et al* 2019), estimation of ice field katabatic winds (Conway *et al* 2021), and landslide hazard assessment (Fenton *et al* 2013). The correct estimation of slope in complex terrain is critical when distributing meteorological inputs (Kienzle 2011) or comparing model output to remotely sensed data (Vionnet *et al* 2021). It is a key secondary characteristic of the landscape that should be estimated correctly. The slope was computed using the gdaldem slope tool with default parameters.

The Topographic Position Index (TPI) is a landscape classification index (Weiss 2001) that measures the relative topographic position of a point as a difference between its elevation and the mean elevation within a predetermined neighbourhood (Weiss 2001, Reu *et al* 2013). TPI is scale-dependent and is highly influenced by the quality of the DEM. By thresholding these TPI values, a landscape can be classified into discrete slope position classes and combined with elevation to produce landforms (Weiss 2001). TPI has had substantial use in, for example, geomorphology (Liu *et al* 2009), soil mapping (Illés *et al* 2011), watershed metrics (Weiss 2001), and ecology (Giroday *et al* 2011). A comprehensive summary can be found in Reu *et al* (2013). TPI is a key landscape attribute that has broad interdisciplinary application. TPI was calculated via the GDAL TPI function using the default (and non-configurable) search radius of the surrounding cells.

2.5. Comparison metrics

The difference between the lidar reference elevation dataset and the candidate elevation dataset, i.e., COP30, FABDEM, SRTM, or MERIT, is described by one of three metrics: difference, root mean squared error, and mean bias. These are described briefly below. Each method is used for elevation, slope, and TPI comparisons.

The difference (δDEM) between the lidar reference dataset ('lidar') and the candidate dataset ('cDEM'), i.e., COP30, FABDEM, SRTM, or MERIT is given by:

$$\delta DEM = \text{lidar} - \text{cDEM}. \quad (1)$$

The Root Mean Squared Error (RMSE) is given as:

$$RMSE = \left(\frac{\sum_1^n (\text{cDEM} - \text{lidar})^2}{n} \right)^{1/2}. \quad (2)$$

The Mean Bias (MB) is given as:

$$MB = \frac{1}{n} \sum_1^n (\text{lidar} - \text{cDEM}). \quad (3)$$

3. Results

3.1. Spatial differences

3.1.1. Elevation

The aligned datasets were compared spatially to the lidar data. The dataset differences (per equation (1)) are shown in figure 6 with Root Mean Squared Errors (RMSE) and Mean Biases summarized in table 1.

The COP30 dataset had the largest negative bias (-1.933 m) and a RMSE of 2.582 m. The error was generally a uniform height bias, with the East-facing slope and Easternmost part of the domain having the largest errors. These two regions correspond to high canopy coverage. The FABDEM had a smaller bias (0.685 m) and RMSE (1.953 m) indicating a slight underestimation of topography likely due to an overestimation in tree height during the canopy removal. The COP30 regions that had the largest error were substantially reduced in the FABDEM. There was a slight positive bias along the Westward facing slope in the middle of the domain. The MERIT dataset had substantial positive and negative biases throughout the domain, with significant over and underestimations present, exceeding 50 m. These areas were in the West-most and northern portions of the basin and were associated with steep slopes and gullies. MERIT had a bias of -1.137 m and an RMSE of 5.095 m. This was the

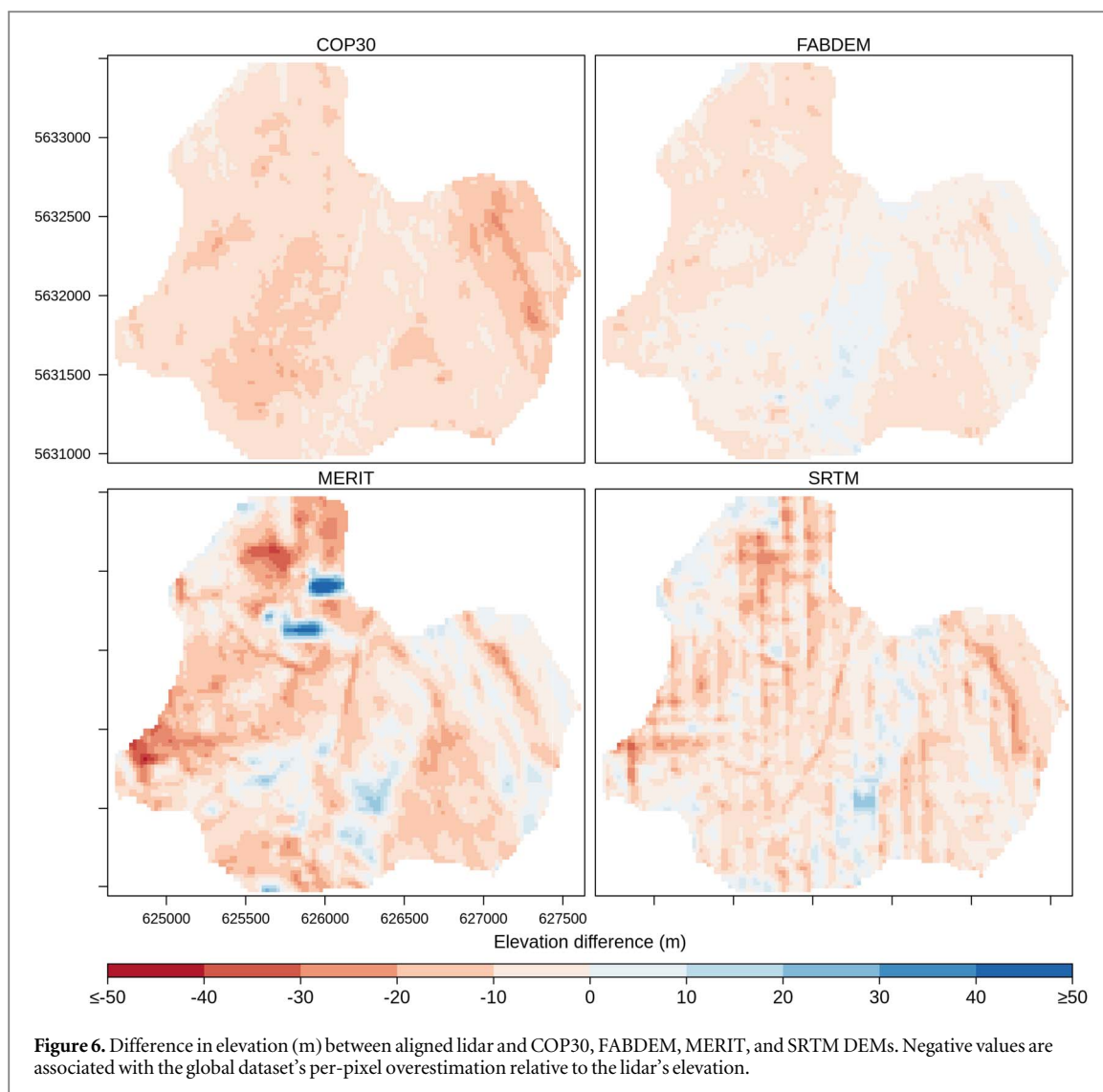


Table 1. Elevation RMSE and mean bias error metrics for the datasets versus the lidar. Negative values are associated with the global dataset's overall overestimation relative to the lidar's elevation.

DataSet	DEM	RMSE (m)	Mean Bias (m)
Elevation	COP30	2.582	-1.933
	FABDEM	1.953	0.685
	MERIT	5.095	-1.137
	SRTM	3.556	-0.326

largest RMSE of the compared DEMs and the second-largest bias. The regions of positive height biases in the southern part of the domain were, paradoxically, present in the locations of the highest vegetation coverage. The SRTM dataset shows the most peculiar results. Substantial striping was present in primarily a North-South orientation.

3.1.2. Slope

Slopes of the aligned datasets were compared spatially to the lidar-derived slope. The slope difference (per equation (1)) is shown in figure 7 with Root Mean Squared Errors (RMSE) and Mean Biases summarized in table 2.

The COP30 and FABDEM slope differences tended to be quite similar, and areas of large slope difference were co-located with the areas with steep gradients and in the transition areas between higher and lower error, as seen in figure 6. The COP30 RMSE (1.122°) and bias (0.121°) are slightly smaller than that of FABDEM's RMSE

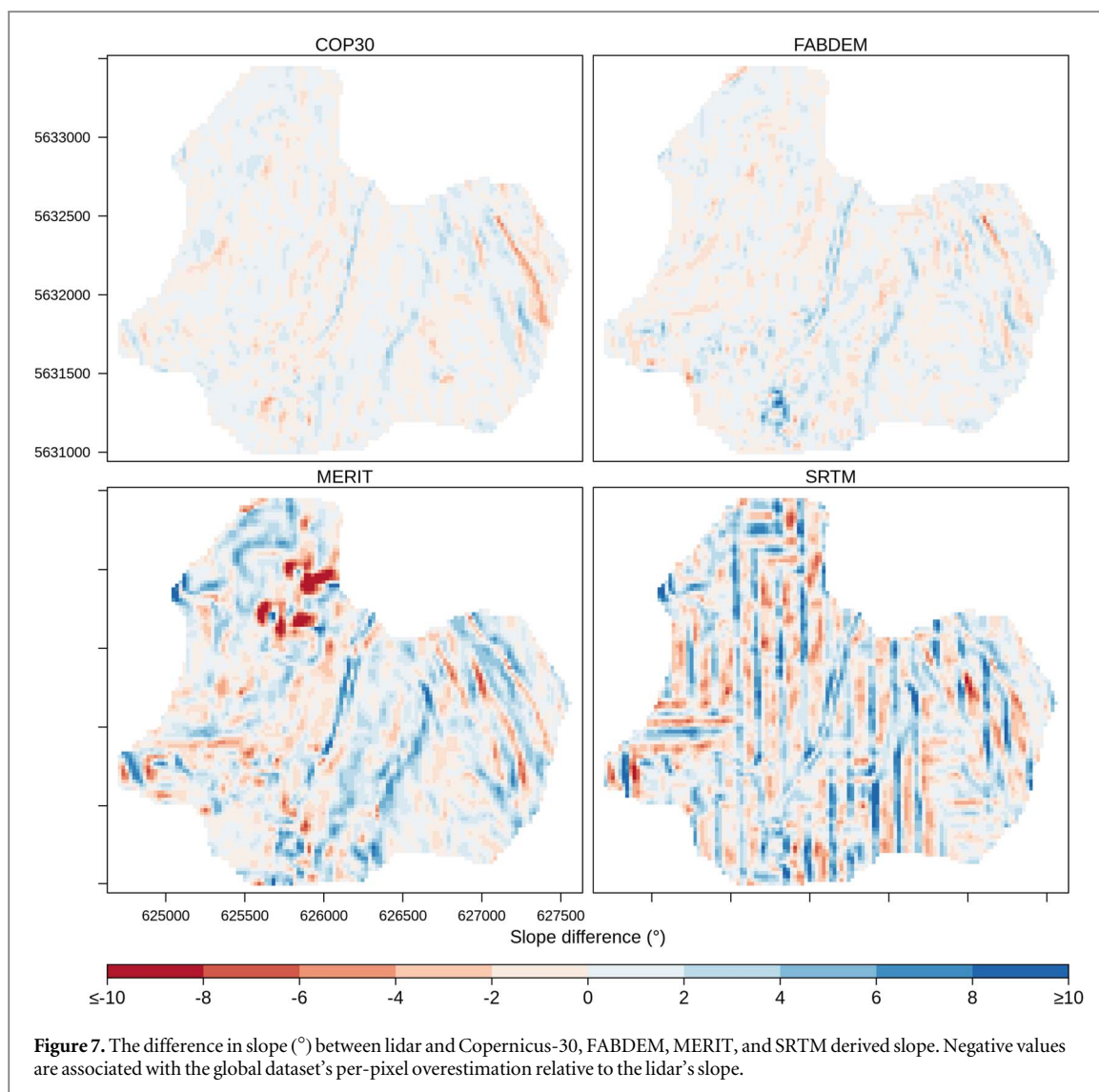


Table 2. Slope RMSE and mean bias error metrics for the datasets versus the lidar. Negative values are associated with the global dataset's overall overestimation relative to the lidar's slope.

DataSet	DEM	RMSE (°)	Mean Bias (°)
Slope	COP30	1.122	0.121
	FABDEM	1.317	0.224
	MERIT	3.179	0.603
	SRTM	3.685	0.755

(1.317°) and bias (0.224°). The MERIT error pattern was similar to that of COP30 and FABDEM, however in increased magnitude. The substantial error in elevation in the northern portion of the domain was reflected in large errors in slope estimates. The RMSE (3.179°) and bias (0.60°) were worse than FABDEM and COP30. The SRTM slope error pattern closely matched that of the elevation errors with substantial North-South striping. The RMSE (3.685°) and bias (0.755°) were the worst.

3.1.3. TPI

The aligned datasets were used to produce the slope and were compared spatially to the lidar-derived Terrain Position Index (TPI). The TPI differences (per equation (1)) are shown in figure 8 with Root Mean Squared Errors (RMSE) and Mean Biases summarized in table 3.

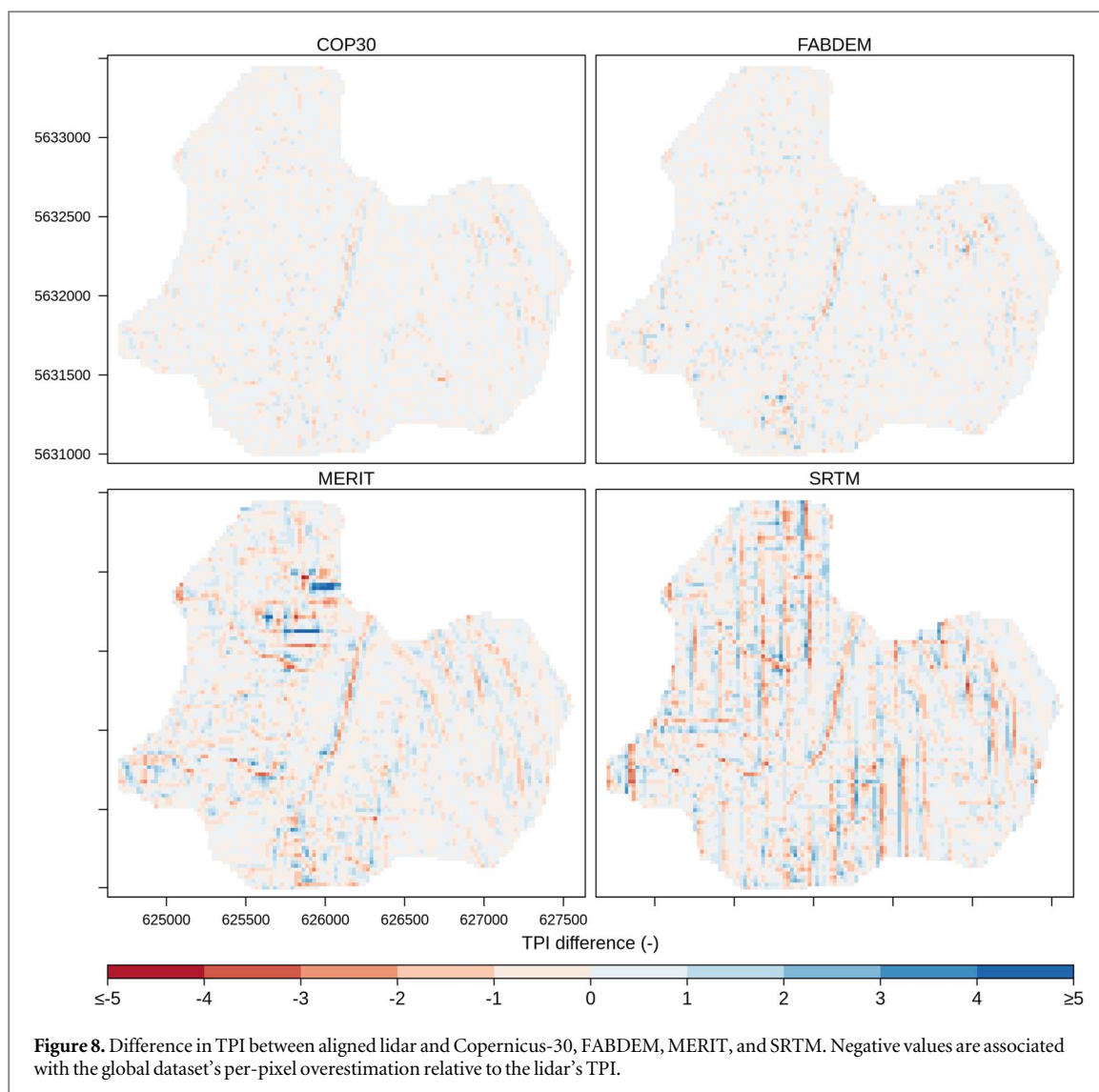


Figure 8. Difference in TPI between aligned lidar and Copernicus-30, FABDEM, MERIT, and SRTM. Negative values are associated with the global dataset’s per-pixel overestimation relative to the lidar’s TPI.

Table 3. TPI RMSE and mean bias error metrics for the datasets versus the lidar. Negative values are associated with the global dataset’s overall overestimation relative to the lidar’s TPI.

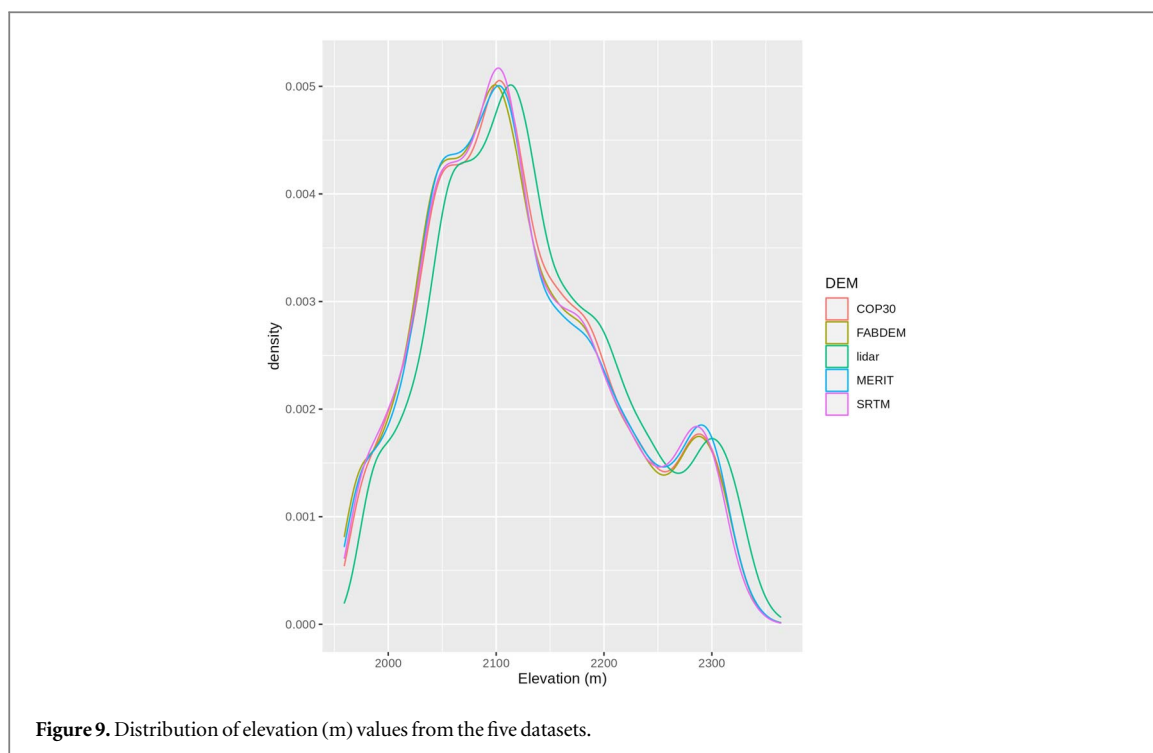
DataSet	DEM	RMSE (-)	Mean Bias (-)
TPI	COP30	0.366	-0.001
	FABDEM	0.480	-0.002
	MERIT	0.869	-0.005
	SRTM	1.088	0.004

The COP30 had the smallest errors in TPI (RMSE = 0.366, bias = -0.001) with the main spatial pattern along the westward-facing slope. The FABDEM had a slightly increased RMSE (0.480) and bias (-0.002), with a similar spatial pattern. Small differences along the eastern portion of the domain corresponded to the steep slopes in this area. The MERIT RMSE (0.869) was the second worst and had the worst bias (-0.005). The terrain features along the North portion of the domain had a further knock-on impact, producing nonsensical TPI values. The SRTM striping resulted in similar spatial error patterns once more, with a North-South striping with some East-West features. This resulted in the worst RMSE (1.088) and the second-worst bias (0.004).

3.2. Distribution

3.2.1. Elevation

The distribution of elevation across the domain is shown below in figure 9. The Copernicus-30 (COP30) is shown in red, FABDEM is shown in yellow, the resampled lidar is shown in green, the MERIT is shown in blue,



and SRTM is in purple. All the datasets had a generally positive bias versus the lidar at low elevations, and a low bias in the mid-elevations. In general, all the datasets have the same shape of the distribution of elevations. The MERIT DEM notably underestimated the occurrence of mid-elevations (≈ 2150 m) and was the most different of the distributions. The SRTM overestimated the occurrence of the 2100 m elevation band. Both MERIT and SRTM both overestimated high elevation (2300 m) occurrences.

3.2.2. Slope

The distribution of the slopes derived from each elevation dataset is shown in figure 10. The Copernicus-30 (COP30) is shown in red, FABDEM is shown in yellow, the resampled lidar is shown in green, the MERIT is shown in blue, and SRTM is in purple. The SRTM had a substantially different distribution than the other datasets. It had a low slope bias, overestimating the occurrence of low slopes ($< 10^\circ$). The SRTM underestimated the second peak at $\approx 17\text{--}20^\circ$ and overestimated the occurrence of large slopes. The FABDEM generally matched the lidar the most closely, although it underestimated slope occurrences in the $\approx 15^\circ$ region. The MERIT DEM is more similar to SRTM and underestimated occurrences of the low slopes and, similar to COP30, overestimated the mid-slopes ($\approx 17^\circ\text{--}20^\circ$). All elevation datasets generally captured the bimodal distribution, with the SRTM being the weakest in doing so.

3.2.3. TPI

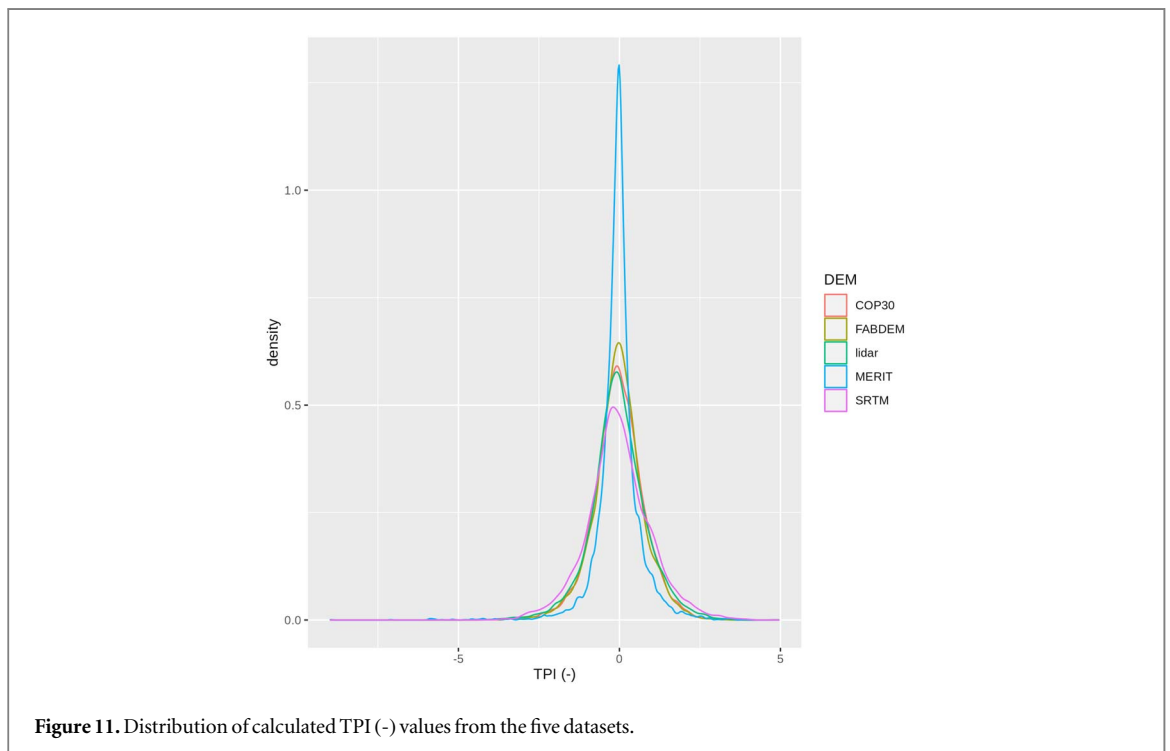
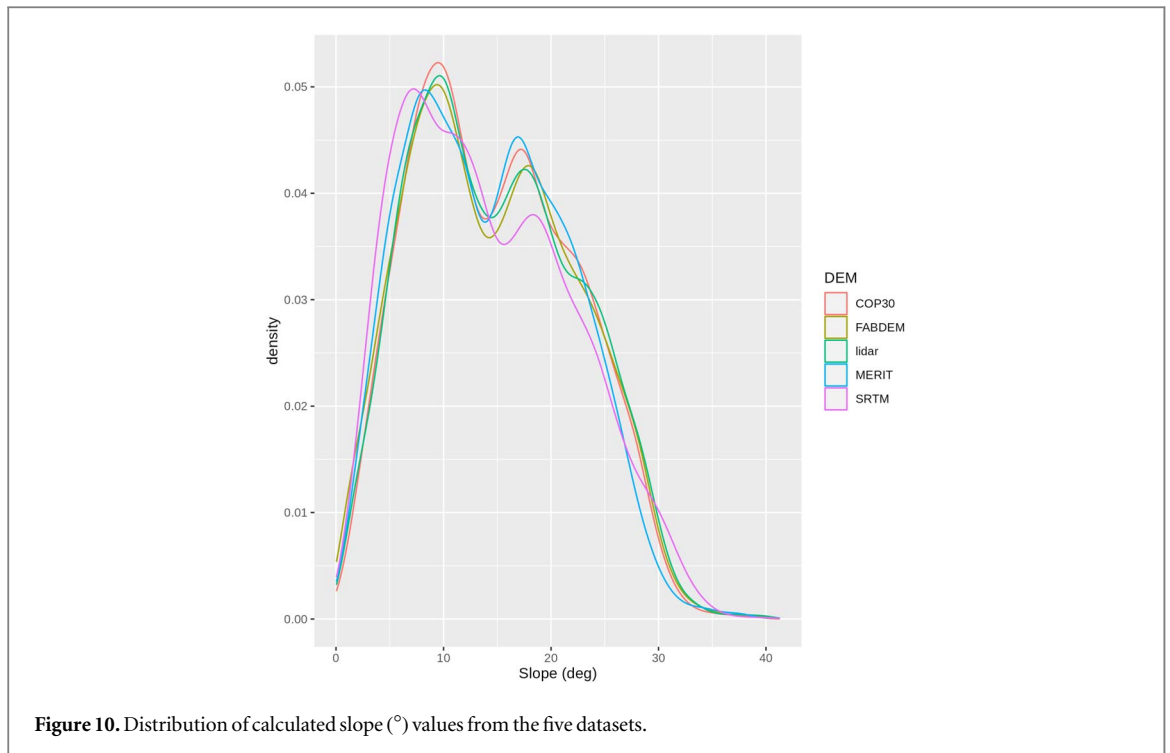
The distribution of TPI derived from each dataset is shown in figure 11 with the TPI derived from the Copernicus-30 (COP30) shown in red, FABDEM shown in yellow, the resampled lidar is shown in green, the MERIT shown in blue, and SRTM in purple. The MERIT DEM produced the most biased distribution of TPI, heavily biased towards small, slightly negative TPI values. SRTM DSM is the opposite, with too many high and low TPI values. COP30 DSM almost exactly matched the lidar values, whereas FABDEM DEM slightly overestimated these small, slightly negative TPI values — i.e., the small depressions in the slopes.

3.3. Distribution of differences

3.3.1. Elevation

The distribution of elevation differences across the domain is shown below in figure 12. The COP30 is shown in red, FABDEM is shown in yellow, the MERIT is shown in blue, and SRTM is in purple. Negative values reflect the dataset being at a higher elevation than the lidar.

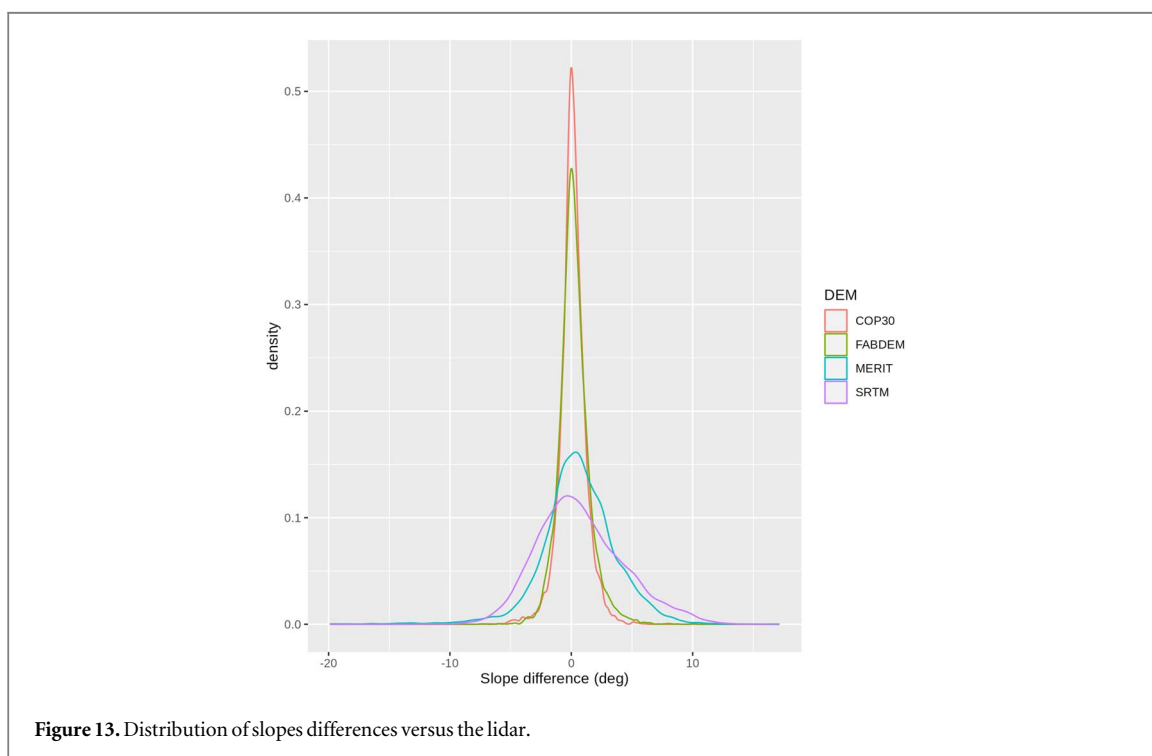
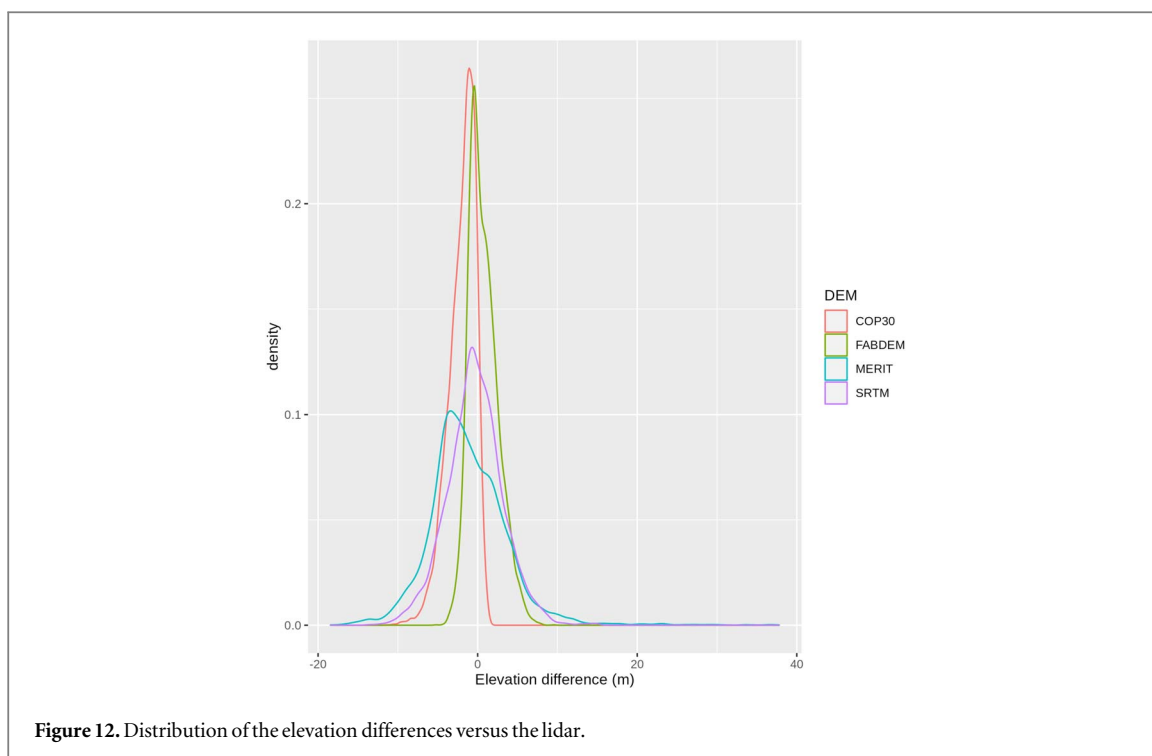
These distributions of differences reflect the range of values informing the previously noted RMSE and bias, with errors due to a positive bias versus the lidar. The SRTM DSM had a more closely centred distribution



around the 0 m elevation, showing a more equal mix of over and underestimation. The MERIT DEM had a notable underestimation bias, whereas the FABDEM DEM had an overestimation bias.

3.3.2. Slope

The distribution of slope differences across the domain is shown below in figure 13. The COP30 is shown in red, FABDEM is shown in yellow, the MERIT is shown in blue, and SRTM is in purple. Negative values reflect the dataset covering a larger slope than the lidar.

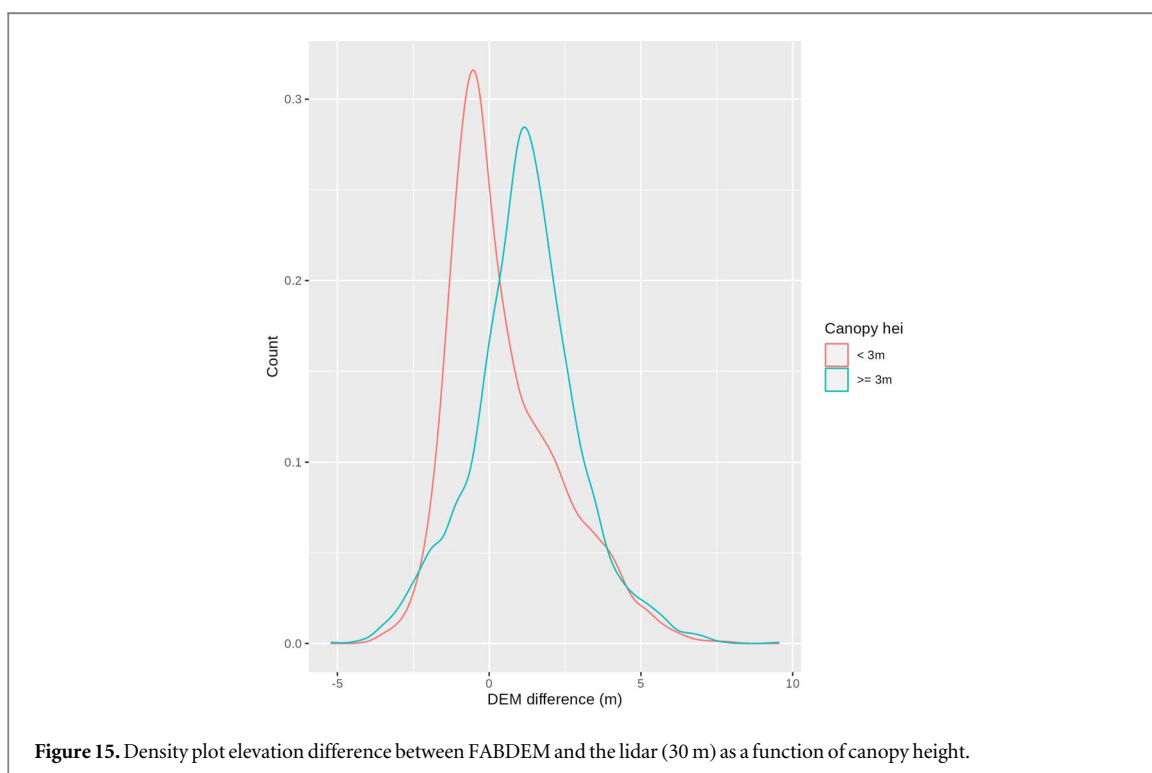
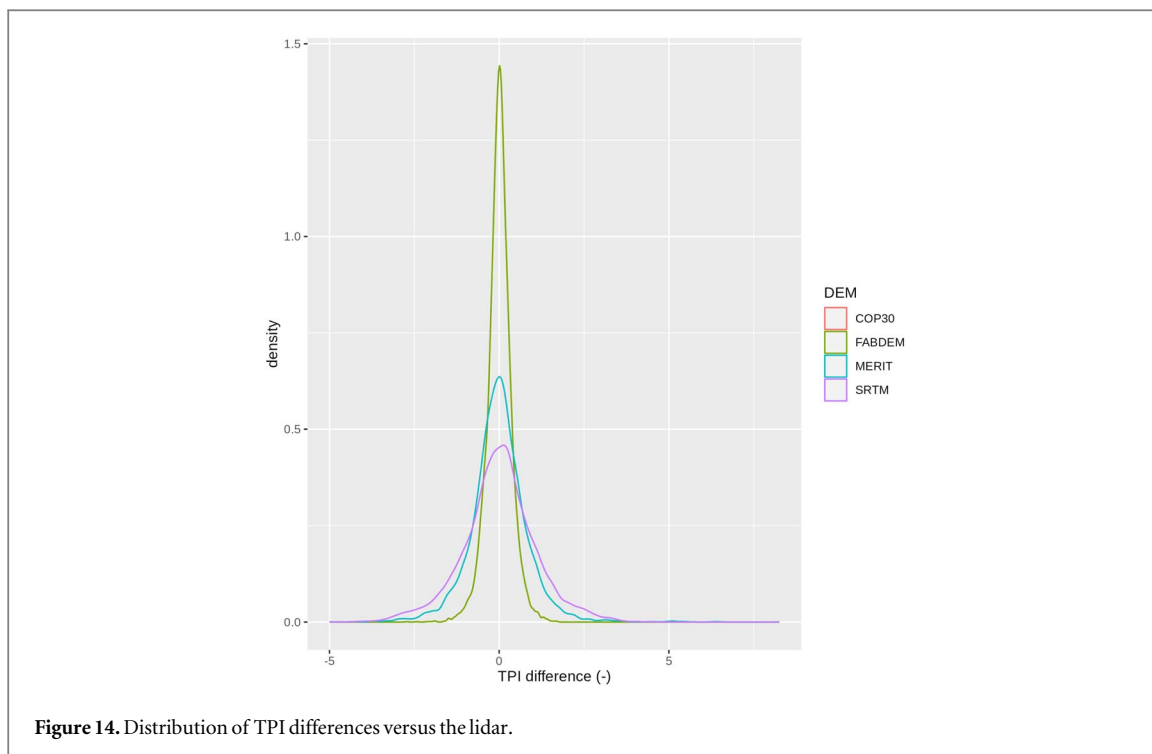


The FABDEM and COP30 distributions were well-centred at 0° . The MERIT DEM and SRTM DSM showed an increasing prevalence of larger slope differences. In these distributions, the SRTM DSM clearly performed the worst of all the other datasets.

3.3.3. TPI

The distribution of TPI differences across the domain is shown below in figure 14. The COP30 is shown in red, FABDEM is shown in yellow, the MERIT is shown in blue, and SRTM is in purple. Negative values reflect the DEM is at a higher elevation than the lidar.

Like the elevation and slope, the COP30 and FABDEM were visually indistinguishable. The MERIT and SRMT error distributions were, respectively, increasingly worse showing a much larger distribution of errors.



3.3.4. Consideration of 3 m threshold

The FABDEM only has bias corrections for canopies exceeding 3 m. The impact of this threshold is shown in figure 15. The errors associated with a location with a canopy below 3 m are shown in red, and the errors associated with a canopy larger than or equal to 3 m are shown in blue. There was an overestimation in the FABDEM elevations in the regions where the vegetation removal was not run. In the areas of tall canopies, there was a slight negative bias in the FABDEM.

4. Discussion

The COP30 DSM is considered one of the best, free, global DSM datasets currently available (Gesch 2018; Grohmann 2018, Guth and Geoffroy 2021, Hawker *et al* 2022). The application of COP30 to environmental forecasting is already providing improved results, e.g., Garrot (2022). The derivative dataset, FABDEM, represents the first 30 m de-forested DEM. Thus, there is a substantial opportunity for FABDEM to provide an improved and usable DEM in complex topography for use in the environmental sciences in locations where the COP30 DSM has too much vertical bias due to tall vegetation canopies. However, the accuracy associated with machine learning approaches in complex mountain topography to de-bias DSM datasets remains an open question. Recent advances in UAV-based lidar systems are providing unprecedented sub-canopy lidar coverage, allowing for detailed canopy height models. In turn, this allows for the creation of low-error reference DEMs that can then be used to validate DEMs.

When compared to the lidar DEM, the FABDEM had the smallest elevation RMSE and mean bias, representing a substantial improvement over existing datasets such as MERIT. Indeed, it also suggests that the vegetation removal over Fortress Mountain Research Basin was generally successful. The areas of largest error in the COP30 dataset almost exactly match the spatial extents of the tallest canopies in this domain. In these tall-canopy areas, the FABDEM substantially improves upon COP30. The shape of the elevation distributions was more similar to that of the resampled lidar (30 m), with a negative bias offset. Without this offset, the two distributions are almost identical, suggesting that the small-scale features are well captured by the FABDEM DEM. This is also reflected in the slope distribution which is almost identical to that of the resampled lidar. The exception is the representation of the $\approx 17^\circ$ slopes whose occurrence is slightly underestimated by FABDEM. The width of the distribution of TPI is similar to that of the lidar, however, there is an over occurrence of low TPI topography versus the lidar. Both these areas of underperformance are likely due to the smoothing passes performed as a post-process step in the creation of FABDEM. Additional improvements in the RMSE of FABDEM are likely attributed to the improved gravitational model used by COP30 (and by extension, FABDEM). COP30 uses the Earth Gravitational Model 2008 (EGM2008) geoid model (Pavlis *et al* 2012) as compared to EGM1996 as used by the other datasets. EGM2008 is substantially improved over EGM1996 and Goyal *et al* (2021) reported differences between the two up to 12 m between the two models.

Due to the hydrological conditioning of the MERIT DEM, it has been widely used in hydrological applications. However, the coarse resolution of 90 m has clear limitations in the complex topography of mountain basins. It is clearly better than the SRTM DSM, however, the artefacts around a few of the gullies in this domain are problematic. In a comparison of SRTM and MERIT in the Rockies, Hirt (2018) found that MERIT had improved void filling with fewer artefacts than SRTM. In the study region herein, the co-location of the artefacts with hydrological features, such as gullies, may attribute these errors to the hydrological conditioning. This is further substantiated as the SRTM data does not have these artefacts. Due to the colocation with steep slopes, it may also be related to void-filling strategies, e.g., Bildirici *et al* (2007). Despite these artefacts, it was found that MERIT is greatly improved versus the SRTM DSM in this area, which agrees with previous work (Hirt 2018). The strong striping in the SRTM dataset was exceptionally problematic. Such patterns have been described in various SRTM evaluations (Yamazaki *et al* 2017, Goyal *et al* 2021). The stripping pattern would have profound impacts on many processes' simulation and quantification, and these data should be used with care in steep mountain basins.

The FABDEM elevation bias and RMSE were the best of all the datasets, improving upon the COP30 dataset, and these improvements are clearly a result of the vegetation debiasing. The SRTM DSM has a clear mid-elevation bias from including the vegetation canopy. However, the FABDEM debiasing improvement comes at the cost of slightly reduced slope and TPI scores, although differences are so small (RMSE differences of 0.2° and 0.1° respectively) that the tradeoff is almost certainly worthwhile. However, smoothing may be problematic in areas of low slopes, such as the interior plains or Arctic coastal plains.

5. Conclusions

Space-based, global-extent digital elevation models are an integral part of quantifying Earth system processes and are a key input for hydrological models. Obtaining high-quality elevation data in steep, forested mountainous terrain is a key challenge for simulating headwater mountain catchments. Without correcting for forest cover, hydrological and ecological process predictions in complex, forested terrain may end up biased. Due to the complex land-atmosphere interactions with various environmental processes, the use of deforested DEMs is strongly motivated. Although aeroplane-based lidar data have been available for mountain basins, the recent deployment of UAV-based lidar systems has further constrained the sub-canopy estimates of elevation in a way previously not possible. This is due to the substantially denser point clouds, with many more sub-canopy

returns helping to better resolve the sub-canopy topography. These sub-canopy elevation data sets provide an optimal dataset with which to evaluate existing (MERIT) and new (FABDEM) deforested DEMs in a steep, representative mountain catchment.

The use of machine learning to improve the fidelity of remote-sensing based, global datasets is a clear opportunity to further improve the inputs to various Earth-science workflows. However, the uncertainties associated with machine learning approaches to complex mountain topography remain an open question. In this study, the evaluation of FABDEM, SRTM, Copernicus-30, and MERIT datasets against high-fidelity sub-canopy lidar data presents a new opportunity for quantifying the error in these datasets.

It was found that the new FABDEM had a 24% reduction in elevation RMSE (2.58 m versus 1.95 m) and a 135% reduction in bias (-1.933 m versus 0.685 m) compared to the Copernicus-30 dataset from which it is derived. The spatial elevation error in Copernicus-30 was collocated with the tall vegetation in the domain, these areas had reduced error in the FABDEM. Further, the FABDEM better preserved the shape of the elevation and derivative parameters such as TPI and slope versus the other datasets of MERIT and SRTM. The FABDEM elevation was the best of all the datasets, clearly a result of the vegetation debiasing although the Copernicus-30 had slightly better slope and TPI RMSE/bias versus the lidar. However, the differences between FABDEM and Copernicus-30 are so small (RMSE differences of 0.2° and 0.1° respective) that the tradeoff is almost certainly worthwhile.

Overall, the FABDEM provides a clear improvement over existing deforested DEM products in complex mountain topography such as the MERIT DEM. The preservation of key aspects of the distribution of derivative landscape products such as slope and TPI suggests that the use in hydrological and ecological models would benefit from these high-resolution data, especially when simulating processes with substantial heterogeneity. The domain evaluated here is representative of the steep, forested mountain headwater basins present in the Canadian Rockies. Further comparisons should be conducted in the future over broader extents as more UAV-lidar data becomes available. However, this study supports the use of FABDEM in forested mountain catchments as the current best-in-class data product.

Acknowledgments

The authors would like to acknowledge support from the Canada First Research Excellence Fund's Global Water Futures Programme, Western Economic Diversification Canada, the Natural Sciences and Engineering Research Council of Canada Discovery Grants and the Canada Research Chairs programme. The field support of Greg Galloway, Robin Heavens, Lindsey Langs, Alistair Wallace, Dong Zhao, and Kieran Lehan in UAV measurements is greatly appreciated.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Christopher B Marsh  <https://orcid.org/0000-0002-1372-4513>

References

- Airbus Defence and Space GmbH 2020 *Copernicus DEM: Copernicus Digital Elevation Model Product Handbook*RFP/RFI-No.: AO/1-9422/18/I-LG pp 1–36
- Axelsson P 2000 DEM generation from laser scanner data using adaptive TIN models *Int. Arch. Photogramm. Remote Sens.* **31** 111–8
- Bernhardt M and Schulz K 2010 SnowSlide: a simple routine for calculating gravitational snow transport *Geophys. Res. Lett.* **37** 1–6
- Beven K 1981 Kinematic subsurface stormflow *Water Resour. Res.* **17** 1419–24
- Beven K and Freer J 2001 A dynamic TOPMODEL *Hydrological Processes*, **15** 1993–2011
- Bildirici İ, Üstün A, Ulugtekin N, Selvi H, Abbak R, Bugdayci I and Dogru A 2007 SRTM data in Turkey: void filling strategy and accuracy assessment *Conf.: 4th Middle East Spatial Technology Conf. and Exhibition (MEST2007) in Bahrain*
- Carabajal C C and Harding D J 2005 ICESat validation of SRTM C-band digital elevation models *Geophys. Res. Lett.* **32**
- Conway J P, Helgason W D, Pomeroy J W and Sicart J E 2021 Icefield breezes: mesoscale diurnal circulation in the atmospheric boundary layer over an outlet of the columbia icefield, canadian rockies *J. Geophysical Research: Atmospheres* **126** 1–17
- Dozier J and Frew J 1990 Rapid calculation of terrain parameters for radiation modeling from digital elevation data *IEEE Trans. Geosci. Remote Sens.* **28** 963–9
- Dubayah R *et al* 2020 The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography *Science of Remote Sensing* **1** 100002
- Dujardin J and Lehning M 2022 Wind-topo: downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning *Q. J. R. Meteorolog. Soc.* **148** 1368–88

- Farr T G *et al* 2007 The shuttle radar topography mission *Rev. Geophys.* **45** 1–33
- Fenton G A, McLean A, Nadim F and Griffiths D V 2013 Landslide hazard assessment using digital elevation models *Canadian Geotechnical Journal* **50** 620–31
- Garnier B J and Ohmura A 1968 A method of calculating the direct shortwave radiation income of slopes *J. Appl. Meteorol.* **7** 796–800
- Garrote J 2022 Free global DEMs and flood modelling—a comparison analysis for the January 2015 flooding event in mocuba city (Mozambique) *Water* **14** 176
- GDAL/OGR contributors 2022 GDAL/OGR geospatial data abstraction software library. [online] Available from <https://gdal.org> (<https://doi.org/10.5281/zenodo.5884351>)
- Gesch D B 2018 Best practices for elevation-based assessments of sea-level rise and coastal flooding exposure *Frontiers in Earth Science* **6** 230
- Gharari S, Clark M P, Mizukami N, Knoben W J M, Wong J S and Pietroniro A 2020 Flexible vector-based spatial configurations in land models *Hydrol. Earth Syst. Sci.* **24** 5953–71
- Giroday H-M C de la, Carroll A L, Lindgren B S and Aukema B H 2011 Incoming! Association of landscape features with dispersing mountain pine beetle populations during a range expansion event in western Canada *Landscape Ecology* **26** 1097
- Goyal R, Featherstone W E, Dikshit O and Balasubramania N 2021 Comparison and validation of satellite-derived digital surface/elevation models over India *J. Indian Soc. Remote Sens.* **49** 971–86
- Grohmann C H 2018 Evaluation of TanDEM-X DEMs on selected Brazilian sites: Comparison with SRTM, ASTER GDEM and ALOS AW3D30 *Remote Sens. Environ.* **212** 121–33
- Guth P L and Geoffroy T M 2021 LiDAR point cloud and ICESat-2 evaluation of 1 second global digital elevation models: Copernicus wins *Transactions in GIS* **25** 2245–61
- Hansen M C *et al* 2013 High-resolution global maps of 21st-century forest cover change *Science* **342** 850–3
- Harder P, Pomeroy J W and Helgason W D 2020 Improving sub-canopy snow depth mapping with unmanned aerial vehicles: lidar versus structure-from-motion techniques *The Cryosphere* **14** 1919–35
- Harding D J and Carabajal C C 2005 ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure *Geophys. Res. Lett.* **32** 1–4
- Hawker L, Uhe P, Paulo L, Sosa J, Savage J, Sampson C and Neal J 2022 A 30 m global map of elevation with forests and buildings removed *Environ. Res. Lett.* **17** 024016
- 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 Sens. Environ.* **207** 27–41
- Illés G, Kovács G and Heil B 2011 Comparing and evaluating digital soil mapping methods in a Hungarian forest reserve *Can. J. Soil Sci.* **91** 615–26
- Isenburg M 2023 LAStools - efficient tools for LiDAR processing. version 210418 [online last accessed 22-03-2023] Available from <http://rapidlasso.com/LAStools>
- Kienzle S W 2011 Effects of area under-estimations of sloped mountain terrain on simulated hydrological behaviour: a case study using the ACRU model *Hydrol. Processes* **25** 1212–27
- Kucharczyk M, Hugenholtz C and Zou X 2018 UAV–LiDAR accuracy in vegetated terrain *Journal of Unmanned Vehicle Systems.* **6** 212–34
- Kulp S A and Strauss B H 2018 CoastalDEM: a global coastal digital elevation model improved from SRTM using a neural network *Remote Sensing of Environment* **206** 231–9
- Langs L E, Petrone R M and Pomeroy J W 2021 Subalpine forest water use behaviour and evapotranspiration during two hydrologically contrasting growing seasons in the Canadian Rockies *Hydrol. Processes* **35** 1–17
- Liu M, Hu Y, Chang Y, He X and Zhang W 2009 Land use and land cover change analysis and prediction in the upper reaches of the minjiang river, China *Environmental Management* **43** 899–907
- Magruder L, Neuenschwander A and Klotz B 2021 Digital terrain model elevation corrections using space-based imagery and ICESat-2 laser altimetry *Remote Sens. Environ.* **264** 112621
- Marsh C B, Pomeroy J W and Spiteri R J 2012 Implications of mountain shading on calculating energy for snowmelt using unstructured triangular meshes *Hydrol. Processes* **26** 1767–78
- Marsh C B, Pomeroy J W, Spiteri R J and Wheeler H S 2019 A finite volume blowing snow model for use with variable resolution meshes *Water Resour. Res.* **56** 1–28
- Moudrý V, Lecours V, Gdulová K, Gábor L, Moudrý L, Kropáček J and Wild J 2018 On the use of global DEMs in ecological modelling and the accuracy of new bare-Earth DEMs *Ecol. Modell.* **383** 3–9
- Natural Resources Canada Precise Point Positioning Online Tool, [online last accessed 22-03-3023] Available from <https://webapp.csrscs.nrcan-rncan.gc.ca/geod/tools-outils/ppp.php?locale=en> 2023
- Neuenschwander A L, Pitts K L, Jelley B P, Robbins J, Klotz B, Popescu S C, Nelson R F, Harding D, Pederson D and Sheridan R 2021 ATLAS/ICESat-2 L3A land and vegetation height, version 4 Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center 1–18
- O’Callaghan J F and Mark D M 1984 The extraction of drainage networks from digital elevation data *Comput. Vis. Graph. Image Process.* **28** 323–44
- O’Loughlin F E, Paiva R C D, Durand M, Alsdorf D E and Bates P D 2016 A multi-sensor approach towards a global vegetation corrected SRTM DEM product *Remote Sens. Environ.* **182** 49–59
- Pavlis N K, Holmes S A, Kenyon S C and Factor J K 2012 The development and evaluation of the earth gravitational model 2008 (EGM2008) *J. Geophysical Research: Solid Earth* **117** 1–38
- Potapov P *et al* 2021 Mapping global forest canopy height through integration of GEDI and Landsat data *Remote Sens. Environ.* **253** 112165
- Reu J D *et al* 2013 Application of the topographic position index to heterogeneous landscapes *Geomorphology* **186** 39–49
- Ryan B C 1977 A Mathematical model for diagnosis and prediction of surface winds in mountainous terrain *J. Appl. Meteorol.* **16** 571–84
- Sampson C C, Smith A M, Bates P D, Neal J C and Trigg M A 2016 Perspectives on open access high resolution digital elevation models to produce global flood hazard layers *Frontiers in Earth Science* **3** 85
- Simard M, Pinto N, Fisher J B and Baccini A 2011 Mapping forest canopy height globally with spaceborne lidar *J. Geophysical Research: Biogeosciences* **116** 1–12
- Uuemaa E, Ahi S, Montibeller B, Muru M and Kmoch A 2020 Vertical accuracy of freely available global digital elevation models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM) *Remote Sensing* **12** 3482
- Vionnet V, Marsh C B, Menounos B, Gascoin S, Wayand N E, Shea J, Mukherjee K and Pomeroy J W 2021 Multi-scale snowdrift-permitting modelling of mountain snowpack *The Cryosphere* **15** 743–69
- Viviroli D, Kumm M, Meybeck M, Kallio M and Wada Y 2020 Increasing dependence of lowland populations on mountain water resources *Nature Sustainability* **3** 917–28

- Wagenbrenner N S, Forthofer J M, Page W G and Butler B W 2019 Development and Evaluation of a Reynolds-Averaged Navier–Stokes Solver in WindNinja for Operational Wildland Fire Applications *Atmosphere* **10** 672
- Weiss A D 2001 Topographic position and landforms analysis [online] Available from (http://jennessent.com/downloads/TPI-poster-TNC_18x22.pdf)
- Yamazaki D, Ikeshima D, Tawatari R, Yamaguchi T, O’Loughlin F, Neal J C, Sampson C C, Kanae S and Bates P D 2017 A high-accuracy map of global terrain elevations *Geophys. Res. Lett.* **44** 5844–53
- Zhao X, Su Y, Hu T, Chen L, Gao S, Wang R, Jin S and Guo Q 2018 A global corrected SRTM DEM product for vegetated areas *Remote Sensing Letters* **9** 393–402