# DETECTION AND CORRECTION OF ASTER GDEM v2 DATA ANOMALIES THROUGH DEM DIFFERENCING

Jojene R. Santillan[1] and Meriam M. Santillan[1, 2]
[1] Caraga Center for Remote Sensing and GIS, College of Engineering and Information Technology, Caraga State University, Ampayon, Butuan City, 8600, Philippines
[2] Division of Geodetic Engineering, College of Engineering and Information Technology, Caraga State University, Ampayon, Butuan City, 8600, Philippines
Email: santillan.jr2@gmail.com, meriam.makinano@gmail.com

**KEY WORDS:** ASTER GDEM v2, SRTM DEM, anomaly detection, DEM correction.

**ABSTRACT:** The ASTER GDEM Version 2 (v2) which was released in 2011 is considered to be the highest resolution and readily available global digital elevation model. Having a spatial resolution of 30-m, the ASTER GDEM v2 contains significant improvements of Version 1 in terms of spatial coverage, refined horizontal resolution, water masking, and inclusion of new ASTER data to supplement the voids and artifacts. Despite of these improvements, data anomalies such as abrupt rise ("bumps") and fall ("pits") in elevation values in the ASTER GDEM v2 still remains. In this paper, we present a simplified and semi-automatic approach of detecting and correcting data anomalies in the ASTER GDEM v2 of Tago River Basin in Surigao del Sur, Mindanao, Philippines. The anomaly detection procedure consisted of calibrating the elevation values in the ASTER GDEM v2 against a resampled SRTM DEM version 4.1, creating a difference image between the calibrated ASTER GDEM and SRTM DEM, applying a low pass filter to the difference image, employing the K-means clustering algorithm to classify the difference image pixels into various classes, and then labeling them as bumps, pits or neither (i.e., not anomaly) using a known set of pixels pre-selected prior to K-means classification. The accuracy of the detection was at 99.47% based on an independent set of randomly selected validation pixels. The detected anomalies were then masked out from ASTER GDEM v2, and the elevation values of these pixels were replaced by elevation values extracted from the resampled SRTM DEM. The corrected DEM was used for hydrological applications such as computations of flow direction and flow accumulation grids, sub-basin delineations and stream network definition. These DEM derivatives are important inputs in hydrological model development for the river basin.

## 1. INTRODUCTION

The ASTER GDEM Version 2 (v2) which was released in 2011 is considered as the highest resolution digital elevation model (DEM) among the free accessible global DEMs (Arefi and Reinartz, 2011). Having a spatial resolution of 30-m, the ASTER GDEM v2 contains significant improvements of Version 1 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 (JPL, 2011). Despite of these improvements, data anomalies such as abrupt rise ("humps/bumps") and fall ("pits") in elevation values in the ASTER GDEM v2 still remains which can produce large elevation errors on local scales (Arefi and Reinartz, 2011). These anomalies can impede effectiveness of the ASTER GDEM v2 for use in certain applications. For example, in hydrologic modeling, anomaly-free DEMs are needed in the generation of flow direction and accumulation grids, basin and sub-basin delineations, and stream network definition. It becomes necessary to detect and correct the data anomalies in the DEM prior to its use to avoid the generation of erroneous flow direction and accumulation grids that can affect the accurate delineation of basin/sub-basin boundaries and stream networks.

Arefi and Reinartz (2011) proposed an algorithm for the removal of local artifacts and anomalies (referred to by the authors as "outliers") in the ASTER GDEM v2 using a segmentation-based method. The first step of the algorithm uses the ASTER GDEM as a "mask" image to create a "marker" image which is generated by subtracting the "mask" image by an offset value ($h$).The authors used a single offset value ($h$) of about 25 m based on their notion that the outliers are in most cases located far beyond the ground elevation level with low height variation of their internal pixels. The mask and the marker images are then used as inputs for image reconstruction using geodesic morphological dilation to extract regional extrema regions (e.g., bumps and pits). The reconstructed image is subtracted from the original ASTER GDEM to generate a normalized DEM which is then segmented using thresholding. The segmented regions are then evaluated as either outlier or not using the local range variation descriptor. The final step of the algorithm employed an Inverse Distance Weighting spatial interpolation procedure to fill the gaps provided by all the different outliers.

In this paper, we present an anomaly detection procedure similar to the concept of "image differencing" as proposed

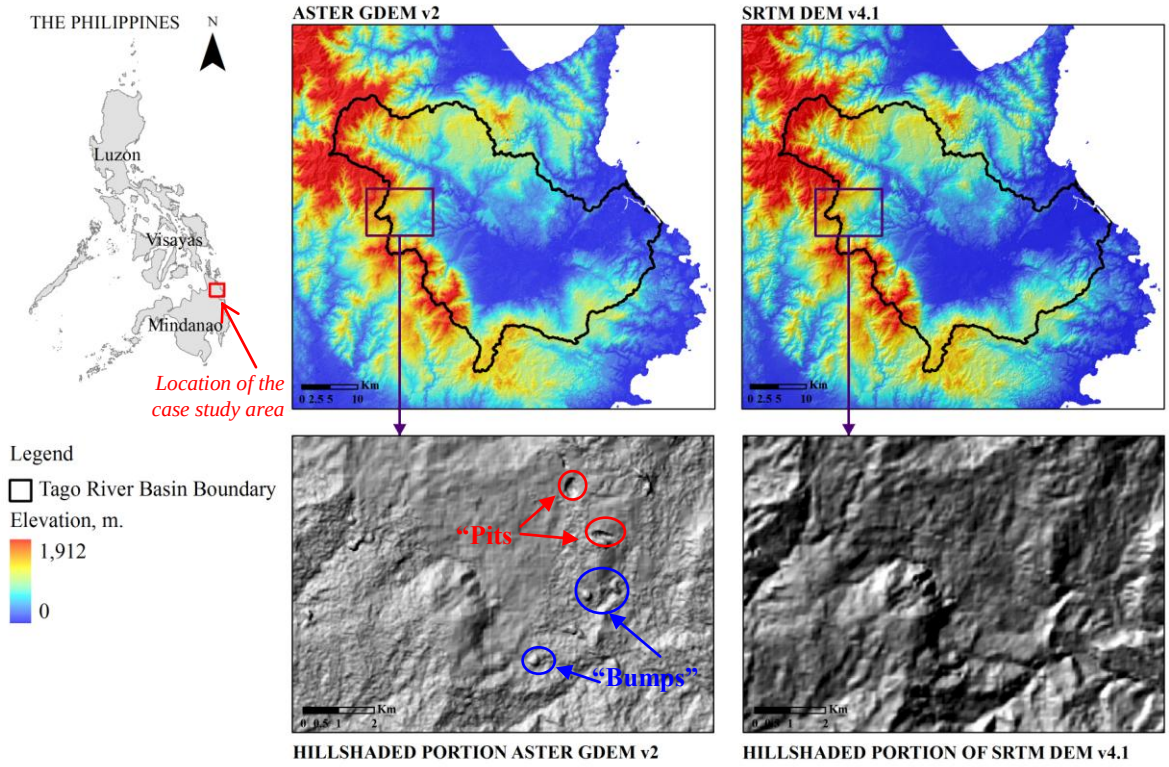


Figure 1. The ASTER GDEM v2 and SRTM DEM of Tago River Basin, Surigao del Sur, Philippines. Also shown are hillshades of the two DEMs in an area with data anomalies (as indicated).

by Arefi and Reinartz (2011). However, the approach does not utilized geodesic morphological dilation of the ASTER GDEM v2 and a marker image to create another image for segmentation, detection and removal. Instead, it involves the classification of a difference image resulting from the subtraction of a resampled SRTM DEM from the ASTER GDEM v2 in order to detect the anomalies, and the replacement of elevation values of the anomalous pixels with those coming from an SRTM DEM. This simplified and semi-automatic approach of detecting and correcting the data anomalies was applied to the ASTER GDEM v2 of Tago River Basin (TRB) in Surigao del Sur, Mindanao, Philippines.

## 2. MATERIALS AND METHODS

### 2.1 Conceptual Basis

The development of the anomaly detection procedure started upon the concept that the elevation of non-anomalous pixels in an ASTER GDEM v2 maybe systematically correlated with their elevations in an SRTM DEM, and that this correlation maybe represented by an equation. Such equation maybe utilized to transform or calibrate the elevation values in an ASTER GDEM v2 to correspond to their elevation in an SRTM DEM. When the SRTM DEM is subtracted from the calibrated ASTER GDEM v2, the computed difference image may highlight data anomalies in the form of high magnitude of differences in elevation values. Very large positive differences may represent bumps while very large negative differences may represent pits. It follows then that this difference image can be analyzed to map the presence of anomalous pixels.

For this study, "pits" are those pixels in the ASTER GDEM v2 that occur as negative elevation anomalies which vary from a few meters to about 100 meters while "bumps" are those pixels that appear as positive elevation anomalies whose magnitude can range from just few meters to more than 100 meters (Arefi and Reinartz, 2011).

### 2.2 Datasets Used

We used ASTER GDEM v2 and SRTM DEM v4.1 datasets of Tago River Basin in Surigao del Sur, Mindanao, Philippines as case study area (Figure 1). The 30-m resolution ASTER GDEM v2 was downloaded from LP DAAC Global Data Explorer (http://gdex.cr.usgs.gov/gdex/) while the 90-m resolution STRM DEM v4.1 was downloaded

from the CGIAR Consortium for Spatial Information (CGIAR-CSI; http://srtm.csi.cgiar.org/Index.asp). The two DEMs were downloaded in Geotiff format with Universal Transverse Mercator (UTM) Zone 51 projection and the World Geodetic System (WGS) 1984 as horizontal datum. They both have the Earth Gravitational Model 1996 (EGM96) as vertical datum. Prior to the analysis, the SRTM DEM was resampled to 30-m resolution using cubic convolution in ArcGIS 9.3. It was then layerstacked with the ASTER GDEM using Envi 5 software to ensure pixel by pixel co-registrations. The two DEM datasets have dimensions of 2514 x 2379 pixels.

### 2.3 Anomaly Detection and Correction Procedure

The first step of the procedure is the calibration of ASTER GDEM v2 elevations to that of the resampled SRTM DEM. It has been reported by several studies that SRTM DEM v4.1 has better accuracy than ASTER GDEM v2 (Suwandana et al., 2012; Changwei, et al., 2013; Athmania and Achour, 2014). The calibration in this case aims to increase the vertical accuracy of the ASTER GDEM elevation values while maintaining its high spatial resolution. The calibration equation was developed through linear regression using a total of 690 points randomly selected within the boundary of both datasets. Points were not collected in areas where there are obvious occurrences of data anomalies. The ASTER GDEM and SRTM DEM elevations of these points were extracted and subjected to linear regression.

The next step is the creation of a difference image wherein the elevation values of pixels in the calibrated ASTER GDEM v2 are subtracted by their corresponding elevation values in the SRTM DEM. A 3x3 low pass filter was then applied to this difference image to preserve and smooth the low frequency connected pixels which are expected to represent the data anomalies.

The low pass filtered difference image was then subjected to K-means clustering to classify the pixels into various classes. Prior to this, two independent sets of pixels representing bumps, pits and non-anomalies were randomly selected from the calibrated ASTER GDEM v2. For both sets, a total of 125 pixels for each class were collected. One set (training) was utilized to label the various classes obtained from the K-means clustering while the other set (validation) was used for accuracy assessment.

All the pixels labeled as either bumps or pits were then masked-out from the calibrated ASTER GDEM v2. The elevation values of these pixels were replaced by their corresponding resampled SRTM DEM v4.1 elevation values through mosaicking. To ensure that there is a smooth transition in elevation values along the boundary of the detected anomalous regions of the calibrated ASTER GDEM v2, a feathering distance of 3 pixels was set during mosaicking.

A combination of Envi 5 and ArcGIS 9.3 software were used throughout the process.

## 3. RESULTS AND DISCUSSION

Figure 2 shows the results of the linear regression of the ASTER GDEM v2 and SRTM DEM v4.1 elevation values of 690 randomly selected points. The average computed difference in elevation values between the two DEM was at 5.16 m with a standard deviation of 11.09 m. It was observed that in the case study area, the ASTER GDEM v2 has higher elevation values compared to SRTM DEM v4.1.

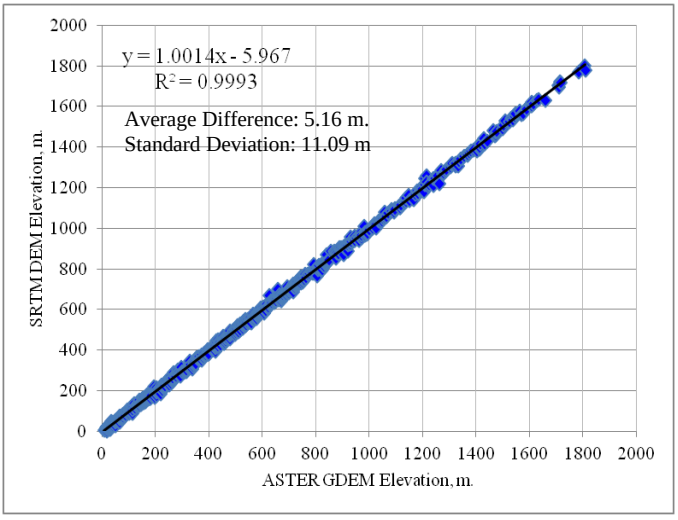


Figure 2. Results of linear regression of ASTER GDEM v2 and SRTM DEM v4.1 elevation values of 690 randomly-selected points.

The linear equation derived from the regression was then used to calibrate the elevation values of the ASTER GDEM v2. The difference image derived from the subtraction of the SRTM DEM v4.1 from the calibrated ASTER GDEM v2 and after application of a low pass filter is shown in Figure 3. It can be observed that the difference image highlighted bumps and pits in the ASTER GDEM v2 which are indicated by large differences in elevations between the two DEMs.

Figure 4 shows the results of the K-means clustering of the low pass-filtered difference image. Initially, there were a

total of 10 classes but after labeling each class using the training set, the clustered difference image was reduced to 3 classes: bumps, pits, and non-anomalies. The overall accuracy of the anomaly detection based on the comparison of the validation set was at 99.47% (Table 1). Of the 2514 x 2379 (or a total of 5,980,806) pixels covered in the analysis, 10.24% were detected as pits while 9.49% were detected as bumps. Although the accuracy is high, it was not clear if there were overestimation or underestimation in the anomaly detection especially that the validation points are on a per pixel/point basis and mostly collected near the center of the pre-identified anomalies. It is possible that the detection procedure may have overestimated by including pixels adjacent to the anomalous pixels. This may have occured during the K-means classification.

The final product of the anomaly detection and correction procedure is shown in Figure 5. All the anomalous pixels have been removed and their elevation values have been replaced by their corresponding SRTM DEM v4.1 elevations. The corrected ASTER GDEM v2 of the study area was used as input to HEC GeoHMS v1.1, an Arcview GIS 3.2 extension, to delineate sub-basins/watershed and stream networks. The results of these delineations will be utilized in the development of the HEC HMS-based hydrologic model of the river basin.

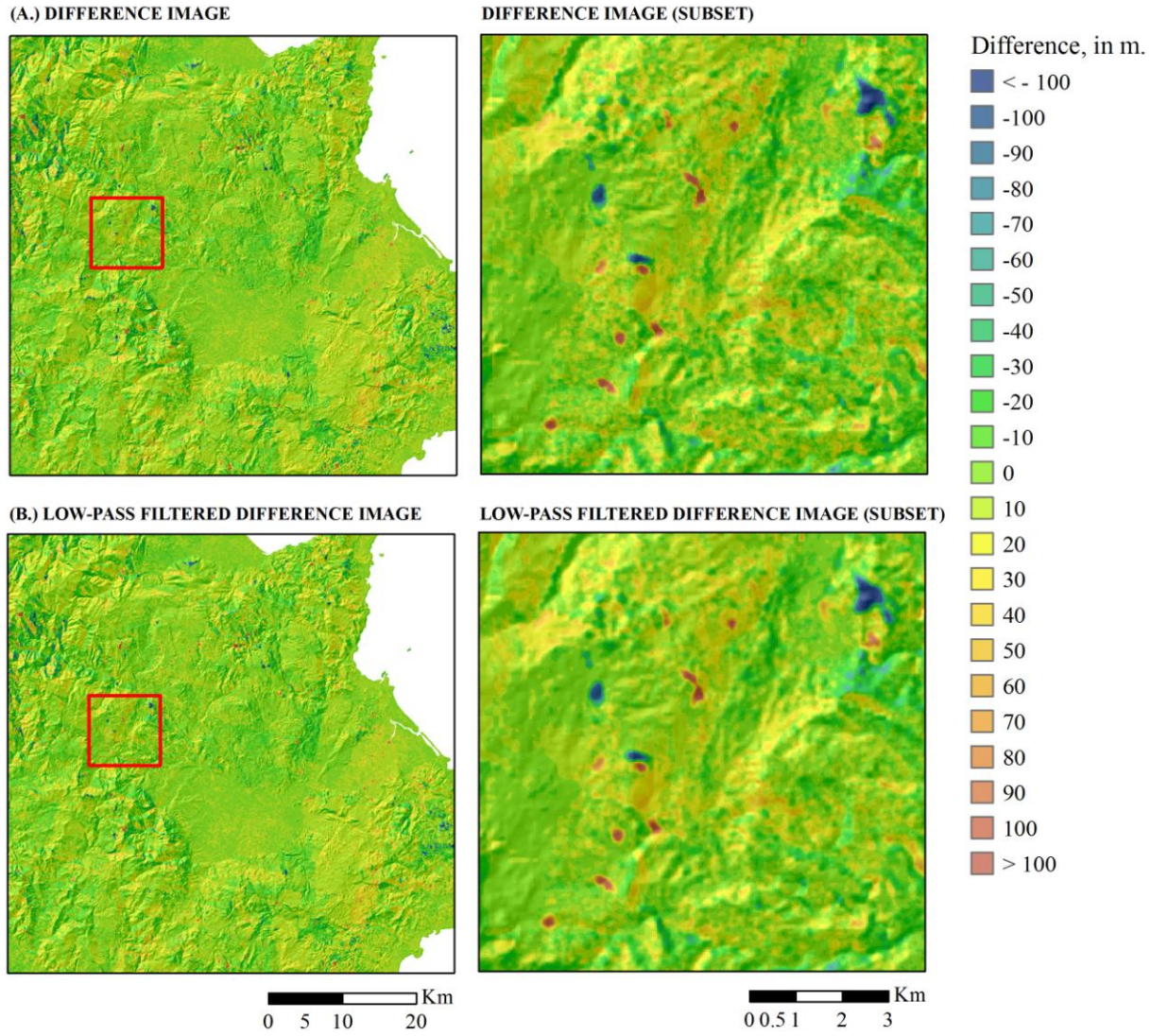


Figure 3. (A) The difference image derived by subtracting the calibrated ASTER GDEM v2 with the SRTM DEM v4.1. (B) The difference image after applying a 3x3 low-pass filter. The images on the right are zoomed-in views of the portion indicated by the red square on the left.

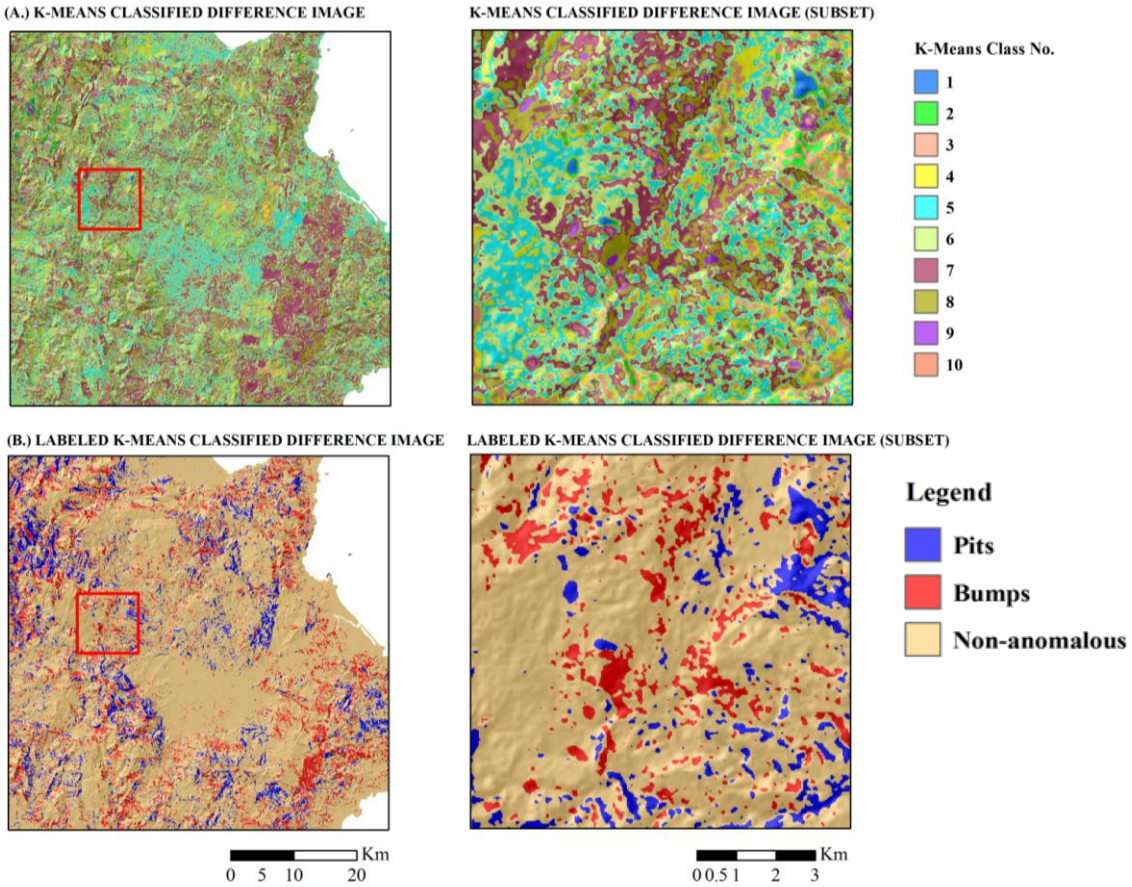


Figure 4. (A) The K-means classified difference image. (B) The labeled K-means classified difference image. The images on the right are zoomed-in views of the portion indicated by the red square on the left.

Table 1. Confusion or error matrix of the ASTER GDEM v2 anomaly detection based on DEM differencing and K-means classification.

| | | Validation Pixels (Actual) | | | | |
|---|---|---|---|---|---|---|
| | | Bumps | Pits | Non-anomalous | Total | User's Accuracy (%) |
| Classified Pixels | Bumps | **125** | 0 | 2 | 127 | 98.43 |
| | Pits | 0 | **125** | 0 | 125 | 100 |
| | Non-anomalous | 0 | 0 | **123** | 123 | 100 |
| | Total | 125 | 125 | 125 | 375 | |
| Producer's Accuracy (%) | | 100 | 100 | 98.40 | | |
| Overall Classification Accuracy (%) | | 99.47 (373/375) | | | | |
| Kappa coefficient | | 0.99 | | | | |

## 4. CONCLUSIONS

In this paper, we have presented a simplified and semi-automatic approach of detecting and correcting data anomalies in the ASTER GDEM v2 of Tago River Basin in Surigao del Sur, Mindanao, Philippines. The anomaly detection procedure consisted of calibrating the elevation values in the ASTER GDEM v2 against a resampled SRTM DEM v4.1, creating a difference image between the calibrated ASTER GDEM v2 and SRTM DEM v4.1, applying a low pass filter to the difference image, employing the K-means clustering algorithm to classify the difference image pixels into various classes, and then labeling them as bumps, pits or neither (i.e., not anomaly) using a known set of pixels pre-selected prior to K-means classification. The accuracy of the detection was at 99.47% based on an independent set of randomly selected validation pixels. The detected anomalies were then masked out from ASTER GDEM v2, and the elevation values of these pixels were replaced by elevation values extracted from the resampled SRTM DEM.

One of the limitations of this study is that we were not able to verify the integrity and accuracy of the SRTM DEM v4.1 used due to unavailability of ground truth elevation data during the time when this study was conducted. It may be possible that the SRTM DEM v4.1 may also have data anomalies that may have affected the calibration of the ASTER GDEM v2 or may have aggravated the data anomalies in the ASTER GDEM v2. These limitations will be addressed in a future study.

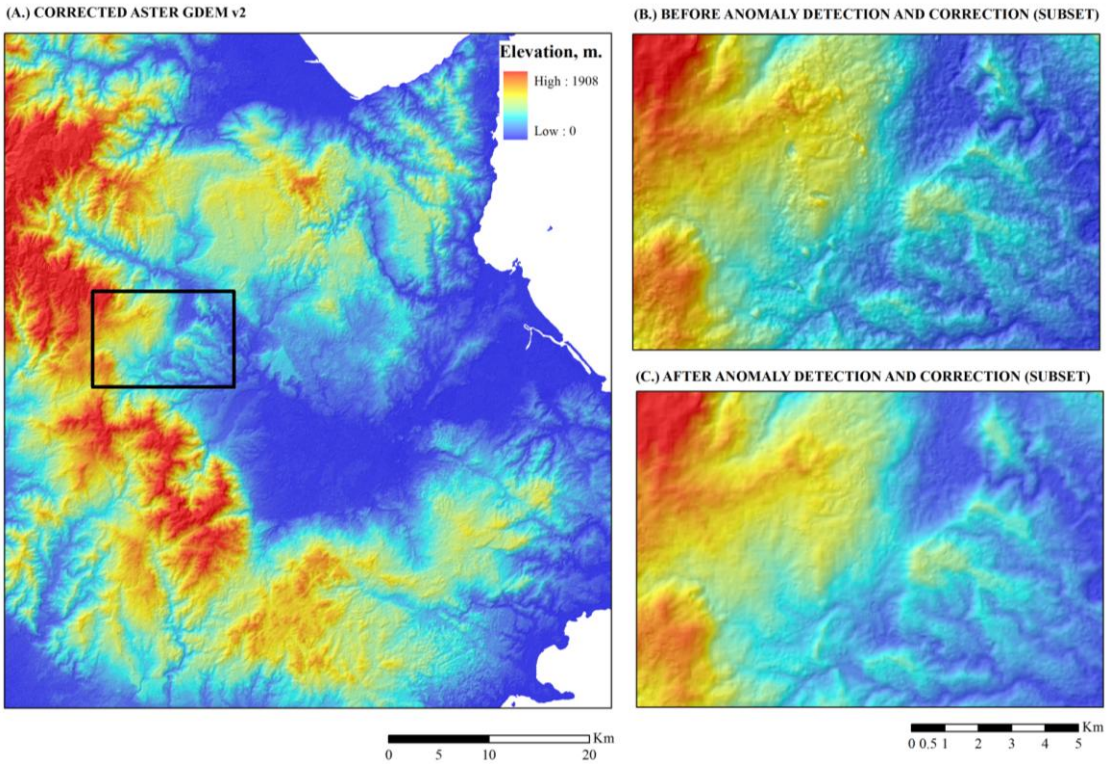


Figure 5. (A) The corrected ASTER GDEM v2 of the case study area. (B) A subset of the ASTER GDEM v2 before applying the anomaly detection and correction procedure. (C) The same subset after the anomalies have been detected, removed and corrected.

## REFERENCES


Arefi, H., Reinartz, P. 2011. Accuracy enhancement of ASTER global digital elevation models using ICESat data. Remote Sensing, 3(7), pp. 1323-1343.

Athmania, D., Achour, H., 2014. External Validation of the ASTER GDEM2, GMTED2010 and CGIAR-CSI-SRTM v4. 1 Free Access Digital Elevation Models (DEMs) in Tunisia and Algeria. Remote Sensing, 6(5), pp. 4600-4620.

Changwei, J. , Shortridge, A., Lin, S., Wu, J., 2013. Comparison and validation of SRTM and ASTER GDEM for a subtropical landscape in Southeastern China. International Journal of Digital Earth, pp. 1-24.

JPL, 2011. ASTER Global Digital Elevation Map Announcement. Retrieved May 14, 2014 from http://asterweb.jpl.nasa.gov/gdem.asp.

Suwandana, E., Kawamura, K., Sakuno, Y., Kustiyanto, E., Raharjo, B., 2012. Evaluation of ASTER GDEM2 in comparison with GDEM1, SRTM DEM and topographic-map-derived DEM using inundation area analysis and RTK-dGPS data. Remote Sensing, 4(8), pp. 2419-2431.


## ACKNOWLEDGEMENTS


This research is an output of the "Phil-LiDAR 1.2.14: LiDAR Data Processing and Validation: Caraga Region" project. We are grateful to the Philippine Council for Industry, Energy and Emerging Technology Research and Development of the Department of Science and Technology (PCIEERD-DOST) for the financial support.