Improving River Bathymetry and Topography Representation of a Low-Lying Flat River Basin by Integrating Multiple Sourced Datasets

Seemanta Sharma Bhagabati∗,†, Akiyuki Kawasaki∗, Wataru Takeuchi∗∗, and Win Win Zin∗∗∗

∗Department of Civil Engineering, The University of Tokyo
†Corresponding author, E-mail: seemanta@hydra.t.u-tokyo.ac.jp
∗∗Industrial Institute of Science, The University of Tokyo, Tokyo, Japan
∗∗∗Yangon Technological University, Yangon, Myanmar

[Received August 9, 2019; accepted February 27, 2020]

1. Introduction

Digital Elevation Models (DEM) are often used in hydrological models to represent basin topography. The DEM is a computer (raster-based) representation of the topography of the target area using a two-dimensional (x-y value) array of grid data. Each of these grids contain elevation information as the z-value [1]. Several different input data sources can be used to generate a DEM for a target area of interest. Although ground survey provides accurate elevation of the region, due to time and budget constraints, this option is not realistic, especially if the target area is large. In recent years, LiDAR data has become increasingly popular as it is able to represent terrain with high resolution (less than a few centimeters) and accuracy. Even though LiDAR data is very accurate, it is expensive and has higher uncertainty for bathymetry data. Other methods of generating DEMs include the use of contour data (Global 30 Arc-Second Elevation – GTOPO 30, USGS), satellite radar interferometry (Shuttle Radar Topography Mission – SRTM, NASA), and optical stereo images (Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model – ASTER GDEM, NASA). For all these data sources, the methods for collecting elevation data are different and/or the principles behind the DEM generation thereof, are different, which in turn may potentially introduce uncertainty in the generated DEM.

The accuracy, applicability, and usefulness of many different DEMs have been evaluated in recent years. The accuracy of DEMs across the globe have been evaluated using elevation reference data from multiple sources such as LiDAR cloud data [2], global positioning system (GPS) [3, 4], DEMs generated using high-resolution photogrammetry [5–7], and existing topographic maps [8–12].

Most developed countries have high resolution DEMs. However, this is not the case for many developing countries, especially Asia and Africa, who must rely on globally available DEMs. Furthermore, it is often found that DEMs from one source alone may not be able to fully represent the topography of the target area with high accuracy. In such cases, it might be possible to improve the overall quality of the available DEM by combining with other sources of data. However, these other sources may not be of the same data type. The technique of combining elevation data from multiple sources of different data types such as point elevation data, raster data, to
Table 1. Characteristics of different DEM source data types.

<table>
<thead>
<tr>
<th>Type</th>
<th>Topography</th>
<th>Bathymetry</th>
<th>Accuracy</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground point survey</td>
<td>Accurate</td>
<td>Accurate</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>LiDAR</td>
<td>Accurate</td>
<td>Poor</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Satellite products</td>
<td>Reliable</td>
<td>Poor</td>
<td>Low</td>
<td>Resolution dependent</td>
</tr>
</tbody>
</table>

Fig. 1. Comparison of different DEMs of the study area: (a) ASTER, (b) SRTM, (c) MERIT, (d) HydroSHEDS, and (e) Bago DEM.

Bhagabati, S. S. et al.

2. Study Area

The Bago River basin in Myanmar was selected as the study area. The Bago River is 335 km long, spanning from the Bago Yoma mountainous region in the north to the Gulf of Mottama in the south, with an approximate catchment area of 5,000 km². The upper basin is mostly hilly terrain, whereas the lower basin is a very flat region and is frequented by annual floods [13, 14]. Because it is excessively flat, the basin also experiences a daily tidal fluctuation as far as 50 km inland from the river mouth.

3. Characteristics of Available Topography Data

There are several freely available DEM types, such as SRTM, ASTER GDEM, and HydroSHEDS (Hydrologic data and maps based on SHuttle Elevation Derivatives at multiple Scales, USGS). Each dataset has a different grid size, however, and in general the grid sizes are relatively large and hence the resolutions are low. As previously mentioned, DEMs may be generated from several different types of data sources. These can be grouped into two main types: (a) in-situ or ground observations, and (b) remotely sensed data. The later can be further sub-divided into (i) close proximity sensors (e.g., LiDAR data), and (ii) distant sensors (e.g., satellite data). The accuracy of a DEM is based on all these input source data, of which ground observations are the most accurate. Table 1 below lists the different sources and their characteristics.

Interestingly, the river channels are almost identical for the upper (hilly) terrain but are drasti-
Improving River Bathymetry and Topography Representation of a Low-Lying Flat River Basin by Integrating Multiple Sourced Datasets

The topography becomes flat in the lower basin. This is due to the different input satellite data as well as the different algorithms used for generating the DEMs. They also show differing elevation values for most of the regions.

The HydroSHEDs DEM, which is a highly modified version of the SRTM DEM, is best able to represent the river channel correctly. Although, the river channel is not completely accurate.

Overall, the lower basin is poorly represented, therefore, the accuracy of elevation data for the region is in question. Some ground elevation data (Height Point – HPt) inside the Bago River basin was available (source: Myanmar Survey Department). After ground-truthing with the available surveyed point data, it was found that almost all the DEMs are unable to accurately represent the flat lower basin, including the Bago DEM. Hence there is a need to improve the available topography data to represent the region correctly. Since the majority of conflicting data are in the larger lower basin (flat region), and given the total area of that region (around 3000 km²), it would not be feasible to conduct a field or LiDAR survey. Therefore, ALOS DSM (hereafter referred to as DSM) data was purchased for the region (representing only the lower Bago River basin, shown in Fig. 2(b)). Due to its very high resolution (5 m), it can best represent the region correctly.

Although the topography of the upper basin (hilly terrain) is correctly represented by all the DEMs mentioned above, the resolution (~90 m) of the data is relatively coarse (as compared to the DSM). The Bago DEM, on the other hand, has a 10 m resolution, and hence has been used for the same.

The Bago DEM was generated using point elevation data from UTM topographic maps, contour data, and surveyed point elevation data (source: Myanmar Survey Department). It represents the upper and mid-basin regions as well as the stream network accurately. However, it struggles to represent the flat region near the river mouth (Fig. 2(a)). Also, the bathymetry of the river is poorly represented.

In order to improve the river bathymetry, cross-sectional survey data was needed. Fortunately, during 2014–2017, multiple cross-sectional surveys were conducted by the University of Tokyo (UTokyo) in collaboration with Yangon Technical University (YTU) and other local stakeholders. These data were obtained and Fig. 2(c) shows all the survey points. A total of 128 cross-sections (CS) were measured from mid-basin to the river mouth (shown by the green colors). Additionally, another 40 CS were surveyed along the Bago-Sittaung canal (blue color).

1. The same Bago DEM is shown in Figs. 1(e) and 2(a). The dissimilarity is due to the stretching techniques used to visualize the data. In the case of Fig. 1(e), the “standard deviation” method has been used. This method provides a good overall stretch of the pixel values, and is the technique that is most often used. In the case of Fig. 2(a), the “histogram equalize” method has been used, which focuses on specific groups of pixel values, which in turn helps to differentiate the small variation of pixels in the target area. The legends for both models appear identical.

![Fig. 2. Different input data used: (a) Bago DEM, (b) DSM, and (c) Cross-section (CS) data.](image)

![Fig. 3. Comparison of elevation between DSM and HPt data.](image)

3.1. Geospatial Analysis and Comparison of the Bago DEM and DSM

As mentioned in the previous section, Bago DEM was generated using surveyed point data, digitized UTM map points, and contour data. Although, input data are correct, it was found that the flat region has many irregularities (as seen in Fig. 2(a)). It seems that these are formed due to the interpolation method used. The mid-upper basin is well represented, but the lower Bago basin, which is a relatively flat region, is very poorly represented. Also, the stream network is represented correctly, but river channel bathymetry is poor.

The DSM has been minimally processed in the PCI Geomatica application (https://www.pcigeomatics.com/) using the morphological filter technique to remove the tree canopy and built-up areas. It represents the lower basin as a very flat region as compared to the multiple irregularities noted in the Bago DEM. However, it cannot suitably represent water bodies (which is a typical limitation of satellite products), hence the river profile and bathymetry data are either missing or incorrect.

Figure 3 shows a comparison of the different eleva-
3.2. Verification of River Network with Cross-Sectional Data

As previously mentioned, a total of 168 river CSs were measured manually during 2014–2017. For a given CS, there were two parts, here referred to as part (a): the ground (riverbank) and part (b) the river (bathymetry). For part (a) total station data was used, whereas for part (b) horizontal location (x-y data) was determined using a GPS receiver, while the depth (z data) was measured using sonar data.

The accuracy of these CSs were checked against the DEM and DSM. Although the CSs matched well for the mid-section of the Bago River, for the lower region that was not the case. Fig. 4 shows the comparison of surveyed CSs and the Bago DEM and DSM. In some cases, the CS data matches either the DEM or the DSM but in the other cases, no similarity can be seen.

The typical method to modify the river bathymetry data would be to increase the width and/or the depth along the river channel by a pre-defined value (usually the difference between surveyed data and DEM/DSM value). This could be accomplished in two ways.

First, the river channel would be identified and the depth of the river along it would be measured. The corresponding elevation value on the DEM would then be identified. The difference between them would be calculated, and these values would be incorporated into the DEM along the length of the river channel. Fig. 5(a) shows one such example.

Alternately, an approximate bathymetry (width and depth) of the river at a given location could be identified, often estimated for more than one measurement. The river channel would be marked on the DEM and then the measured bathymetry is imposed onto the DEM along the length of the marked channel. This would give an output channel bathymetry which is both smooth and uniform (Fig. 5(b)).

Although very commonly used and relatively easy to perform, this output product would not be entirely accurate. This method is generally used when there is very little information about the river bathymetry (such as CS data) available. However, CSs are often conducted for many rivers, especially when bridges or other structures are constructed. In those cases, detailed bathymetry data of the affected regions of the river are available. In this research, we aim to improve the river bathymetry by using CS data. Ultimately, we hoped to generate the river channel with accurate bathymetry and update this information onto the DSM first, then transfer it onto the final DEM for the Bago River basin. If the dataset were to be used as is, we would not get a smooth representation of the riverbank because at most points along the river, the bank slopes do not match (as shown in Fig. 6). Hence, a special algorithm would be required in order to generate a smooth, accurate output.

4. Previous Research

As discussed in the previous sections, it is not possible to follow the typical approach for modifying bathymetry data, which is to burn the river network into the DEM. Historically, many methods have been developed to improve the river bathymetry of a DEM such as generating the output using sonar data [17], deriving it from multispectral imagery and aerial photography [18, 19], and other image analysis techniques [20]. As mentioned by Lane et al. [21], by using triangulation, a DEM can be easily generated, provided the distributed bathymetry data
are readily available. However, due to this extremely time-consuming procedure, such extensive surveys are rarely carried out. In contrast, CS data are more commonly available, as they are less time consuming and suitably accurate. Such data are often used as inputs for one-dimensional stream flow models.

Tate et al. [22] tried to interpolate CS data points along a river channel using triangulation. However, due to different spacing between the data points, they encountered few issues and the output product was not satisfactory.

A method based on spline interpolation was proposed by Flanagin et al. [23]. It includes a spline interpolation for a riverbed coupled with the triangulated mesh of a nearby floodplain. Although the results were generally suitable, in the case of CS profile locations which were spread farther apart, this method caused overshooting of the interpolated elevations.

In recent years, as LiDAR data became popular, some researchers tried to integrate surveyed river CSs with LiDAR data to generate smoother river bathymetry [24–26]. Schäppi et al. [25] developed a method to integrate river CS data with a LiDAR-based DEM using linear interpolation. However, for complex river systems the river network had to be divided into smaller sections. By defining “breaklines,” they tried to rectify the merging errors. Even still, issues remained for the heavily meandering river systems.

Caviedes-Voullième et al. [26] used a combination of spline and linear interpolation to combine CS data with a DEM. Spline interpolation was used to define the river trajectory, followed by linear interpolation applied in the vertical direction. Although applicable for flat, meandering rivers, a drawback of this method is that all the CSs must have the same number of observation points, even if the river’s width varies significantly (as it generally does from upstream to downstream). This is a significant limitation and is not practical for rivers such as the Bago River, as the width in the mid-basin is around 100 m, while it is more than 2000 m near the river mouth. This limitation also causes errors in extremely meandering river systems.

5. Methodology

To address these issues, a new methodology has been developed. Fig. 7 shows the complete methodology. There are three main steps involved in this method.

(a) Intermediate DEM

In the first step, the DEM and DSM were stitched together along a region of “no difference.” Both the DEM and DSM were resampled into the same grid size (10 m) and projected into UTM Zone 47 for Myanmar. Then, the difference of elevation between the two DEMs was determined by performing spatial calculations. Next, the region of no difference was identified by plotting the contour line (value = 0). The longest contour line was then selected as the boundary (Fig. 7(a)(i)). The DEM and DSM were then clipped along this boundary, with a buffer of ±250 m on either side. The output of this step was the “Intermediate DEM.” Fig. 7(a)(ii) shows the input and output of this step.

(b) River DEM

In the second step, the river CSs were interpolated to obtain the River DEM. Because the field-generated CSs were at least 1 km apart, the original interpolated DEM was of poor quality and had low accuracy, resulting in a very irregular representation of the riverbed. To achieve a smoother profile, CSs were extended on both sides in the upstream and downstream directions. It was assumed that the characteristics of the river CSs did not change for 500 m in either direction. Based on this assumption, the CSs were replicated every 100 m, and orientated perpendicularly to the river network. Figs. 7(b)(i) and (ii) show both the original and the manually created CSs, respectively, for a small portion of the Bago River.

As mentioned earlier, the riverbank slope did not match with the surveyed CS data. In addition, the datasets are of different types (the riverbank is a raster and the CS data are vectors). To overcome this issue, the following steps were performed.

First, a boundary for each of the CSs was created by joining the end points of adjacent CSs. Then, a 200 m buffer was created (Fig. 7(b)(iii)). This region was masked out of the Intermediate DEM and converted to point data (Fig. 7(b)(iv)). Since both the river CSs and the masked region were now point datasets, it was possible to interpolate them together. By using the “topo-to-raster” interpolation tool of ArcGIS, the “River DEM” was generated. Fig. 7(b)(v) shows the input and output of this step.

(c) Final DEM

In the final step, the Intermediate DEM and River DEM were then mosaicked using the “blend” tool to get a smoother topography along the boundary. The output of this process is shown in Fig. 7(c). From here onwards, the “Final DEM” will be referred to as the “Enhanced DEM (EnDEM).”
Fig. 7. Methodology to combine different DSM and DEM with cross-sections: (a) Intermediate DEM, (b) River DEM, and (c) Final DEM.

Fig. 8. Comparison of DSM with EnDEM.

Figure 8 shows a comparison of the surface topography of the DSM and the EnDEM near the river mouth. Originally, in the DSM, there was no data for the river channel in the lower Bago basin near the river mouth. In the EnDEM, not only is the river well defined but the overall quality of the DEM along the riverbank has improved. Along the region of mosaicking there is a smooth transition between the river profile and the adjoining areas.

6. Discussion

6.1. Characteristics of Cross-Section Data

As mentioned in Section 5, the river CSs were replicated. This sampling interpolation was done to achieve a denser series of CSs.

Originally there were 168 CSs. The distance between two CSs and the number of data points per CS varied upstream to downstream from 500 m to 2500 m, and 30 to 75, respectively. River width varied from 100 m (upstream) to 2200 m (downstream).

For a non-meandering river, a distance of 1000 m between CS profile locations seems acceptable. However, in the case of a meandering river, especially one such as the Bago River with tight meanders every 500 m to 2000 m, shorter intervals are necessary to successfully generate the river network. Fig. 9 shows the comparison for a small section of the river generated using 500 m, 200 m, and 100 m intervals. Although the river network pattern is generally suitable with the 500 m and 200 m intervals, the bathymetry data is completely lost. However, as the interval is reduced down to 100 m, the bathymetry data is retained.

Regarding regions of heavy meandering, by using only the originally measured CSs, it was not possible to generate an accurate representation of the streambed bathymetry. Fig. 10 shows one such example. In this section, only seven CSs were originally measured (shown in green color). Using these alone, the ability to generate the river network, let alone the bathymetry, would be severely compromised. However, by using a 100 m CS interval, the output was able to better represent the meandering region.
Improving River Bathymetry and Topography Representation of a Low-Lying Flat River Basin by Integrating Multiple Sourced Datasets

These results show that a CS interval similar to the minimum river width is needed for better representation of river network and bathymetry.

6.2. Verification of the EnDEM

By following the method explained in Section 5, the EnDEM was generated. The upper part was derived from the Bago DEM and the lower part from the DSM (ALOS product). Since the raster data had different resolutions, the DSM was up-scaled to 10 m in order to generate the EnDEM of 10 m resolution. Fig. 11 shows a comparison of the SRTM, ASTER GDEM, HydroSHEDS, and EnDEM for three regions: the upper, mid-, and lower basins. The approximate location of the river channel is shown using red dashed lines. The graphs show the vertical profile of each DEM for the region marked by a blue oval shape. As expected, the different DEMs all represent the upper hilly terrain in a similar manner. In the mid-basin however, the ASTER GDEM is inaccurate, while the other three DEMs show similar results, with the EnDEM showing the most accurate elevation data. In the lower basin (near the river mouth), the ASTER GDEM is again inaccurate. Even the SRTM is very poor, as the depth of river channel is shown to be a mere 2–3 m. The HydroSHEDS, which is a highly modified version of the SRTM, is able to suitably represent the riverbed and bank elevation but it smooths out and mildly slopes towards the river, thereby creating a valley-like region. The actual topography is represented only by the EnDEM. In fact, unlike the HydroSHEDS, the area near the river mouth is extremely flat as shown in the EnDEM. From these graphs, we can conclude that the EnDEM is able to best represent the upper, mid-, and lower basin regions correctly as compared to the other three DEMs.

In order to conduct a quantitative analysis of the improvements shown by the EnDEM, one method could have been to divide the available 168 CSs into a 70/30 ratio of two groups as follows: (1) a larger group of 118 CSs used to generate the EnDEM, and (2) a smaller group of the remaining 50 CS to be used as a test group to validate the quality and accuracy of the EnDEM.

However, as explained in the previous section, these 168 CS are not enough. Moreover, the characteristics of the CS along the river channel (such as width and number of data points) are not constant. Hence in this case, in order to maintain a high quality of DEM output (achieved by using all the 168 CSs), quantitative validation was deemed inappropriate and was not conducted.

7. Conclusion

Globally available DEMs often cannot accurately represent terrain, especially in the case of flat deltaic river systems. These DEMs may be generated from differing sources and based on the method of use, the individual characteristics, quality, and accuracy of the output products may vary. Certain products may be able to represent a particular style of terrain correctly (e.g., mountainous terrain), while at the same time be unable to suitably represent another style (e.g., flat terrain, water bodies). Additionally, nearly all remotely sensed terrain data are unable to obtain river bathymetry. In cases where river bathymetry data are required, river CS survey is essential. Manual ground survey does provide accurate data; however, due to time, cost, and accessibility limitations, it is not practical. Since different terrain products have variable advantages and disadvantages, the best approach would be to try to integrate them by only considering the advantages of each product. In that way the output prod-
uct generated will be of much higher accuracy.

To achieve the same, this paper presented a new method to integrate differing terrain data types (i.e., raster: DEM and DSM) with surveyed CS data (i.e., vector, point elevation). The output, known as EnDEM, contains the best parts of all the input data. It was found that a CS profile interval similar to the minimum river width was required for better representation of the overall river network, bathymetry, and meandering characteristics. Unlike previous studies, this method can be simultaneously applied to the entire length of a river.

CS surveys are conducted for many rivers. However, due to the technical gap or resources, those survey data are not fully utilized. Sometimes, only accurate bathymetry data is essential (e.g., studies on riverbed morphology). This method, although primarily developed to combine multiple data sources and be used with complex deltaic river basins, may also be used simply to integrate the surveyed data onto the available DEM.

As a recommendation for future work, more CS surveys should be conducted at a much finer spatial frequency, which will help in measuring the quantitative improvements of the EnDEM.

Acknowledgements

This research was supported by the Japan Science and Technology Agency (JST, JPMSIA1407)/Japan International Cooperation Agency (JICA), and the Science and Technology Research Partnership for Sustainable Development Program (SATREPS).

References:

With reference to the contents of the provided document:

**Name:** Seemanta Sharma Bhagabati  
**Affiliation:** Project Researcher, Department of Civil Engineering, The University of Tokyo

**Address:** 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

**Brief Career:**  
- 2010-2012 Master Degree program, Asian Institute of Technology, Thailand  
- 2012-2014 Research Associate, Asian Institute of Technology, Thailand  
- 2014-2018 Doctoral Degree program, The University of Tokyo  
- 2017-2018 Technical Assistant, The University of Tokyo

**Selected Publications:**  

**Academic Societies & Scientific Organizations:**  
- Sustainability Science, Editor  
- Turkish Journal of Geographical Sciences, Editorial Advisory Board  
- Science Council of Japan (SCJ)

---

**Name:** Akiyuki Kawasaki  
**Affiliation:** Project Professor, Department of Civil Engineering, The University of Tokyo

**Address:** 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

**Brief Career:**  
- 2008-2009, 2019- Visiting Scholar, Harvard University  
- 2010-2015 Visiting Associate Professor, Asian Institute of Technology (AIT)  
- 2010-2017 Project Associate Professor, The University of Tokyo  
- 2018- Project Professor, The University of Tokyo

**Selected Publications:**  

**Academic Societies & Scientific Organizations:**  
- Sustainability Science, Editor  
- Turkish Journal of Geographical Sciences, Editorial Advisory Board  
- Science Council of Japan (SCJ)

---

**Name:** Win Win Zin  
**Affiliation:** Professor, Department of Civil Engineering, Yangon Technological University

**Address:** Gyogone, Insein Road, Yangon 11011, Myanmar

**Brief Career:**  
- 1996- Assistant Lecturer, Yangon Technological University  
- 2001- Lecturer, Yangon Technological University  
- 2009- Associate Professor, Yangon Technological University  
- 2017- Professor, Yangon Technological University

**Selected Publications:**  

**Academic Societies & Scientific Organizations:**  
- Myanmar National Committee on Large Dam (MNCOLD)  
- Myanmar Engineering Society (MES)