Digital Elevation Model (DEM) Extraction from Google Earth: a Study in Sungai Muar Watershed

Noradila Rusli [a] and M. Rafee Majid [b]

[a] Faculty of Architecture, Planning & Surveying, Universiti Teknologi MARA, 40450 Shah Alam, Selangor Darul Ehsan, Malaysia,- radhiutm@gmail.com

[b] Department of Urban & Regional Planning, Universiti Teknologi Malaysia, 81310 Skudai, Malaysia- rafee@utm.my

KEY WORDS: Digital Elevation Model (DEM), Google Earth, ArcGIS 9.3.1, ArcHydro9, Terrain Zonum Solutions, Sungai Muar Watershed, Watershed Delineation

ABSTRACT:

Topographic data and DEMs are commonly perceived as providing loaded information for hydrological modelling. However, this information is severely limited, and most users rely on published topographic maps or DEMs produced by government agencies. This paper proposes an alternative method of generating DEM by using online source of elevation data which is easier and free. The result shows that, the difference in percent of the watershed boundary is 0.19% by using Google Earth's elevation data as it compared with the same watershed area which is delineated using 20 meter interval contour data obtained from JUPEM (Department of Survey and Mapping Malaysia). Firstly, the location of the study area identified in Google Earth, was divided into thirty-six areas and marked using provided mark functions in Google Earth. The coordinates were exported to an online application tool named Terrain Zonum Solutions for extracting 5000 elevation points for each thirty-six areas. These points were uploaded and processed in ArcGIS software (version 9.3.1) to generate the DEM. As to compare the reliability of the elevation values that extracted from Google Earth, the DEM was utilized and integrated with river line data to delineate Sungai Muar's watershed boundary using ArcHydro version 1.4 tool. Based on the results, this study proposes an alternative method in obtaining a DEM data for a wide area which is traditionally time consuming and costly.

1 Introduction

In hydrologic studies, it is vital to obtain topographic information for many purposes such as to delineate watershed boundary and to identify the flow direction. Generally, there are several of methods to extract elevation data such as topography map, on-site measurement or data application from government agencies. However, the existing methods are traditionally tedious, time consuming and costly especially for wider area. In Malaysia, it is common for lacking of accessible data for free and accurate. Fortunately, by the extensive development of information technology there are many free source tools that can be manipulated to gather the topographic information and later processed into DEM (Digital Elevation Model). The DEMs commonly perceived as providing a means for encoding the spatial distribution of the hydrological process (Abbot et al., 1986; Moore et al., 1990). In real practice, the topographic information normally gathered from those topographic maps which can be bought from government agency; JUPEM (Department of Survey and Mapping Malaysia). Then, the map scanned and digitized its contour lines to extract the elevation values. Later, these values are used and processed to generate the DEM. Due to extremely tedious work, the digitizing process becomes unpractical and produces less accurate results if it is implemented in wider area. Human error factor may contribute to an action to simplify the contour line while digitizing process. Another option is by downloading DEM SRTM (Shuttle Radar Topography Mission) data which is available for free, provided by USGS (U.S Geological Survey) with the vertical accuracy of 30 m in US and 90 m worldwide. However, for a small area, the accuracy provided by the SRTM may not suitable for some hydrological modelling process because they require more accurate DEM. Therefore, this paper presents an alternative method to extract elevation data using online free web tool and to generate it into DEM by the aid of GIS application tools.

2 Data and Methods

1.1 DEM Extraction from Google Earth

Sungai Muar's watershed area has been selected in this study to be extracted the elevation data. The basic principle for determining the watershed boundary is by identifying the main river (Sungai Muar) and its tributaries. As the limited source of data in this study, the river line for Sungai Muar was digitized using ArcGIS 9.3.1 from existing river map published by the Department of Irrigation and Drainage of Malaysia. These lines then converted to .kml (Keyhole Markup Language) file and viewed in Google Earth to identify the approximate area of the Sungai Muar's watershed. The conversion process can be conducted using Shape2Earth extension in Map Window GIS tool. Sungai Muar flows in Muar, Segamat, Gemas and Gemencheh in Negeri Sembilan to Seri Menanti and Batu Kikir, Negeri Sembilan. The extent of Sungai Muar watershed area are from 1.926702°N, 102.056472°E in lower left corner and 2.972698° N, 103.222606°E in upper right corner. Then, these areas were divided into thirty-six areas using a marking tool provided in Google Earth. Each of these marks contained latitude and longitude values which are in WGS84 (World Geodetic System) projection. Next, these coordinates were exported to extract elevation data in an online tool named Terrain Zonum Solution that can be accessed for free in http://www.zonums.com/gmaps/terrain.php. This tool is easy to understand and to use as it only required some basic information such as the location (latitude, longitude) and numbers of points sampling. The maximum points would be 5000 points for both type of sampling either randomly or uniform grid. The tool automatically marks the elevations points which depend on the desired requirement of the detail for the data. Reduce the extent area but increase the number of sampling points to obtained detail elevation data. The output of the extraction process are the coordinates (x,y) values (WGS84 projection) including the elevation (z) in meters. These values are then imported into a text file for further process in ArcGIS 9.3.1. Figure 1 shows the overall process of extracting the elevation data in Google Earth.

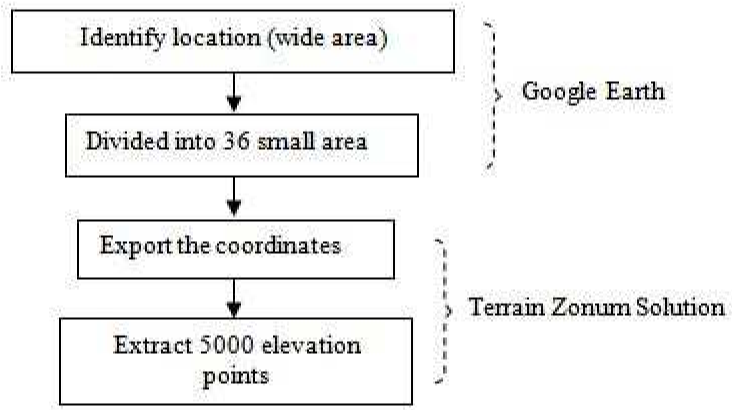

Figure 1: The flow of extracting processs

1.2 Watershed Delineation

This section explains the process to delineate Sungai Muar's watershed boundary by using the previous elevation data extracted from Google Earth (later, GE elevation data will be used in this paper).

The elevation points are then exported to ArcGIS 9.3.1 and converted into shapefiles (.shp). Before conducting the next process, the elevation value should be checked. Google Earth gives -32768 m value for any points obtained from water area such as sea. Therefore, these values should be eliminated as they become outliers for the data and disturb the accuracy for the next process. As the data cleaned from the outliers, the TIN (Triangulated Irregular Network) was then generated using 3D Analyst tool in ArcGIS 9.3.1. According to Peucker et al., (1978) and Mark (1975), TINs usually the samples of surface specific-points, such as peaks, ridges and break in slope and form an irregular network of points store as a set of x,y, and z coordinates together with pointers to their neighbours in the net. From TIN, it then converted to raster as it becomes the DEM data which is the most widely used terrain model for drainage delineation. A DEM is a grid of square cell; and each cell value represents the elevation of the land surface. By determining how water flows from cells to cells, the set of cells whose drainage flows through the cells at the outlet point location can be identified and thus the watershed area determined (Maidment, 2002). The ArcHydro version 1.4 was utilized to delineate Sungai Muar's watershed boundary. It is an extension tool for water resource management in ArcGIS and this version is compatible with ArcGIS 9 or 9.3.1. The depression in the DEM data firstly 'filled' by increasing the cells value in depression to depression's spill point value. The next process is to compute the flow direction for each cell in the 'filled' DEM. After that, the next step is followed by the computation of the flow accumulation value for each cell. This is simply the count for each cell of how many upstream cells would contribute based on their flow directions. After these pre-processing phases, the datasets can be further processed to delineate watersheds. This process is simple by integrating flow direction data with digitized Sungai Muar river line that have been converted into raster. The overall process of manipulating the extraction of elevation points until the watershed delineated. The processed displayed in Figure 2.

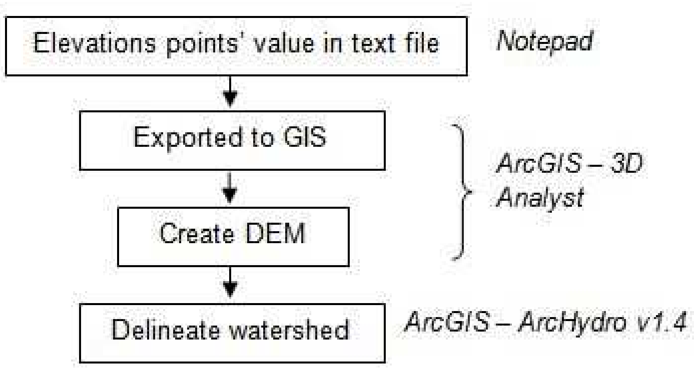

Figure 2: The brief process of watershed delineation

2 Results and Analysis

In this section, the height comparison between GE's elevation and 20 metres interval contour will be discussed using a normal distribution statistical comparison. Later, the delineated boundaries from both data were compared and analyzed.

2.1 GE Compare with 20m Interval Contour

To analyze the reliability of GE's elevation data, approximately, 6521 hectares of flat area and hilly area were extracted their height values. These GE's elevation points then overlapped with 20 metres of contour interval data and analyzed using normal distribution statistical method. Based on the results, the difference of mean and standard deviation value for GE's elevation data is close to 20 metres of contour interval data (see Table 1).

Area	Source	Mean	Stand. Dev
Flat	Google Earth	12.87	6.41
	20 m interval	14.18	7.36
Hilly	Google Earth	62.54	13.15
	20 m interval	61.94	12.30

Table 1 Height comparisons for three sources elevation data

Later, the mean and standard deviation values, used to calculate the normal distribution of both data and displays in Figure 3. According to the bell shape formed, the GE's elevation data fell in the range of 20 metres contour interval data for its normal distribution in hilly and flat area.

2.2 Further GE Comparison

There was a hydrologic study conducted by NAHRIM (National Hydraulic Research Institute of Malaysia) in Sungai Muar watershed area (San & Selamat, 2010) which involved 20 meter interval contour of data for watershed delineation. Hence, their boundary area was digitized and compared with the results of watershed delineation using GE's elevation data. Based on the results (see Table 2), the watershed boundary delineated from GE's elevation data produced an (i) area difference as it compared to NAHRIM's delineated boundary. The finding is really interesting due to the source of NAHRIM's elevation data by using 20 meter interval of contour data. Additionally, a (ii) ratio measurement also conducted to compare the boundary shape of GE with 20 m interval of contour data. This calculation conducted by divided boundary's perimeter over area. In addition, the GE boundary also compared the (iii) percentage of overlay area with NAHRIM'S boundary by intersecting both boundaries in ArcGIS. Lastly, these boundaries are also compared with their location of mean centre which is 2.1 km difference in distance between two mean points (see Figure 4).

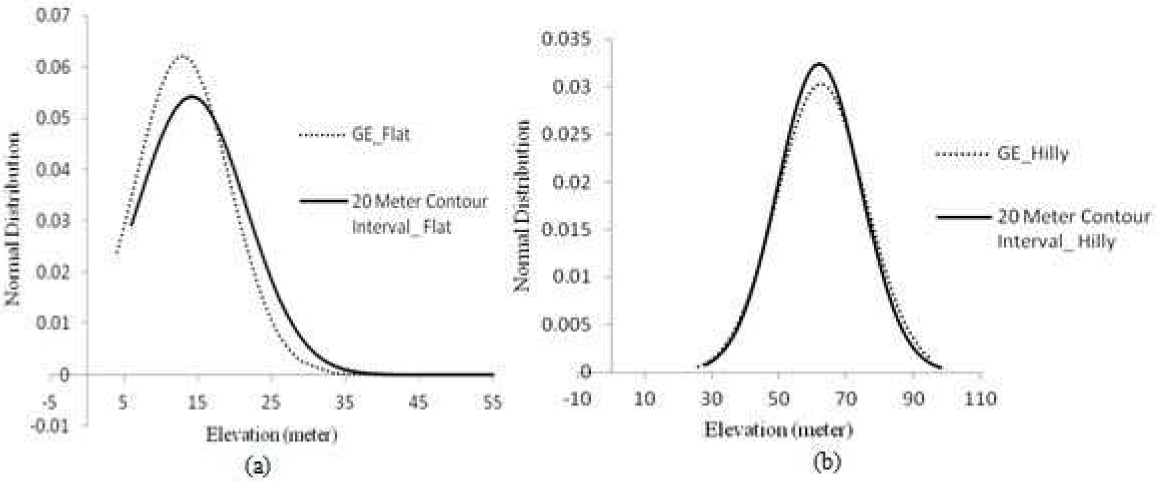

Figure 3(a): Normal distribution graph in flat area, (b) Normal distribution graph in hilly area

Data Source	Area (km^2)	Perimeter (m)	(i) % Difference	(ii) Ratio (Perimeter/Area)	(iii) % Overlay	(iv) Mean centre distance
20 m interval	6149.00	590.15	0.19%	0.09	76 %	2.1km
Google Earth	6137.46	614.20		0.10		

Table 2 Comparison results of both

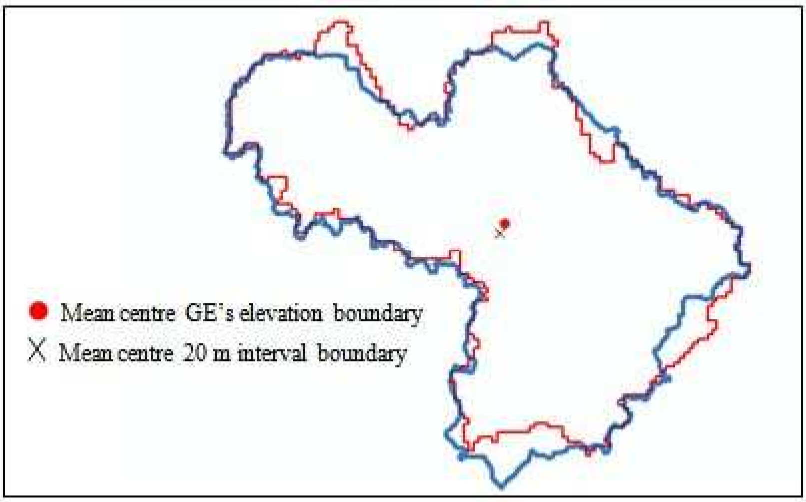

Figure 4 Mean centre location for both boundaries

There are wide online forum discussions on the accuracy of elevation data in Google Earth. However, there is no official statement from Google on their elevation source of data including its accuracy. The early reliable information from Google Earth's personnel, that the elevation provided by GE is originated from ten (10) different sources of elevation data. However, they do not mention about the sources. Among the GE's user, they speculated that one of the elevation source in GE is from SRTM data which its resolution vertically 30 metres in US and 90 metres worldwide. However, starting from June 2009, NASA and Japan both released ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) satellite data on the internet for free. The data covers 99% of the entire Earth's landmass terrain compared to SRTM which is only 80% coverage. In addition, ASTER's data providing better resolution compared to SRTM. It has a base resolution of 30 metres, and can be extended to 7m-10m. Therefore, there is a probability that GE is utilizing this free data as one of their elevation source. In fact, the GE's users had tested the accuracy of elevation data by comparing the site measurement. Many of them agree that GE's elevation is more accurate in flat area with error range less than 5m and some findings less than 1m. However, in mountainous terrain, the data is less accurate due to GE's elevation where the data appears to be an average elevation data for an area centred over the desired point.

3 Conclusions

The results from the watershed's delineation process have indicated that the elevation data extracted from GE is usable. Moreover, this finding has given some hopes to the researchers that are having difficulties in collecting elevation data for their studies. In fact, researches do not need to spend money to conduct extremely tedious processing of data collection if the same accuracy can be easily obtained from Google Earth.

References

Abbott, M.B., Bathurst, J.C., Clinge, J.A., O'Connell, P.E. & Rasmussen, J. 1986. An introduction to the European Hydrological System - Systeme Hydrologique Eruopean, 'SHE', 1: History and philosophy of a physically-based, distributed modelling system; 2: Structure of a physically-based, distributed modelling system. Journal of Hydrology 87, 45-90,61-79

Maidment, D.R., 2002. *Arc Hydro: GIS for Water Resources*. ESRI Press, Redlands CA. 220 pages.

Mark. D. M., 1975. Computer analysis of topography: B comparison of terrain storage methods. Geografiska Annaler, 57A, 179- 188.

Moore, I. D., Grayson, R. B., & Wilson, J. P. 1990. Runoff modelling in complex three-dimensional terrain. *Proc. International Conference on Water Resources in Mountainous Regions, XXlI Congress of the International Association of Hydrogeologists*, Lausanne, Switzerland.

Peucker, T. K., Fowler, R. J., Little, J. J., & Mark, D. M. 1978. The triangulated irregular network. *Proc. Auto Carto III, Congress on Surveying and Mapping*, Fallschurch Virginia, 516-540

San, L.Y & Selamat, Z. 2010. Flood Hazard Mapping Using Two-Dimensional River Modelling for Sungai Muar, Johor in *Colloqium on R & D in Flood Plain Management*, pp. 102-136.

Acknowledgements

The authors would like to convey a gratitude to Universiti Teknologi Malaysia for funding this research. This paper is a part from research entitled "Community Level Management of the Effects of Land Use/Land Cover Changes in the Magnitude and Frequency of Flooding" (Vot 00G11).