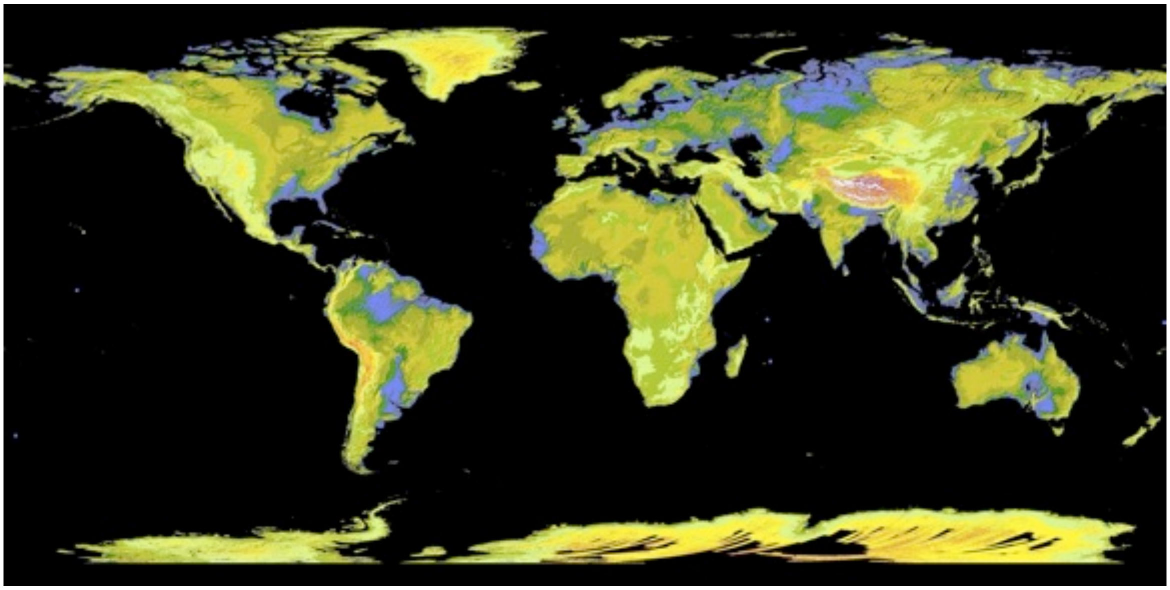
ASTER Global Digital Elevation Model Version 2 – Summary of Validation Results

August 31, 2011



by the ASTER GDEM Validation Team, with contributions by:

Tetsushi Tachikawa[1], Manabu Kaku[2], Akira Iwasaki[3]
Dean Gesch[4], Michael Oimoen[5], Zheng Zhang[6], Jeffrey Danielson[4]
Tabatha Krieger[7], Bill Curtis[7], Jeff Haase[7]
Michael Abrams[8], Robert Crippen[8]
Claudia Carabajal[9]

Compiled by Dave Meyer[4] on behalf of the NASA Land Processes Distributed Active Archive Center and the Joint Japan-US ASTER Science Team

1 Earth Remote Sensing Data Analysis Center (ERSDAC)
2 Mitsubishi Material Techno Corporation, under contract to ERSDAC
3 University of Tokyo
4 US Geological Survey, Earth Resource Observation and Science Center
5 Stinger, Ghaffarian Technologies, Inc. under contract to USGS/EROS
6 Earth Resources Technology, Inc , under contract to USGS/EROS
7 US National Geospatial Intelligence Agency/SNAT
8 Jet Propulsion Laboratory, California Institute of Technology
9 Sigma Space Corp. under contract to NASA/GSFC

Table of Contents

Executive Summary

On June 29, 2009, NASA and the Ministry of Economy, Trade and Industry (METI) of Japan released a Global Digital Elevation Model (GDEM) to users worldwide at no charge as a contribution to the Global Earth Observing System of Systems (GEOSS). This “version 1” ASTER GDEM (GDEM1) was compiled from over 1.2 million scene-based DEMs covering land surfaces between 83°N and 83°S latitudes. A joint US-Japan validation team assessed the accuracy of the GDEM1, augmented by a team of 20 cooperators. The GDEM1 was found to have an overall accuracy of around 20 meters at the 95% confidence level. The team also noted several artifacts associated with poor stereo coverage at high latitudes, cloud contamination, water masking issues and the stacking process used to produce the GDEM1 from individual scene-based DEMs (ASTER GDEM Validation Team, 2009). Two independent horizontal resolution studies estimated the effective spatial resolution of the GDEM1 to be on the order of 120 meters.

A second version of the ASTER GDEM (GDEM2) is scheduled for release by NASA and METI in mid-October, 2011. Improvements in the GDEM2 result from acquiring 260,000 additional scenes to improve coverage, a smaller correlation kernel to yield higher spatial resolution, and improved water masking. As with the GDEM1, the GDEM2 validation was performed by the U.S. and Japanese partners. Vertical accuracy assessments included a comparison of the GDEM2 against absolute geodetic references over the Conterminous US (CONUS), against national elevation grids over the US and Japan, against the Shuttle Radar Topography Mission (SRTM) 1 arc-second elevation grids over the US and 20 sites around the globe, and against space borne laser altimeter data globally. Horizontal accuracy assessments were conducted as part of the Japan and the global SRTM studies, and horizontal resolution studies were conducted in both Japan and the US. Each group documented changes in artifacts in GDEM2 due to processing improvements.

The absolute vertical accuracy study found the GDEM2 to be within -0.20 meters on average when compared against 18,000 geodetic control points over the CONUS, with an accuracy of 17 meters at the 95% confidence level. The Japan study noted the GDEM2 differed from the 10-meter national elevation grid by -0.7 meters over bare areas, and by 7.4 meters over forested areas. Similarly, the CONUS study noted the GDEM2 to be about 8 meters above the 1 arc-second NED over most forested areas, and more than a meter below NED over bare areas. The global altimeter study found the GDEM2 to be on average within 3 meters of altimeter-derived control, and also documented sensitivity to tree canopy height. The Japan study noted that the horizontal displacement in GDEM1 of 0.95 pixels was reduced to 0.23 pixels in GDEM2. Both teams noted improvements in horizontal resolution, between 71 and 82 meters, comparable to the SRTM 1 arc second elevation model, but at the cost of some increased noise. The number of voids and artifacts noted in GDEM1 were substantially reduced in GDEM2, and in some areas virtually eliminated.

Based on these findings, the GDEM validation team recommends the release of the GDEM2 to the public, acknowledging that, while vastly improved, some artifacts still exist which could affect its utility in certain applications.

Background

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) on NASA's Terra spacecraft collects in-track stereo using nadir- and aft looking near infrared cameras. Since 2000, these stereo pairs have been used to produce single-scene (60 x 60 km) digital elevation models having vertical (root-mean-squared-error) accuracies generally between 10 m and 25 m. On June 29, 2009, NASA and the Ministry of Economy, Trade and Industry (METI) of Japan released a Global Digital Elevation Model (GDEM) to users worldwide at no charge as a contribution to the Global Earth Observing System of Systems (GEOSS). This "version 1" ASTER GDEM (GDEM1) was compiled from over 1.2 million scene-based DEMs covering land surfaces between 83°N and 83°S latitudes. GDEM1 is a 1 arc-second elevation grid distributed as 1°-by-1° tiles.

A joint US-Japan validation team assessed the accuracy of the GDEM1, augmented by a team of 20 cooperators selected through an Announcement of Opportunity (AO). In summary, the GDEM1 was found to have an overall accuracy of around 20 meters at the 95% confidence level. The team also noted several artifacts associated with poor stereo coverage at high latitudes, cloud contamination, water masking issues and the stacking process used to produce the GDEM1 from individual scene-based DEMs (ASTER GDEM Validation Team, 2009). Two independent horizontal resolution studies estimated the effective spatial resolution of the GDEM1 to be on the order of 120 meters (Crippen, 2009; Tachikawa et al. 2009).

NASA and METI are scheduled to release a second version of the ASTER GDEM (GDEM2) in mid-October, 2011; this report extends the validation methods used for GDEM1 to the new GDEM. The GDEM2 has the same gridding and tile structure as GDEM1, but benefits from the inclusion of 260,000 additional scenes to improve coverage, a smaller correlation kernel (5x5 versus 9x9 for GDEM1) yielding higher spatial resolution, and improved water masking. Also, a negative 5 meter overall bias observed in the GDEM1 was removed in newer version. As with the GDEM1, the GDEM2 validation was the joint responsibility of U.S. and Japanese partners. The U.S. validation team included the U.S. Geological Survey (USGS, in cooperation with NASA), NASA's Jet Propulsion Laboratory (JPL), the National Geospatial-Intelligence Agency (NGA) and the NASA Goddard Space Flight Center (GSFC). The Japanese validation was conducted by the Earth Remote Sensing Data Analysis Center (ERSDAC) in cooperation with the University of Tokyo and Mitsubishi Materials Techno Corporation (under contract to ERSDAC). As before, the GDEM2 will be distributed at no charge to users through ERSDAC on behalf of METI, and at the Land Processes Distributed Active Archive Center (LP DAAC), located at the USGS Earth Resource Observation and Science Center (EROS), on behalf of NASA.

Methods & Reference Elevation Data Sets

Japan/ERSDAC. The Japanese validation team's methods for evaluating the GDEM2 is documented in detail by Tachikawa et al. (2011b), but is briefly summarized here. The primary reference used for the Japan study is the 10-m mesh DEM produced by the Geographical Survey Institute (GSI) of Japan. The study focused on 4 GDEM2 tiles in central Honshu Island (figure 1), spanning elevations from sea level to peaks exceeding 3000 meters. The impact of land cover on GDEM2 elevation errors was determined by stratifying the GDEM2 against the GSI's "Subdivision Land Use Data of Digital National Land Information", a 100-m land cover grid derived from satellite, aerial photography and field measurements. This land cover dataset was most recently updated in 2007. The Japan assessment included horizontal and vertical accuracy assessment against the GSI DEM, a horizontal resolution estimate against the GSI DEM decimated to variable resolutions, and an assessment of artifacts.

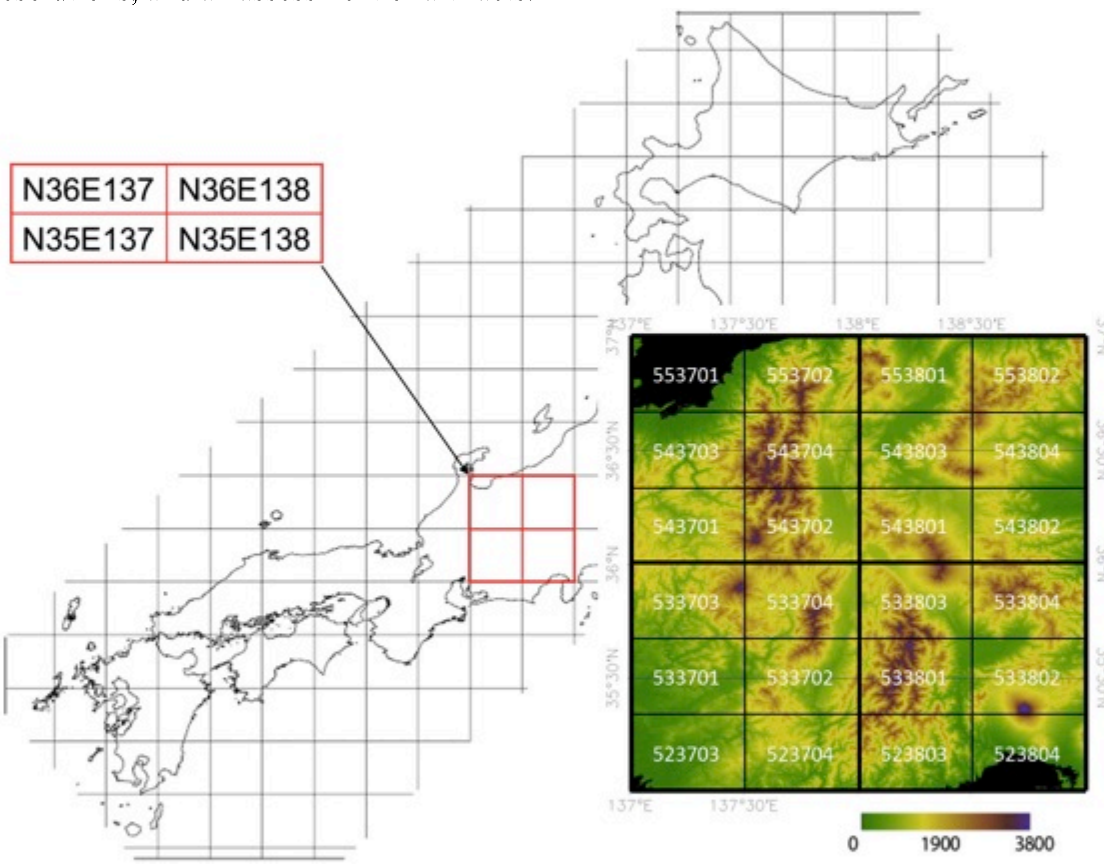


Figure 1 - Japan study site.

CONUS/USGS. Similar to the GDEM1 validation, the GDEM2 was evaluated over the CONUS against the "GPS on Bench Marks" data set (Gesch et al., 2011) as an absolute geodetic reference (figure 2). GDEM2 postings were bi-linearly interpolated over the benchmarks for the comparison. The GDEM2 was also evaluated against other 1 arc-second elevation grids over the CONUS: the National Elevation Dataset (NED), the Shuttle Radar Topography Mission (SRTM) DEM, and the original GDEM1. The NED data set has a published vertical accuracy of 2-3 meter root-mean-squared error (RMSE), and thus was used as the primary reference for pixel-by-pixel grid comparisons. As with the Japan study, the CONUS assessments were segmented by land cover classes to look for relationships between accuracy and cover type. The land cover data set used is the 2006 update of the National Land Cover Database (NLCD), which includes 19 classes derived from 30-meter Landsat data. Like the GDEM1 validation study, all 934 1x1 degree GDEM2 tiles covering the CONUS were included in this validation effort. The US effort focused on vertical accuracies and land cover effects.

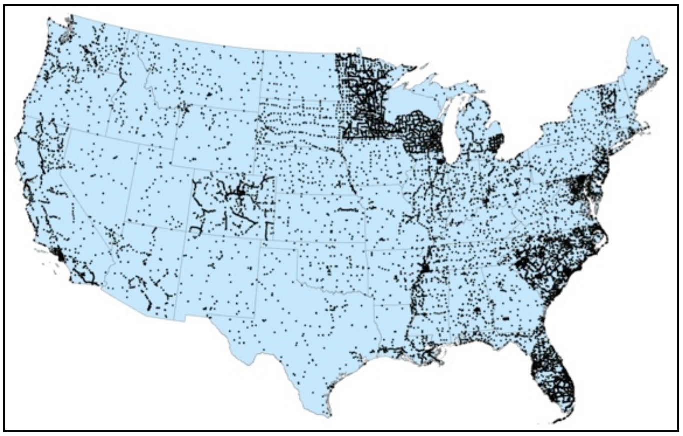


Figure 2 - The 18,207 GPS benchmarks used for absolute accuracy determination over CONUS.

Global: NGA. The NGA reproduced much of the work done for GDEM1, using the same 284 GDEM tiles as before, located at 20 geographic areas globally (Krieger et al., 2011). The results from the current GDEM2 validation are based on either a comparison with global 1 arc-second SRTM ("DTED level 2", or "DTED2"), or with the GDEM1. The NGA also did an extensive visual identification of artifacts in the GDEM2.

Global/ICESat. The NASA Planetary Geodynamics group at the Goddard Space Flight Center (GSFC) evaluated the GDEM2 against data collected by the Geoscience Laser Altimeter System (GLAS) on board the Ice, Cloud and land Elevation satellite (ICESat) (Zwally et al., 2002, Schultz et al., 2005). The results were stratified against Globcover land cover data derived from the Medium Resolution Imaging Spectrometer (MERIS) on the European Space Agency's Environmental Satellite (ENVISAT), to assess the correlation of errors with land cover. Results were stratified against the Vegetation Continuous Field (VCF) product derived from NASA Moderate Resolution Imaging Spectro-radiometer (MODIS) (Hansen et al., 2003; Hansen et al., 2006) to determine the effect of canopy height and density on the comparisons. These studies were conducted for North America, South America, Eurasia, Western Europe, Africa, Australia, New Zealand and Greenland (Carabajal, 2011).

Resolution/JPL. The JPL team focused largely on horizontal resolution determination and artifact identification. This study was based on comparisons to very high resolution (2 and ~3 meter) DEMs derived from LIDAR and non-LIDAR sources.

Each of the groups evaluated the GDEM for anomalies and artifacts as identified in the GDEM1. Many of the artifacts were identified and characterized through visual inspection, although the ERSDAC, NGA and GSFC groups also provided quantitative evaluations of the extent and impact of both voids and fills on GDEM quality.

This document presents the results of the joint validation team's efforts, with reference to the results of the Japanese team (Tachikawa et al., 2011b), the USGS study (Gesch et al., 2011), the NGA evaluation (Krieger et al., 2011), and the ICESat study (Carabajal, 2011). These are included as Appendices in this document.

Results

The joint ASTER GDEM2 validation team addressed different elements of this study. The vertical accuracy assessments were conducted regionally by the Japan and USGS teams, and globally by the NGA and NASA/GSFC. The Japan and NGA teams conducted horizontal accuracy assessments, while the Japan and JPL teams estimated the horizontal resolution of the GDEM2.

Vertical assessment using geodetic references

The approached used by the USGS team to estimate vertical errors from GPS benchmarks is described in detail in Gesch et al. (2011). Point-to-pixel differences were computed through bilinear interpolation of the GDEM2 postings at the precise latitude/longitude location of the GPS benchmark. Positive differences represent locations where the interpolated GDEM2 elevation exceeded the GPS point elevation, and, conversely, negative errors occur at locations where the GDEM2 elevation was below the GPS elevation. No horizontal error determinations were made, thus any horizontal errors would be subsumed within the vertical error assessment.

Figure 3 plots all of the errors against GPS benchmark elevations to determine if any correlations exist between error and elevation – none are apparent, and the errors are uniformly distributed about the zero error axis (e.g., the errors appear to be unbiased). Table 1 summarizes the results. The RMSE for the GDEM2 is 8.68 meters (compared to 9.34 meters for v1), and the absolute vertical accuracy, expressed as a linear error at the 95% confidence level (LE95), is 17.01 meters (compared to 18.31 meters for v1). The mean error provides an estimate of overall biases between the GDEM2 and the reference GPS benchmarks; the mean error for v2 is -0.20 (compared to -3.69 for v1). Each of these metrics was also computed for 1 arc-second NED and SRTM elevations, using the same methodology.

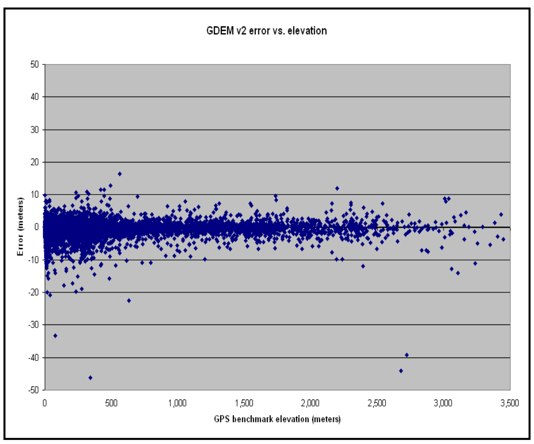


Figure 3 – Absolute GDEM2 error as a function of elevation.

DEM	Minimum	Maximum	Mean	Standard Deviation	RMSE	LE95
GDEM2	-137.37	64.80	-0.20	8.68	8.68	17.01
NED	-46.21	16.42	-0.33	1.81	1.84	3.61
SRTM	-28.67	28.58	0.73	3.95	4.01	7.86
GDEM1	-127.74	105.41	-3.69	8.58	9.34	18.31

Table 1 – Results from the CONUS absolute vertical accuracy assessment (in meters).

Land cover analysis on geodetic results.

The GDEMs (both versions) are maps of Earth's surface derived from ASTER, an optical stereo instrument that observes the landscape, including its land cover. Consequently, GDEM generally maps the tops of dense land covers such as forests and urban areas (buildings). This is in contrast to NED, which is a "bare Earth" model by design, and in contrast to GPS benchmarks, which always represent the bare Earth surface. (Note that SRTM, a radar sensor, also generally mapped the upper surface of land cover, although - unlike ASTER - it did so exclusively in February during leaf-off conditions for northern deciduous forests.) Relative to the bare Earth, land cover effects can be considered errors, and Figure 4 sorts the GDEM2 mean errors by their magnitudes, stratified by 14 NLCD classes. Positive mean errors are (for the most part) associated with "tall" land cover classes such woody wetlands, various forest types, and developed urban areas. The negative errors are largely associated with near-ground cover types that should approximate bare-earth and thus exhibit little or no bias. However, these cover types exhibit a negative bias, as noted on the right side of the plot – the average error associated with "open" land cover uniformly have negative values. Thus, the overall bias of -0.20 meters in table 1 above is interpreted as an overall negative bias of the GDEM2, averaged with the positive errors associated with "taller" cover types.

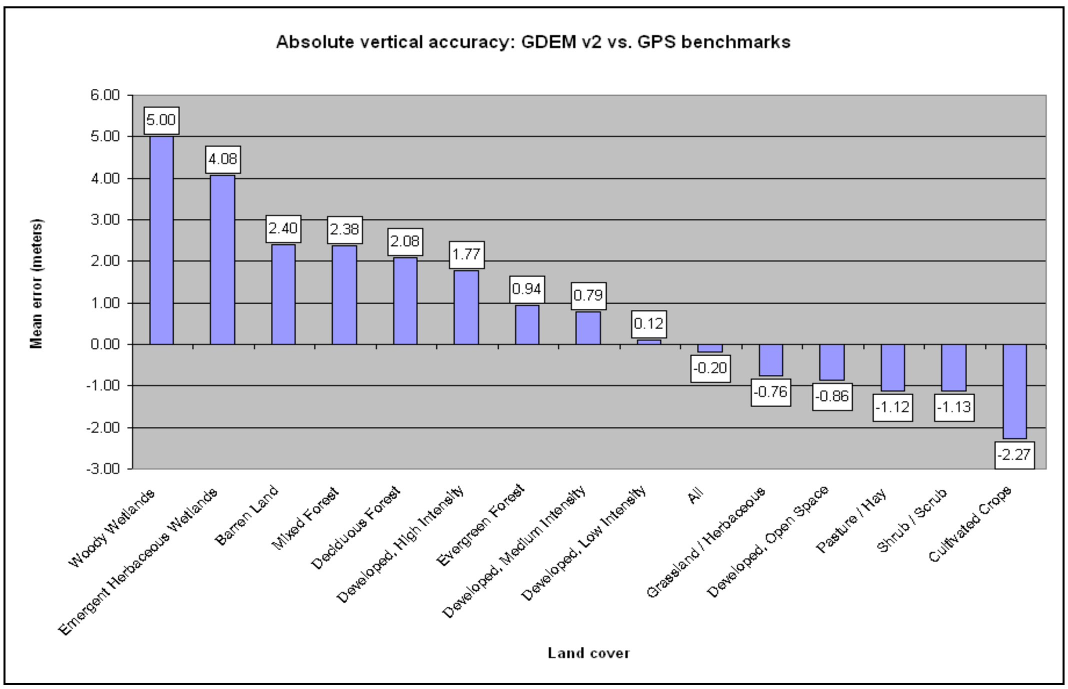


Figure 4 - Mean errors between the GDEM2 and the GPS benchmarks by land cover class. The plot is sorted by the magntidue of the mean error, largest (most positive) errors on the left to smallest (most negative) errors on the right.

Scene number analysis.

Figure 5 illustrates the relationship between the number of scenes (stereo pairs, or "NUM") in the "stack" used to estimate a given GDEM2 elevation value, with mean and RMS errors. Larger errors are associated with fewer than ten scenes, and especially fewer than three scenes. Also, mean errors are positive (too high an elevation) for fewer scenes, probably indicating cloud contamination to be the cause of the elevation error and possibly cloud persistence as the cause of the lack of usable scenes. Clearly, an increase in the number of scenes (higher NUM count) reduces error significantly between 1 and 10 scenes, but there is little improvement after about 15 scenes. Thus, it would appear a NUM count of between 10-15 could serve as a form of quality control filter when using the GDEM2 for applications requiring stable elevation uncertainties.

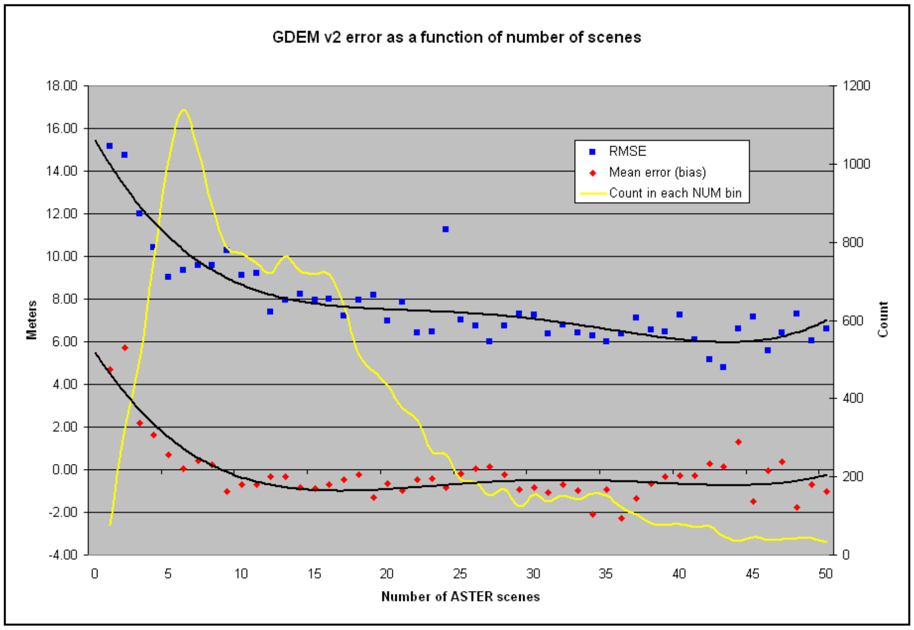


Figure 5 - The relationship between mean and RMS error, and the "NUM" stacking number.

Horizontal and vertical error assessment using reference elevation grids

The GDEM2 is evaluated against DEMs used for both national (Japan, US) and global applications. As described above, the GSI DEM is the reference used for the Japan study and the NED and SRTM DTED2 were used in the US study. Globally, the NGA used the SRTM DTED2. Note that, unlike the GDEM1 validation, the 3 arc-second SRTM DEM was not used in the CONUS study. These studies taken together address horizontal and vertical errors between both versions of the GDEM with the aforementioned reference DEMs, and estimate the horizontal resolution of the GDEM2.

Horizontal error.

As discussed in Tachikawa et al. (2011b), the validation over Japan was based on the 10 meter GSI DEM. Preprocessing of the GSI DEM ensured comparability to the GDEM2 in terms of format, datum and grid sampling. Both the GDEM2 and GSI DEM were resampled to 0.04 arc-second cells for comparison. Horizontal errors were estimated by shifting the GDEM2 by integer sample increments, in both x and y directions, over the GSI DEM, then computing the standard deviation of the differences at each shifted location. The shift associated with the minimum standard deviation was taken to be the mis-registration expressed to the nearest sample interval; sub-sample shifts were then calculated through interpolation. The average error of the 24 areas used for this analysis yielded an average East/West shift of -0.13 arc-seconds (positive is East), and an average North/South shift was estimated to be -0.19 arc-seconds (positive is North). Figure 6 depicts the computed shifts for each of the 24 areas for both the v1 and v2 GDEMs – an overall improvement in the GDEM2 is noted.

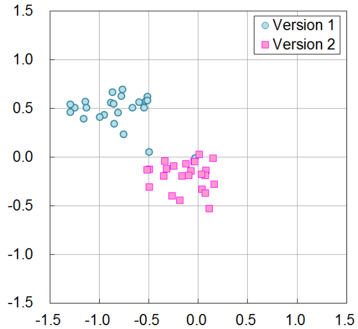


Figure 6 – Horizontal shift errors for the Japan study (north is up).

As a point of comparison to the Japan horizontal error estimates, the NGA study compared both the GDEM1 and GDEM2 to the global SRTM DTED2 at each of the aforementioned 20 global sites. Referring to table 2 below, the average East/West shift (relative to the SRTM) was estimated to be 0.104 arc-seconds for GDEM2 (compared to 0.0759 for GDEM1), and the average North/South shift was 0.175 arc-seconds for GDEM2 (compared to 0.187 for GDEM1). The horizontal shifts between the GDEM1 and GDEM2 were reported as sub-pixels. The differences between the NGA and Japan results are possibly due to the coarser SRTM used in the NGA study (as compared to the 10 m GSI DEM), yielding a less sensitive estimate.

	(in arc-seconds)	E-W Shift	N-S Shift	Horizontal shift magnitude
GDEM1	AVERAGE	0.0759	0.187	0.616
	MINIMUM	0	0	0
	MAXIMUM	1.333	0.777	2.236
GDEM2	AVERAGE	0.104	-0.175	0.601
	MINIMUM	0	0	0
	MAXIMUM	1.0877	1.00	2.37

Table 2 – Horizontal accuracy estimates from the global SRTM DTED2 comparison.

Vertical error.

The average elevation error for the Japan study is -7.4 meters overall, ranging from -5.58 to +15.45 meters over the 24 study areas (Tachikawa et al. 2011b), using the GSI DEM as a reference. Similar to the CONUS study, these results were stratified against the GSI land cover map described above (figure 7). The impact of “tall” cover types on the GDEM2 is clear - elevations within the forest class exhibit a 8.68 meter mean error, compared to 3.10 meters for the USGS study, and -1.32 and -1.09 meters for rice & farm (“flat”) classes, compared to the

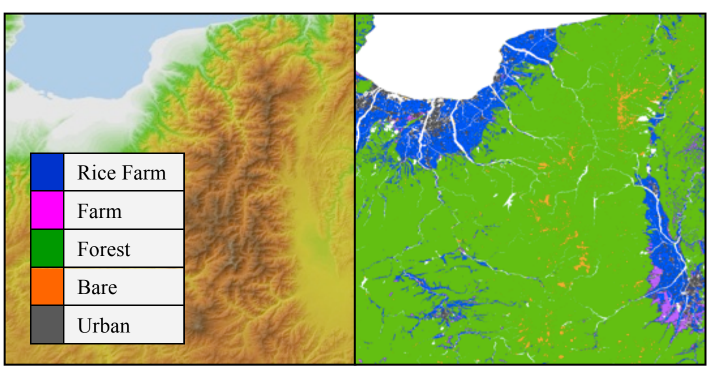


Figure 7 - The north west GDEM2 tile of the 4-tile Japan study site (left), the GSI Land cover used in the study (right). See figure 1 to orient to the entire study area.

USGS estimate of -0.96 for "open" classes (see figure 8 below and figure 4 in Gesch et al. (2011). The difference in magnitude for forest class errors between the US and Japan assessments could be due to differences in the definition of "forest" in each of the classification schemes, differences in the characteristics of the forests themselves, and/or differences in reference elevations and/or methods used for comparison (interpolation to GPS points versus comparisons to the 10 meter GSI grid).

All grids used in the CONUS study have 1 arc-second resolution and geographic coordinates, thus the NED, SRTM DTED2 and both versions of the GDEM could be compared on a pixel-by-pixel basis. Like the comparison to CONUS GPS benchmarks, positive differences indicate the GDEM2 exceeds the comparison DEM, negative differences indicate lower GDEM2 elevations. For this evaluation, the NED elevations were converted from the NAVD88 vertical datum to the EGM96 geoid vertical reference frame (the SRTM DTED2 and the GDEM2 are both natively referenced to EGM96). As with the absolute accuracy assessment, the differences between the GDEM2 and both NED and SRTM DTED2 were segmented by land cover class.

Figure 8 plots the mean errors between the GDEM2 and both NED and SRTM DTED2, arranged from maximum GDEM2 - NED difference (left) to minimum (right). It is clear that the remote sensing derived DEMs (GDEM2, SRTM DTED2) are not bare earth models like the NED, in that the GDEM2 - SRTM DTED2 differences are relatively small for the "taller" classes, where the GDEM2 is substantially higher than the NED

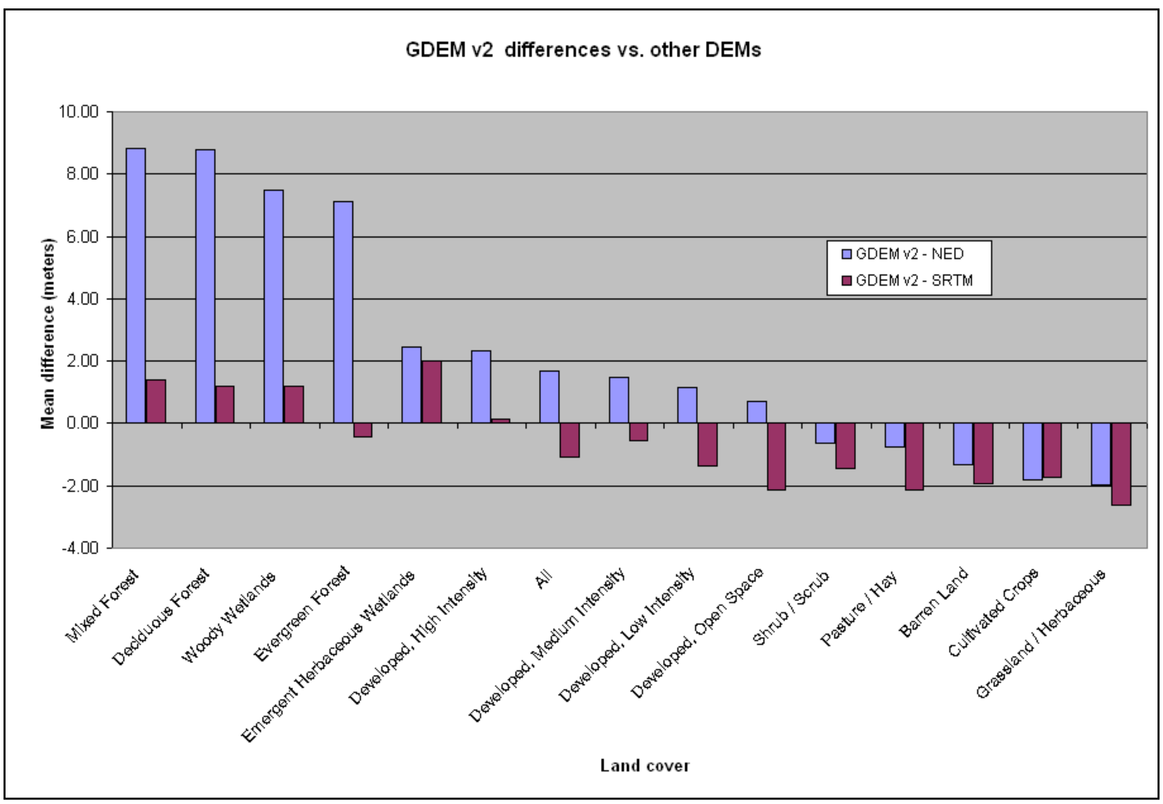


Figure 8 - Intercomparison of the GDEM2 to the SRTM DTED2 and NED DEMs over CONUS.

(left side of the figure). On the right side of the figure (corresponding to more “open” classes), the GDEM2 is somewhat lower than both the NED & the SRTM, reinforcing the previously discussed observation that a slightly negative bias may exist in the GDEM2.

The global SRTM DTED2 study conducted by the NGA is summarized in table 3 below. Overall, the GDEM2 more closely matches the SRTM elevations, where previously the negative 5-meter bias was observed in the older GDEM1. The horizontal errors were discussed in the previous section. Figure 9 maps the average differences between the GDEM2 and SRTM DTED2 at each of the 20 global sites. For the comparison of the old and new versions of the GDEM, it was found that aforementioned bias in the older version had been removed. Refer to Krieger et al. (2011) for a more detailed discussion.

	In meters	Difference	STD	Median	Max	Min
GDEM1	**AVERAGE**	-6.087	8.673	-14.365	180.765	-215.77
	MIN	-15.488	1.282	-140	14	-900
	MAX	2.535	22.171	2	2333	-32
GDEM2	**AVERAGE**	-1.572	8.826	-9.004	127.465	-161.05
	MIN	-10.620	1.065	-132	23	-763
	MAX	9.036	20.204	6	425	-32

Table 3 - Vertical errors from the global SRTM DTED2 study.

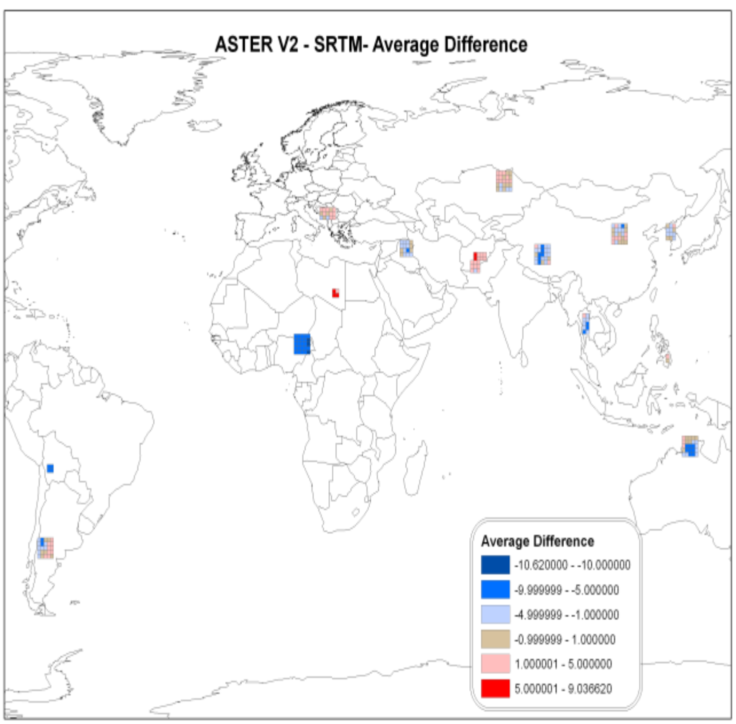


Figure 9 – Test sites used in the global SRTM DTED2 study.

Horizontal resolution.

While random noise and structured artifacts are features that limit the utility of a digital elevation model (DEM), an additional and largely independent feature is its inherent horizontal resolution. The concept of horizontal resolution in DEMs is not well defined, and it is often mistakenly stated as the spacing of postings (which is analogous to pixel spacing or pixel size in images). But just as a blurry image misses spatial detail, a DEM can appear blurry and its utility can be limited when its level of detail does not match the spacing of its postings.

GDEM2, like GDEM1 and full-resolution SRTM, has 1-arc-second (approximately 30m) postings, but none of them have 30m horizontal resolution. Previous studies found that horizontal resolution for SRTM exceeds 60 meters (Smith and Sandwell, 2003; Guth, 2006). Later, the Japanese and US ASTER Science Teams independently found the horizontal resolution of GDEM1 to be on the order of 120 meters (Crippen, 2009; Tachikawa et al., 2009). This was determined by statistical comparisons to much-higher resolution DEMs for which the horizontal detail was degraded until it best matched that of GDEM1.

For the current Japan study, the GDEM2 was compared against the GSI 10-m DEM, decimated to coarser grids from 1 to 9 arc-seconds in 1-arc-second intervals (9 different resolutions). Figure 10 shows some of the decimated DEMs used in the evaluation. Similar to the horizontal error study, the standard deviations of the elevation differences

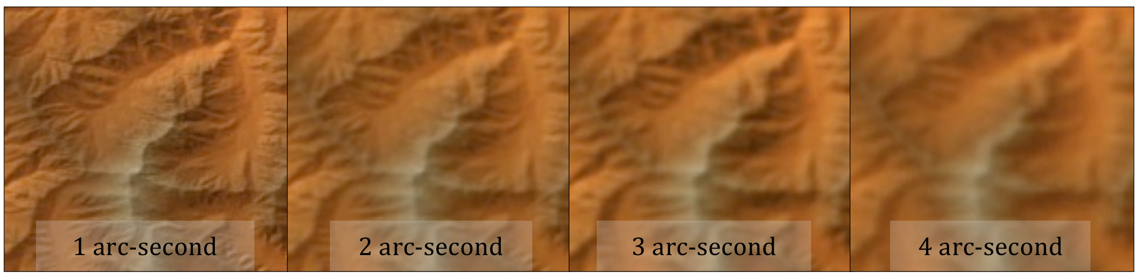


Figure 10 - Subsets from some of the decimated GSI DEMs used in the horizontal resolution estimation over Japan study areas.

between the GDEM2 and each of the nine decimated GSI DEMs were computed, and the low point of a parabolic curve fit to the points was taken as the horizontal resolution (Figure 11). The 2.4-arc-second resolution depicted for the GDEM2 corresponds to about 72 meters. This compares to 3.8 arc-seconds for GDEM1, or around 114 meters. (In this report, arc-seconds are converted to meters, using 30 meters/arc-second, regardless of latitude.)

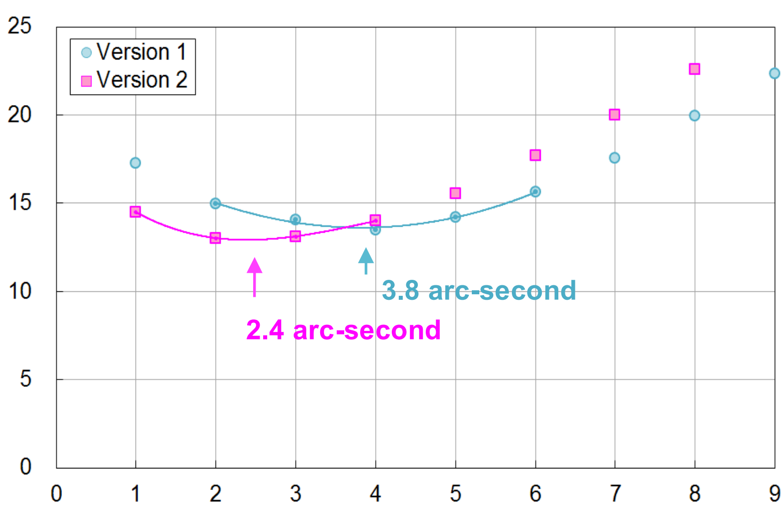


Figure 11 – Horizontal resolution study from the Japan study; vertical exist is standard devation of the difference in meters, horizontal axis is ground resolution of the Reference DEM in arc-seconds.

The US team similarly measured the horizontal resolutions of GDEM1 and GDEM2 and also SRTM. The results are presented in Table 4 along with the results from Japan. The resolution improvement from GDEM1 to GDEM2 is about as expected for a change of the stereo correlation kernel from 9x9 pixels (9x15m = 135m) to 5x5 pixels (5x15m = 75m). Resolution is limited by (but not to) the kernel size, and resolution can be improved by oversampling in analog-to-digital conversions. In this case, the analog terrain is observed in 135m-wide (GDEM1) or 75m-wide (GDEM2) kernels and is oversampled at 30m postings in the digital elevation model. Thus, the measured horizontal resolutions near or superior to the kernel size are as expected.

DEM	Japan Non-LIDAR	West Virginia Non-LIDAR	Utah LIDAR	California LIDAR	Average Non-LIDAR	Average LIDAR
GDEM-1	114	118	119	124	116	121
GDEM-2	72	70	81	83	71	82
SRTM 1-arc-sec	--	72	76	79	72	77
SRTM 3-arc-sec	--	97	101	103	97	102

Table 4 – Estimated horizontal resolutions for GDEM and SRTM DEMs, measured in arc-seconds but presented in meters, using 30m per arc-second. Japan: GSI 10m DEM within quad N35E137. West Virginia: Keyser-Romney area, 1/9 arcsec NED. Utah: Oquirrh Mountains, 2m State of Utah LIDAR. California: Northeast Ventura County 1/9 arcsec NED. (Japan: Tachikawa et al., 2009; WV, UT, and CA: Crippen, 2011, work in progress).

The US ASTER Science Team has also visually compared GDEM2 to full-resolution SRTM at numerous sites around the world and found their horizontal resolutions to be very similar in many cases. This can be evident in standard shaded relief or height-as-brightness displays (figure 12), but a simple difference image generated from the two DEMs can further highlight their relative horizontal resolution differences, if any. In sharp relief terrain, the poorer resolution DEM will have lower ridges and higher gullies (but similar slopes) compared to a higher resolution DEM, resulting in a distinct ridge and gully pattern in the difference image. This pattern was commonly found in comparisons of GDEM1 to full-resolution SRTM, but it is generally not found in similar comparisons using GDEM2. Thus we conclude that GDEM2 generally has a horizontal resolution that is substantially improved from GDEM1 and quite similar to full-resolution SRTM. As shown in Table 4, it is also about 20% superior to the publicly released SRTM 3-arc-second global DEM.

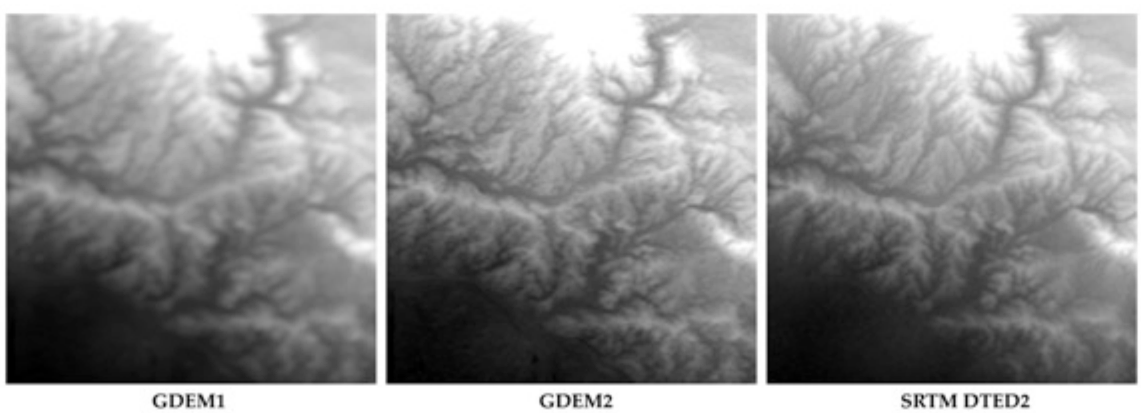


Figure 12 - Visual comparisons of horizontal resolution from three DEMs.

Vertical error assessment using ICESat altimetry

The GSFC study involved comparisons between the GDEM2 and the ICESat data described above, converted to WGS84/EGM96. The ICESat data resolves about 50 meters at the ground, with footprints spaced every 170 meters along the satellite track.

Differences were estimated between the ICESat elevations and the nearest-neighbor GDEM2 elevation to the ICESat footprint. These differences were compared to various elevations within the ICESat vertical profile: highest (H), centroid (C), and lowest (L) elevations, as well as ground (G) elevations where distinct, lower peaks are found under vegetation canopies, or in bare areas (Carabajal and Harding, 2005, Carabajal and Harding 2006). The GSFC group applied stringent editing criteria (Carabajal et al., 2011) and cloud screening to yield quality ground control points (GCPs) for the comparisons. The report includes an overall description of methods, and results for each of the regions described earlier, including error analyses with respect to relief, elevation, land cover, vegetation cover density, and the number of ASTER stereo pairs used to construct an elevation at a given GDEM2 grid cell.

Table 5 summarizes the ICESat results for each of the regions analyzed in the study. With the exception of Greenland, the mean differences for all regions are within +/- 3 meters. The standard deviations and RMSEs are consistent with other studies, all less than 12 meters; these results are discussed in more detail later. Greenland exhibits anomalously high values due to contamination associated with ice cover – special attention will be given shortly to these results.

Area	N	Mean (m)	Median (m)	STD (m)	RMSE (m)	Min. (m)	Max. (m)
Africa	14661568	1.6	0.325	11.61	11.72	-267.0	1802.16
S. America	2283947	2.17	1.84	8.51	8.78	-376.38	1242.94
N. America	5410981	-2.11	-1.96	11.73	11.92	-514.4	2761.32
Australia	4349145	-2.83	-2.97	7.08	7.62	-122.49	168.23
New Zealand	16836	0.08	-.028	8.89	8.89	-52.05	132.79
W. Europe	1714027	-2.77	-2.77	10.71	11.06	-339.45	2436.01
Eurasia	15264903	-1.60	-1.65	11.76	11.87	-496.43	2347.37
Greenland	4190411	235.70	109.04	535.00	584.62	-3606.7	4152.07

Table 5 – Summary of global ICESat results. Note that the sign of the differences are reversed from GSFC study for consistency with the other studies summarized here. The convention used above is GDEM2 – ICEsat centroid elevation.

Land cover analysis.

The ICESat /GDEM2 Globcover analysis results varied from region to region, with no strong correlations evident over the entire global dataset. Table 6 lists, the results over those areas classified as "bare" (Globcover class 200). Globally, the ICESat and GDEM2 elevations are within +/- 2 meters (New Zealand and Greenland have too few samples to draw meaningful conclusions).

Although both the ICESat GLAS & ASTER instruments are optical in nature, the former is an active system designed to penetrate plant canopies to a significant degree. Thus, one would expect deviations to occur between the ICESat and GDEM2 that increase with canopy height and density. As discussed earlier, the MODIS VCF data set was used to isolate these effects. The full ICESat report (Carabajal, 2011) includes complete assessments of the differences as a function of bare, herbaceous and tree cover percentages. Figure 13 plots the results for percent tree cover. The differences between ICESat and GDEM2 increase with tree cover for each region, indicating the sensitivity of

Region	N	Mean (m)	Median (m)	STD (m)	RMSE (m)	Min (m)	Max (m)
Africa	3601586	2.11	0.97	10.66	10.86	-198.99	361.90
Australia	243066	-1.64	-1.78	6.64	6.84	- 66.72	63.34
Eurasia	4049072	0.58	0.10	10.36	10.38	-389.23	590.38
N. America	7172	-1.96	-2.60	5.86	6.18	-31.93	89.41
S. America	157484	0.86	0.53	7.92	7.97	-155.08	141.41
N. Zealand	111	4.25	1.18	10.59	11.41	-21.03	35.77
W. Europe	107217	-1.56	-1.57	6.34	6.53	-83.00	573.56
Greenland	6	-0.93	-0.48	6.80	6.86	-9.91	7.41

Table 6 - ICESat results over areas identified as "bare" (Globcover class 200). Note once again the convention of GDEM2-ICESat is used for consistency with the other studies in this report.

the GDEM2 to top-of-canopy elevations as compared to the penetrating ICESat altimeter. Interestingly, those regions including temperate and boreal forests (North America, Eurasia, Western Europe) indicate an inversion of that relationship at complete tree canopy cover (e.g., the difference is somewhat less at 100% tree cover than at lower cover percentages).

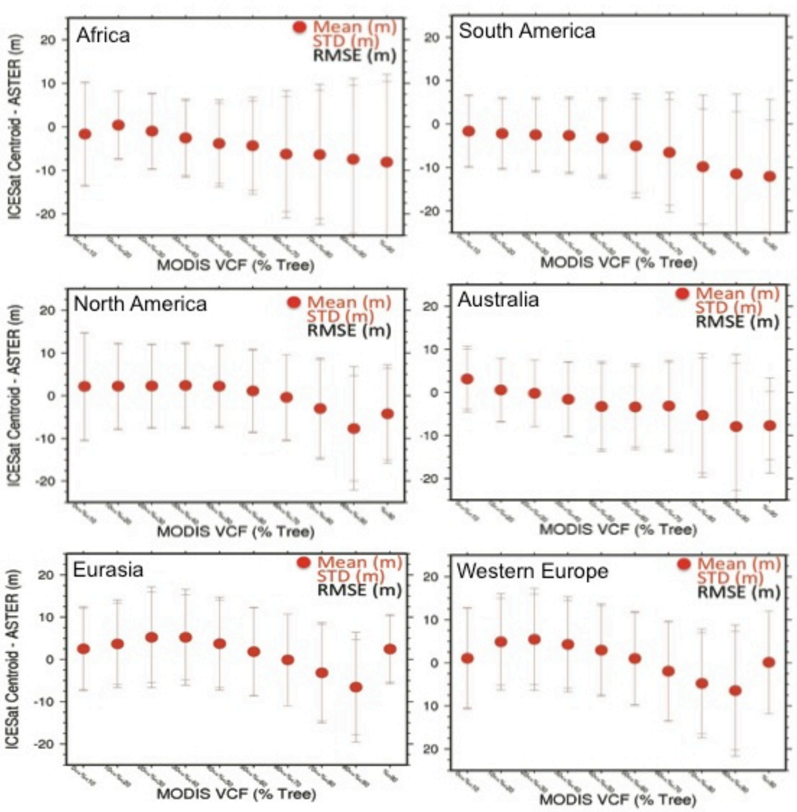


Figure 13 - GDEM2 differences as a function of VCF percent tree cover (increments in 10% cover from 0 to 100%).

Scene number/fill analysis.

As with the CONUS absolute accuracy study, the ICESat – GDEM2 differences were evaluated against the number of ASTER scene pairs used to construct a given GDEM2 elevation. Figure 14 shows this relationship for 6 of the areas. In most cases, the differences are stable between roughly 8 to 40 scenes, comparable to the CONUS results (see figure 5) The errors at higher scene pair numbers become more variable because these values occur very infrequently (note figure 5 from the CONUS study, which plots mean error against the number of scene pairs used). The differences were also evaluated against the “fill” values in the GDEM2 (e.g., where the value of the scene pair number is less than 0) – these indicate areas where other sources of information were used to replace poor quality or missing ASTER-derived values. Overwhelmingly, the replacement values had larger differences with ICESat elevations than those derived from the ASTER Instrument itself (see Carabajal, 2011).

Greenland analysis.

Greenland was isolated for special attention because it is (1) an area where ICESat and ASTER are among the few DEMs available, and (2) an area where cloud cover associated with perennial ice at high latitudes impedes the remote retrieval of elevations at near-infrared wavelengths. ASTER is at a particular disadvantage in that it is also requires adequate seasonal illumination for imaging. Also, the often-featureless terrain complicates scene-pair correlations required to measure parallax displacement (and thus terrain height) using stereo methods. Not surprisingly, large errors can occur where the GDEM2 hasn't been filtered to exclude areas free of perennial ice and/or lacking sufficient observations for correlation. However, as discussed in the land cover analysis above, for those few areas classified as bare (n = 6 samples, Globcover class 200), ICESat and GDEM2 compared quite well (see table 6). For those areas determined to be herbaceous (Globcover class 150, n = 2496), the ICESat – ASTER statistics show a mean difference of 1.60 meters, a standard deviation of 12.90 meters, and an RMSE of 13.00 meters. Also, for those areas having scene pair counts between 15 and 40, the difference is largely stable, with ICESat being about 5 meters higher than GDEM2; above 15 scene pairs the standard deviation falls below 12 meters. However, nearly 75% of the sampled data had below 5 scene pairs per GDEM2 pixel. Thus, although large errors exist in the GDEM2 over Greenland, useful elevations can be retrieved using higher scene pair counts and ice-free land cover as quality filters. It is assumed similar results could be expected over Antarctica, although this was beyond the scope of the current study.

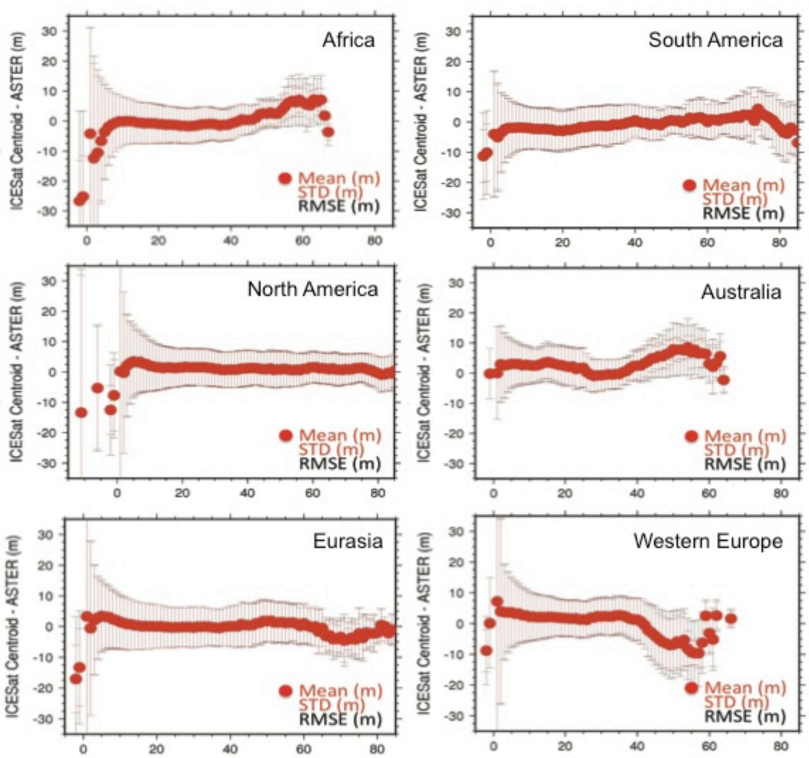


Figure 14 - ICESat - GDEM2 differences as a function the number of ASTER scene pairs ("NUM") used in generating elevations at a given pixel. Recall "NUM" < 0 corresponds to elevations filled from non-ASTER sources.

GDEM quality and artifacts

The initial validation study identified a number of artifacts in the GDEM1, due to residual cloud anomalies, elevation "steps" at scene boundaries, irregular stack number boundaries, and noise associated with inland water bodies. Overall, these artifacts can be attributed to cloud screening, inadequate number of observations, water masking, and residual mis-registration between the individual DEMs used to construct the GDEM. The initial GDEM1 distribution included specific statements regarding these "residual anomalies", accompanied by cautionary statements regarding its use for visual interpretation and quantitative application. The GDEM1 Validation Report includes an extensive discussion of these artifacts (ASTER GDEM Validation Team, 2009).

Due primarily to increasing the number of scenes used in the GDEM2 (260,000 additional scenes) and improved water masking, these artifacts have been substantially reduced in the new version, and in many places nearly eliminated. Particular attention was given to high-latitude anomalies associated with poor coverage – the large number of new acquisitions used in GDEM2 addressed this issue. The Japan study team discussed at length the improvement in coverage (reduction in "voids") (Appendix A), as did the

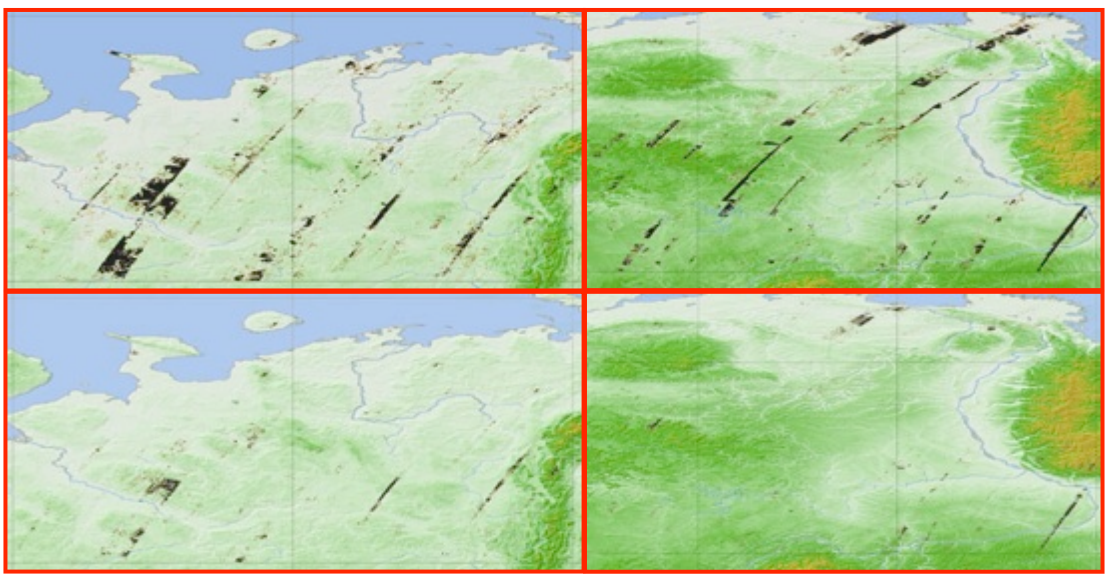


Figure 15- Removal of voids at high latitudes at two sites over northern Eurasia due to increased acquisitions, GDEM1 (top), GDEM2 (bottom).

NGA study (Appendix C). Figure 15 is an excerpt from the Japan study indicating the reduction of voids in northern Eurasia due to increased coverage. Comparing GDEM2 to GDEM1, the NGA team noted the number of artifacts ("spikes" and "wells") was reduced by an average of 7119 per tile over the 284 tiles evaluated. This study also noted an increase in the number of voids in some GDEM2 tiles, possibly due to a sampling bias from the 20 areas used in the evaluation, and improvement in quality control (e.g., detection and removal of artifacts, and replacing these artifacts by void values).

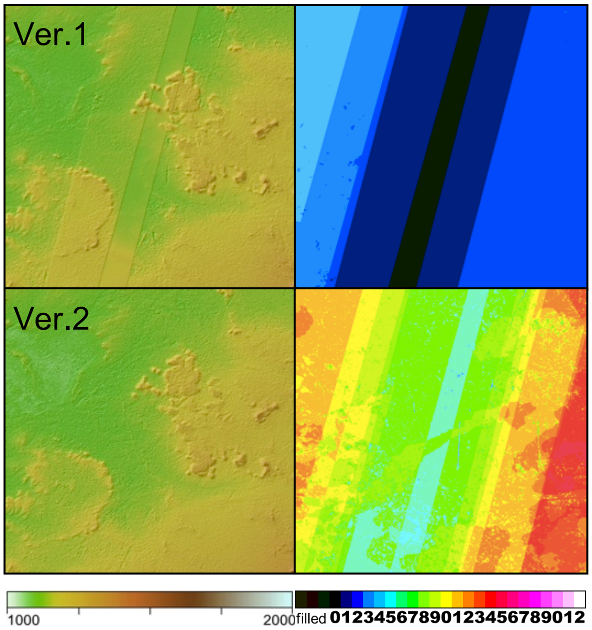


Figure 16 - Elevation and stacking in S31E023 (Southern Africa); GDEM1 above, GDEM2 below, elevation on the left, stacking number on the right.

Both the Japan and NGA teams provided extensive visual examples of the improvements in the GDEM2; a few examples are provided here for illustration. Figure 16 is taken from Tachikawa et al. (2011b), illustrating improvements in elevation quality derived from the increased number of observations. The "step boundaries" are largely gone from the GDEM2. Similarly, all of the validation groups noted a substantial reduction in the GDEM1 artifacts known as "pits", "bumps", and "mole runs" in nearly all areas observed. Clearly, there remain a smaller number of regions in the GDEM2 (compared to GDEM1) where acquisitions haven't been increased substantially, often due to persistent cloud cover, and the artifacts remain.

The NGA study focused considerable attention on the impacts of water masking, a problem for GDEM1. There were notable improvements from the new water mask. Figure 17 illustrates masking impacts over 3 areas. The top image is in Bolivia, showing typical improvements resulting from the new mask – both topology rendering (blue) and hydrology (red) are mapped more accurately in GDEM2 (left) than GDEM1 (right). The middle figure (Canada) shows improved resolution on land (blue), but introduced artifacts over water (red). The bottom figure (also Canada) shows improvements to both land and water features, although the GDEM2 exhibits high-frequency noise at waters edge.

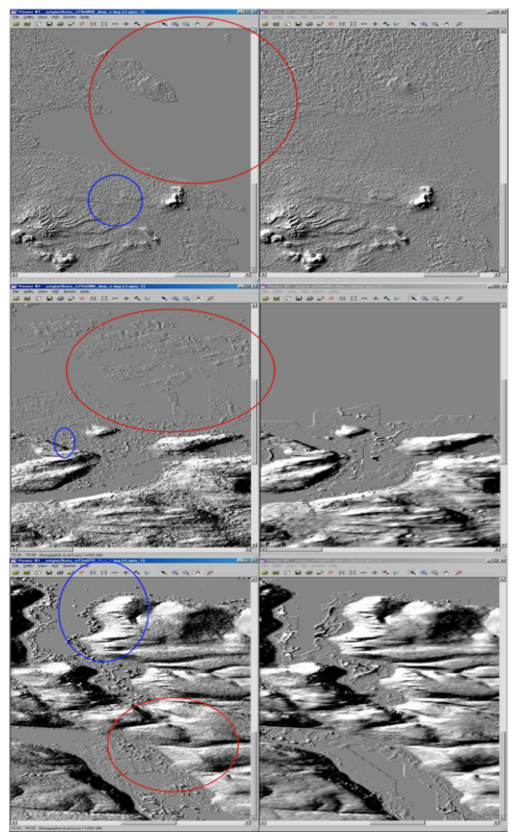


Figure 17 - Changes between GDEM1 (right) and GDEM2 (left) resulting from improvements in water masking. Top is Bolivia (tile s19w068), middle is Canada (tile n72w080), bottom is also Canada (tile n72w079).

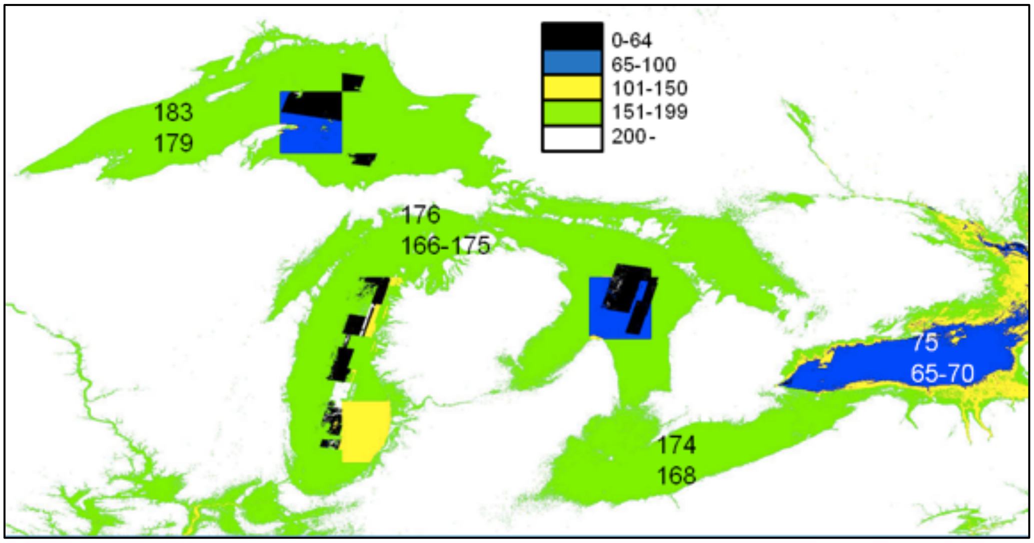


Figure 18 - Large lake anomalies over the North America Great Lakes.

The JPL team also evaluated the affect of the masking on inland water bodies. Examination of the GDEM2 revealed anomalous values in large lakes: the US Great Lakes and the Caspian Sea. Figure 18 illustrates the problem. The paired numbers are true elevations (top) and GDEM2 elevations for four of the Great Lakes. The agreement is quite close, with deviations of about 5 meters. However, within the three largest lakes (Michigan, Superior, Huron), interior tiles have quite errant values: Superior and Huron show blocks with values of 0-64, and 65-100 (black and blue colors); Michigan has tiles with values of 0-64, 101-150, and >200 (black, yellow, and white colors). Mostly where tiles include shoreline and land, the GDEM values are accurate. The same characteristics are seen in the Caspian Sea, with similar anomalies. No effort was made to post-process the GDEM and edit out the anomalies. Interestingly, the Black Sea does not exhibit this anomalous characteristic.

It is important to note that using a smaller stereo correlation kernel to enhance horizontal resolution (high frequency topographic "signal") can also increase high frequency noise. Thus, not unexpectedly, the improvement in horizontal resolution in GDEM2 as compared to GDEM1 came at the cost of added high-frequency noise. This added noise explains the higher standard deviation of GDEM differences from benchmark elevations (USGS; 8.58 to 8.68 m) and from SRTM (NGA; 8.673 to 8.826 m) despite the general reduction of glitches (such as pits, spikes, and clouds) and despite the improved GDEM horizontal resolution. The increased noise is visually apparent in shaded relief images (e.g., Figure 19, left column) and in images that isolate the noise by subtracting almost noise-free, bare Earth topography derived from a much higher resolution NED (1/9 arc second) DEM, degraded to GDEM and SRTM horizontal resolutions (Figure 19, right column). Land cover effects (forest versus open landscape) are evident in the noise images, most clearly by SRTM, which has a standard deviation versus the resolution-scaled NED of 3.26 m for this site. The increase in high-frequency noise here from GDEM1 to GDEM2 corresponds to standard deviations versus NED of 5.83 m and 7.20 m, respectively, after the NED was degraded to the horizontal resolution of each GDEM.

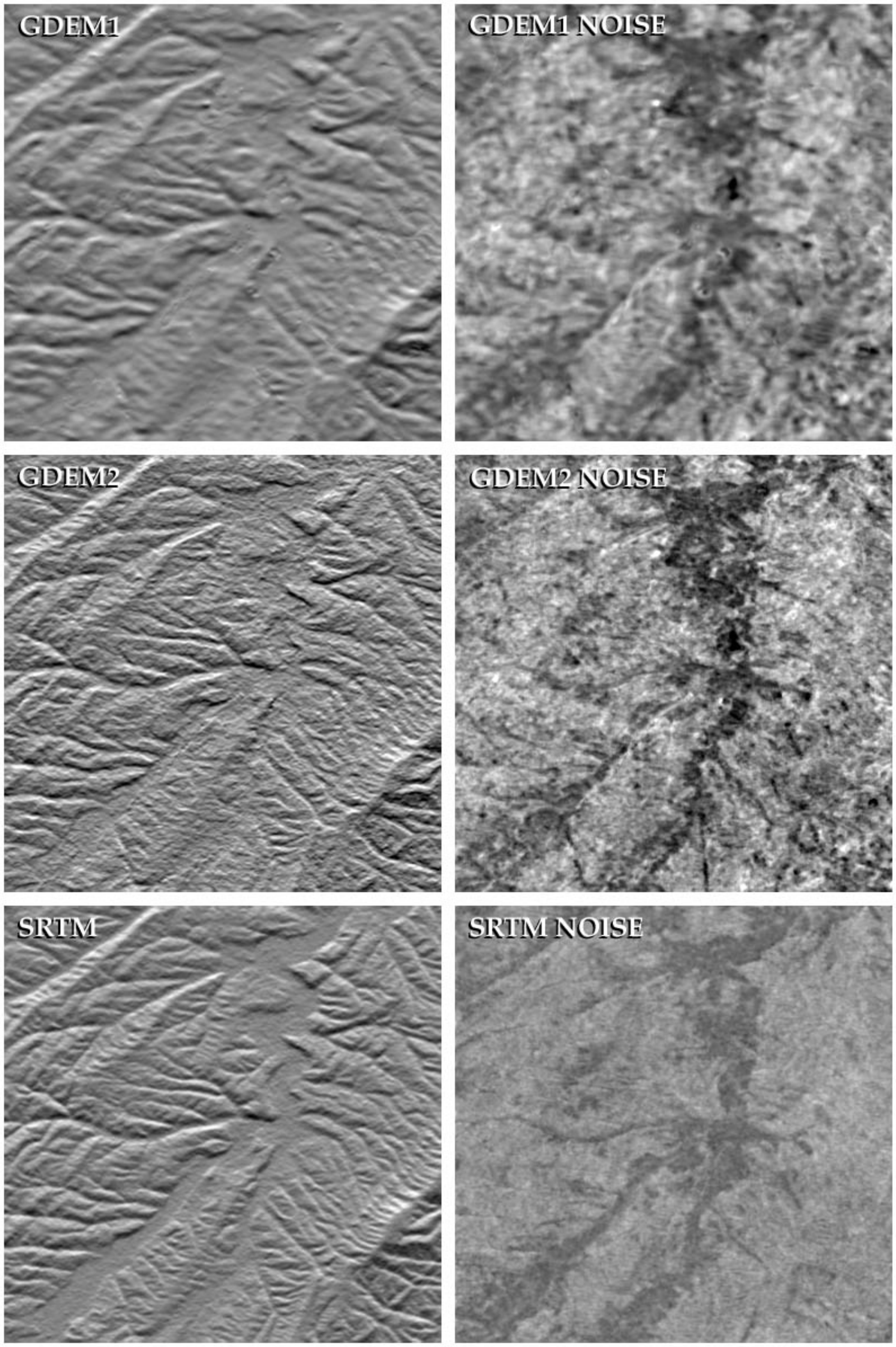


Figure 19 - Left: Shaded relief views of GDEM1, GDEM2, and SRTM 1-arc-second data, southeast of Keyser, West Virginia, 388x388 arc-seconds, equal contrast stretches. Note the similar horizontal resolutions for GDEM2 and SRTM, both superior to GDEM1. Right: Differences between each DEM and a high-resolution, bare Earth DEM (1/9 arc-sec NED) degraded to each corresponding DEM's horizontal resolution, with equal contrast stretches (not shaded).

Conclusions

In summary, changes in the number of acquired ASTER stereo pairs (1.5 million) and improvements in processing (water masking, smaller correlation kernel size, bias removal) have produced significant improvements in GDEM2 as compared to GDEM1. These improvements include increased horizontal and vertical accuracy, as compared to both GPS benchmarks and standard DEMs (GSI, NED, STRM DTED2), and improved horizontal accuracy and resolution (similar to the SRTM DTED2).

The Japan study is summarized in table 7:

<table>
<tr><th colspan="3"></th><th>Version 1</th><th>Version 2</th></tr>
<tr><td colspan="3">Horizontal Error</td><td>0.82 arc-sec. to west
0.47 arc-sec. to south</td><td>0.13 arc-sec. to west
0.19 arc-sec. to north</td></tr>
<tr><td rowspan="6">Elevation Error</td><td rowspan="3">Flat and open area (rice farm)</td><td>offset</td><td>-4.8 m</td><td>-0.7 m</td></tr>
<tr><td>SD</td><td>6.2 m</td><td>5.9 m</td></tr>
<tr><td>RMSE</td><td>-</td><td>6.1 m</td></tr>
<tr><td rowspan="3">Mountainous area largely covered by forest</td><td>offset</td><td>+2.2 m</td><td>+7.4 m</td></tr>
<tr><td>SD</td><td>15.4 m</td><td>12.7 m</td></tr>
<tr><td>RMSE</td><td>-</td><td>15.1 m</td></tr>
<tr><td colspan="3">Horizontal Resolution</td><td>3.8 arc-sec. (114m*)</td><td>2.4 arc-sec. (72m*)</td></tr>
</table>

Table 7– Validation results from the Japan study (one arc-second corresponds to 30 meters).

This study determined:

- The voids in northern area have decreased due to new ASTER acquisitions.
- The artifacts mostly disappear as a result.
- All lakes in the Japan study are perfectly flat by new water body detection algorithm (although inland water body problems exist elsewhere, see figure 18).

The US/CONUS validation raised several important observations about the quality of elevation measurements contained in GDEM2:

- There is an improvement in overall RMSE of nearly two-thirds of a meter (8.68 m vs. 9.34 m) when comparing the measured accuracies of GDEM2 and GDEM1. Likewise, there has also been an improvement in overall mean error (bias) in GDEM2 when compared with GDEM1 (-0.20 m vs. -3.69 m).

- It is clear that GDEM2 is influenced by above ground features (tree canopies and built structures), as observed in both the comparison of GDEM2 with GPS benchmarks, which represent ground level elevations, as well as in the GDEM2-NED differencing, with NED representing ground level elevations. This agrees with the Japan study that also noted a positive bias over forest cover types.

- In many forested areas, GDEM2 has elevations that are higher in the canopy than SRTM. This observation is based on both the comparison of GDEM2 with GPS benchmarks, as well as the GDEM2-SRTM differencing. Once again, this finding was reinforced in the Japan study, although the latter had a larger bias for tall cover types: 8.68 meters, compared to 3.10 meters for the CONUS study.

- An analysis of the number of ASTER individual scene DEMS that are stacked and averaged to derive the elevation value for every pixel in GDEM2 shows that improvements to mean error and RMSE are minimal beyond about 15 scenes.

- GDEM2 exhibits an apparent "true" negative elevation bias of about 1 meter, which was revealed through an analysis of mean error by land cover type. The overall mean error of -0.20 m is certainly an improvement over the mean error of -3.69 for GDEM1, but it somewhat masks the true performance of ASTER in measuring the elevation in open terrain conditions (non-vegetated, non-built-up). The overall mean error is dampened by the positive elevation biases contributed by forested and built-up land cover. While the true negative elevation bias of about 1 meter for GDEM2 is a significant improvement over the true negative elevation bias of about 5 meters for GDEM1, it is nonetheless a condition that users of GDEM2 data should be aware of and factor into decisions regarding application of the product. This is also reflected in the Japan study as well.

The horizontal resolution estimates were similar from the US and Japan teams when using non-LIDAR reference data sets, with GDEM2 estimated at 70m and 72m, respectively, and GDEM1 estimated at 118 and 114, respectively. The US team's additional estimates using LIDAR-derived high-resolution reference data averaged 82m for GDEM2 versus 121m for GDEM1. These higher estimates were the expected result of using a more precise and accurate reference DEM. The same LIDAR sites also produced average estimates of 77m for SRTM 1-arc-second data and 102m for SRTM 3-arc-second data. It is therefore concluded that (1) GDEM2 horizontal resolution is improved about 35% as compared to GDEM1, (2) GDEM2 horizontal resolution nearly matches that of SRTM 1-arc-second DEM, and (3) GDEM2 horizontal resolution is about 20% superior to the publicly available SRTM 3-arc-second global DEM.

Unfortunately, the addition of higher-frequency topographic signal in GDEM2 as compared to GDEM1 came at the cost of added, nearly ubiquitous, high frequency noise, as is visually apparent and as indicated by the higher standard deviation of differences from benchmark elevations (USGS) and from SRTM postings (NGA) despite the general reduction of artifacts such as pits and spikes. However, noisy signal is generally better than missing signal, and fine noise can be suppressed by filtering if critical, so even from a signal and noise tradeoff perspective we conclude that GDEM2 is more versatile than GDEM1.

The conclusions from the 20 global validation sites using SRTM DTED2 from the NGA group include:

- The GDEM2, in general, has slightly more void than GDEM1.

- Voids were mainly introduced in areas in Russia where large numbers of spikes/wells were removed.
  - Some voids in Australia have been successfully filled.
  - This result may differ from the Japan study in that it focused on higher latitudes where many artifacts (spikes/wells) were removed and replaced with voids. The Japan study clearly showed significant reductions in voids over large regions at high latitudes.
- GDEM2, in general, has fewer artifacts (spikes/wells) than GDEM1.
  - Large numbers of spikes/wells were removed north of 60 degrees of latitude.
  - Small numbers of spikes/wells were introduced in certain tiles due to the steepening of cliff lines (e.g., land/water transitions). Cliff lines in GDEM1 appear to match SRTM DTED2 better than in GDEM2 (possibly due to changes in water masking in GDEM2?).
- GDEM2 has been raised on average about 4 m above GDEM1.
- No significant horizontal shifts between GDEM1, GDEM2, and SRTM DTED2 were observed using statistical methods (however, vertical profile comparisons were not made). The variance with the Japan study is likely due to the coarse SRTM DEM used in the global study, compared to the 10 meter DEM used over Japan.
- Statistically, GDEM2 more closely matches SRTM DTED2 when comparing elevations post to post.
- The resolution / depiction of non-hydrology surfaces in GDEM2 has improved.
- Hydrology surfaces that fall “within” the SRTM water mask coverage footprint have improved.
- Hydrology surfaces that fall “outside” the SRTM water mask coverage footprint in general have improved, however shoreline noise, non-containment issues and artificial islands (ice?) are still present.
- Non-hydrology artifacts (pits, spikes, mole runs/steps) identified in GDEM1 generally are either diminished or removed in GDEM2.

Overall conclusion from the Global SRTM DTED2 study: While it is fairly clear from this high-level review that the quality of GDEM2 is superior to GDEM1, especially above 60 degrees north, NGA believes the data would still have to be assessed and edited on a case-by-case basis before use in specific applications.

The ICESat study concluded that globally (with the exception of Greenland), the GDEM2 elevations are on average within 3 meters of highly edited altimeter measurements, with standard deviations and RMSEs under 12 meters. For bare ground, the GDEM2 was on average within around 2 meters to the altimeter measurements (with the exception of New Zealand), having standard deviations and RSMEs under 12 meters. Although the GDEM2 exhibits large errors over much of Greenland, for those areas classified as either bare or herbaceous, the errors are on average within 2 meters of the ICESat elevations.

Based on the above, the joint ASTER validation team recommends the release of the GDEM version 2 to the public, noting substantial improvements in the quality of this product over the original GDEM. In many aspects (horizontal resolution, vertical accuracy) the GDEM2 is comparable to the SRTM DTED2, while extending to higher latitudes (83 degrees versus 60 degrees).

Acknowledgements

The LP DAAC is operated by the USGS on behalf of NASA under Interagency Agreement No. NNG09HP16I which funded the compilation of this report. The authors would like to thank Tom Sohre and Jason Stoker, both of the USGS, for their review and commentary on this document.

The development of the ICESat Geodetic Control Data Base is being supported by NASA's Earth Surface and Interior (ESI) Program, managed by Dr. John Labrecque, Contract NNH09CF42C: "Building an ICESat Geodetic Control Data Base for Global Topographic Mapping and Solid Earth Studies." Dr. David J. Harding is a Co-I in this proposal. Vijay Suchdeo helped with processing of the ICESat data. Processing of ASTER evaluation was done in collaboration with Dr. Jean-Paul Boy (University of Strasbourg) while visiting NASA/GSFC at the Planetary Geodynamics Laboratory.

References

ASTER GDEM Validation Team (2009). ASTER global DEM validation summary report. METI & NASA, 28pp.

Carabajal, C.C., and D. J. Harding (2005). ICESat validation of SRTM C-band digital elevation models, *Geophys. Res. Let.,*32, L22S01, doi:10.1029/2005GL023957.

Carabajal, C.C. and D. J. Harding (2006). SRTM C-band and ICESat laser altimetry elevation comparisons as a function of tree cover and relief, *Photogram. Eng. and Rem. Sens*., 72(3), 287-298.

Carabajal, C.C., D.J. Harding, J.-P. Boy, J.J. Danielson, D.B. Gesch and V.P. Suchdeo (2011). Evaluation of the global multi-resolution terrain elevation data 2010 (GMTED2010) using ICESat geodetic control, SPIE Proceedings, International Symposium on LIDAR and Radar Mapping: Technologies and Applications (LIDAR & RADAR 2011), Nanjing, China.

Carabajal, C.C. (2011) ASTER global DEM version 2.0 evaluation using ICESat geodetic ground control. Report to the ASTER GDEM Version 2 Validation Team.

Crippen, R.E. (2009). Spatial resolution of the ASTER global elevation model (GDEM). Presentation at the 35th ASTER Science Team Meeting, Kyoto, Japan.

Gesch, D., M. Oimoen, Z. Zhang, J. Danielson, D. Meyer (2011). Validation of the ASTER Global Digital Elevation Model (GDEM) Version 2 over the Conterminous United States. Report to the ASTER GDEM Version 2 Validation Team.

Guth, P.L. (2006). Geomorphometry from SRTM: comparison to NED: Photogrammetric Engineering and Remote Sensing, v. 72, no. 3, pp. 269-277.

Hansen, M., R. DeFries, J.R. Townshend, M. Carroll, C. Dimiceli, and R. Sohlberg (2006). Vegetation continuous fields MOD44B, 2001 percent tree cover, Collection 4, University of Maryland, College Park, Maryland.

Hansen, M., R.S. DeFries, J.R.G. Townshend, M. Carroll, C. Dimiceli, and R.A. Sohlberg (2003). Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm", Earth Interactions, Vol 7, No 10, pp 1-15.

Krieger, T., W. Curtis, and J. Haase (2011). Global Validation of the ASTER Global Digital Elevation Model (GDEM) Version 2. Report to the ASTER GDEM Version 2 Validation Team.

Schultz, B. E., H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio (2005). Overview of the ICESat mission, *Geophys. Res. Lett.*, 32, L21S01, doi:10.1029/2005GL024009.

Smith, B., and Sandwell, D. (2003). Accuracy and resolution of shuttle radar topography mission data: Geophysical Research Letters, v. 30, no. 9, p. 20-1 - 20-4.

Tachikawa, T., M. Kaku, and A. Iwasaki (2009). ASTER GDEM validation. Presentation at the 35th ASTER Science Team Meeting, Kyoto, Japan.

Tachikawa, T., M. Kaku, A. Iwasaki (2011a). ASTER GDEM Version 2 Validation. Presentation at the 39th ASTER Science Team Meeting, Tokyo, Japan.

Tachikawa, T., M., Kaku, A. Iwasaki (2011b) ASTER GDEM Version 2 Validation Report. Report to the ASTER GDEM Version 2 Validation Team.

Zwally, H.J., R. Schutz, W. Abdalati, J. Abshire, C. Bentley, J. Bufton, D. Harding, T. Herring, B. Minster, J. Spinhirne and R. Thomas (2002). ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics*, 34(3-4), 405-445