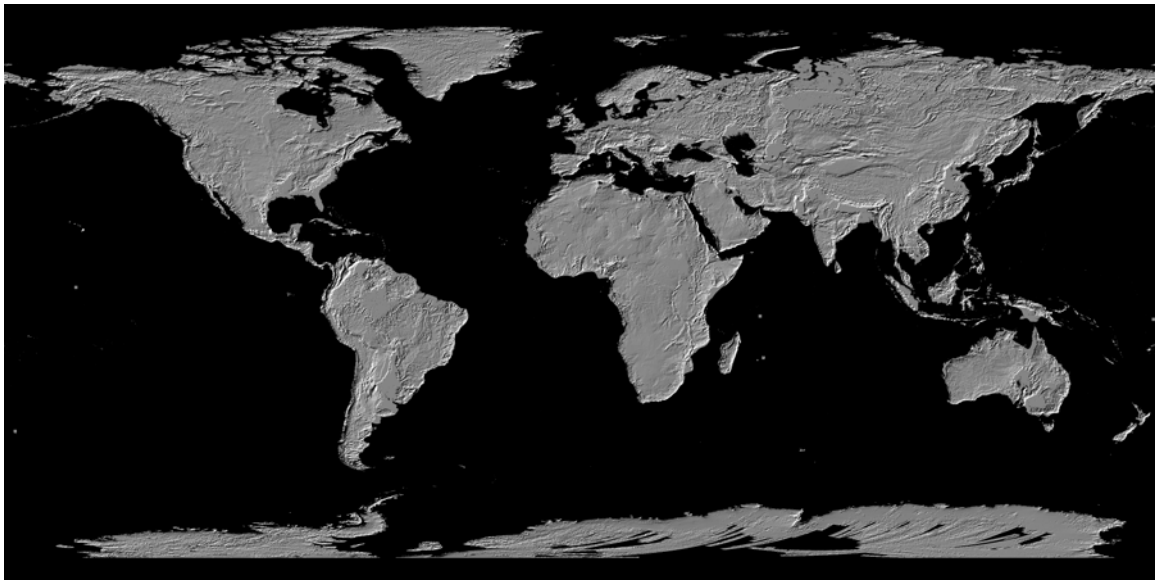
ASTER Global DEM Validation

Summary Report



Prepared by

ASTER GDEM Validation Team:
METI/ERSDAC
NASA/LPDAAC
USGS/EROS

In cooperation with
NGA and Other Collaborators

June 2009

ASTER GDEM Validation Summary Report

I. Introduction and Background.

This report presents highlights and the most relevant results from initial studies conducted to validate and otherwise characterize the new global digital elevation model (DEM) produced by the Ministry of Economy, Trade and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA) from optical stereo data acquired by the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). The ASTER Global DEM (GDEM) was released to the public on June 29, 2009. Results and findings presented in this report were instrumental in the decision by METI and NASA to release the ASTER GDEM.

A. ASTER

ASTER is an imaging instrument built by METI and operates on the NASA Terra platform. Images are acquired in 14 spectral bands using three separate telescopes and sensor systems. These include three visible and near-infrared (VNIR) bands with a spatial resolution of 15 meters (m), six short-wave-infrared (SWIR) bands with a spatial resolution of 30m, and five thermal infrared (TIR) bands that have a spatial resolution of 90 m. VNIR Band3 also is acquired using a backward-looking telescope, thus providing along-track stereo coverage from which high-quality digital elevation models (DEMs) are generated as one of a suite of ASTER standard data products. ASTER DEM standard data products are produced with 30m postings, and have Z accuracies generally between 10 m and 25 m root mean square error (RMSE).

B. ASTER GDEM

The methodology used to produce the ASTER GDEM involved automated processing of the entire 1.5-million-scene ASTER archive, including stereo-correlation to produce 1,264,118 individual scene-based ASTER DEMs, cloud masking to remove cloudy pixels, stacking all scene-based DEMs, removing residual bad values and outliers, averaging selected data to create final pixel values, and then correcting residual anomalies before partitioning the data into 1°-by-1° tiles.

The ASTER GDEM is being distributed by METI and NASA through the Earth Remote Sensing Data Analysis Center (ERSDAC) and the NASA Land Processes Distributed Active Archive Center (LP DAAC) at no charge to users worldwide as a contribution to the Global Earth Observing System of Systems (GEOSS). It is packaged in 1°-by-1° tiles, and covers land surfaces between 83°N and 83°S with estimated accuracies of 20 meters at 95 % confidence for vertical data and 30 meters at 95 % confidence for horizontal data. The ASTER GDEM is in geotiff format with geographic lat/long coordinates and a 1 arc-second (30m) grid. For each 1°-by-1° tile, two files are delivered: a) DEM data file; and b) a quality assessment (QA) file, which is a file that shows the number of scene-based DEMs contributing to the final DEM value at each pixel or the location of data anomalies that have been corrected and the data source used for the correction.

II. Validation Approach and Methodologies

Validation and characterization of the ASTER GDEM was the joint responsibility of the U.S. and Japanese partners. On the U.S. side, validation studies were led by the U.S. Geological Survey in

cooperation with NASA, the U.S. National Geospatial-Intelligence Agency (NGA), and 20 cooperators selected through an Announcement of Collaborative Opportunity (AO). On the Japan side, validation studies were conducted by ERSDAC in cooperation with the University of Tokyo and Mitsubishi Materials. The approach of each side was similar in that both derived statistical accuracies of the ASTER GDEM using various reference DEMs and a variety of more accurate ground control points. In addition, both sides attempted to characterize the ASTER GDEM in terms of features, artifacts, and residual anomalies that affect the overall quality of the product and may impact user application of the ASTER GDEM.

A. U.S. Validation Approach

The U.S. validation approach employed two primary components. The first derived a detailed accuracy assessment for the conterminous United States (CONUS), and the second sought to extend the detailed CONUS results to the rest of the globe using a sampling strategy.

The vertical accuracy of the ASTER GDEM was determined for CONUS by comparing the 934 ASTER GDEM tiles that comprise the CONUS area with corresponding 1-arc-second data from the USGS National Elevation Dataset (NED). Vertical accuracy of NED data is approximately 2-3 m RMSE. ASTER GDEM data were differenced with the corresponding NED data on a pixel-by-pixel basis, and minimum, maximum, mean, standard deviation, and RMSE values were computed on a pixel-by-pixel basis for each tile. The statistics were aggregated for all of CONUS, and they were further analyzed on the basis of cover type, number of individual DEMs contributing to each GDEM pixel value, and relief. Similar calculations and analyses also were performed between NED and Shuttle Radar Topography Mission 1-arc-second (SRTM1) data, and between GDEM and SRTM1 data, to facilitate comparisons with ASTER GDEM tiles from Alaska and continents other than North America.

Absolute vertical accuracies of ASTER GDEM tiles were measured by comparison with geodetic control point dataset from the National Geodetic Survey (NGS) that includes more than 13,000 ground control points (GCPs) distributed throughout the CONUS (Fig. 1). These points have centimeter-level accuracy in their horizontal and vertical coordinates, as they are produced by high-precision GPS observations on established survey benchmarks. Vertical accuracies of the NED and SRTM data also were assessed using the NGS reference points to give improved context to the ASTER results.

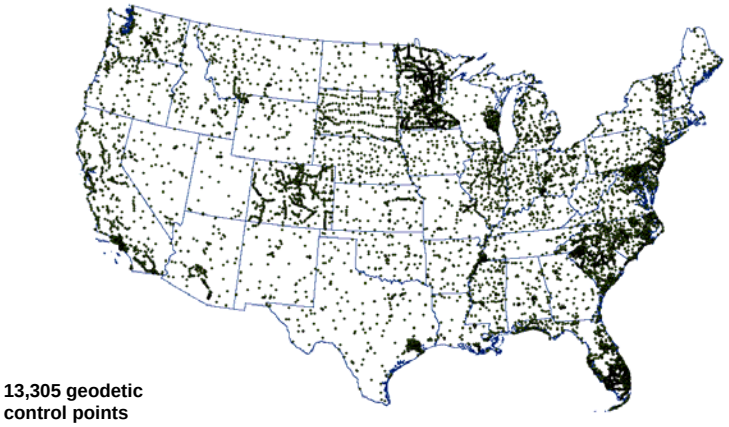


Figure 1. National Geodetic Survey GPS on Benchmark GPS locations.

In addition to the 934 tiles covering the conterminous U.S., approximately 350 ASTER GDEM 1ºx1° tiles covering land surface areas distributed across the globe were evaluated in cooperation with individuals selected through the AO process and in cooperation with colleagues at NGA. The 350 additional 1ºx1° tiles represent about 1.5% of the non-CONUS tiles.

AO collaborators were selected based on availability of both high-quality DEM reference data sets and absolute control covering international and Alaska sites that otherwise would not be available to ASTER GDEM assessment. NGA colleagues were invited to participate in ASTER GDEM validation studies because of the high quality reference data available to them and because of their interest in the ASTER GDEM. All collaborators were specifically asked to assess the accuracy of the ASTER GDEM using both reference DEM and absolute control. Results received from AO collaborators varied in quality and completeness, as well as in the degree to which assessment guidelines were adhered.

Virtually all investigators participating in the U.S.-side studies also observed and described various anomalies and artifacts in the ASTER GDEM, which are summarized in this report.

B. Japan Validation Approach

The Japan validation approach focused on measuring the accuracy of ASTER GDEM tiles throughout the country of Japan. Japanese efforts also focused on locating and characterizing various artifacts and residual anomalies in ASTER GDEM tiles from Japan and other parts of the world.

GDEM vertical accuracies were determined based on comparisons with two high-resolution DEMs produced by Japan's Geographical Survey Institute, GSI 5m and GSI 10m, which have 5m and 10m postings, respectively. GSI 5m was derived from laser profiles and aerial photogrammetry, and the effects of buildings and trees have been removed to create a "bare-earth" DEM. The GSI 10m is based on digitized topographic maps with scales between 1:5,000 and 1:25,000. ASTER GDEM data were differenced with the corresponding GSI 5m and GSI 10m data on a pixel-by-pixel basis, and minimum, maximum, mean, and standard deviation values were computed on a pixel-by-pixel basis for each tile. These statistics were further analyzed on the basis of cover type and terrain type. Similar calculations and analyses also were performed between GSI 5m and GSI 10m data and Shuttle Radar Topography Mission 3-arc-second (SRTM3) data.

Vertical accuracies also were measured against nearly 300 GCPs collected from throughout Japan (Fig. 2). Two collections of GCPs were used. One collection, consisting of 82 GCPs with <3cm accuracy, was developed for periodic testing of the geolocation accuracy of ASTER standard data products. In addition, 205 GCPs with <1cm accuracy were used from the GEONET: GPS Earth Observation Network System, which is a continuous GPS observation system established by Geographical Survey Institute. GEONET consists of approximately 1200 control points across Japan.

In addition to characterizing residual anomalies and artifacts in the ASTER GDEM, Japanese investigators assessed the horizontal or positional accuracy of the ASTER GDEM, and they assessed its effective ground resolution.

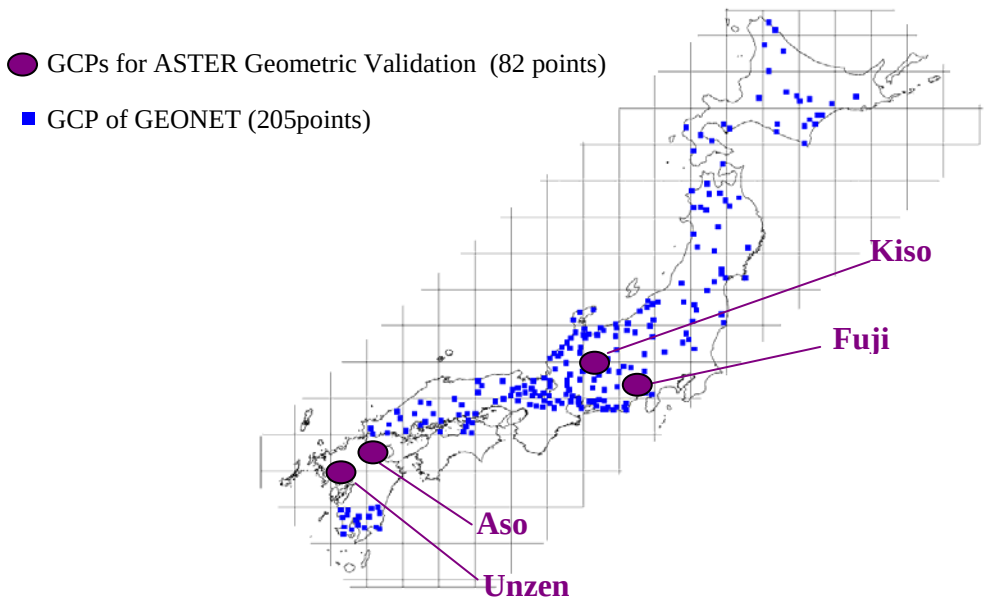


Figure 2. Distribution of Japan ground control points used in ASTER GDEM validation.

III. Summary and Discussion of Validation Results

This section of the ASTER GDEM Validation Summary Report presents the most relevant results of the various statistical accuracy tests performed during the course of these validation studies. Also presented and described are examples of the more common and/or significant residual anomalies and artifacts found to be present in the ASTER GDEM.

A. Raster-Based Accuracy Results

Table 1 presents, in three sub-tables, the mean elevation difference, standard deviation, and R**MSE** calculated for all approximately ***11 billion pixels*** (**ALL CONUS**) that comprise the ASTER GDEM of the conterminous U.S. as compared to NED, SRTM1 and SRTM3. Similar calculations are included for SRTM1 and SRTM3 compared with NED. In addition, means, standard deviations, and RMSEs are presented for the NLCD water class, for three aggregated NLCD land cover type classes (urban, forest, and open), and one additional category that seeks to reduce the effects of water and snow/ice. Unless otherwise noted, vertical accuracy results presented in this report include contributions from geolocation errors of the ASTER GDEM.

Mean Difference						
	ALL CONUS	Water	Urban	Forest	Open	**Excluding Water and Ice & Snow**
GDEM minus **NED** by NLCD	**-3.64**	**-1.32**	**-4.06**	**1.72**	**-6.40**	**-3.77**
GDEM minus **SRTM1** by NLCD	**-5.71**	**-0.53**	**-6.09**	**-4.91**	**-6.49**	**-5.99**
GDEM minus **SRTM3** by NLCD	**-5.64**	**-0.57**	**-6.12**	**-4.64**	**-6.49**	**-5.92**
SRTM1 minus **NED** by NLCD	**2.07**	**-0.95**	**2.02**	**6.65**	**0.14**	**2.23**
SRTM3 minus **NED** by NLCD	**1.87**	**-0.92**	**2.04**	**6.59**	**0.16**	**2.23**

Standard Deviation						
	ALL CONUS	Water	Urban	Forest	Open	**Excluding Water and Ice & Snow**
GDEM minus **NED** by NLCD	**8.75**	**15.71**	**6.94**	**9.93**	**7.31**	**8.19**
GDEM minus **SRTM1** by NLCD	**7.59**	**11.17**	**6.48**	**8.24**	**6.95**	**7.35**
GDEM minus **SRTM3** by NLCD	**8.09**	**11.11**	**6.69**	**8.99**	**7.43**	**7.88**
SRTM1 minus **NED** by NLCD	**4.41**	**2.24**	**3.71**	**6.10**	**3.56**	**4.50**
SRTM3 minus **NED** by NLCD	**6.18**	**2.39**	**4.15**	**7.93**	**4.78**	**5.90**

RMSE						
	ALL CONUS	Water	Urban	Forest	Open	**Excluding Water and Ice & Snow**
GDEM minus **NED** by NLCD	**10.87**	**16.53**	**9.06**	**10.93**	**10.33**	**10.46**
GDEM minus **SRTM1** by NLCD	**9.91**	**9.30**	**9.50**	**10.01**	**9.95**	**9.94**
GDEM minus **SRTM3** by NLCD	**10.28**	**9.34**	**9.65**	**10.49**	**10.30**	**10.33**
SRTM1 minus **NED** by NLCD	**6.20**	**3.21**	**4.85**	**9.53**	**4.11**	**6.32**
SRTM3 minus **NED** by NLCD	**7.10**	**3.32**	**5.21**	**10.76**	**5.20**	**7.36**

Table 1. Raster-based ASTER GDEM vertical accuracy results for CONUS, including the NLCD water class and three aggregated land cover type classes. All values are in meters.

Observations. Table 1 summarizes a wealth of information and reduces the results of voluminous data processing in such a way that one must guard against oversimplifying or trivializing the results of the ASTER GDEM comparisons with NED and SRTM data for 934 CONUS tiles investigated. The ASTER GDEM final validation report will delve much deeper into these results. Several relevant observations of the data are presented below.

ASTER DEMs previously have been observed to show a general negative bias, which seems to average approximately 5 meters. That observation is confirmed in the CONUS ASTER GDEM data, where elevations average a bit less than 4 meters lower than NED data. The bias exceeds 6 meters for the "Open" land cover classes. That the bias varies depending on cover type also is demonstrated in the results. For example, as just mentioned "Open" classes, where the land cover is rather reflective of "bare earth," average 6.4 meters lower for GDEM data than for NED data, but forest classes average 1.7 meters higher for GDEM than for NED. The latter reflects the fact that in forested areas, the ASTER GDEM, like other DEMs derived from optical sensors, measures near-top-of-canopy, not the bare earth. The CONUS results suggest that SRTM DEMs also are reporting canopy elevations.

The RMSE values are the best measure of the statistical accuracy of the ASTER GDEM for the 934 CONUS tiles, because the RMSE incorporates both bias and variation. The CONUS overall RMSE of 10.87 would equate to a CONUS GDEM accuracy at 95% confidence of **21.31 meters**, or a little more than the 20 m accuracy at 95% confidence estimated for the ASTER GDEM prior to its production. Even when inland water and snow/ice are removed from the calculation, the 10.46 RMSE still **slightly exceeds 20 m at 95% confidence.**

However, it is interesting to note that for a majority (558) of the 934 CONUS tiles, the RMSE is less than 10, which means the 95% confidence accuracy of each of those tiles is less than 20m. The likely reason for this is that there exist in the ASTER GDEM for CONUS a small number of rather significant "outliers" or anomalously high or low elevations. These outliers unduly affect the overall accuracy results for the CONUS tiles when viewed in their total.

It is interesting to note the substantial difference in RMSE of the water class for the ASTER GDEM (**16.53**) vs. NED compared with the SRTM1 RMSE (**3.21**) vs. NED. For Version 1 of the ASTER GDEM, **no inland water mask has been applied**, and all values for lakes and rivers are those that were automatically calculated in the generation of the ASTER GDEM. Consequently, most lakes and rivers have ranges of elevations in the ASTER GDEM, rather than a "flattened" single elevation for lakes and continuously decreasing elevations for rivers. The SRTM DEMs have been edited using a surface water mask that results in substantially more accurate water surface elevation in those data sets than for the ASTER GDEM, and that is reflected in the RMSEs for the water class.

Accuracy results for SRTM1 and SRTM3 vs. NED are presented in Table 1 for general interest and because they will be important in extrapolating CONUS results to certain non-CONUS tiles. Both SRTM1 and SRTM3 have smaller RMSEs compared to NED for CONUS than the ASTER GDEM.

Table 2 presents the mean difference, standard deviation, RMSE, and other information for 16 National Land Cover Data (NLCD) cover type classes for the CONUS tiles of the ASTER GDEM compared with NED data. The table primarily shows similar relationships as the aggregated land cover classes presented in Table 1, but at significantly greater detail.

ASTER GDEM minus NED								
Land Cover Type Name	**Land Cover Type %**	**Size (n)**	**Min. Value**	**Max. Value**	**Mean**	**Std. Dev.**	**RMSE**	**Land Cover Type Name**
All CONUS	100.00%	11,009,538,826	-717.1	3,934.0	**-3.64**	**8.75**	**10.87**	All CONUS
Water	5.25%	577,779,074	-473.9	3,360.8	-1.32	15.71	16.53	Water
IceSnow	0.02%	2,270,190	-466.5	300.5	-3.47	19.87	21.19	IceSnow
DevOpen	3.02%	332,207,863	-570.6	675.5	-4.33	7.47	9.68	DevOpen
DevLow	1.35%	148,143,563	-508.4	417.6	-3.89	6.23	8.30	DevLow
DevMed	0.52%	57,732,880	-421.4	198.4	-3.43	5.69	7.65	DevMed
DevHigh	0.18%	19,868,351	-364.3	197.5	-2.58	5.94	7.50	DevHigh
Barren	1.19%	130,930,015	-652.7	571.9	-5.53	14.39	16.02	Barren
DecidFor	11.28%	1,241,518,950	-716.6	3,934.0	2.73	9.05	10.38	DecidFor
EvGrnFor	12.43%	1,369,027,336	-713.7	3,208.7	0.66	11.45	12.13	EvGrnFor
MixedFor	2.09%	229,592,234	-717.1	3,152.8	2.35	8.64	9.95	MixedFor
ShrubSc	20.55%	2,262,790,180	-666.5	971.0	-5.75	8.29	10.50	ShrubSc
Grass	14.79%	1,628,834,364	-561.4	3,926.9	-7.14	6.76	10.49	Grass
Pasture	6.83%	751,591,087	-456.3	761.8	-5.55	6.02	8.81	Pasture
Crop	15.86%	1,746,327,462	-598.6	639.8	-7.30	6.21	10.15	Crop
WoodWet	3.42%	376,764,125	-410.7	3,210.8	1.85	7.07	8.41	WoodWet
HerbWet	1.22%	134,161,152	-222.3	1,227.4	-2.17	5.57	8.44	HerbWet

Table 2. Raster-based vertical accuracy results for CONUS by 16 NLCD cover type classes.

Observations. In addition to providing further, more detailed illustration of the effects of various cover types on mean elevation, the difficulty in achieving good correlation for areas of snow and ice are reflected in the high RMSE (**21.19**) of the ice/snow land cover class. The comparable RMSE for SRTM1 for the ice/snow class also is relatively high, but still about 10 meters less than for the ASTER GDEM.

Table 3 presents accuracy results for regional areas from seven different ASTER GDEM tiles located in the southern half of Japan. Two of the tiles have been divided into four sub-areas and one tile into two sub-areas for reporting. These results are particularly significant because the means and standard deviations were calculated after a geolocation error correction was applied to the ASTER GDEM. RMSEs were estimated from the means and standard deviation values.

Area		Elevation Error (m)				
		Min	Max	Mean	S.D.	RMSE (est.)
Fukuoka		-41	45	-2.41	5.16	6.37
Kochi		-59	60	1.87	7.18	8.12
Kyoto		-50	50	-1.85	5.95	6.88
Noubi	NW	-50	63	1.10	4.51	5.16
	NE	-48	45	-2.50	4.13	5.38
	SW	-56	62	4.18	6.66	8.75
	SE	-25	58	2.33	4.46	5.63
Osaka	N	-31	37	1.36	3.86	4.54
Saitama	W	-15	31	1.30	3.49	4.14
	E	-26	31	5.43	2.77	5.49
Tokyo	NW	-23	51	-0.75	5.87	6.25
	NE	-26	48	6.13	3.36	6.43
	SW	-23	37	1.11	5.22	5.78
	SE	-20	52	4.47	4.52	6.76

RMSE is estimated using (RMSE = 0.5 * | mean | + StdDev)

Table 3. Raster-based vertical accuracy results for 14 Japan ASTER GDEM test sites. Results are for ASTER GDEM – GSI 5m DEM. RMSE is estimated from mean and standard deviation.

Observations. In comparing results from Table 3 with those from Table 2, it is important to recognize there is no direct link between the reference DEMs used in each case, but both reference data sets have vertical accuracies of less than 3m RMSE, so the results should be quite comparable. The RMSEs for the Japan GDEM tiles are distinctly smaller than those for the CONUS GDEM tiles. It is quite possible the improvement shown in the Japan tiles is a result of the geolocation correction applied to those data, and the results of this comparison provide at least an initial quantitative hint at the possible effects of geolocation errors on the vertical accuracy of the ASTER GDEM.

Table 4 presents the geolocation errors calculated for the ASTER GDEM tiles whose elevation errors are reported in Table 3. Composite errors are reported for the three tiles whose elevations errors were reported by sub-region in Table 3.

	Fukuoka	Kochi	Kyoto	Noubi	Osaka	Saitama	Tokyo
Geolocation Error E-W (m)	**-19.25**	**-16.55**	**-23.63**	**-15.24**	**-8.33**	**-17.25**	**-14.23**
Geolocation Error N-S (m)	**-5.40**	**20.68**	**13.04**	**13.96**	**57.05**	**27.63**	**17.82**

Table 4. Geolocation errors for seven ASTER GDEM tiles from Japan.

Observations. The ASTER GDEM geolocation errors for these seven tiles clearly vary some, but the direction of error is pretty consistently to the northwest, with an average offset of North 20.68m; East -16.35m. This sampling is too small to draw general conclusions about overall ASTER GDEM geolocation error. However, these horizontal errors would seem to be sufficiently large that their correction could reasonably be expected to reduce elevation errors in the corrected tile, as likely is the case with the results reported in Table 3.

Table 5 presents raster-based accuracy results for the ASTER GDEM and SRTM3 for three tiles in Japan, which are characterized by steep mountainous terrain and heavily forested slopes. Forest cover exceeds an estimated 80% in each of the three tiles.

1°-by-1° Tile	ASTER GDEM			SRTM3		
	Mean	Std. Dev.	RMSE (est.)	Mean	Std. Dev.	RMSE (est.)
N32 E131	0.31	13.13	13.29	3.84	8.55	10.47
N33 E133	7.68	16.26	20.10	9.28	8.88	13.52
N36 E137	1.45	15.58	16.31	4.61	9.21	11.52

RMSE is estimated using (RMSE = 0.5 * |mean| + StdDev)

Table 5. Raster-based vertical accuracy results for three Japan ASTER GDEM tiles. Results are for ASTER GDEM – GSI 10m DEM and SRTM3 – GSI 10m DEM. RMSE is estimated from mean and standard deviation values.

Observations. Table 5 reports larger RMSEs than reported in Table 3 for three ASTER GDEM tiles located rather close to the tiles represented in Table 3. There are a number of possible reasons for the lower accuracy for these three tiles, and all likely are factors. First, no geolocation error correction was applied to Table 5 tiles. Secondly, the results may reflect the effect that high relief/steep terrain has on decreasing the accuracy of the ASTER GDEM, as demonstrated in Figure 3 for one of the tiles from Table 5. Thirdly, the larger RMSEs also likely reflect the influence of the dominant forest cover that is characteristic of these three tiles. Similar effects are seen for the SRTM3 data, but the RMSEs of the SRTM3 data are consistently lower than for the ASTER GDEM, as they were for CONUS.

Table 6 presents aggregated raster-based accuracy results for CONUS tiles based on terrain relief. Mean, standard deviation, and RMSE values are plotted for five ranges of relief for GDEM – NED. For these calculations, the "relief window" was set at 1500m-by-1500m, or just under 1 square mile. For CONUS, 62% of the tiles have relief of 40m or less.

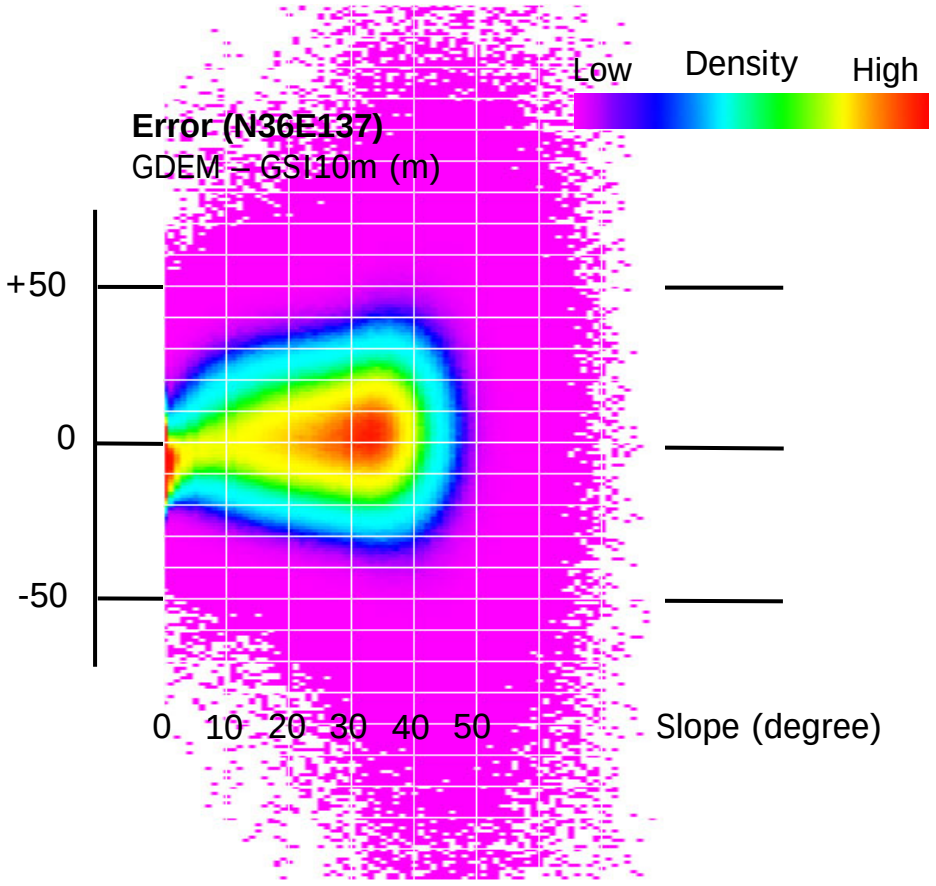


Figure 3. Graph of ASTER GDEM vertical error plotted against slope for GDEM tile N36 E137.

Relief (m)	Mean	Std. Dev.	RMSE
0 – 20	-4.86	6.23	9.19
20 – 40	-4.76	6.32	8.94
40 – 200	-3.35	7.85	9.76
200 - 600	-1.12	13.58	14.76
>600	-1.95	23.32	24.27

Table 6. Raster-based vertical accuracy results for CONUS based on terrain relief.

Observations. Table 6 results are consistent with Table 5 results and with Figure 3 in that they illustrate a decrease in ASTER GDEM accuracy when terrain relief becomes high (larger average slope). However, interestingly, relief does not seem to be much of a factor in GDEM accuracy at low to intermediate relief. Table 6 data also reflect the negative mean bias in ASTER GDEM data, and the tendency toward more positive mean difference elevations between GDEM and NED shown in Table 5 may be a function of a relative increase in forest cover type at higher elevations where greater relief typically is more common.

Table 7 presents raster-based accuracy results from CONUS ASTER GDEM tiles as compared to both NED and SRTM1 data, but as a function of the number of scene-based ASTER GDEMs (stack number) that contributed to each individual pixel in the CONUS

ASTER GDEM tiles. Theoretically, the ASTER GDEM RMSE should decrease with increasing stack number.

Mean Difference							
	ALL	Exclude fills	NUM 0 to 4	NUM 5 to 8	NUM 9 to 15	NUM 16 to 30	NUM > 30
GDEM minus **NED** by NUM	**-3.48**	**-3.51**	**-0.51**	**-2.63**	**-3.88**	**-4.19**	**-4.34**
GDEM minus **SRTM1** by NUM	**-5.36**	**-5.39**	**-3.55**	**-4.99**	**-5.58**	**-5.81**	**-5.79**

Standard Deviation							
	ALL	Exclude fills	NUM 0 to 4	NUM 5 to 8	NUM 9 to 15	NUM 16 to 30	NUM > 30
GDEM minus **NED** by NUM	**8.94**	**8.93**	**15.10**	**9.02**	**8.19**	**7.88**	**7.22**
GDEM minus **SRTM1** by NUM	**7.75**	**7.73**	**14.45**	**7.79**	**6.80**	**6.55**	**6.01**

RMSE							
	ALL	Exclude fills	NUM 0 to 4	NUM 5 to 8	NUM 9 to 15	NUM 16 to 30	NUM > 30
GDEM minus **NED** by NUM	**10.80**	**10.78**	**16.88**	**10.83**	**10.07**	**9.85**	**9.27**
GDEM minus **SRTM1** by NUM	**9.72**	**9.73**	**12.02**	**9.97**	**9.47**	**9.38**	**9.06**

Table 7. Raster-based ASTER GDEM vertical accuracy results for CONUS, showing variation by aggregated stack number.

Observations. Table 7 results confirm that ASTER GDEM RMSEs decrease with increasing stack number, both as compared to NED and SRTM1 as reference data sets. It is interesting to note, but perhaps not surprising, that exclusion of fill data does not affect RMSE in a significant way. The fact that the RMSE (vs. NED) actually improves slightly when fill values are removed seems a bit surprising. That may reflect smoothing that is part of the fill process, or it may be a result of removing 131,000 pixels of Canadian Digital Elevation Data (CDED) that somehow were included in the CONUS calculations. Those 131,000 pixels of CDED data had an RMSE of 253, so their removal would improve slightly the overall CONUS RMSE, as well.

Perhaps most significant point about Table 7 results is the fact that when the stack number is less than 4, the accuracy of the ASTER DEM is significantly less than when the stack number is greater than 4. Unfortunately, on a global basis, a rather high percentage of pixels have stack numbers less than 4, which is the primary factor that makes a straightforward extrapolation of CONUS results to the rest of the world rather difficult.

The most extensive assessment of ASTER GDEM tiles from regions outside CONUS was performed by colleagues at NGA, using SRTM1 data and other reference DEMs available to them. Figure 4 shows the general distribution of the 20 geographic areas and 284 GDEM tiles evaluated by NGA.

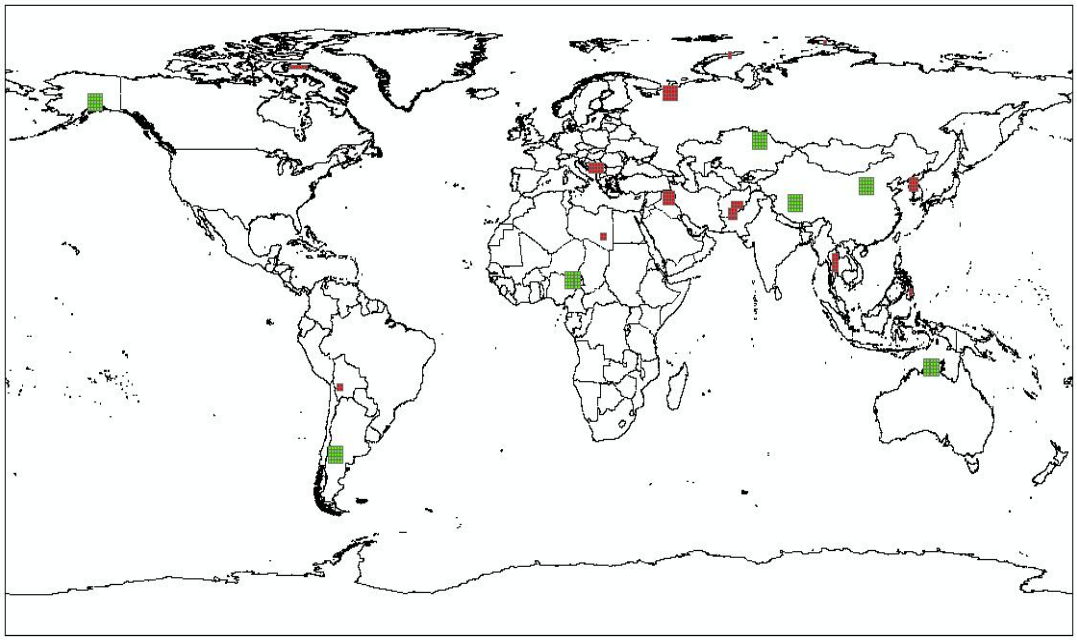


Figure 4. Locations of ASTER GDEM tiles evaluated by NGA colleagues. Sites using SRTM as the reference data set are shown in green.

Geographic Area	No. of Tiles	Mean Elevation Difference ASTER DEM – Reference DEM			
		SRTM DTED® 2.		Other Source	
		Mean (m)	90% L.E. (m)	Mean (m)	90% L.E. (m)
Afghanistan-1	8	-0.6	9.8	0.6 (a)	7.2
Afghanistan-2	9	-1.7	5.4	-0.9 (b)	3.7
Argentina	25	-5.2	6.5		
Australia	25	-4.7	7.9		
Bolivia	4	-10.9	8.1		
Bosnia	15	-4.0	9.1		
Canada	6			-14.0 (b)	16.2
China-1	25	-7.9	15.2		
China-2	25	-4.1	11.1		
Iraq	16	-5.0	6.3		
Kazakhstan	25	-5.5	6.8		
Korea	12	-6.0	10.8	0.0 (b)	14.9
Libya	4	0.0	11.0		
Nigeria	25	-13.3	4.7		
Philippines	2	-6.6	14.6		
Russia-1	2			-4.1 (b)	17.9
Russia-2	1			4.2 (b)	159.3
Russia-3	20			106.9 (b)	548.1
Thailand	10	-8.6	8.8		
U.S.A.-Alaska	25			-10.0 (c)	25.1

(a) IFSAR DEM DTED® 2.
(b) Photogrammetrically derived DTED® 2.
(c) Cartographic source DTED® 2 based on USGS NED.

Table 8. Raster-based vertical accuracy results for 284 ASTER GDEM tiles from 19 international test sites and 1 Alaska test site evaluated by NGA.

Table 8 presents the raster-based accuracy results generated by NGA for 284 tiles located in South America, Africa, Asia, Australia, Europe, and Alaska. Most results were produced by comparing ASTER GDEM data with the 1-arc-sec SRTM DTED® 2. Some results are based on comparisons with DTED® 2 data derived from other sources. Accuracies are expressed as 90% linear error (L.E.) in meters

Observations. Table 8 results go a long way toward answering the question of whether or not GDEM accuracy results obtained from the detailed evaluation of CONUS tiles can at least generally be extrapolated to the rest of the world. The answer, based on Table 8 results, is they clearly can be. CONUS results demonstrated that the ASTER GDEM RMSEs generally increase by only one meter or less when derived based on SRTM1 data as compared to NED data. Consequently, Table 8 results are very comparable to CONUS results.

When converted from 90% L.E. to accuracy at 95% confidence, ASTER GDEM accuracies reported by NGA for 16 of 20 sites studied meet or exceed the 20m accuracy at 95% confidence estimate of SILC. In addition, in at least two of the four cases that fail to meet the 20m accuracy at 95% confidence, the poor accuracy results likely are functions of residual cloud anomalies that were not removed in the cloud screening or anomaly removal steps applied during GDEM production. The poor result from Alaska may or may not be due, at least in part, to poorer quality DTED® 2 data available for those tiles.

Cooperators selected by AO who participated in ASTER GDEM validation activities also provided important information about GDEM accuracy for tiles located outside CONUS. However, synthesizing those results for presentation in a consistent and congruent way proved rather difficult for a variety of reasons. First, the reference data sets used varied substantially, and the stated accuracies for them were not always verifiable. In addition, the exact methodologies applied by investigators generally were not described in the report of results submitted. In some instances, we ran separate calculations on tiles where the RMSE reported seemed unusually high. In more than half of those instances, in which we differenced SRTM DTED 2 (30 m) data from ASTER GDEM data, our results provided lower RMSEs than those reported by the AO cooperator.

Table 9 is a compilation of virtually all of the raster-based accuracy results provided by those collaborators, including those that were unexpectedly high and some whose accuracy may be in question. Table 10 presents the examples of where lower RMSEs were obtained using SRTM reference data compared to AO investigator results for the same ASTER GDEM tiles.

AO Collaborator Raster-Based Accuracy Results

PI	Tile	Reference	Est. Ref. Accuracy	Min	Max	Mean	Std Dev	RMSE
ID17	57N154W	SRTM		-97	50	-0.90	4.18	4.18
ID14	N71 157W	InSAR DEM	2m	-14	31	2.40	3.30	4.5*
ID20	39N 009W	SRTM	10m DEM	-42	44	-1.99	5.25	5.57
ID10	34S 058W	SRTM		-64	443	1.83	5.99	6.26
ID10	33S 057W	SRTM		-108	87	0.55	6.32	6.34
ID19	31N 024E	HRS Ref3D	5m	-77	68	-4.98	4.79	6.91
ID10	34S 057W	SRTM		-103	86	0.89	6.95	7.01
ID10	35S 057W	SRTM		-95	87	0.16	7.60	7.60
ID19	31N 024E	SRTM	5m	-76	67	-6.33	4.38	7.70
ID15	12N 012W	10 m SPOT-5 DEM	3.87 m	-176	156	-4.13	7.71	7.75

ID19	43N 001E	FRA Nat. DEM	5m	-146	47	-9.56	6.06	8.00
ID02	36N 002E	SRTM	7m, 90% Conf	-305	203	-0.79	8.16	8.20
ID02	48N 008E	IRS-P5 DEM	8m, 90% Conf	-20	20	-4.10	7.29	8.36
ID13	49N 007E	Lidar DEM	0.1 m	-20	18	-4.94	7.07	8.61
ID15	12N 012W	SRTM		NA	NA	NA	NA	8.75
ID07	52N 010E	DTM Germany	4m, 95% Conf	-117	488	-1.46	8.16	8.89*
ID12	51N 001W	Lidar DEM	<1m	-76	106	-2.49	8.58	8.93
ID06	36N 140E	Prism DSM	6m	-97	143	-3.71	8.14	8.95
ID12	51N 001W	BlueSky DEM	1m	-155	131	-3.27	8.50	9.10
ID13	48N 008E	SRTM		-143	105	-2.84	9.03	9.46
ID15	05N 001W	SRTM		NA	NA	NA	NA	9.56
ID02	48N 008E	SRTM	5m, 90% Conf	-143	105	-2.90	9.19	9.63
ID13	48N 009E	SRTM		-198	615	-3.04	9.70	10.12
ID06	36N 140E	Lidar DSM	<1m	-83	47	-5.91	8.68	10.51
ID14	N70 157W	InSAR DEM	2m	-37	95	-2.80	9.20	10.60
ID03	50N 006E	Star3i DTM (IFSAR)	0.6 m	-251	410	-3.97	10.00	10.75
ID13	48N 009E	Lidar DEM		-182	41	5.48	9.61	11.06
ID01	05N 002W	SRTM		-122	110	-4.65	8.82	11.15*
ID13	49N 007E	DEM20	0.5 m	-66	99	1.43	11.37	11.45
ID07	48N 013E	DTM Germany	4m, 95% Conf	-88	71	-7.48	7.94	11.68*
ID12	52N 002W	BlueSky DEM	1m	-180	2512	-5.38	10.65	11.94
ID13	48N 008E	BW-50 DEM	3 m	-139	126	0.97	12.11	12.04
ID12	43N 005E	IGN DEM	1m	-316	226	0.84	12.03	12.06
ID13	48N 009E	BW-50 DEM	3 m	-192	632	-0.51	12.11	12.14
ID07	50N 010E	DTM Germany	4m, 95% Conf	-148	151	-1.71	10.62	12.33*
ID13	49N 007E	SRTM		-146	205	-9.34	8.37	12.54
ID19	43N 001E	SRTM	5m	-144	33	-11.34	5.63	12.66
ID17	31S 122E	SRTM		-28	14	-12.14	4.51	12.69
ID05	41N 002E	ICC DTM	1.1m	-256	85	-9.11	8.83	12.69
ID15	05N 001W	30m DEM	7.62 m	-279	318	-3.87	12.13	12.73
ID18	56N 134W	Lidar	.15m	-164	58	2.70	12.58	12.85
ID19	43N 001E	HRS Ref3D	5m	-45	70	-12.24	5.79	13.54
ID03	50N 006E	Star3i DSM (IFSAR)	0.6 m	-251	410	-10.30	9.74	14.17
ID12	41N 002E	ICC DEM		-243	106	-8.69	11.24	14.21
ID19	24N 120E	HRS Ref3D	10m	-295	210	-6.23	12.84	14.27
ID07	51N 010E	DTM Germany	4m, 95% Conf	-127	3623	-3.10	12.99	14.54*
ID07	47N 010E	DTM Germany	4m, 95% Conf	-269	383	-4.84	12.81	15.23*
ID05	41N 001E	ICC DTM	1.1m	-194	140	-12.99	8.21	15.36
ID17	18S 065W	SRTM		-91	47	-4.44	14.60	15.38
ID10	35S 056W	SRTM		-83	486	-0.29	7.02	16.32
ID12	43N 005E	UCL DEM	1m	-85	187	1.22	16.35	16.40
ID04	46N 007E	DHM25	1.5 to 3m	-256	550	1.17	16.36	16.40
ID06	35N 136E	Prism DSM	6m	-386	352	-8.43	14.31	16.61
ID04	46N 007E	SRTM C-band 90m		-411	373	-0.57	16.74	16.75
ID18	65N 150W	NED		-185	141	-10.80	14.20	17.80
ID01	45N 008E	SRTM		-443	251	-1.70	16.98	17.83*
ID04	46N 007E	Airborne LIDAR	0.5 to 1.5m	-250	269	-1.13	18.07	18.10
ID19	24N 120E	SRTM DTED1	16m	-250	189	-6.63	17.27	18.50
ID19	69N 031E	HRS Ref3D	10m	-265	3216	-15.65	10.78	19.01
ID18	68N 150W	IFSAR	3m	-201	250	2.63	19.63	19.80
ID06	46N 007E	Prism DSM	6m	-452	863	0.85	21.73	21.75
ID19	60N 149W	HRS Ref3D	10m	-395	355	0.14	22.13	22.13
ID11	46N 010E	SRTM		-518	526	-0.78	22.17	22.18

ID09	46N 007E	DHM25 (SwissTopo)	2-6m	-514	640	3.50	21.56	22.39
ID01	14N 087W	SRTM		-327	238	-9.61	17.60	22.4*
ID18	66N 156W	NED		-220	323	-14.10	17.40	22.40
ID06	35N 136E	Lidar DSM	<1m	-87	38	-11.88	19.07	22.46
ID06	28N 090E	Prism DSM	6m	-419	476	1.94	23.07	23.15
ID04	46N 007E	SPOT Reference 3D	10m slope<20d	-841	1295	-2.56	23.20	23.34
ID18	60N 154W	IFSAR	3m	-491	453	-5.54	23.48	24.11
ID16	11N 108E	DEM 1:50000 100m		-171	271	-9.03	26.90	26.90
ID18	60N 154W	IFSAR		-474	495	-5.30	27.30	27.80
ID14	30N 080E	SRTM		-782	343	0.40	28.70	28.9*
ID14	29N 080E	SRTM		-200	200	-10.90	25.20	30.65*
ID04	46N 008E	OrbiSAR X-Band DSM	1 to 2m	-1251	883	-2.65	30.80	30.91
ID04	46N 007E	Ikonos	3m	-309	766	-4.33	30.82	31.12
ID04	46N 008E	OrbiSAR P-Band DTM	2 to 3m	-1248	892	0.64	31.42	31.42
ID14	29N 081E	SRTM		-462	501	-8.10	28.90	32.95*
ID04	46N 008E	Swiss DUDES DEM	7 to 15m	-2420	1840	1.42	35.61	35.64
ID18	60N 154W	NED		-490	1033	-10.30	36.80	38.20
ID12	31N110E	China DEM		-598	588	-0.88	38.78	38.26
ID04	46N 008E	SRTM		-2408	881	3.17	43.81	43.93
ID11	46N 010E	Lidar DSM	5m	-194	60	-19.44	26.71	50.19
ID01	40N 023E	SRTM		-249	829	-8.51	49.00	53.25*
ID06	N27 086E	Prism DSM	6m	-1515	1306	-17.64	51.78	54.70
ID01	02S 079W	SRTM		-3591	3066	-13.18	51.02	57.61

*** Indicates where RMSE is estimated (RMSE = 0.5 * | mean | + StdDev)**

Table 9. Raster-based vertical accuracy results reported by 20 AO investigators for international and Alaska ASTER GDEM tiles.

PI	Tile	Reference	Min	Max	Mean	Std Dev	RMSE
ID19	24N 120E	HRS Ref3D	-295.00	210.00	-6.23	12.84	14.27
ID19	24N 120E	SRTM DTED1	-250.00	189.00	-6.63	17.27	18.50
GDEM Val Team	24N 120E	SRTM DTED2	-291.00	210.00	-1.78	6.77	**7.00**
ID12	43N 005E	IGN DEM	-316.00	226.00	0.84	12.03	12.06
ID12	43N 005E	UCL DEM	-85.00	187.00	1.22	16.35	16.40
GDEM Val Team	43N 005E	SRTM DTED2	-228.00	161.00	-6.72	6.82	**9.58**
ID12	41N 002E	ICC DEM	-243.35	105.58	-8.69	11.24	14.21
GDEM Val Team	41N 002E	SRTM DTED2	-261.00	117.00	-5.75	8.00	**9.85**
ID12	31N110E	China DEM	-598.41	588.29	-0.88	38.78	38.26
GDEM Val Team	31N110E	SRTM DTED2	-562.00	580.00	2.20	27.48	**27.57**
ID01	40N 023E	SRTM DTED1	-249.00	829.00	-8.51	49.00	53.25*
ID01	40N 023E	IceSat					46.10
GDEM Val Team	40N 023W	SRTM DTED2	-169.00	109.00	-8.26	8.07	**11.54**
ID01	45N 008E	SRTM DTED1	-443.00	251.00	-1.70	16.98	17.83*
ID01	45N 008E	IceSat					43.42
GDEM Val Team	45N 008E	SRTM DTED2	-428.00	182.00	-3.27	9.79	**10.32**
ID14	29N 080E	SRTM DTED1	-200.00	199.60	-10.90	25.20	30.65*
GDEM Val Team	29N 080E	SRTM DTED2	-3512.00	604.00	-9.20	21.10	**23.02**
ID06	N27 086E	Prism DSM	-1515.00	1306.00	-17.64	51.78	54.70
GDEM Val Team	N27 086E	SRTM DTED1	-1101.00	603.00	-12.24	30.19	**30.79**

*** Indicates where RMSE is estimated (RMSE = 0.5 * | mean | + StdDev)**

Table 10. Examples of improved results, compared with AO investigator results, obtained by ASTER GDEM Validation Team for the same tiles using mostly SRTM DTED2 data. Table includes comparisons with both raster reference data sets and absolute control data sets.

Observations. Taken as a whole, results obtained by the AO investigators for international and Alaska tiles show a lower accuracy for the ASTER GDEM than the results obtained by NGA for different international and Alaska tiles. Still, 25% of AO investigator tile vertical accuracy results show the ASTER GDEM meeting the estimated 20m accuracy at 95% confidence relative to the reference data sets, and the RMSEs for 70% of AO investigator tiles were reported at <20m. These numbers would improve further if the influence of inaccurate results was removed. Also, as was the case with the NGA results, some tiles evaluated by AO collaborators have an overall RSME that greatly exceeds the estimated accuracy specification because they contain residual anomalies (e.g., cloud top elevations) that can greatly skew the statistics. For many such tiles, the vast majority of the elevation values would be close to or well within the estimated 20 m at 95% confidence specification.

B. Absolute Control-based Accuracy Results

In addition to assessing the statistical accuracy of the ASTER GDEM against various reference DEMs, ASTER GDEM validation studies assessed GDEM accuracy against a large number of absolute ground control points, most with accuracy substantially less than 1 m. Table 11 presents the mean difference and RMSE for the ASTER GDEM, as well as for NED, SRTM1, and SRTM3, for more than 13,000 highly accurate GCPs located on surveyed benchmarks throughout CONUS (Fig. 1).

(NN = nearest neighbor; I = interpolated)	Number of Benchmarks	Min.	Max.	Mean	RMSE	Average Mean	Average RMSE
GDEM minus Benchmark Elevations (NN)	13,193	-128.64	55.31	**-3.71**	**9.33**		
GDEM minus Benchmark Elevations (I)	13,193	-127.74	105.41	**-3.69**	**9.37**	**-3.70**	**9.35**
NED minus Benchmark Elevations (NN)	13,193	-65.20	16.78	**-0.45**	**2.24**		
NED minus Benchmark Elevations (I)	13,193	-65.20	17.85	**-0.46**	**2.23**	**-0.45**	**2.24**
SRTM1 minus Benchmark Elevations (NN)	13,193	-63.66	37.23	**0.76**	**3.91**		
SRTM1 minus Benchmark Elevations (I)	13,193	-63.66	32.99	**0.77**	**3.84**	**0.76**	**3.88**
SRTM3 minus Benchmark Elevations (NN)	13,193	-63.66	28.46	**0.80**	**4.20**		
SRTM3 minus Benchmark Elevations (I)	13,193	-63.66	28.46	**0.81**	**4.04**	**0.81**	**4.12**

Table 11. Absolute-control-based ASTER GDEM vertical accuracy results for CONUS

Observations. The overall ASTER GDEM RMSE of 9.35 meters for more than 13,000 CONUS GCPs converts to 95% accuracy of 18.33 meters, which is within the estimated 20m accuracy at 95% confidence for the ASTER GDEM. Because of the higher accuracy of the GCPs on benchmarks than NED, **these results may be more accurate than the CONUS results presented in Table 1.** On the other hand, GCPs on benchmarks tend to be located along roads rather than along high ridges where the ASTER GDEM is less accurate than in areas with low to moderate relief. Consequently, it is difficult to say with certainty which set of validation data present the most accurate results. However, the two sets of CONUS results are consistent with each other and confirm that the ASTER GDEM generally, or approximately, meets the estimated accuracy specification of 20 m at 95% confidence for CONUS tiles.

Table 12 presents the mean difference and RMSE for the ASTER GDEM for nearly 300 GCPs collected from throughout Japan (Fig. 2).

	Number of GCPs	Min.	Max.	Mean	RMSE	Average Mean	Average RMSE
GDEM minus ASTER Geolocation GCPs	82	-20.2	29.8	**-5.26**	**9.95**	**-6.01**	**8.58**
GDEM minus GEONET GCPs	205	-18.5	5.35	**-6.31**	**7.97**		

Table 12. Absolute-control-based ASTER GDEM Accuracy Results for Japan GCPs

Observations.

Vertical accuracy of ASTER GDEM data based on absolute control for Japan tiles is very consistent with CONUS results based on absolute control, although the Japan results showed a larger negative bias for the GDEM data. This difference may be due, at least in part, to the difference in sample size. The 8.58 m average GDEM RMSE determined for the 287 GCPs from Japan converts to an average vertical error at 95% confidence of 16.81 meters. This easily meets the ASTER GDEM estimated vertical error specification of 20m at 95% confidence, lending further evidence that this specification may be met on a global basis. However, it also must be noted that the density of ASTER acquisitions over Japan and the United States are much greater than the global average, so the lowest ASTER GDEM vertical errors would be expected for Japan and CONUS parts of the data set.

Table 13 presents the GCP-based accuracy results generated by NGA from more than 45,000 GCPs from tiles located in South America, Africa, Asia, Australia, Europe, and Alaska. The GCPs applied by NGA are not as accurate as those applied in achieving the CONUS and

Geographic Area	No. of Tiles	No. of GCPs	ASTER-GCP 90% L.E. (m)	GCP 90% L.E. (m)	ASTER GDEM 90% L.E. (m)	ASTER GDEM 95% Confidence
Afghanistan-1	8	677	17.4	10	14.2	16.90
Afghanistan-2	9	2506	10.8	10	4.1	4.88
Argentina	25	1950	11.5	10	5.7	6.78
Australia	25	2821	12.9	10	8.1	9.64
Bosnia	15	4853	17.9	10	14.8	17.61
Canada	6	16	33.6	10	32.1	38.20
China-1	25	3774	22.6	10	20.3	24.16
China-2	25	4823	20.0	10	17.3	20.59
Iraq	16	5653	13.3	10	8.8	10.47
Kazakhstan	25	2118	15.5 (a)	10	11.8	14.04
Korea	12	7190	18.9	10	16.0	19.04
Libya	4	868	14.5 (a)	10	10.5	12.50
Nigeria	25	2384	16.5 (a)	10	13.1	15.59
Philippines	2	848	25.0 (a)	10	22.9	27.25
Russia-1	2	128	34.1	10	32.6	38.79
Russia-2	1	43	26.9	10	25.0	29.75
Russia-3	20	1214	24.8	10	22.7	27.01
Thailand	10	2640	17.4	10	14.2	16.90
U.S.A.-Alaska	25	2000	24.1	10	21.9	26.06

Table 13. Absolute-control-based ASTER GDEM accuracy results provided by NGA for 20 non-CONUS sites.

Japan results reported in Tables 11 and 12. Consequently, reported accuracies, expressed as 90% L.E. in meters, were adjusted by NGA from the GDEM-GCP results that also are reported in Table 1. The ASTER GDEM accuracy at 95% confidence column was calculated by the ASTER validation team by multiplying the 90% L.E. value by 1.19.

Observations. For the raster-based accuracy results reported by NGA, the ASTER GDEM met the accuracy estimate of 20m at 95% confidence for 16 of 20 geographic areas, relative to the DTED2 reference data sets used by NGA. In the case of the absolute-control-based accuracy results, the ASTER GDEM met absolute 20m accuracy at 95% confidence for 12 of the 20 geographic areas evaluated. Accuracy for no area exceeded 40 m at 95% confidence, and only two exceeded 30 m. In addition, possible explanations exist for some of the less accurate results. For example, the high result for the Canada site was based on only 16 GCPs, and the suspected residual cloud issues that affected the raster-based results for the same Russian tiles also could have been a factor, albeit much less so, in the absolute-control-based results. All things considered, the results reported in Table 13 would seem to support the earlier conclusion that detailed accuracy results obtained for CONUS tiles generally can be extrapolated to other tiles around the globe, caveated by the recognition that residual cloud anomalies will degrade the accuracy of any tile in which those anomalies occur.

Table 14 is a compilation of virtually all of the absolute-control-based accuracy results submitted by AO collaborators. More than 100,000 GCPs with variable accuracy and acquired by a variety of methods are represented in the table.

PI	TILE	Reference	Est. Ref. Accuracy	# of Pts.	Min	Max	Mean	Std Dev	RMSE
ID15	12N 012W	GPS Trimble Pro	1 m	41	-8	6	-0.89	3.14	3.23
ID10	33S 057W	GCP	< 1m	20	-4	10	1.46	4.15	4.40
ID10	34S 058W	GCP	< 1m	20	-7	10	1.17	4.52	4.67
ID17	57N 154W	ICESat GLAS	NA	2837	-43	79	1.63	4.99	5.25
ID10	35S 056W	GCP	< 1m	20	-24	2	-3.08	5.38	6.20
ID16	11N 108E	Nat Benchmarks	NA	3	-7	11	1.49	8.9	7.2
ID10	35S 057W	GCP	< 1m	20	-15	24	-0.63	7.24	7.26
ID02	36N 002E	GPS Kinematic	3m	NA	-124	66	-3.67	6.55	7.51
ID04	46N 007E	GPS ground control	0.2 to 0.4	131	-19	11	-5.23	5.51	7.58
ID09	46N 007E	GCP SwissTopo	< .10 m	622	-23	96	-1.35	7.59	7.71
ID17	31S 122E	ICESat GLAS	NA	4657	-32	18	-6.68	4.19	7.88
ID19	31N 024E	ICESat GLAS	2m	886	-19	5	-7.42	2.92	7.97
ID19	31N 024E	GPS Points	10m	42	-24	3	-3.23	7.43	8.10
ID10	34S 057W	GCP	< 1m	20	-17	28	-0.10	8.94	8.94
ID03	50N 006E	GPS	NA	36	-2	18	-8.51	4.32	9.55
ID19	24N 120E	ICESat GLAS	2m	73	-33	32	-1.89	9.58	9.70
ID15	05N 001W	GPS Trimble Pro	1 m	13	-4	18	6.33	6.55	10.01
ID05	41N 002E	ICESat points	NA	513	-49	13	-8.79	4.87	10.04
ID19	43N 001E	ICESat GLAS	2m	381	-22	10	-10.23	4.17	11.04
ID08	10N 76E	ICESat GLAS	8.1 m	13627	-230	138	-2.51	11.38	11.65
ID19	43N 001E	GPS Points	5m	44	-27	4	-10.62	6.59	12.50
ID18	65N 150W	ICESat GLAS	NA	1587	-55	34	-8.60	9.10	12.50
ID19	24N 120E	GPS Points	10m	29	-40	40	-2.29	12.81	13.01
ID08	63N 134E	ICESat GLAS	8.1 m	18735	-205	40	-10.38	8.00	13.10
ID19	69N 031E	ICESat GLAS	2m	524	-47	21	-11.23	8.01	13.79
ID05	41N 001E	ICESat GLAS	NA	2729	-49	19	-12.73	5.37	13.82
ID12	52N 002W	KGPS Survey	1 m	2970	-47	48	-9.70	9.85	13.83

ID17	66N 051W	ICESat GLAS	NA	1042	-102	215	-6.20	12.50	13.95
ID08	65N 151W	ICESat GLAS	8.1 m	12416	-149	920	-8.63	12.79	15.43
ID19	60N 149W	GPS Points	5m	24	-38	28	-9.28	12.80	15.81
ID17	18S 065W	ICESat GLAS	NA	405	-69	81	5.05	15.22	16.02
ID01	14N 087W	ICESat GLAS	NA	4172	NA	NA	NA	NA	16.67
ID12	51N 001W	KGPS Survey	1m	6495	-119	167	-8.88	15.27	17.66
ID18	66N 156W	ICESat GLAS	NA	859	-69	166	-10.60	16.20	19.40
ID18	60N 154W	ICESat GLAS	1m	1685	-86	182	1.80	20.30	20.30
ID19	60N 149W	ICESat GLAS	2m	400	-87	179	-7.06	21.65	22.75
ID17	78S 162E	ICESat GLAS	NA	1852	-224	193	-1.50	24.90	24.94
ID01	05N 002W	ICESat GLAS	NA	3545	NA	NA	NA	NA	27.47
ID01	02S 079W	ICESat GLAS	NA	5205	NA	NA	NA	NA	30.75
ID01	45N 008E	ICESat GLAS	NA	7673	NA	NA	NA	NA	43.42
ID01	40N 023E	ICESat GLAS	NA	5127	NA	NA	NA	NA	46.10

Table 14. Absolute-control-based accuracy results reported by 20 AO investigators for international and Alaska ASTER GDEM tiles.

Observations. In spite of some concerns similar to those previously expressed about uncertainty in the accuracy of reference data sets applied and lack of detailed information about the methodologies used, results presented in Table 14 are very consistent with results previously presented for non-CONUS tiles. Nearly half of the RMSEs reported convert to accuracies of 20 m or less at 95% confidence. In addition, most of those RMSEs that are higher than 10.2 m are for results from assessments that used ICESat GLAS data as the absolute control. When ICESat GLAS data are properly edited for spurious data points, they can provide very accurate absolute control. However, when the data are not properly edited before using as reference, they can yield biased and inaccurate results, as demonstrated in Table 10.

C. GDEM Artifacts and Residual Anomalies

An important objective of ASTER GDEM validation efforts was to characterize the ASTER GDEM in terms of specific features, such as artifacts and residual anomalies, that may affect the overall accuracy of the data, impede its use for certain applications, or just render it cosmetically unappealing. Indeed, it was determined that the **ASTER GDEM does contain residual anomalies and artifacts that most certainly degrade its overall accuracy, represent barriers to effective utilization of the GDEM for certain applications, and give the product a distinctly blemished appearance in certain renditions.** This section identifies, illustrates, and describes a variety of anomalies and artifacts that negatively impact the accuracy, utility, and overall appearance of the ASTER GDEM.

Residual Cloud Anomalies.

In spite of the fact that elevations from more than 1,250,000 individual, scene-based ASTER DEMs contributed to ASTER GDEM, there remain places on the Earth's land surface for which no cloud-free ASTER have been acquired. As a result of a combination of dominant weather patterns and ASTER acquisition scheduling, these areas occur dominantly in the tropics and in the extreme northern and southern latitudes (Fig. 5). In the case of the tropics, as well as many mid- and some high-latitude areas, most residual anomalies caused by the presence of clouds in the DEM source scene were replaced by SRTM DTED1, NED, or Canadian DEM data as part of the ASTER GDEM generation process. However, where such data were not available, most notably north of 60° north latitude and south of 56° south

latitude, residual cloud anomalies are present in the ASTER GDEM. These anomalies typically take the form of linear patterns that are "sub-swath" in width, but may be up to dozens of kilometers in length (Fig 6). The associated elevation anomalies (GDEM errors) may be several thousand meters in magnitude (Fig. 7). Since completion of initial validation studies, elevation anomalies caused by residual clouds have been replaced with -9999 values for those anomalous values detected on the Eurasian continent north of 60° north latitude. However, results presented in this report, still include effects of such residual clouds where they were present.

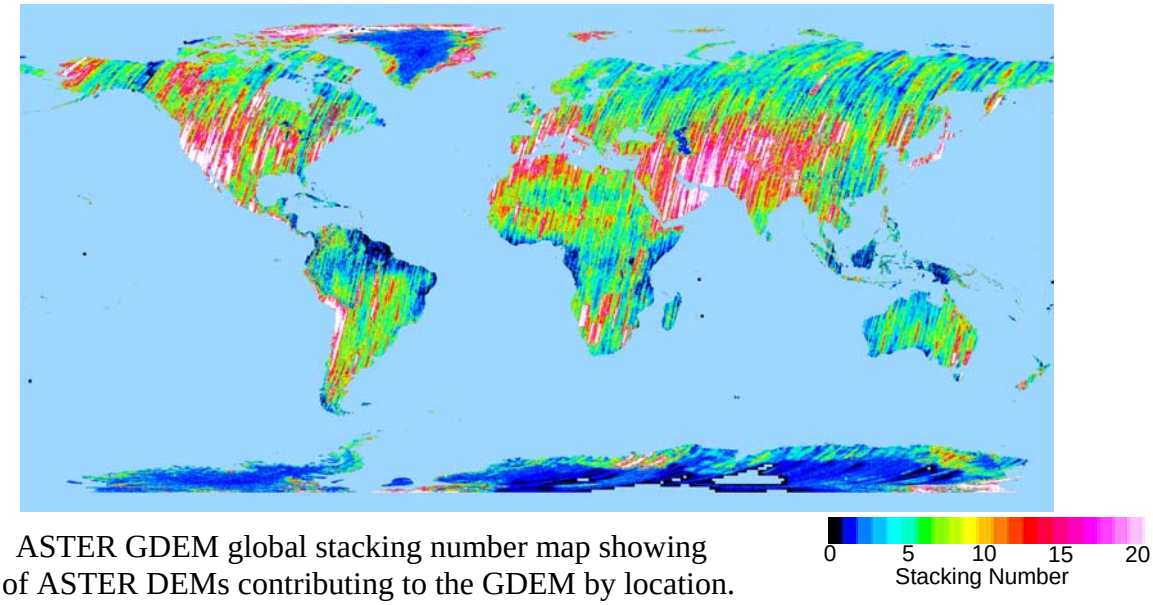


Figure 5. ASTER GDEM global stacking number map showing numbers of ASTER DEMs contributing to the GDEM by location.

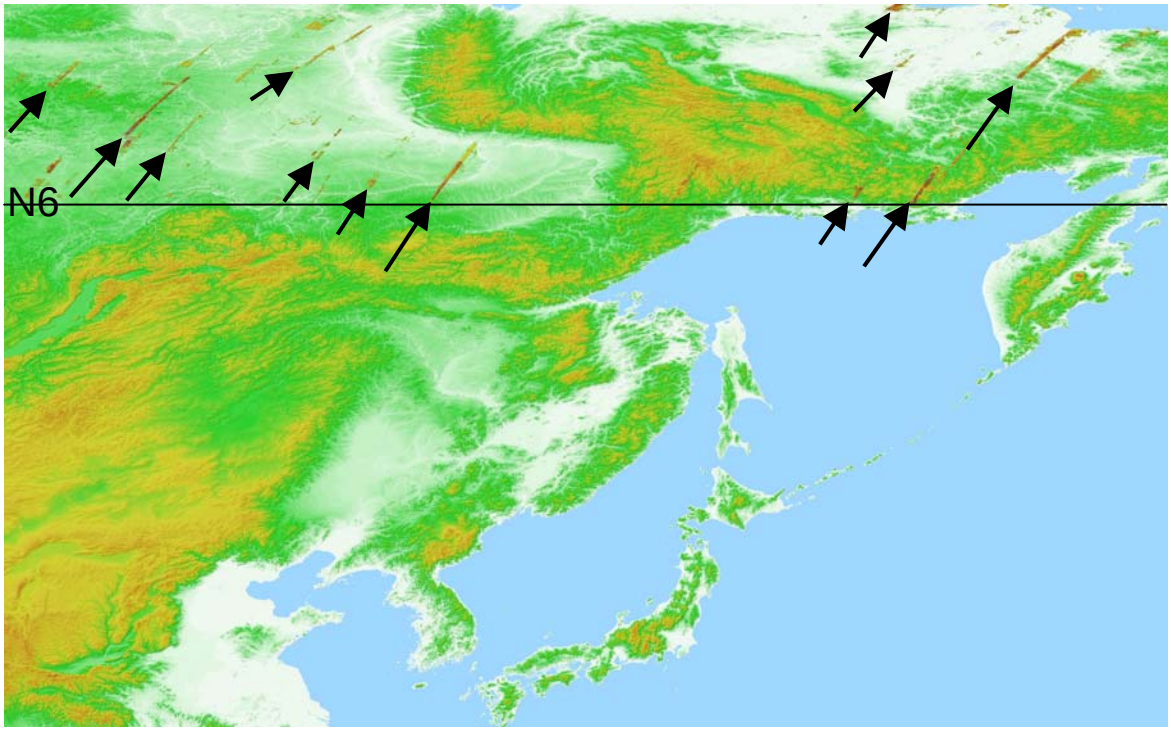


Figure 6. Examples of linear, residual cloud anomalies in the ASTER GDEM.

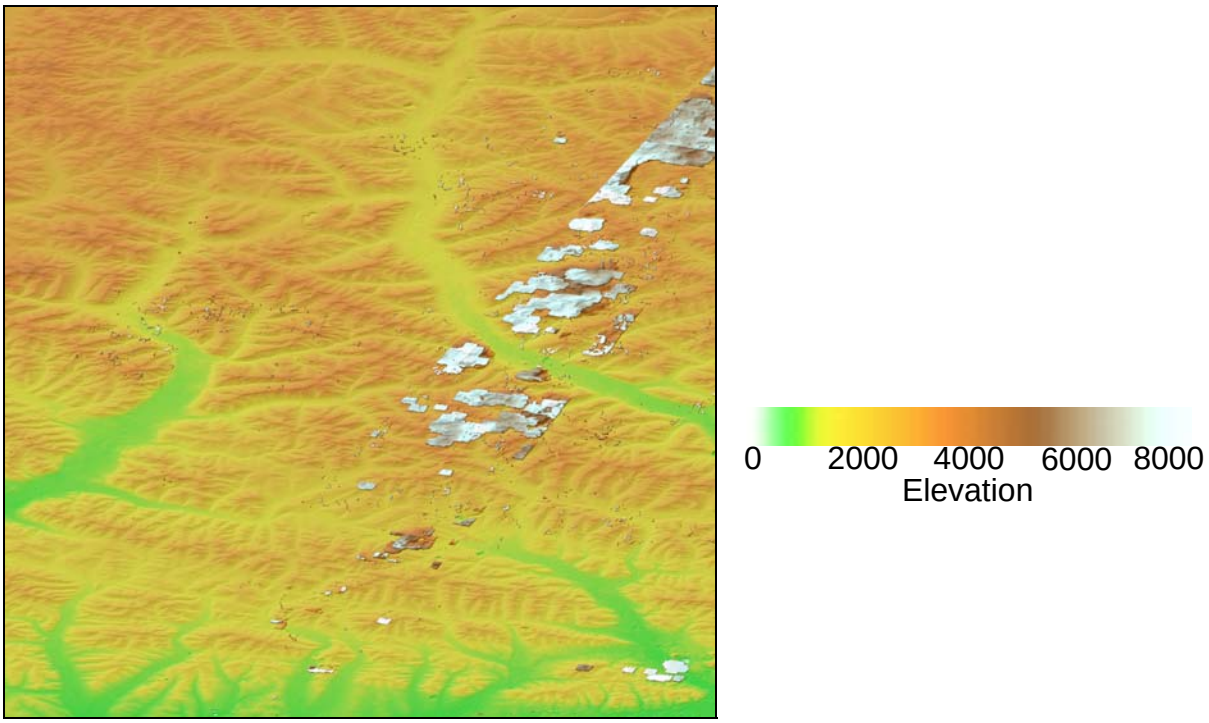


Figure 7. Residual cloud anomaly over shaded-relief surface.

Steps at Scene Boundaries.

The linear boundaries that exist between swath-oriented zones of two different stack numbers are very commonly the sites of "step anomalies" in the corresponding GDEM (Fig 8). Step anomalies are pervasive throughout the ASTER GDEM, but the magnitude of the offset at the step typically is around 10 meters (Fig. 9). It can be more in places.

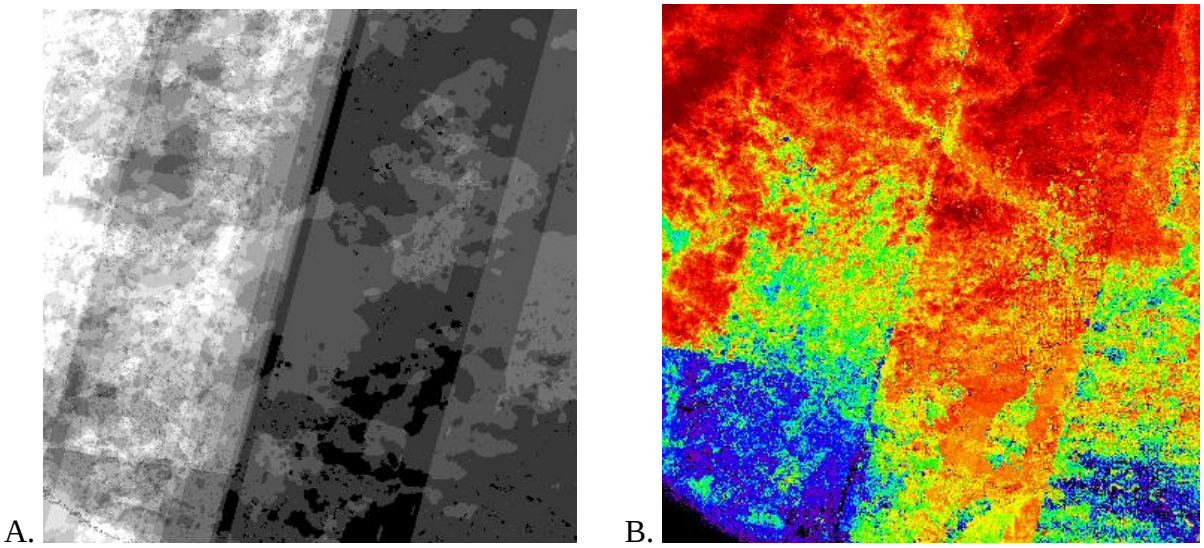


Figure 8. Examples of linear boundaries between swath-oriented stack number zones (A) and associated "step anomaly" in the corresponding ASTER GDEM tile (B).

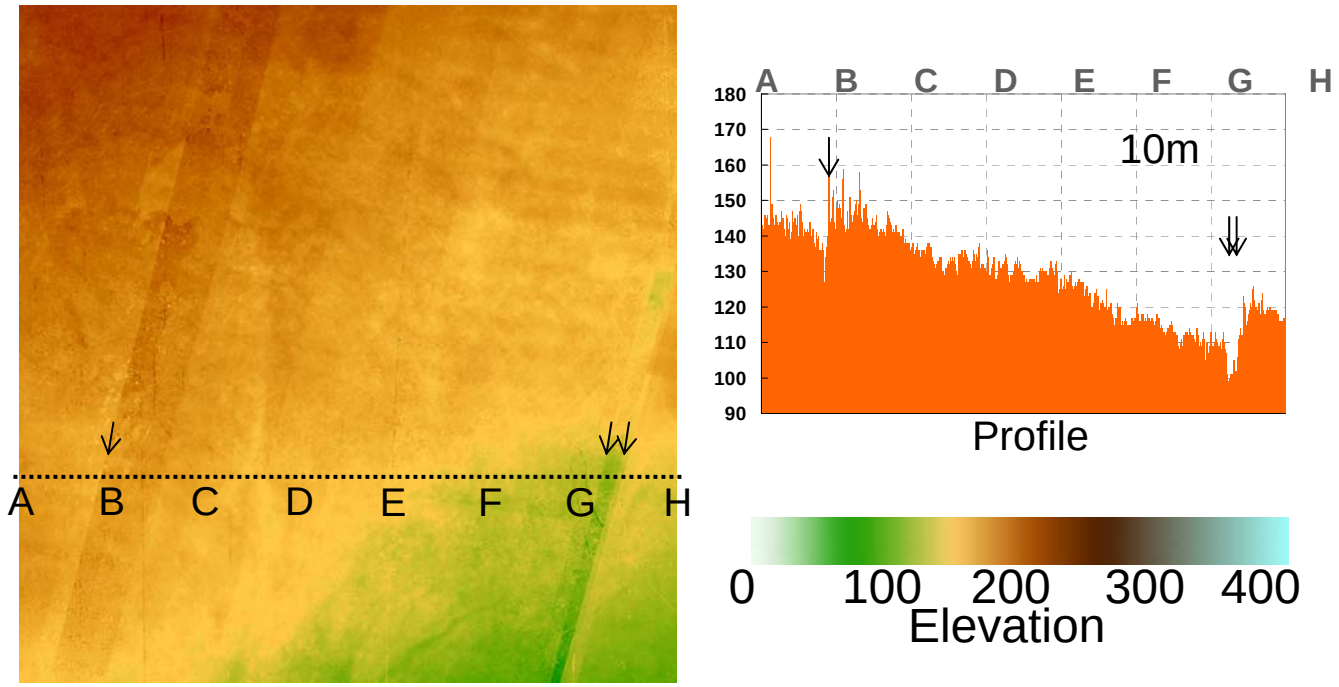


Figure 9. Example of "step anomaly" offset in an ASTER GDEM tile.

Artifacts Related to Irregular Stack Number Boundaries.

The vast majority of boundaries between different stack number zones, or areas, are irregular or more precisely, curvilinear. This is well illustrated in Figure 8B, where every grey level change represents a different stack number area. Unfortunately, these curvilinear boundaries seem to be the source of the vast majority of troublesome artifacts in the ASTER GDEM. The artifacts manifest themselves in several ways, three of which are illustrated below.

Pits, or small negative anomalies, occur with regularity and often high frequency in virtually all ASTER GDEM tiles examined during these validation exercises. The magnitude of the negative elevation anomaly associated with the pits varies dramatically from just a few meters to 100 meters or more. Figure 10 illustrates examples of the pit artifacts and their association with stack number boundaries.

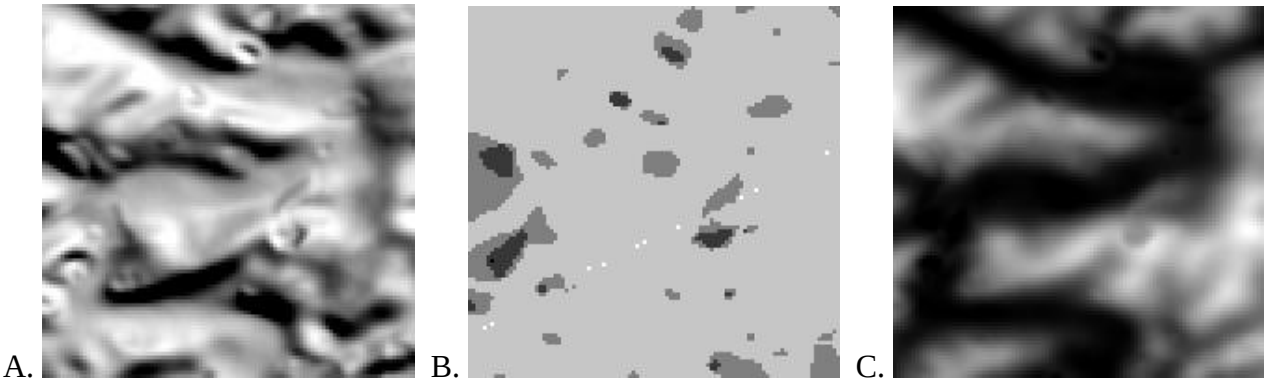


Figure 10. Examples of "pit" artifacts in an ASTER GDEM shaded-relief image (A) that are clearly related to the stack number boundaries (B). Pits typically are less apparent in the normal intensity ASTER GDEM images (C).

"Bumps" appear to be the positive elevation-anomaly-artifact-equivalent to pits. They occur with similar frequency as pits, and they also were found to be present in virtually every ASTER GDEM tile examined during the validation studies. Bumps often exhibit a small central depression, or a small pit. The magnitude of the positive elevation anomaly associated with bumps can range from just a few meters to more than 100 meters. Figure 11 illustrates examples of bump artifacts and their association with stack number boundaries.

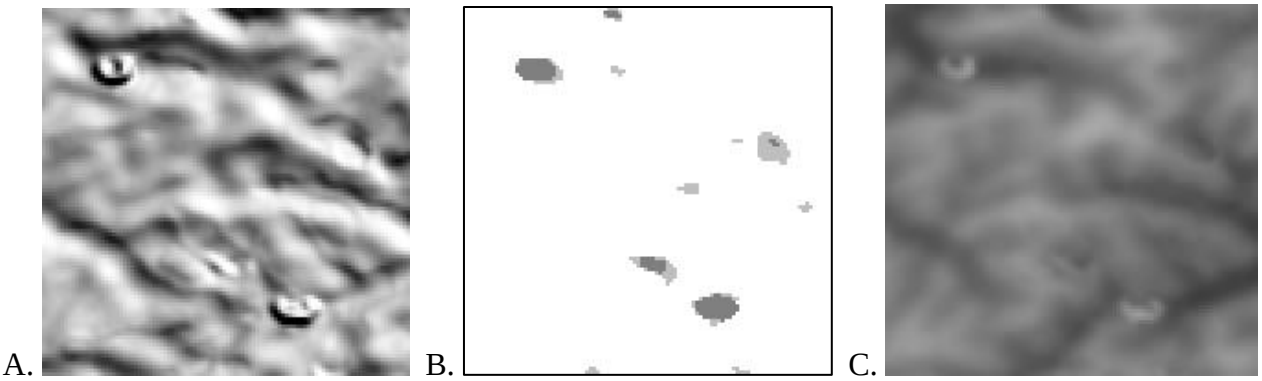


Figure 11. Examples of "bump" artifacts in an ASTER GDEM shaded-relief image (A) that are clearly related to the stack number boundaries (B). Bumps typically are less apparent in the normal intensity ASTER GDEM images (C), though in this example they are rather apparent.

"Mole Runs" are positive curvilinear anomalies that take the extended shape of some curvilinear stack number boundaries. Mole runs are less common than pits and bumps, and they occur most frequently in relatively flat terrain. The magnitude of the elevation anomalies associated with mole runs is typically much less than those associated with pits and bumps, ranging from barely perceptible to a few meters, and rarely more than 10 meters. Figure 12 illustrates examples of mole run artifacts and their association with stack number boundaries.

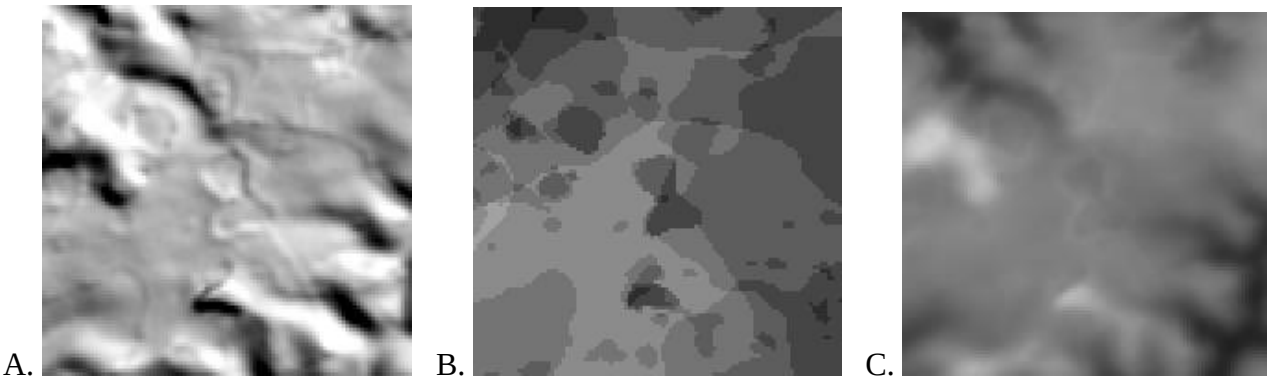
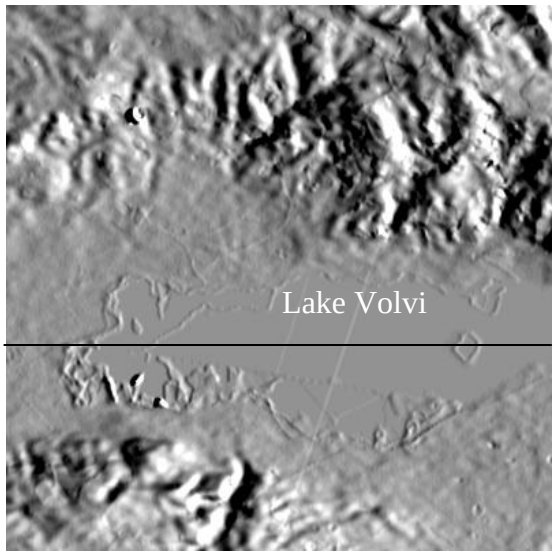
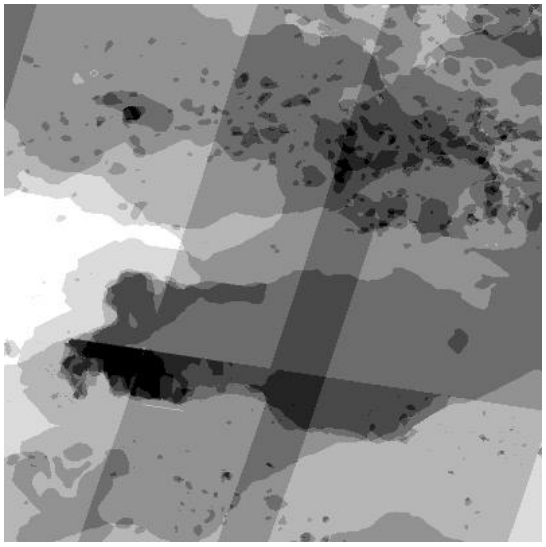
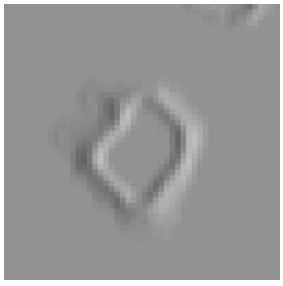
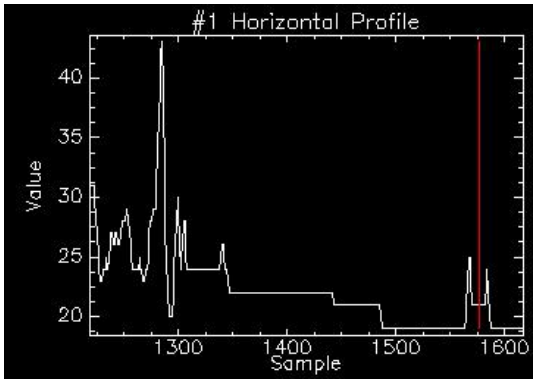


Figure 12. Examples of "mole run" artifacts in an ASTER GDEM shaded-relief image (A) that are clearly related to the stack number boundaries (B). Mole runs, particularly, are less apparent in the normal intensity ASTER GDEM images (C).

Figure 13 illustrates many of the artifacts described above in one rather small portion of the ASTER GDEM covering Lake Volvi in Greece. Step-anomalies, mole runs, pits, and bumps are present in Figure 13, and an elevation profile through some of them is included.

A. 

B. 

C. 
D. 

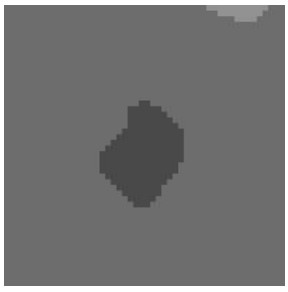
E. 

Figure 13. ASTER GDEM shaded-relief image (A) and corresponding stack number image (B). Image C is a close-up of the "pit-in-bump" artifact at the east end of the black horizontal profile line in image A, and image E is the close-up of the corresponding stack number area from image B. D is the ASTER GDEM elevation profile along the black profile line in image A.

Observations. Most of the linear and curvilinear artifacts apparent in Figure 13A also are elevation anomalies that are directly related to the stack number patterns in Figure 13B. The horizontal profile reveals that the elevation error associated with the "pit-in-bump" (Fig 13A & C) ranges between 3-6 meters. However, at the west end of Lake Volvi a number of elevation anomalies related to the stack number patterns have elevation errors of more than 20 meters. The linear "step-at-scene-boundary" anomalies are only 1-2 meters, at least where they cross Lake Volvi.

Interestingly enough, Lake Volvi is an example of where the ASTER GDEM portrays inland water body elevations as being reasonably consistent. The primary variations, actually steps, in Lake Volvi elevation occur at linear boundaries that exist between swath-oriented zones of two different stack numbers.

Inland Water Body Elevations and Noise.

Figure 14 illustrates the more typical situation with respect to ASTER GDEM portrayal of inland water body elevations. Figure 14(a) is the shaded-relief image of the ASTER GDEM for an area on the far north slope of Alaska, and Figure 14(b) is the shaded-relief image of the same area made from an airborne InSAR DEM. Figure 14 also illustrates one example of anomalous "grainy" textures in the ASTER GDEM, which numerous investigators referred to as "noise."

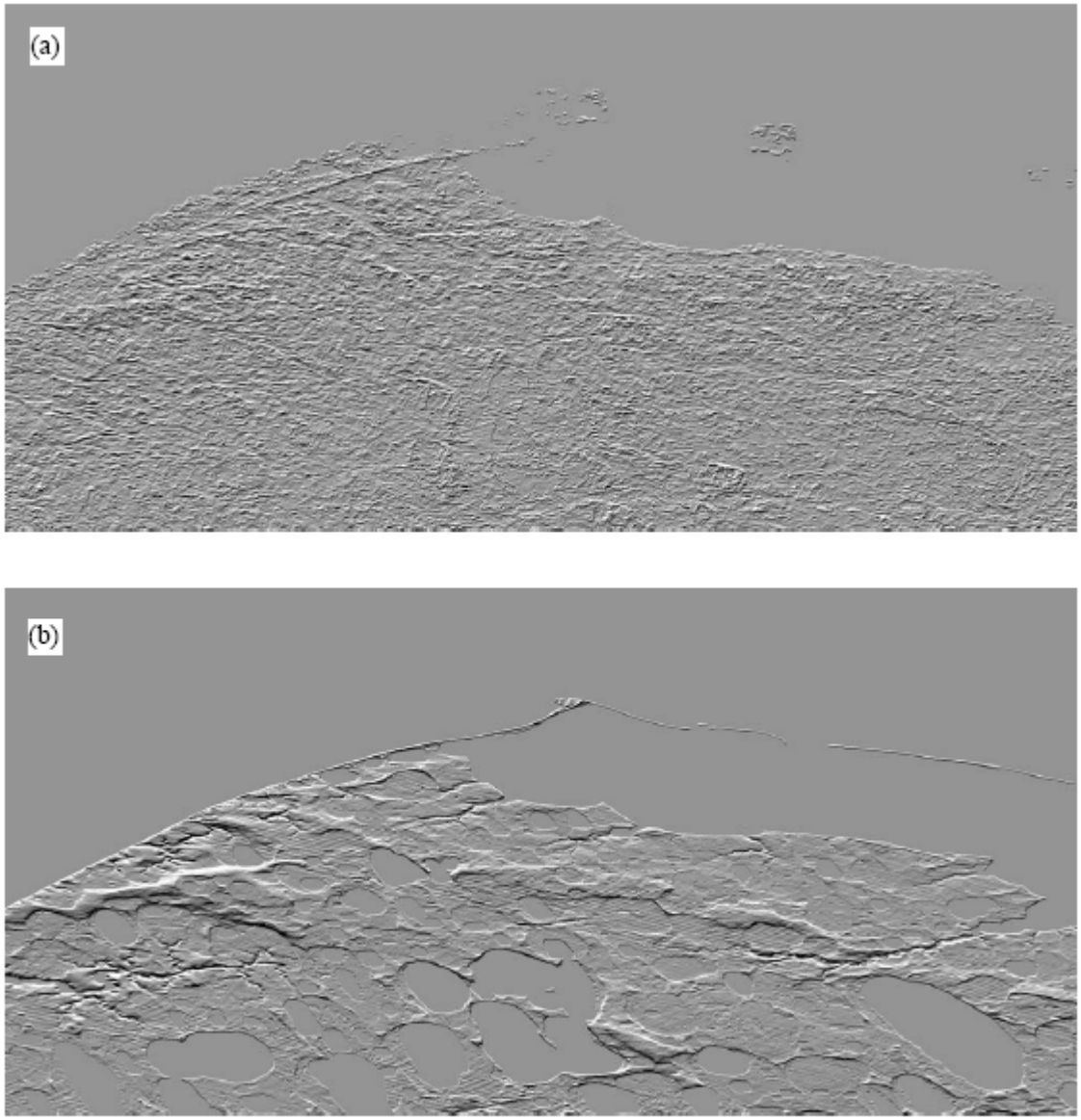


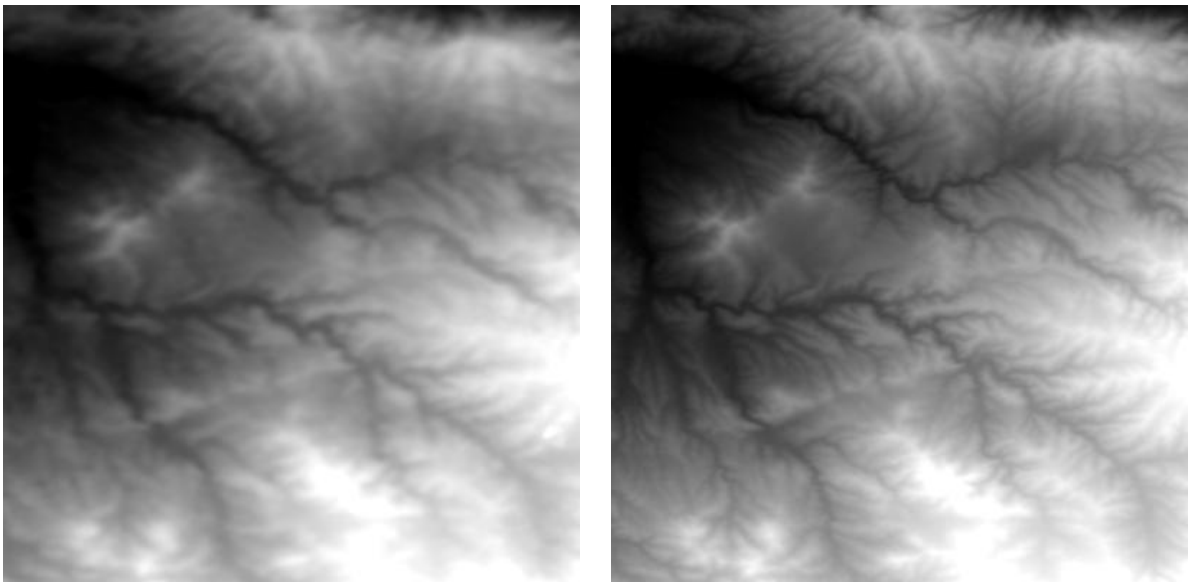
Figure 14. ASTER GDEM shaded relief image (a) of North Slope Alaska terrain compared with an airborne InSAR DEM (b) of the same area.

Observations. The area shown in Figure 14 is on the Alaska North Slope, where the tundra forms generally flat lying terrain underlain by permafrost and dotted with hundreds of lakes. The high-resolution DEM shown in Figure 14(b), produced from airborne InSAR data, has vertical accuracy of better than 2m and does a very good job of portraying this terrain. The

ASTER GDEM shows very much less detail, and lake locations are not readily apparent. Little topographic detail is visible in the ASTER GDEM, which here is characterized mostly by noise, and which likely is a function of low stereo correlation over this terrain. What is not known is when the ASTER images used to generate this part of the ASTER GDEM were acquired. If they are scenes with abundant snow cover, that would help explain the poor GDEM quality over this low relief area. Other artifacts previously described seem to be largely missing from Figure 14(a).

Spatial Detail of Topographic Expression Resolved by the ASTER GDEM.

While not really constituting an artifact or anomaly, the ability of the ASTER GDEM to resolve topographic features is a final topic of importance addressed in this Report. The ASTER GDEM has been described as a 30m DEM, as has the ASTER DEM standard data product, which also has 30m postings. However, various investigators have concluded that the ASTER DEM standard data product does not have 30m ground resolution. Similarly, various investigators have concluded that the SRTM DTED2 likewise does not have 30m ground resolution, but rather that the SRTM DTED2 ground resolution is approximately 50m. Figure 15 compares subsets of the ASTER GDEM with SRTM DTED2 data for the same area.



ASTER GDEM **SRTM DTED2**

Figure 15. ASTER GDEM and SRTM DTED 2 data covering the exact same area in Greece.

It is quite clear from even a casual examination the images in Figure 15 that the ASTER GDEM is not as sharp as the SRTM DTED2 and appears to contain less spatial detail. Studies conducted by Japanese investigators as part of the ASTER GDEM validation attempted to measure the spatial detail of topographic expression resolvable in the ASTER GDEM by comparing it with a series of degraded GSI 10m DEMs and calculating the standard deviations of the difference images. The plot of the standard deviations optimized at 3.8 arc seconds, meaning that the spatial detail resolvable by the ASTER GDEM, at least of the data tested, is slightly better than 120m.

IV. Summary and Conclusions

Studies were conducted by a large group of international investigators, working under the joint leadership of U.S and Japan ASTER Project participants, to validate the estimated accuracy of the new ASTER Global DEM product and to identify and describe artifacts and anomalies found in the ASTER GDEM. Detailed studies of CONUS ASTER GDEM data calculated an overall vertical RMSE for the 934 CONUS GDEM tiles of **10.87** meters, as compared to NED data. When compared with more than 13,000 GCPs from CONUS, the RMSE drops to **9.35** meters. These values convert, respectively, to vertical errors of just over and just under the estimated ASTER GDEM vertical error of 20 meters at 95% confidence. Studies conducted by Japanese investigators, colleagues from NGA, and other colleagues from around the globe confirm that detailed vertical accuracy results obtained for CONUS can, in general, be extrapolated with confidence to other parts of the world. Various factors effect local ASTER GDEM accuracy, so RMSEs for individual tiles vary from much better than the average CONUS results to considerably worse. However, the overall accuracy of the ASTER GDEM, on a global basis, can be taken to be approximately 20 meters at 95 % confidence.

The ASTER GDEM was found to contain significant anomalies and artifacts, which will affect its usefulness for certain user applications. There are two primary sources of these anomalies and artifacts. One is residual clouds in the ASTER scenes used to generate the ASTER GDEM, and the other is the algorithm used to generate the final GDEM from the variable number of individual ASTER DEMs available to contribute to the final elevation value for any given pixel. The latter class of artifacts and anomalies are the more significant, because they occur in virtually every ASTER GDEM tile, and the magnitude of the associated elevation error can be relatively large. In addition, the negative effects of residual clouds in the contributing ASTER scenes have been reduced in Version 1 of the ASTER GDEM, because major cloud anomalies over the Eurasian continent north of 60° north latitude recently were removed by replacing most of the anomalous elevations with -9999 values.

Another shortcoming of the current ASTER GDEM Version 1 is the fact that no inland water mask has been applied. Consequently, the elevations of the vast majority of inland lakes are not accurate, and the existence of most water bodies is not indicated in the ASTER GDEM.

After careful review and consideration of the results and findings presented in this Validation Summary Report, METI and NASA decided to release the ASTER GDEM for public use and further evaluation. METI and NASA acknowledge that Version 1 of the ASTER GDEM should be viewed as “experimental” or “research grade.” However, they have decided to release the ASTER GDEM, because they believe its potential benefits outweigh its flaws and because they hope the work of the user community can help lead to an improved ASTER GDEM in the future.

V. Participation and Acknowledgements

The ASTER GDEM validation and characterization studies that formed the basis of this Summary Report were led by the joint ASTER GDEM Validation Team with members from the United States and Japan. The Joint Validation Team also was responsible for preparation of this report. U.S. team members, representing USGS/EROS and/or NASA/LP DAAC, were G. Bryan Bailey, Jeff Danielson, Norman Bliss, Dean Gesch, Kenneth A. Duda, Gayla Evans, and Zheng Zhang. Various Japanese team members, representing ERSDAC, contributed significantly to the validation studies.

In addition, the Joint Validation Team wishes to acknowledge the important contributions of the following collaborators from NGA, James A. Slater, Barry Heady, George Kroenung, William Curtis, Jeffrey Haase, Daryl Hoegemann, Casey Shockley, and Kevin Tracy.

We also wish to acknowledge the participation of the investigators who all contributed to ASTER GDEM validation through studies conducted at their own expense in response to the Announcement of Collaborative Opportunity. They included: Don Atwood, University of Alaska Fairbanks, Dave Maune and Tim Blak, Dewberry; Laurent Cunin and Roland Gachet, IGN, Marc Bernard, SPOT; Claudia Carabajal and David Harding, Sigma Space Corporation; Peter G. Chirico, USGS; Pablo d'Angelo and Hossein Arefi, German Aerospace Center; Volker Hochschild, Jan Kropacek, and Felix Bachofer, University of Tuebingen; Michael Hovenbitzer, German Federal Agency for Cartography and Geodesy; André Jalobeanu and Matthieu Ferry, Portuguese Geographic Institute; Alexander Klaus, Sebastian Carl, Michael Vollmar, and Yvonne Fischer, GAF AG; Joshua Lyons, UNITAR/UNOSAT, Holger Heisig, Swiss Office of Topography; Rodolfo Mendez-Baillo and Carlos López, National Mapping Agency of Uruguay; Bryan Mercer, Qiaoping Zhang, and Michael Denbina, Intermap; Jan-Peter Muller and Shih-Yuan Lin, University College London; Robert Ryan and Mary Pagnutti, Innovative Imaging and Research; Emmanuel Baltsavias and Haris Papasaika, Swiss Institute of Geodesy and Photogrammetry, Paolo Pasquali and Alessio Cantone, sarmap s.a.; Hannes I. Reuter and Andrew Nelson, GISexperts, Peter Strobl and Wolfgang Mehl, JRC; Andrew Jarvis, International Center for Tropical Agriculture; Yongwei Sheng, UCLA; Takeo Tadono and Masanobu Shimada, JAXA, Junichi Takaku, RESTEC; and Pham Thi Mai Thy and Lam Dao Nguyen, Hochiminh City Institute of Resources Geography, Le Trung Chon, Hochiminh City University of Technology. Finally, we thank the CEOS Working Group on Calibration/Validation, the GEO DA-09-03d "Global DEM" task team, and the ISPRS Working Group on "Global DEM interoperability" for assistance in circulating our Announcement of Collaborative Opportunity.