An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery

David E. Shean [a,*], Oleg Alexandrov [b], Zachary M. Moratto [b], Benjamin E. Smith [a], Ian R. Joughin [a], Claire Porter [c], Paul Morin [c]

[a] Polar Science Center, Applied Physics Lab, University of Washington, Box 355640, 1013 NE 40th St., Seattle, WA 98105, USA
[b] Intelligent Robotics Group, NASA Ames Research Center, M/S 269-3, Moffet Field, CA 94035, USA
[c] Polar Geospatial Center, University of Minnesota, R280 Learning & Environmental Sciences, 1954 Buford Ave, St Paul, MN 55108, USA

ABSTRACT

We adapted the automated, open source NASA Ames Stereo Pipeline (ASP) to generate digital elevation models (DEMs) and orthoimages from very-high-resolution (VHR) commercial imagery of the Earth. These modifications include support for rigorous and rational polynomial coefficient (RPC) sensor models, sensor geometry correction, bundle adjustment, point cloud co-registration, and significant improvements to the ASP code base. We outline a processing workflow for ∼0.5 m ground sample distance (GSD) DigitalGlobe WorldView-1 and WorldView-2 along-track stereo image data, with an overview of ASP capabilities, an evaluation of ASP correlator options, benchmark test results, and two case studies of DEM accuracy. Output DEM products are posted at ∼2 m with direct geolocation accuracy of <5.0 m CE90/LE90. An automated iterative closest-point (ICP) co-registration tool reduces absolute vertical and horizontal error to <0.5 m where appropriate ground-control data are available, with observed standard deviation of ∼0.1–0.5 m for overlapping, co-registered DEMs ($n = 14, 17$). While ASP can be used to process individual stereo pairs on a local workstation, the methods presented here were developed for large-scale batch processing in a high-performance computing environment. We are leveraging these resources to produce dense time series and regional mosaics for the Earth's polar regions.

© 2016 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

The archive of very high-resolution (<1 m) satellite optical imagery for Earth has grown exponentially in the past decade. Commercial vendors can collect on-demand, sub-meter imagery anywhere on the planet from multiple platforms, with revisit times of less than a day at higher latitudes. Resampled (0.25 m minimum GSD as of June 2014, formerly 0.5 m) image products are now available to United States federal employees and federally-funded civilian researchers through the NextView license—a partnership between the National Geospatial-Intelligence Agency (NGA), commercial vendors, and federal agencies supporting scientific research (Neigh et al., 2013). Access to these data is enabling exciting new research across many disciplines, including wildlife ecology, forestry, geology, volcanology, and cryospheric sciences.

In 2009, the Polar Geospatial Center at the University of Minnesota initiated a campaign to obtain commercial imagery for the Earth's polar regions, specifically Antarctica, Greenland, and Alaska. This effort has grown significantly, with millions of scenes, many acquired as stereo pairs, now available for polar research and logistical applications. This growing catalog contains multiple years of cloud-free observations (Fig. 1), including hundreds of repeat observations for many high-priority science targets (e.g. Jakobshavn Isbrae, Greenland (Shean et al., 2014)).

1.1. Instrument description

DigitalGlobe currently offers high-resolution optical imagery from six spacecraft. The high inclination and relatively short period of these polar-orbiting satellites is ideal for repeat high-latitude data collection, where competition for commercial tasking is limited. Here, we focus our discussion on the WorldView-1 and

---

∗ Corresponding author. Tel.: +1 (206) 221 8727; fax: +1 (206) 543 6785.
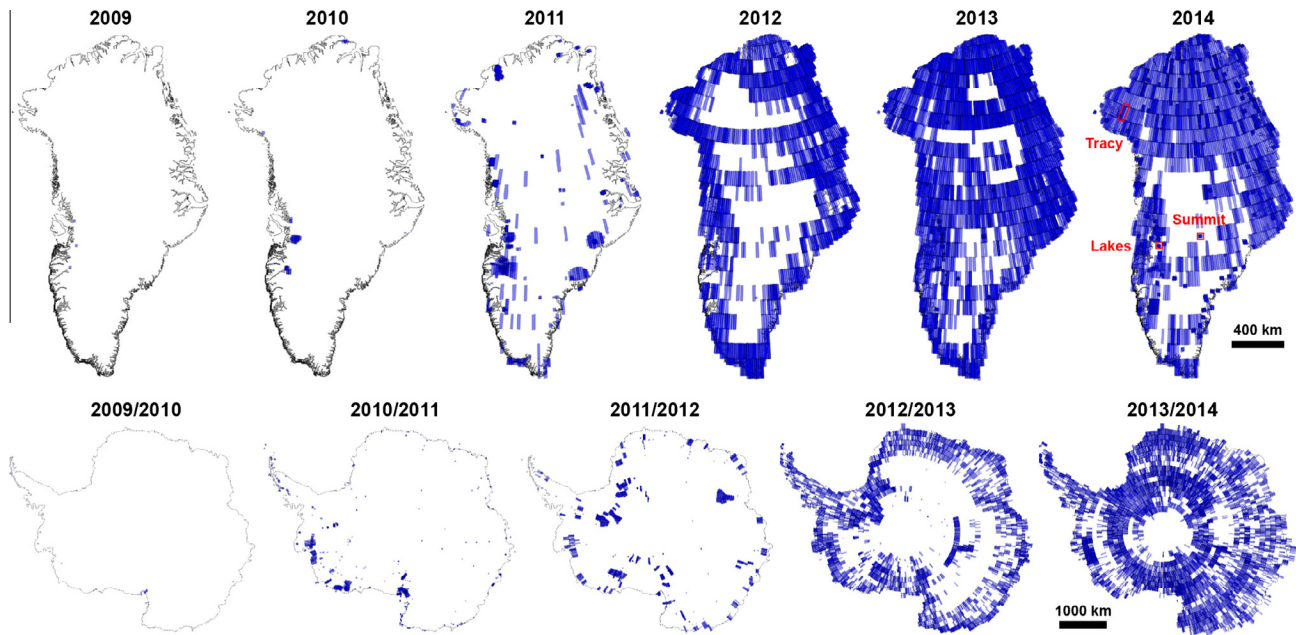*E-mail address:* dshean@uw.edu (D.E. Shean).

Fig. 1. Annual stereo coverage (cloud cover <75%) for Greenland (top row) and Antarctica (bottom row) in DigitalGlobe archive. Labels (Tracy, Summit, Lakes) and outlines on 2014 Greenland map indicate case study site locations.

WorldView-2 platforms, which constitute the bulk of the archived commercial stereo imagery from 2009 to 2015.

WorldView-1 and WorldView-2 share a similar pushbroom linescan camera with 11-bit dynamic range and 8–64 line time-delayed integration (TDI). The WorldView-1 focal plane includes 50 panchromatic (450–800 nm) charge-coupled devices (CCDs, also known as Detector Sub-Arrays [DSAs]) with 8-μm pixel size, arranged in two adjacent rows (see Fig. 1 of Updike and Comp, 2010). This provides an effective swath width of 35,840 pixels, corresponding to ∼17.6 km at ∼0.5 m GSD for nadir acquisitions. WorldView-2 has a similar panchromatic array, with ∼16.4 km swath width and ∼0.46 m nadir GSD.

Numerous high-resolution satellite platforms are capable of acquiring stereo imagery (e.g., Deilami and Hashim, 2011). However, only sub-meter GSD images resolve small-scale surface features (e.g., sastrugi, crevasses) that are not apparent at lower resolution (e.g. ∼15–30 m GSD Landsat imagery). This high-frequency texture enables precise image correlation for feature tracking and/or surface reconstruction.

The WorldView-1 and WorldView-2 spacecraft acquire images at off-nadir angles from 0° to >45°, with the ability to acquire two or more images of the same target (>90% overlap) in a single orbital pass, forming an along-track stereo pair with typical convergence angles of ∼30–60°. The relatively short time interval (∼60–90 s) between acquisitions of subsequent along-track images almost always ensures repeat observation of an effectively identical scene with similar illumination. In addition, sensor position/attitude errors for along-track stereo pairs are correlated, which generally improves relative accuracy (Dolloff and Theiss, 2012).

While the analysis presented here focuses on along-track stereo pairs, it also is possible to form across-track or "coincident" stereo pairs from two images acquired on different orbits, as long as they have appropriate convergence angle, solar illumination, and time separation (e.g., Becker et al., 2015). The maximum allowable time separation depends on surface displacement rates and/or textural change rates (e.g. snow accumulation or melt rates). An interval of a few hours is generally appropriate for fast-flowing outlet glaciers, whereas the threshold for a static desert landscape might be several years.

1.2. Data description

Nearly all WorldView data available from the NGA archives are Level 1B (L1B) products—seamless, geometrically- and radiometrically-corrected mosaics of sub-images from DSAs in the focal plane array. Images are typically delivered in the National Imagery Transfer Format (NITF), which consists of a comprehensive header and compressed JPEG2000 subdatasets.

The L1B images are often split in the along-track direction and delivered as a set of overlapping, ∼30 K-line (∼14 km) subscenes, rather than a single long image strip (Fig. 2). Stereo L1B images are often delivered in larger tiled row/column subscenes (R1C1, R1C2, etc.). Adjacent subscenes are prepared with at least 1.8 km (∼3600 lines) of overlap, and all subscenes include a right and bottom border of "empty" pixels with DN values of ∼1–3.

All DigitalGlobe L1B products have associated XML metadata files that include two sensor models to transform interior sensor coordinates to exterior world coordinates for higher-level image processing. The first is a generalized, industry-standard model involving rational polynomial coefficients (e.g., Grodechi and Dial, 2003) for ratios of 20-term, 3rd-degree, two-variable polynomials relating image coordinates (row, column) to geodetic coordinates (latitude, longitude, height). The second is a synthetic, linearized (1-D), rigorous sensor model that relates image line number to time in ephemeris/attitude tables. In practice, the rigorous model is more computationally expensive but more accurate than the RPC model, especially for scenes with significant topographic relief. The Ames Stereo Pipeline supports both models, with a default hybrid approach for optimized performance and quality—initial processing steps are performed with the RPC model, and final triangulation with the rigorous model.

The horizontal geolocation accuracy specification (direct sensor orientation, <30° off nadir, excluding terrain effects) for WorldView-1 and WorldView-2 L1B products is 5.0 m CE90 (90% of circular error, ∼1.6-sigma) and 2.3 m root-mean-square error (RMSE) (DigitalGlobe, 2014). A study involving 979 WorldView-1 images and 4412 WorldView-2 images provided observed horizontal accuracy estimates of 4.0 m and 3.5 m CE90, respectively (DigitalGlobe, 2014). Independent studies estimate <2–3 m

horizontal RMSE without ground control, and ~1 m RMSE after zero-order RPC correction with only 1–2 ground control points (GCPs) (Cheng and Chaapel, 2008).

The vertical accuracy specification for WorldView-1 and WorldView-2 L1B products (<30° off nadir, excluding terrain effects) is 5.0 m LE90 (90% of linear error) (DigitalGlobe, 2014), with observed vertical accuracy for WorldView-1 (181 stereo pairs) and WorldView-2 (160 stereo pairs) estimated at 3.7 m and 3.6 m LE90 (DigitalGlobe, 2014), respectively. An independent analysis of 50 WorldView-1 stereo pairs found 90% of measured errors for individual stereo pairs to be <3.4 m horizontal and <4.5 m vertical at 101 reference points (Dolloff and Settergren, 2010). We refer the reader to Aguilar et al. (2014, 2013) for a review of other relevant case studies.

2. Data processing

Many commercial, GUI-based software packages are capable of processing DigitalGlobe stereo image data. While sophisticated and proven, these options can be prohibitively expensive, especially for non-academic users. In addition, these options are less amenable to batch processing, as they require a trained operator to perform manual tasks (e.g. picking tie points) between "black box" processing steps. Our approach leverages mature, open-source, command-line software to process these data, which enables fully-automated processing of thousands of images in a high-performance computing environment.

The NASA Ames Stereo Pipeline (ASP) (Broxton et al., 2009; Broxton and Edwards, 2008; Moratto et al., 2010) was developed by the Intelligent Robotics Group (IRG) at the NASA Ames Research Center with sensor models for NASA planetary missions available from the open source USGS Integrated Software for Imagers and Spectrometers (ISIS) (Anderson, 2008). The planetary community has used ASP for numerous mapping applications (e.g., Beyer et al., 2010; Broxton et al., 2009; Fassett, 2016; Re et al., 2012; Shean et al., 2011; Watters et al., 2015).

The ASP code is written in C++, and leverages many 3rd-party open-source libraries. Several core algorithms are implemented in the NASA Vision Workbench (VW) library—an image-processing and computer-vision library designed to efficiently work with extremely large images. This efficiency is accomplished using lazy evaluation and a thread-safe caching system for parallel processing of image blocks (default 256 × 256 pixels). As a result, most ASP utilities are multi-threaded with limited memory usage.

In 2012, we began integrating support for rigorous DigitalGlobe sensor models and generic RPC models into the existing ASP codebase. As a result of these efforts, many new tools and improvements have been implemented in ASP/VW. The methods presented here outline our preferred mass-production workflow for WorldView-1/2 along-track stereo imagery. The pipeline has been tuned for L1B imagery of the Antarctic and Greenland ice sheets, and test cases in Washington State and Alaska. The terrestrial Cryosphere community has embraced these tools, with several recent publications leveraging ASP-derived DEMs for scientific investigations (e.g., Pope et al., 2015; Stevens et al., 2015; Willis et al., 2015). The information presented here is based on the October 2014 release of ASP v2.4.2.

Official documentation and precompiled ASP binaries for Linux and Mac OS X are available from the IRG website. The documentation contains detailed background information, sample commands, and recommended parameters for all supported sensors. The latest development source code is available in a public GitHub repository (NeoGeographyToolkit/StereoPipeline and visionworkbench/visionworkbench) under an Apache 2 license.

3. Methods: ASP processing workflow

The ASP workflow consists of several modular, command-line utilities (Fig. 2). This design involves the creation of intermediate files, which increases storage requirements, but allows the user to resume interrupted processing or bypass time-consuming steps when reprocessing. The primary processing steps (utilities) include image preprocessing (stereo_pprc), integer image correlation (stereo_corr), sub-pixel disparity refinement (stereo_rfne), disparity filtering (stereo_fltr), stereo triangulation (stereo_tri), and gridded DEM generation (point2dem). Python scripts (stereo, parallel_stereo) offer wrappers to run the full pipeline with a single command.

3.1. Input image preparation

ASP currently supports stereo processing for two input images, which are referred to as "left" (reference or "master") and "right" (source or "slave") images. All processing is performed in the original coordinate system of the "left" image.

3.1.1. L1B correction

As mentioned earlier, each WorldView-1 and WorldView-2 L1B product is derived from many separate Level 0 (L0) sub-images acquired by individual CCDs (DSAs) in the focal-plane array. These sensors are organized in two rows that are physically offset by ~0.5–1.0 cm (actual dimensions are proprietary). This geometry requires both optical distortion and along-track parallax corrections to derive the rigorous sensor model distributed with the data, which models the 2-D sensor array as a single "synthetic" 1-D line of pixels. The requisite corrections depend on acquisition parameters (i.e., TDI, image scan direction) and product resampling.

DigitalGlobe does not distribute raw L0 data products from individual DSAs and does not disclose the details of the L1B mosaic generation (e.g., seam locations, optical distortion parameters). To the naked eye, the L1B products appear seamless. Sub-pixel disparity maps derived from L1B products, however, consistently reveal sub-pixel offsets of ~0.1–0.5 pixels at DSA boundaries, with largest offsets for low TDI settings. These offsets create systematic DEM artifacts with alternating ± errors of approximately ~0.1–0.5 m (Fig. 3). Similar artifacts are observed for other sensors with CCD arrays (e.g., SPOT5, Leprince et al., 2008).

Using a large sample (~1500) of WorldView-1 and WorldView-2 along-track stereo pairs, we derived corrections to remove these offsets from L1B images. The corrections involve DSA boundary locations (initial sample number and period in pixels) and x and y offsets that minimize the total sub-pixel disparity variance for each combination of spacecraft, TDI setting, and scan direction.

The ASP wv_correct utility applies this correction to input L1B images. We apply the correction to individual subscenes before mosaicking to avoid complications that can arise from differences in subscene resampling (see Section 7.7). While this correction may not be necessary for most applications involving qualitative image analysis, it is essential for production of precise displacement maps (e.g., feature tracking to derive surface velocities), and DEM difference maps with subtle elevation change signals.

3.1.2. Subscene processing

As mentioned earlier, L1B images longer than ~30 K lines typically are delivered as multiple subscenes with unique sensor model parameters. Two processing approaches can be used to generate full orthoimages/DEMs from these split products: piecemeal and mosaicked.

The first step of the piecemeal approach involves identifying valid, overlapping "left" and "right" subscene combinations. This
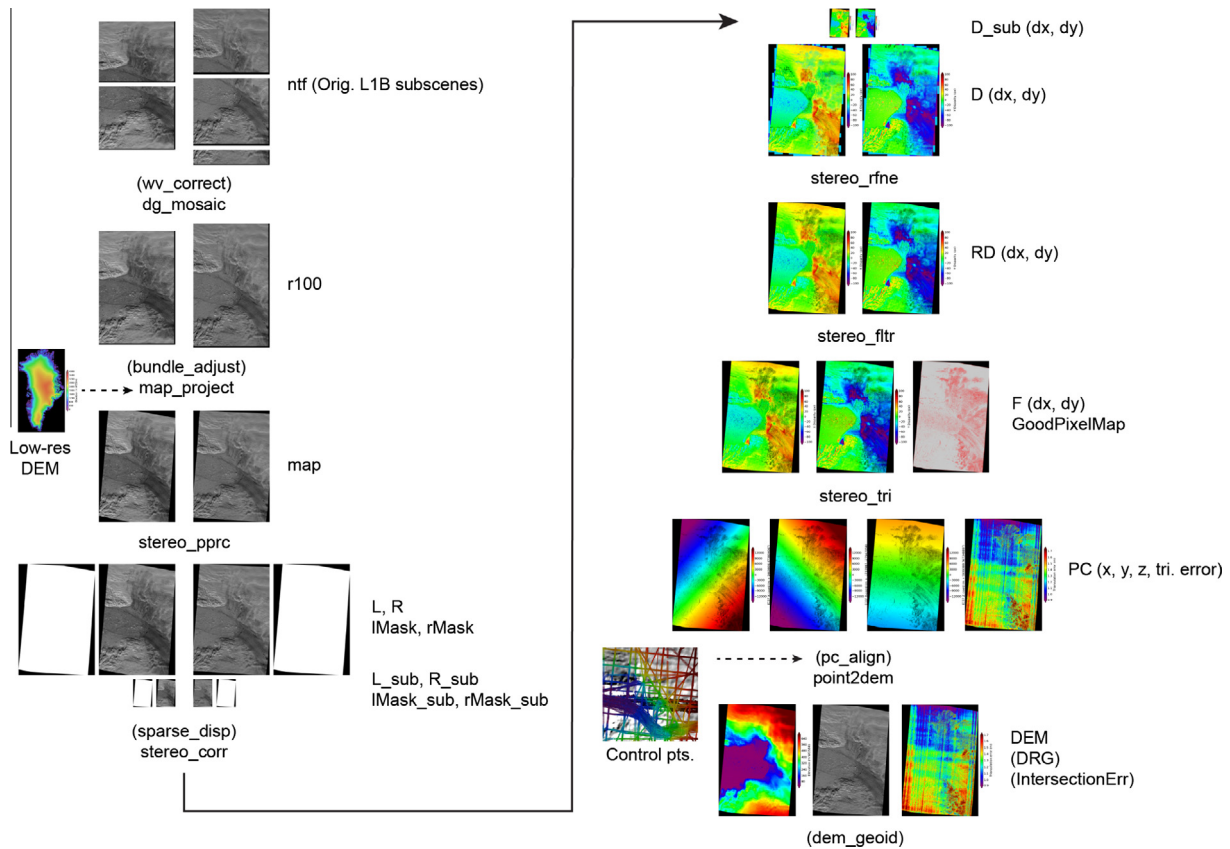
Fig. 2. ASP v2.4 workflow for mosaicked, orthorectified WorldView input imagery. Commands are listed in line with thumbnails with optional commands in parenthesis; default product extensions are listed to the right of each step, with product type listed for each band (e.g. RD band 1 = x disparity, band 2 = y disparity). Dashed arrows indicate external data input (i.e., low-res DEM for orthorectification, control points for co-registration).

is nontrivial, as the subscene boundaries are not coincident, and often a single "left" subscene will need to be processed with two or more "right" subscenes. This yields many possible subscene pair combinations, often with redundant overlap. To overcome this issue, we compute intersection area for all possible overlapping "left" and "right" subscenes, then process only those with overlap area above some threshold (e.g., ~30 km²). This is more efficient than a brute force approach to batch process all possible combinations, but can result in small residual gaps in coverage. The piecemeal approach generates output products for all subscene pairs, which must be mosaicked to reconstruct the full L1B image extent. Seamless results can be obtained by coregistering subscene point clouds (Section 5) before DEM generation (Section 3.6).

Our preferred methodology combines the individual L1B subscenes before stereo processing. The ASP dg_mosaic utility can mosaic multiple input L1B subscenes from the same parent image (with common DigitalGlobe catalog ID) to produce a single output image with combined xml metadata, including updated sensor models. While the output mosaics often have large file sizes and image dimensions (e.g., ~36,000 × 220,000 pixels for 1° geocell images), mosaicking eliminates the need to process redundant image data where individual subscenes overlap, and ASP's tile-based processing can efficiently handle these large images. We find that a nodata threshold value of ~5 successfully eliminates the empty (but nonzero) border from input subscenes.

The dg_mosaic utility generates a new RPC model for the mosaicked L1B product, which can be used for subsequent orthorectification. It should be noted, however, that RPC accuracy decreases with increasing image size, since a polynomial of the same degree is used to characterize a much larger geographic

extent. Thus, the piecemeal approach may provide improved results if the RPC sensor model is used during triangulation. Such a limitation does not apply when using the rigorous sensor model for triangulation, which is the ASP default for DigitalGlobe imagery.

3.1.3. Bundle adjustment

An optional bundle adjustment tool (bundle_adjust) can update sensor ephemeris/attitude information for two or more input images. Interest points are identified and matched for all input images, and valid matches are forward-projected to triangulate 3D points using the rigorous or RPC sensor model. The user also can provide known ground control point coordinates and corresponding image pixel locations. The 3D point locations and sensor positions/orientations are optimized using one of several solvers and robust cost functions to minimize reprojection error. The updated ephemeris/attitude information can then be used during triangulation of dense matching results.

Considering the inherent geolocation accuracy of WorldView-1 and WorldView-2, and limited observed triangulation error variance (e.g., Fig. 3), we typically bypass this step. We have found that automated co-registration of triangulated point clouds or final DEM products using a rigid-body transformation (see Section 5) can accomplish similar results, with reduced processing time and no need for manual identification of control points in input images. However, bundle adjustment before stereo reconstruction can be essential for other sensors (e.g., early planetary orbiters).

3.1.4. Input orthorectification

ASP currently supports two types of input images: (1) L1B images in original sensor coordinates (image line, sample) and
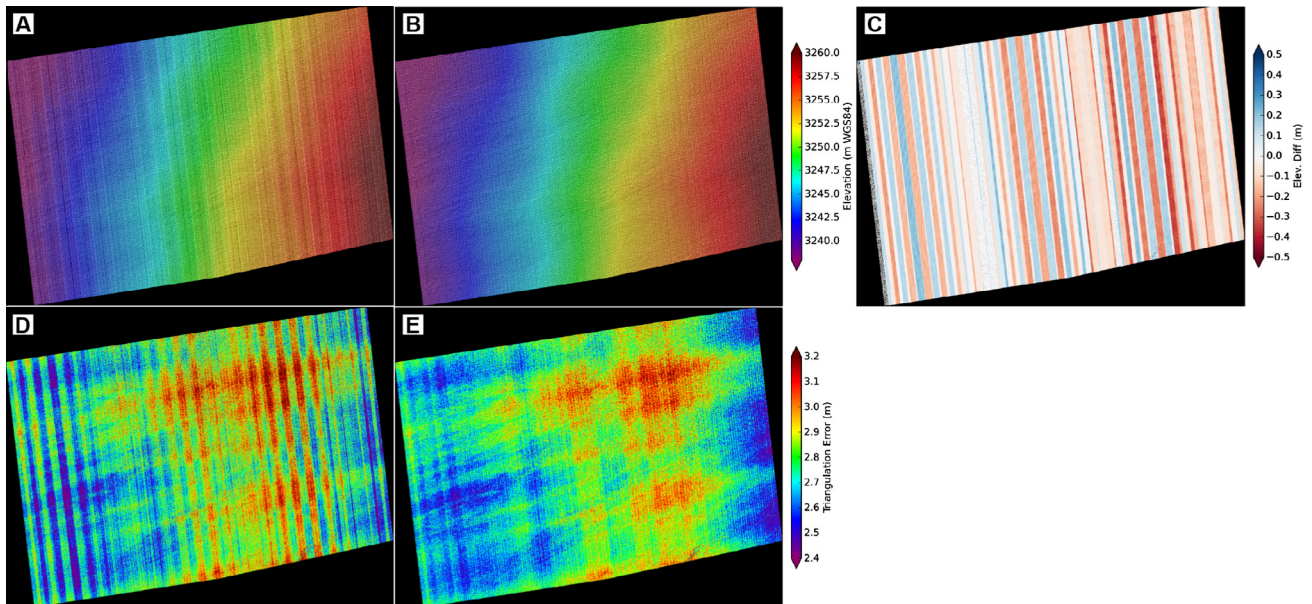
Fig. 3. DEM derived from (A) uncorrected, and (B) corrected WorldView-2 L1B input images near Summit station on July 13, 2011. (C) Elevation difference map showing magnitude of artifacts. (D + E) Triangulation error maps for DEMs in A and B.

(2) orthorectified images in real-world, projected coordinates (e.g. UTM, polar stereographic).

The unmodified L1B images preserve original image GSD, do not require existing knowledge about surface topography, and are not susceptible to distortion caused by geolocation error and/or errors in the DEM used for orthorectification. During ASP preprocessing, interest point matching is used to align the "right" image to the "left" image via a simple transformation to reduce disparity offsets. Currently supported transformations include generalized projective (Homography) and affine epipolar (AffineEpipolar). These alignment options generally work well, but may fail for scenes with significant nonplanar relief (e.g. a large, isolated stratovolcano) or limited image contrast/texture.

When an existing DEM covering the extent of input images is available, the left and right L1B images can be aligned via multi-threaded orthorectification. This process removes most of the terrain disparity signal, and the subsequent stereo processing effectively uses residual feature offsets to refine the existing DEM. Our mass-production workflow utilizes the RPC sensor model to orthorectify mosaicked images, with an appropriate projection automatically determined by input image latitude.

The optimum orthoimage resolution depends on the frequency content of the source images and the application-specific requirements for output products. Our workflow uses the minimum GSD (highest resolution, smallest off-nadir angle) of the two input images to preserve as much high-frequency texture as possible. We have found, however, that a subsampled orthoimage resolution of ∼1.0 m GSD often produces comparable results with significantly reduced data volume and processing time for WorldView imagery. For typical meter-scale ice-sheet texture, correlation success decreases substantially for subsampled input image resolutions >2.0 m GSD.

Our workflow pre-computes orthoimage extent using the intersection of corner coordinates listed in the "left" and "right" xml metadata. We also smooth the input DEM used for orthorectificaion to avoid introducing artifacts and distortion. We have successfully processed WorldView-1/2 imagery using the USGS National Elevation Dataset DEM (Gesch et al., 2002), Greenland Ice Mapping Project (GIMP) DEM (Howat et al., 2014), BEDMAP2

Antarctic Surface DEM (Fretwell et al., 2013), and various gridded airborne LiDAR DEMs.

3.2. Correlation

Correlation in ASP is performed by a generalized image correlator that attempts to match a "reference" image chip (default 21 × 21 pixels) from the "left" image with similar "source" chips extracted from a specified 2-D search window in the "right" image. The correlation is computed using an efficient caching scheme in the spatial domain rather than the frequency domain to allow correlation with missing data. Multiple cost metrics are available, but the default normalized cross-correlation offers the best results for most applications. To improve match confidence, the correlation is computed both forward (L → R chip matching) and reverse (R → L) with a configurable difference threshold (default 2 pixels).

The correlator produces dense disparity maps, where an integer disparity (x and y pixel offsets for matched chips) is computed for each valid pixel in the input image. Correlator performance is closely tied to the search window dimensions—runtime is proportional to the number of pixels in the image times the number of offsets in the search window. This scheme is efficient for small search ranges and scenes with limited relief, but can require long runtimes if significant offsets remain following initial image alignment, or if the DEM used to orthorectify the input images contains significant errors. To overcome this issue, ASP uses a Gaussian pyramid approach, iteratively performing the correlation on subsampled versions of the input images, and using the low-resolution disparities to seed finer-resolution correlation. Over ice-sheet surfaces, however, images subsampled to resolutions coarser than ∼2–8 m can appear nearly featureless, potentially causing this hierarchical search scheme to fail.

Search window dimensions can be defined globally (constant over the entire input image), or locally (spatially-variable). The latter is accomplished by initially seeding the correlator with a low-resolution map of the optimal x and y search window offsets and search window dimensions (Fig. 2). Proper seeding can effectively limit the local search window size to only a few pixels for all pyramid levels and all locations in the input image, offering significant

correlator performance improvement. The low-resolution search window maps also include a mask defining regions to be correlated, which prevents time-consuming searches over regions lacking sufficient texture (e.g., clouds and open water).

Depending on input frequency content, our workflow seeds the correlation using one of two methods: (1) dense correlation of subsampled input images (seed-mode 1), or (2) sparse, local disparities computed from full-resolution input images (seed-mode 3). The former performs well for input images with substantial low-frequency texture (e.g. scenes with significant relief and/or well-distributed albedo differences in the subsampled images), while the latter is necessary for input images with limited low-frequency contrast/texture (e.g., ice-sheet interior). Regardless of seeding method, a tile timeout option is leveraged to prevent problematic blocks from slowing overall progress.

3.3. Sub-pixel refinement

The ASP integer correlation step generates dense grids of discrete integer disparity offsets. This is a relatively crude estimate, and more precise disparity maps, potentially up to $\sim$0.1 pixel resolution, can be generated via sub-pixel refinement (e.g., Heid and Kääb, 2012). Three sub-pixel refinement algorithms are available in ASP (Broxton et al., 2009): parabolic (subpixel-mode 1), affine adaptive Bayes expectation–maximization (subpixel-mode 2), and affine adaptive (subpixel-mode 3).

With parabolic refinement, a 2-D parabola is fit to correlation cost scores within a 3×3 window around each valid pixel in the integer disparity map, and new sub-pixel disparity values are estimated at the parabola minimum. This approach is very efficient, but can suffer from "pixel-locking" artifacts in the output DEM (e.g., Shimizu and Okutomi, 2002). For highest-quality results, ASP offers a more robust sub-pixel refinement—affine adaptive with Bayesian expectation–maximization (Nefian et al., 2009). This "BayesEM" sub-pixel refinement computes a 2-D affine transformation to match the "right" and "left" image chips within a Bayesian expectation-maximum framework, offering improved results for images distorted by steep topography and noise. Additionally, ASP now includes an affine adaptive refinement option (e.g., Stein et al., 2006) without the computationally-intensive BayesEM framework, offering a compromise between output quality and processing time.

The choice of sub-pixel refinement approach is application-specific. BayesEM provides superior results for scientific analyses of small-scale topographic features with steep slopes (e.g. moraines, crevasses). Parabolic refinement should be adequate for applications that require DEM products with high accuracy but relatively coarse resolution over surfaces with limited relief. In practice, we begin with parabolic sub-pixel refinement, and preserve the integer disparity maps (D.tif) for later reprocessing with BayesEM if desired. It is also possible to limit refinement to a user-specified sub-region, which can significantly reduce runtimes.

3.3.1. Sub-pixel refinement comparison
We performed systematic tests to evaluate the performance and output quality of available ASP refinement methods. The first involved 12 WorldView-1/2 stereo pairs over volcanoes in the Pacific Northwest, with input image dimensions of $\sim$36,000 $\times$ 36,000 pixels. Wall time for BayesEM refinement was $\sim$41–114 times greater than for Parabolic refinement on a dedicated computing node (dual 8-core 2.60 GHz Intel Xeon E5-2670, 16 ASP threads). A second test was performed for a $\sim$13,000 $\times$ 13,000 pixel section from two stereo pairs acquired over supraglacial lakes on the Greenland ice sheet (Fig. 4). All refinement options were run using 12 ASP threads on a server with eight 4-core 2.27 GHz Intel Xeon X7560 processors. Results are summarized in Table 1 and Fig. 5.

Comparison of output DEMs produced with different refinement algorithms and posting (Fig. 5) shows variation in morphologic detail, especially over steep slopes and regions with increased roughness. These tests show that the BayesEM and affine adaptive refinement offer superior results over parabolic refinement, but at the expense of a $\sim$50–150$\times$ and $\sim$10–30$\times$ increase in CPU time, respectively. We observe limited differences amongst refinement methods for smooth surfaces, especially for >8–32 m posting. We also note that reducing output DEM posting by a factor of $\sim$8$\times$ eliminates most parabolic refinement artifacts and fills small data gaps. See Section 3.6 for further discussion of output product generation.

3.4. Filtering

Regardless of sub-pixel refinement approach, it is inevitable that the output disparity map will include spurious matches that lead to artifacts ("blunders") in the triangulated point cloud and gridded DEM products. The ASP workflow includes filtering algorithms to remove these problematic disparity values before triangulation. We have found that the erosion of small, isolated clusters ($\sim$32–1024 pixels) surrounded by missing data removes many problematic disparity values, with minor loss of detail. A subsequent mean difference to neighbors (filter-mode 1, default) or thresholding (filter-mode 2) can help remove residual outliers, with additional filtering options available during DEM generation (Section 3.6).

3.5. Triangulation

Triangulation combines spacecraft ephemeris/attitude information with a sensor model and known image disparity offsets to generate a 3D point cloud. The ASP triangulation routine computes 3D coordinates (in a Cartesian Earth-Centered Earth-Fixed [ECEF] coordinate system) for the closest intersection of forward-projected rays originating from the physical locations (i.e., sensor orbital position) of all matched pixels. For the rigorous DG sensor model, this is accomplished by querying the ephemeris/attitude tables and interpolating camera pose at linetimes corresponding to "left" and "right" matched pixel location in the refined, filtered disparity map. If the input images were initially orthorectified using a low-resolution DEM (Section 3.1.4), an inverse transformation (using the same external DEM from the orthorectification step) is applied to determine original L1B image coordinates of the disparities, which are then triangulated as described above.

The output is a 4-band raster point cloud file format (PC.tif). For every successfully-matched pixel in the input "left" image, band 1 contains the triangulated ECEF x-coordinate, band 2 the y-coordinate, and band 3 the z-coordinate (Fig. 2). Band 4 provides a triangulation error metric (distance, in meters, between the two rays at closest intersection) that can be used to evaluate the quality of the disparity matches, the sensor model, and ephemeris/attitude data.

3.6. Output product generation

The ASP point2dem utility converts an input point cloud (native PC.tif format, CSV, or LAS) to a gridded DEM. Multiple filters are available for outlier removal, including a triangulation error filter (percentile or absolute threshold) and threshold filter that removes outliers relative to the median of a rolling window. We have found the default percentile triangulation error filter (3 * value at 75th percentile) to be effective at removing common blunders for WorldView-1/2, eliminating the need for the threshold median filter. After filtering, a final elevation value is calculated for each output grid cell using a Gaussian weighted average of points within a
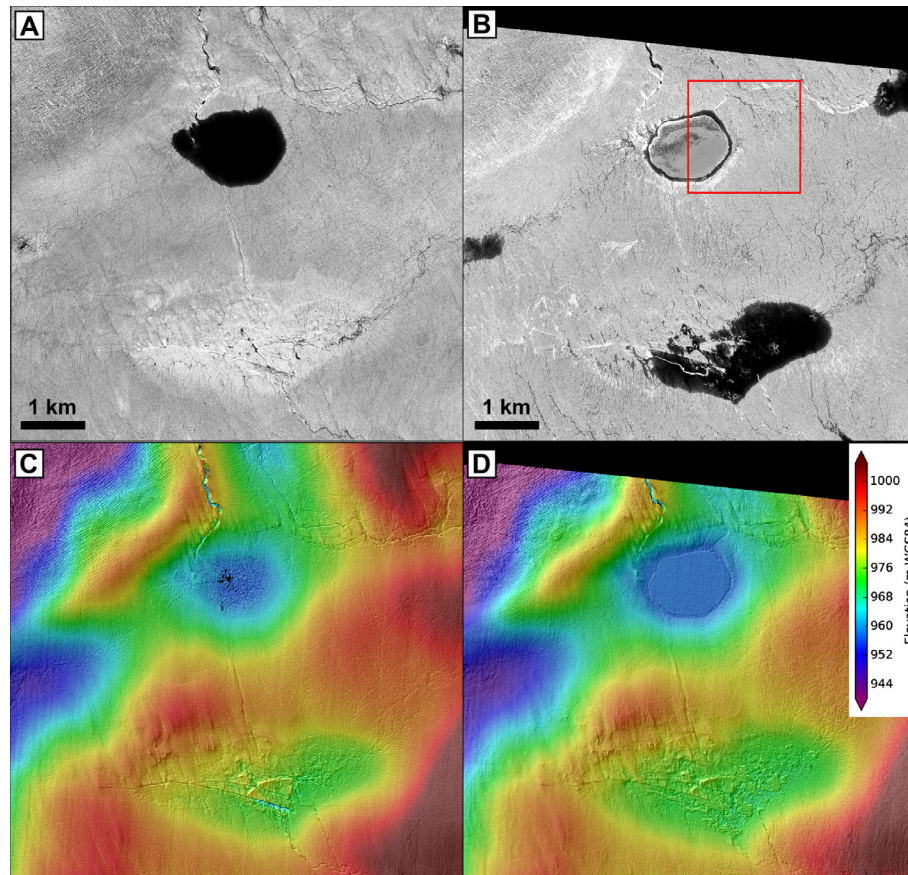
Fig. 4. WorldView-1 orthoimages and DEMs of supraglacial lakes site in West Greenland from (A + C) July 28, 2009 and (B + D) July 17, 2011. Note variable lake levels and lake ice cover, with successful correlation/triangulation of shallow supraglacial lake bathymetry and floating ice on lake surface. Outline in panel B shows extent of Fig. 5. Images © 2015 DigitalGlobe, Inc.

Table 1
Runtime comparison for ASP refinement methods. The time for "None" offers an estimate of baseline disk input/output requirement. Integer correlation wall times for the July 28, 2009 and June 17, 2011 pairs were 154 and 141 s, respectively.

Refinement method	July 28, 2009 pair wall time (s)	June 17, 2011 pair wall time (s)	Avg. ratio to Parabolic
None	46	47	0.4
Parabolic	110	110	–
Affine	2573	2737	23.4
BayesEM	16,556	16,231	150.5

specified radius (default, one cell width). Output elevation values are computed relative to the WGS84 ellipsoid and the ASP dem_geoid utility can be used to apply a geoid correction (e.g., EGM96, EGM2008) to obtain orthometric heights.

Since the along-track WorldView stereo products only involve two images, steep slopes within an acquisition-dependent aspect range are susceptible to occlusion. The point2dem utility includes optional gap-filling routines that can improve the aesthetics of output DEM and orthoimage products. Our production workflow does not use these options, however, as we wish to limit scientific analysis to elevation values derived from triangulated points.

The default point2dem output posting is similar to the "left" image resolution. Adjacent points are not necessarily independent, however, with the degree of spatial correlation dependent on input image texture content and correlator chip size. Fig. 5 shows that native (~0.5 m) DEM posting does not offer any improvement over the "4×" (~2 m) posting. Thus, we reduce output DEM posting by a factor of at least ~4, which results in artifact mitigation, noise reduction, and reduced output file size.

4. Benchmark tests

We deployed ASP on the NASA Pleiades Supercomputer, and performed benchmarking and profiling tests for the workflow outlined in Section 3. Fig. 6A shows a breakdown of runtimes for 149 WorldView-1/2 stereo pairs (typical dimensions ~36,000 × 220,000 pixels) over West Antarctica, processed using 12 ASP threads on Pleiades Westmere nodes (dual 6-core 2.93 GHz Intel Xeon X5670). These tests show that the correlation and refinement steps require the longest wall time with greatest spread due to input image variability (e.g., image quality, surface texture, clouds) and problematic tile issues discussed below.

Fig. 6 also shows the results of a benchmarking test (on a Pleiades Bridge node, 8 quad-core 2.27 GHz Xeon X7560 processors) with variable number of cores/threads (1–32) for the clipped ~13,000 × 13,000-pixel images shown in Fig. 4A. Additional cores/threads significantly reduce runtime, with diminishing performance improvements beyond ~8–10 cores/threads.

Several ASP processing steps (e.g., orthorectification, correlation, refinement) can efficiently utilize all available CPU resources for extended periods of time. However, some stages are limited by the speed at which data can be read from or written to the disk, and some steps cannot be parallelized (e.g. computing global statistics during preprocessing). Also, problematic (often featureless) tiles can delay further processing and cause temporary drops in CPU utilization due to a sequential tile-writing requirement for the tif file format. This can be mitigated with the tile timeout option, tuned VW tile cache parameters, and/or a wrapper that splits input images for parallel processing (parallel_stereo).
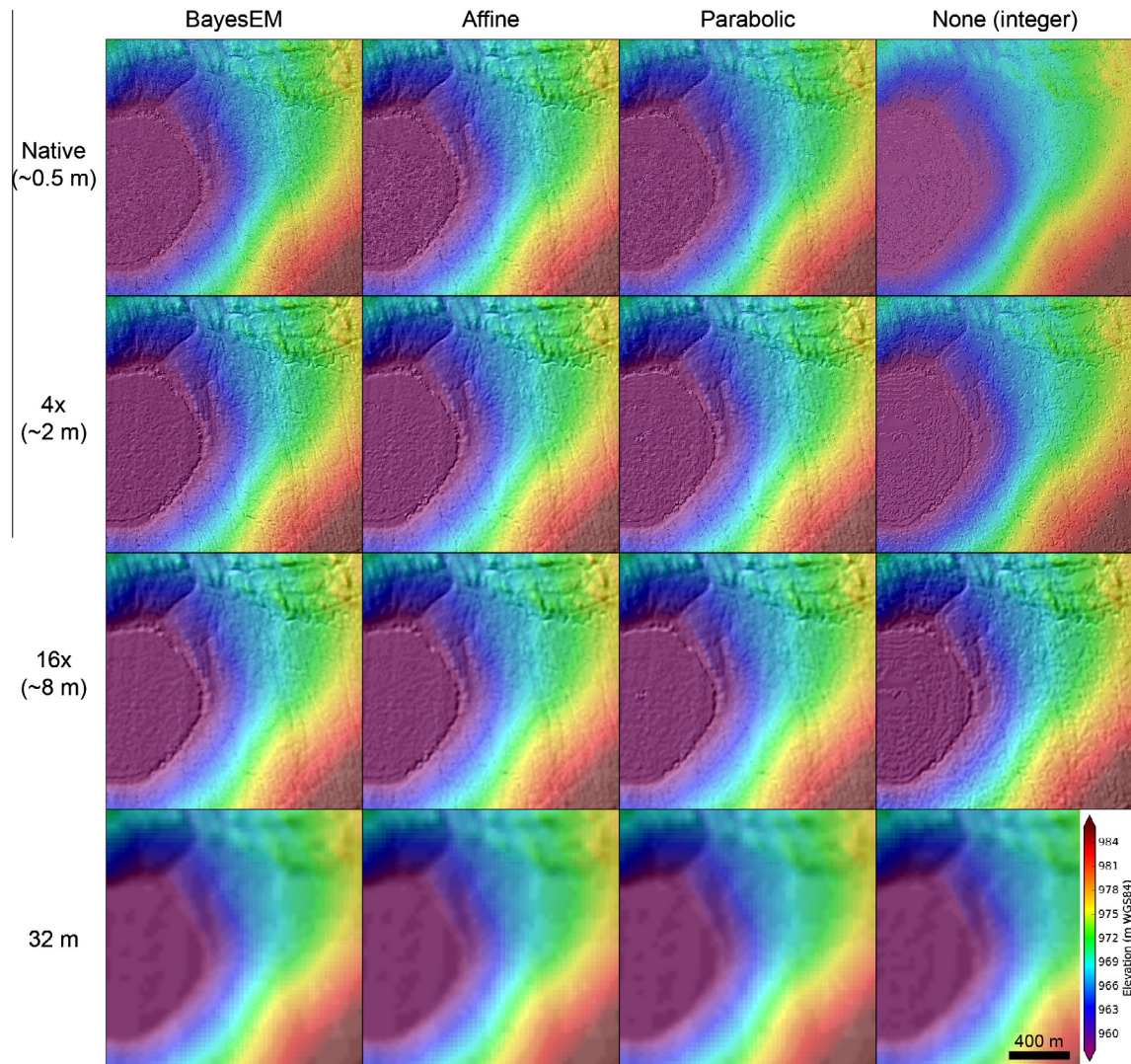
Fig. 5. Comparison of ASP sub-pixel refinement methods (columns) and output DEM posting (rows) for 1.7 × 1.7 km section of the July 17, 2011 DEM over ice-covered supraglacial lake and stream channels (see Fig. 4 for context).

5. DEM co-registration with control data

While the <5.0 m geolocation accuracy of the L1B WorldView-1/2 products is impressive, it is insufficient for many precise geodetic applications. We now consider approaches to further improve horizontal and vertical accuracy of DEM products.

Traditional photogrammetric workflows involve manual identification of tie points and control points in input imagery that are used to improve accuracy during bundle adjustment. Unfortunately, this approach does not scale for extremely large datasets, and it assumes either near-simultaneous image acquisition or a static surface—a situation that does not hold for WorldView data of the Earth's dynamic ice sheets. Numerous alternative approaches have been developed to remove offsets between overlapping gridded DEMs (e.g., Berthier et al., 2007; Noh and Howat, 2014; Nuth and Kääb, 2011). Here, we present a generalized, automated co-registration workflow for the ice sheets that relies on accurate, temporally- and spatially-coincident control data (i.e., GPS, airborne/terrestrial LiDAR point clouds, existing gridded DEMs).

The ASP pc_align utility automatically co-registers an input "source" file (point cloud or gridded DEM) to available "reference" control data. This is accomplished using a point-to-plane or point-to-point iterative closest point (ICP) algorithm (Pomerleau et al., 2013) that iteratively improves the transformation required to minimize offsets. The default point-to-plane algorithm works best for control data with adequate spatial distribution over surfaces with sufficiently-variable slope and aspect. Alternatively, the point-to-point ICP algorithm can offers superior co-registration results for smooth, planar surfaces like those over the ice sheet interior. The ICP output is a 3-D transformation (3 translation and 9 rotation terms) in the ECEF coordinate system, which can optionally be limited to a 3-parameter translation without rotation. The utility optionally applies this transformation to export a corrected 4-band ASP point cloud.

We have found that a simple translation (i.e. removal of constant horizontal and vertical bias) is almost always sufficient to correct WorldView-1 or WorldView-2 PC/DEM products. Based on this finding, it follows that only a small number of control points are required for co-registration (e.g., Cheng and Chaapel, 2008). We have also observed, however, that a limited number of the larger DEM products (e.g., ~17 × 110 km geocell pairs) display an along-track, planar tilt of ~1–3 m ($\sim$0.5–1.0 × 10^{-3} degrees) and/or an occasional ~1 m cross-track tilt. For these situations, control data should be well distributed throughout the scene to constrain an appropriate rotation correction.
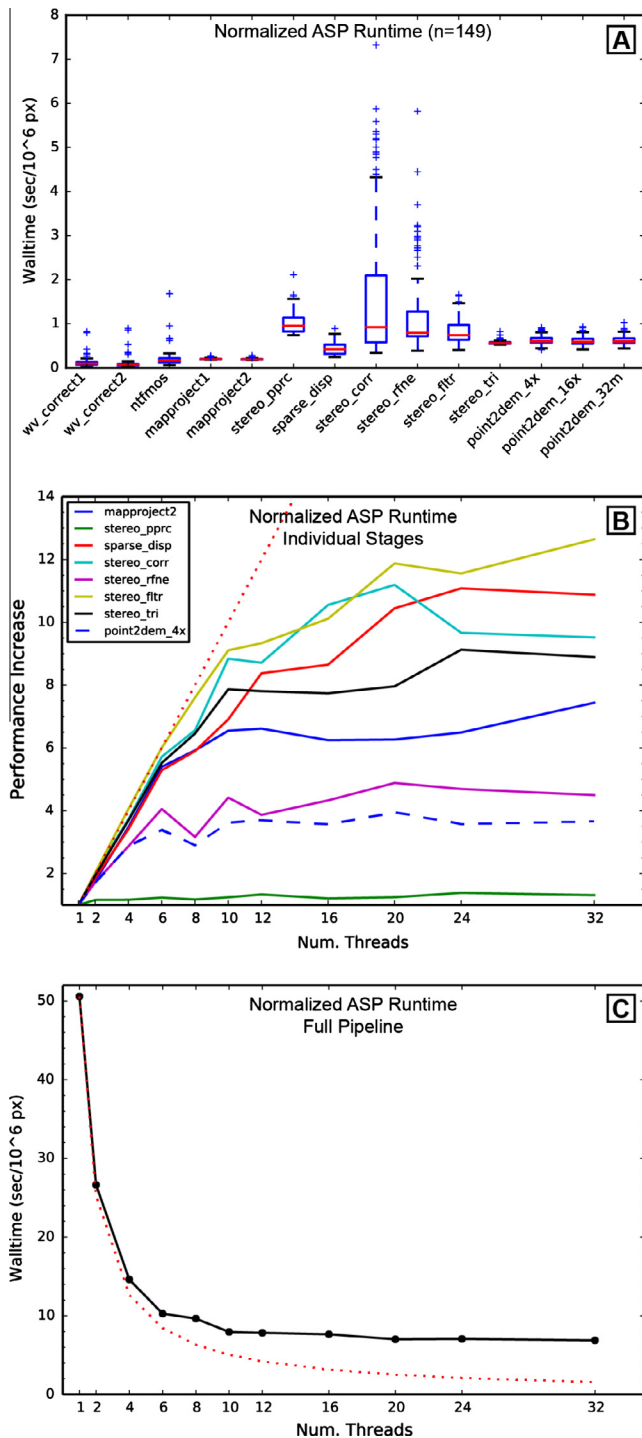
Fig. 6. Benchmark test results. (A) Area-normalized runtime of individual ASP stages for 149 pairs with variable dimensions/quality. Box and whisker plot shows median runtime, inner quartile (box) and 1.5*IQR whiskers (ticks) with fliers as crosses. (B) Area-normalized performance increase (multiplicative factor relative to single-threaded case) vs. number of ASP threads (set to number of physical CPU cores) for each ASP stage on clipped July 28, 2009 input images (Fig. 4A). Dotted line is idealized linear relationship. Each thread count test was run once on shared hardware resources, which explains some of the variability. (C) Area-normalized performance vs. number of ASP threads when running full pipeline on clipped July 28, 2009 input images. Dotted line is idealized 1/n relationship.

5.1. Ice sheet control points

We have compiled a comprehensive database of available control data for Antarctica and Greenland (Table 2). The primary

sources include NASA ICESat-1 Geoscience Laser Altimeter System (GLAS) data (Schutz et al., 2005; Zwally et al., 2002), NASA Airborne Topographic Mapper (ATM, (Krabill et al., 2002; Martin et al., 2012)) and Land, Vegetation, and Ice Sensor (LVIS, (Blair et al., 1999; Hofton et al., 2008)) airborne LiDAR data, all available from the National Snow and Ice Data Center. These data are typically collected as annual or seasonal campaigns spanning relatively short time windows (weeks–months), with data acquisition for a particular site typically occurring over ~1–3 days.

We filter the dense L1B (qfit format) ATM data to remove unwanted returns (e.g., crevasse floors). Median elevation values are computed for 10-m grid cells after removing outliers (20–80th percentile pass for rolling 20-m window). For each of these points, a plane is fit to all points within a 50 m radius, and the point is preserved if it falls <10 m from this plane. Additional filters for GLAS data remove points with low uncorrected reflectivity (reflctUC < 0.025) and high waveform misfit (iceSVar > 0.04 mV).

Our automated workflow queries the filtered control point database for an input DEM extent and extracts independent surface velocity data (Joughin et al., 2010; Rignot et al., 2011) for the same extent. For each valid control point within this extent, an estimated displacement ($\|\mathbf{v}\| \cdot dt$) is computed from the sampled 2-D velocity vector $\mathbf{v}$ and the time offset dt between the control point timestamp and DEM timestamp (Fig. 7). The point is discarded if the estimated displacement exceeds some threshold (e.g., 5.0 m), or the time offset exceeds a fixed threshold (e.g., 120 days). The remaining points tend to be clustered over static surfaces (e.g. exposed bedrock) and "dynamic control surfaces" (e.g., low-velocity ice with limited surface slope and roughness).

An absolute elevation filter (e.g., >10 m above sea level) and ice-shelf mask removes points over floating ice, while an absolute elevation difference filter (e.g., $|z_{GCP} - z_{DEM}| < 30$ m) excludes outliers. Site-specific filtering parameters are set according to local control point availability and prior knowledge of local ice sheet dynamics. If a sufficient number of control points with sufficient spatial distribution remain, they are used to correct the DEM and estimate accuracy.

Looking forward, the NASA ICESat-2 mission (Abdalati et al., 2010) is slated for launch in 2017. The multi-beam laser altimetry data will offer near-contemporaneous, global control data with ~10 cm accuracy for all subsequent WorldView stereo DEMs. These data will eliminate the need for complex control point filtering algorithms and provide robust accuracy estimates for DEMs over dynamic surfaces, regardless of static control surface availability.

6. Accuracy analysis

We now present two case studies to estimate relative and absolute horizontal/vertical accuracy for WorldView-1 and WorldView-2 DEMs generated using ASP. Each involves a different setting, with variable surface conditions and control data availability.

6.1. Summit, Greenland case study

We obtained all available 2010–2013 WorldView stereo image pairs with <75% cloud-cover for Summit, Greenland, where the National Science Foundation maintains Summit Station (Figs. 7 and 8). This location is just west of the ice-sheet divide and has extremely low surface slopes (<0.2°), surface velocities (~3–5 m/yr, Fig. 7B), and annual accumulation rates (~0.2 m water equivalent per year). Recent studies of accumulation rates, firn compaction rates, and ice dynamics near Summit show that surface elevations remained effectively constant from ~2010 to 2013 (R.L. Hawley, personal communication, 2015). These

Table 2
Elevation control sources for the Earth's ice sheets.

Instrument (data product)	Instrument type	Observation period	Shot diameter (m), swath width	Shot density	Horizontal accuracy (m)	Vertical accuracy (m)
GLAS (L2,GLA12)	Satellite laser altimeter	2003–2009	~65 m shot	172 m along-track	~6–20	~0.1–0.5
ATM (L1B)	Conical-scanning airborne LiDAR	1993-present	30–45° swath	1 pt/10 m^2	~0.75	<0.1
LVIS (L2)	Swath-scanning airborne LiDAR	2009-present	~10–25 m shot, 12° swath	1 pt/400 m^2	<2	<0.1

Fig. 7. Control point filtering. (A) Summit DEM from April 9, 2013. (B) Surface velocity map (Joughin et al., 2010) with all ~3.7 × 10^6 available ground control points from 1999 to 2014 (black). Observed velocity variability is within measurement noise for this location. (C) Culled ground control points (~3.3 × 10^5) with colors indicating expected displacement ($\|v\| \cdot dt$) for time offset dt between DEM and control point acquisition. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
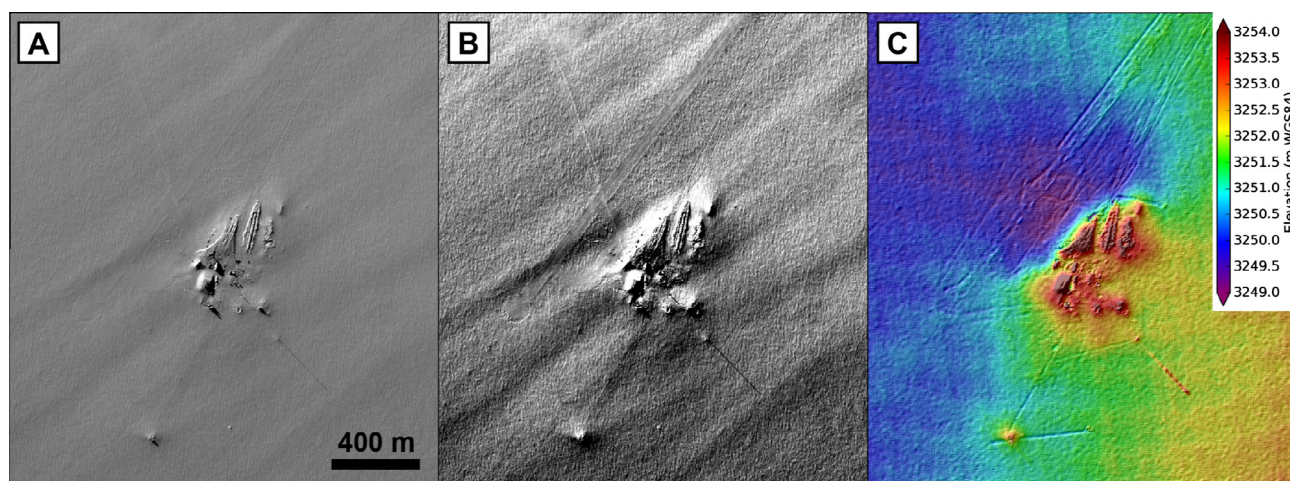

Fig. 8. WorldView-1 image of Summit Station, Greenland on July 14, 2013 with (A) 0.1–99.9% stretch and (B) 2–98% stretch. Note presence of meter-scale texture (sastrugi). (C) DEM generated with BayesEM refinement and posted at ~2 m. Note color ramp range of 5 m vertical meters, and the presence of a broad depression to the northwest of the station. Linear artifacts are the result of residual pixel locking on snowmobile tracks, runway margins, and tower shadow. Images © 2015 DigitalGlobe, Inc. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

characteristics make Summit an ideal calibration site for satellite observations of the Greenland ice sheet (e.g., Siegfried et al., 2011).

We processed WorldView-1 (n = 3) and WorldView-2 (n = 11) stereo pairs using the workflow outlined in Section 3, with seed-mode 3, parabolic refinement, and ~2 m output DEM posting. A total of 3.73 × 10^6 control points spanning 1999–2014 were extracted for the area covered by these DEMs (Fig. 7B). The control points were filtered with a maximum time offset dt of 1.5 years and a maximum $\|\mathbf{v}\| \cdot dt$ displacement of 10 m. Final control point samples included 8.4 × 10^4 to 3.3 × 10^5 points, with >2–5 non-parallel flightlines available for each DEM (Fig. 7C). A random sample of 10^5 points was used for co-registration, with 75% of these points considered inliers during ICP. The points not used for co-registration can serve as independent check points, although their

spatial distribution is nearly identical along flightlines, and we compute final error estimates using all original filtered control points for each DEM.

Co-registration was performed with point-to-point ICP, a maximum displacement setting of 20 m, and final transformation limited to a translation (no rotation). The high-density DEMs (~0.25 pts/m^2) were set as the "reference" for co-registration, with low-density control points (~0.01 pts/m^2) as the moveable "source." The DEMs were then corrected using the inverse of the final ICP solution. These ECEF translation vectors were converted to a local stereographic projection (Fig. 9A) and the resulting horizontal and vertical offsets were used to compute CE90 and LE90 for the 14 DEM sample using standard formulas (Federal Geographic Data Committee, 1998):

$$RMSE_{j=\{x,y,z\}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\Delta j_i^2} \qquad (1)$$

$$CE90 = 2.146 \cdot (RMSE_x + RMSE_y)/2 \qquad (2)$$

$$LE90 = 1.6449 \cdot RMSE_z \qquad (3)$$

Additionally, errors $\Delta z_i = (z_{GCP_i} - z_{DEM_i})$ and absolute errors $|\Delta z_i|$ at all control points were computed for each DEM, both before and after the transformation. These errors rarely display normal distributions and outliers can skew traditional accuracy measures (e.g., RMSE, standard deviation), so robust statistics (e.g., Höhle and Höhle, 2009) were computed to further characterize ASP/WV uncertainty. For each DEM, the median (50th percentile), 16th and 84th percentile of signed errors was computed (Fig. 9B), as was the normalized median absolute deviation (NMAD):

$$m_{\Delta z} = median(\Delta z_i) \qquad (4)$$

$$NMAD = 1.4826 \cdot median(|\Delta z_i - m_{\Delta z}|) \qquad (5)$$

which provides a robust estimate of standard deviation.

6.1.1. Absolute vertical accuracy

The observed 5.22 m LE90 for uncorrected DEMs slightly exceeds the 5.0 m CE90/LE90 specification published by DigitalGlobe (Table 3), although 6 of these pairs contain at least one image acquired at off-nadir angles >30° (Fig. 9C). One DEM (April 4, 2013) acquired with a combination of relatively high off-nadir

angles (37.7°, 26.1°) and low convergence angle (33°) displayed vertical bias of +8.83 m (Fig. 9C). We note that if this DEM is excluded, LE90 drops to 3.61 m (Table 3) for the remaining 13 DEMs acquired with more favorable geometry.

The sample of Summit DEMs shows an apparent vertical bias of +2.1 m, with uncorrected WorldView DEMs higher than control points (Fig. 9A and B, Table 3). There are many possible factors that could contribute to such a bias (e.g., sensor model error, preferred acquisition geometry for a particular geographic location, a small systematic error in the ASP code). Fortunately, we have run enough test cases to confirm that this +2 to +3 m vertical bias appears systematic, and can be systematically removed from uncorrected DEM products before co-registration and accuracy analysis. Removing this bias reduces LE90 to 3.91 m (Table 3).

Fig. 9B shows that ICP co-registration successfully removes vertical bias for all Summit DEMs (Table 3). The NMAD values remain the same, as horizontal corrections are small, with limited potential for improvement over near-planar surfaces with limited slope/aspect variance. After co-registration, both RMSE and NMAD are ~0.2 m for all DEMs (Table 3).

6.1.2. Relative vertical accuracy

We now consider residual errors in the co-registered Summit DEMs. The standard deviation of elevation values computed for every pixel in a "stack" (n = 14) of overlapping DEMs ranges from 0.1 to >0.5 m (Fig. 10B), with a mean value of 0.19 m. If we assume that residual translation offsets are negligible and the surface is not changing over time, then these values provide a measure of
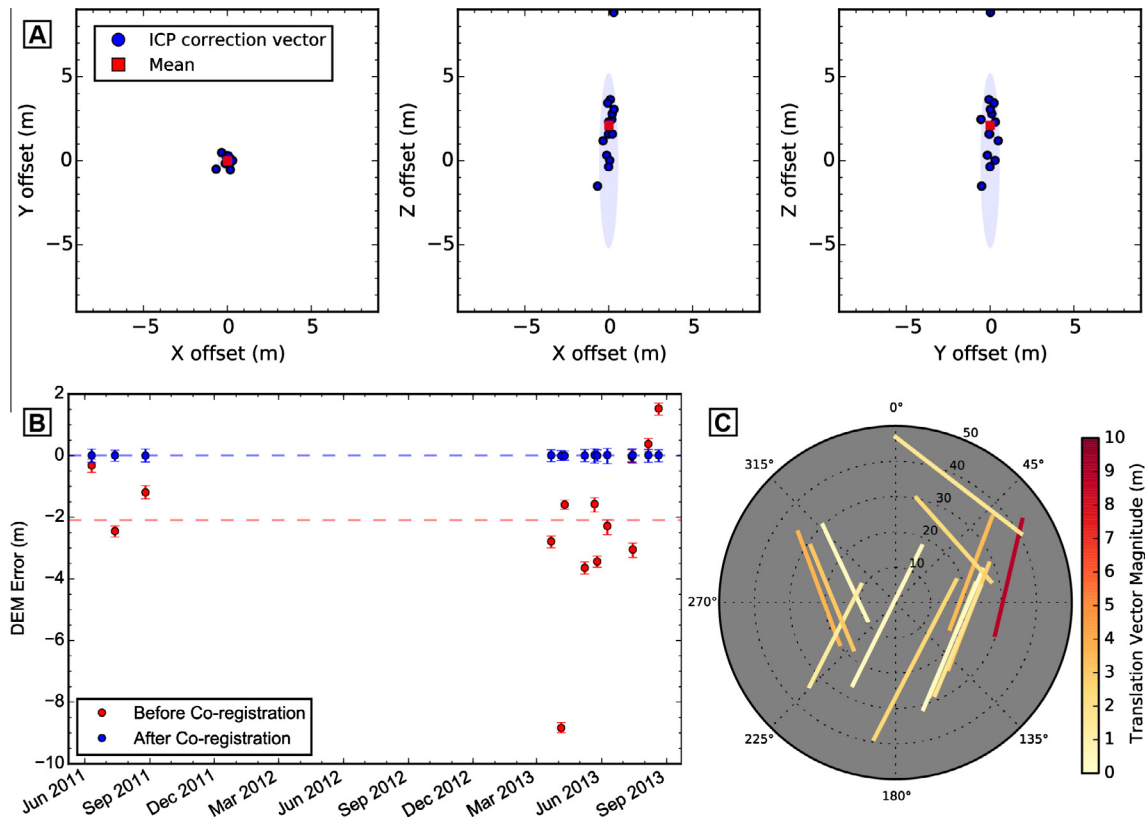

Fig. 9. (A) Components of ICP translation vectors required to correct each Summit station DEM to filtered control point data. Mean vertical offset for all 14 DEMs is +2.1 (DEMs 2.1 m above control data). Shaded ellipses show CE90 and LE90. Note presence of April 4, 2013 outlier with significant z translation of 8.8 m. (B) Median of error ($\Delta z_i = z_{GCP_i} - z_{DEM_i}$) sampled at all ($\sim 8.4 \times 10^4$–$3.3 \times 10^5$) control points, before and after ICP co-registration. Error bars show 16th–84th percentile spread (robust estimate of ±1-sigma) and dashed horizontal lines show mean of n = 14 sample. (C) Stereopair baseline geometry for Summit DEMs. The endpoints of each line represent the spacecraft azimuth and elevation of the two stereo images comprising the pair. Color represents the magnitude of the ICP translation vector required to co-register the DEM with the control data. Note relatively short baseline and high off-nadir angles of outlier. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3

Co-registration results and error analysis for Summit and Tracy sites. LE90 values are computed before/after removing mean vertical offset (see text for discussion). The Summit $n = 13$ values are computed after removing April 4, 2013 DEM with poor acquisition geometry. The final four rows provide an average of the specified metric for all DEMs in the sample.

	Summit		Tracy
Sample size	14	13 (outlier removed)	17
Mean vertical offset (m)	+2.10	+1.58	+2.52
CE90 (m)	0.57	0.58	3.61
LE90 before/after removing bias (m)	5.22/3.91	3.61/2.51	5.04/2.86
RMSE before co-reg (m)	2.40	1.90	2.83
RMSE after co-reg (m)	0.20	0.21	0.96
NMAD before co-reg (m)	0.20	0.20	0.48
NMAD after co-reg (m)	0.20	0.20	0.44

relative DEM accuracy. Residual error is related to "jitter" artifacts, DSA boundary artifacts, pixel-locking artifacts, "blunders" due to spurious disparities, and other sources of measurement noise.

Qualitatively, the map of standard deviation displays several characteristics that warrant discussion. The first involves primarily along-track undulations that we attribute to independent spacecraft "jitter" artifacts in each DEM (Fig. 10B). The orientation of these artifacts is variable, as the available DEMs include both ascending and descending acquisitions, and artifact amplitude appears dampened over areas with higher sample count. Section 7.6 offers further discussion of these artifacts.

Standard deviation values appear to increase near the lateral margins of some input DEMs (Fig. 10B). This likely involves increased error in the sensor geometry and/or optical distortion correction near the edges of the sensor, especially for certain TDI/scan-direction combinations. This variability also affected the derivation of L1B corrections applied by the wv_correct utility (Section 3.1.1), which could compound observed error near lateral DEM margins.

6.2. Tracy Glacier region, Greenland case study

The Summit analysis involved smaller DEMs over near-planar surfaces with extremely low slopes and surface velocities, effectively offering a "best case" scenario for DEM accuracy evaluation.

The following analysis includes larger DEMs over an area near the Greenland coast with variable relief, slope, and aspect.

We obtained all available 2012–2013 WorldView stereo image pairs with cloud-cover <75% in a 1°-latitude geocell near Tracy Glacier in Northwest Greenland (Fig. 11). This location covers several outlet glaciers, fjords, small ice caps, and exposed bedrock. We processed 17 overlapping ∼17–110 km WorldView-1 ($n = 9$) and WorldView-2 ($n = 8$) stereo pairs using the same methodology described in the Summit case study.

A total of 4.4×10^7 filtered database control points spanning 1999–2014 were available for co-registration (Fig. 12B). For this analysis, we limited control data to static bedrock surfaces. A "RockMask" of ice-free surfaces was generated from the 90 m Greenland Ice Mapping Project (GIMP) IceMask and OceanMask datasets (Howat et al., 2014). The extent of this mask was further reduced with three 1-pixel erosion iterations to exclude pixels near ice margins and shorelines. After applying this RockMask and clipping to individual DEM extents, the total number of control points per DEM ranged from ∼8.5×10^5 to ∼1.5×10^6, with broad spatial distribution over bedrock surfaces (Fig. 12B).

We expect some surface elevation variability due to snow accumulation on bedrock surfaces during control point and/or DEM acquisition. Snow depth measurements at Thule airport varied from 0 to ∼0.5 m from 2012 to 2013 (Czimczik, 2014). Additionally, valleys tend to accumulate windblown snow and can preserve snow longer into the melt season. Thus, we might expect increased elevation variance near steep valley walls. To mitigate these effects, we did not apply a time offset filter—final control data included all available points from 1999 to 2014, with acquisition during different times of year (Fig. 12C). This approach, combined with the 25% outlier removal during ICP, should limit the influence of points acquired when seasonal snow was present on bedrock surfaces.

6.2.1. Results

As with Summit, the ICP co-registration successfully removes vertical bias for all input DEMs (Fig. 13, Table 3). In addition, horizontal corrections reduce the Δz_i error spread for all DEMs (Fig. 13B), with sample ($n = 17$) average NMAD of 0.44 m after co-registration (Table 3). Standard deviation of elevation values in the stack of overlapping DEMs range from ∼0.1–0.5 m over flat bedrock, ∼1–2 m over ice-covered lakes, and >2–5 m over dynamic
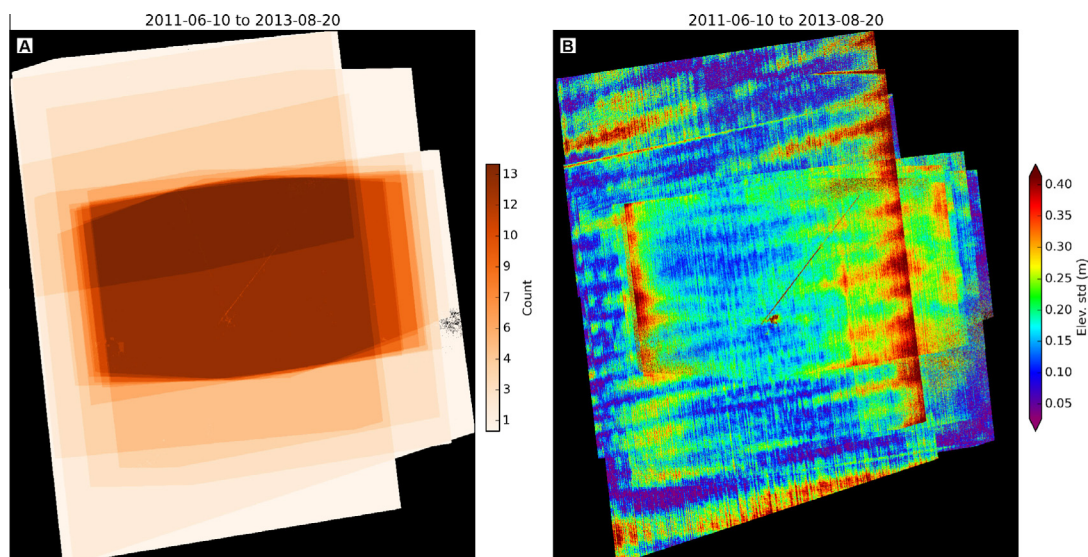

Fig. 10. Maps of (A) pixel count, and (B) standard deviation for stack of co-registered Summit DEMs ($n = 14$).
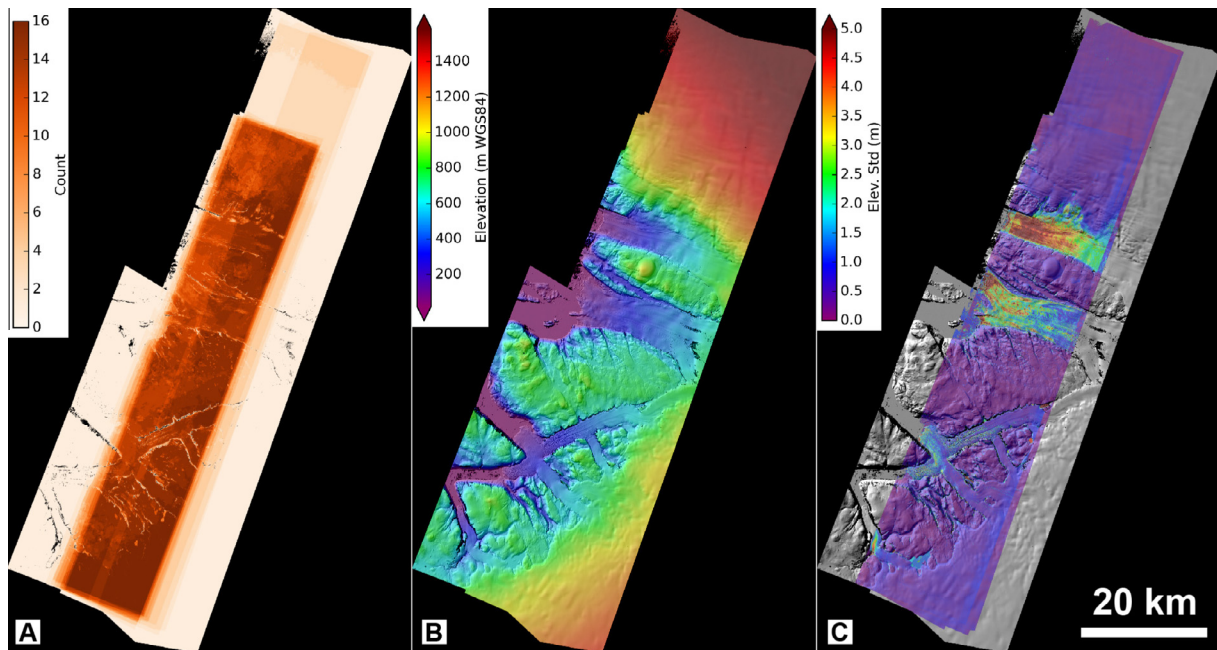
Fig. 11. Maps of (A) pixel count, (B) median elevation, and (C) standard deviation for stack of co-registered Tracy Glacier geocell DEMs (*n* = 17). Note high standard deviation over dynamic outlet glaciers. See Fig. 14 for further analysis of error over bedrock.
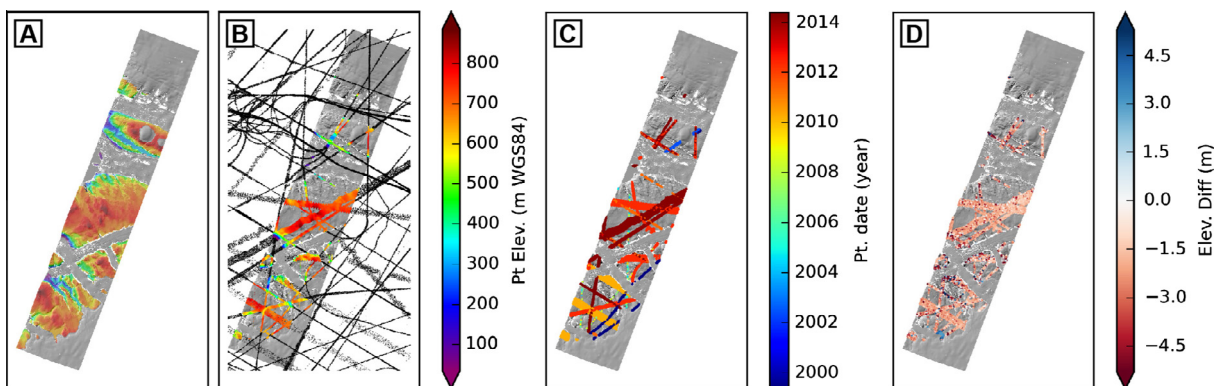

Fig. 12. Tracy Glacier geocell control point filtering for April 25, 2013 DEM. (A) Shaded relief map with colored elevation values over bedrock surfaces. (B) All available ∼4.4 × 10⁷ control points (black) and 1.3 × 10⁶ filtered points over bedrock (color). (C) Filtered point timestamps (min = 5/10/1999; max = 5/20/2014). (D) Elevation difference between control points and DEM before co-registration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

outlet glaciers (Fig. 11C). The latter provides a sample of the seasonal and interannual ice thickness change signals motivating much of this work.

6.2.2. Slope-dependent accuracy

The mean of standard deviation values for the stack of co-registered DEMs is 0.46 m over bedrock (Fig. 14B). An analysis of stack standard deviation values that fall within discrete 1.0° slope bins shows an apparent linear relationship between DEM error and surface slope (Fig. 14C). The bin median for slopes up to ∼35° is <1.0 m, although the spread within each bin increases with increasing slope. Bin median values of ∼0.2–0.4 m are observed for <10° slopes, which should be representative of most ice sheet surfaces. These results are consistent with a similar error vs. slope analysis for photogrammetrically-derived DEMs over mountainous terrain (Müller et al., 2014).

We note that the input DEMs for this test were generated with ASP's Parabolic sub-pixel refinement, which can introduce "step" artifacts over steep slopes (Fig. 5). Reprocessing with BayesEM

refinement would likely decrease observed error over steeper slopes, which would decrease the slope of the linear fit in Fig. 14C.

7. Limitations and discussion

Along-track stereo WorldView imagery offers an exciting, high-resolution dataset for Earth science applications. As with any remote sensing technology, however, these data and methods are subject to several limitations, which we now address.

7.1. Atmosphere

First and foremost, these are optical data, and successful image correlation requires a clear view of the surface. Opaque clouds in the scene cause DEM data gaps, but we have found that partial atmospheric obfuscation (e.g. thin clouds, smoke, haze, etc.) has essentially no impact on output DEM quality, as long as sufficient surface texture is visible through the clouds to allow correlation.
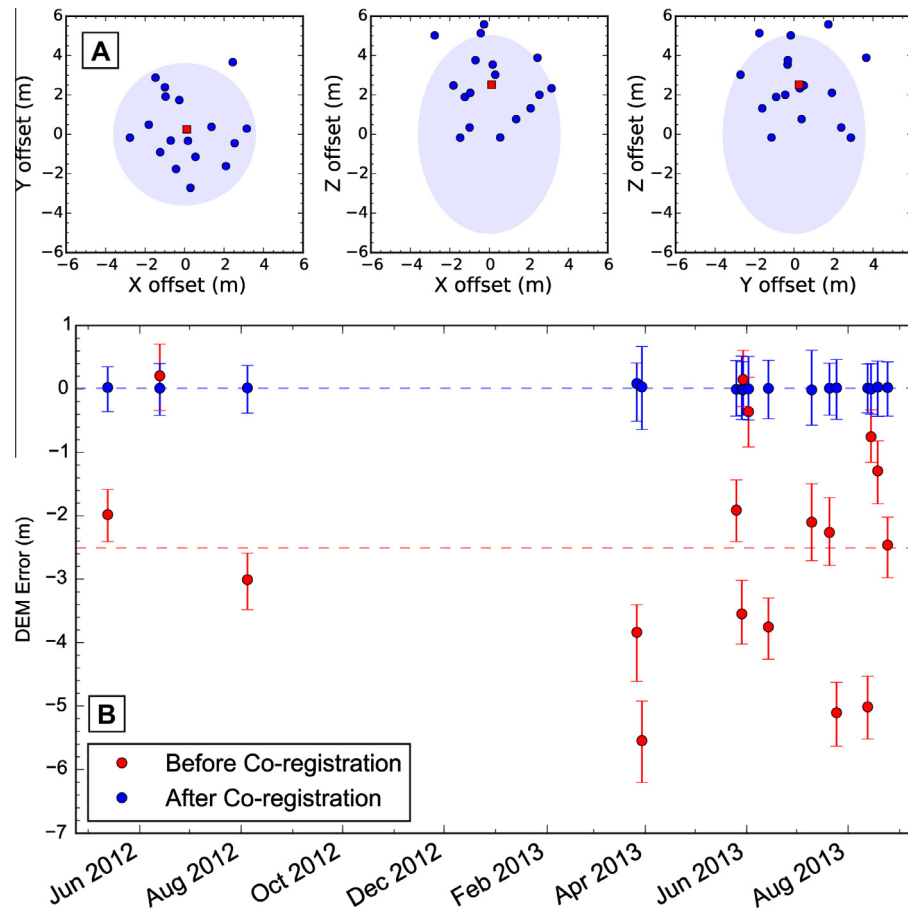
Fig. 13. (A) ICP translation vector components, and (B) median error (with 16th–84th percentile spread) for Tracy Glacier DEMs. See Fig. 9 caption for details.

Unlike commercial photogrammetric software options (e.g. SOCET SET), ASP does not currently include corrections for atmospheric refraction, as it was originally developed for NASA planetary orbiters around airless (or nearly airless) bodies. However, this effect should be negligible for WorldView-1/2 altitude and typical off-nadir angle range.

7.2. Water

We have found that it is possible to correlate bathymetric surfaces in shallow (∼1–5 m), clear water (e.g., surface meltwater lakes on the Greenland ice sheet, Fig. 4). Accurate triangulation of subaqueous surfaces, however, requires a localized refraction correction—functionality that has not yet been implemented in ASP v2.4. Without this correction, shallow subaqueous surfaces will have an apparent positive vertical offset, with magnitude dependent on water depth and pair geometry.

Deep open water will almost always fail to correlate, especially when surface waves or sun glint are present. The primary exception involves floating surface features such as sea ice or flotsam, which can provide acceptable correlation results (e.g., Fig. 4). One must use caution during interpretation, however, as these floating features may experience significant (>0.1–5.0 px) displacements due to waves or surface winds/currents over the ∼60–90 s interval between along-track image acquisition. If ignored, these displacements will produce anomalous topographic signals, sometimes 10's to 100's of meters depending on displacement magnitude and pair geometry. While this situation is not ideal for surface reconstruction, it may be valuable for other applications, as it is possible to measure these relatively rapid displacements through feature-tracking (e.g., Kääb and Leprince, 2014).

7.3. Vegetation

Samples of WorldView-1/2 stereo data over glaciers in the Pacific Northwest with nearby forests and meadows confirm that correlation success is typically near-perfect over exposed rock and ice, with increased noise and data gaps over vegetated surfaces. Stereo geometry and vegetation characteristics (e.g., density, height, shape, and spacing) will affect correlation success. For sparse, low-lying vegetation (e.g. winter brush), it often is possible to image the underlying ground surface, ultimately providing a digital terrain model (DTM) of surface elevations. For dense vegetation, successful correlation is still possible, but the resulting digital surface model (DSM) will include vegetation/canopy. While undesirable for many geoscience applications, these measurements have value for forestry and other biomass inventory applications, especially when existing bare-earth DTM data are available.

7.4. Image saturation and shadows

The 11-bit WorldView-1 and WorldView-2 sensors offer excellent dynamic range and image contrast. With appropriate sensor TDI/gain settings during acquisition, high signal-to-noise ratios are possible for both high (e.g., snow, ice) and low (e.g., basaltic rock) albedo surfaces in the same scene. The same is also true for scenes with significant relief, where both illuminated slopes and shadows contain sufficient contrast for successful correlation. While continuous DEMs can be produced for these scenes (excluding occluded areas), we have noted artifacts along some shadow edges (e.g., Fig. 8) due to the pixel-locking phenomenon discussed in Section 3.3.
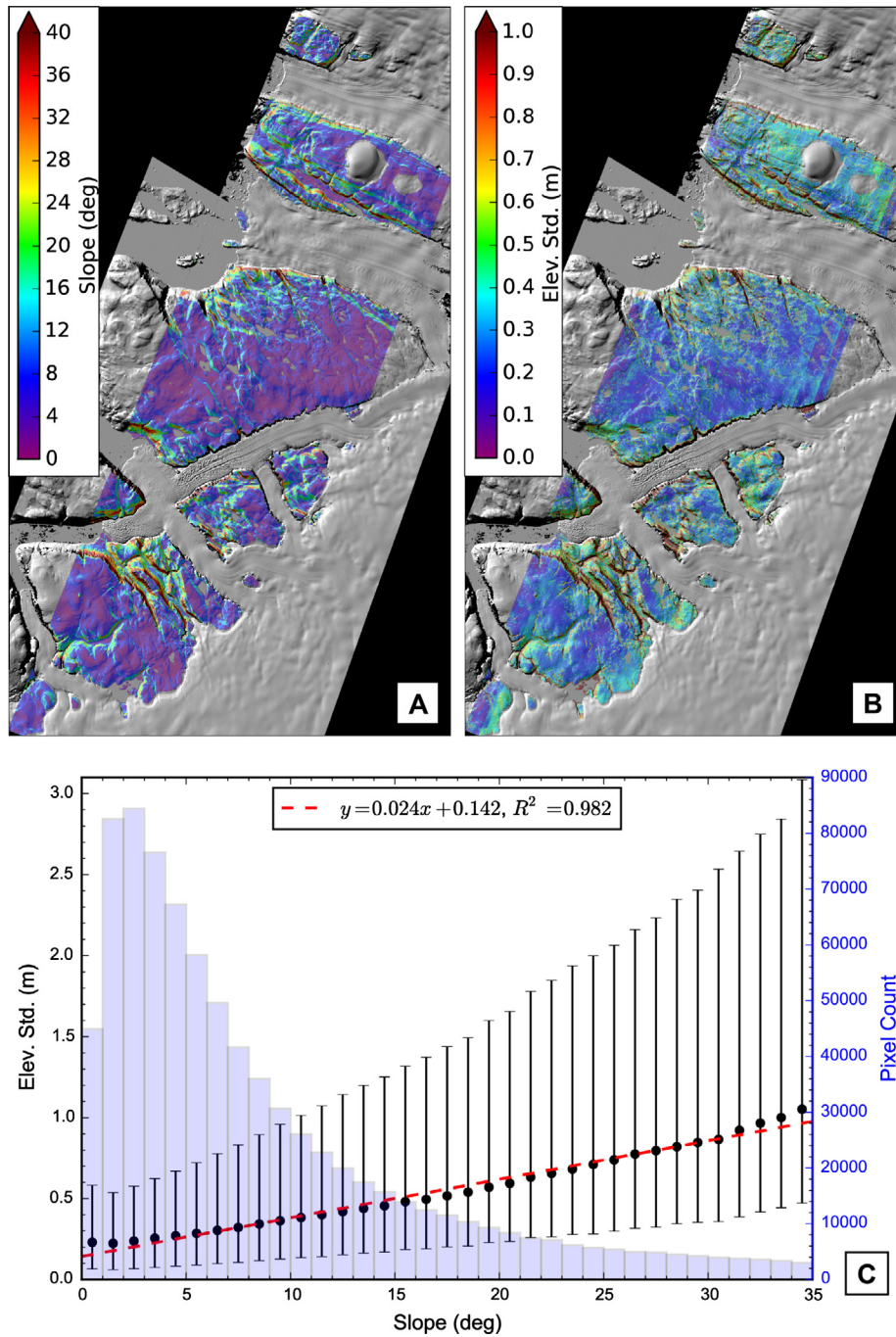
Fig. 14. Analysis of Tracy Glacier DEM stack (n = 17) over bedrock (~827 km²). (A) Surface slope computed from median elevation (Fig. 11B), (B) standard deviation, and (C) standard deviation vs. surface slope. Points represent median (with 16th–84th percentile spread) of all standard deviation values that fall within each 1.0° slope bin. Shaded bars show bin pixel counts (right axis). Dashed line shows linear fit to median values within each bin.

Many early and late season images at high latitudes have very high solar incidence angles (>80° from nadir). If sensor TDI/gain settings are not properly set, partial saturation of extremely bright or dark regions can occur. While this often produces substantial gaps in disparity maps, we have found that successful correlation is still possible for saturated regions when limited surface texture is visible.

7.5. DSA boundary artifacts

As described earlier, the sub-pixel DSA offsets in the L1B mosaics can produce alternating ±0.1–0.5 m vertical errors in output DEMs (Fig. 3). The wv_correct utility mitigates these artifacts, but corrected images will inevitably contain residual linear artifacts, especially near image margins (Fig. 3, 10B). These artifacts are highlighted when differencing WorldView DEMs over surfaces with little or no elevation change. In some cases, an empirical correction for the full scene can be derived from along-track statistics over these surfaces (e.g., Nuth and Kääb, 2011), although this becomes more challenging for scenes with significant relief due to distortion of initially linear artifacts.

7.6. Jitter artifacts

In Section 6.1, we identified along-track elevation errors with magnitude of ~0.1–0.5 m (Fig. 10B) due to "jitter"—slight variations in spacecraft orientation that are not captured by the 50 Hz

attitude tables provided with the L1B rigorous sensor model. Comparable artifacts and corrections (e.g., Mattson et al., 2009) are well documented for the High Resolution Imaging Science Experiment (HiRISE) camera aboard the NASA Mars Reconnaissance Orbiter (McEwen et al., 2007), which shares a similar heritage with WorldView-1/2. In theory, similar corrections could be derived for WorldView-2 using the physically offset panchromatic and multispectral sensors. Unfortunately, unlike HiRISE, the raw L0 image data from individual WorldView-2 DSAs are not available for additional processing. As with the L1B DSA artifacts, these "jitter" artifacts are most apparent in DEM difference products, and custom corrections can potentially be derived for some applications (e.g., Berthier et al., 2007; Nuth and Kääb, 2011).

7.7. Resampled subscene artifacts

Prior to June 2014, federal license restrictions required commercial image vendors to downsample L1B data to $\geq$0.5 m GSD for civilian customers. Due to evolving off-nadir angles for longer image acquisitions, many WorldView-2 products contain some full-resolution subscenes with collected/product GSD of >0.5 m and some subscenes with collected GSD <0.5 m and resampled product GSD of 0.5 m. This selective resampling also affects the rigorous sensor model parameters (e.g., synthetic detector pixel size). The dg_mosaic utility handles these cases by scaling all input subscenes to generate a seamless image with uniform sensor model parameters. We have noticed, however, that some output DEMs derived from these images can display residual artifacts and elevation offsets along subscene boundaries. In some cases, an along-track tilt of ~1 m is observed for each individual subscene, creating a "sawtooth" profile along the full DEM. Despite multiple efforts to develop a dg_mosaic workaround, we have found that it is often simpler to reorder these problematic catalog IDs, or to generate DEMs using the piecemeal processing workflow described in Section 3.1.2.

8. Additional sensors and future work

While this work focuses on WorldView-1 and WorldView-2 data, the DigitalGlobe archive contains VHR stereo pairs acquired by the IKONOS, QuickBird-2, and GeoEye-1 sensors, with data acquisition beginning in ~1999. Although the resolution and quality of ephemeris/attitude information for IKONOS and QuickBird-2 is inferior to WorldView-1/2 and GeoEye-1, the methodology and tools described here can also be used to generate high-quality DEMs from these data. In addition, WorldView-3 now offers improved resolution (~0.31 m nadir GSD) on top of the existing WorldView-2 stereo acquisition capabilities and accuracy specifications.

The updated ASP tools can process commercial VHR image products from other vendors (e.g., Airbus DS, SkyBox) that include standardized RPC sensor models. The ASP source code also includes several sensor templates (e.g., linescan, framing, pinhole) that can be adapted to support additional rigorous sensor models.

The ASP codebase is presently under active development, and will continue to evolve in the coming years, especially as the terrestrial developer/user community continues to grow. Proposed future work includes improved multi-view stereo and Structure from Motion support, integration of more correlation algorithms, and various performance improvements. Finally, our ongoing research efforts continue to leverage ASP and WorldView DEMs to study outlet glaciers in Greenland, ice streams/shelves in Antarctica, and glaciers/snowpack in the Pacific Northwest.

References

Abdalati, W., Zwally, H.J., Bindschadler, R., Csatho, B., Farrell, S.L., Fricker, H.A., Harding, D., Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinhirne, J., Webb, C., 2010. The ICESat-2 laser altimetry mission. Proc. IEEE 98, 735–751. http://dx.doi.org/10.1109/JPROC.2009.2034765.

Aguilar, M.A., del Mar Saldana, M., Aguilar, F.J., 2014. Generation and quality assessment of stereo-extracted DSM from GeoEye-1 and WorldView-2 imagery. IEEE Trans. Geosci. Remote Sens. 52, 1259–1271. http://dx.doi.org/10.1109/TGRS.2013.2249521.

Aguilar, M.A., Saldaña, M.del M., Aguilar, F.J., 2013. Assessing geometric accuracy of the orthorectification process from GeoEye-1 and WorldView-2 panchromatic images. Int. J. Appl. Earth Obs. Geoinf. 21, 427–435. http://dx.doi.org/10.1016/j.jag.2012.06.004.

Anderson, J.A., 2008. ISIS camera model design. In: 39th Lunar and Planetary Science Conference Abstracts, p. 2159.

Becker, K.J., Archinal, B.A., Hare, T.M., Kirk, R.L., Howington-Kraus, E., Robinson, M.S., Rosiek, M.R., 2015. Criteria for automated identification of stereo image Pairs. In: 46th Lunar and Planetary Science Conference Abstracts, p. 2703.

Berthier, E., Arnaud, Y., Kumar, R., Ahmad, S., Wagnon, P., Chevallier, P., 2007. Remote sensing estimates of glacier mass balances in the Himachal Pradesh (Western Himalaya, India). Remote Sens. Environ. 108, 327–338. http://dx.doi.org/10.1016/j.rse.2006.11.017.

Beyer, R.A., Archinal, B., Chen, Y., Edmundson, K., Harbour, D., Howington-Kraus, E., Li, R., McEwen, A., Mattson, S., Moratto, Z., Oberst, J., Rosiek, M., Scholten, F., Tran, T., Robinson, M.LROC Team, 2010. LROC stereo data – results of initial analysis. In: 41st Lunar and Planetary Science Conference Abstracts, p. 2678.

Blair, J.B., Rabine, D.L., Hofton, M.A., 1999. The laser vegetation imaging sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. ISPRS J. Photogramm. Remote Sens. 54, 115–122.

Broxton, M.J., Edwards, L.J., 2008. The Ames stereo pipeline: automated 3D surface reconstruction from orbital imagery. In: 39th Lunar and Planetary Science Conference Abstracts, p. 2419.

Broxton, M.J., Nefian, A.V., Moratto, Z., Kim, T., Lundy, M., Segal, A.V., 2009. 3D lunar terrain reconstruction from apollo images. In: Advances in Visual Computing. Springer, pp. 710–719.

Cheng, P., Chaapel, C., 2008. DigitalGlobe's WorldView-1 satellite: increased image collection opportunities. Geoinformatics, 16–21.

Czimczik, C., 2014. Climate Data Thule Airport; Collaborative Research: Environmental Changes Alter the Carbon Cycle of High Arctic Ecosystems: Shifts in the Ages and Sources of CO_2 and DOC https://www.aoncadis.org/dataset/THU_climate.html. ACADIS Gateway.

Deilami, K., Hashim, M., 2011. Very high resolution optical satellites for DEM generation: a review. Eur. J. Sci. Res. 49, 542–554.

DigitalGlobe, 2014. Geolocation Accuracy of WorldView Products (White Paper).

Dolloff, J., Settergren, R., 2010. An assessment of WorldView-1 positional accuracy based on fifty contiguous stereo pairs of imagery. Photogramm. Eng. Remote Sens. 76, 935–943. http://dx.doi.org/10.14358/PERS.76.8.935.

Dolloff, J.T., Theiss, H.J., 2012. Temporal correlation of metadata errors for commercial satellite images: representation and effects of stereo extraction accuracy. In: ISPRS XXII Congress, pp. 215–223.

Fassett, C.I., 2016. Ames stereo pipeline-derived digital terrain models of mercury from MESSENGER stereo imaging. Planet. Space Sci. (submitted for publication)

Federal Geographic Data Committee, 1998. Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy (No. FGDC-STD-007.3-1998).

Fretwell, P., Pritchard, H.D., Vaughan, D.G., Bamber, J.L., Barrand, N.E., Bell, R., Bianchi, C., Bingham, R.G., Blankenship, D.D., Casassa, G., Catania, G., Callens, D., Conway, H., Cook, A.J., Corr, H.F.J., Damaske, D., Damm, V., Ferraccioli, F., Forsberg, R., Fujita, S., Gim, Y., Gogineni, P., Griggs, J.A., Hindmarsh, R.C.A., Holmlund, P., Holt, J.W., Jacobel, R.W., Jenkins, A., Jokat, W., Jordan, T., King, E.C., Kohler, J., Krabill, W., Riger-Kusk, M., Langley, K.A., Leitchenkov, G., Leuschen, C., Luyendyk, B.P., Matsuoka, K., Mouginot, J., Nitsche, F.O., Nogi, Y., Nost, O.A., Popov, S.V., Rignot, E., Rippin, D.M., Rivera, A., Roberts, J., Ross, N., Siegert, M.J., Smith, A.M., Steinhage, D., Studinger, M., Sun, B., Tinto, B.K., Welch, B.C., Wilson,

D., Young, D.A., Xiangbin, C., Zirizzotti, A., 2002. Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. Cryosphere 7, 375–393. http://dx.doi.org/10.5194/tc-7-375-2013.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., Tyler, D., 2002. The national elevation dataset. Photogramm. Eng. Remote Sens. 68, 5–32.

Grodechi, J., Dial, G., 2003. Block adjustment of high-resolution satellite images described by rational polynomials. Photogramm. Eng. Remote Sens. 69, 59–68.

Heid, T., Kääb, A., 2012. Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery. Remote Sens. Environ. 118, 339–355. http://dx.doi.org/10.1016/j.rse.2011.11.024.

Hofton, M.A., Blair, J.B., Luthcke, S.B., Rabine, D.L., 2008. Assessing the performance of 20–25 m footprint waveform lidar data collected in ICESat data corridors in Greenland. Geophys. Res. Lett. 35. http://dx.doi.org/10.1029/2008GL035774.

Höhle, J., Höhle, M., 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. ISPRS J. Photogramm. Remote Sens. 64, 398–406. http://dx.doi.org/10.1016/j.isprsjprs.2009.02.003.

Howat, I.M., Negrete, A., Smith, B.E., 2014. The Greenland Ice Mapping Project (GIMP) land classification and surface elevation datasets. Cryosphere Discuss. 8, 453–478. http://dx.doi.org/10.5194/tcd-8-453-2014.

Joughin, I., Smith, B.E., Howat, I.M., Scambos, T., Moon, T., 2010. Greenland flow variability from ice-sheet-wide velocity mapping. J. Glaciol. 56, 415–430.

Kääb, A., Leprince, S., 2014. Motion detection using near-simultaneous satellite acquisitions. Remote Sens. Environ. 154, 164–179. http://dx.doi.org/10.1016/j.rse.2014.08.015.

Krabill, W.B., Abdalati, W., Frederick, E.B., Manizade, S.S., Martin, C.F., Sonntag, J.G., Swift, R.N., Thomas, R.H., Yungel, J.G., 2002. Aircraft laser altimetry measurement of elevation changes of the Greenland ice sheet: technique and accuracy assessment. J. Geodyn. 34, 357–376.

Leprince, S., Muse, P., Avouac, J.-P., 2008. In-flight CCD distortion calibration for pushbroom satellites based on subpixel correlation. IEEE Trans. Geosci. Remote Sens. 46, 2675–2683. http://dx.doi.org/10.1109/TGRS.2008.918649.

Martin, C.F., Krabill, W.B., Manizade, S.S., Russell, R.L., Sonntag, J.G., Swift, R.N., Yungel, J.K., 2012. Airborne Topographic Mapper Calibration proCedures and Accuracy Assessment (Technical Memorandum No. NASA/TM–2012-215891). National Aeronautics and Space Administration, Goddard Space Flight Center.

Mattson, S., Boyd, A., Kirk, R.L., Cook, D.A., Howington-Kraus, E., 2009. HiJACK: correcting spacecraft jitter in HiRISE images of Mars. In: European Planetary Science Congress Abstracts, pp. EPSC2009-604-1.

McEwen, A.S., Eliason, E.M., Bergstrom, J.W., Bridges, N.T., Hansen, C.J., Delamere, W. A., Grant, J.A., Gulick, V.C., Herkenhoff, K.E., Keszthelyi, L., Kirk, R.L., Mellon, M.T., Squyres, S.W., Thomas, N., Weitz, C.M., 2007. Mars reconnaissance orbiter's high resolution imaging science experiment (HiRISE). J. Geophys. Res. 112. http://dx.doi.org/10.1029/2005JE002605.

Moratto, Z.M., Broxton, M.J., Beyer, R.A., Lundy, M., Husmann, K., 2010. Ames stereo pipeline. NASA's open source automated stereogrammetry software. In: 41st Lunar and Planetary Science Conference Abstracts, p. 2364.

Müller, J., Gärtner-Roer, I., Thee, P., Ginzler, C., 2014. Accuracy assessment of airborne photogrammetrically derived high-resolution digital elevation models in a high mountain environment. ISPRS J. Photogramm. Remote Sens. 98, 58–69. http://dx.doi.org/10.1016/j.isprsjprs.2014.09.015.

Nefian, A.V., Husmann, K., Broxton, M., To, V., Lundy, M., Hancher, M.D., 2009. A Bayesian formulation for sub-pixel refinement in stereo orbital imagery. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 2361–2364.

Neigh, C.S., Masek, J.G., Nickeson, J.E., 2013. High-resolution satellite data open for government research. Eos, Trans. Am. Geophys. Union 94, 121–123.

Noh, M.-J., Howat, I.M., 2014. Automated coregistration of repeat digital elevation models for surface elevation change measurement using geometric constraints. IEEE Trans. Geosci. Remote Sens. 52, 2247–2260. http://dx.doi.org/10.1109/TGRS.2013.2258928.

Nuth, C., Kääb, A., 2011. Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. Cryosphere 5, 271–290. http://dx.doi.org/10.5194/tc-5-271-2011.

Pomerleau, F., Colas, F., Siegwart, R., Magnenat, S., 2013. Comparing ICP variants on real-world data sets: open-source library and experimental protocol. Auton. Robots 34, 133–148. http://dx.doi.org/10.1007/s10514-013-9327-2.

Pope, A., Scambos, T.A., Moussavi, M., Tedesco, M., Willis, M., Shean, D., Grigsby, S., 2015. Estimating supraglacial lake depth in western Greenland using Landsat 8 and comparison with other multispectral methods. Cryosphere 9, 3257–3292. http://dx.doi.org/10.5194/tcd-9-3257-2015.

Re, C., Cremonese, G., Dall'Asta, E., Forlani, G., Naletto, G., Roncella, R., 2012. Performance evaluation of DTM area-based matching reconstruction of Moon and Mars. In: Bruzzone, L. (Ed.), Image and Signal Processing for Remote Sensing XVIII, Proc. of SPIE, p. 85370V. http://dx.doi.org/10.1117/12.974524.

Rignot, E., Mouginot, J., Scheuchl, B., 2011. Ice flow of the Antarctic ice sheet. Science 333, 1427–1430. http://dx.doi.org/10.1126/science.1208336.

Schutz, B.E., Zwally, H.J., Shuman, C.A., Hancock, D., DiMarzio, J.P., 2005. Overview of the ICESat mission. Geophys. Res. Lett. 32. http://dx.doi.org/10.1029/2005GL024009.

Shean, D.E., Fahle, J., Malin, M.C., Edwards, L.J., Posiolova, L., 2011. MRO CTX stereo image processing and preliminary DEM quality assessment. In: 42nd Lunar and Planetary Science Conference Abstracts, p. 2646.

Shean, D., Joughin, I., Floricioiu, D., Smith, B., Moratto, Z., Alexandrov, O., Morin, P., Porter, C., 2014. A 2008–2013 time series of TanDEM-X & WorldView-1/2 stereo DEMs for Jakobshavn Isbrae, Greenland. In: International Geoscience and Remote Sensing Symposium: Technical Program, p. 3553.

Shimizu, M., Okutomi, M., 2002. Precise subpixel estimation on area-based matching. Syst. Comput. Jpn. 33, 1409–1418.

Siegfried, M.R., Hawley, R.L., Burkhart, J.F., 2011. High-resolution ground-based GPS measurements show Intercampaign bias in ICESat elevation data near summit, Greenland. IEEE Trans. Geosci. Remote Sens. 49, 3393–3400. http://dx.doi.org/10.1109/TGRS.2011.2127483.

Stein, A.N., Huertas, A., Matthies, L., 2006. Attenuating stereo pixel-locking via affine window adaptation. Proceedings IEEE International Conference on Robotics and Automation, 914–921.

Stevens, L.A., Behn, M.D., McGuire, J.J., Das, S.B., Joughin, I., Herring, T., Shean, D.E., King, M.A., 2015. Greenland supraglacial lake drainages triggered by hydrologically induced basal slip. Nature 522, 73–76. http://dx.doi.org/10.1038/nature14480.

Updike, T., Comp, C., 2010. Radiometric Use of WorldView-2 Imagery. DigitalGlobe Technical Note.

Watters, W.A., Geiger, L.M., Fendrock, M., Gibson, R., 2015. Morphometry of small recent impact craters on Mars: size and terrain dependence, short-term modification. J. Geophys. Res.: Planets 120, 226–254. http://dx.doi.org/10.1002/2014JE004630.

Willis, M.J., Herried, B.G., Bevis, M.G., Bell, R.E., 2015. Recharge of a subglacial lake by surface meltwater in northeast Greenland. Nature. http://dx.doi.org/10.1038/nature14116.

Zwally, H.J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., Bufton, J., Dezio, J., Hancock, D., Harding, D., 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. J. Geodyn. 34, 405–445.