Deciphering the Precision of Stereo IKONOS Canopy Height Models for US Forests with G-LiHT Airborne LiDAR

Christopher S. R. Neigh 1,*, Jeffrey G. Masek 1, Paul Bourget 2, Bruce Cook 1, Chengquan Huang 3, Khaldoun Rishmawi 3 and Feng Zhao 3

1 Biospheric Sciences Laboratory, Code 618, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; E-Mails: jeffrey.g.masek@nasa.gov (J.G.M.); bruce.cook@nasa.gov (B.C.)
2 Geography-Anthropology Program, Muskie School of Public Service, University of Southern Maine, P.O. Box 9300, Portland, ME 04104, USA; E-Mail: paulbourget737@gmail.com
3 Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA; E-Mails: cqhuang@umd.edu (C.H.); rishmawi@umd.edu (K.R.); zhao26@umd.edu (F.Z.)

* Author to whom correspondence should be addressed; E-Mail: christopher.s.neigh@nasa.gov; Tel.: +1-301-614-6681; Fax: +1-301-614-6699.

Received: 5 November 2013; in revised form: 7 February 2014 / Accepted: 14 February 2014 / Published: 26 February 2014

Abstract: Few studies have evaluated the precision of IKONOS stereo data for measuring forest canopy height. The high cost of airborne light detection and ranging (LiDAR) data collection for large area studies and the present lack of a spaceborne instrument lead to the need to explore other low cost options. The US Government currently has access to a large archive of commercial high-resolution imagery, which could be quite valuable to forest structure studies. At 1 m resolution, we here compared canopy height models (CHMs) and height data derived from Goddard’s airborne LiDAR Hyper-spectral and Thermal Imager (G-LiHT) with three types of IKONOS stereo derived digital surface models (DSMs) that estimate CHMs by subtracting National Elevation Data (NED) digital terrain models (DTMs). We found the following in three different forested regions of the US after excluding heterogeneous and disturbed forest samples: (1) G-LiHT DTMs were highly correlated with NED DTMs with $R^2 > 0.98$ and root mean square errors (RMSEs) < 2.96 m; (2) when using one visually identifiable ground control point (GCP) from NED, G-LiHT DSMs and IKONOS DSMs had $R^2 > 0.84$ and RMSEs of 2.7 to 4.1 m; and (3) one GCP CHMs for two study sites had $R^2 > 0.7$ and RMSEs of 2.6 to 3 m where data were collected less than four years apart. Our results suggest that IKONOS stereo data are a useful LiDAR alternative where high-quality DTMs are available.
1. Introduction

Forests contain the largest proportion of above ground biomass in terrestrial ecosystems and are subject to natural and human induced disturbances that can reduce their carbon (C) storage [1]. The ongoing rise of anthropogenic C emissions has influenced ecosystem functioning [2,3], and forests should be monitored and evaluated for changes in above ground biomass, canopy cover, and other structural parameters. Forest canopy height is an important structural metric that relates directly to stand age for even aged forests, life cycle, and C sequestration potential when combined with existing allometric relationships [4,5]. Many remote sensing approaches exist to document the structural state of forests. RA dio Detection And Ranging (RADAR), Light Detection And Ranging (LiDAR), and photogrammetric methods using stereo imagery, have all been used to measure forest structure and each approach has varying accuracies and implementation costs [4,6]. We evaluated two existing stereo forest canopy height model (CHM) approaches, one without ground control points (GCPs) and one with GCPs as others have shown substantial improvement in IKONOS mapping accuracy with GCPs [7,8]. We used IKONOS Geo stereo imagery and build upon prior methodologies developed in Quebec [9,10] and compared results against an airborne instrument, Goddard’s LiDAR, Hyperspectral and Thermal Imager (G-LiHT) [11]. We analyzed the precision of IKONOS Geo stereo data with G-LiHT acting as truth, and estimated costs to complete surveys per hectare.

Traditionally, collection of forest management and reporting information in the continental US (CONUS) has been performed with ground plot surveys through the United States Forest Service (USFS) Forest Inventory and Analysis (FIA) program [12]. Field plots are re-measured on a rotating basis every 5-years in the Eastern US (10-years in the Western US), and these data provide a wealth of information including tree height, diameter at breast height, stem volume, species type, etc. This forest plot information can be used for C stock estimates and change analysis [13]. FIA surveys are then used to constrain aerial photography interpretation to produce maps from field samples. Now with open US government access to large volumes of high-resolution satellite imagery (HRSI) [14], forest monitoring studies can use remote sensing data to augment field sample methods to derive wall-to-wall mapping assessments at 1 m resolution [15].

Panchromatic HRSI has been found to be a suitable alternative to airborne imagery in remote locations. Aerial stereo photogrammetry has been successfully used in the past retrieving forest canopy structure digital surface models (DSMs) when combined with accurate digital terrain models (DTMs) [16–21]. Using multi-angle aerial photography, Gong et al. [22] reported overall accuracies of 94% and 90% for tree height and crown radius measurements, respectively. However, only a few have used stereo HRSI to measure forest canopy height [5,9,10,23–25]. Baltsavias et al. [26] has provided limitations to achieving good results through image matching for vegetation including: (1) limited texture; (2) distinct object discontinuities; (3) repetitive objects; (4) occlusions; and (5) multilayered objects. Accurate co-registration of imagery, high resolution DTM s, and airborne field measurements have fostered improvements for forest applications. Maps of forest canopy height can be generated by
combining the HRSI derived DSM with a co-registered DTM, such as those provided by the United States Geological Survey (USGS) National Elevation Data (NED). NED DTMs based on airborne LiDAR have rapidly expanded in coverage over recent years enabling more forest canopy heights to be mapped.

All DigitalGlobe satellites can collect within track stereo imagery and the volume of these data has grown exponentially. Now over a decade of stereo measurements from multiple HRSI sensors exist. WorldView-1 currently has the largest stereo archive of sub-meter panchromatic data beginning in 2008, with IKONOS having the second largest archive beginning in 2000. These data could provide beneficial information on the state of forests.

Most studies using stereo HRSI have focused on urban feature extraction with successful results [27–30]. Geo stereo products which comprise most of the stereo IKONOS archive have a reported horizontal accuracy of 25 m CE90 and vertical accuracy of 22 m LE90 by DigitalGlobe. This reported error would limit these products to DTM generation, however with high precision GCPs Wang et al. [8] demonstrated that Geo stereo product accuracies can be enhanced from ~5 to 1.5 m in horizontal and from 7 to 2 m in vertical directions. Poon et al. [23] found Geo stereo accuracy to be within 3 pixels (1 m panchromatic resolution), or 5.3 m RMSE for forest canopies. Fraser et al. [31] found that with a few high quality GCPs, IKONOS Geo stereo data can achieve a vertical accuracy of 0.7 m. These studies imply that IKONOS stereo products would be useful for measuring vegetation canopy heights if GCPs are used to refine rational polynomial coefficients (RPCs) geolocation.

Due to the amorphous shape of forest canopies, their accuracy is more difficult to quantify. Often best results have been found relating field measurements to small footprint LiDAR, and HRSI has been found to be the second best option [6]. From prior successful outputs of these studies in urban environments and a handful of forest canopy studies [5,10,32], we performed an assessment of the capabilities of existing remote sensing data to map canopy heights in three different biogeographic forested regions of the CONUS including Appalachian, Atlantic Coastal Plain, and North American Pacific Maritime. From prior studies [8,23] we believe comparing DSMs derived from IKONOS will enable the assessment of the best approach to extract CHMs.

Our objective was to determine the accuracy of stereo IKONOS for mapping tree height in three different biogeographic regions of the US. We evaluated whether forest height estimates could be improved with GCPs derived from NED DTM data in bare earth locations and compared results to airborne LiDAR. We produced IKONOS DSMs with no-GCPs, one GCP and 16 GCPs to determine the accuracy of IKONOS DSMs with different processing methods and compared results to G-LiHT DSMs. We also compared DTMs and CHMs from NED, IKONOS and G-LiHT to understand were the error is introduced and how it could be reduced.

2. Study Areas

We selected three forested regions of study where G-LiHT and archived IKONOS Geo stereo data overlapped. These sites included from east to west: (1) Harvard Forest in central Massachusetts; (2) Jamison in central South Carolina; and (3) Hoquiam on the central west coast of Washington state (Figure 1). Generally, IKONOS stereo images encompass a total area of ~100 km² and we subset IKONOS outputs to the G-LiHT footprint for comparison. Harvard Forest has low terrain variability
with heights varying from 280 to 360 m (earth gravitational model 1996 geoid). Few forest disturbances have occurred in this area over the past 30 years. Jamison is generally flat as well, with elevation ranging from 55 to 85 m. It has agro-forestry and many homogenous patches of even age stands that have been harvested over the past 30 years. The Hoquiam study region extends from sea level to rolling hills at 120 m elevation with some slopes greater than 20°. Its forest was harvested in large patches throughout the image extent. Each study region represented different forest cover types with different disturbance patterns and management practices. Conifers dominated the Washington and South Carolina study regions, and mixed hardwoods are the dominant species in Massachusetts.

**Figure 1.** Upper left shows the NASA’s Blue Marble 500 m imagery of the A–C study area locations. Panels (A) Harvard Forest, Massachusetts; (B) Hoquiam, Washington; and (C) Jamison, South Carolina show I-cubed 1 m seamless color mosaic of commercial and government imagery with tiled G-LiHT products indicated with white quadrangles and IKONOS stereo footprints indicated with red quadrangles.
3. Data and Methods

To generate CHM estimates and decipherer error contributing components, we compiled data from three sources to test in this study, including: (1) NASA G-LiHT airborne instrument; (2) IKONOS Geo stereo products from GeoEye (now DigitalGlobe); and (3) USGS NED DTMs 3 m bare earth LiDAR where available, and 10 m elsewhere.

3.1. G-LiHT Airborne Laser Scanner

G-LiHT is a compilation of available off the shelf instrument components of LiDAR, imaging spectroscopy, and thermal instruments that when combined enables data fusion [11]. It provides highly accurate and discrete foliage and canopy measurements. G-LiHT is a compact system that can be placed on most small aircraft, and has collected data over a diverse range of cover types and has undergone calibration/evaluation studies [11]. We used LiDAR data from the VQ-480 onboard G-LiHT. The VQ-480 uses a high-performance laser rangefinder with a rotating polygon mirror with three facets that deflect a 1550 nm Class 1 laser beam toward the ground. The pulse repetition rate is user selectable up to 300 kHz and provides a measurement rate up to 150 kHz along a 60° swath perpendicular to the flight direction. A laser beam divergence of 0.3 mrad produces a 10 cm diameter footprint at the nominal operating altitude of 335 m. Data from our three study regions was collected leaf-on in the summers of 2011 and 2012 with identical flight parameters. Processed LiDAR data are available online [33], including classified returns and heights, and 1 m CHMs and DTMs. LiDAR returns were resampled and processed as ~7 km transect segments for efficient processing. Classification of ground returns was performed with a progressive morphological filter [34]. Delaunay triangulation is used to create a Triangulated Irregular Network (TIN) of ground hits, and the TIN is used to linearly interpolate DTM elevations on a 1 m raster grid. Examples of products and details about how these data have been processed are available in Cook et al. [13]. G-LiHT data on average have more than six shots per m², and pixels with fewer than this were excluded from our analyses. CHMs were derived from differencing rasterized first returns from ground classified last returns. Our study focused on using LiDAR data acquired from G-LiHT as the validation dataset. We used G-LiHT DSMs, CHMs and DTMs to validate IKONOS DSMs, CHMs and NED DTMs, respectively.

3.2. IKONOS Stereo Imagery

IKONOS Geo stereo data were collected from the archives of GeoEye under the NextView licensing agreement at no direct cost to the investigation [14]. NextView is a contract between the US government and US commercial vendor GeoEye (now merged with DigitalGlobe) to provide HRSI to Federal, State, and local government agencies and organizations that support US government interests. Those who receive research funding from US government organizations can acquire these data at no direct cost. IKONOS Geo stereo reference products were acquired from the USGS Commercial Remote Sensing Space Policy (CRSSP) Imagery-Derived Requirements (CIDR) tool [35]. These data products were collected from the vendor at the reference accuracy level with standard geometric processing. IKONOS is a pushbroom imager, and each pixel has its own time dependent attitude angles and perspective center position. GeoEye does not provide per-pixel look angles, instead a
sensor orientation model, or rational function model (RFM) is used to describe the object-to-image space transformation mathematically. RFM parameters derived from a rigorous sensor model are supplied with the imagery and are termed RPCs [36,37].

Launched in 1999, IKONOS is the first high-resolution commercial satellite instrument with a 1 m panchromatic band, and four multispectral 4 m bands. Within track stereo data are collected less than two minutes apart. Typically an image is collected at an elevation angle (greater than 60° from the Earth’s horizon) and a second image is collected at a higher elevation angle (above 72°) with 30 to 45° convergence resulting in base-to-height ratios of 0.54 to 0.83. Typically ratios from 0.5 to 0.9 yield high height precision [8]. Within track stereo has an advantage over cross track stereo (multiple passes) because of similar atmospheric and surface conditions, so these data can have a systematic correction between image pairs. We did not apply atmospheric correction or stretch the data for image matching enhancement. All imagery was cloud free. Specific data characteristics are provided in Table 1.

Table 1. Goddard’s airborne LiDAR Hyper-spectral and Thermal Imager (G-LiHT) and IKONOS dates of acquisition with IKONOS Geo stereo viewing geometry, base-to-height ratio and National Elevation Data (NED) digital terrain model (DTM) resolution by study region.

<table>
<thead>
<tr>
<th>Study Region</th>
<th>G-LiHT Acquisition Date mm/yr</th>
<th>G-LiHT Data ID</th>
<th>IKONOS Acquisition Date mm/dd/yr</th>
<th>IKONOS Nominal Collection Elevation</th>
<th>B/H Ratio</th>
<th>NED DTM Resolution (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harvard Forest MA</td>
<td>08/2011</td>
<td>AMIGACarb_5S_Aug2011</td>
<td>10/12/2000</td>
<td>64.40°–72.87°</td>
<td>0.64</td>
<td>10</td>
</tr>
<tr>
<td>Jamison SC</td>
<td>09/2011</td>
<td>AMIGACarb_13s_Sep2011</td>
<td>10/31/2010</td>
<td>60.75°–87.04°</td>
<td>0.55</td>
<td>3</td>
</tr>
<tr>
<td>Hoquiam WA</td>
<td>08/2012</td>
<td>AMIGACarb_G03_Aug2012</td>
<td>09/14/2008</td>
<td>69.03°–77.94°</td>
<td>0.58</td>
<td>10</td>
</tr>
</tbody>
</table>

3.3. NED DTMs

To estimate the height of forest canopies, DSMs must have DTMs subtracted from them to produce CHMs. NED DTMs from the USGS [38] have a national resolution of 10 m for the CONUS and they are comprised of data from multiple sources including cartographic, photogrammetric, auto-correlated maps, imagery, and airborne LiDAR [39]. NED is also produced at 3 m resolution with airborne LiDAR. NED is primarily based upon digital elevation model (DEM) ortho quadrangles although the LiDAR derived 3 m DTMs are rapidly being acquired over large areas of the CONUS. NED LiDAR data are primarily collected leaf-off (winter) with shot densities that are not optimized for measuring forest structure (~1 shot per m²). These data measure bare earth topography and are optimized to do so as they are the best available DTM source data to use with HRSI DSMs to estimate CHMs in the CONUS. NED DTMs were resampled with cubic convolution to match the 1 m resolution of the IKONOS DSMs. Horizontal and vertical error of these data have been estimated by the National Geodetic Survey (NGS) with 9187 reference control points throughout the CONUS to be 1.67 m vertical RMSE for the entire 1 arc second (30 m) data [40]. The 3 m LiDAR ground classified
last-return derived data are expected to have improved accuracy although to our knowledge currently no summary statistics exist.

3.4. Processing Stereo HRSI

Remote sensing data have been processed and subset to the domain of overlap for comparison. We describe how we process the data herein and focus on the development of the IKONOS stereo DSMs. No co-registration procedures were applied, thus we relied on each datasets geolocation in this comparison. Techniques in photogrammetry, geodesy, and cartography have been developed to account for the complete viewing geometry. Models can be either simple (rational functions, polynomial, or thin plate spline) that build correlation between the ground and pixels from ground control points (GCP), or rigorous (Toutin satellite model) transformations of satellite orbital viewing geometry [36]. Satellite orbital math models rely on co-linearity conditions transforming image space into ground space. These models have evolved to compensate for the swath width of high-resolution systems, which have inherent errors that need correction from platform position, velocity, orientation, sensor orientation, integration time, instantaneous field of view, and earth surface characteristics, i.e., terrain relief. Some difficulties remain, and accuracies vary based upon data available to generate epipolar pairs-image pairs that have excluded projection distortion from various math models. Estimates of height from these model results vary by location depending upon available GCPs and physical topography.

We provide our schema in Figure 2, and data processing details include:

(1) Stereo products were processed by using IDL/ENVI Version 5.0 DEM extraction module software. These data contain RPC files with geolocation information. DSMs were produced with RPCs alone and with GCPs to refine RPCs;

(2) Identical features in each image pair were manually collected, generating fifty or more stereoscopic parallax tie points. Points were collected with an even distribution throughout the image overlap and they had a maximum y-parallax error of 0.75 m. Typically, man-made features (buildings, road intersections, parking lots, etc.) with easily identifiable corners or intersections were selected between images. Tie points were used to generate left and right epipolar images (stereo images that overlap, and when displayed together produce an anaglyph, or 3-D viewable image) that are useful for solving the external image orientation. By overlaying epipolar images, one can produce an anaglyph or 3-D image (with 3-D red/blue glasses) so that height through parallax can be extracted. Stereo tie points then guide a moving window to develop correlated tie points between epipolar images to produce posts of height or grid points in a TIN from which stereo DSMs are rasterized from. Using epipolar images reduces one dimension of variability and increases the processing speed of image matching. Images are matched through successive iterations starting at coarse pyramid levels that are predefined (starting at 264), moving downward with each successive TIN toward full resolution (128, 64, 32, 8, 4, 2, 1). User input cannot be provided with ENVI’s DEM extraction module through the matching procedure and no preprocessing filters were applied to optimize images for feature extraction;

(3) GCPs were identified from bare areas in IKONOS and NED DTM data. GCPs were placed where identical image features existed between stereo pairs, and height was estimated from the same approximate location in the DTM. For each GCP the NED DTM height was recorded for IKONOS
DSM processing. Locations were selected based on low topographic relief to reduce vertical error in the horizontal plane. We acknowledge that this process was straightforward and easily accomplished with LiDAR derived NED DTMs as similar features in IKONOS and NED are pronounced when comparing 3 m and 1 m data. This task was more difficult with NED 10 m DTM data as it has less pronounced features to select visually identifiable co-located points. Previous studies have reported large increases in accuracy in both vertical and horizontal error when using GCPs. A notable improvement of this approach is high-resolution DTMs (3-m) that enable the identification of similar features in stereo HRSI. This could result in a large cost savings and improve accuracy of CHMs compared to collecting \textit{in-situ} GCPs;

(4) GCPs were then used to improve pixel to ground relationships in building epipolar images for image correlation analysis in development of DSMs. We produced three types of DSMs from the IKONOS Geo stereo data to understand and quantify errors in stereo CHM development. These included: (a) DSMs based on RPCs alone, requiring a geoid calculation; (b) DSMs derived with one GCP from NED DTM in a bare earth location; and (c) DSMs derived with 16 evenly distributed GCPs in bare earth locations throughout the image.

Figure 2. Schema of how data were compared to decipherer components that impact IKONOS canopy height model (CHM) error. The left portion describes data used in analysis, center describes derived data products, and the right portion describes the comparison of products.

To generate DSMs without GCPs the geoid height was added to the DSM and then the DSM was subtracted from the DTM to calculate the CHM. The CHM without GCPs was calculated with the following equation per pixel:

\[
CHM = (DSM_i + G) - \text{NED}
\]

where \text{NED} is either the 3 m where available, or 10 m bare earth DTM, \(G\) is the geoid height estimated at image center coordinate and derived from the National Geodetic Survey earth gravitational model.
1996 (EGM96) geoid calculator, and $DSM_i$ is the IKONOS digital surface model without GCPs. The geoid is used to convert between the ellipsoid (IKONOS) to mean sea level elevations (NED).

The CHM with GCPs was calculated with the following equation per pixel:

$$CHM = DSM_{igcp} - NED$$

where $DSM_{igcp}$ is the IKONOS digital surface model with one or 16 GCPs.

The parallax between stereo pairs allows estimates of height. The ENVI DEM extraction software uses cross correlation between the grayscale of targets within a moving window to estimate the y-parallax or 3-D stereoscopic difference between identical points within 1 m panchromatic imagery. We used a 5 × 5 moving correlation window, with a correlation minimum threshold of 0.7, and the highest possible terrain processing level available to produce DSMs so that topographic features would not be smoothed.

3.5. Valid Data Ranges

We developed a data acquisition strategy to obtain and compare our estimates of DTMs, DSMs, and CHMs for three study regions through automated quality-screening thresholds applied to G-LiHT, IKONOS and Landsat disturbance history (using criteria provided in Table 2). We developed automated approaches to minimize quality problems while limiting the degree of manual intervention in processing. Image analysts did not edit IKONOS DSMs, and height anomalies exist in both CHMs from IKONOS and G-LiHT primarily over water because specular reflection can overwhelm LiDAR sensors and create ranging errors. Also, IKONOS parallax errors are high over water due inaccuracies of correlated points for TIN posts used for DSM development. To exclude water and urban features from sampling we used the normalized difference vegetation index (NDVI) derived from Gram-Schmidt pan-sharpened multi-spectral bands from IKONOS with the following equation:

$$NDVI = (Band 4 - Band 3)/(Band 4 + Band 3)$$

where NDVI of water, urban features (roads, buildings, concrete, etc.), and clouds typically have negative values and vegetation has positive values due to the high reflectance of near-infrared wavelengths interacting with the internal mesophyll structure of healthy leaves.

To compare forested area for top of canopy heights (IKONOS DSMs and G-LiHT first returns) and non-forested areas (NED DTMs and G-LiHT ground classified last returns) we used G-LiHT range corrected and instrument calibrated apparent reflectance at 1550 nm. Typically, bare ground has high reflectance values and vegetation has low reflectance values at this wavelength. We used this LiDAR attribute to screen data for the appropriate comparison of height measurements. DTM estimates were compared between G-LiHT ground classified last returns and NED DTMs. We sampled 100,000 pixels for DSMs and CHMs, and 25,000 pixels for DTMs from a standard normal distribution of vectorized data to insure samples were consistent. To remove forest edges and ensure samples were acquired for homogenous areas we selected two pixels on either side vertically from the randomly selected pixel to produce a five-pixel sample. We then calculated the median and standard deviation for each 5-pixel sample for DTMs, DSMs, and CHMs. If the standard deviation of the five height samples was greater than two for any dataset compared, we excluded that sample from analysis. We then used the median height value of the cluster of pixels for data comparison. This approach removed a large portion of
pixels from our sample within steep slopes and also removed vegetation with complex vegetation height structures. We believe this approach reduces resolution differences between data and potentially reduces co-registration error that would enhance canopy height variability in heterogeneous forests.

Landsat time series stacks are being processed for the entire CONUS to understand annual forest disturbance using the vegetation change tracker (VCT) algorithm [41–43]. These 30 m resolution records of forest disturbance date back to 1984, have been composited to remove cloud cover, document persisting undisturbed forest, and are a useful tool to understand differences in canopy height estimates from persisting undisturbed and disturbed forest. More information about the algorithm is provided in Huang et al. [41]. We used VCT data to exclude forest disturbances that occurred after the IKONOS acquisition, since disturbances can significantly alter forest structure height. This would create incomparable measurements between IKONOS and G-LiHT CHMs. A bias will result from acquisition differences, and this bias will be enhanced in areas with intensive agroforestry, since young forests grow at faster rate than old forests [44]. This difference could result in false comparison of values due to changes in canopy height and gap formation from harvest between dates of data collection. We also acknowledge that forest degradation such as thinning or partial harvest not captured by VCT could have an impact to forest canopy height, hence rendering this metric a less effective tool for screening.

### Table 2. G-LiHT, IKONOS, and Landsat valid data ranges for DTM, digital surface model (DSM), and CHM comparisons.

<table>
<thead>
<tr>
<th></th>
<th>Valid Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHM height</td>
<td>&gt;0.1 m</td>
</tr>
<tr>
<td>DSM height</td>
<td>0 &gt; 500 m</td>
</tr>
<tr>
<td>DTM height *</td>
<td>0 &gt; 500 m</td>
</tr>
<tr>
<td>IKONOS NDVI</td>
<td>&gt;0.3</td>
</tr>
<tr>
<td>G-LiHT instrument calibrated apparent reflectance first returns</td>
<td>&lt;0.15</td>
</tr>
<tr>
<td>G-LiHT instrument calibrated apparent reflectance last returns *</td>
<td>&gt;0.35</td>
</tr>
<tr>
<td>G-LiHT DTM Slope *</td>
<td>&lt;10°</td>
</tr>
<tr>
<td>G-LiHT and IKONOS DSM 5 pixel sample</td>
<td>&lt;2 standard deviations</td>
</tr>
<tr>
<td>Landsat VCT</td>
<td>All data pre and non disturbed forest post IKONOS acquisition.</td>
</tr>
</tbody>
</table>

* = values used to exclude poor quality data for DTM comparisons.

### 3.6. Data Comparison

Each data source was sampled using a standard normal distribution of pixels extracting height from DTMs, DSMs, and CHMs and compared to G-LiHT reporting correlation ($R^2$) and root mean square errors (RMSEs). We believe this provided an unbiased and quantitative way to identify errors that may propagate into IKONOS CHMs. The precision of LiDAR measurements ranging is submeter but due to limitations in geolocating points and sampling density G-LiHT CHM accuracy is degraded to ~1 m [6]. We believe comparing DTM and DSM stereo results to G-LiHT is best approach to evaluate the accuracy of these products.
4. Results

To decipher error in IKONOS CHMs we compared three different approaches to generate IKONOS DSMs in three regions of the CONUS. Comparing methods to develop IKONOS DSMs and CHMs to G-LiHT enabled understanding of error propagation in the creation of IKONOS CHMs.

Figure 3. Scatter plots of three regions using a standard normal distribution of 25,000 randomly sampled points after filtering from Table 2 comparing NED DTM vs. G-LiHT DTM. The black line indicates a one-to-one relationship and the thin gray line indicates a least squares linear fit. The total number of samples is displayed in the plot title. Black points indicate no filtering, red points indicate data filtered with high LiDAR reflectance and blue points indicate data filtered with high LiDAR reflectance and slopes less than 10°.

4.1. NED DTMs vs. G-LiHT DTMs

We found G-LiHT DTMs to be highly correlated to NED DTMs. Through successive filtering of DTM data, $R^2$ improved and RMSEs declined. We provide filtering results in Figure 3, showing the impact of no filtering, reflectance filtering, and reflectance filtering with G-LiHT DTM slope filtering. These results decompose the error between DTM products to help clarify the cause of errors. We found the largest RSME of 6.42 m to exist in Hoquiam Washington where variance in terrain slope of the 1 m resolution G-LiHT DTM is the largest of all three-study regions. The NED DTM samples from Hoquiam had the largest amount of error in the NED 10 m DTM and existed in locations with steep terrain. This implies NED 10 m DTM error will propagate into CHM estimates for Hoquiam and
Harvard Forest. When areas of slope greater than 10° were excluded, RMSE reduced from 4.13 m to 2.96 m in Hoquiam. This study site had the largest absolute reduction in error of all three investigated from 6.42 to 2.96 m. Jamison also had a reduction in error with an RMSE declining from 2 to 0.2 m. Flight parameters were slightly different in Jamison where the aircraft was allowed to briefly fly higher than protocol. The LiDAR has limited ranging capability and the aircraft was flying at an altitude near the edge of that range. Off-nadir shots could not reach the ground but could reach the top of trees. Without any ground returns at the swath edge, the G-LiHT algorithm used to create the DTM wrongly assumed the tree tops (the only returns sensed) were ground. Without screening for high altitude edge anomalies, errors can be introduced to the G-LiHT DTM. We found reflectance filtering could also remove these anomalies.

4.2. IKONOS DSMs vs. G-LiHT DSMs

Three types of DSMs for all three study sites were compared to G-LiHT DSMs displayed in Figure 4 with criteria from Table 2. Table 3 provides GCP quality information for each region. GCP errors are based on the difference between model predicted GCPs based on RPCs and analyst selected GCPs in meters. The largest errors occurred in Jamison were more non-model predicted GCPs were used as compared to the other two study regions. Overall the median $X$ and $Y$ GCP error was less than 3 m and all sites had $R^2$ greater than 0.79 with RMSEs varying from 2.7 m in Jamison to 4.3 m in Harvard Forest. Filters applied from Table 2 reduced samples sizes by 77 to 91%, and remaining values from the sample of 100,000 are provided in figure titles.

All sites showed strong agreement between IKONOS DSMs and G-LiHT DSMs, but we found a vegetation median height offset from G-LiHT in the one GCP Hoquiam DSM sample that required more investigation. We provide Figure 5 to describe the median height offset of the standard normal distribution sample of pixels between IKONOS DSMs and G-LiHT first returns. We found that Harvard Forest and Jamison one GCP IKONOS DSMs (red line) had a good overall fit to G-LiHT DSMs, with median offset height values of 0.34 m and 0.48 m, respectfully. Note the spread of the sample is distributed more in Harvard Forest as compared to Jamison, and we believe this is due to the difference in data acquisition periods. From these results we find optimal results are achieved with one GCP to improve the absolute height value in DSMs in Jamison and Harvard Forest. In Hoquiam (with high relief) the one GCP IKONOS DSM had the largest median offset of 6 m compared to the 16 GCP IKONOS DSM which had a median height offset from G-LiHT of 0.02 m. In this case 16 GCPs provided optimal results. We believe the 6.42 m RMSE from the 10 m NED DTM found and displayed in Figure 3, is compounded here when only one GCP is selected for IKONOS DSM development. From these results, we believe adding multiple GCPs in a region with high terrain variability could potentially minimize this error. Furthermore, the outlier samples in Hoquiam where the NED DTM values are higher than G-LiHT DTM are related to slopes greater than 10°. Additionally when 3 m LiDAR derived NED DTM data become available, we believe it could substantially reduce one GCP error because the vertical error from the 10 m DTM is propagated to the bundle adjustment error. As indicated in Toutin [7], when GCPs have an accuracy poorer than 3 m, 20 GCPs spread over the entire image are a compromise to obtain a 3 to 4 m accuracy in the bundle adjustment; when GCP accuracy is better than 1 m, ten GCPs are enough to decrease the bundle adjustment error of either
panchromatic or multiband images to 2 to 3 m. This implies 3 m NED DTMs with sub-meter RMSEs would improve the bundle adjustment.

Figure 4. Scatter plots of three study regions using a standard normal distribution of pixels randomly sampled with totals indicated in plot title. Three types of DSMs from IKONOS were compared to G-LiHT LiDAR first returns. Values not meeting criteria in Table 2 were excluded from analysis. The colors from left to right indicate no ground control points (GCPs) in black, one GCP in red, and 16 GCPs in blue. The black line indicates a one-to-one relationship and the gray line indicates a least squares linear fit.

Table 3. Errors for sixteen GCPs used for each study region based on the difference between actual and model predicted GCP locations in meters.

<table>
<thead>
<tr>
<th>Harvard Forest—Massachusetts</th>
<th>Elev</th>
<th>Error X</th>
<th>Error Y</th>
<th>Error Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>320.0</td>
<td>0.00003</td>
<td>0.00011</td>
<td>18.26</td>
</tr>
<tr>
<td>Minimum</td>
<td>231.9</td>
<td>-0.00006</td>
<td>-0.00015</td>
<td>12.47</td>
</tr>
<tr>
<td>Maximum</td>
<td>353.8</td>
<td>0.00032</td>
<td>0.00021</td>
<td>22.84</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>34.4</td>
<td>0.00008</td>
<td>0.00010</td>
<td>2.39</td>
</tr>
</tbody>
</table>
Table 3. Cont.

<table>
<thead>
<tr>
<th></th>
<th>Elev</th>
<th>Error X</th>
<th>Error Y</th>
<th>Error Z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Jamison</strong>—South Carolina</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>84.2</td>
<td>0.25675</td>
<td>2.83890</td>
<td>33.00</td>
</tr>
<tr>
<td>Minimum</td>
<td>55.0</td>
<td>−1.62250</td>
<td>−3.95360</td>
<td>18.81</td>
</tr>
<tr>
<td>Maximum</td>
<td>103.5</td>
<td>2.50230</td>
<td>37.50100</td>
<td>40.16</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>11.6</td>
<td>0.98625</td>
<td>9.20157</td>
<td>4.85</td>
</tr>
<tr>
<td><strong>Hoquiam</strong>—Washington</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>40.9</td>
<td>−0.00001</td>
<td>0.00001</td>
<td>25.30</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.8</td>
<td>−0.00048</td>
<td>−0.00015</td>
<td>22.81</td>
</tr>
<tr>
<td>Maximum</td>
<td>99.4</td>
<td>0.00004</td>
<td>0.00010</td>
<td>28.52</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>30.9</td>
<td>0.00012</td>
<td>0.00008</td>
<td>1.65</td>
</tr>
</tbody>
</table>

**Figure 5.** Height difference of LiDAR DSMs minus IKONOS DSMs. Data are presented as a histogram count percent of total from the standard normal distribution of randomly sampled pixels. The colors indicate no GCP DSMs with dashed black lines, one GCP DSMs with solid red lines, and 16 GCP DSMs are shown with dashed blue lines. Values not meeting criteria in Table 2 were excluded from analysis.

4.3. **IKONOS CHMs vs. G-LiHT CHMs**

G-LiHT and IKONOS CHMs derived from one GCP IKONOS DSMs had $R^2$ 0.71 for Jamison, $R^2$ 0.70 for Hoquiam, and a $R^2$ 0.24 for Harvard Forest, as shown in Figure 6. The latest result was anticipated due the eleven-year difference in data acquisition for the Harvard Forest study site and the potential inefficiency of Landsat VCT to detect non-stand clearing events and/or small forest canopy gap changes. Filter thresholds of slope and DSM standard deviation from Table 2 had the largest impact on CHM correlation. Relaxing these values any further to produce a larger sample size would reduce correlation. Jamison had the highest $X, Y$ registration error as indicated in Table 3, and we believe this could have propagated to CHM comparison results. The median height offset for the Hoquiam CHM was due to the one GCP DSM offset displayed in Figure 5 and the relative high RMSE error (6.42 m for unfiltered and 2.96 m for filtered data) for the DTM displayed in Figure 3 compared to the other study sites. The quality of the one GCP used contributed to this error. Subtracting the estimated 6 m median height offset from the one GCP DSM or using 16 or more GCPs would reduce the Hoquiam CHM height offset and place it into closer agreement with the G-LiHT CHM.
Figure 6. CHM scatter plots of three study regions with total number of random samples indicated in plot title. Values not meeting valid criteria listed in Table 2 were excluded from analysis. The solid black line indicates a one-to-one relationship and the solid gray line indicates a least squares linear fit. Colors from left to right indicate no GCPs in black, one GCP in red, and 16 GCPs in blue.

We used Landsat forest disturbance data to remove samples that would bias a comparison of airborne LiDAR to IKONOS due to temporal differences in data acquisition. These temporal data are useful for more than screening disturbed sites in our analysis. They can be used for cross validation of disturbed patches of forest. In Figure 7 we provide a graphic that displays how multiple datasets with good co-registration can be used for cross comparison of vegetation height. The resolution difference between CHMs is evident where linear road features are easily identifiable in the IKONOS multi-spectral image (MSI) and the G-LiHT CHM with values of zero. Although, many of the roads are lost in the IKONOS CHM. Future studies that acknowledge the limitations of each dataset could be fused to perform a space-for-time analysis to understand rates of forest regrowth for C-cycle science or be used to optimize large area canopy height mapping where wall-to-wall stereo data exist.
5. Discussion

Landsat VCT has the greatest accuracy in detecting stand clearing events and has limited abilities in capturing storm disturbance, understory thinning, insect outbreak, etc. We used VCT to exclude major differences in forest structure that would be observable in both G-LiHT and IKONOS stereo data that could bias our results. We used VCT to exclude disturbed stands because forest growth is logarithmic with young stands growing at much faster rates (~10 m over 10 years) than old growth stands (~1.5 to 2.5 m over 10 years) in CONUS forests [45]. Currently Landsat VCT products end in 2010, and do not include tropical storm Irene that occurred in 2011 in Harvard Forest. G-LiHT data were acquired before tropical storm Irene struck the area with wind gusts of 16.4 m/s (37 mph). No maps of wind damage are available for the 2000 to 2011 period [46] and we did not consider wind damage when filtering data for any of our study regions.

Currently HRSI stereo data are available to US federal agencies and those who perform research that is funded by federal agencies (Universities, non-governmental organizations, etc.) at “no direct cost” through the National Geospatial-Intelligence Agency (NGA). Our results suggest NED DTMs and IKONOS DSMs have accuracies acceptable to produce CHMs with RMSEs less than 3.9 m for temporally disparate data and RMSEs less than 2.6 m for near co-incident data. We anticipated the smallest and largest errors in Jamison and Harvard Forest, respectively, because of IKONOS acquisition differences to G-LiHT. We assumed that a non-disturbed vegetation growth signature would appear in this comparison. Correlation and RMSE in Jamison and Hoquiam implies IKONOS stereo data are a sufficient alternative to LiDAR in locations with even aged forest stands and low topographic relief for biomass mapping. Our study compared closed canopy forests and does not
evaluate the ability of stereo HRSI to map tree height of sparse canopy forests. Future studies should evaluate forest canopy cover fraction in relationship to stereo HRSI and LiDAR measurements to understand when stereo HRSI may underestimate tree height. Additionally, the USDA forest service resource photography has collected airborne 16-bit digital stereo imagery at high spatial resolution (25 to 50 cm) in four bands over the western U.S. These data may serve as an additional resource when conducting forest height studies in the US The large archive of stereo HRSI could provide valuable forest height information to large area biomass mapping studies that use stratify and multiply sampling approaches [47–49] and/or direct remote sensing modeling approaches [50–52].

Costs to conduct commercial LiDAR surveys vary in price depending upon the extent of the study area. Typically, firms do not wish to perform projects less than 25 to 50 km$^{-2}$ with costs starting at $4 per hectare. For costs to be significantly reduced the study area must be substantially larger, over 1000 km$^{-2}$, or pooled with other firms and projects to avoid having to meet a $25,000 minimum project size [53]. This profitability analysis by firms excludes small regions from LiDAR mapping because of high data acquisition costs. When compared to the cost of software to process HRSI stereo imagery (~$5000 to greater than $25,000) and open US Government access to HRSI stereo data from NGA license agreements [14], the advantages of this approach become profound when conducting a sample study over a large area. Additionally, automated procedures using open access software (e.g., AMES Stereo Pipeline) could reduce the cost of conducting large-scale studies even further. At the same time, it should be noted the use of airborne LiDAR remains the standard for vertical accuracy in CHM construction, and airborne LiDAR would also be preferred for higher relief areas.

6. Conclusions

We found IKONOS Geo stereo data to be a useful low cost LiDAR alternative for CHM generation in the CONUS with $R^2$ 0.71 and of RMSE 2.6 m to G-LiHT LiDAR acquisitions when these data are collected one-year apart, and where a 3 m NED DTM was available. The high $R^2$ for DSMs in all three study regions was indicative of including terrain with a broad height measurement range of 40 to 100 m as compared to CHMs with a range of 0 to 30 m. Harvard Forest had the largest DSM range of 100 m, and had the largest acquisition difference of eleven-years. This broad range of values obscured the eleven-year ~10 m height differences in data acquisitions. However, this acquisition difference dominated the CHM signal producing as expected low $R^2$. Site characteristics of slope, topographic range, and variable forest growth rates from edaphic conditions impacted our results. Quality of GCPs also impacted our results due to differences in GCP errors between sites. Variable rates of regrowth between stands (young vs. older stands, different forest types, etc. can be a source of random error (not only bias) when the two dates of acquisitions are different and must be considered when using these data in future studies.

The dense archive of stereo HRSI, primarily from WorldView-1, over large portions of the globe could provide a much-needed tool to understand forest C storage and change when coupled with Landsat. These data may also be a desirable alternative to airborne LiDAR in remote regions of the Earth that are difficult to access. The current lack of a spaceborne LiDAR has prompted the use of stereo HRSI, with the sacrifice of absolute precision in forest canopy height measurements. However, WorldView-1 offers a resolution advantage over IKONOS of 42 cm resolution at nadir and 5 m
geo-location accuracy from ephemeris without GCPs, and could prove to be more effective in DSM generation for forest canopies. A comparison of WorldView-1 DSMs processed at multiple resolutions with different image processing and photogrammetric software packages that have different image matching routines (Geomatica, IDL/ENVI, Socketset GXP, SAT-PP, etc.) and G-LiHT products is needed to understand WorldView-1 accuracy. Future studies could task commercial instruments for regions of interest and automate HRSI CHMs into sampling methods with FIA assessments to monitor and model C stocks that are critical to meet United Nations climate change treaty reporting goals.

Acknowledgments

This study was made possible by NASA’s Terrestrial Ecology program under grants NNH08ZDA001N-TE and NNH10ZDA001N-CARBON. We would also like to thank three anonymous reviewers who enhanced the quality and clarity of our manuscript.

Author Contributions

Christopher Neigh proposed and developed the research design, collected stereo IKONOS data, performed the data comparison analysis, results interpretation, manuscript writing, and coordinated the revision activities. Jeffrey Masek obtained funding, assisted with developing the research design, results interpretation, and manuscript writing. Paul Bourget collected NED DTM data, produced DSMs and CHMs, assisted with manuscript writing and helped refine the research design. Bruce Cook provided G-LiHT data, interpreted results and assisted with manuscript writing. Chengquan Huang interpreted results and assisted with manuscript writing. Khalidoun Rishmawi assisted with refining the research design, interpretation of results and manuscript revision. Feng Zhao provided Landsat VCT and assisted with interpretation of results.

Conflict of Interest

The authors declare no conflict of interest.

References


© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).