Urban Drainage Network Generation with Geographic Information Systems Using Remotely Sensed Data

N. Attygalla and N. T. S. Wijesekera

Abstract: In Sri Lanka most of the existing urban drainage systems are not functioning properly. Establishing a proper drainage system in an urban area is extremely important since it could avoid floods and inundation which causes damage to property and inconvenience to city dwellers; improve health conditions of the urban community by avoiding stagnant water; give an aesthetically pleasing environment, etc. Therefore, in Sri Lanka, establishing a proper drainage system and maintaining the same has become a major challenge in urban areas. Urban drainage network identification is a primary requirement for developing a proper drainage system. Urban drainage network identification is a tedious task, which is usually done through engineering surveys. The computer aided Geographical Information Systems (GIS) can be utilized to generate drainage network from the terrain data. The main objective of this work is the study of stream network generation in GIS using data extracted from aerial photographs, carryout accuracy comparison with surveyed drainage network and then to identify the parameters that affect the accuracy of generating streamlines. In this study, two urban watersheds from Colombo and one watershed from Moratuwa were studied to identify the stream network generation and associated terrain indicators using data extracted from aerial photographs.

The data extraction for the study was done using three methods. Existing drainage network and watershed boundary were digitized from hard copy maps of Sri Lanka Land Reclamation and Development Corporation (SLLRDC). Contours, spot heights, buildings and road data in digital form were extracted from the digital archives of the National Survey Department (NSD). Some of the existing streamlines and culvert locations were surveyed specifically using GPS. Then the generation of the Triangulated Irregular Network (TIN) was carried out using the extracted contours and spot heights. Using the TIN, Digital Elevation Models (DEM) for spatial resolutions of 2m, 5m, 10m, 20m & 50m were generated. For each DEM, flow direction grids and flow accumulation grids were generated as components of stream network generation. Once the streamlines are generated from the flow accumulation grid, it is necessary to give a threshold value to separate the stream network. After comparing several vector and grid based methods, it was identified that Grid Based Comparative Squares (GBCS) method could be effectively used for comparing the accuracy of generated and extracted streamlines. In the GBCS method the streamlines generated and observed are matched for fitting only within identified square areas. In the squares, comparison is done using zones of distance from observed streamlines created using buffering capability of GIS, and giving an error code for the generated lines which fall into buffers that represent the deviation. Therefore, once the squares are selected spatially, then the degree of fitting of the computed and the observed streamline are compared with a system of buffers drawn to each observed streamline within a square. The squares and a buffer zoning of the extracted streams enable the identification of deviations of the generated streams from those of engineering survey sheets. Deviations were compared using RMSE (Root Mean Square Error) as the numerical indicator.

The extracted stream network of catchments indicated a strong human influence in deviating the stream network and hence buildings were combined to the DEM to identify whether the results would show a difference. The same accuracy assessment method for five spatial resolutions was used for comparison of stream network. Also Flatness, Stream Order and Surface Slope based on streamline network of each watershed were computed for comparison.

The study also identified that the accuracy levels were lower in the areas of lower stream order. This shows that in the upper reaches the...
streamlines would vary significantly probably due to less contributing cells and also as a result of human interventions.

The stream network identification to a reasonable accuracy using remotely sensed data of 0.4m accurate contours and additional spot heights is not a straightforward operation in the case of selected watersheds. The significant flat terrain poses problems in identifying the flow directions to find a unique stream network.

1.0 Introduction

In Sri Lanka most of the existing urban drainage systems are not functioning properly. Establishing a proper drainage system in an urban area is extremely important since it could avoid floods and inundation which causes damage to property and inconvenience to city dwellers; improve health conditions of the urban community by avoiding stagnant water; give an aesthetically pleasing environment, etc. Increase of population in urban areas leads to the reduction of available land and this results in encroaching and closing of drainage canal reservations and land filling of water retention & detention areas. Lack of a proper canal system to ensure good drainage and having inadequate low land to act as retention and detention areas, generates a series of problems related to one another. Due to this reason, flooding of urban areas in Sri Lanka has become a constant problem.

Therefore, in Sri Lanka, establishing a proper drainage system and maintaining the same has become a major challenge in urban areas. Urban drainage network identification is a primary requirement for developing a proper drainage system. Urban drainage network identification is a tedious task, which is usually done through engineering surveys. Location and elevation data from these surveys are then plotted to identify physical features and terrain features. Plotting to different scales, maintaining and checking accuracy, interpolation of values etc., is a time consuming difficult task. In the present system, the stream network is also surveyed and superimposed on to the contour map. The computer aided Geographical Information Systems (GIS) can be utilized to generate drainage network from the terrain data. In most of the occasions, a drainage network designer or a municipal engineer needs to obtain stream network and watershed information, depending on the engineering applications that need to be implemented. Present day GIS enable the identification of terrain details such as slope and aspects, the watershed boundary, drainage networks and thereafter changes to drainages due to change of terrain etc.

The urban area terrain and stream network generation is a task, which is well known as difficult because of the surface flatness, which is a common feature in most urban areas. In Sri Lanka, the application of GIS and aerial photography to identify the stream network along with associated accuracy is yet to be carried out. As such, in this study, two urban watersheds from Colombo and one watershed from Moratuwa were studied (Figure 1) to identify the stream network generation and associated terrain indicators using data extracted from aerial photographs.

![Figure 1: General Outline of Study Area](image)
Table 1: Details of Input Data

| Type of Data                              | Source                                    | Stated Accuracy (H - Horizontal) | (V - Vertical) |
|-------------------------------------------|-------------------------------------------|----------------------------------|----------------|
| Existing Streamline                       | Paper maps of SLLRDC                      | H 0.6 m                          |                |
| Watershed boundary                        | Paper maps of SLLRDC                      | H 0.6 m                          |                |
| Culvert locations                         | Paper maps of SLLRDC                      | H 0.6 m                          |                |
| Existing streamline                       | Mosaic Image of Aerial photograph         | H 0.6 m                          |                |
| (Katubedda watershed)                    | Digital archives of NSD                   | H 0.4 m                          | V 0.4 m        |
| Contours                                  | Digital archives of NSD                   | H 0.4 m                          | V 0.3 m        |
| Spot heights                              | Digital archives of NSD                   | H 0.6 m                          |                |
| Building                                  | Digital archives of NSD                   | H 0.6 m                          |                |
| Roads                                     | GPS field Survey (Trimble)                | H 0.3 m                          |                |
| Existing minor streamlines not indicated in maps | GPS field Survey (Hand held)            | H 5 m                            |                |
| Culvert locations not indicated in maps   |                                           |                                  |                |

2.0 Objectives

The main objective of this work is to study stream network generation in GIS using data extracted from aerial photographs, carry out accuracy comparison with surveyed drainage network and then to identify the parameters that affect the accuracy of generating streamlines. Other issues addressed are the data resolutions, comparison techniques and the identification of the use of GIS outputs with remotely sensed data for engineering applications. The present study was carried out for the three watersheds stated above and a comparison was done with the surveyed stream networks.

3.0 Study Area & Data

3.1 Description of Study Area

Torrington watershed is located in central Colombo containing a mixture of residential areas, roads, playgrounds, marshy areas etc. The watershed area is approximately 450 ha. The watershed drains to the sea at Wellawatte. Attidiya watershed which is located in the southern part of Colombo which is approximately 190 ha contains a mixture of residential areas and other users. The watershed drains to the sea at Wellawatte. Katubedda watershed which is located in the southern part of Greater Colombo containing a mixture of playground, important buildings like higher educational institutes, marshy areas etc. The watershed drains to the sea at Panadura through Bolgoda Lake.

3.2 Description of Data

The data extraction for the study was done using three methods. Existing drainage network and watershed boundary were digitized from hard copy maps of Sri Lanka Land Reclamation and Development Corporation (SLLRDC). Contours, spot heights, buildings and road data in digital form were extracted from the digital archives of the National Survey Department (NSD). These data were generated by NSD using Remotely sensed Data with reasonable ground verification. Some of the existing streamlines and culvert locations were surveyed specifically for this study using GPS. In the Katubedda watershed some of the streamlines could not be extracted from the mosaic image of the aerial photographs. Therefore, a GPS field survey was carried out to obtain this information. A trimble GPS instrument was used and its stated horizontal accuracy is ±0.3m. Culvert locations of the Torrington watershed were identified using low the accurate Garmin GPS and its stated accuracy is ±5m. The details of input data are given in Table 1.

4.0 Methodology

The study methodology for the streamline generation in an urban watershed using remotely sensed data is as shown in Figure 2. Initially the Torrington watershed data set was used for application and the other two
4.1 Database Preparation

The digital database required for this study was prepared by extracting data of contours, spot heights, buildings, roads, culvert locations and existing streamlines directly from the database of NSD. NSD spatial data are in vector format having elevation as an attribute. The existing streamlines of Torrington & Attidiya were extracted from the maps of SLLRDC which had been prepared by joining of 1:1000 maps and reducing to 1:4000 scale. This map indicated joining and reducing errors, which affect the overall accuracy of stream network extraction. This map scanned and cropped to several 1:1000 files avoiding pasting edges, was geo-referenced using building and road features as reference points. Then, streamlines were extracted by considering each 1:1000 files at a time. In the Katubedda watershed, existing streamlines were extracted using aerial photographs since SLLRDC maps were not available. The aerial photographs, which were scanned and geo-referenced, were combined to form a single image through mosaic from which streamlines were digitized.

4.2 Terrain Generation

The generation of the Triangulated Irregular Network (TIN) was carried out using the extracted contours and spot heights. Using the TIN, Digital Elevation Models (DEM) for spatial resolutions of 2m, 5m, 10m, 20m & 50m were generated. For each DEM, flow direction grids and flow accumulation grids (Kohler 1999) were generated as components of stream network generation. Direction of flow in a cell is called “flow direction” and the flow accumulated at a particular cell due to contributing cells from upstream is called “flow accumulation”.

Once the streamlines are generated from the flow accumulation grid, it is necessary to give a threshold value to separate the stream network. The primary objective of streamline generation is to develop a set of streams, which follow a similar pattern of streamlines in the engineering survey sheets. This is done by assuming that the remotely sensed data could be used for the development of a data set which can be utilized for comparing the presently used engineering survey sheets. Therefore, the threshold value has to be selected such that the streamline patterns of maps match with the generated streamlines. In this study it was assumed that the patterns match, if the total numbers of branches are the same in both generated and extracted streamlines since the direction of generated and extracted streamlines are almost the same due to the influence of morphology in
the streamline development process. The concept of "outlet stream order" which indicates the level of branching off streams at the outlet of a watershed, gives an indicator of how well the streamlines are branched off within the watershed.

Therefore, the threshold value for flow accumulation grid was identified under the condition that stream order of generated and extracted streamlines at the outlet become equal. In this case, for all three watersheds, stream order at the outlet for both generated and existing streamlines was determined by changing the threshold value. Table 2 indicates how this variation occurs in the Torrington watershed. Stream order value for both generated and actual streamlines are done manually by counting the branches up to the outlet based on the "Shreve Order" method.

The GIS generated threshold values for different resolutions were varied by trial and error until the extracted streams matched in branching off. Then, the respective threshold value that fulfills this requirement was identified for each watershed. This threshold value was divided by the total number of 2 m² size pixel of the respective watershed, to get the per cell threshold value for the watershed. The per cell threshold value gives a more generic representation of a watershed, related to streamline generation. The per cell threshold value for each watershed is given in Table 3. The average per cell threshold value comes as 0.016 for the three watersheds. Therefore, for geographically similar areas, meaning for areas having similar topography and land use, an average per cell threshold value of 0.016 would reasonably match the generated streamline network with the stream network from engineering surveys.

| Threshold Value | Generated Stream Order | Actual Stream Order |
|-----------------|------------------------|---------------------|
| 6000            | 30                     | 14                  |
| 8000            | 27                     | 14                  |
| 10000           | 32                     | 14                  |
| 12000           | 25                     | 14                  |
| 14000           | 18                     | 14                  |
| 15000           | 14                     | 14                  |
| 16000           | 12                     | 14                  |

**Table 3: Per Cell Threshold Value for study Watersheds**

|                      | Torrington | Attidiya | Katubedda |
|----------------------|------------|----------|-----------|
| Threshold value/cell | 0.014      | 0.015    | 0.019     |

4.3 Accuracy Assessment

After comparing several vector and grid based methods (Attygalla 2005), it was identified that Grid Based Comparative Squares (GBCS) method could be effectively used for comparing the accuracy of generated and extracted streamlines. In the GBCS method the streamlines generated and observed are matched for fitting only within identified square areas. In the squares, comparison is done using zones of distance from observed streamlines created using buffering capability of GIS, and giving an error code for the generated lines which fall into buffers that represent the deviation. Therefore, once the squares are selected spatially, then the degree of fitting of the computed and the observed streamlines are compared with a system of buffers drawn to each observed streamlines within a square (Attygalla 2005). The squares and a buffer zoning of the extracted streams enable the identification of deviations of the generated streams from those of engineering survey sheets. The layout of the squares and an example of the buffers are shown in Figure 3.
The sizes of buffers and the squares were changed systematically to identify the deviations corresponding to the outputs of each watershed. As a result, the square sizes assigned to Torrington, Attidiya and Katubedda watersheds are 200m, 400m and 180m respectively. The degrees of deviations were measured using the grid cell sizes and the number of cells away from the observed line. As for example, in order to identify the degree of deviation in the results, the grid cells in each buffer were identified using numerical labels. The average buffer distances of 1, 3.5, 7.5, 15, 27.5 and 70 values were assigned respectively for the Torrington watershed which has an average deviation of 35 m. Deviations were compared using RMSE (Root Mean Square Error) as the numerical indicator (William 2001). The RMSE value indicates the average deviation of generated stream lines from those in the hard copy maps.

RMSE is defined as,

$$\text{RMSE} (m) = \sqrt{\frac{\sum Ni (Xi)^2}{\sum Ni}}$$

Where Ni - Number of cells for each buffer
Xi - Average buffer deviation

### 4.4 Spatial Resolution

The study for all three watersheds was done for five spatial resolutions i.e. 2m, 5m, 10m, 20m & 50m. The base dataset accuracy of ±0.6m indicated that a spatial resolution of 1m would not improve the quality of the outputs.

### 4.5 Incorporation of Buildings

The extracted stream network of catchments indicated a strong human influence in deviating

---

**Figure 4: Torrington Watershed with Building and Streams**

the stream network and hence buildings were combined to the DEM to identify whether the results would show a difference (Figure 4). The same accuracy assessment method for five spatial resolutions was used for comparison of stream network.

### 4.6 Catchments Parameters

Flatness, Stream Order and Surface Slope of each watershed was computed for comparison. The respective values are given in Table 4.

### 5.0 Results

### 5.1 Accuracy

The generated stream networks using spot heights and contours along with the extracted network for three watersheds is shown in Figure 5. The mean RMSE computed using RMSE values for each square for each of the three watersheds is shown in Table 5.

---

**Table 4: Mean RMSE and Other Parameters for Watersheds**

| Watershed   | Mean RMSE for 2m Resolution | Catchments Parameters |
|-------------|-----------------------------|-----------------------|
|             | Contours & Spot Heights     | Contours, Spot Heights and Building Incorporation | Flatness | Stream Order | Surface Slope |
| Torrington  | 31                          | 28                    | 21%       | 14           | 0.169        |
| Attidiya    | 78                          | 74                    | 48%       | 5            | 1.36         |
| Katubedda   | 40                          | 29                    | 22%       | 7            | 0.893        |

---

47 ENGINEER
Table 5: Average RMSE variation with spatial resolution.

| Resolution | RMSE (m) | Torrington | Attidiya | Katubedda |
|------------|----------|------------|----------|------------|
| 2          | 31.2     | 74.2       | 40.9     |            |
| 5          | 34.2     | 72.5       | 41.1     |            |
| 10         | 32.1     | 69.0       | 41.2     |            |
| 20         | 35.1     | 57.0       | 42.8     |            |
| 50         | 35.8     | 69.2       | 46.7     |            |

5.2 Combining the Building layer

The DEM modified with buildings indicated, that in some areas there were visible improvements but in general it had no significant clearly represented effect (Attygalla 2005). The average RMSE value for Torrington was 28.9 after incorporating buildings.

5.3 Catchments Characteristics

The selected watershed parameters computed for the three study areas were compared with error values are shown in Table 4.

6.0 Discussion

6.1 Stream Network Accuracy

The average RMSE value variation with spatial resolution did not indicate a general trend of deterioration of quality with spatial resolution getting coarser. The RMSE on average for a 2m spatial resolution indicated a minimum value of 31m for Torrington watershed whereas for Katubedda and Attidiya watersheds the same was 41 and 74 respectively. This shows that the generated stream network deviated significantly from that which had been extracted from engineering surveys. The study identified that the main reason for this is the high percentage of areas in each watershed having zero or near zero slopes thereby creating difficulties in the generation of unique streamlines. When the flatness increased the matching also rapidly deteriorated (Table 4).

The study also identified that the accuracy levels were lower in the areas of lower stream order. This shows that in the upper reaches the streamlines would vary significantly, probably due to less contributing cells and also as a result of human interventions. A comparison of the RMSE error with percentage area under each slope class is shown in Table 6.

6.2 Incorporation of Buildings

The incorporation of buildings did not much improve the stream network because the changes to the DEM by the buildings and the stream generation methodology enabled only a slight deviation, which was quite different to the actual situation. In ground reality the urban dwellers modify the streams to suit their land parcels or housing complex layouts. Unless such land cluster information is incorporated and the DEM is modified accordingly, it may be difficult to improve the stream network accuracy. This needs the incorporation of field surveys of existing main drainages, minor roads and other critical land parcel details along with remotely sensed data. This further modification to the DEM would generate more accurate streamline network which can be utilized for urban planning. However it has been noted that incorporation of buildings is important to obtain a stream network that would appear close to that picked by ground surveying.

7.0 Conclusions

1. The stream network identification to a reasonable accuracy using remotely sensed data of 0.4m accurate contours and additional spot heights is not a straightforward operation in the case of selected watersheds. The significant flat terrain poses problems in identifying the flow directions to find a unique stream network.

Table 6: Percentage of Watershed Area in each Different Slope Class

| Watershed     | slope class | Average RMSE |
|---------------|-------------|--------------|
|               | 0-0.01      | 0.01-0.5     | 0.05-1      | 1-5         | 5-40         | |
| Torrington    | 20.3        | 1.59         | 52.8        | 23.5        | 1.63         | 31.27        |
| Attidiya      | 48          | 0.1          | 5.1         | 36.9        | 9.5          | 74.25        |
| Katubedda     | 21.3        | 1.05         | 27          | 42.5        | 7.9          | 40.95        |
2. For datasets having contours and spot heights of approximately 0.4 m accuracy, it was identified that the best set of streamlines could be generated with 2m spatial resolution and the best average accuracy was 48.7m.

3. In general, the study area indicated that generated streamline accuracy is improved when the surface slope and the stream order increases, and when the flatness of the watershed is decreased.

4. The key parameters affecting accuracy of a general urban drainage network could be clearly identified as 1). Flatness of surface and 2). Building and other artificial terrain modifications.

5. Evaluation done with the incorporation of buildings revealed that the addition of better topographic information by critically looking at the terrain could in general improve the accuracy of generated streamlines.

6. For watersheds similar to those in the case studies, the flow accumulation per cell threshold value for stream network generation is in the order of 0.016.

Acknowledgement

The Authors wish to thank the National Water Supply & Drainage Board and its staff members for their support in report preparation. A special thanks is given to the University of Moratuwa and its Civil Engineering department for providing opportunities for the research.

References

1. Johansen, K. 2000. NWSDB Project Document: MEMO Technical Support, NSD & NWSDB Sri Lanka, Colombo, Blom & Halcrow Ltd.

2. Kohler, M. 1999. User Guide of the Arcflow Surface Water Analysis., USA.

3. Attygalla, N. 2005. Urban Drainage Network Generation using Geographical Information System, MEng Thesis submitted to University of Moratuwa Sri Lanka September (Unpublished)

4. William, W.H & Douglas, C.M & David, M.G. & Connie, M.B. 2001. Probability and statistic in engineering. John Willey & Sons (Asia) Pte Ltd, Singapore