The SRTM Data “Finishing” Process and Products

James A. Slater, Graham Garvey, Carolyn Johnston, Jeffrey Haase, Barry Heady, George Kroenung, and James Little

Abstract

The Shuttle Radar Topography Mission (SRTM) successfully acquired terrain elevation data for 80 percent of the Earth’s landmass in February 2000. The radar system and data collection scheme designed by NASA’s Jet Propulsion Laboratory (JPL) met the global requirements of the U.S. Department of Defense for Level 2 Digital Terrain Elevation Data (DTED®). JPL processed the raw data into unfinished DTED® 2 and other products that were delivered to two contractors of the National Geospatial-Intelligence Agency. The contractors edited the unfinished DTED® 2, updated the associated products, and generated finished products for distribution. Automated processes were developed by each contractor to identify, delineate and set heights for lakes, rivers, and ocean coastlines in conformance with an extensive set of editing rules created to maintain consistency and uniformity in the final products. The finished DTED® is significantly better than the 16 m vertical accuracy required by the original specification.

Introduction

Topographic data and digital terrain models are critical components of any physical description of the Earth’s surface, and in conjunction with ellipsoid and geoid models, provide surface structures required for myriad applications. Mapping, navigation, military mission planning and simulation, search and rescue operations, agricultural planning, flood modeling, and orthorectification of satellite imagery are some of the uses of terrain data. Despite many years of data collection and production of Digital Terrain Elevation Data (DTED®), the National Geospatial-Intelligence Agency’s (NGA) predecessor agencies, the Defense Mapping Agency and the National Imagery and Mapping Agency, had managed to cover only a small percentage of the world with the medium-scale terrain elevation data required for many U.S. military applications. The U.S. Department of Defense requirement for global terrain elevation data was certified in 1995 and gave impetus to the data acquisition strategy that became the Shuttle Radar Topography Mission (SRTM) in February 2000.

The primary goal of the SRTM project was to produce DTED® Level 2 for the land area between 56 degrees South and 60 degrees North latitudes, constituting about 80 percent of the Earth’s landmass. The intent was to generate a uniform, self-consistent depiction of the Earth’s topography that would be common to all users and applications. The SRTM fulfilled this goal by acquiring all the data over one 10-day period with one sensor: a single-pass, dual-antenna, C-band interferometric synthetic aperture radar system. NASA’s Jet Propulsion Laboratory (JPL) designed the instrumentation and data collection scheme, and generated a set of *unfinished* products from the raw radar data using software developed at JPL (Farr and Kobrick, 2000). These products were forwarded to NGA contractors who developed standardized procedures for doing the final data editing, or *finishing*, required to bring the products into full DTED® compliance. The suite of finished products was provided to NGA for distribution. This paper focuses on the finishing process, beginning with the specific requirements for SRTM DTED® 2, outlining the editing rules and quality assurance procedures, and then discussing at some length the water body editing which was the most challenging aspect of the finishing process. The remainder of the paper presents statistics on the quality of the finished data.

Finishing Requirements

SRTM DTED® 2

SRTM DTED® 2 is defined as a uniform grid of elevation values, spaced at one-arcsecond (approximately 30 m) intervals in latitude and longitude between 50° N and 50° S latitudes. At latitudes between 50 and 60 degrees, the longitudinal grid spacing is two arcseconds, while the latitudinal spacing remains one arcsecond. DTED® are divided into 1° × 1° cells for processing, storage, and retrieval purposes. Elevations are rounded to the nearest integer meter and are referenced to mean sea level as defined by the WGS84 EGM96 geoid. SRTM elevations represent the elevations of the reflective surface (e.g., tree canopy, building roof, bare ground) for the radar return and have not been reduced to bare Earth. Specified accuracy requirements for the SRTM DTED® 2 were 16 m absolute vertical error (90 percent linear error, with respect to the reflective surface), 20 m absolute horizontal error (90 percent circular error) and 10 m relative vertical error (90 percent linear error). All the SRTM DTED® conform to the DTED® Performance Specification (NIMA, 2000).

The general DTED® specification requires the identification, delineation, and elevation determination of water bodies that meet minimum size criteria. Ocean coastlines must be defined and ocean elevations are set to 0 meters. Lakes greater than 600 m in length must be captured and

James A. Slater, Jeffrey Haase, Barry Heady, George Kroenung, and James Little are with the National Geospatial-Intelligence Agency (James.A.Slater@nga.mil).

Graham Garvey is with Intermap Technologies.

Carolyn Johnston is with Vexcel Corporation.

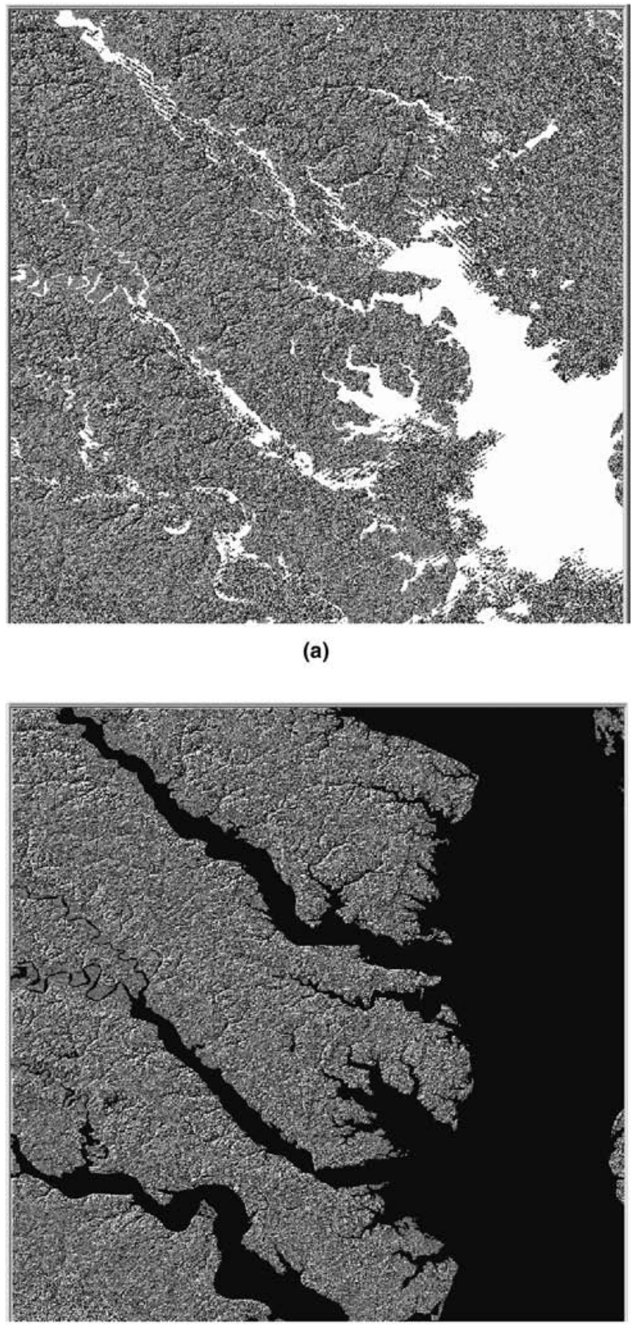
Photogrammetric Engineering & Remote Sensing
Vol. 72, No. 3, March 2006, pp. 237–247.

0099-1112/06/7203–0237/$3.00/0

flattened to a constant elevation. Double-line drains (rivers) greater than 183 m in width must be delineated and stepped down monotonically in elevation. Islands are delineated if greater than 300 m in length or, for smaller islands (down to 14,400 square meters), if at least 10 percent of the island's elevations are 15 m or more above the surrounding water. Shoreline elevations are set one meter higher than the adjacent water to ensure containment of the water. After the editing is complete, all adjacent cell edges must match for continuity.

Due to the unique characteristics of the SRTM data, additional finishing requirements were added to the general specifications for the DTED® 2 product. Anomalous points (spikes or wells) in the elevation data were removed and voided out if they exceeded the mean elevation of the surrounding (eight) neighboring elevation posts by 100 meters or more. Voids of 16 or fewer contiguous posts were filled by interpolation of surrounding elevations. All larger voids were left in the data. (Missing elevations are denoted in the SRTM DTED® by a null value of −32767; a Partial Cell Indicator field in the DTED® header record gives the percent completeness for the cell, where 00 is defined as 100 percent complete, 05 is 5 percent complete, 95 is 95 percent complete, etc.)

The water editing was the most time-consuming and difficult part of the finishing processing; however, the result is a much more realistic depiction of the terrain. It also significantly reduced the number of voids inherent in the unfinished JPL data due to the typically weak backscatter from water bodies. The key requirement here was that all water bodies and shorelines were to be depicted as they were in February 2000 at the time of the shuttle flight, not as they may have appeared at an earlier or later date based on other reference material. (The methods used to do this are described later.) Figure 1 illustrates the difference between unfinished and finished SRTM DTED® 2.



(a)

(b)

Figure 1. Shaded relief depiction of unfinished (a) and finished (b) SRTM DTED® 2 for a cell on the Chesapeake Bay (W077N37). White areas in the unfinished image are data voids; black areas in the finished image depict the edited water surface.

SRTM DTED® 1

Production of a Level 1 version of the SRTM DTED® was also a requirement. SRTM DTED® 1 requires essentially the same format as DTED® 2, but it is at a lower density. Elevation values are spaced at three-arcsecond intervals (approximately 90 m) between 50° N and 50° S. At latitudes between 50 and 60 degrees, the longitudinal spacing is six arcseconds, while the latitudinal spacing remains three arcseconds. The SRTM DTED® 1 was generated only after the DTED® 2 was finished. It was created by sampling the finished DTED® 2 at even three-arcsecond intervals, thereby forcing the SRTM DTED® 1 and SRTM DTED® 2 to have identical values at coincident horizontal grid locations, and, consequently, the same absolute vertical and horizontal accuracies at those locations. Note that because of this sampling scheme, it is possible that residual voids smaller than 16 contiguous posts may be present in the DTED® 1. In addition, some water bodies that were contained by higher elevations in the DTED® 2 may not be completely contained in the finished DTED® 1 (Although water containment is a desirable attribute in the DTED®, it is not explicitly required in the specification. It was decided that the extra processing to locate and adjust a small number of uncontained water bodies in the SRTM DTED® 1 was unwarranted at the time.).

Other SRTM Data Products

In addition to the unfinished SRTM DTED® 2, JPL generated four other associated data products for each cell: an ascending and a descending orthorectified image mosaic (AOIM, DOIM), Terrain Height Error Data (THED), and a Seam Hole Composite Map (SHCM). No finishing was required for the image mosaics and they were passed through to NGA in their original form. The THED and SHCM were updated if changes were made to the associated DTED® 2 during the finishing process.

The AOIM is a composite of the optimal pixels (generally those closest to a 45° look angle) from all the ascending passes (orbits) over a cell and contains a scaled radar intensity value for each pixel. It has a ground sampling distance of approximately 30 m. The DOIM is a similar product for the descending passes. The THED product contains a random error estimate at 90 percent confidence for each elevation value in the SRTM DTED® 2 as well as an estimate of residual long wavelength systematic error for sub-cells within a cell. The random error component is due principally to the radar system, while the long wavelength systematic component is thought to be due to un-modeled error in the attitude and orbit determination system (Rodriguez *et al.*, 2004; Salamonowicz, 2003). THED values corresponding to water body elevations in the DTED® are set to 0 meters and those corresponding to interpolated DTED® posts are set to a null value. The SHCM is a raster bit map, co-registered to the SRTM DTED® 2 file, indicating the location of radar image strip seams and voids in the data, and where voids have been filled during finishing. All four of these products conform to the National Imagery Transmission Format Standard (NITFS) Version 2.1 and are described in detail in the SRTM Data Products Specification (NIMA, 2001).

As a by-product of the water body editing, a water mask was generated, at essentially a 30 m resolution, corresponding exactly to the SRTM DTED® 2 for each cell. An SRTM Water Body Data (SWBD) file, produced for each cell in ESRI® Shapefile format (ESRI, 1998), provides a vector representation of all the water body shorelines captured during the finishing process.

Data Processing Strategy

An overview of the data processing scheme is shown in Figure 2. As noted earlier, data from the Space Shuttle were processed in two stages to produce the finished SRTM data products. JPL performed the first stage using its Ground Data Processing System, which took in the raw radar measurements, positioning and attitude determination data and some ground control, and output a suite of matching products: the AOIM, DOIM, DTED® 2, THED, and SHCM. These products were forwarded directly to NGA contractors who performed all the finishing processing including an independent validation and verification (IV&V) prior to delivering the finished products to NGA. JPL and the contractors spent a great amount of time developing automated processing systems that could handle the more than 14,000 cells of data collected during the mission. Much of the development work had to be done prior to the shuttle launch without any real data with which to work. Immediately after the successful data acquisition, JPL produced prototype products over a small number of cells for the contractors to use for testing and evaluating their production systems.

In order to accomplish the finishing task in a reasonable time period and to insure against possible contractor

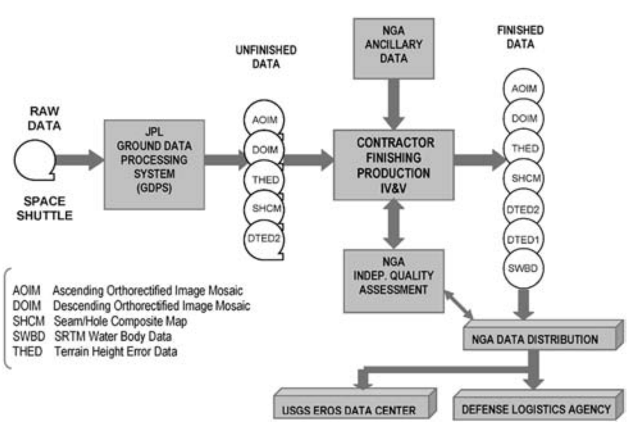


Figure 2. Overview of SRTM data processing.

performance problems, NGA employed two independent contractor teams to finish the SRTM data. BAE Systems and Vexcel Corporation constituted one team; Boeing (formerly Autometric, Inc.), Intermap Technologies, and PixSell® formed the second team. They were each given the same functional requirements and the production work was divided between them. BAE based its system development on its SOCET SET® software, while Boeing based its development on its SoftPlotter® software. The production target was an average of 20 hours manual processing time per cell. The work was done at three production sites and at its peak employed a total of 80 to 90 personnel working two shifts. During the peak production periods, the contractors (combined) delivered 800 finished cells per month to NGA.

Ten percent of the delivered cells were randomly inspected, but the rest went directly from the contractors into the NGA data distribution system. All DTED® 1, U.S.-only DTED® 2 and the SWBD were forwarded to the U.S. Geological Survey's Earth Resources Observation and Science (EROS) Data Center for public dissemination. The rest of the data products are at present restricted to U.S. Department of Defense use and are only distributed to others on a case-by-case basis.

Quality Assurance

Key attributes of the SRTM data are its homogeneity and consistency. The challenge was to preserve these qualities through the finishing process. Thus, a significant effort was expended to try to minimize and control the differences between contractors, sites, shifts and operators. Computational methodology and editing procedures were standardized wherever possible for things such as interpolation of voids, setting lake and river elevations, feature representation, and the use of ancillary data sources. With input from the contractors, NGA developed an extensive set of editing rules (NIMA, 2003a; NIMA, 2003b) for all contractor staff to use for the data processing. The rules dealt primarily with the water editing, but also with updating the THED and SHCM during finishing, and were continually revised as new editing issues were identified.

Multiple quality control checks were inserted into the processing. Immediately upon arrival at the contractor sites, all unfinished JPL data products were screened for blunders, omissions, inconsistencies, and other gross errors. Automated editing was always followed by an operator's manual review. If this review resulted in any manual editing, a second review was required by another operator. Extensive sets of software tools were developed by the contractors to facilitate the data review and editing. These included pseudo-stereo and shaded relief displays of the DTED®, statistical comparisons with reference data, and graphical displays of water and shoreline elevations. Self-consistency checks were built into the automated processing to assure a one-to-one correspondence for the DTED®, THED, and SHCM. All finished products for a cell were subjected to independent verification and validation (IV&V) procedures conducted by a reviewer from the contractor who had not previously worked on that cell. Finally, NGA reviewed in depth a random sample of 10 percent of the finished cells for specific and systemic production problems as cells were delivered each month. Immediate feedback was given to the contractors to correct errors or to revise or clarify editing rules. Regardless of the source of a problem, comments were always given to both contractors to maintain a common baseline.

Ancillary Reference Data Provided for Finishing

To assist with the quality control and editing of the finished data, four types of reference data were provided to each contractor. Photogrammetrically-derived ground control points (5 to 10 m vertical accuracy) and limited reference DTED® 2 (10 m vertical accuracy) from optical imagery sources were used to check the validity of SRTM elevations prior to and after the finishing process. In support of the water body editing, the contractors were supplied with Compressed Arc-Digitized Raster Graphics (CADRG) and "Landcover" water masks for most of the cells. The CADRG are digitized versions of NGA maps and charts from 1:50 000 to 1:1 000 000 scales. Generally, 1:250 000 or larger scale maps proved useful as an indicator of the possible presence or absence of water in an area. They were not used for delineating water bodies, but served as backup support for the OIMs and Landcover.

The Landcover water layer used as a reference in the SRTM finishing was one of 13 Landcover classes of orthorectified Landsat Thematic Mapper (TM) data obtained under an NGA contract with Earth Satellite Corporation. The Landsat scenes from which the water layer was derived dated from the late-1980s to 1993, so they were significantly older than the SRTM data. This water layer intentionally excluded ice, snow, wetlands, agricultural and rice fields, and mangrove swamps. The planimetric accuracy of the data was specified as 50 m (RMSE), but the actual accuracy may be closer to 19 to 25 m (RMSE) based on a NASA evaluation. The Landcover had a pixel resolution of 28 m, approximately the same as SRTM imagery. For the United States only, the water layer was generated from the U.S. Multi-Resolution Land Cover (MRLC) dataset derived from Landsat TM data from 1990 to 1994.

Editing Rules

Editing rules were developed by NGA to build uniformity and consistency into the SRTM finishing process. The water body delineation, in particular, required detailed specifications for treating generic situations that would recur during the course of editing the 14,000 cells. The OIMs were the primary sources for water identification and delineation. All water bodies and shorelines were delineated as much as possible as they appeared at the time of the data collection in February 2000. The ancillary data sources were used only as indicators and confirmation of the presence of water in an area. In cases where the land/water interface was indiscernible in the OIMs, or when SRTM data were missing, the Landcover may have been used for water boundary delineation. In the rare cases where only Landcover was used to delineate water bodies in an entire cell, the text "Landcover only water processing" is entered into the DTED® Data Set Identification (DSI) record. When water existed in a cell but the OIMs were unusable and no Landcover was available, the DTED® DSI record contains the entry "No water processing."

From an operational perspective, the general requirements and the DTED® specification left too much uncertainty as to how to deal with specific editing situations, so the Edit Rules tried to address these shortcomings. For example, since SRTM data represent the reflective surface, causeways, harbor facilities and piers are left in the finished DTED® product provided they are substantiated by the OIMs, Landcover, or CADRG, and the structures are not less than 90 m wide. On the other hand, bridges, power lines, and pipelines were removed over regions classified as water in the finished data in order to maintain the integrity of the water body.

As a general rule for the determination of elevations during the JPL phase of the data processing, all ocean elevations were defined as 0 meters. This served as a baseline for the finishing. The contractors edited the cells from the coastline inland to ensure consistency with the ocean. In a cell where the ocean shorelines appear to be at high tide in one of the pair of OIMs and low tide in the other OIM, high tide was selected for delineation in the DTED®.

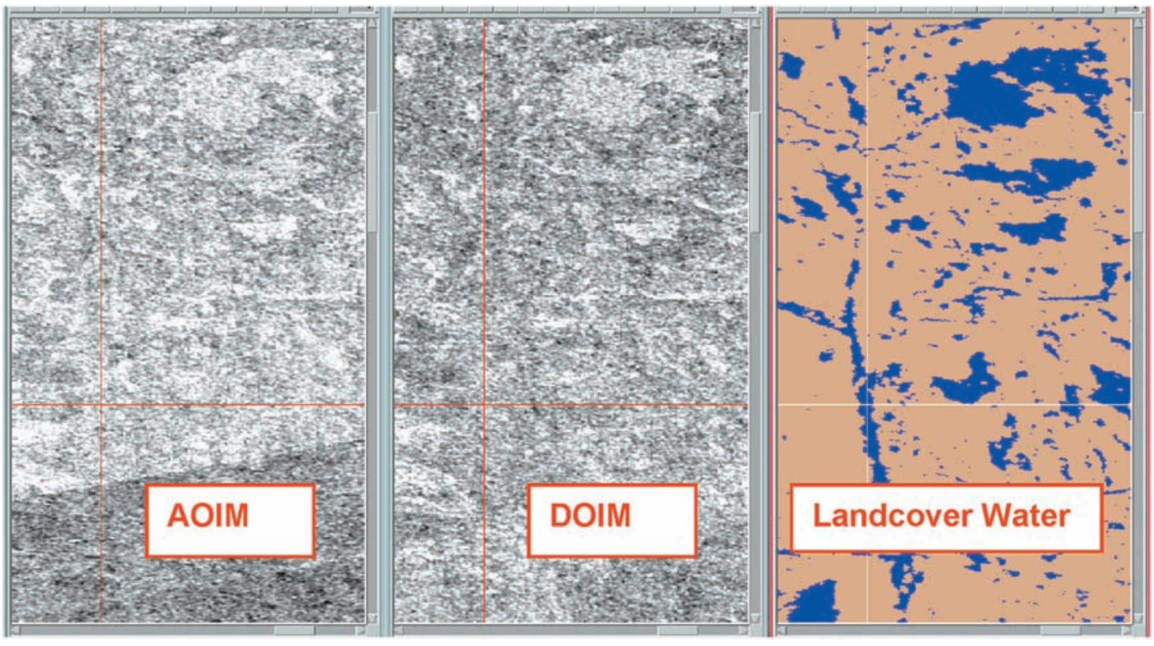


Plate 1. Example of snow and frozen water bodies in AOIM and DOIM for cell W079N58 with the older Landcover water layer as a reference. The lack of contrast due to the ice and snow in the images made automated and manual water body identification and delineation extremely difficult. Editors relied on the Landcover reference for guidance and in a few worst cases as a default for locating larger water bodies.

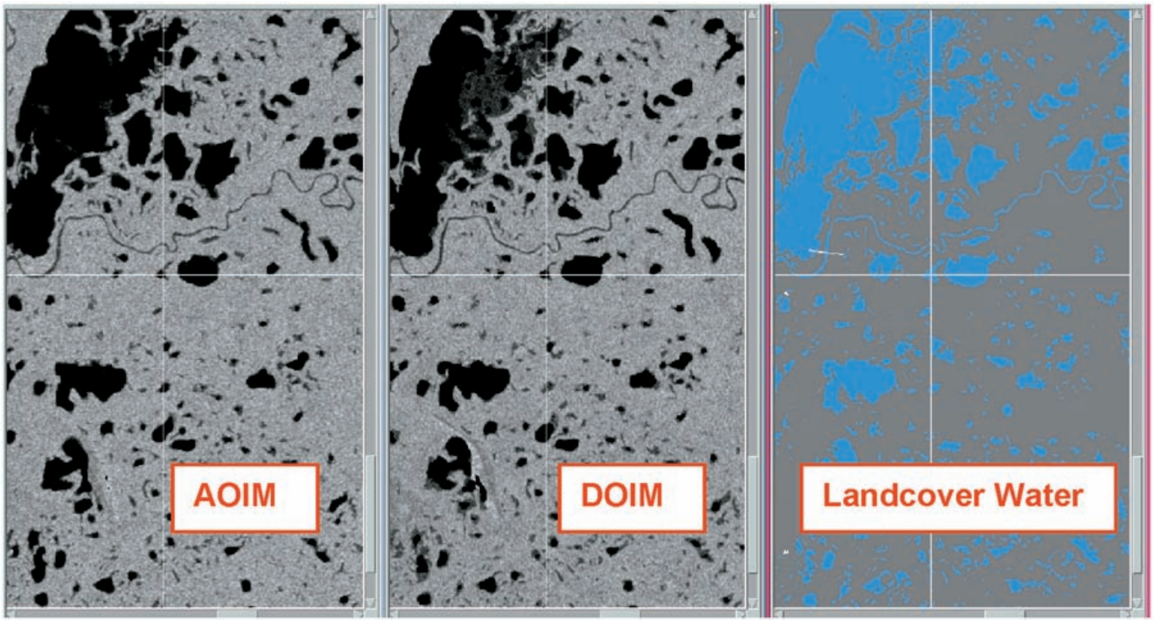


Plate 2. Example of water bodies that are well defined in the AOIM and DOIM and in good agreement with the older Landcover water layer. The large number of water bodies that meet the DTED® capture criteria add complexity to the data editing.

Practical issues addressed by the Edit Rules include the following examples. What is the minimum length of a river that meets the minimum width criterion for it to be included as a river? The rule is 600 m. Should a river continue to be depicted as its width narrows below 183 m, and where should it be cut off? The rules call for the depiction to stop when the river narrows to less than 90 m, but also allow for continuing the depiction if the river widens again, in order to maintain continuity of the drainage. Within a cell, rivers were delineated from upstream to downstream in order to eliminate some of the trailing ends of double-line drains as they narrow down below 90 m. How should agricultural areas that appear to be covered with water in the OIMs be classified: as land or water? They are depicted as water only if also supported by the Landcover or CADRG. In some cases, islands may not appear in the SRTM data due to poor radar returns on the water, which limited JPL's ability to extract the islands from the data. In these situations, if an island was detected in the Landcover data or in the CADRG and was large enough to meet capture criteria, then the island is portrayed as a void surrounded by water in the DTED®.

In many instances, water bodies appear all or in part as voids in the OIMs and as such are indistinguishable from voids due to radar shadow or layover or noisy data rejected during JPL's processing. This is especially true in high relief areas. Care was taken to confirm that the suspected water bodies were consistent with the surrounding terrain, thus avoiding illogical features such as lakes sitting on steep slopes.

Methodology for Automated Water Identification and Delineation

Based on the overall finishing requirements, the editing rules and the unique characteristics of the SRTM radar image data, each contractor designed automated methods and tools for identifying and delineating water bodies and for determining the appropriate elevations. Automating this process was the only viable way to handle the volume of data, standardize the extraction of water, and minimize the dependency on "cartographic judgment" of operators. The two approaches developed by the contractors worked fairly well, but were not totally reliable even in ideal conditions. Frozen lakes in the higher latitudes (Plate 1) and extremely complex water networks in some areas (Plate 2) complicated the problem.

Boeing/Intermap Water Classification Approach

Intermap Technologies used a probabilistic approach to evaluate the distribution of the scaled radar intensity values in the OIMs and from this determined which pixels were most likely water and which were not. Radar signals exhibit low or no returns over water, resulting in dark or void regions in the OIMs. Even in regions where there are useable radar returns, the resolved heights are corrupted, as the signal is dominated by noise or water motion, and thus the DTED® is very noisy in these areas. JPL was able to retrieve enough data during the interferometric processing to partially or completely define many water bodies. Figures 3a and 3b show typical ascending and descending image mosaics.

Intermap Technologies converted the image intensity data (grey values) into histograms: one for the AOIM and one for the DOIM in each cell. As expected, the water exhibited a lower signal-to-noise ratio than most land-based targets, typically creating a separable bimodal distribution with a varying degree of mixed modalities (see Figure 3c). Initial testing indicated that each mode in a given OIM tended to be normally distributed, with significant overlap between the peaks. This phenomenon allowed the OIMs to be segmented into three data types representing water, land, and void (defined as no data areas). This was accomplished using an optimum thresholding technique (see, for example, Sonka *et al.*, 1993) that generates an independent probability image for each of the OIMs. In order to provide interactive processing speeds, two assumptions were made when selecting the

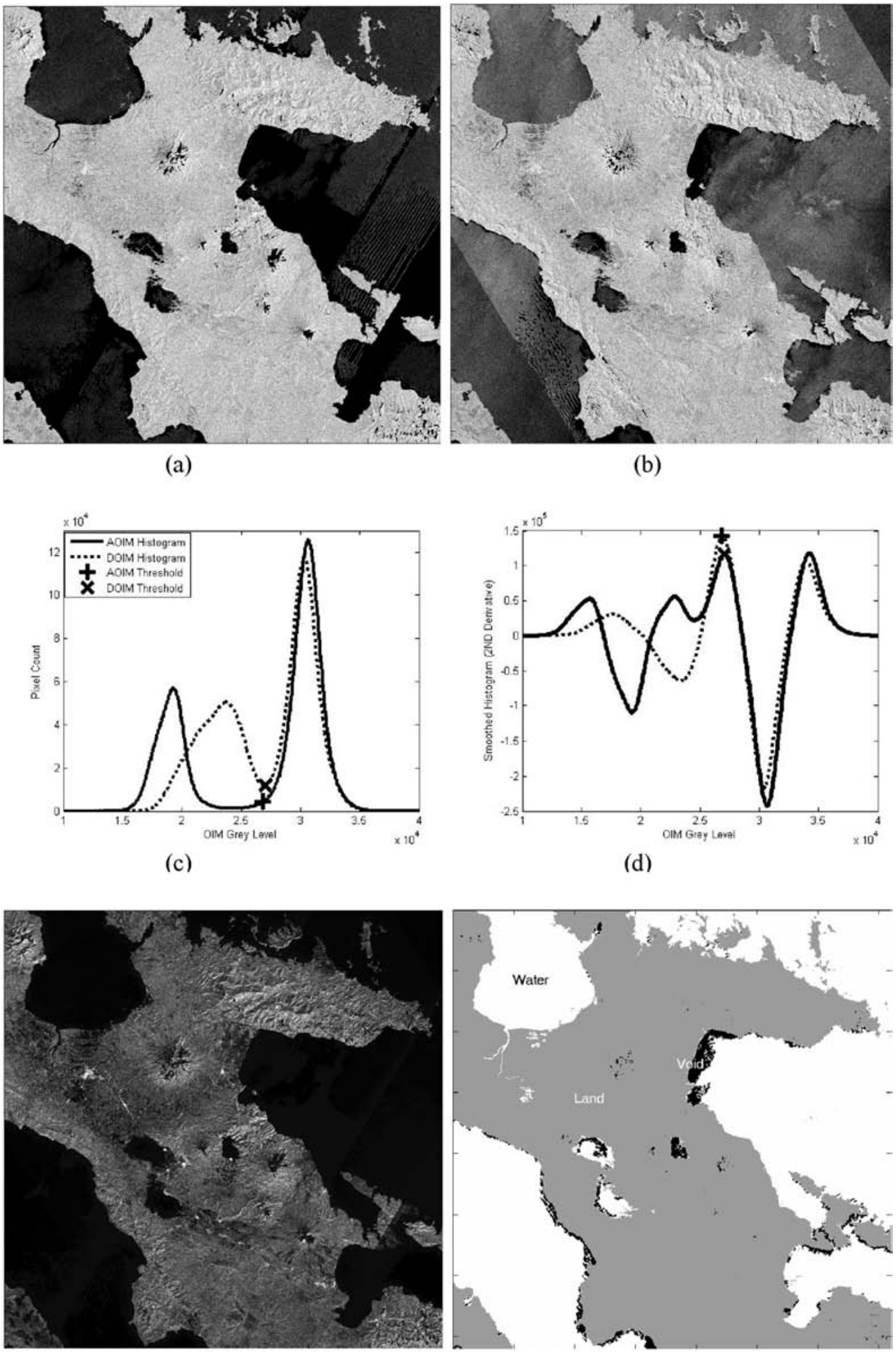


Figure 3. Illustration of identification and delineation of water bodies using the thresholding technique for the AOIM (a) and DOIM (b) for cell E123N13. Bimodal histograms (c) of AOIM and DOIM grey level values show the threshold values used to distinguish land from water as derived from the smoothed 2[nd] derivative curves (d). After scaling the grey-level values in each OIM by their respective threshold values, a single probability image is created by summing the scaled AOIM and DOIM (e). The raster classification mask (f) for land (grey), water (white) and voids (black) is generated from this.

threshold: (a) water and land exist within the cell, and (b) the rightmost, highest grey-level histogram mode (peak) represents land. For each image histogram, a corresponding histogram of its smoothed second derivatives was generated, and the second maximum value going from highest to lowest grey value (or right to left) in the latter histogram was selected as the optimum land/water threshold (see Figure 3d). The AOIM and DOIM threshold values were then subtracted, respectively, from the AOIM and DOIM intensity values to create two separate probability images, where values above zero represent land and those below zero represent water. These adjusted images were summed (Figure 3e) and an adaptive floodfill algorithm was used to segment the probability image into a raster mask of land, water, and voids (Figure 3f). The adaptive floodfill technique utilized a conservative seeding technique to initially fill regions over posts with a high probability of being water. A liberal filling procedure followed that allowed water bodies to breach the threshold by a user-defined percentage, when connected to a water body that has met the seeding requirements.

Following the delineation of a base classification layer of water, land, and voids, a set of rule-based operations was performed to prune and reclassify the results in accordance with the Edit Rules described earlier. Upon completion of the automated classification, an operator reviewed and corrected the raster mask for errors using a combination of the DTED® 2, AOIM, DOIM, THED, and the ancillary reference data provided by NGA. The water mask then was vectorized using a skeletonization operation to create a continuous topological representation (water network) of each water system.

Lake elevations were determined first using a two-step process. An average elevation of a lake was computed directly from sampling elevations of the water itself when these data existed. The average lake elevation was then compared against the lowest 5^{th} percentile of the lake's shoreline elevations. The lower of the two was then selected as the elevation for the lake.

River elevations were determined by sampling the water elevations perpendicularly from the centerline of the water network out to the neighboring riverbanks. The average of these samples was stored at each point within the water network along with the associated shoreline elevations. An iterative fitting mechanism was then utilized to suppress erroneous elevation values resulting from systematic errors or correlated noise within the river. All possible paths were determined from the start to the end of a river system, resulting in numerous permutations if the river system contained islands and branching. From these, the most direct path was selected and processed to ensure monotonicity and containment within the majority of the river's banks. Following completion of this path, the algorithm filled all connected paths ensuring that the selected elevations were tied to previously selected elevations. This process ensured that drainages were monotonic, topologically correct, and stepped in one-meter increments.

After the automated estimation of the water elevations, an operator reviewed the proposed heights and could, if needed, manually alter the selections. When an acceptable representation was achieved, a final editing operation set lakes to one elevation, stepped down rivers in one-meter increments as dictated by the water network, and made sure that all the water was contained appropriately within the surrounding relief.

BAE/Vexcel Water Processing Approach

The BAE/Vexcel Corporation suite of editing tools for SRTM was built as a set of add-on modules to BAE System's SOCET SET® product. Their approach used an edited version of the Landcover water mask as a guide for the SRTM water identification and delineation.

To begin, the Landcover raster water mask was processed in three steps to provide water body outlines, island outlines, and cell-edge boundary markers to the subsequent processing modules. In the first step, the Landcover preprocessing software filtered the Landcover mask to remove small water bodies, aggregate closely located clusters of disconnected water pixels into clusters of connected water pixels, and smooth the edges of large water bodies, using morphological operators on the original mask. Water bodies on the cell boundaries required special handling; all were included in the output from Landcover preprocessing, regardless of their size, since boundary water bodies might belong to a larger body of water on the other side of the boundary. Boundary markers (two-point vectors describing the precise intersection of a water body along the cell boundary) were then generated for each water body. When the Landcover water mask was not available for a cell, this preprocessing step was skipped.

After the filtering process, automated water body and island sizing procedures were used to cull out features in the Landcover mask that were too small for inclusion in the edited data. Marginal features were kept, since later processing might increase their size and make them candidates for inclusion. The final step in the Landcover preprocessing was to convert the water bodies, islands and boundary markers from raster images to ESRI® Shapefile format vectors so the outlines could be displayed over the OIM raster images for manual editing in the next part of the process. Here, the water body outline polygons (islands included) created by the Landcover preprocessing were presented to the operator for interactive editing, laid over the AOIM and DOIM. The point of this *mask editing* step was to correct large-scale errors in water body topology in the Landcover mask, that were uncorrected in Landcover preprocessing; in particular, to ensure that networks of rivers were appropriately connected, since these were often unclearly represented in the Landcover data. The operator had the capability to disconnect or reconnect inappropriately represented water bodies, but had to delineate the water bodies only in the coarsest sense. A double-line drain, for example, might be coarsely delineated with only a few line segments; automatic delineation would later adjust the data to more accurately fit the water body boundary in the SRTM data. It was the operator's task only to ensure that the topology of the water network was correct.

Automated delineation of the SRTM water bodies followed. The algorithms used Gaussian-smoothed AOIM and DOIM data, and were derived from the theory of regularized energy functionals (Nordstrom, 1990). The best delineation of the water body was the polygon defined by the set of vertices, $\{\vec{v}_i\}_{i=1}^{n}$, minimizing the following energy functional:

$$E = \sum_i f(v_i) + g(v_i).$$

The first term f is a maximum likelihood estimate for the contour that bounds the feature, given the local data. The second term g is a term which forces the contour to be smooth, and not to self-intersect. A third term was added to the sum when large regions of bright ice were encountered in data from the northern parts of Northern America, permitting the favoring of contours that enclosed water bodies that were either brighter or darker than their surroundings. The minimization of this mathematical expression for each polygon was an optimization in high-dimensional parameter space; a gradient descent algorithm was

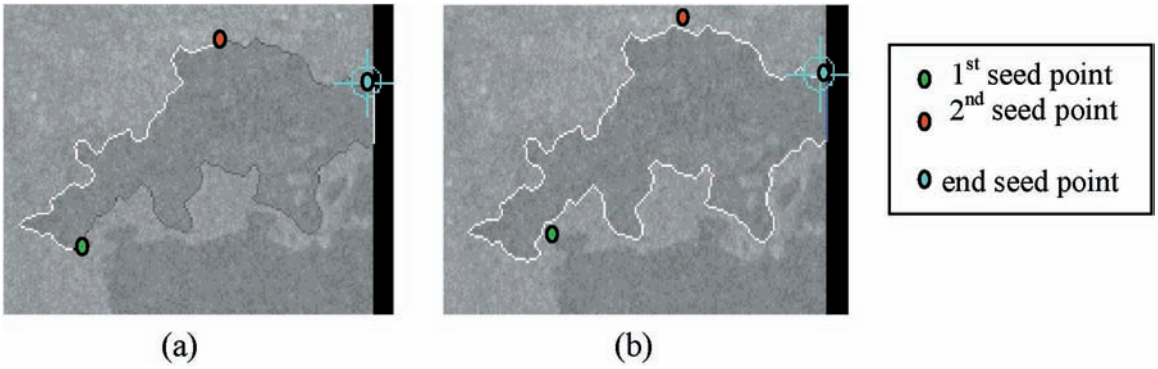


Plate 3. Fast Trace feature extraction of a lake with only three user-input points (a) resulting in the shoreline delineation (b).

used to minimize the expression. The resulting polygon was often a perfect fit to the water body in the OIM data, provided that the initial estimate (provided either by the Landcover data or by the operator in the Mask Edit step) was physically close to the final contour.

Following the auto-delineation step, the operator used the water body delineation and editing tool, *SITE*, to manually finish the task of defining the water body, estimating the water elevation(s), and editing the feature into the SRTM DTED®. In the manual delineation step, an algorithmic tool called *Fast Trace* was routinely used to add and adjust water bodies in the output from the auto-delineation. Fast Trace is an efficient and accurate boundary extraction method using minimal operator input. It used features from the AOIM and DOIM as input to calculate the optimal paths from each pixel to a seed point set by the user. The optimal path is defined as the minimum cumulative cost path from a start seed point to an end seed point. Both the start and end seed points must be known; often Fast Trace requires only two points on the boundary of a lake to be identified by the operator to extract a perfect boundary representation of the feature as illustrated in Plate 3.

After delineating a water body, the processing path diverged for double-line drains and lakes. The core problem in both cases was to estimate the elevation at a point or points within the water body. However, while SRTM DTED® elevations were usually available over lakes to aid in the estimation process, the DTED® was typically void over the centerlines of double-line drains (usually because of flowing water).

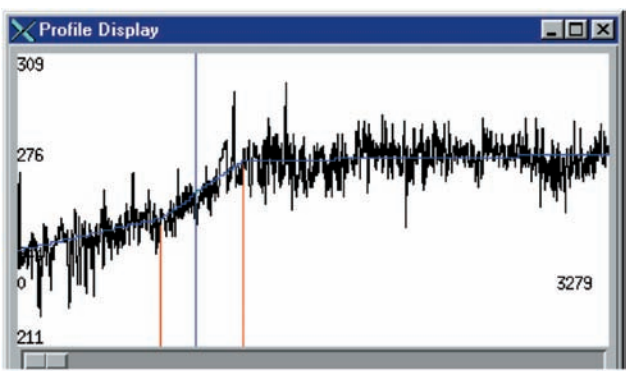


Plate 4. Profile display for double-line drain elevation editor, showing the estimated centerline elevation (in black), and the adjusted step function estimate (in blue).

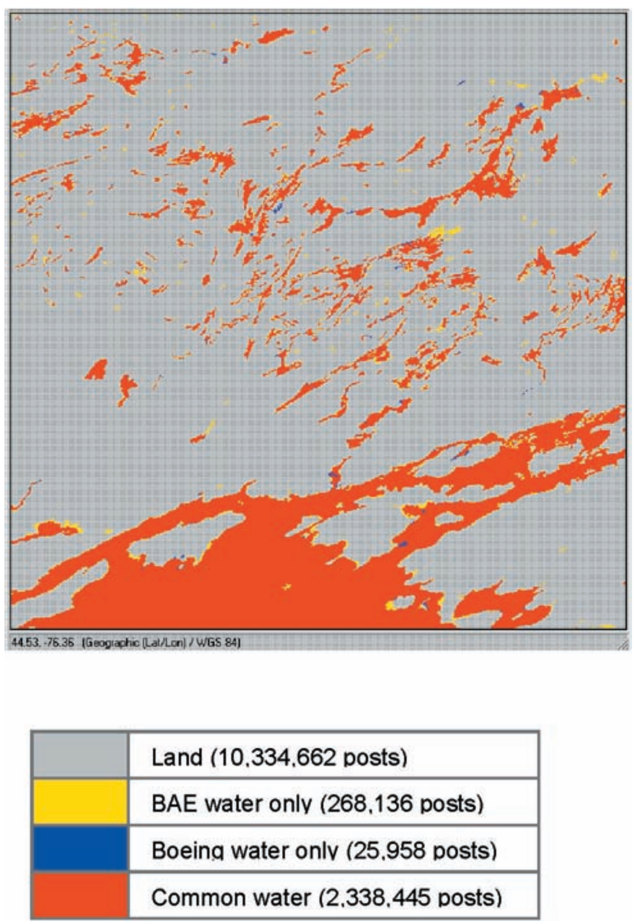


Plate 5. Differences between BAE and Boeing in identification and delineation of water in sample cell W077N44.

For lakes, the elevation estimation was fairly straightforward. Two estimates were made of the elevation of the lake: one derived from over-water elevation estimates if they were available, and one derived from the shoreline of the water body. The water-derived elevation was taken to be a trimmed mean based on 24 water samples; the water samples were ordered, and the top and bottom four samples were eliminated to get rid of the worst outliers. The remaining 16 samples were averaged to obtain the over-water estimate, regardless of the standard deviation of their sample distribution.

The shoreline estimate was made by extracting all the elevations along the shoreline. In order to eliminate the worst outliers, a median filter was applied along the shoreline elevation function. The shoreline estimate was taken to be the minimum of the smoothed shoreline function. In practice, this estimate was sometimes too low, and the water-based estimate proved more reliable; therefore, the water-based estimate was routinely used as the default elevation. Note that the water-based and shore-based estimates were usually within five meters of each other, providing a high level of confidence for the final estimate.

In the double-line drain case, a graph of *estimated* elevation versus the centerline was generated and displayed to the user, along with an automatically calculated monotonic function describing the steps to be edited into the double-line drain. This estimate was derived at a point on the centerline, using samples from a narrow strip oriented normal to the tangent line at the centerline. Samples were used from the river itself if available, and were also taken from the shoreline a short distance onto the land, and biased downward to estimate the water elevation. The resulting elevation estimation function was usually quite noisy, but overall robust enough to use as a reference in setting a monotonic step-function as the edited double-line drain elevation estimate (see Plate 4).

Intercomparison of BAE and Boeing Water Editing

As a part of NGA's test and evaluation plan during the finishing production, 25 contiguous cells, in the Northeastern United States and Canada, were chosen for both contractors to finish. NGA wanted to quantitatively compare the finished DTED® 2 and in particular the water processing from the two contractors in order to understand and perhaps control the differences being introduced into the final products by the independent production systems. The North American cells were chosen because they were the first cells delivered by JPL and they contained varying degrees of water coverage ranging from less than 1 percent to about 88 percent of the cell. In retrospect, these cells were not ideal because many water bodies were frozen at the time of the mission. Production problems at one of the contractors delayed its completion of the test cells, thereby introducing about a six-month difference in the experience levels of the operators at the two contractors and the maturity of their respective production systems. Despite this, there were a number of interesting results that were derived from the comparison. Table 1 shows the number of water posts in common and unique to each contractor for ten representative cells. In 24 of the 25 cells, BAE classified more posts than Boeing as water. This may be due to the use of Landcover by BAE in its initial water identification procedure versus a more conservative approach by Boeing. The sum of the BAE-only and Boeing-only water posts is an indication of the contractor-induced uncertainty in the amount of water in a cell since these posts may or may not be classified as water depending on which contractor did the finishing. As a percentage of the water in common in each cell, the differences between the contractors in the total water defined in each cell ranged from 0.03 percent (in cell W073N40, the cell with the most water) to 26.23 percent (in cell W075N42, one of the cells with the least overall water). In terms of the entire cell (land + water), these contractor differences are a fraction of one percent of the total posts in a cell, with the exception of cell W077N44 where the total water difference amounts to 1.9 percent of the whole cell. The differences between the contractors are due to a difference in the number of lakes identified and differences in delineation of shorelines, no doubt complicated by the frozen nature of some of the water bodies.

Plate 5 illustrates the contractor differences in the location of water in cell W077N44, which had the largest difference in the number of water posts between the two contractors. The overall consistency is extremely good, despite the fact that BAE identified considerably more lakes than Boeing in the cell and that there are slight delineation differences throughout. Since the land elevations were not adjusted by the contractors, differences in elevations of one meter or more are attributable to water editing differences between the contractors. In cell W077N44, the water elevations determined by Boeing and BAE agree to within ±2 m for the vast majority of posts (see Plate 6).

Finished SRTM DTED® 2 Cell Statistics

The finished SRTM DTED® 2 and THED incorporate completeness and accuracy statistics in the final products. The editing process eliminates many of the voids in the unfinished data through interpolation and by defining water bodies. Table 2 gives a regional breakout of the percent of complete coverage (non-void areas) in the finished data.

TABLE 1. SAMPLE CELLS FROM A TEST COMPARISON OF WATER EDITING BY THE TWO SRTM PRODUCTION CONTRACTORS

	Water in Common Between Contractors		Water Differences Between Contractors		
Cell ID	No. of Posts	Percent of Entire Cell	BAE-Only Water (No. of Posts)	Boeing-Only Water (No. of Posts)	Total Water Difference Between BAE & Boeing (% of Common Water)
W076N40	61,671	0.48	18,330	5,625	20.60
W075N42	73,051	0.56	22,065	2,907	26.23
W073N43	102,375	0.79	34,492	8,527	25.36
W074N42	202,676	1.56	27,160	24,284	1.42
W075N44	418,006	3.22	61,803	27,812	8.13
W076N44	510,050	3.93	108,480	11,256	19.06
W074N41	1,344,585	10.37	52,694	17,231	2.64
W077N44	2,338,445	18.03	268,136	25,958	10.36
W077N43	6,152,311	47.45	47,044	11,479	0.58
W073N40	11,399,991	87.91	14,076	10,268	0.03

Table 2. Void Statistics for Finished SRTM DTED® 2 by Region (Excluding Island Cells)

Percent Full	North and South America No. of Cells	North and South America Cumul. % of Total	Eurasia No. of Cells	Eurasia Cumul. % of Total	Australia No. of Cells	Australia Cumul. % of Total	Africa No. of Cells	Africa Cumul. % of Total	Total No. of Cells	Total Cumul. % of Total
100	1,174	28.7	1,846	32.2	380	35.8	527	16.2	3,927	27.8
99	2,666	93.7	3,531	93.8	677	99.7	2,288	86.6	9,162	92.6
98	84	95.8	102	95.6	2	99.9	117	90.2	305	94.8
95–97	108	98.4	117	97.7	1	100.0	123	94.0	349	97.2
90–94	44	99.5	67	98.8	0		81	96.5	192	98.6
85–89	7	99.7	27	99.3	0		29	97.4	63	99.0
80–84	2	99.7	16	99.6	0		22	98.1	40	99.3
70–79	5	99.8	19	99.1	0		33	99.1	57	99.7
50–69	4	99.9	5	100.0	0		20	99.7	29	99.9
<50	3	100.0	0		0		9	100.0	12	100.0
Total	4,097		5,730		1,060		3,249		14,136	

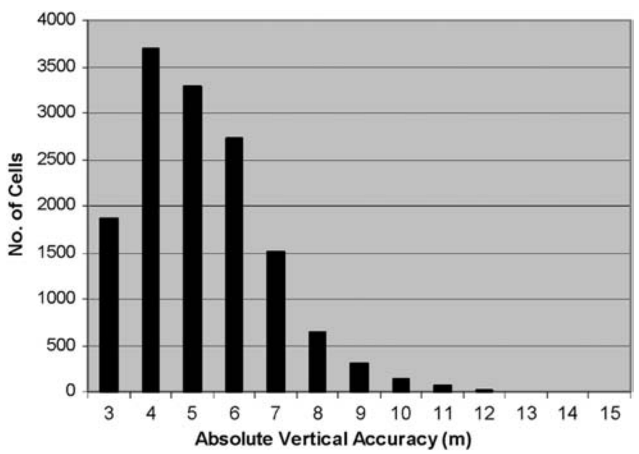


Figure 4. Global SRTM Absolute Vertical Accuracy estimates (90 percent confidence) showing 62 percent (8,846 cells) of the cells at 5 m or less, 91.7 percent (13,096 cells) at 7 m or less, and 99.4 percent (14,193 cells) at 10 m or less.

About 93 percent of the cells in the western hemisphere and Eurasia, and over 99 percent of the Australia cells have less than one percent voids, compared to 87 percent of the cells in Africa. The lower completion percentage is primarily due to the poor radar returns (low signal-to-noise ratio due to soil conditions) in some of the northern areas of the African continent.

Each SRTM DTED® cell also contains an absolute vertical accuracy estimate that is a composite of the random error and the systematic error estimates for the cell as a whole at the 90 percent confidence level. The distribution of these estimates for all 14,277 cells is shown in Figure 4. All the estimated vertical accuracy values fall between 3 and 15 meters, with 99.4 percent of the cells having absolute vertical accuracies of 10 m or less. The original 16 m vertical accuracy requirement appears to have been met and surpassed by a significant margin.

Summary

The challenge of the finishing phase of production was to "do no harm" to the SRTM terrain elevation data that had been carefully acquired and processed by JPL. The mission was designed to produce a global data set that was internally consistent and homogeneous, and met the accuracy and feature specifications for DTED® 2. The need to maintain this consistency and the sheer volume of data demanded the use of automated procedures for product finishing. NGA provided detailed functional specifications and editing rules to the two finishing contractors, but each developed production systems independently. Procedures for filling voids and removing spikes (and wells) were easy to implement the same way at both contractors, but water editing was much more complicated. Despite the common rule set, the differences in the processing software and manual editing applied by the two contractors caused some noticeable differences in the delineation of water bodies. Based on the limited comparison of test cells, these differences are relatively small with respect to the entire cell in most cases. Since the majority of "land" posts remained untouched by the finishing, their elevation values remain unchanged in the final product.

Many of the voids in the unfinished DTED® have been eliminated in the final products leaving only a few percent of the data missing. The THED and SHCM files were updated in the production process to match the finished DTED® 2, and DTED® 1 and SWBD files were created after these other files were finalized to complete a suite of matching products that includes the image mosaics for each cell. Overall, the finishing work added value to the SRTM products by more realistically depicting topography and cleaning up some of the small data anomalies. The accuracy of the DTED® was preserved, if not improved slightly, during the finishing and is better than the original requirement by as much as a factor of two.

Acknowledgments

The authors wish to acknowledge the work of many people from Boeing, BAE, Intermap Technologies, Vexcel Corporation, and NGA who contributed to the development, testing and evaluation of the production methodology described in this paper. We particularly want to cite Richard Carande, Marty Mara and Zhaoqiang Bi of Vexcel, and Tom Spoering and Keith Tennant of Intermap Technologies for their efforts.

References

ESRI, 1998. *ESRI Shapefile Technical Description, An ESRI White Paper*, July.

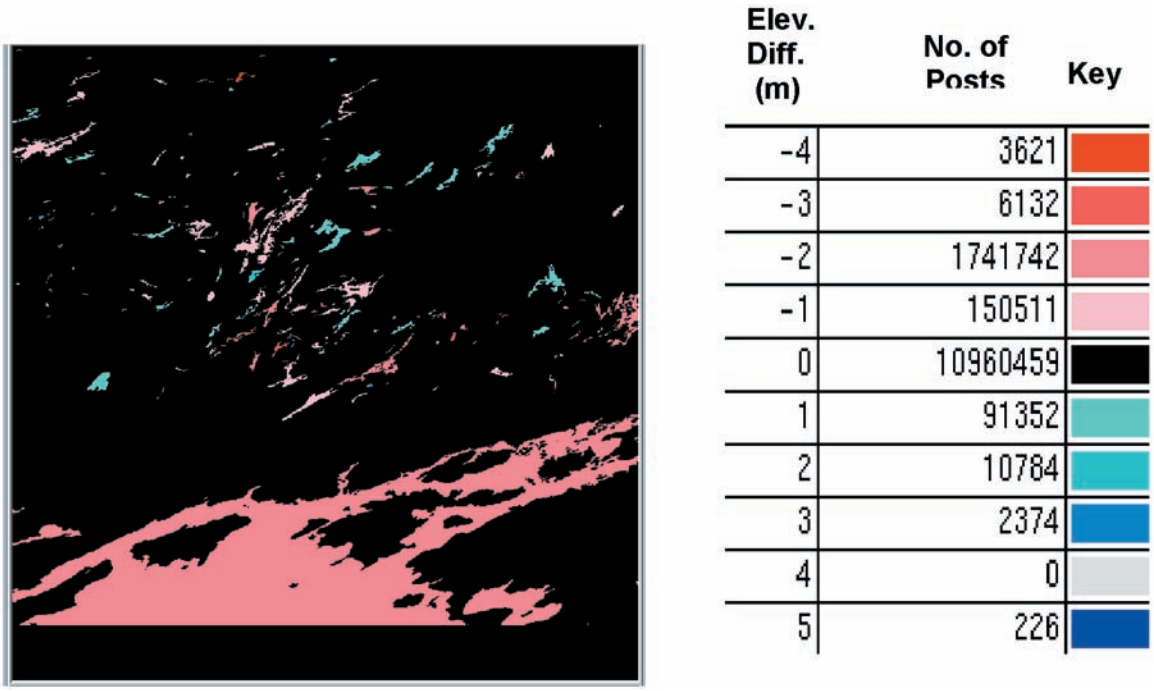


Plate 6. Edited water elevation differences (Boeing minus BAE) for sample cell W077N44. Note that zero difference (black) includes land as well as water.

Farr, T.G., and M. Kobrick, 2000. Shuttle Radar Topography Mission produces a wealth of data, *EOS Transactions*, American Geophysical Union, 81(4):583–585.

NIMA, 2000. *Performance Specification, Digital Terrain Elevation Data (DTED®)*, MIL-PRF-89020B, 23 May.

NIMA, 2001. *Shuttle Radar Topography Mission (SRTM) Data Products Specification, Specification for DTED® and NITFS Formatting of SRTM Data Products*, National Imagery and Mapping Agency, Version 1.3, 19 April.

NIMA, 2003a. *SRTM Data Editing Rules*, Version 2.0, 12 Mar.

NIMA, 2003b. *SRTM Water Body Data Product Specific Guidance*, Version 2.0, 12 March.

Nordstrom, K.N., 1990. Biased anisotropic diffusion: a unified regularization and diffusion approach to edge detection, *Image Vision Computing*, 8(4): 318–327.

Rodriguez, E., C.S. Morris, J.E. Belz, E.C. Chapin, J.M. Martin, W. Daffer, and S. Hensley, 2004. *An Assessment of the SRTM Topographic Products*, NASA Jet Propulsion Laboratory, Report No. D-31639.

Salamonowicz, P., 2003. Accuracy assessment of elevation data from the Shuttle Radar Topography Mission, *Proceedings of 2003 ASPRS Conference*, 05-09 May, Anchorage, Alaska.

Sonka, M., V. Hlavac, and R. Boyle, 1993. *Image Processing, Analysis and Machine* Vision, First edition, Chapman & Hall, pp. 118-119.