VERTICAL ACCURACY ASSESSMENT OF 30-M RESOLUTION ALOS, ASTER, AND SRTM GLOBAL DEMS OVER NORTHEASTERN MINDANAO, PHILIPPINES

J. R. Santillan*, M. Makinano-Santillan*\textsuperscript{a,b}

\textsuperscript{a} CSU Phil-LiDAR 1 Project, Caraga Center for Geo-Informatics College of Engineering and Information Technology, Caraga State University Ampayan, Butuan City, Agusan del Norte, Philippines - jrsantillan@carsu.edu.ph

\textsuperscript{b} Division of Geodetic Engineering College of Engineering and Information Technology, Caraga State University Ampayan, Butuan City, Agusan del Norte, Philippines - mmsantillan@carsu.edu.ph

\textsuperscript{∗} Commission IV, WG IV/3

KEY WORDS: AW3D30, ASTER GDEM2, SRTM-30m, Digital Elevation Model, Accuracy Assessment, Philippines

ABSTRACT:

The ALOS World 3D - 30m (AW3D30), ASTER Global DEM Version 2 (GDEM2), and SRTM-30m are Digital Elevation Models (DEMs) that have been made available to the general public free of charge. An important feature of these DEMs is their unprecedented horizontal resolution of 30-m and almost global coverage. The very recent release of these DEMs, particularly AW3D30 and SRTM-30m, calls for opportunities for the conduct of localized assessment of the DEM’s quality and accuracy to verify their suitability for a wide range of applications in hydrology, geomorphology, archaeology, and many others. In this study, we conducted a vertical accuracy assessment of these DEMs by comparing the elevation of 274 control points scattered over various sites in northeastern Mindanao, Philippines. The elevations of these control points (referred to the Mean Sea Level, MSL) were obtained through 3rd order differential levelling using a high precision digital level, and their horizontal positions measured using a global positioning system (GPS) receiver. These control points are representative of five (5) land-cover classes namely brushland (45 points), built-up (32), cultivated areas (97), dense vegetation (74), and grassland (26). Results showed that AW3D30 has the lowest Root Mean Square Error (RMSE) of 5.68 m, followed by SRTM-30m (RMSE = 8.28 m), and ASTER GDEM2 (RMSE = 11.98 m). While all the three DEMs overestimated the true ground elevations, the mean and standard deviations of the differences in elevations were found to be lower in AW3D30 compared to SRTM-30m and ASTER GDEM2. The superiority of AW3D30 over the other two DEMs was also found to be consistent even under different landcover types, with AW3D30’s RMSEs ranging from 4.29 m (built-up) to 6.75 m (dense vegetation). For SRTM-30m, the RMSE ranges from 5.91 m (built-up) to 10.42 m (brushland); for ASTER GDEM2, the RMSE ranges from 9.27 m (brushland) to 14.88 m (dense vegetation). The results of the vertical accuracy assessment suggest that the AW3D30 is more accurate than SRTM-30m and ASTER GDEM2, at least for the areas considered in this study. On the other hand, the tendencies of the three DEMs to overestimate true ground elevation can be considered an important finding that users of the DEMs in the Philippines should be aware of, and must be considered into decisions regarding use of these data products in various applications.

1. INTRODUCTION

The ALOS World 3D - 30m (AW3D30), ASTER Global DEM Version 2 (GDEM2), and SRTM-30m are Digital Elevation Models (DEMs) that have become available to the general public free of charge. An important feature of these DEMs is their unprecedented horizontal resolution of 30-m and almost global coverage. The very recent release of these DEMs, particularly AW3D30 and SRTM30, calls for opportunities for the conduct of localized assessment of the DEM’s quality and accuracy to verify their suitability for a wide range of applications in hydrology, geomorphology, archaeology, and many others. On the other hand, assessments of the DEM’s accuracy in many different locations throughout the world are critical for improving the next generation of global DEMs (Suwandana et al., 2014).

Although numerous studies have been carried out for accuracy assessments of DEMs in different parts of the world using various kinds of reference data and reference DEMs (e.g., Arefi and Reinhartz, 2011; Hirt et al., 2010; Gomez, et al., 2012; Li et al., 2013; Athmania and Achour, 2014; Suwandana et al., 2014; Jing et al., 2014; Ioannidis et al., 2014; Satge et al., 2015), very few have been conducted in the Philippines (e.g., Fabila and Paringit, 2012; Meneses III, 2013). This is despite the fact that DEMs such as those from SRTM and ASTER are being used as major sources of topographic information for many applications including hydrological analysis and simulations (e.g., Jaranilla-Sanchez et al., 2011; Santillan et al., 2011; Sarmiento et al., 2012; Cituario and David, 2014; Chen and Senarath, 2014), flood modelling and hazard mapping (e.g., Abon et al., 2011; Ignacio and Henry, 2013), geological hazard analysis (e.g., Lagmay et al., 2012), and landslide mapping characterization (e.g., Evans et al., 2006; Oh and Lee, 2011). The quality and accuracy of the DEMs used and their suitability for these applications were not adequately assessed.

In this paper, we present the results of our vertical accuracy assessment of the AW3D30, ASTER GDEM2 and SRTM-30m DEMs covering Northeastern Mindanao, Philippines (Figure 1). The assessment aims to characterize the accuracy of the DEMs using such measures as the Root Mean Square Error (RMSE) and Mean Error. The effect of varying land-cover on elevation accuracy was also assessed.

To our knowledge, this paper is the first to report a vertical accuracy assessment of these specific DEMs covering Mindanao, Philippines.
Figure 1: Series of maps showing the study area, the location of the ground control points, and the three DEMs subjected to vertical accuracy assessment. Numerical values in (a.) indicate transect numbers as described in Table 1.
2. BRIEF DESCRIPTIONS OF THE DEMS

2.1 AW3D30

The AW3D30 was released in 2015 by the Japan Aerospace Exploration Agency (JAXA), and can be downloaded free of charge from http://www.eorc.jaxa.jp/ALOS/en/aw3d30/. The AW3D-30 is actually a resampling of the 5-meter mesh version of the World 3D Topographic Data, which is considered to be the most precise global-scale elevation data at this time (JAXA, 2015). AW3D-30 was generated using the traditional optical stereo matching technique as applied to images acquired by the Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM) sensor onboard the Advanced Land Observing Satellite (ALOS) (Takaku et al., 2014). Details on how the DEM was generated are discussed in the papers of Tadono et al (2014) and Takaku et al (2014). Due to its very recent release, studies assessing the vertical accuracy of AW3D30 are few or are yet to be reported. On the other hand, the accuracy of the 5-m mesh version of this DEM (AW3D-5m) have been reported in a few studies. Tadono et al. (2014) found the AW3-5m to have height accuracies better than 5m in four test sites with varying terrain features while Takaku et al. (2014) found the same DEM version to have a Root Mean Square Error (RMSE) of almost 4 m based on comparisons with various datasets including airborne LiDAR Digital Surface Model (DSM) and ground control points (GCPs). Recent assessment conducted by Tadono et al. (2015) confirmed an RMSE of 4.10 m.

2.2 SRTM-30m

The SRTM-30m (“SRTM V3.0, 1 arcsec”) is an enhancement to the low resolution SRTM topographic data having 90-m (3 arc-seconds, which is 1/1200th of a degree of latitude and longitude) resolution covering regions outside the United States (US) which was released publicly in 2003. The new data, released in September 2014, increase the detail to 30-m (or 1 arc-second), revealing the full resolution of the world’s landforms as originally measured by SRTM in the year 2000 (NASA JPL, 2014). Before this release, the best available 90-m SRTM DEMs for regions outside the US were: (i.) SRTM Version 3 (also called “SRTM Plus”) released by the National Aeronautics Space Administration (NASA) in November 2013 (NASA LP DAAC, 2013); and (ii.) CGIAR-CSI SRTM Version 4.1 released by the Consultative Group for International Agricultural Research – Consortium for Spatial Information (CGIAR -CSI) in 2008 (Jravits et al., 2008). According to its mission objectives, SRTM DEMs are expected to have linear vertical absolute height error of less than 16 m, linear vertical relative height error of less than 10 m, circular absolute geolocation error of less than 20 m, and circular relative geolocation error of less than 15 m (Farr et al., 2007). SRTM-30m accuracy assessments conducted by NIMA, the USGS, and the SRTM project team have shown the absolute vertical error to be much smaller, with the most reliable estimates being approximately 5 m (Kellndorfer et al., 2004).

2.3 ASTER GDEM2

The ASTER GDEM Version 2 was considered to be the highest resolution DEM among the free accessible global DEMs during its release in 2011 (Arefi and Reinfartz, 2011). The ASTER GDEM v2 contains significant improvements of Version 1 (released in 2009) in terms of spatial coverage, refined horizontal resolution, increased horizontal and vertical accuracy, water masking, and inclusion of new ASTER data to supplement the voids and artifacts (NASA JPL, 2011). Although vastly improved, some artifacts still exist in the form of abrupt rise (humps/bumps) and fall (pits) which can produce large elevation errors on local scale (Arefi and Reinfartz, 2011). Compared to AW3D30 and SRTM-30m, studies assessing the quality and vertical accuracy of ASTER GDEM v2 are many (e.g., Takikawa et al., 2011; Gesch et al., 2012; Athmania and Achour, 2014; Suwandana et al., 2014). In Japan, the ASTER GDEM2 was reported to have an RMSE of 6.1 m in flat and open areas, and 15.1 m in mountainous area largely covered by forest (Tachikawa et al., 2011). In the conterminous US, the RMSE computed for GDEM2 was 8.68 m based on the comparison with more than 18,000 independent reference ground control points (Gesch et al., 2012). An external validation conducted by Athmania and Achour (2014) shows the GDEM2 to have RMSE of 5.3 and 9.8 m in test sites located in southern Tunisia and in northeastern Algeria, respectively. In Banten province, Indonesia, RMSE values ranging from 4.543 to 7.759 m was computed by Suwandana et al (2014). The results of these example studies show that the accuracy of ASTER GDEM2 varies from one location to another. Hence, localized or site-specific accuracy assessment of the ASTER GDEM2 is very important.

3. DATASETS AND METHODS

3.1 DEMs

The AW3D30 DEM of northeastern Mindanao (Figure 1b) was downloaded from http://www.eorc.jaxa.jp/ALOS/en/aw3d30/. The downloaded data is a Beta Version (V15.05) that was released by JAXA in May 2015. It consisted of four 1x1 degree lat/long tiles in GeoTIFF format: N008E125, N008E126, N009E125 and N009E126. For each tile, the DEM was provided in two types: AVE and MED according to the method used when resampling from the 5-meter mesh version (AVE = average, MED = median). We opted to use the AVE tiles. All the tiles were mosaicced and saved in GeoTIFF format using Global Mapper software, and reprojected from WGS 1984 geographic coordinate system to Universal Transverse Mercator Zone 51 projection (retaining WGS 1984 as its horizontal datum) using ArcGIS 9.3 software. The elevation values in the AW3D30 are considered “height above sea level” (JAXA, 2015). Missing data due to cloud cover is evident in the AW3D30 DEM (shown as white gaps in Figure 1b).

The SRTM-30m and ASTER GDEM2 (Figures 1c and 1d) were both downloaded from LP DAAC Global Data Explorer (http://gdex.cr.usgs.gov/gdex/). In GeoTIFF format with Universal Transverse Mercator Zone 51 (UTM51) projection and the World Geodetic System (WGS) 1984 as horizontal datum. They both have the Earth Gravitational Model 1996 (EGM96) as vertical datum.

3.2 Reference Elevation Data

Reference data used in the analysis consisted of 12 transects consisting of 274 ground control points or GCPs (Table 1) located in various sites in northeastern Mindanao, Philippines (Figure 1a). The GCPs have elevations ranging from 1.76 to 61.14 meters from the Mean Sea Level (MSL). For each transect, the ground elevations at the control points were obtained through 3rd order differential leveling using a high precision digital level (FOIF EL302A). Differential levelling is a vertical surveying technique of measuring vertical distances from a known elevation point to determine elevations of unknown points (Anderson and Mikhail, 1998). For this study, we implemented a
closed-loop leveling survey, i.e., starting from a known elevation point and closing or returning to the same known elevation point. We used vertical control points/benchmarks established by the Philippines National Mapping and Resource Information Authority (NAMRIA) as starting/closing reference points in the conduct of our levelling surveys. The surveys strictly followed the procedures, standards and specification for Third Order Geodetic Levelling set upon by the Federal Geodetic Control Committee (FGCC). Details of these standards and specification can be viewed at http://www.ngs.noaa.gov/FGCS/tech_pub/1984-stds-specs-geodetic-control-networks.htm. The accuracy of the leveling survey conducted for each transect was assessed by checking the maximum loop closures (MLC) not to exceed 12mm/√D (Anderson and Mikhail, 1998), where D is the loop distance in km (or approximately twice the transect length). The MLC is computed by getting the difference between actual and survey-derived elevation values of the closing reference control point.

The horizontal positions (WGS84 latitude and longitude) of the GCPs were determined using a Garmin 550t handheld global positioning system (GPS) receiver. At each GCP, the geographic coordinates were measured through time-based averaging (minimum of 2 minutes observation time) until the positional accuracy indicated in the receiver is less than 10 m. A shapefile was generated from the gathered GCPs, and it was re-projected to UTM51 using ArcGIS 9.3 software.

The control points were established in relatively stable areas (e.g., roads, pavements, bridges, and other similar concrete structures located within a particular land-cover type) which are assumed to have been present from year 2000 onwards and have not changed through time. The control points are representative of five (5) land-cover classes namely brushland (45 points), built-up (32), cultivated areas (97), dense vegetation (e., forests, palm vegetation, and mangroves; 74), and grassland (26). Since the DEMs were generated using data gathered in the year 2000 onwards, we find it appropriate to use the best available land-cover map for the entire Philippines produced by the NAMRIA for the year 2003 (scale of 1:250,000) to group the GCPs according to land-cover types.

### 3.3 Vertical Accuracy Assessment

Similar to the accuracy assessment procedures implemented by Gesch et al. (2012), vertical accuracies of the three DEMs were assessed by comparing the DEM elevations with those of the GCPs. At each point, the DEM elevations were extracted using ArcGIS 9.3 software. Then, the differences in elevation were computed by subtracting the GCP elevation from its corresponding DEM elevations, and these differences are the measured errors in the DEMs. For a particular DEM, positive errors represent locations where the DEM was above the GCP elevation, and negative errors occur at locations where the DEM was below the control point elevation. From these measured errors, the mean error and RMSE for each DEM were calculated, including standard deviations of the mean errors. The mean error (or bias) indicates if a DEM has an overall vertical offset (either positive or negative) from true ground level (Gesch et al., 2012). Finally, accuracy assessment results were analyzed by land-cover types to look for relationships between vertical accuracy and cover type.

### 4. RESULTS AND DISCUSSION

#### 4.1 Vertical Accuracies of the DEMs

Shown in Figure 2 are the calculated errors of the DEMs plotted with the actual elevation of the GCPs, while the summary of computed error statistics are listed in (Table 2).

In general, there is no clear relationship between the calculated errors and elevation for all DEMs. It cannot be said that the errors in the DEM increases with elevation or otherwise. On the other hand, the calculated errors are not uniformly distributed on both sides of the error axis. In fact, majority of the errors are greater than zero (i.e., biased positively). This means that all the DEMs overestimated ground elevations in majority of the GCPs. This is confirmed by the positive values of mean errors for all DEMs.

Among the three DEMs, AW3D30 exhibited the lowest mean error and RMSE values of 4.36 and 5.68 m, respectively. AW3D30’s errors also have the lowest standard deviation of 3.66 m. Majority of the AW3D30 errors are within the 0-10 m range.

The SRTM-30m DEM is next to AW3D30 in terms of accuracy, with mean error and RMSE values of 6.91 and 8.28 m, respectively. Looking at the error plots (Figure 2), the distribution of SRTM-30m errors have almost similar pattern to that of AW3D30. One distinguishing characteristic is the wider range of errors compared to AW3D30, with majority of the SRTM-30m’s errors within the 0-20 m range.

Among the three DEMs, ASTER GDEM2 has the highest mean error and RMSE values of 8.37 and 11.98 m, respectively. Majority of the ASTER GDEM2 errors ranges from 0-30 m, with a large standard deviation of 8.58 m. Regardless of elevation, ASTER GDEM2’s errors were -11.92 m at the minimum, and 39.27 m at the maximum.

#### 4.2 Land-cover Effects on DEM Accuracy

The mean error and RMSE values of the DEMs grouped according to land-cover type are shown in Figure 3. The mean error and RMSE reflect the effects of land-cover on the measurement of elevation by the three DEMs. It is noticeable that there is an almost linear relationship between the mean error and RMSE regardless of land-cover type.

For AW3D30, high mean errors were found for grassland followed by brushland and dense vegetation. However, looking at the error bars which represents the 95% confidence interval of the mean, it can be said that the errors associated to these land-cover types are not unique due to overlaps in the 95% CI values. On the other hand, low mean errors and RMSEs were found in relatively-open terrains represented by built-up and cultivated areas. Again, the errors in these two land-cover types cannot be uniquely differentiated due to overlapping 95% CI of the means. The positive values of these mean errors regardless of land-cover type mean that the overestimation of the true ground elevation by AW3D30 is consistent across different land-cover types. For ASTER GDEM2, elevation errors are more pronounced in densely vegetated areas, grassland and built-up, but the differentiation between the effects of these land-cover types is hard to determine due to overlapping 95% CI of the means; relatively low errors were found for brushland and cultivated areas. For SRTM-30m, the effects of land-cover is similar to that of AW3D30 (i.e., high errors in brushland, grassland and dense vegetation; low errors in built-up and cultivated areas). Again, due to overlapping 95% CI values of the mean error, we cannot pinpoint which among the land-cover types have the greatest effect on SRTM-30m’s elevation accuracy.

For easier comparison, the mean errors and RMSEs are plotted such that error values of each DEM is plotted side-by-side of each other (Figure 4). Among the three DEMS, AW3D30 has the lowest mean errors and RMSEs in all land-cover types, while
Table 1: Ground control points used in the DEM vertical accuracy assessments, grouped by transect. Refer to Figure 1 for their locations.

<table>
<thead>
<tr>
<th>Transect No.</th>
<th>Transect Length, m.</th>
<th>No. of Points</th>
<th>Average Distance Between Points, m.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9,973</td>
<td>22</td>
<td>475</td>
</tr>
<tr>
<td>2</td>
<td>3,097</td>
<td>6</td>
<td>619</td>
</tr>
<tr>
<td>3</td>
<td>2,767</td>
<td>6</td>
<td>553</td>
</tr>
<tr>
<td>4</td>
<td>3,227</td>
<td>7</td>
<td>538</td>
</tr>
<tr>
<td>5</td>
<td>7,172</td>
<td>12</td>
<td>652</td>
</tr>
<tr>
<td>6</td>
<td>14,204</td>
<td>32</td>
<td>458</td>
</tr>
<tr>
<td>7</td>
<td>12,518</td>
<td>42</td>
<td>305</td>
</tr>
<tr>
<td>8</td>
<td>20,883</td>
<td>52</td>
<td>409</td>
</tr>
<tr>
<td>9</td>
<td>3,681</td>
<td>11</td>
<td>368</td>
</tr>
<tr>
<td>10</td>
<td>6,515</td>
<td>21</td>
<td>326</td>
</tr>
<tr>
<td>11</td>
<td>5,629</td>
<td>32</td>
<td>182</td>
</tr>
<tr>
<td>12</td>
<td>11,268</td>
<td>31</td>
<td>376</td>
</tr>
</tbody>
</table>

ASTER GDEM2 has the highest except in the case of brushlands where the mean error and RMSE of SRTM-30m were higher than that of ASTER GDEM2.

4.3 Discussion

An important finding of this study is that the three DEMs overestimated the true ground elevations regardless of land-cover types. The magnitude of overestimation varies according to the DEM. In terms of vertical accuracy, it is very clear that AW3D30 outperformed SRTM-30m and most especially ASTER GDEM2 due to the former’s lower mean errors and RMSE values compared to the latter DEMs. On the average, AW3D30 overestimates ground elevation by 4.36 m, 6.91 m by SRTM-30m, and 8.37 m by ASTER GDEM2. These overestimations can be expected as ALOS, ASTER and SRTM are first return systems that measure aboveground elevations (Tadono et al., 2014; Gesch et al., 2012). The computed 5.68 m RMSE of AW3D30 is slightly higher than the expected vertical accuracy of the ALOS World 3D which is 5 m (RMSE). The computed mean errors are also slightly higher than the errors computed by Takaku et al (2014) in their preliminary assessment of the DEM where they calculated an average error of 2.08 m and RMSE of 3.94 based on 122 GCPs. How-
ever, these expected accuracy and the computed errors are for the 5-m version of the DEM. While it is not yet clear how decreasing the spatial resolution to 30-m affected the vertical accuracy of AW3D, the computed mean error and RMSE are indications that indeed this DEM offers a more accurate depiction of the ground elevation compared to SRTM-30m and ASTER GDEM2.

The results for SRTM-30m shows that its accuracy is better than the expected mean error of 10 m and RMSE of 16 m (Farr et al., 2007). This also confirms earlier assessments conducted by NIMA, the USGS, and the SRTM project team showing the absolute vertical error to be much smaller (Kellendörfer et al., 2007).

The results for ASTER GDEM2 add to the many literatures reporting the low accuracy of this DEM (e.g., Suwandana et al., 2012; Athmania and Achour, 2014). The high mean error, standard deviation and RMSE computed in this study maybe indications of the presence of voids and artifacts in the DEM that may have been captured by the GCPs used in the analysis.

The results of the analysis on the effects of land-cover on DEM elevation accuracy were found to be inconsistent with what have been published in DEM accuracy assessment studies, particularly those focused on SRTM and ASTER GDEM2. For example, in the assessment conducted by Gesch et al. (2012), a clear relationship between land-cover types and ASTER GDEM and SRTM-30m accuracies were found, i.e., errors in elevation increased as the land-cover changes from unvegetated to fully vegetated. In their study, positive bias was found in GCPs locations dominated by forests. Moreover, as land cover becomes more open, the ASTER GDEM2 and SRTM-30m RMSE values were nearly equivalent as these DEMs are measuring near ground level elevations (Gesch et al., 2012). In the present study, these findings were not encountered. While there are indications that indeed land-cover types affected DEM accuracy, a clear relationship between the two appears to be inexistent. However, this does not mean that this relationship cannot exist at all for DEMs covering the Philippines due to the study’s limitations in the number of GCPs and the relatively coarse land-cover map used in grouping the GCPs.

5. CONCLUSIONS

Our vertical accuracy assessment using 274 GCPs in northeastern Mindanao, Philippines shows that AW3D30, ASTER GDEM2 and SRTM-30m overestimated true ground elevations. The tendencies of the three DEMs to overestimate elevation can be considered an important finding that users of the DEMs in the Philippines should be aware of, and must be considered into decisions regarding application of these data products.

Among the three, AW3D30 was found to be the most accurate in depicting true ground elevations as this DEM has the lowest mean error, RMSE and standard deviation among the three. It is followed by SRTM-30m and ASTER GDEM2. The superiority of AW3D30 over the other two DEMs was also found to be consistent even under different landcover types.

A limitation of this study is the use of transect points instead of spatially-distributed points such that a comprehensive evaluation of the DEMs’ accuracies was not fully done. Another limitation is the narrow range of elevations of the GCPs considered in the analysis, which only ranged from 1.76 to 61.14 m. It is not yet clear if the present findings will remain valid if GCPs with elevations greater than 61.14 m are used. To address these limitations, a follow-up study is needed and should involve establishment of additional number of GCPs that are spatially distributed over the study area, and with a wider range of elevation values.

ACKNOWLEDGEMENTS

This work was conceptualized and implemented during the early stages of the Caraga State University (CSU) Phil-LiDAR 1 project under the "Phil-LiDAR 1. Hazard Mapping of the Philippines using LiDAR” program funded by the Department of Science and Technology (DOST). We acknowledge the support of the CSU Phil-LiDAR 1 Field Survey Team headed by Engr. Ronald M. Makinano in the conduct of vertical surveys. ASTER GDEM is a product of METI and NASA.

We are thankful to the DOST and the Philippine Council for Industry, Energy and Emerging Technology Research and Development (PCIEERD) for the financial assistance that allowed us to present this paper during the XXIII ISPRS Congress 2016.

REFERENCES

Abon, C.C., David, C.P.C., Pellejera, N. E. B., 2011. Reconstructing the tropical storm Ketsana flood event in Marikina


