CoastalDEM v3.0: Improving fully global coastal elevation predictions through a convolutional neural network and multi-source DEM fusion

Scott A. Kulp[1]*, Benjamin H. Strauss[1]

In Brief
Here we present CoastalDEM v3.0, the latest update to Climate Central's global coastal digital elevation model (DEM) powered by an artifical convolutional neural network. In this version, we have refined the neural network architecture and incorporated a number of new input datasets. In particular, instead of a single base elevation model as most past work has employed, we use a fusion of multiple global DEM sources, which helps to counter spatially-autocorrelated vertical errors. As a result, CoastalDEM v3.0 not only extends its boundaries to all latitudes, but also outperforms the vertical accuracy of all other publicly-available global scale elevation models, including CoastalDEM v2.1, DiluviumDEM, FABDEM, and COPDEM, as measured by a global ground-truth dataset.

[1] *Climate Central, Princeton, NJ, USA*
***Corresponding author**: skulp@climatecentral.org

1. Introduction

Digital elevation models (DEMs) represent the foundation of coastal flood vulnerability assessments. In certain places, such as the coastal United States [1], Australia [2], the UK [3], and other select developed countries, high-accuracy DEMs derived from airborne lidar exist and are publicly available. In most other parts of the world, we rely on global-scale elevation models derived from satellite-based measurements, such as NASADEM [4], COPDEM [5], and AW3D30 [6]. However, all of these DEMs suffer from large vertical errors with a positive bias, especially in areas with dense vegetation or urban development [7, 8, 9].

Over the last ten years, improving satellite-derived global elevation models has become a topic of increasing effort and interest in the scientific community, especially as it relates to flooding impacts analysis. Early attempts were limited to very small areas of interest [10, 11] or only sought to reduce errors caused by vegetation, and not urban development. [12, 13, 14, 15]. CoastalDEM v1.1 [16] was the first DEM to use an artificial multilayer perceptron neural network to correct global-scale coastal elevations present in NASA's SRTM, but was trained on ground truth data in the US alone, and so offered less certainty in areas with dissimilar vegetation, architecture, and population density. CoastalDEM v2.1 [7] resolved this limitation by employing data from NASA's ICESat 2, version 3 as ground truth. The ICESat 2 satellite uses lidar technology to collect a lattice of global transects of elevation measurements[17]. CoastalDEM v2.1 also utilized a more sophisticated convolutional neural network (CNN) architecture to further improve results.

Since then, the European Space Agency developed their Copernicus digital elevation model (COPDEM) [5], based on data collected from the TanDEM-X mission [18]. While still a digital surface model and so positively biased in developed or vegetated areas, recent work indicates that it offers higher vertical precision and more consistent performance than SRTM/NASADEM [19, 20]. Accordingly, the most recent efforts in global DEM correction have used COPDEM as their base elevation source, including FABDEM [8] and DiluviumDEM [21]. These datasets, however, suffer many of the same limitations as CoastalDEM v1.1. First, they employ weaker prediction models: random forests and decision trees, respectively. While much less computationally intensive than convolutional neural networks to train and use, these models are weaker and primarily optimized for tabular data, and cannot take full advantage of the spatial gridded structure of the input data employed in this task, as CNN's were specifically invented to do. Use of the more limited approaches can lead to substantial noise, such as is clearly visible in FABDEM's and DiluviumDEM's outputs (Figure 6, below). Furthermore, these datasets were trained only on lidar-derived elevation models in a handful of countries where they are available. While these DEMs therefore perform well in those particular regions, their apparently high accuracies do not necessarily extend to developing countries where lidar data is not available - and it is in these exact regions where high-accuracy global DEMs are most needed.

The main purpose of this white paper is to rigorously assess the error in CoastalDEM v3.0 and other leading DEM's with global or near-global coverage of coastal areas. Most published validations of coastal DEMs have relied on error assessment in small and few areas [22, 23, 24, 25]. Here, we

take pains to make a thorough global assessment, and also to compute and publish error rates in every coastal nation in the world with enough ground truth data to achieve a sufficient sample size.

2. Technological Advances in CoastalDEM v3.0

- **A fusion of base elevation sources.** CoastalDEM v2.1 was based off of the NASADEM global elevation model which, while more accurate than the older SRTM dataset, still contained important errors including extensive striping artifacts. CoastalDEM v3.0 instead incorporates not one but several of the most recent and advanced global elevation models. Such a fusion helps reduce local errors unique to any single base layer, while also combining the particular strengths of all three datasets.
- **New and updated input variables.** CoastalDEM v3.0 incorporates updates from most of the input variables used in CoastalDEM v2.1, and includes new ones. These improved inputs further help produce more accurate elevation estimates relative to ground truth.
- **Complete global coastal coverage** CoastalDEM v2.1 was limited to latitudes between S56 and N60, as these were the extents of NASA's SRTM mission. CoastalDEM v3.0 now instead fully extends to the poles.
- **Deeper inland coverage** While CoastalDEM v2.1 corrected land elevations up to 120m, it was still limited to coastal tiles. CoastalDEM v3.0 offers corrections everywhere in the world up to the 120m limit, even in tiles that are deeper inland. This more than doubles CoastalDEM's extent.
- **Trained on updated high-quality global elevation data.** CoastalDEM v2.1 was trained using version 3 of NASA's ICESat-2 mission, which collects space-based lidar data along a global lattice of transects, too sparse for direct generation of a high-resolution DEM, but useful for the purposes employed here. CoastalDEM v3.0 uses the newest update to these data, currently at version 6. Our internal comparisons to more accurate airborne-lidar-derived elevation models in the US suggest that this newer version of ICESat-2 offers large improvements to vertical accuracy compared to its predecessor. Furthermore, we made additional improvements to ICESat-2 in areas of very high population density to improve performance in major global cities, discussed in the Methods section, below.

3. Results

3.1 Validation against ICESat-2

Here we use land elevation measurements from NASA's ICESat-2 v6, with adjustments made in extremely high-density areas described in Methods below, as ground truth to assess the global accuracy of global DEMs. We include five recent products – CoastalDEM v3.0, CoastalDEM v2.1, DiluviumDEM, FABDEM, and COP-DEM. We disregard all ICESat-2 points flagged as being covered by clouds or snow. Additionally, all error values exceeding 50 m are treated as outliers and removed from the assessment (fewer than 0.005% of points have a discrepancy this large).

We have empirically found that DEM performance varies by elevation. Since CoastalDEM's intended purpose is for coastal flood modeling on land presently above sea level and especially in populated areas, we develop error assessments for different elevation and population density bands and their combinations. We consider three low-lying elevation bands: 0-2 m, which covers most current high-frequency flood risks today, and most sea-level-rise projections through 2100; 0-5 m, which covers most low-frequency flood risks today, plus the combination of most sea-level-rise scenarios and flood risks this century; and 0-10 m, which covers the most extreme flood and sea level scenarios [26, 27]). More specifically, when assessing vertical accuracy of a DEM, we consider only grid cells where the "true" (ICESat-2) *or* the "estimated" (DEM) elevations are greater than zero and lower than the given maximum elevation. For brevity, for the rest of this report we only list the upper elevation bounds assessed ($<$2 m, $<$5 m, or $<$10 m), with the lower bound of 0 m left implied. We also consider three population density bands: $>$10,000 people per square kilometer (high-density urban), $>$1,000 people per square kilometer (urban), and any population density. All available data points present in ICESat-2 that meet the above requirements and given filters are used in the following assessments.

Note that while the $<$2 m threshold is highly relevant to coastal flooding vulnerability assessments, the number of such ICESat 2 points in very high-population areas is much lower in most countries than the number below the $<$5 m threshold. Accordingly, while our globally-aggregated results prominently present numbers at the $<$ 2 m threshold, which is the immediate highest-risk threshold, the figures presenting results by country or smaller units typically avoid that band because of the low sample size.

In general, we find that in the global aggregate, all corrected global DEMs (both versions of CoastalDEM, DiluviumDEM, and FABDEM) offer low bias under most point filtering parameters, though the uncorrected Digital Surface Model (DSM), COPDEM, contains high vertical bias in populated areas. That is, there is not a clear and consistent over- or under-performer among the corrected DEMs with this metric. Accordingly, in this report, we use RMSE and LE90 as the main distinguishing factors.

In each of the elevation and population bands assessed, we find CoastalDEM v3.0 offers the lowest RMSE and LE90 across all other global DEMs by a considerable margin, with DiluviumDEM being its consistently closest competitor (Table 1). Across all land below 2 m, regardless of population

density, CoastalDEM v3.0 contains an RMSE of 1.35 m, 29-47% lower than its corrected competitors, and an LE90 of 1.25m, 34-51% lower than the others. Similar results are seen at the 5 m and 10 m thresholds, with DiluviumDEM never falling below 1.78 m in RMSE or LE90, and CoastalDEM v3.0 never exceeding 1.53 m. CoastalDEM v3.0 thus shows the highest global accuracy when evaluated with these criteria.

Interestingly, each of the global DEMs (including COPDEM) see somewhat improved RMSE performance but higher LE90 in moderately-developed areas (greater than 1,000 people per square kilometer, where roughly half of the world's total population lives). This may be explained by large outliers present in ICESat-2 itself in heavily vegetated areas [28, 29, 30], skewing RMSE results (highly sensitive to outliers) in very low density areas when using ICESat-2 as ground truth. Regardless, relative performance between the global DEMs are consistent with CoastalDEM v3.0 again offering less than <1.3 m RMSE and < 1.6m LE90 in low-elevation areas - over 10% and 21% lower than DiluviumDEM, respectively. In segments of coastline with very high population density (greater than 10,000 people per square km, where errors caused by tall buildings and dense development are most severe), CoastalDEM's v3.0 improved performance is even more notable, with 1.27 m RMSE and 1.81m LE90 at the 2 m elevation threshold - 32% lower than DiluviumDEM under both metrics.

DEMs can contain spatially-autocorrelated errors even when they exhibit strong global performance, so it is important to also assess bias and RMSE at smaller spatial scales. Here we employ the GADM 2.0 dataset [31], a collection of global administrative unit spatial boundaries. Vertical error distributions are computed at the smallest-available administrative units for each area (such as state- and county-equivalents), which are then aggregated to wider spatial scales, including across countries. We then use these error distributions to estimate all relevant error metrics, including the median bias, RMSE, and LE90. Detailed error statistics by nation are presented in Supplementary Dataset S1.

In Figures 1 and 2, we present choropleth maps of nations' median biases and RMSE's under CoastalDEM v3.0, as well as DiluviumDEM and FABDEM, using the <5 m elevation threshold and >1000 ppsk population density threshold. We again find that across each of these DEMs, median biases are consistently low, with none being clearly superior to the others under this metric. Instead, we once more find the DEMs are more clearly differentiated by their error scatters. CoastalDEM v3.0 offers highly consistent RMSE's across nearly all coastal nations, improving upon DiluviumDEM and FABDEM each in 50 of the 53 countries with at least 1,000 ICESat-2 points within this domain. Additionally, CoastalDEM v3.0's LE90 outperforms the other DEMs in 52 of these 53 nations (Egypt representing the single exception).

Figure 3 provides further evidence of consistent performance across small spatial scales. Here we assess error across smaller ("level 1") administrative units, roughly equivalent to US counties. We applied the same domain filtering as the preceding figures (>1,000 people per square kilometer, <5 m elevation). This figure presents median bias and RMSE density plots based on all qualifying small regions with at least 100 such points (roughly 1,000 administrative units in count). Results for each of the global DEMs are represented by the colored curves, with steeper curves closer to 0 m corresponding to more consistent and accurate results. Again we find CoastalDEM v3.0 outperforms each of the competing DEMs, especially in terms of RMSE.

Elevation profiles in select cities comparing ICESat-2 (or NOAA Coastal Lidar, where available in the US), CoastalDEM v3.0, DiluviumDEM, and FABDEM are presented in Figures 4 and 5. We can see more clearly here that ICESat-2, even after applying the improvements as previously described, is still an imperfect truth set - there are substantial noise and "spikes" in these measurements that can exceed multiple meters. Regardless, we see that CoastalDEM v3.0 generally appears to generate a smoother and more well-fitting curve that the other DEMs. In Houston and Charleston, where NOAA Coastal Lidar is available, CoastalDEM v3.0 and DiluviumDEM produce more or less matching curves. Severe noise becomes clearly visible within DiluviumDEM and FABDEM in global cities, especially Amsterdam, Dhaka, Jakarta, and Shanghai.

We also explore a qualitative assessment of the DEM rasters themselves in Figure 6. While we do not have a lidar-derived DEM to directly assess accuracy in these locations, patterns and clear artifacts are visible within these rasters. For instance, we see that DiluviumDEM appears to contain very high-frequency noise, likely a direct result of using a decision tree as the computational model. FABDEM uses a random forest instead and so appears more robust to such high frequency random errors, but lacks sharp features and appears to contain distinct regions of land that are anomalously too high (i.e., low-frequency spatially autocorrelated noise), a reflection of what we have already seen in the elevation beam profiles in Figures 4 and 5. Meanwhile, with CoastalDEM v3.0 employing a powerful convolutional neural network and utilizing multiple input base DEMs, the resultant rasters are overall smooth, contain minimal high- or low-frequency noise, and are even able to clearly resolve fine details such as some streets.

Table 1. Global error statistics across each DEM, three elevation thresholds (2 m, 5 m, and 10 m), and three population density bands (any density (Any), more than 1,000 people per km^2 (>1K), and more than 10,000 people per km^2 (>10K)). ICESat-2 is used as ground truth. For each row, only pixels are included whose elevation falls below the elevation threshold (according to ground truth or the DEM), and whose population density falls within the given band. Rows presenting CoastalDEM v3.0 statistics are in bold. All units are in meters except for population density, which is people per km^2.

DEM	Max Elev	Pop Density	Mean Bias	Median Bias	RMSE	LE90
CoastalDEM v3.0	**2**	**Any**	**0.15**	**0.07**	**1.35**	**1.25**
CoastalDEM v2.1	2	Any	0.21	0.16	1.91	2.62
DiluviumDEM	2	Any	0.33	0.07	2.26	1.89
FABDEM	2	Any	0.65	0.08	2.61	2.53
COPDEM	2	Any	1.10	0.09	4.58	4.81
CoastalDEM v3.0	**5**	**Any**	**-0.02**	**-0.03**	**1.34**	**1.26**
CoastalDEM v2.1	5	Any	-0.04	-0.01	2.01	2.77
DiluviumDEM	5	Any	0.27	0.07	1.88	1.80
FABDEM	5	Any	0.50	0.05	2.22	2.17
COPDEM	5	Any	1.05	0.09	3.90	3.99
CoastalDEM v3.0	**10**	**Any**	**-0.13**	**-0.10**	**1.53**	**1.34**
CoastalDEM v2.1	10	Any	-0.22	-0.12	2.20	3.03
DiluviumDEM	10	Any	0.25	0.09	1.78	1.84
FABDEM	10	Any	0.41	0.04	2.06	2.02
COPDEM	10	Any	0.99	0.09	3.61	3.64
CoastalDEM v3.0	**2**	**>1K**	**0.26**	**0.23**	**1.25**	**1.54**
CoastalDEM v2.1	2	>1K	0.30	0.32	1.88	2.83
DiluviumDEM	2	>1K	0.23	0.12	1.42	1.97
FABDEM	2	>1K	0.52	0.26	1.73	2.29
COPDEM	2	>1K	1.57	0.86	3.00	4.52
CoastalDEM v3.0	**5**	**>1K**	**-0.07**	**-0.02**	**1.25**	**1.49**
CoastalDEM v2.1	5	>1K	0.00	0.08	1.89	2.75
DiluviumDEM	5	>1K	0.13	0.05	1.41	1.89
FABDEM	5	>1K	0.36	0.17	1.60	2.12
COPDEM	5	>1K	1.60	0.88	3.01	4.71
CoastalDEM v3.0	**10**	**>1K**	**-0.18**	**-0.08**	**1.36**	**1.62**
CoastalDEM v2.1	10	>1K	-0.27	-0.13	2.16	3.07
DiluviumDEM	10	>1K	0.11	0.04	1.50	2.00
FABDEM	10	>1K	0.34	0.16	1.67	2.20
COPDEM	10	>1K	1.64	0.90	3.09	4.83
CoastalDEM v3.0	**2**	**>10K**	**0.38**	**0.40**	**1.27**	**1.81**
CoastalDEM v2.1	2	>10K	0.93	1.09	2.33	3.44
DiluviumDEM	2	>10K	0.11	0.15	1.85	2.73
FABDEM	2	>10K	0.55	0.29	2.06	2.85
COPDEM	2	>10K	3.02	2.35	4.43	6.95
CoastalDEM v3.0	**5**	**>10K**	**-0.21**	**-0.07**	**1.51**	**1.84**
CoastalDEM v2.1	5	>10K	0.46	0.67	2.37	3.20
DiluviumDEM	5	>10K	0.23	0.16	1.98	2.71
FABDEM	5	>10K	0.56	0.31	2.14	2.79
COPDEM	5	>10K	3.10	2.39	4.56	7.08
CoastalDEM v3.0	**10**	**>10K**	**-0.42**	**-0.18**	**1.93**	**2.29**
CoastalDEM v2.1	10	>10K	-0.11	0.17	2.84	3.86
DiluviumDEM	10	>10K	0.15	0.12	2.24	3.15
FABDEM	10	>10K	0.54	0.30	2.37	3.26
COPDEM	10	>10K	3.21	2.44	4.78	7.52

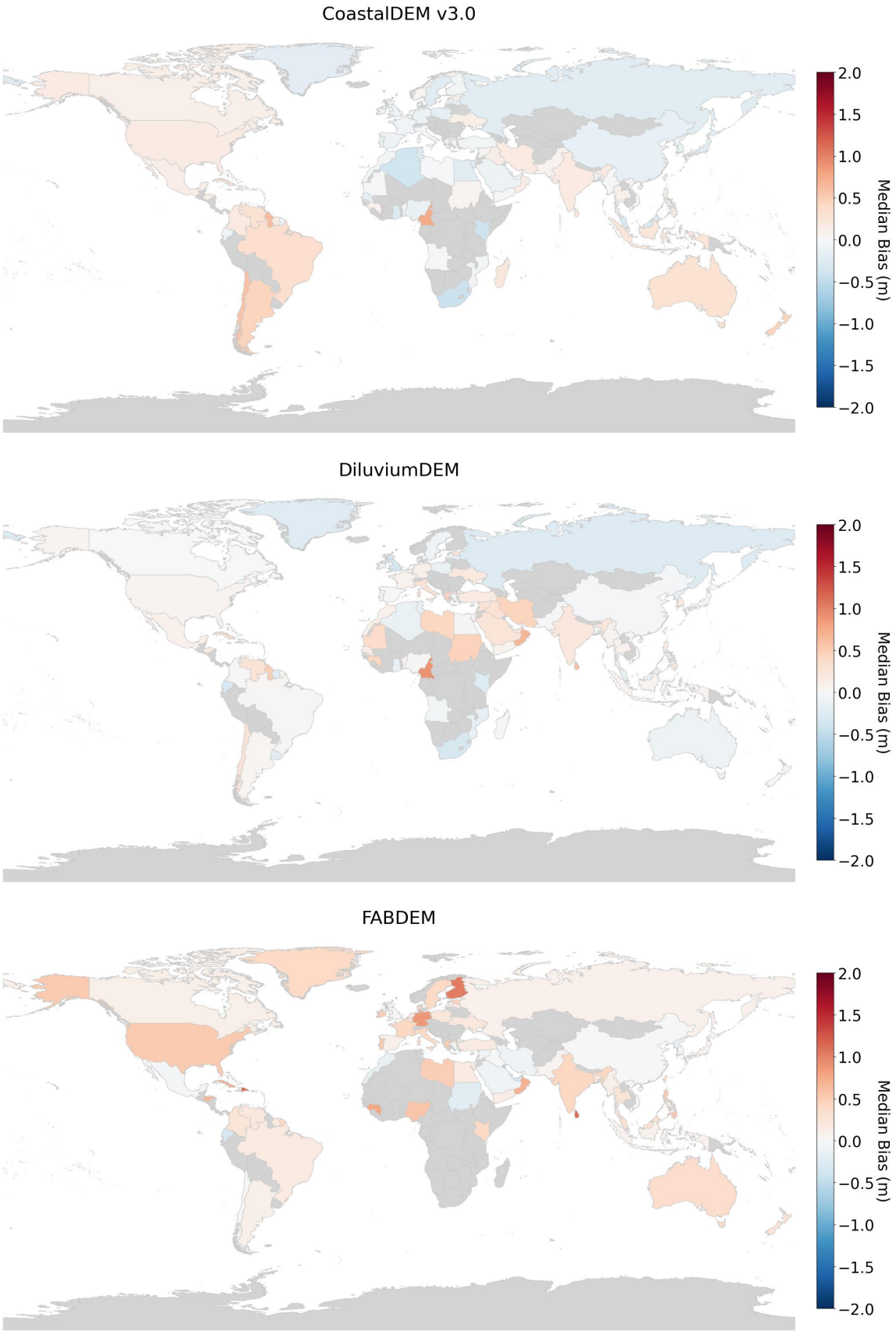

Figure 1. Choropleths presenting median bias under CoastalDEM v3.0, DiluviumDEM, and FABDEM in low-elevation regions across coastal nations, using ICESat-2 as ground truth. Only grid cells with elevation <5 m and population density >1000 people per km^2 are considered, and only nations with $n \geq 1000$ of these grid cells are evaluated.

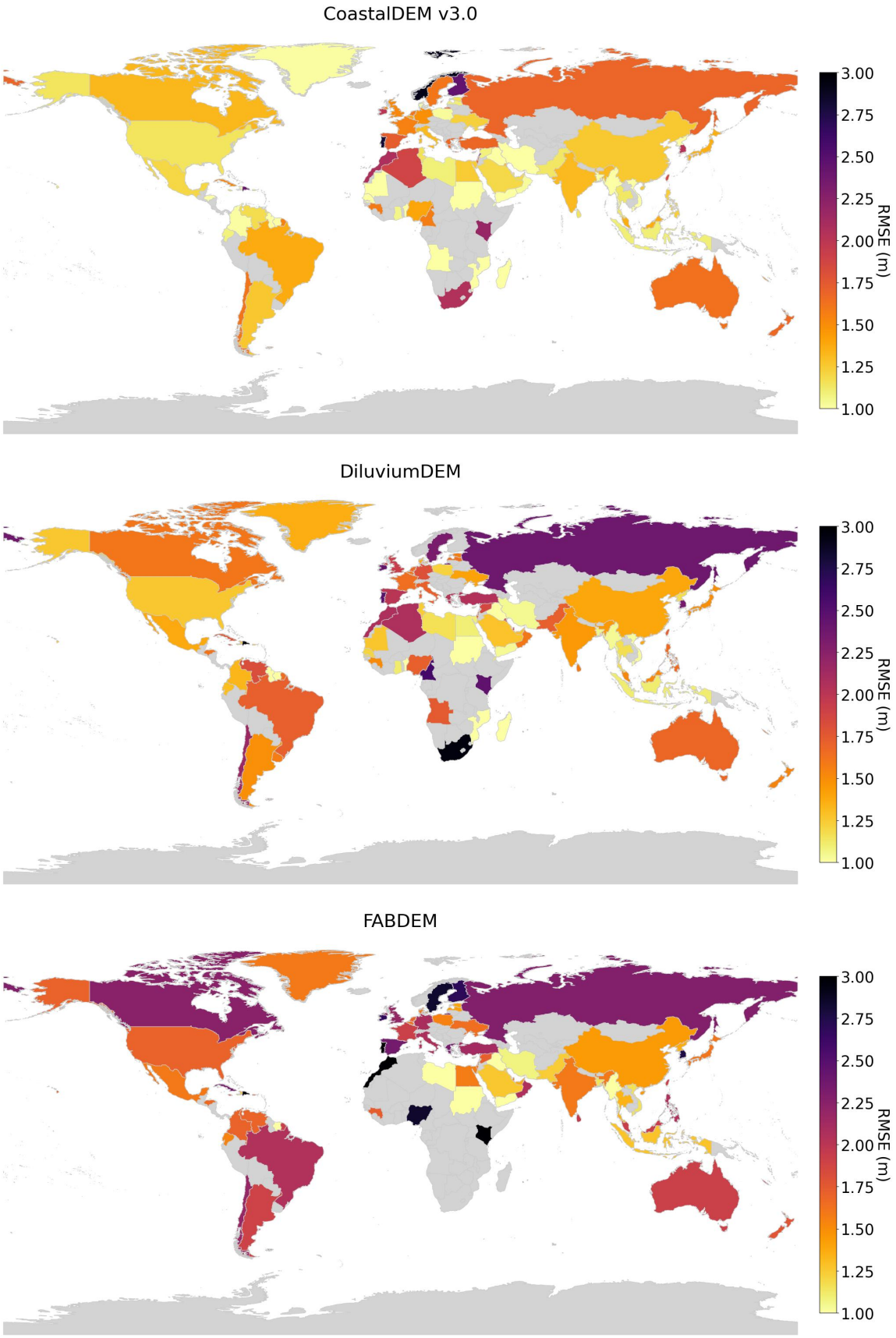

Figure 2. Choropleths presenting RMSE under CoastalDEM v3.0, DiluviumDEM, and FABDEMin low-elevation regions across coastal nations, using ICESat-2 as ground truth. Only grid cells with elevation <5 m and population density >1000 people per km^2 are considered, and only nations with $n \geq 1000$ of these grid cells are evaluated.

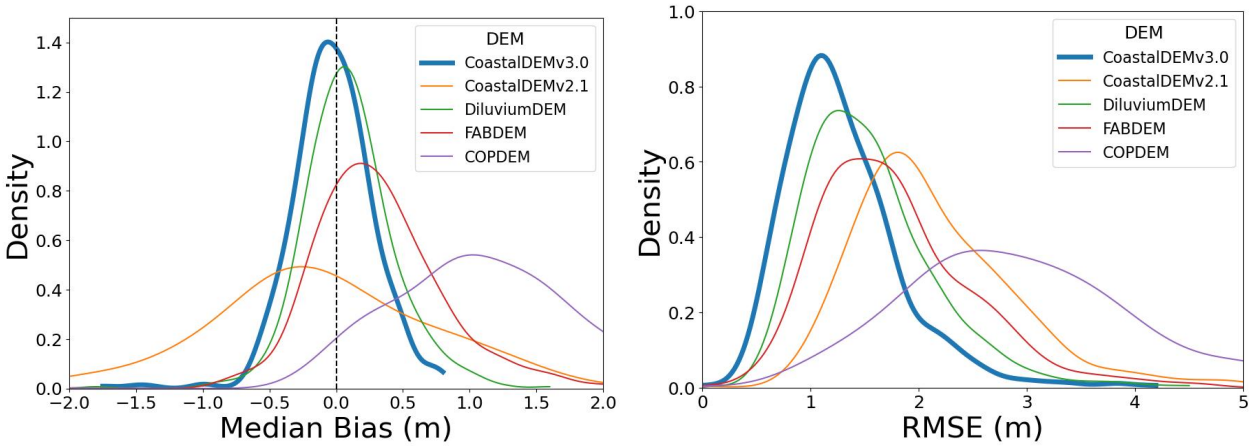

Figure 3. Density plots of median bias (left) and RMSE (right) for each of the global DEMs across level-1 administrative units (GADM 2.0), using ICESat-2 as ground truth. CoastalDEM v3.0 is highlighted in blue. Only grid cells whose elevations are lower than 5 m and contain >1000 people per square km are considered. Taller, sharper peaks closer to zero on on the x-axis represent higher, more consistent global accuracy.

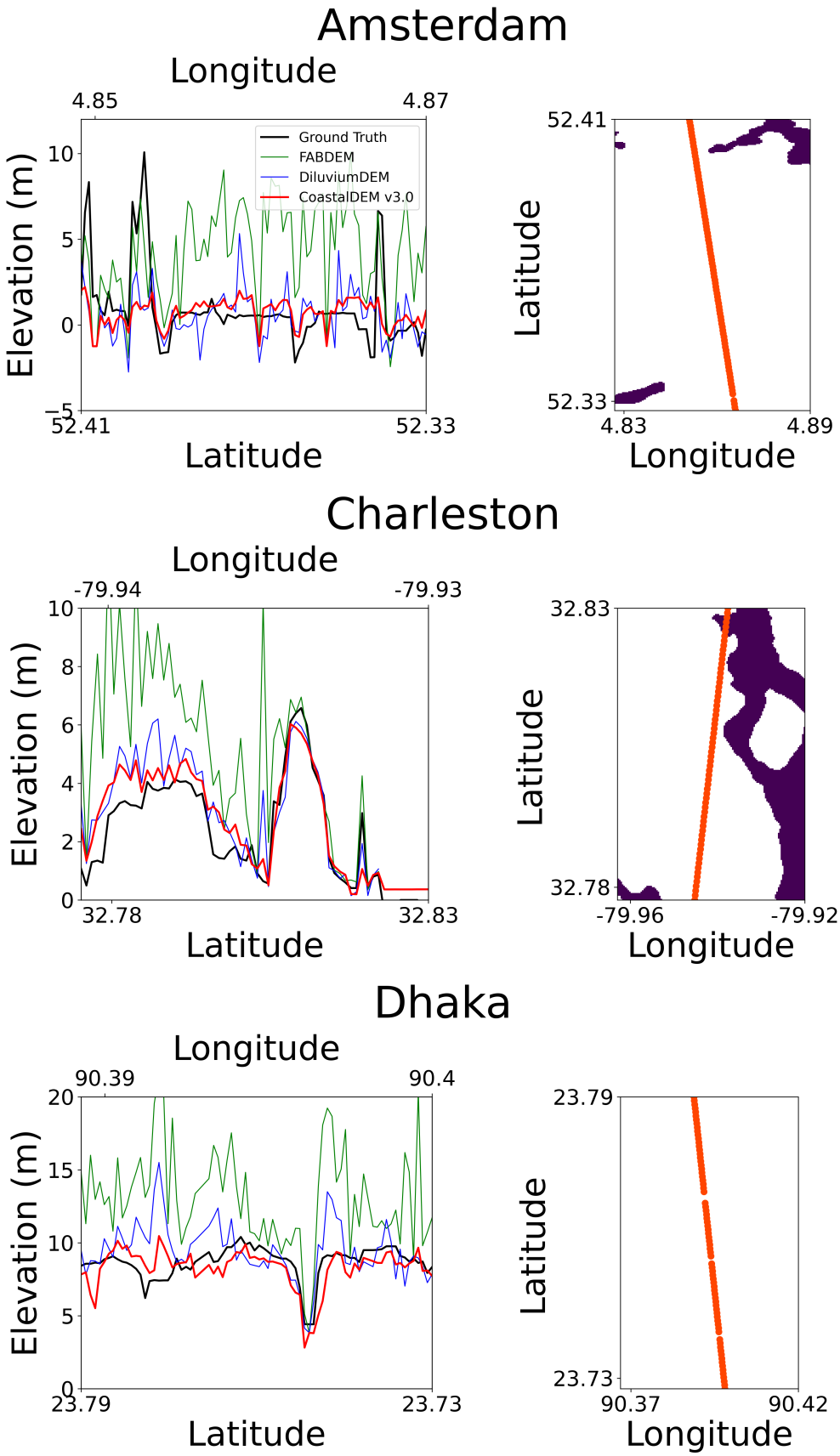

Figure 4. Elevation profiles under CoastalDEM v3.0, DeluviumDEM, and FABDEM in Amsterdam, Charleston, and Dhaka along an ICESat-2 beam path. For each city, the left panel presents estimated elevation along the path according to each dataset, with ground truth (NOAA lidar in the US, and ICESat-2 elsewhere) and CoastalDEM v2.1 highlighted in black and red, respectively. The right panel shows a map view where the path lies on the city in red, with water bodies highlighted in purple.

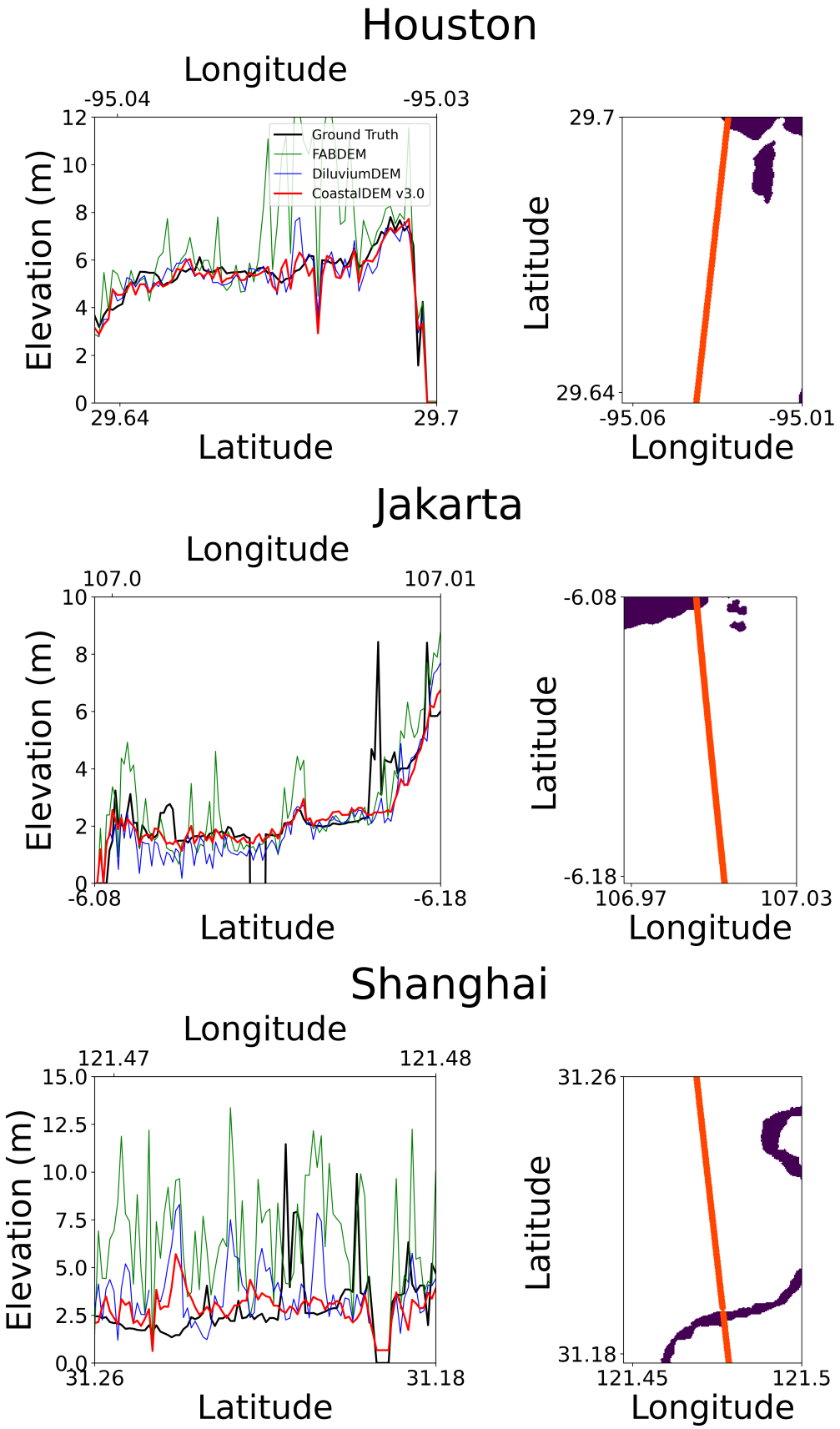

Figure 5. Elevation profiles under CoastalDEM v3.0, DeluviumDEM, and FABDEM in Houston, Jakarta, and Shanghai along an ICESat-2 beam path. For each city, the left panel presents estimated elevation along the path according to each dataset, with ground truth (NOAA lidar in the US, and ICESat-2 elsewhere) and CoastalDEM v2.1 highlighted in black and red, respectively. The right panel shows a map view where the path lies on the city in red, with water bodies highlighted in purple.

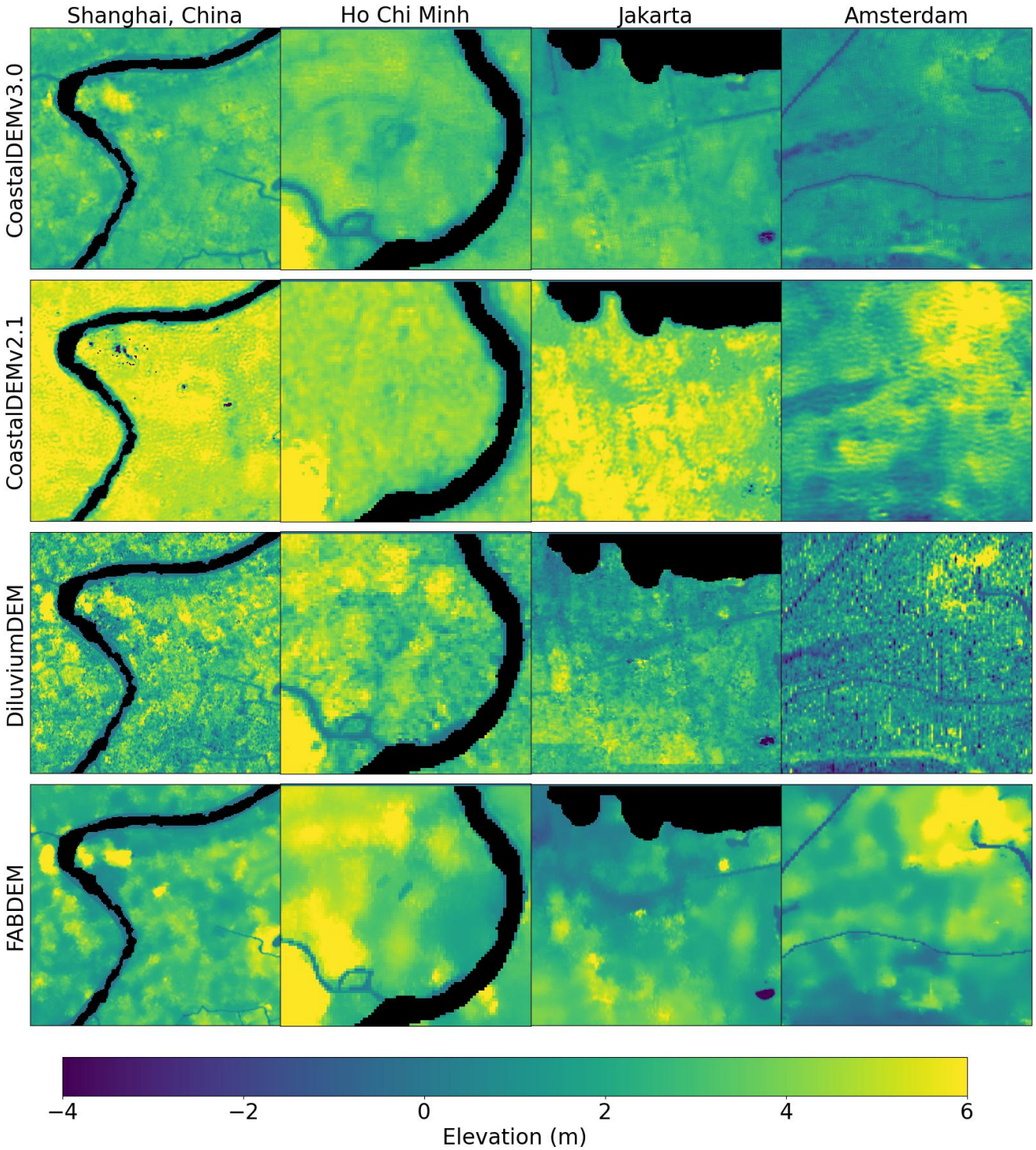

Figure 6. Maps of select international cities presenting zoomed image samples of the different global DEMs (CoastalDEM v3.0, CoastalDEM v2.1, DiluviumDEM, and FABDEM). Black areas represent existing water bodies. Note the color scale starts at a negative value to reflect the full distribution of elevation measurements across the DEMs.

3.2 Validation against airborne lidar-derived DEMs

While ICESat-2 v6 is the best elevation data source with global coverage presently available, the fact that we train the CNN using it as ground truth means we risk misstating accuracy if ICESat-2 is our only validation. For instance, systematic errors present in ICESat-2 measurements could potentially have been learned by the neural network and propagated across the output dataset. Further, while we use all available and applicable ICESat 2 measurements to assess the DEMs, a small fraction (under 20%) of them was also used to train the CNN model, potentially skewing the results. To resolve these concerns, we additionally use two high-accuracy elevation DEMs derived from airborne lidar as ground truth in separate error assessments.

As previously discussed, in the United States, NOAA makes publicly available high-quality DEMs across the entire US coastline, which are classified to bare earth elevation, with vertical errors <20 cm RMSE [32]. These data are released at about 5 m horizontal resolution, which we downsample to 1 arc-second (about 30 m) using median filtering. Meanwhile, in Australia, Geospace Australia [2] collected and publicly released bare-earth lidar-derived elevation data along much of their coastlines. These data offer <16 cm vertical RMSE [33] at roughly 25 m horizontal resolution, which we also downsample to 1-arcsecond to match CoastalDEM v3.0.

National results for both the US and Australia are presented in Table 2. We focus on grid cells with population densities exceeding 1,000 per square kilometer. In general we see that CoastalDEM v3.0 performs consistently well when assessed against these lidar-derived models. In the US, CoastalDEM v3.0 shows a large improvement when compared to CoastalDEM v2.1, FABDEM, and COPDEM, with RMSE/LE95 at 1.15 m/1.25 m at the 5 m elevation threshold, and 1.36 m/1.44 m at the 2 m threshold. These numbers are also lower than DiluviumDEM by 5-20cm, though CoastalDEM v3.0 contains a somewhat elevated bias here at the 2 m threshold at about 20 cm. That said, DiluviumDEM was specifically trained using this same lidar-derived elevation data in the US, so its strong performance here is to be expected. That CoastalDEM v3.0 is at parity with DiluviumDEM in the US, even though it was trained with noisier global data, is notable and is promising evidence that CoastalDEM is not overfitting on the ICESat-2 data.

Similar results are seen when comparing against the Australian lidar-derived data, though we caution that the domain size is considerably smaller here, given the substantially smaller high-density and low-elevation area. Again we see CoastalDEM v3.0 and DiluviumDEM are the two best performers, with CoastalDEM v3.0 offering lower LE90 by 20-40 cm across all elevation bands, though with worse RMSE below the 2 m elevation threshold, and roughly equivalent RMSE at the other thresholds.

Directly comparing these results in Table 2 against those presented in Supplementary Dataset 1 (the country-level vertical assessment using ICESat-2 as ground truth), we find overall good agreement with the same conclusions reached. In the US across each of the DEMs, the lidar-derived validations tend to reveal lower bias though higher RMSE than the ICESat-based analysis. Both tables show CoastalDEM v3.0 contains a positive bias roughly 10-20 cm higher than DiluviumDEM depending on elevation threshold. However, both also show that CoastalDEM offers a slightly (1-15%) lower RMSE and LE90 than DiliviumDEM. Across the board, deviations between ICESat-2 and lidar-derived error assessments do not exceed 30 cm, and in most cases do not exceed 20 cm.

In Australia, the deviations between the ICESat-2 and lidar-based validations are more pronounced but again present the same conclusions. There are not enough ICESat-2 points within the <2 elevation threshold for a valid comparison, but below 5 m and 10 m, ICESat-derived RMSE's for CoastalDEM and DiluviumDEM are about 13-25 cm higher than the lidar-derived values, and ICESat-derived LE90's are 26-55 cm higher. Regardless, the conclusions drawn in Australia are alone are essentially the same whether using ICESat or lidar: RMSE's for both CoastalDEM and DiluviumDEM are very close, while CoastalDEM offers a consistently though moderately improved LE90.

Figure 7 presents error maps in select cities in the US and Australia. Colors represent the difference between elevation according to the designated global DEM and the corresponding lidar-derived DEM. We can see how CoastalDEM v3.0 performs strongly relative to the other DEMs overall, with it clearly outperforming CoastalDEM v2.1 and FABDEM. In the US locations, we see that the results are excellent, and overall very similar to DiluviumDEM. In Brisbane, the differences between CoastalDEM v3.0 and DiluviumDEM show a clearer difference, with CoastalDEM offering lower errors in the great majority of the city, but a distinct sharp positive error region directly adjacent to the river's border. This may be an explanation for CoastalDEM's elevated Australian RMSE in the <2 m domain previously seen in Table 2.

Finally, US state-level choropleths of median bias and RMSE for each global DEM can be found in Figures 8 and 9. Again considering points below 5 m and with >1,000 people per square kilometer, we find that CoastalDEM v3.0 offers consistent performance across smaller administrative areas.

These error statistics derived from DEMs based on airborne lidar are overall similar to the global results using data based on ICESat-2 satellite lidar. The airborne lidar ground-truth values were not used in computing CoastalDEM v3.0. The consistency in error assessment across testing approaches mitigates concerns about potential overfitting of our neural network model, and suggests that we can reasonably accept that results based from our computed global error statistics should hold.

Table 2. Error statistics in the USA and Australia across each DEM and three elevation thresholds (2 m, 5 m, and 20 m). Airborne lidar-derived elevation data are used as ground truth. For each row, only pixels are included whose elevation falls below the elevation threshold (according to ground truth or the DEM), and whose population density exceeds 1K per square kilometer. Rows presenting CoastalDEM v3.0 statistics are in bold. All units are in meters

Nation	DEM	Max Elev	Mean	Median	RMSE	LE90
USA	**CoastalDEM v3.0**	**2**	**0.18**	**0.23**	**1.36**	**1.44**
USA	CoastalDEM v2.1	2	0.23	0.36	1.98	2.72
USA	DiluviumDEM	2	0.03	0.02	1.38	1.59
USA	FABDEM	2	0.61	0.46	1.93	2.41
USA	COPDEM	2	2.14	1.59	3.55	5.06
USA	**CoastalDEM v3.0**	**5**	**0.04**	**0.10**	**1.15**	**1.25**
USA	CoastalDEM v2.1	5	-0.14	-0.07	1.93	2.83
USA	DiluviumDEM	5	0.00	0.00	1.20	1.48
USA	FABDEM	5	0.66	0.49	1.69	2.29
USA	COPDEM	5	2.32	1.72	3.61	5.25
USA	**CoastalDEM v3.0**	**10**	**-0.02**	**0.04**	**1.15**	**1.32**
USA	CoastalDEM v2.1	10	-0.29	-0.21	2.11	3.09
USA	DiluviumDEM	10	-0.01	-0.02	1.20	1.55
USA	FABDEM	10	0.69	0.52	1.70	2.31
USA	COPDEM	10	2.45	1.79	3.77	5.51
Australia	**CoastalDEM v3.0**	**2**	**0.08**	**0.38**	**2.16**	**1.87**
Australia	CoastalDEM v2.1	2	-0.12	0.44	3.05	4.08
Australia	DiluviumDEM	2	-0.34	-0.38	1.75	2.27
Australia	FABDEM	2	0.72	0.40	1.88	2.45
Australia	COPDEM	2	1.85	1.29	3.28	4.85
Australia	**CoastalDEM v3.0**	**5**	**0.10**	**0.20**	**1.50**	**1.57**
Australia	CoastalDEM v2.1	5	-0.28	0.06	2.49	3.64
Australia	DiluviumDEM	5	-0.08	-0.18	1.43	1.81
Australia	FABDEM	5	0.59	0.33	1.51	2.04
Australia	COPDEM	5	1.79	1.30	2.86	4.09
Australia	**CoastalDEM v3.0**	**10**	**0.00**	**0.11**	**1.52**	**1.69**
Australia	CoastalDEM v2.1	10	-0.78	-0.38	3.01	4.55
Australia	DiluviumDEM	10	-0.06	-0.16	1.55	1.97
Australia	FABDEM	10	0.52	0.26	1.57	2.07
Australia	COPDEM	10	1.81	1.31	2.92	4.08

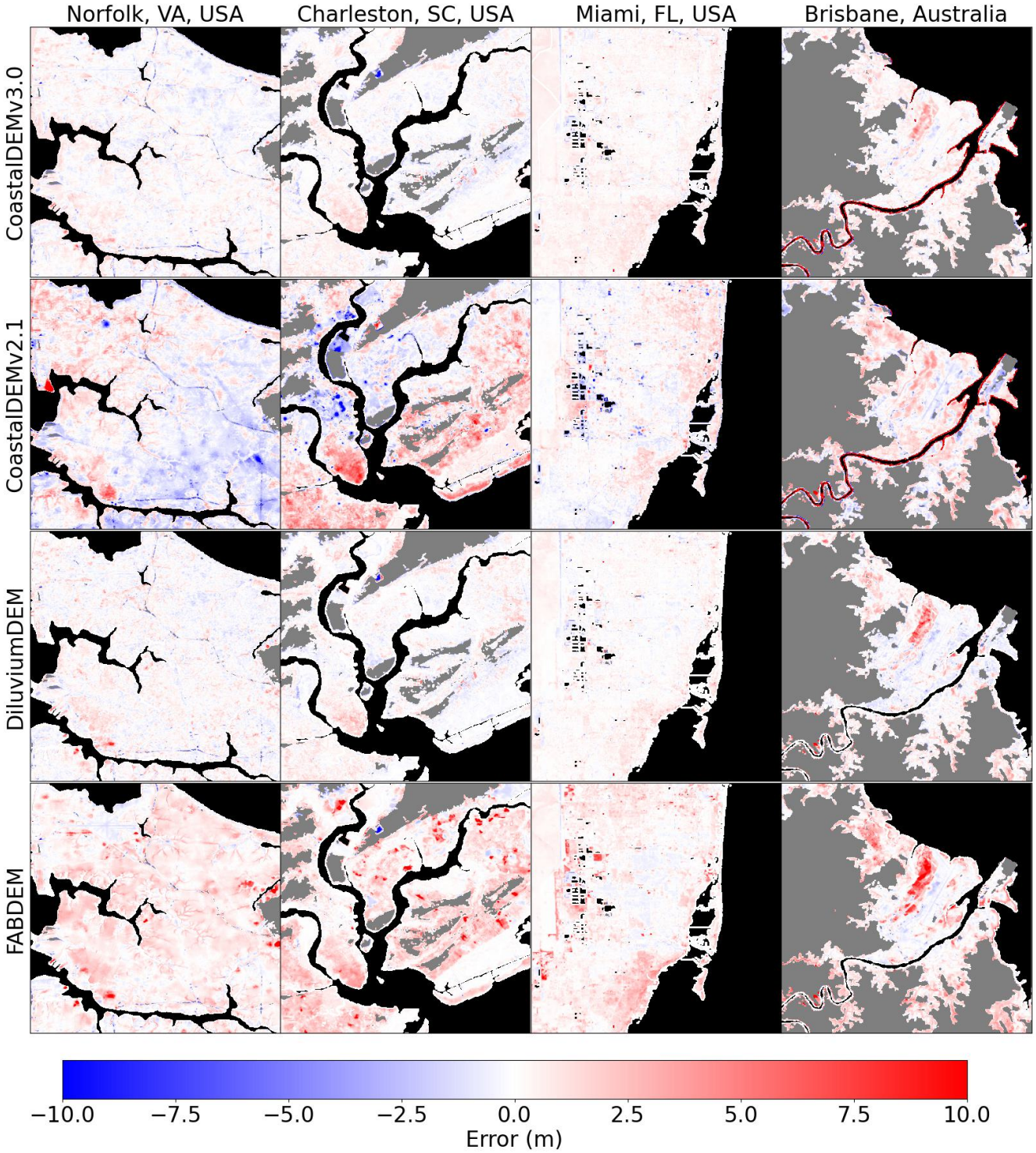

Figure 7. Maps of select US and Australian cities presenting the difference between global DEMs (CoastalDEM v3.0, CoastalDEM v2.1, DiluviumDEM, and FABDEM) and a lidar-derived DEM. Black areas represent existing water bodies, and gray areas represent pixels whose elevation exceeds 20m.

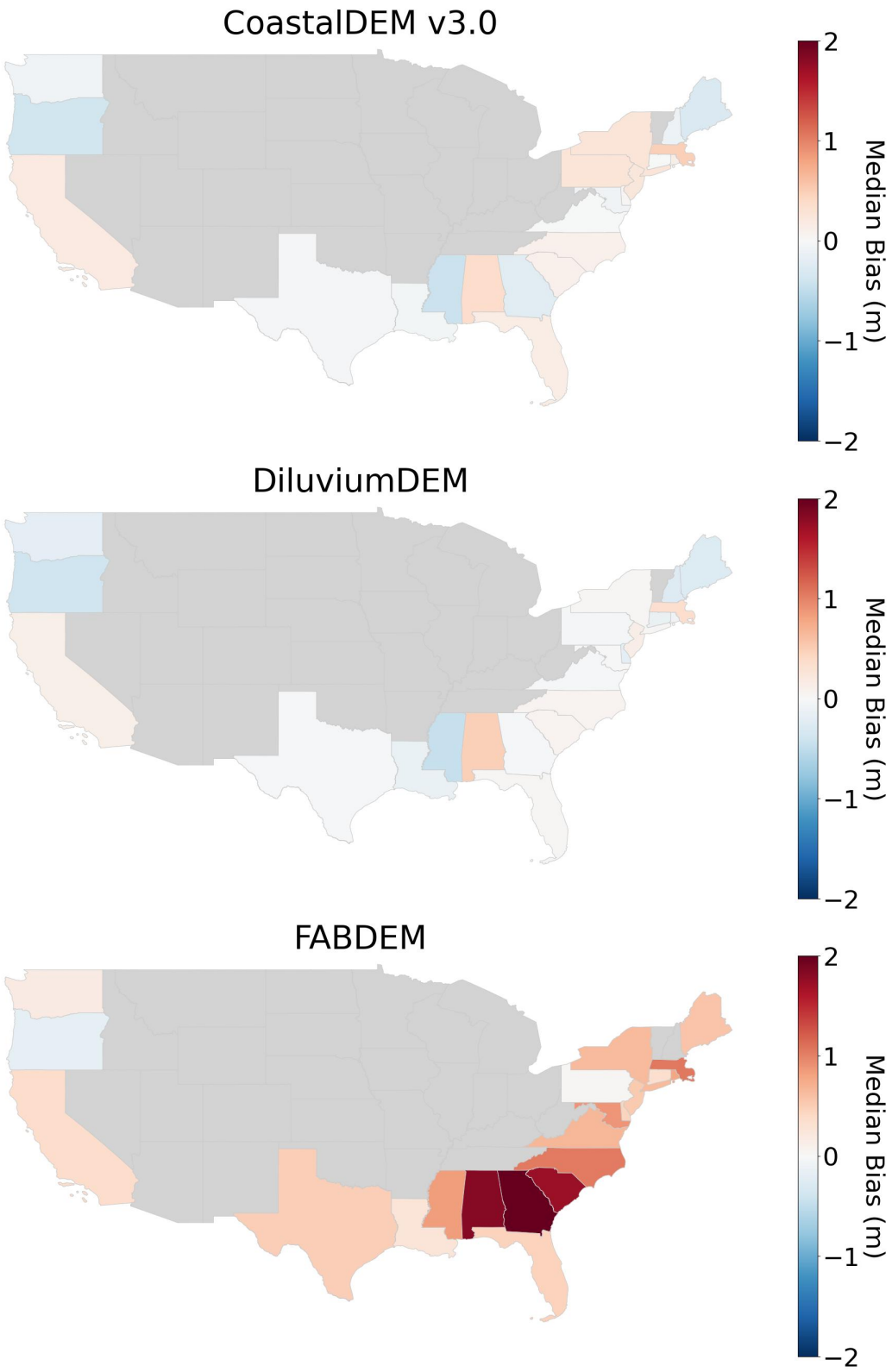

Figure 8. Choropleths presenting median bias under CoastalDEM v3.0, DiluviumDEM, and FABDEM in low-elevation regions across US states, using elevation data from NOAA's coastal lidar as ground truth. Only pixels whose elevations are lower than 5 m are considered. Only areas with population densities above 1,000 people per square kilometer are included.

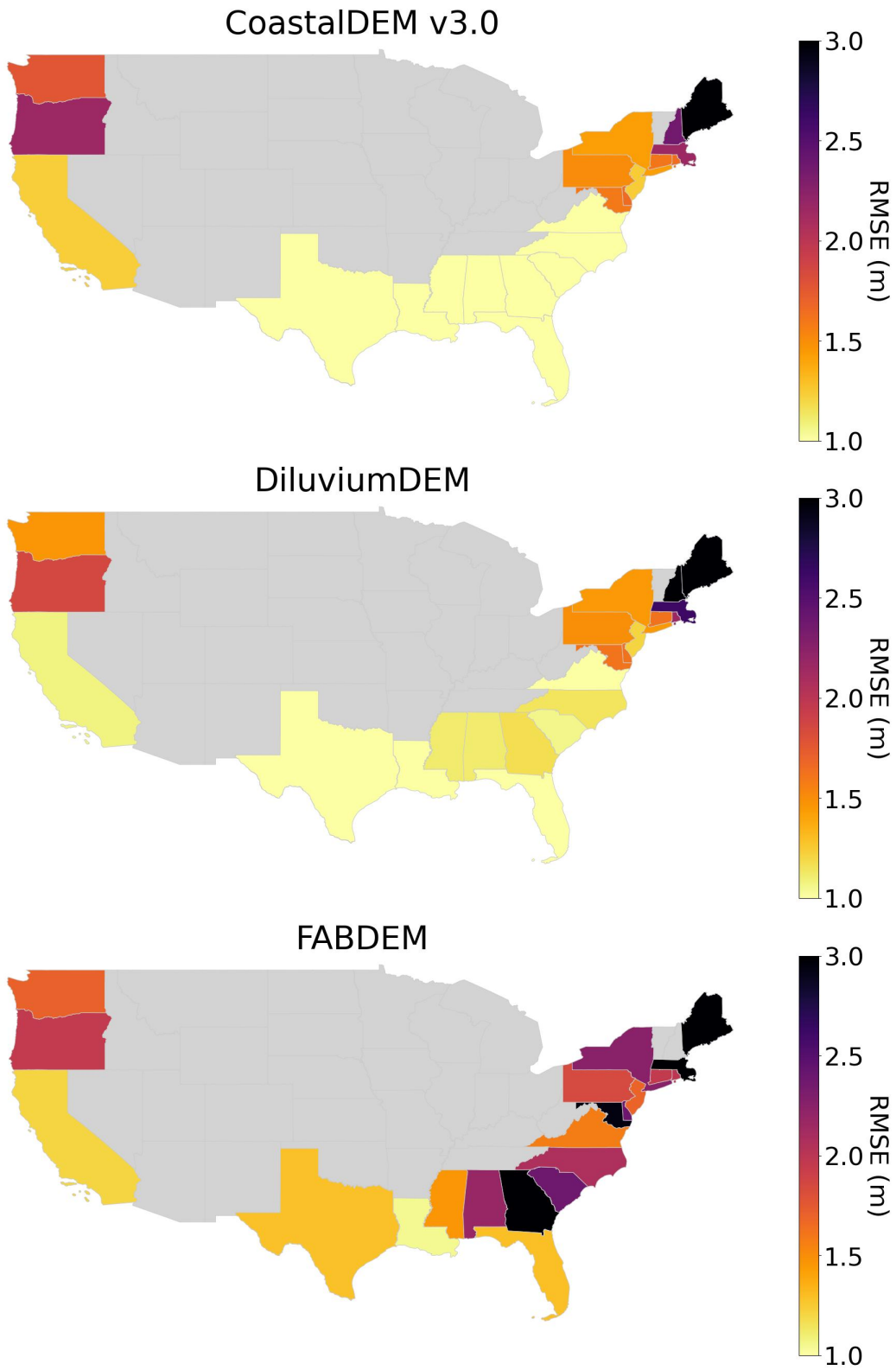

Figure 9. Choropleths presenting median RMSE under CoastalDEM v3.0, DiluviumDEM, and FABDEM in low-elevation regions across US states, using elevation data from NOAA's coastal lidar as ground truth. Only pixels whose elevations are lower than 5 m are considered. Only areas with population densities above 1,000 people per square kilometer are included.

4. Discussion

Climate Central has invested and will continue to invest significant resources and energy into improving CoastalDEM. As more and improved additional data sets become available, we intend to add them in improving the neural network.

At the same time, we acknowledge that neither CoastalDEM nor any global product is likely to ever outperform DEMs based on high-quality airborne lidar elevation data. Probably the main reason that airborne lidar data have not been collected over most areas, and DEMs have not been generated, is the high cost. Coastal jurisdictions able to develop lidar-based DEMs will improve their sea level rise and coastal flooding risk assessments, and we strongly encourage this development.

We also acknowledge that the original DEM data from which CoastalDEM is derived was collected before year 2015. The surface of the earth is changing with time, especially in areas prone to subsidence due to high rates of groundwater or fossil fuel extraction, or river-delta-sediment compaction. In addition, artificial earth works have the potential to alter the coastal risk profiles represented by all currently-available global DEMs. This temporal quality calls for more up-to-date and regular refreshes of coastal DEMs with airborne lidar and new remote sensing capabilities that may become available.

5. Conclusion

CoastalDEM was developed to provide an improved, widely available, near-global digital elevation model for the primary purpose of evaluating coastal flood risk considering storms and sea level rise. With this use case in mind, elevations below 5 m are of particular interest as they span the range of most tides, storms, and projected sea-level-rise scenarios through the year 2100.

In addition, coastal areas with high population density are both areas where accurate vulnerability assessments are especially important and areas where the urbanized, built environment has challenged remote sensing technologies intended to measure ground elevations, resulting in material vertical bias that negatively impacts coastal flood risk assessments.

Performance data indicate error scatter is consistently and substantially reduced with CoastalDEM v3.0. CoastalDEM v3.0 is particularly strong in the elevation range below 2 m and 5 m where coastal flood risk is acute and in densely populated regions where buildings and the built environment adversely affect other global DEMs. Near-zero bias and low noise and error scatter means smaller elevation errors propagating into coastal flood analysis so critical to understanding the threat posed by sea level rise.

6. Availability

CoastalDEM v3.0, referenced to EGM96 or EGM2008, is available at 30 m and 90-m horizontal resolution by license from Climate Central via `https://go.climatecentral.org/coastaldem/`. No-cost, non-commercial licenses at 90 m horizontal resolution are available to qualified academic and research organizations. Rasterized ICESat-2 data processed to reduce the effects of buildings in high-density areas are also available upon request. With no-cost licenses available and superior RMSE's, CoastalDEM v3.0 is the lowest-error global DEM available for sea level rise and coastal flood risk research and assessments.

7. Methods

7.1 ICESat-2 v6

NASA distributes ICESat-2 v6 measurements as a large collection of HDF5 files. Here, we download the entirety of the L3A Land and Vegetation Height Version 6 (ATL08) dataset [34], which contains a number of elevation metrics at points 12 m apart along six beam tracks. For each point, we extract the fields $h\_te\_best\_fit$, $latitude$, $longitude$, and $layer\_flag$. The variable $h\_te\_best\_fit$ refers to the best fit height returned by photons at the midpoint within the point's footprint, and $layer\_flag$ is a binary variable that is 1 if the point is likely covered by snow or clouds (points flagged as such are removed). All points in the entire ICESat-2 dataset meeting the given requirements and filters described in this report were used in the assessments.

7.2 Building Removal

For each segment in the ATL08 data set, we first classify each segment on whether their midpoints are on land using COPDEM's water body mask, as well as whether their midpoints are on a building using data from the 2019 World Settlement Footprint dataset [35]. Here, we skip any segments that do not cover a building, or whose population density is below 10,000 people per square kilometer, and leave them unaltered in the final product.

If a segment in a population-dense area does contain a building, then its terrain height measurements are unreliable. For such cases, we propose a straightforward operation to improve its accuracy. Given a segment s, we construct an adjusted elevation function, E^*:

$$E^*(s,n,i) = \texttt{percentile}(E(\texttt{neighbors}(s,n)),i), \tag{1}$$

where $E(\texttt{neighbors(s,n)})$ is the set of the n closest neighboring measurements to s (including s itself) along its track, and i is the desired percentile of that set of measurements. For instance, if $n = 7$, and $i = 50$, $E^*(s,7,50)$ represents the 50th percentile of the terrain height measurements of s and the three segments on both of its sides. We use linear interpolation to estimate intermediate percentiles. In effect, this approach represents a "sliding percentile" function along a given beam track. Note that this function is calculable where a point's nearest 6 neighbors are valid (i.e., not flagged nor removed).

We then must determine an optimal value of n and i that minimizes adjusted elevation error. We have found that in cases where buildings are fairly sparse, simply using the 7-point median ($n = 7$, $i = 50$) performs well. However, in dense urban environments, it is often the case that more than half of the 7 points cover buildings, making even this median value biased too high. In these situations, we have empirically found that setting $i = 45$ produces more accurate results. In both cases, we tuned these parameters based on accuracy assessments within the US, detailed below.

In summary, terrain height measurements for any given segment at locations whose population density exceed 10,000 per square kilometer are set to:

$$E^{(final)}(s) = \begin{cases} E(s), & \text{if } s \text{ does not contain a building or any of its 6 neighboring points are invalid} \\ E^*(s,7,50), & \text{if } s \text{ and fewer than half of its neighbors contain a building} \\ E^*(s,7,45), & \text{if } s \text{ and at least half of its neighbors contain a building} \end{cases}$$

Results of these adjustments compared to US and Australia lidar are presented in Table 3. Since the land area exceeding 10,000 ppsk is small compared to other coastal areas (generally less than 2% of all populated land), the overall error statistics for land >0 or >1000 ppsk do not meaningfully change. However, there is a clear and important difference in extremely high density but low elevation areas, cutting vertical bias by at least 75% and reducing RMSE by 10-23%.

Location	ICESat Metric	Max Elev	Min Pop Dens (ppsk)	Num Pts	Original Bias	Adjusted Bias	Linear Bias Change	Original RMSE	Adjusted RMSE	RMSE % Change
USA	best_fit	5m	>0	51774	0.03	0.02	**-0.01**	1.51	1.49	**-1%**
USA	best_fit	5m	>1000	25936	-0.05	-0.07	0.02	1.55	1.52	**-2%**
USA	best_fit	5m	>10000	691	0.40	0.10	**-0.30**	2.34	1.81	**-23%**
USA	best_fit	20m	>0	120697	0.11	0.11	0.00	1.44	1.43	**-1%**
USA	best_fit	20m	>1000	61212	0.04	0.03	**-0.01**	1.36	1.35	**-1%**
USA	best_fit	20m	>10000	2367	0.50	0.41	**-0.09**	2.27	2.12	**-7%**
Australia	best_fit	5m	>0	4705	0.22	0.21	**-0.01**	2.18	2.18	0%
Australia	best_fit	5m	>1000	2468	0.11	0.10	**-0.01**	2.08	2.08	0%
Australia	best_fit	5m	>10000	9	1.30	0.21	**-1.09**	2.30	2.08	**-10%**
Australia	best_fit	20m	>0	17786	0.36	0.36	0.00	1.84	1.84	0%
Australia	best_fit	20m	>1000	10790	0.24	0.24	0.00	1.65	1.64	**-1%**
Australia	best_fit	20m	>10000	70	2.02	1.88	**-0.14**	4.81	4.79	0%

Table 3. ICESat 2 terrain height vertical error before and after adjustment, assessed by population density. Negative change values represent improvement in the adjusted terrain height data.

7.3 CoastalDEM v3.0

Like CoastalDEM v2.1, CoastalDEM v3.0 uses a large convolutional neural network architecture [36] to predict errors present in other global DEM(s), using a number of global datasets as inputs. CNNs are specifically designed for and are widely used in tasks involving imagery, making them a good fit for the raster datasets used here. n artificial neural network to predict errors present in another global DEM (here, NASADEM), using a number of global datasets as inputs. These inputs include elevation, population density, and vegetation density and height metrics. In total, CoastalDEM v3.0 ingests 10 independent input datasets to feed the model, compared to CoastalDEM v2.1's seven datasets.

Both CoastalDEM v3.0 and v2.1 were trained on global ICESat-2 elevation models, though ICESat-2 has been upgraded from v3 to v6, which from our assessments offers considerably better accuracy. The building removal adjustments described above further enhance this truth set to optimize CoastalDEM v3's final performance.

Acknowledgments

The research to upgrade from CoastalDEM v2.1 to v3.0 was supported by Climate Central.

References

[1] NOAA, "Digital Coast Coastal Lidar," 2015. [Online]. Available: http://coast.noaa.gov/digitalcoast/data/coastallidar

[2] Geoscience Australia, "Digital Elevation Model (DEM) 25 Metre Grid of Australia derived from LiDAR," jan 2015. [Online]. Available: http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_89676

[3] UK Environment Agency, "SURVEY/LIDAR Composite 1m DTM 2022 Elevation," 2022. [Online]. Available: https://environment.data.gov.uk/image/rest/services/SURVEY/LIDAR_Composite_1m_DTM_2022_Elevation/ImageServer

[4] NASA JPL, "NASADEM Merged DEM Global 1 arc second V001. Distributed by OpenTopography." 2021. [Online]. Available: https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.032021.4326.2

[5] European Space Agency, "Copernicus Digital Elevation Model," 2022. [Online]. Available: https://spacedata.copernicus.eu/collections/copernicus-digital-elevation-model

[6] JAXA, "ALOS Global Digital Surface Model (AW3D30) Version 4.0," 2023. [Online]. Available: https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm

[7] S. A. Kulp and B. H. Strauss, "CoastalDEM v2.1: A high-accuracy and high-resolution global coastal elevation model trained on ICESat-2 satellite lidar," 2021. [Online]. Available: https://go.climatecentral.org/coastaldem/clkn/https/www.climatecentral.org/coastaldem-v2.1

[8] L. Hawker, P. Uhe, L. Paulo, J. Sosa, J. Savage, C. Sampson, and J. Neal, "A 30 m global map of elevation with forests and buildings removed," *Environmental Research Letters*, vol. 17, no. 2, p. 024016, feb 2022. [Online]. Available: https://iopscience.iop.org/article/10.1088/1748-9326/ac4d4f

[9] K. G. Nikolakopoulos, "Accuracy assessment of ALOS AW3D30 DSM and comparison to ALOS PRISM DSM created with classical photogrammetric techniques," *European Journal of Remote Sensing*, vol. 53, no. sup2, pp. 39–52, jul 2020. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/22797254.2020.1774424

[10] D. Wendi, S.-Y. Liong, Y. Sun, and C. D. Doan, "An innovative approach to improve SRTM DEM using multispectral imagery and artificial neural network," *Journal of Advances in Modeling Earth Systems*, vol. 8, no. 2, pp. 691–702, jun 2016.

[11] M. Meadows and M. Wilson, "A Comparison of Machine Learning Approaches to Improve Free Topography Data for Flood Modelling," *Remote Sensing*, vol. 13, no. 2, p. 275, jan 2021.

[12] D. Yamazaki, D. Ikeshima, R. Tawatari, T. Yamaguchi, F. O'Loughlin, J. C. Neal, C. C. Sampson, S. Kanae, and P. D. Bates, "A high-accuracy map of global terrain elevations," *Geophysical Research Letters*, vol. 44, no. 11, pp. 5844–5853, jun 2017.

[13] C. A. Baugh, P. D. Bates, G. Schumann, and M. A. Trigg, "SRTM vegetation removal and hydrodynamic modeling accuracy," *Water Resources Research*, vol. 49, no. 9, pp. 5276–5289, sep 2013.

[14] Y. Su, Q. Guo, Q. Ma, and W. Li, "SRTM DEM Correction in Vegetated Mountain Areas through the Integration of Spaceborne LiDAR, Airborne LiDAR, and Optical Imagery," *Remote Sensing*, vol. 7, no. 9, pp. 11 202–11 225, sep 2015.

[15] F. O'Loughlin, R. Paiva, M. Durand, D. Alsdorf, and P. Bates, "A multi-sensor approach towards a global vegetation corrected SRTM DEM product," *Remote Sensing of Environment*, vol. 182, pp. 49–59, 2016.

[16] S. A. Kulp and B. H. Strauss, "CoastalDEM: A global coastal digital elevation model improved from SRTM using a neural network," *Remote Sensing of Environment*, vol. 206, pp. 231–239, mar 2018.

[17] A. Neuenschwander, S. Popescu, R. Nelson, D. Harding, K. Pitts, J. Robbins, D. Pederson, and R. Sheridan, "ICE, CLOUD, and Land Elevation Satellite (ICESat-2) Algorithm Theoretical Basis Document (ATBD) for Land-Vegetation Along-track products (ATL08) Contributions by Land/VEG SDT Team Members and ICESAt-2 Project Science Office," Tech. Rep., 2018. [Online]. Available: https://icesat-2.gsfc.nasa.gov/sites/default/files/files/ATL08_15June2018.pdf

[18] German Aerospace Center (DLR), "TanDEM-X - Digital Elevation Model (DEM) - Global, 90m," 2018. [Online]. Available: https://doi.org/10.15489/ju28hc7pui09

[19] M. A. Ghannadi, S. Alebooye, M. Izadi, and A. Ghanadi, "VERTICAL ACCURACY ASSESSMENT OF COPERNICUS DEM (CASE STUDY: TEHRAN AND JAM CITIES)," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. X-4/W1-202, pp. 209–214, jan 2023. [Online]. Available: https://isprs-annals.copernicus.org/articles/X-4-W1-2022/209/2023/

[20] S. Franks and R. Rengarajan, "Evaluation of Copernicus DEM and Comparison to the DEM Used for Landsat Collection-2 Processing," *Remote Sensing*, vol. 15, no. 10, p. 2509, may 2023. [Online]. Available: https://www.mdpi.com/2072-4292/15/10/2509

[21] D. Dusseau, Z. Zobel, and C. R. Schwalm, "DiluviumDEM: Enhanced accuracy in global coastal digital elevation models," *Remote Sensing of Environment*, vol. 298, p. 113812, dec 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425723003632

[22] C. J. Okolie, J. P. Mills, A. K. Adeleke, J. L. Smit, M. V. Peppa, A. O. Altunel, and I. D. Arungwa, "Assessment of the global Copernicus, NASADEM, ASTER and AW3D digital elevation models in Central and Southern Africa," *Geo-spatial Information Science*, pp. 1–29, feb 2024. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/10095020.2023.2296010

[23] C. B. Marsh, P. Harder, and J. W. Pomeroy, "Validation of FABDEM, a global bare-earth elevation model, against UAV-lidar derived elevation in a complex forested mountain catchment," *Environmental Research Communications*, vol. 5, no. 3, p. 031009, mar 2023. [Online]. Available: https://iopscience.iop.org/article/10.1088/2515-7620/acc56d

[24] H. Li, J. Zhao, B. Yan, L. Yue, and L. Wang, "Global DEMs vary from one to another: an evaluation of newly released Copernicus, NASA and AW3D30 DEM on selected terrains of China using ICESat-2 altimetry data," *International Journal of Digital Earth*, vol. 15, no. 1, pp. 1149–1168, dec 2022. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/17538947.2022.2094002

[25] E. Uuemaa, S. Ahi, B. Montibeller, M. Muru, and A. Kmoch, "Vertical Accuracy of Freely Available Global Digital Elevation Models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM)," *Remote Sensing*, vol. 12, no. 21, p. 3482, oct 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/21/3482

[26] C. Tebaldi, R. Ranasinghe, M. Vousdoukas, D. J. Rasmussen, B. Vega-Westhoff, E. Kirezci, R. E. Kopp, R. Sriver, and L. Mentaschi, "Extreme sea levels at different global warming levels," *Nature Climate Change*, vol. 11, no. 9, pp. 746–751, sep 2021.

[27] B. Fox-Kemper, H. T. Hewitt, G. C. Xiao, S. S. Aðalgeirsdóttir, T. L. Drijfhout, N. R. Edwards, M. Golledge, R. E. Kopp, G. Krinner, A. Mix, D. Notz, S. Nowicki, I. Nurhati, J.-B. L. Ruiz, A. B. A. Sallée, and Y. Y. Slangen, "Ocean, Cryosphere and Sea Level Change." in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, Eds. Cambridge University Press, 2021.

[28] A. Neuenschwander, E. Guenther, J. C. White, L. Duncanson, and P. Montesano, "Validation of ICESat-2 terrain and canopy heights in boreal forests," *Remote Sensing of Environment*, vol. 251, p. 112110, dec 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425720304831?via%3Dihub

[29] V. Moudrý, K. Gdulová, L. Gábor, E. Šárovcová, V. Barták, F. Leroy, O. Špatenková, D. Rocchini, and J. Prošek, "Effects of environmental conditions on ICESat-2 terrain and canopy heights retrievals in Central European mountains," *Remote Sensing of Environment*, vol. 279, p. 113112, sep 2022. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0034425722002267?via%3Dihub

[30] J. Zhu, P.-f. Yang, Y. Li, Y.-z. Xie, and H.-q. Fu, "Accuracy assessment of ICESat-2 ATL08 terrain estimates: A case study in Spain," *Journal of Central South University*, vol. 29, no. 1, pp. 226–238, jan 2022. [Online]. Available: https://link.springer.com/10.1007/s11771-022-4896-x

[31] University of Berkeley, Museum of Vertebrate Zoology, and International Rice Research Institute, "Global Administrative Areas (Boundaries)," 2012. [Online]. Available: http://www.gadm.org/

[32] NOAA, "Lidar 101: An Introduction to Lidar Technology, Data, and Applications," 2012. [Online]. Available: https://coast.noaa.gov/data/digitalcoast/pdf/lidar-101.pdf

[33] Intergovernmental Committee on Surveying and Mapping, "ICSM LiDAR Acquisition Specifications and Tender Template," 2010. [Online]. Available: https://www.icsm.gov.au/sites/default/files/2017-03/NZ-LiDAR_Specifications_and_Tender_Template.pdf

[34] A. L. Neuenschwander, K. Pitts, B. P. Jelley, J. Robbins, J. Markel, S. C. Popescu, R. F. Nelson, D. Harding, D. Pederson, B. Klotz, and R. Sheridan, "ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 6," 2023. [Online]. Available: https://nsidc.org/data/atl08/versions/6

[35] DLR, "World Settlement Footprint 2019," 2021. [Online]. Available: https://geoservice.dlr.de/data-assets/twg5xsnquw84.html

[36] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," in *The handbook of brain theory and neural networks*, M. Arbib, Ed. MIT Press, 1995, pp. 255–258.