A Global Assessment of the SRTM Performance

Ernesto Rodríguez, Charles S. Morris, and J. Eric Belz

Abstract

*The NASA/NGA Shuttle Radar Topography Mission (SRTM) collected interferometric radar data which has been used by the Jet Propulsion Laboratory to generate a near-global topography data product for latitudes smaller than 60°. One of the primary goals of the mission was to produce a data set that was **globally consistent** and with **quantified errors**. To achieve this goal, an extensive global ground campaign was conducted by NGA and NASA to collect ground truth that would allow for the global validation of this unique data set. This paper documents the results of this SRTM validation effort using this global data set. The table shown below summarizes our results (all quantities represent 90 percent errors in meters).*

	Africa	Australia	Eurasia	Islands	N. America	S. America
Absolute Geolocation Error	11.9	7.2	8.8	9.0	12.6	9.0
Absolute Height Error	5.6	6.0	6.2	8.0	9.0	6.2
Relative Height Error	9.8	4.7	8.7	6.2	7.0	5.5

In the paper, we present a detailed description of how the results in this table were obtained. We also present detailed characterizations of the height and planimetric components of the error, their magnitudes, geographical distribution, and spatial structure.

Introduction

Interferometric SAR is a technique that, for each point on the Earth's surface, measures the distance between two known points (the interferometric antennas) and the surface point. Given that the vector between the two antennas (the "interferometric baseline") and their absolute position is known through *in situ* measurements, the range to the point and the range difference (which is estimated from the interferometric phase difference between the two antenna signals) can be used to triangulate the location of the surface point in space. The surface height can then be inferred by relating this location to the appropriate datum. The details of the technique are subtle, and the reader is referred to the technical literature for further details (Rodriguez and Martin, 1992; Rosen *et al.*, 2000).

The Shuttle Radar Topography Mission (SRTM) (van Zyl, 2001; Farr and Kobrick, 2000) was a collaboration between NASA's Jet Propulsion Laboratory (JPL) and the National Geospatial-Intelligence Agency (NGA) to use a radar interferometer (Rodriguez and Martin, 1992, Rosen *et al.*, 2000) to generate a *globally consistent* digital elevation map (DEM) for latitudes smaller than 60°. As part of the SRTM mission, an extensive ground campaign was conducted by NGA and NASA to collect ground-truth that would allow for the global validation of this unique data set.

One of the primary goals of SRTM was to generate DEMs that have globally consistent and known errors. The purpose of this paper is to document the analysis that was undertaken to characterize the SRTM height and position errors. Other assessments of the SRTM performance have examined the spatial resolution of SRTM (Smith and Sandwell, 2003), or compared its performance against lidar for vegetated regions (Sun *et al.*, 2003). Unlike more specialized studies, we endeavor to characterize all components of the SRTM errors on a global basis.

In the next two sections, we review the ground-truth data used in this study and summarize the characteristics of the residual errors after calibration and continental least squares error adjustments. The remaining sections of this paper present the details showing how summary numbers were derived. Comparisons are made to verify that SRTM met the 16 m (90 percent) mission absolute height accuracy requirement followed by studies in greater detail the different spatial components in the SRTM errors, and looks in some detail at the long wavelength errors and the resolution capabilities of the instrument. The final sections concentrate on the high-frequency error components and validation of the Terrain Height Error Data (THED), a SRTM mission product available through NGA to selected investigators, and details of the methodology used to calibrate the system timing and geolocation and verification of the geolocation accuracy using radar identifiable kinematic GPS transects. We conclude by synthesizing our results into continental assessments of the SRTM height and position accuracies. The results presented in this paper are an extract of a longer JPL report (Rodriguez *et al.*, 2005) to which the reader is referred for greater details and further results.

Verification Data Sets

The global verification of the SRTM DEMs presented a unique challenge since the DEM accuracy was expected to be better than the accuracy of most topographic data sets which were globally available. To solve this problem, extensive ground truth data were collected with the understanding that the quality of these data might not be sufficient for a final assessment of the SRTM accuracy.

Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109 (ernesto.rodriguez@jpl.nasa.gov).

Photogrammetric Engineering & Remote Sensing
Vol. 72, No. 3, March 2006, pp. 249–260.

0099-1112/06/7203–0249/$3.00/0
© 2006 American Society for Photogrammetry and Remote Sensing

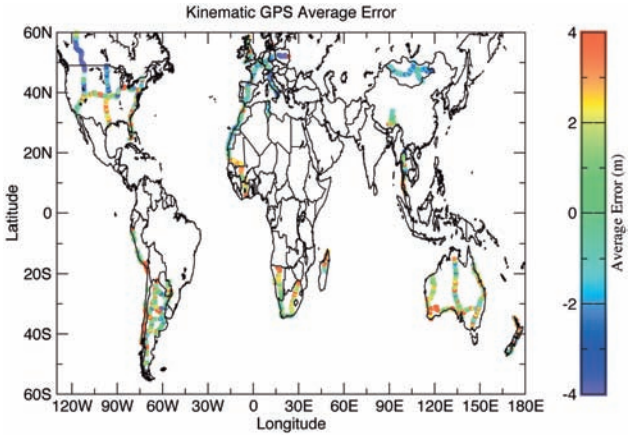

Plate 1. Map of SRTM minus kinematic GPS height differences averaged over 0.5° of latitude and longitude. Long wavelength trends in the error at continental scales can be seen from this figure, as well as isolated regions of higher errors.

The most extensive ground-truth effort was the collection of a globally distributed set of ground control points (GCPs) using kinematic Global Positioning System (KGPS) transects. The distribution of the data is shown in detail in Plate 1. A typical KGPS transect spanned a substantial part of a continent, thus allowing for the characterization of errors at all length scales.

The KGPS data were collected by driving vehicles carrying Global Positioning System (GPS) receivers along radar-identifiable roads, thereby producing a set of data representing a non-uniform sampling of the latitude, longitude and height. The GPS data were processed using the JPL GIPSY software (**http://gipsy.jpl.nasa.gov**). The total data collection yielded nearly 9.4 million samples covering six continents with a general accuracy of approximately 50 cm (1σ).

Regions with significant discrepancy between ground-truth and the SRTM heights (near over-passes and bridges, or in very forested regions) were eliminated during a visual quality assurance process. GPS data with a standard deviation of more than 1 meter within a 30 m pixel, or more than 5 meters over a 3 by 3 pixel box (90 m × 90 m) were removed. Furthermore, data within 3 pixels of a void posting were removed, since these areas were usually associated with SRTM phase unwrapping anomalies.

Multiple GPS samples within a single SRTM 30 m data pixel were averaged so that each SRTM pixel was equally weighted. The data were then divided into two data sets: one for validation of absolute height accuracy, the other for estimation of geolocation accuracy. The distribution of kinematic GPS GCPs and image ground-truth points for each continent is shown Table 1.

Table 1. Kinematic GPS Data Summary

"Continent"	No. of Kinematic GPS Tracks	No. of Cells With Data	No. of GPS Samples ×1000	No. of Truth Points ×1000
N. America	6	178	1750	402
S. America	5	196	1005	328
Africa	4	143	2439	395
Eurasia	11	192	2622	445
Australia	4	140	1145	381
New Zealand	2	31	393	64

Table 2. Summary of the Geographic Distribution of DTED® 2 Patches

Continent	Number
Africa	5
Australia	0
Eurasia	26
Islands	0
N. America	6
S. America	5

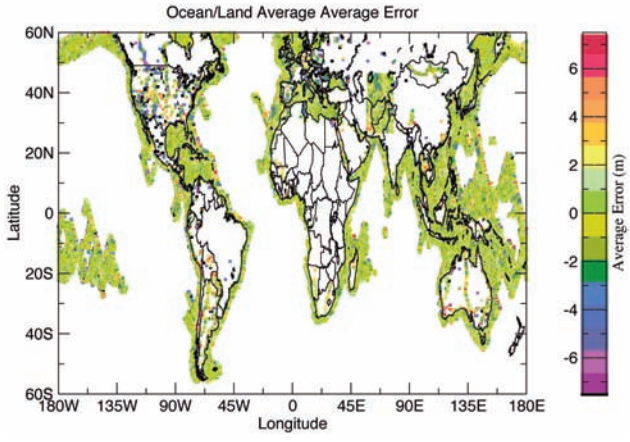

Plate 2. Map of residual long wavelength errors for land and ocean GCPs after mosaic corrections are applied. Notice that the errors over water are basically unbiased. Isolated regions of higher error, usually associated with GCPs or DTED® Level-2 patches, can also be seen.

As part of its Digital Terrain Elevation Data (DTED®) Level 2 generation, NGA produced a set of 42 one-degree cells whose accuracy was expected to be similar to, or better than the SRTM data, while its spatial resolution was similar to SRTM. The geographical distribution of DTED 2 cells is presented in Table 2. The locations of the DTED® data can be seen in as the isolated data patches over land in Plate 2.

In addition to the high-quality DTED® 2 cells, NGA generated a set of height "chips," and we used 141 in our analysis. The process used to generate these chips (or patches) was similar to the one used to generate DTED® 2 data, but the final data sets did not go through the same rigorous quality assessment and editing. The distribution of these patches was selected to be as random as possible, while concentrating on regions of some, but not extreme, relief.

The ocean surface topography, which is a combination of the geoid, the ocean circulation, and tidal change is very well known (Fu and Cazenave, 2001). In the deep ocean, altimeters have determined the ocean surface to centimetric levels and tidal and geoid models (e.g., EGM96) are accurate to similar accuracy. In coastal regions, both tidal and geoid models are less precise; nevertheless, the error in the topography is much smaller than 50 cm, which is more than sufficient to gauge the SRTM accuracy and provide GCPs for calibration purposes. For the SRTM mission calibration and validation, the EGM96, the TOPEX mean sea surface, together with a tidal model tuned for coastal accuracy (Tierny *et al.*, 2000), were used to generate ocean ground-truth surfaces. These GCPs were used for instrument calibration and assessment of the residual errors. GCPs were generated with a separation of 10 km for all the SRTM ocean data, which extended for at least 90 km from the coast.

In addition to previous data sets, NGA provided approximately 70,000 GCPs which had been obtained using various techniques over the years. Over 10,000 GCPs obtained from the JPL automated GPS processing database were also added to the GCP database. Finally, over 6,000 GCPs were derived from the GPS transects. The total number of GCPs available was 86,774.

The use of this GCP database for verification was problematic at times. The distribution of GCPs was non-random with the majority of GCPs densely packed in a small number of geographic areas. And while the majority of GCPs were in reasonable agreement with SRTM and other DEM sources, there was a small, but significant, percentage of GCPs that were obviously in error by many tens of meters (over 100 meters in some cases). Unfortunately, there was no useful error assessment on individual GCPs. Given this lack of reliability and the vast quantity of other verification data sets, the analysis presented below does not dwell in great detail on the differences between this data set and the SRTM data set. Investigations by NGA, subsequent to the analysis in this report, have shown that in all cases of major discrepancy, the GCP data were in error, rather than SRTM [J. Slater, personal communication, 2005].

All of the data comparisons presented below were done using heights located relative to the WGS84 ellipsoid. Although the final archived data product contains heights relative to the EGM96 geoid, to minimize datum and geoid induced errors, we used the raw JPL height product which reports heights relative to the WGS84 ellipsoid. All data used in the comparisons were translated to this datum, given the available metadata. It is possible, especially on the case of older GCP data, that the assumed datums may have been in error resulting in some of the isolated outliers.

SRTM Error Sources

The error characteristics for interferometric SARs are well understood and have been summarized in the open literature in (Rodriguez and Martin, 1992; Rosen *et al.*, 2000). Rosen *et al.* (2000) also reviewed interferometric SAR up to the year 2000, and the reader is referred there for greater detail on the technique. A detailed description of all the error sources that were present for the SRTM mission is presented in (Rodriguez *et al.*, 2005). The dominant error sources are summarized below.

Interferometric errors can be divided into static and time-varying errors. Static errors are those that can be regarded as having been constant over the data collection. Since these errors are constant, they can be calibrated by means of natural or man-made targets with know position and height. Time varying errors are due to motion of the interferometric mast and changes in the interferometric phase due to changes in the electronic beam steering. These errors can be partially compensated by dynamic calibration and least squares adjustments using GCPs, as described below.

Contributors to the SRTM error

Baseline Roll Errors

An error in knowledge of the baseline roll angle will induce a cross-track slope error in the estimated topography whose magnitude is equal to the roll error. The SRTM instrument used a sophisticated metrology system (AODA) coupled with post-flight filtering and estimation of the baseline position. The main components of the baseline motion are due to the natural modes of oscillation of the mast. These motions can be modeled and removed so that they do not constitute a dominant error source. In addition, the baseline position is affected by the Shuttle's attitude maneuvers. The time scale for residual roll errors is long, resulting in spatial errors with wavelength on the order of thousands of kilometers. They constitute the primary source for long-wavelength residual errors. Rodriguez *et al.* (2005) show a representative example of this residual long-wavelength error estimated by subtracting the sea surface topography, which is known to centimetric accuracy from the SRTM estimated topography. The peak values of this residual error are ± 10 m.

Phase Errors

These errors are due to two sources: random thermal or differential speckle noise and systematic phase changes due to antenna pattern mismatches or long term drift of the instrument electronics. The noise contamination results in height errors which are random and which exhibit short spatial correlation lengths. These errors cannot be compensated during ground processing. The antenna patterns for each of the channels do not have identical far-field phase characteristics. If uncompensated, this phase mismatch results in a net systematic phase error (called the phase screen) which is a function of the look angle. Due to the stability of the antenna far-field phase, this phase screen can be estimated by binning the height errors over the ocean as a function of look angle and applying the mean phase bias as a function of angle as a phase correction to the interferogram.

To estimate this phase screen, SRTM collected data prior and after each continental crossing, as well as for a small number of basin-wide ocean data takes. Estimates of the phase screen were obtained as a function of time and for each of the four SRTM elevation beams and all beam positions. The phase screen correction was not observed to change significantly over the mission lifetime; the estimated changes in the height error correction were estimated to be below 10 cm.

Finally, a small slow drift of the differential phase was observed over the lifetime of the mission, probably due to slow changes in instrument temperature. The residual phase errors induce cross-track tilts which are practically indistinguishable from the residual roll errors discussed above.

Beam Differential Errors

The SRTM swath consisted of four subswaths, each produced by a different SRTM elevation beam steering position. Systematic phase differences between the SRTM beams induce height differences at the beam overlaps. These differences can be time-varying since the beam steering angles are changed according to the topography to preserve the swath. This dynamic effect is corrected during ground processing by using the beam overlap height differences to estimate beam phase offsets. Since the instrument was designed so that all beams overlapped continuously, a time varying along-track phase calibration could be estimated and applied to the data. This ensured that all beams had a consistent phase error. The residual error due to this phase mismatch was negligible (below 10 cm) after along-track calibration.

Timing and Position Errors

These errors are induced by uncompensated delays in the system or errors in the estimated baseline position and result in geolocation errors. They are calibrated by using targets whose position is known and which can be identified in the radar image or topography, such as corner reflectors or kinematic GPS transects. The calibration procedure and geolocation are discussed in the Geolocation Errors Section.

Error Compensation

Along-Track Calibration

Estimate the beam-to-beam height discontinuities, as described above, and make corrections which vary with orbit position. This process insures the consistency of the SRTM DEM over each 220 km swath.

Dynamic Calibration
Use the known ocean topography (including tidal effects) to estimate static phase screens and time changing residual phase and roll errors and to generate ocean GCPs. Use these GCPs to constrain the least squares adjustment.

Mosaic or Least Squares Adjustment
Use the calibration data, together with other GCPs (including a subset of the kinematic GPS GCPs) and pass-to-pass tie-points to perform a continent scale-weighted least squares adjustment to reduce residual height errors. The mosaic correction estimates system phase (or, equivalently, roll) errors, and corrects the data in height and position given the estimated phase errors. The adjustment uses a sensor error model to constrain the geometrical characteristics of the error. In practice, the interferometric errors can be reduced to effective roll and range errors. The temporal evolution of these errors is constrained by requiring continuity at fitting region boundaries, *a priori* constraints at the ocean crossings, and simple polynomial evolution of the errors between boundaries. The adjustment implies that residual errors will have some correlation over long scales. Due to the lack of an extensive GCP database in continent interiors, the errors are best constrained at the oceans and less constrained in the interiors.

Rodriguez *et al.* (2005) showed the errors after static calibration, but prior to least-squares adjustments. The dominant errors are swath-to-swath differences, which are evident from the obvious swath patterns in the data. After calibration adjustments most swath discontinuities have disappeared, and the overall error variance has been reduced, although small residual long wavelength errors remain.

Absolute Height Error

In this section, we examine the difference between the SRTM height estimates and those from other sources. No attempt is made to co-register the data sets planimetrically. The spatial characteristics of the errors will be examined in the next section.

The most accurate measure of the height performance of SRTM over typical land surfaces is given by comparing with the kinematic GPS transects. Table 3 summarizes the performance, while Figure 1 shows the histograms and cumulative distribution function for the height error and its magnitude.

Examination of these results shows a very consistent behavior for the means, standard deviations, and absolute errors. The main exception is New Zealand, which has a much larger standard deviation and 90 percent error. The discrepancy may be due to the rugged terrain encountered by the New Zealand transect. Note that the main drawback of the kinematic GPS data is that, since it is limited to roads, it will selectively avoid very rugged terrain. Thus, the estimates presented in Table 3 may be optimistic for rough regions.

Table 4 summarizes the continental average results for the DTED® 2 comparisons, while Figure 2 summarizes the global performance in a histogram. The results for all the continents are roughly consistent, with the exception of South America, which shows a significantly higher mean deviation.

In order to investigate the source of large 90 percent errors, we plot in Plate 3 the behavior of the 90 percent error against the median and standard deviation of the height difference. The largest differences are due to mean shifts between the two data sets. Since the SRTM data is continentally adjusted against the ocean, and because the kinematic GPS data did not show these large deviations, it is possible that the average errors may be present in the DTED® 2 data, rather than the SRTM data.

On the other hand, Plate 3 shows that large errors can also occur due to large standard deviations. These may be due to degraded SRTM performance over regions of large slopes.

Table 5 summarizes the patch comparison results. The histograms, given in Rodriguez *et al.* (2005), are similar to the DTED® 2 shown in Figure 2. The results show no consistent mean shift, but the 90 percent error is larger than that for either the kinematic GPS or the DTED® 2 comparisons.

To examine the additional error source, Plate 4 presents the behavior of the 90 percent error as a function of the median difference and the standard deviation. These results show that a large part of the differences can be attributed to mean shifts between the two data sets. These shifts tend to average out over a continent, indicating that the variability may be due to errors in the mean height for selected patches.

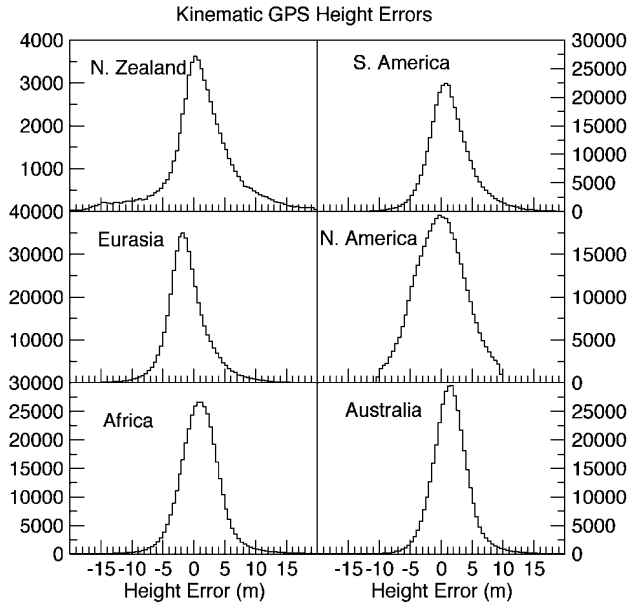

Figure 1. Histograms of SRTM minus kinematic GPS height differences averaged over 0.5° of latitude and longitude.

TABLE 3. SUMMARY OF KINEMATIC GPS GCP COMPARISON WITH SRTM DATA. ALL QUANTITIES ARE IN METERS

Continent	Mean	Standard Deviation	90% Absolute Error
Africa	1.3	3.8	6.0
Australia	1.8	3.5	6.0
Eurasia	−0.7	3.7	6.6
North America	0.1	4.0	6.5
New Zealand	1.4	5.9	10.0
South America	1.7	4.1	7.5

TABLE 4. STATISTICS FOR THE HEIGHT DIFFERENCE BETWEEN SRTM AND DTED© LEVEL 2. ALL QUANTITIES ARE IN METERS.

Continent	Average	Median	90% Diff	STD	RMS
Africa	2.44	2.30	8.80	4.68	5.53
Eurasia	−0.07	−0.09	8.07	4.50	5.36
N. America	1.38	1.40	7.67	4.01	4.88
S. America	11.36	11.20	16.88	4.60	12.77

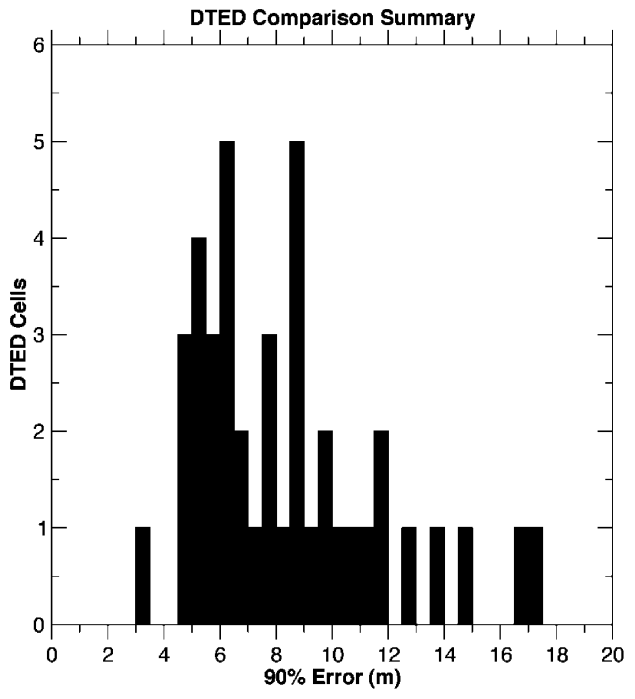

Figure 2. Histogram of the 90 percent error values for the comparison of the SRTM height data against the DTED® Level-2 data.

TABLE 5. STATISTICS FOR THE HEIGHT DIFFERENCE BETWEEN SRTM AND HEIGHT PATCHES. ALL QUANTITIES ARE IN METERS

Continent	Min	Max	Average	Median	90% Diff	STD	RMS
Africa	−39.79	42.13	4.13	3.31	10.15	4.26	7.63
Australia	−25.28	21.73	2.10	2.25	7.67	3.44	5.04
Eurasia	−40.17	36.44	−1.79	−1.36	12.48	5.99	8.28
N. America	−40.01	38.79	−0.61	−0.54	9.99	4.86	6.74

Even though the mosaic procedure uses the ocean and land GCPs for error reduction, the number of parameters estimated by the least squares adjustment (on the order of thousands) is much smaller than the number of ground control points (on the order of millions). Therefore, the post-fit residuals of the GCP data are still a valid measure of the instrument performance. In general, the ocean is darker than land, and we can expect higher random noise in the residuals. On the other hand, the slopes for the ocean are negligible, and slope induced errors will be minimized. The land GCPs used for this data set include a subset of the kinematic GPS GCPs, as well as a specially collected set of GCPs over Afghanistan, which is a region with extreme topography.

Table 6 summarizes the performance for each continent and surrounding ocean, while Figure 3 presents the height error histograms. Notice that these results are consistent with those obtained using the kinematic GPS transects. For most of the data, the probability density function is well characterized by a Gaussian distribution, but for some continents (e.g., South America) a number of outliers exist which significantly deviate from a Gaussian distribution in the tails of the distribution (due to the large number of points in the histogram, this is not easily seen in Figure 3).

To examine further the deviation from Gaussian behavior, we separate the land only GCPs and re-compute the histograms. The summary results are presented in Table 7 and in Figure 4. These results show that most of the outliers in the distribution do indeed come from the land GCPs. There is no mean shift in the bias, however, and the increase in 90 percent height error is due mainly to the increased proportion of outliers. The height error performance is still superior to that predicted by the height patches and by some of the DTED® 2 cells.

For all of the data sets examined, the SRTM data meets and exceeds the 16 m (90 percent) performance goal, often by a factor of 2. There is a broad agreement between all the

TABLE 6. SUMMARY OF LAND AND OCEAN GCP COMPARISON WITH SRTM DATA. ALL QUANTITIES ARE IN METERS

Continent	Mean	Standard Deviation	90% Absolute Error
Africa	0.4	4.8	7.5
Australia	0.1	4.4	6.5
Eurasia	0.2	5.0	7.0
North America	−0.2	4.6	6.5
South America	0.0	5.1	6.5

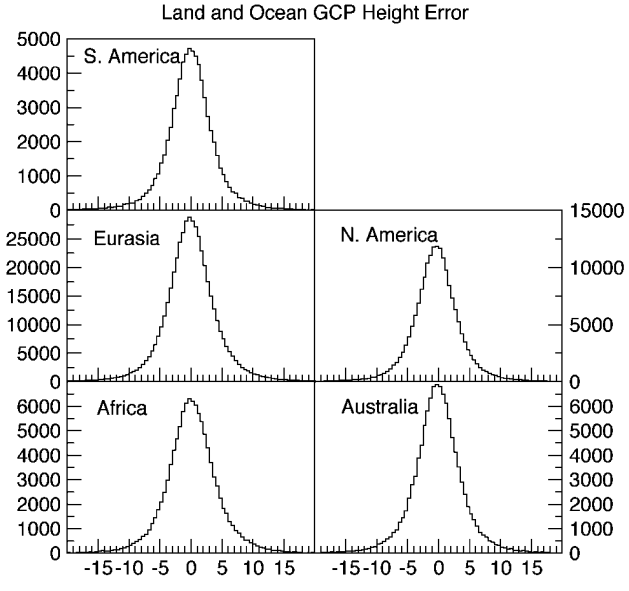

Figure 3. Histograms of SRTM minus ocean and land GCP height differences averaged over 0.5° of latitude and longitude. The histograms are approximately Gaussian. Although some outliers are present, they are too few to be seen in these histograms.

TABLE 7. SUMMARY OF LAND GCP COMPARISON WITH SRTM DATA. ALL QUANTITIES ARE IN METERS

Continent	Mean	Standard Deviation	90% Absolute Error
Africa	1.0	5.4	8.0
Australia	0.7	6.3	7.5
Eurasia	0.5	6.1	8.5
North America	−0.8	8.3	8.5
South America	−0.5	9.6	8.5

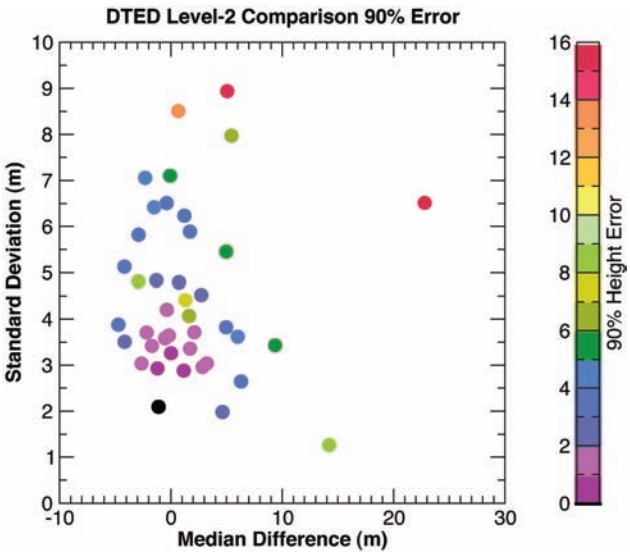

Plate 3. 90 percent Height errors plotted against height difference median and standard deviations for the SRTM to DTED® Level-2 comparisons. Note the largest differences are due to a mean shift between the two data sets.

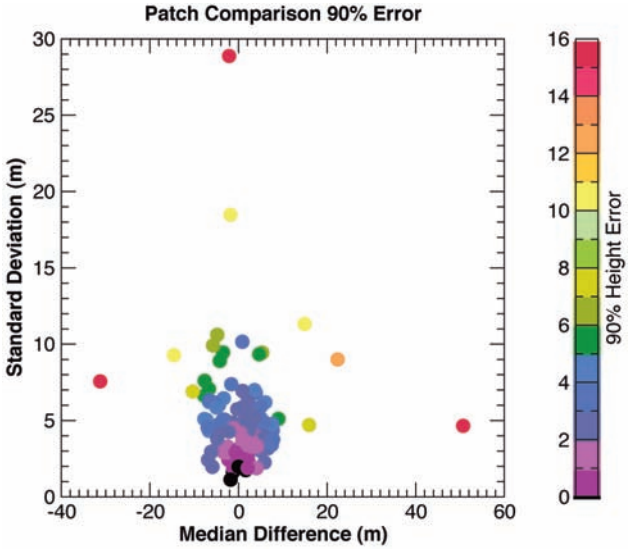

Plate 4. 90 percent Height errors plotted against height difference median and standard deviations for the SRTM to height patch comparisons.

ground truth data sources, but the largest deviations are exhibited by the height patches and some of the DTED® 2 patches and land GCPs. Given the overall consistency in the mean error shown by the SRTM data, it is probable (J. Slater, NGA, private communication) that the mean shifts between the data sets are not due to SRTM.

Spatial Characteristics of the Error

In order to separate long wavelength errors from topographic errors, we compute the average of the difference between the SRTM DEM and the kinematic GPS data using boxcar low-pass filters (i.e., for any given latitude and longitude, all the differences of SRTM height minus the KGPS height within degrees of latitude or longitude of the desired point are averaged). The resulting global average errors are shown in Plate 1.

Examination of this Plate 1 shows that the average errors have coherent long wavelength signatures that can vary significantly over a continent. These errors are consistent with residual motion errors of the interferometric baseline. There seems to no apparent correlation of the error to the underlying topography, or a preferred scale or direction of the error which persists across continents. The fact that the kinematic GPS data are essentially unbiased relative to SRTM shows that the mosaic has removed the mean long wavelength error over the continent. The magnitude of the error is on the order of approximately 2 m.

To further visualize the nature of the long wavelength errors, we apply the same boxcar low-pass smoothing filter to the land and ocean GCP residuals after the mosaic corrections. The results are presented in Plate 2, which shows that there are no detectable residual tilts over the continents, and, though there are still some detectable residual long wavelength modulations over the water, their magnitude is small. The magnitude of the long wavelength error then grows as the distance from the coast increases. The preeminent example of this behavior is the residual height error for GCPs in Central Asia, which can be seen in Plate 2. One also notices in these figures that GCP outliers (black) tend to occur in isolated points, indicating probable errors in the ground-truth data.

To further study the spatial characteristics of the error, it is useful to introduce the "structure function" as a tool for examining the spatial dependence of the data. This function is defined as

$$D(\Delta) = \langle(\delta h(x) - \delta h(x + \Delta))^2\rangle \quad (1)$$

where $\delta h(x)$ is the height error as a function of position, Δ is the spatial separation between two points, and angular brackets denote ensemble averaging. The structure function can be viewed as an estimate of the relative height error between two points, as a function of the separation between them. For homogeneous statistical processes, it can be shown (Goodman, 1985) that the structure function can be related to the height error correlation function, $C(\Delta)$, by

$$C(\Delta) = 1 - \frac{D(\Delta)}{2\sigma^2} \quad (2)$$

where σ is the standard deviation of the process (which is constant in space). Thus, if the SRTM error could be viewed as a single random process with a given correlation length, we would expect that the structure function would start at zero, increase with distance until finally achieving an asymptotic value for distances much greater than the correlation length.

In Figure 5, we present the computed structure function for the height error along the kinematic GPS tracks. These results show that, when averaged over an entire continent, the structure function can either approach an asymptotic value (South America), show a slightly increasing trend with separation (North America, Australia), or show non-stationary behavior (Africa and Eurasia). These results show the presence of long wavelength errors that can be coherent over continental scales, as one would expect for residual motion errors. On the other hand, even though fluctuations and long wavelength correlations exist, the magnitude of the relative RMS error is almost always smaller than 6 m, which corresponds to a relative 90 percent height error of about 10 m.

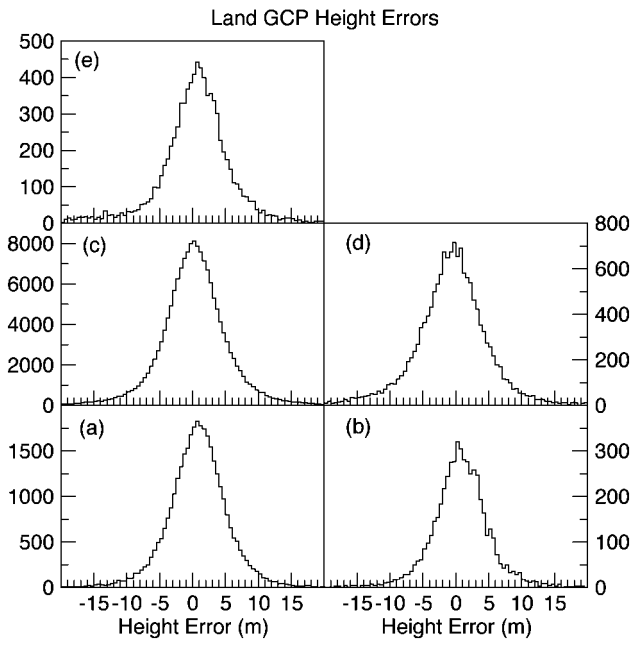

Figure 4. Histograms of SRTM minus land GCP height differences averaged over 0.5° of latitude and longitude. Again, the histograms are approximately Gaussian, but in this case the presence of outliers can be seen more readily.

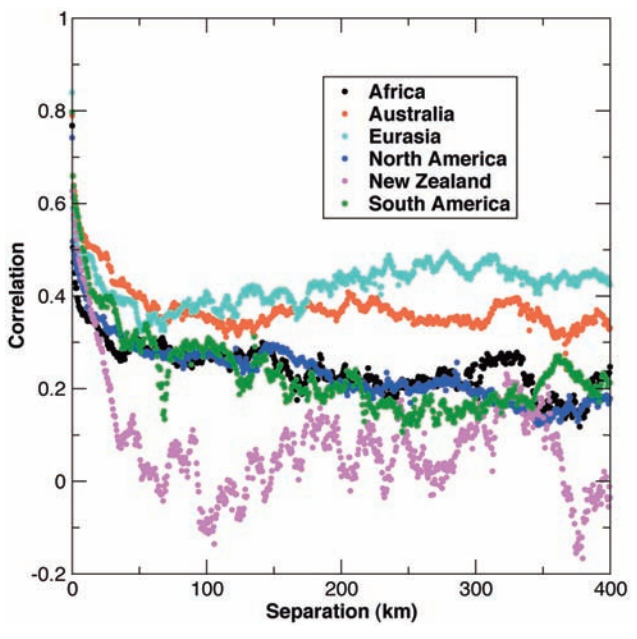

Plate 5. Correlation function for all continents over the scale of two SRTM swaths. Notice that the correlation drops quickly in all cases (see Plate 6), but there is no consistent behavior across the different continents. This is probably due to different long wavelength error characteristics associated with the platform roll knowledge error and differences in the distribution of continental topography.

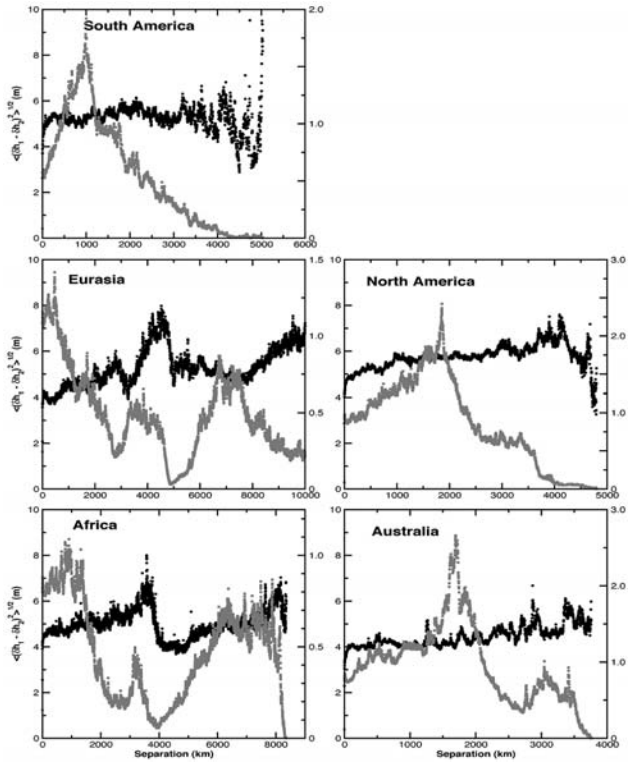

Figure 5. Structure function for the Kinematic GPS GCP errors (black, left axis) and number of points (divided by 1×10^6) used in the estimate (grey, right axis). The RMS height errors computed using all points are: Africa: 3.8 m; Australia: 3.5 m; Eurasia: 3.7 m; North America: 4.0 m; South America: 4.1 m.

To study the short wavelength behavior of the error, we compute the height error correlation function using Equation 2. Plate 5 presents the behavior of the correlation function on separation scales from 0 to 400 km, or about twice the SRTM swath. The generic behavior is the same for all continents: a rapid decrease from a peak value followed by a plateau of correlation. Three continents (South America, North America, and Africa) exhibit almost identical behavior, although their structure function differs at larger scales. Australia and Eurasia show significantly higher correlation plateaus. This is probably due to the fact that, since the kinematic GPS tracks were collected over flatter regions, the long wavelength error component makes up a larger fraction of the total error compared to regions that have greater topography. The results over New Zealand can be similarly explained as being caused by large topographic variations. However, more work needs to be done to verify these conjectures.

All of the correlation functions exhibit a very similar behavior near the origin. Plate 6 presents a zoom of the previous figure around the origin. It is clear from this figure that in all cases, the correlation drops from its peak value of 1 at the origin to the plateau region for separations greater than approximately 100 m to 200 m, while significant correlations are still apparent for separations of 50 m. This is consistent with the SRTM spatial averaging scheme, which applied a three-pixel boxcar filter (approximately 45 m correlation length) to reduce speckle noise.

Although the previous figures present the average behavior for a continent, it should be remembered that the

short-wavelength behavior within a continent will show variability reflecting the local random errors. Figure 6 shows the variability in the short wavelength structure function for nearby kinematic GPS tracks which were processed as part of the same least squares adjustment. As can be seen, significant variability in the short wavelength behavior is present due to changes in the local conditions.

We conclude that the SRTM error can be thought of as consisting of three parts: first, a long wavelength component, due to residual roll errors, with a magnitude of about 2 m and a spatially non-stationary behavior; the second component consists of random (i.e., medium to short wavelength) errors that add an additional spatially varying error component and finally, at the shortest scales, speckle noise decorrelates for separations on the order of one to two pixels. The next section examines the nature of the random error in greater detail.

Random Error

Random errors (i.e., zero-mean errors with very short spatial correlation lengths) in SRTM data are caused by instrument thermal noise and, after filtering, residual geometric decorrelation effects (Rodriguez and Martin, 1992; Rosen *et al.*, 2000).

Estimating the random terrain height error directly from the SRTM data required determination of the channel-to-channel correlation, γ, which is estimated from the radar interferogram. The interferometric phase error is then computed from

$$\delta\Phi = \frac{1}{\sqrt{2N_L}} \frac{\sqrt{1-\gamma^2}}{\gamma} \tag{3}$$

where N_L is the number of independent radar realizations, or "looks." The height error can then be obtained from the phase error by using the height-phase sensitivity equation (Rodriguez and Martin, 1992; Rosen *et al.*, 2000). The theoretical estimates obtained by using this process were made into the "Terrain Height Error Data" (THED), which is SRTM NGA data product (currently available to selected investigators).

To validate the theoretical predictions we estimated the true, high frequency, random error by high-pass filtering the difference between known topography and the SRTM estimate. The high-pass filter was implemented by selecting a box size, removing a mean error plane by comparing against DTED® Level-2 ground truth, and computing the root mean squared (RMS) of the residual. The selection of the optimal box size is a compromise: if the box is too small, the error will be underestimated; if too large, topography will alias into the estimated error. After some experimentation, it was found that an average box size of 11 pixels (about 330 m) represented a suitable compromise between these two effects.

Plate 7 shows an example of the estimated random error for the topography around Twentynine Palms, California, shown in Plate 8. The greatest errors are associated with regions of higher relief, but also with regions of reduced radar brightness. An image of the THED data for the same region (Rodriguez *et al.*, 2005) shows a high degree of correlation with this image, although the THED seems to underestimate the measured errors for large errors (or low correlation).

In order to formalize this relationship, we used 22 of the DTED® 2 cells and repeated the procedure outlined above. The results of the comparisons are shown in Figure 7. There is a tight correlation between the two data sets, with a small deviation at the higher error values. This underestimation at larger errors has been confirmed by a more recent study (Salamonowicz, 2005) which showed that the deviation from

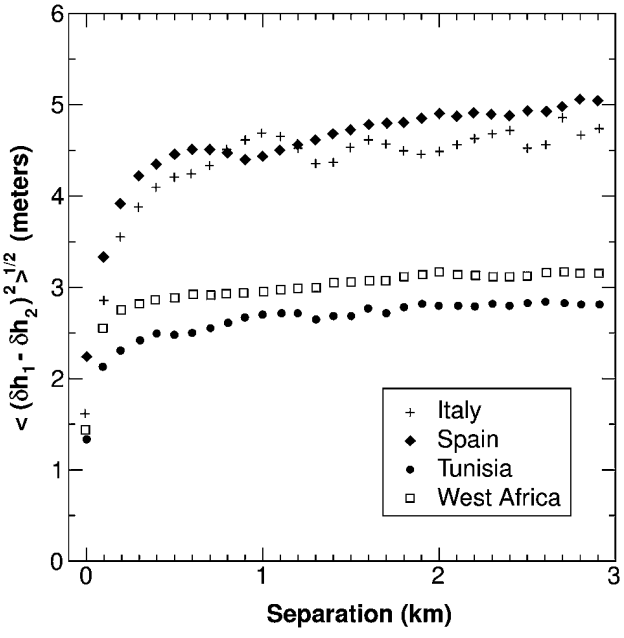

Figure 6. Variability of the structure function for different kinematic GPS tracks. Notice that the behavior is common for distances less than about 100 meters, but can vary substantially after that reflection the non-stationarity of the errors.

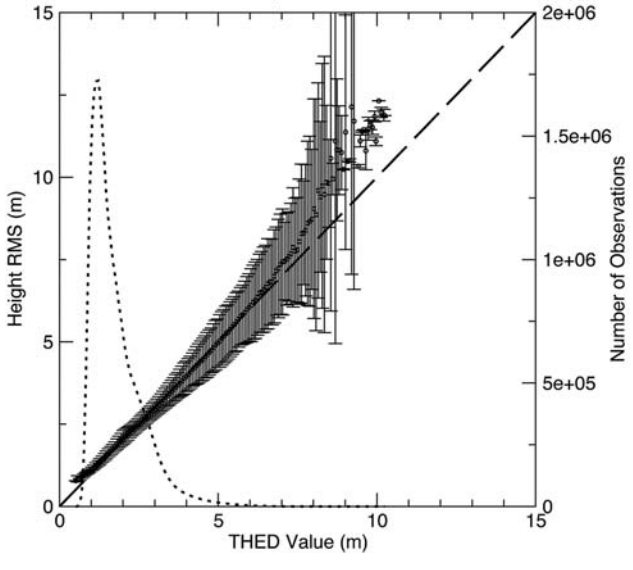

Figure 7. Statistical relationship between the THED values and the estimated random error. The short-dashed curve represents the (arbitrarily normalized) histogram of the values used to derive the relationship. Notice the excellent agreement for most of the data range. A slight underestimate of the validation error estimates can be seen for large error values.

the predicted THED values continues to increase at very high errors. This underestimate of the error may be due to the high-pass filtering process aliasing or biases in the estimation of correlation for low correlation values. However, it must be stressed that for the vast majority of the SRTM data, the correlation is quite good and the THED values are a reasonable estimate of the random error. The results presented in Figure 7 validate the accuracy of the SRTM THED using globally distributed DTED® 2 data with about 31 million observations.

To visualize the global distribution of height errors, we assume that the random height errors at different height postings are independent, zero-mean, Gaussian distributed variables and estimate the cell height variance by averaging the estimated height variances over each cell. The 90 percent error is obtained by multiplying the resulting standard deviation by 1.64.

The resulting global assessment of the SRTM random errors is presented in Plate 9. This map shows how the increased number of looks which are obtained at higher latitudes reduces the random error. Because the number of looks varied at lower latitudes, a crosshatched pattern is seen in some areas. This again reflects the fact that areas with additional data will have lower random errors.

The map also shows the relationship between topography and random error (Himalayas, Andes, Rockies, etc.) and between radar brightness and random error (Sahara, Saudi Arabia). Except in areas of high relief or low cross section, the error is typically <5 m. At higher latitudes or for flat regions with good coverage, the magnitude of the random error can be <2 m (90 percent). In the areas of highest error, values are on the order of 10 m.

Geolocation Errors

We evaluated the SRTM geolocation (i.e., position) errors using both radar corner reflectors and kinematic GPS tracks.

Radar corner reflectors, which can be identified in the radar image, were used to estimate and correct for the system delay, a single parameter for each of the four SRTM beams; these were not changed during the processing of the data. Errors in system timing lead to a positioning error along the radar look direction and thus to a geolocation error (as well as a height error) (Rosen *et al.*, 2000).

Since all the corner reflector data were used for the calibration of only four parameters, one can examine the geolocation errors of the entire corner reflector data set to assess the stability of the system timing over the lifetime of the mission. The corner reflectors were surveyed to an accuracy of 10 cm (absolute) using GPS-based surveying. Thus, any difference between the measured corner reflector range and the range calculated based on orbit and surveyed corner reflector position can be related to geolocation errors in the instrument look direction.

For the corner reflector calibration/validation experiment, we deployed standard 1.8 meter (6-foot) and 2.4 meter (8-foot) trihedral corner reflectors at calibration sites in California and the Northern Territory, Australia for the SRTM mission. The total extent of the corner reflector positions extended over 400 km for both sites, allowing for the calibration and validation of all SRTM beams for at least two independent passes (the SRTM single pass swath was about 200 km). Both sets of reflectors were observed by ascending and descending passes.

The corner reflector positions were extracted from the single-look complex (SLC) data products using a 32-fold oversampling. The optimum common range delays were determined for each beam by measuring range offsets for each reflector at five time delays and performing a linear fit to estimate the optimum delay. The standard deviation of the computed range offsets for each beam, which represents the range variability of the corner reflector positions, is given in Table 8. The average value is 1.3 m, which corresponds to a maximum geolocation error in the range direction of less than one meter.

After correcting each beam for an optimal constant range offset, a linear fit of the range offset versus orbit gives a derived slope is -5.9 ± 4.2 mm/orbit, which indicates very little change over the mission. Dividing the measurements into descending and ascending orbits (which also represent early and late times in the mission) gives the time-dependence change of -0.37 ± 0.32 m, as shown in Table 9. Dividing the corner reflectors into Northern Hemisphere (California) and Southern Hemisphere (Australia) yields no systematic differences.

In conclusion, the common range delays, estimated from the corner reflector positions, are well understood and stable. The geolocation errors resulting from errors in these delays are small (typically less than two meters).

The corner reflector data, although very precise, was geographically limited. To extend the assessment of the positioning accuracy of SRTM, we examine the kinematic GPS collected along roads that could be identified in the radar imagery. To assess the geolocation accuracy accuracy of the SRTM, we compared both the shape of the GPS track against the image of the roads along which the data were taken, and the topographic heights measured by the kinematic GPS and SRTM, as described below.

Independent estimates of geolocation errors using topography were derived by minimizing the RMS height difference between kinematic GPS topography to the SRTM DEM, on a 1-degree cell basis. Each GPS track was shifted from its nominal position over a grid of ± 2 arc-seconds in 0.1 arc-second steps, in both the North/South, East/West directions. For each grid point, the standard deviation of the height error (i.e., the difference between the 32-oversampled SRTM heights and the GPS heights) was recorded. The best-fit geolocation error was determined by the minimum of the surface over the grid.

The image geolocation validation was based on the assumption that roads are darker than their surroundings. The image brightness, at a GPS point, was defined by a cubic convolution interpolation of the image at the latitude and longitude of the GPS data point. Data were analyzed on a one-degree cell basis. In order to avoid contamination of the

TABLE 8. STATISTICS FOR THE ESTIMATED SRTM RANGE OFFSETS FOR EACH OF THE SRTM BEAMS

Beam	No. of GCPs	(δR) (m)	σ (m)
1	27	-0.11 ± 0.23	1.2
2	22	-0.24 ± 0.23	1.1
3	21	$+0.31 \pm 0.39$	1.8
4	10	$+0.16 \pm 0.49$	1.6

TABLE 9. TEMPORAL AND SPATIAL DEPENDENCE OF RANGE OFFSETS

Data Qualifier	No. of GCPs	Mean Range Offset (m)
Descending (1st half of mission)	39	$+0.21 \pm 0.27$
Ascending (2nd half of mission)	31	-0.16 ± 0.17
California	43	-0.02 ± 0.20
Australia	27	$+0.09 \pm 0.22$

measured mean brightness by bright objects such as buildings or vehicles, any data point within three pixels of a point brighter than 8 times the median brightness was rejected.

The entire GPS track over a 1-degree cell was shifted from its nominal position over a grid of ±2 arcseconds in 0.1 arcsecond steps, in both the North/South, East/West directions. For each grid point, the mean brightness of the interpolated image was recorded. The best-fit geolocation error was determined by the minimum of the surface over the grid.

For each continent, the individual 1 degree by 1 degree cells were combined into separate 10 degree by 10 degree super-cells in which mean geolocation errors were calculated from a weighted average of the aggregate cells. This was done for the ascending image, the descending image, and the height (DEM) analysis independently.

Finally, the combined bias estimates were constructed by forming the weighted average of the ascending image, descending image, and height analysis. An example of the combined results is shown in the Figure 8, which exhibits the results for North America; other continents show a similar behavior, as can be seen in Rodriguez *et al.* (2005). The summary of the horizontal accuracy for each continent appears in Table 10. These values were derived from the weighted average of the combined results (black points in Figure 8), and is denoted by the red point in the sample figure.

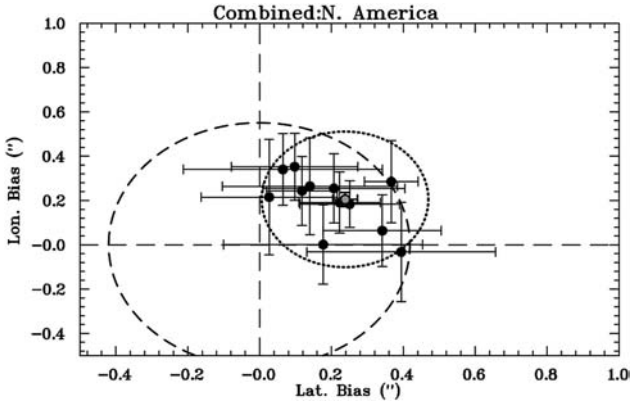

Figure 8. Best fit geolocation biases using combined height and image data for 10 by 10 super-cells in North America. The 90 percent absolute (relative) limits are represented by the long-dash (short-dash) circles.

TABLE 10. SUMMARY OF SRTM HEIGHT PERFORMANCE. ALL QUANTITIES REPRESENT 90% ERRORS IN METERS

	Africa	Australia	Eurasia	Islands	N. America	S. America
Absolute Geolocation Error	11.9	7.2	8.8	9.0	12.6	9.0
Absolute Height Error	5.6	6.0	6.2	8.0	9.0	6.2
Relative Height Error	9.8	4.7	8.7	6.2	7.0	5.5
Long Wavelength Height Error	3.1	6.0	2.6	3.7	4.0	4.0

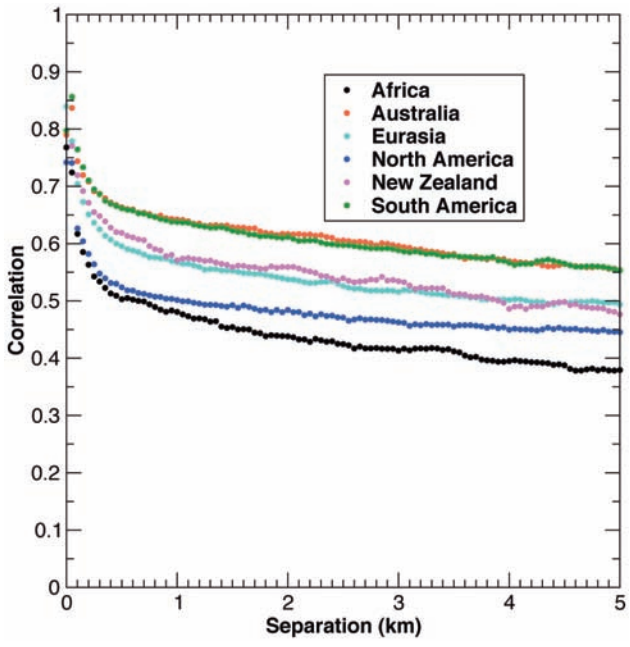

Plate 6. Zoom on the correlation function presented in Plate 5 for small distances from the origin. Notice that the correlation drops significantly within the first 100 meters, consistent with the assessment of the spatial resolution capabilities of SRTM.

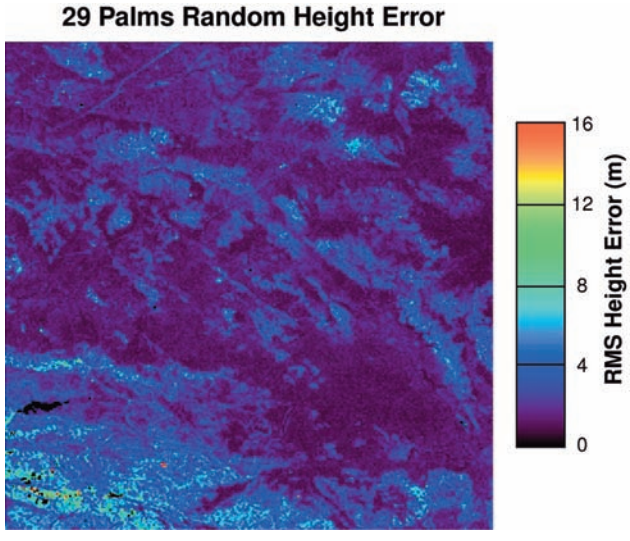

Plate 7. Height RMS calculated for the topography shown in Plate 8 (Twentynine Palms, California) using an 11 pixel averaging window. This is an estimate of the height random error.

In addition to the error sources listed above, Table 10 lists the "relative height error." This error is defined as the expected magnitude of the height difference between two points in a 1-degree cell. Since within this cell one can assume that long wavelength errors are essentially constant, the variance of the height error difference will be the sum of the individual random error variances. The relative error

over the cell is estimated by averaging this quantity and assuming that the random errors follow Gaussian statistics.

Performance Summary

We have presented an assessment of the SRTM topographic data using a variety of globally distributed data sets. The results of this assessment show that the SRTM has an absolute height error which exceeds the mission goal of 16 m (90 percent) by almost a factor of two. An analysis of the spatial characteristics of the data show that the SRTM height errors can be divided into two components: the random error, due to measurement noise, is the dominant error source, and exhibits a very short correlation length (between 45 m and 90 m). This error source exhibits geographical variability due to changes in the surface brightness, the surface slope, and the number of observations. The long wavelength error, which has scales much greater than the SRTM swath (approximately 220 km), has a magnitude of about two meters and is uncorrelated with the random error. This error is due to residual errors in the estimation of the SRTM mast vector and small secular variations of the system phase. Although globally distributed, the validation data used, especially the kinematic GPS data, were confined to regions with moderate topography and low vegetation. However, the comparisons also included a substantial data set over the Hindu Kush, which presented extreme topography and was generally consistent with the other results.

As a final step, we provide a summary of the performance by estimating the continental scale absolute height error. To do this, we rely mainly on the kinematic GPS data set since, as described above, there are significant uncertainties associated with the other (non-ocean) GCPs, or high quality data may not be available for the continent.

The SRTM continent-wide absolute height error is estimated by subtracting the DEM height from the GPS ground-truth, after interpolating the DEM to the GPS point by cubic interpolation. The continent-wide height total error distribution is estimated from these data, and the continent-wide 90 percent relative and absolute errors derived directly from it. Table 10 presents the estimated total errors for each continent.

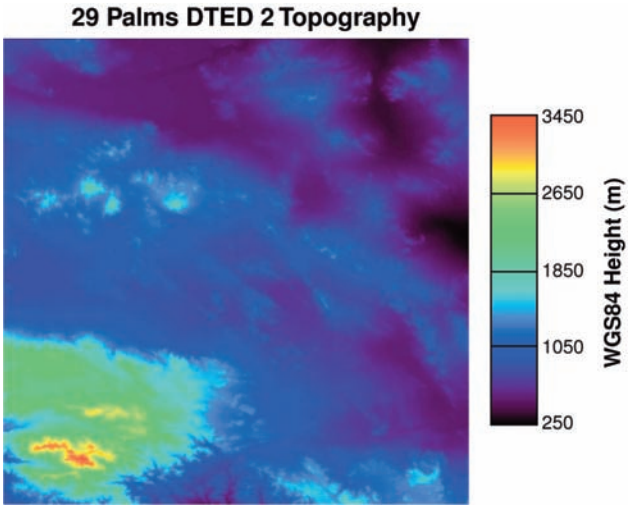

Plate 8. Topography for the Twentynine Palms DTED® 2 cell used to illustrate random errors. The area is sparsely vegetated, but has significant topography and some significant vegetation at higher elevations.

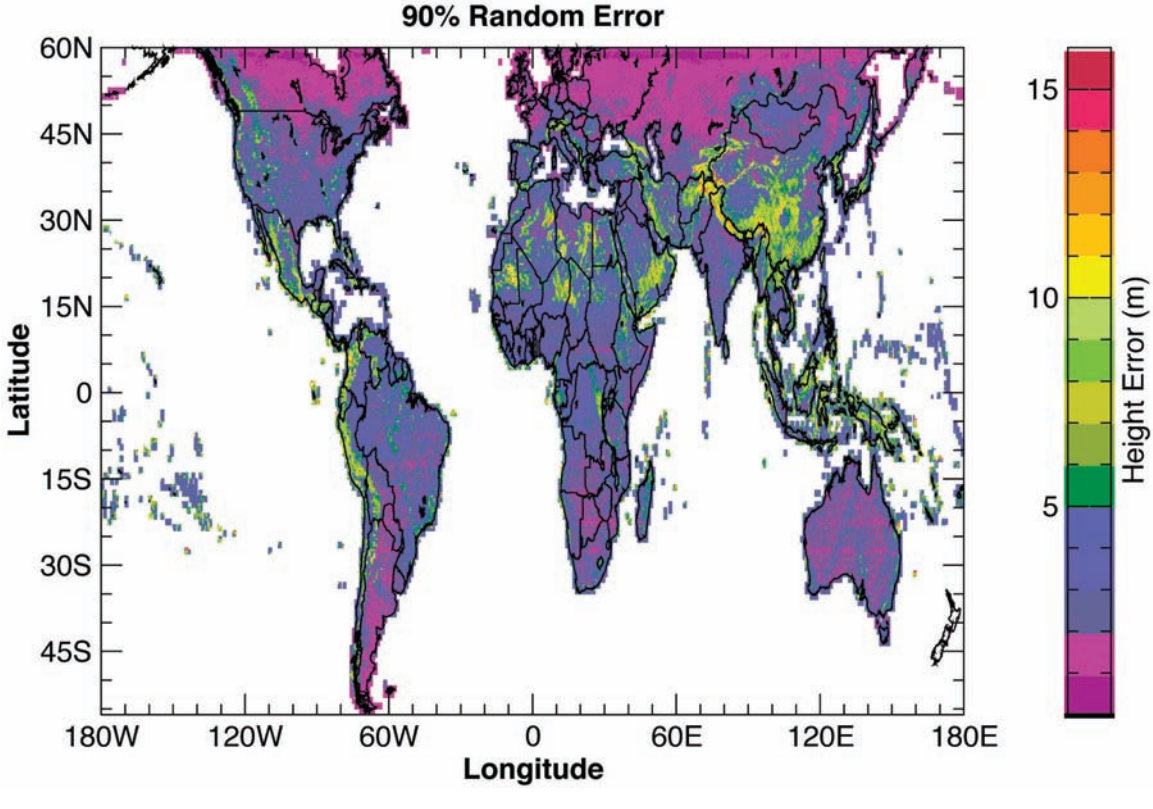

Plate 9. Global SRTM random error map at 1/8° resolution computed using the SRTM THED data product.

Since the continent-wide error contains both short and long wavelength error components, we attempt to separate the random error by averaging the kinematic GPS errors over scales which are expected to be larger than the correlation distance of the random error, which is mainly dependent on topography or surface brightness. Although there is no clear-cut separation between residual long wavelength error and surface topography, we select the SRTM product cell ($1° \times 1°$) as a convenient scale. The results for the 90 percent residual long wavelength error for each continent are presented in Table 10.

The geolocation error was estimated using image and height matching of the kinematic GPS transects. The details are presented in the Geolocation Errors section, and the results summarized in Table 10.

Since the random error is dependent on the topography and radar brightness, continent-wide averages are of limited usefulness. An estimate of the random error can be obtained from the interferometric correlation. This estimate is stored in the THED product, which is available for each height post.

In addition to the THED and the continent-wide error estimate, it is desirable to estimate the combined error characteristics on a cell ($1° \times 1°$) and sub-cell ($1/8° \times 1/8°$) basis. The results are shown in Plate 9. This information is stored in the SRTM height data headers.

If the long wavelength error is defined as that error which has scales greater than the a nominal cell size (such as, 110 km, or 1 degree at the Equator), an estimate of the energy contained in the long wavelength error can be obtained from the ground truth by averaging the observed error over scales smaller than the nominal cell size, and computing the variance of the residual. These results form the last entry to Table 10.

Acknowledgments

This work was conducted by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA). A data comparison of this magnitude is the work of many people across many continents, of whom we cannot acknowledge only a few since. In addition to the many field workers, we would like to give special thanks to AUSLIG, Anthony Milne (UNSW), Catherine Ticehurst (CSIRO), Alex Held (CSIRO), John Homer (U. Queensland) in Australia for their invaluable help in collecting ground truth data in adverse conditions. In Argentina, we would like to acknowledge the Instituto Geográfico Militar for their help in collecting Kinematic GPS data. In NGA, we would like to thank all the people who made the kinematic GPS data collections possible and provided us with ground truth data. In particular, we would like to thank Paul Salamonowicz, James Slater, and Barry Heady for their help and support. At Ohio State University, we would like to thank C.K. Shum and his team with their help regarding ocean tides. At JPL, we would like to thank our other coauthors in the companion report, and all the other members of the SRTM project, without whom none of this could have happened.

References

Farr, T., and M. Kobrick, 2000. Shuttle radar topography mission produces a wealth of data, *American Geophysical Union Eos*, 81:583–585.

Fu, L., and A. Cazenave (editors), 2001. *Satellite Altimetry and Earth Sciences*, Academic Press.

Goodman, J., 1985. *Statistical Optics*, New York, Wiley-Interscience.

Rodriguez, E., and J.M. Martin, 1992. Theory and design of interferometric synthetic aperture radars, *IEEE Proceedings-F*, 139(2):147–159.

Rodriguez, E., C.S. Morris, J. Belz, E. Chapin, J. Martin, W. Daffer, and S. Hensley, 2005. *An Assessment of the SRTM Topographic Products*, Technical Report JPL D-31639, Jet Propulsion Laboratory, Pasadena, California.

Rosen, P., S. Hensley, I. Joughin, F. Li, S. Madsen, E. Rodriguez, and R. Goldstein, 2000. Synthetic aperture radar interferometry, *Proceedings of the IEEE*, 88(3):333–382.

Salamonowicz, P., 2005. *A Comprehensive Assessment of the Shuttle Radar Topography Mission Elevation Data Accuracy*, USGS, The Shuttle Radar Topography Mission: Data Validation and Applications.

Smith, B., and D. Sandwell, 2003. Accuracy and resolution of shuttle radar topography mission Data, *Geophysical Research Letters*, 30(9):1467.

Sun, G., K. Ranson, V. Khairuk, and K. Kovacs, 2003. Validation of surface height from shuttle radar topography mission using shuttle laser altimeter, *Remote Sensing of Environment*, 88(4):401–411.

Tierny, C., L. Kantha, and G. Born, 2000. Shallow and deep water global ocean tides from altimetry and numerical modeling, *Journal of Geophysical Research*, 105(C5):11259–11277.

van Zyl, J., 2001. The shuttle radar topography mission (SRTM): A breakthrough in remote sensing of topography, *Acta Astronautica*, 48(5–12):559–565.