# Fusion of DEM and Satellite Imagery for Surface Area Calculation: Development of an Algorithm and a Tool for GRASS-GIS

Santosh Tripathi
RS&GIS FoS, Asian Institute of Technology
POBox: 4, Klong Luang, Pathumthani, 12120, Thailand
Santosh.Tripathi@ait.ac.th

Junichi Susaki
RS&GIS FoS, Asian Institute of Technology
POBox: 4, Klong Luang, Pathumthani, 12120, Thailand
susaki@ait.ac.th

Xiayong Chen
RS&GIS FoS, Asian Institute of Technology
POBox: 4, Klong Luang, Pathumthani, 12120, Thailand
xychen@ait.ac.th

**Abstract:** Multiplying the number of pixels of satellite imagery with the resolution gives the palnimetric area covered by the pixels. At many times, such as in calculating the area covered by a particular class in classified satellite imagery, however, the slope surface area is a more useful parameter. To calculate slope surface area we need to take into consideration the undulations in the ground which can be done by using the DEM. DEMs come in different forms. However, with the availability of the SRTM-DEM, grid DEMs are most easily found DEMs. This paper compares different approaches for calculating the surface area based upon grid DEM and suggests a procedure for more accurate area calculation. This procedure involves increasing the resolution of the DEM by interpolating it before calculating the area. This approach has been found to give more accurate results. Moreover, a procedure to calculate the surface area covered by each pixel of satellite imagery by fusing the DEM and the satellite imagery is given. Accuracy assessment for different approaches was done by comparing the calculated areas with the ‘actual surface area’. Source code for the area calculation based on the interpolation procedure is available which can be used from the GRASS-GIS. A novel feature of this study is that it uses synthetic terrains and synthetic DEMs in the analysis.
**Keywords:** Slope Surface Area Calculation.

## 1. Introduction

In many situations slope surface area is a better indicator of the surface area than the planimetric area. For example, of the two plots with same palnimetric area populated with equally dense forest, plot in mountainous area will have more trees than the plot in a flat terrain. So, under such condition, to find the actual amount of resource available one has to resort to slope surface area.

In this study we focus on grid DEMs and term ‘area’ will, unless stated otherwise, mean the slope surface area. Using grid DEMs for area calculation is not new. Kundu and Pradhan [4] suggests using SHR (Surface to Horizontal Area Ratio) for calculating the surface area. It uses a 3 by 3 roving window operator. Height of the nine pixels covered by the window is used to calculate the area of the eight triangular faces. Ultimately one fourth of this area is taken to be the area of the central pixel in the window. In the method proposed by Jennses [2] the length of the lines connecting the 3-dimensional center point of a pixel to the 3-dimensional center point of its eight neighbors is calculated. Length is also calculated of the lines which connect the 3-dimensional center point of contiguous neighbors. All these lines when connected produce eight triangles. Area of the pixel under consideration is calculated by halving the length of all the lines and calculating the area of resulting eight triangles. Although the procedure is different, both the papers cited above essentially employ the same formula and hence give the same result.

For the accuracy assessment of the method proposed, Kundu and Pradhan [4] compares the area calculated by the SHR method to the area calculated by using the simple slope method. Similarly, in case of the Jenness [2], the area calculated is compared with the area calculated using the TIN DEM. This approach for accuracy assessment, though can give a relative comparison, cannot be used to understand the absolute accuracy of the procedure for area calculation. For assessing the accuracy, calculated area has to be compared to the actual surface area of the landscape of which the area has been calculated using the DEM.

In this study, we try to eliminate this problem by using synthetic terrains at high resolution. Moreover, in addition to the above cited procedure we investigate a four triangle method and two triangle method. Similarly, the effect of increasing the resolution of the DEM by interpolation prior to calculating the area is also tested. Based upon the results it is concluded that accuracy of surface area calculation is improved when interpolation is done. Moreover, a procedure for calculating the surface area covered by satellite imageries using grid DEM is given.

## 2. Methodology

To check accuracy-wise the performance of different approaches for surface area calculation, one naturally needs actual surface area of the test site. Actual surface area acts as the benchmark against which all the calculated areas can be compared. However, at one hand getting such a data requires resources and at the other hand the required data can be reasonably generated using the computers. Moreover, using computers can allow the experimenter to experiment with the situations that only occasionally arise in real world. This motivated us to use synthetic terrain and synthetic DEMs in the present study. Such numerical experiments are not new. For example, Ines and Honda used such approach in [1].

### 1) Synthetic Terrain

Fractal techniques allow generation of feature rich terrains [5]. In this study, we used the diamond-square algorithm for terrain generation. This algorithm works by iteration and adds details with each iteration. By specifying the values for first few iterations, it is possible to generate desired landform. The algorithm was implemented in MATLAB as a MATLAB function. Following figure shows code snippet from the function which implements the diamond step of the algorithm.

```
for(x_j=1:sections)
   for(y_j=1:sections)
      if(op((x_j*sec_size+1)-sec_size/2, (y_j*sec_size+1)-sec_size/2)==0)
         op((x_j*sec_size+1)-sec_size/2, (y_j*sec_size+1)-sec_size/2)...
          =...
               (...
               op(x_j*sec_size+1, y_j*sec_size+1)...
               +op(x_j*sec_size+1-sec_size, y_j*sec_size+1)...
               +op(x_j*sec_size+1, y_j*sec_size-sec_size+1)...
               +op(x_j*sec_size-sec_size+1, y_j*sec_size-sec_size+1)...
               )/4+roughness;
       end
    end
```

**Fig. 1 Code snippet showing the implementation of diamond step of the diamond-square algorithm**

Terrains so generated are in the form of grids. We used the cell resolution of 1 m in our study. Source code for this and all the other algorithms implemented during the course of this study is available from the first author upon request.

### 2) Synthetic DEM

In case of the SRTM-DEM, height noted for each cell is the average height of a number of points on the cell [3]. Though the actual model for generation of the synthetic DEM from the synthetic terrain that would mimic SRTM-DEM would be overly complex, in this study we find the representative height by averaging the height location of all the cells in the terrain falling within the given cell of the DEM. This algorithm too was implemented in MATLAB.

### 3) 'Actual Area'

To the area of any continuous two dimensional functions integration has to be done with respect to both dimensions over the whole area. Numerical integration proceeds by assuming that over a very small area the function can be represented by planar segments. Accuracy of this assumption increases as the size of 'very small area' is decreased. We have too used the same assumption in calculating the 'actual area'. The resolution of each cell of terrain is taken to be 1m and the corresponding DEMs have been created with resolution of 100m. Thus, each DEM cell represents 10000 cells of the original terrain and with respect to the DEM cells each cell of the terrain can be regarded as a surface with zero curvature. Even after we assume a zero curvature, we still have to identify the sides of the surface that a cell of the terrain represents. This is fundamentally the same problem as finding the area from the DEM. However, the size of the cells being very small with respect to the DEM cells, approximations can be accepted. Approach$_{21}$ (see next subsection) has been used to calculate the 'actual area'. It should be noted that though this area, in comparison to the area calculated using the DEM, is a better approximation of the area of the surface; it is not the actual area.

### 4) Area from the DEM

As written in the introduction section, the procedures proposed by Kundu and Pradhan [4] and Jennesse[2] are essentially the same. For this reason, the Jennsse[2] approach has been implemented as one of the approaches being investigated. Hereafter this will be denoted by Approach$_{11}$.

To find the area covered by a pixel of the DEM ( i, j) three by three window of pixels centered at i, j can be taken. The height of points 1, 2, 3 and 4 can be calculated as the average of all the pixels surrounding each point. For example, for point 1in fig 1, the height can be taken as,

$$H_1 = \frac{\begin{pmatrix} H_{i-1,\,j-1} + H_{i-1,\,j} \\ + H_{i,\,j-1} + H_{i,\,j} \end{pmatrix}}{4} \tag{1}$$

Here, $H_1$ is the height of the point 1 and the other subscripted Hs represent the height registered for corresponding pixel in the DEM.

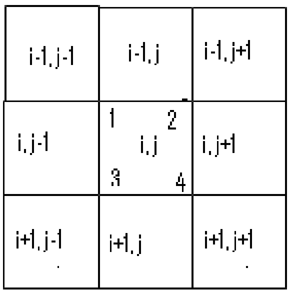


**Fig. 2 A 3 by 3 window for calculating the heights of points 1, 2, 3 and 4**

Once the heights have been calculated, the area can be calculated by finding the distance between the three dimensional representations of points 1-2, 2-4, 4-3, 3-1 and 2-3 and calculating the area of two triangles that are formed. This approach will, hereafter be known as Approach$_{21}$.

Similarly, after calculating the heights of the points 1, 2, 3 and 4 as above, one can calculate the sides of the four triangles that are formed by the four points and the center of the cell (i, j). The area represented by the cell (i, j) is then the sum of the area of four triangles. Hereafter, this approach is termed Approach$_{31}$.

All these approaches, assume that the value represented in the cell is the elevation of the center of the area represented by the cell. All of them interpolate the changes of the elevation linearly. However, actual elevation change may occur in any form. To capture these variations, two-dimensional interpolation of the elevations can be done to find the elevation values and then the above listed approaches can be used for each interpolated value. Approach$_{i2}$ represents this approach for each of the above approaches respectively.

### 5) Information Fusion between Satellite Imagery and DEM

Given the satellite imagery and the DEM, area covered by each of the pixels of the satellite imagery can be found by calculating the area using the DEM cells that correspond to the particular pixel in question.

We used altogether 9 terrains of size 2000m by 2000m. Thus, each terrain had 4000000 cells and produced 400 cells of the DEM. The terrains have been chosen with different landscape: from flat surface to highly undulated. Some of the terrains generated and used are shown in the figures below. For all the figures, the vertical axis shows the height in meters, the horizontal axes are x and y and each unit is 25 meters. The images have been rotated to give better view of the terrains. These plots were generated using the MATLAB surf function after generating the numerical values using the diamond-square fractal algorithm..

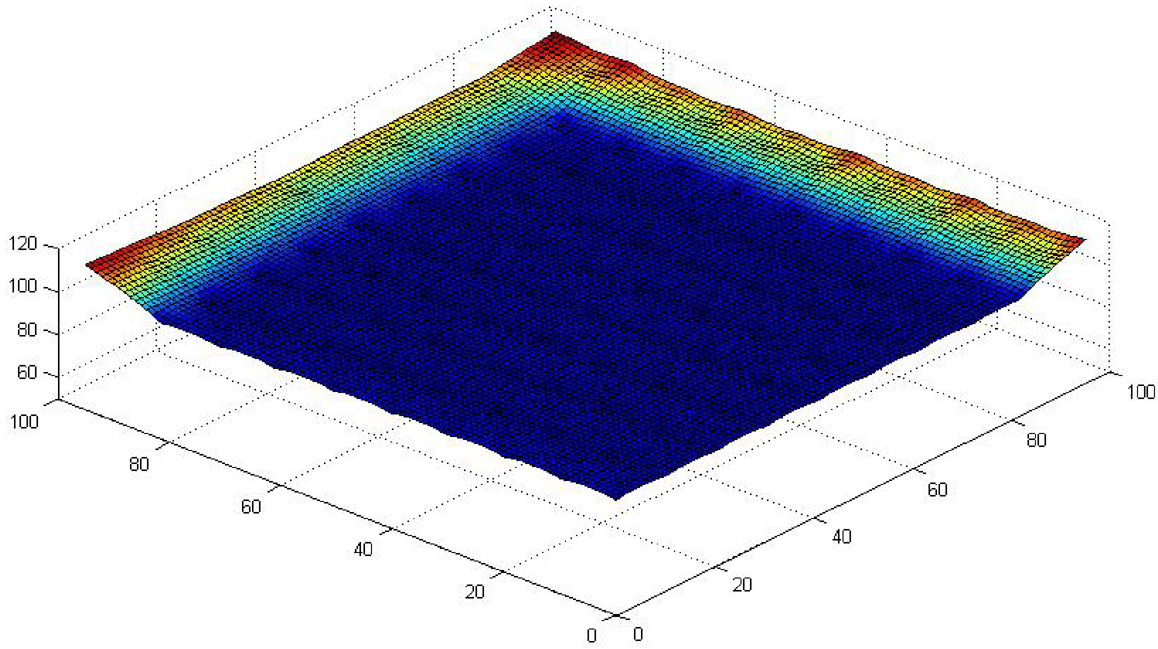


**Fig. 3 Terrain No. 1**

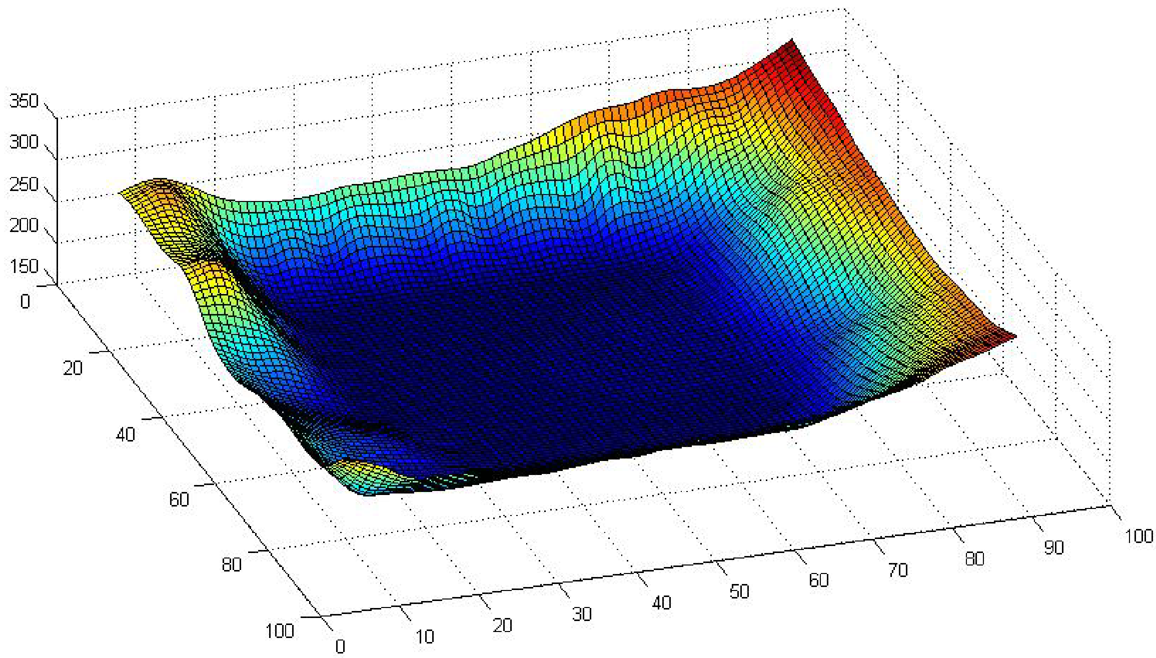


**Fig. 4 Terrain No. 2**

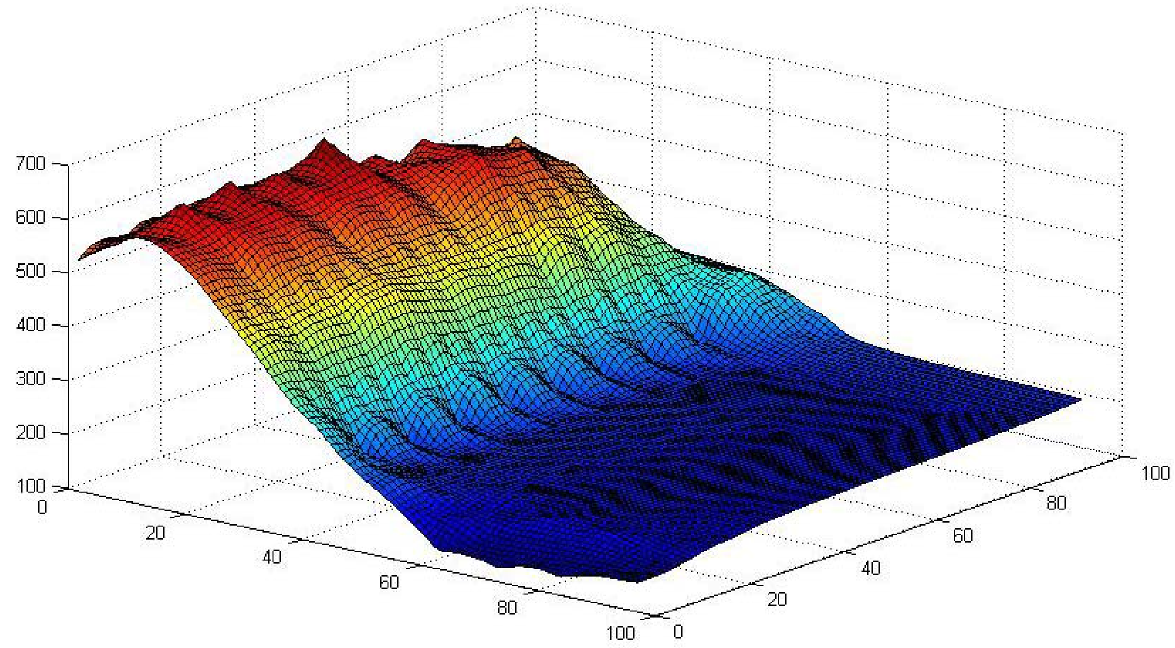


**Fig. 5 Terrain 3**

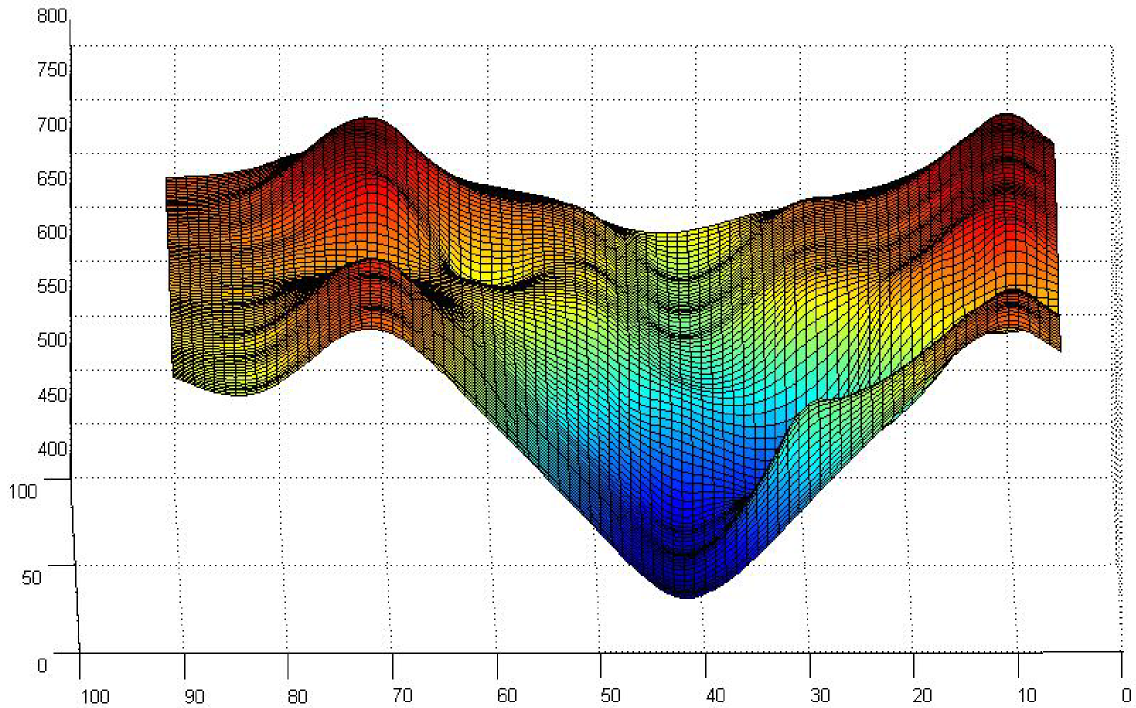


**Fig. 6 Terrain No. 7**

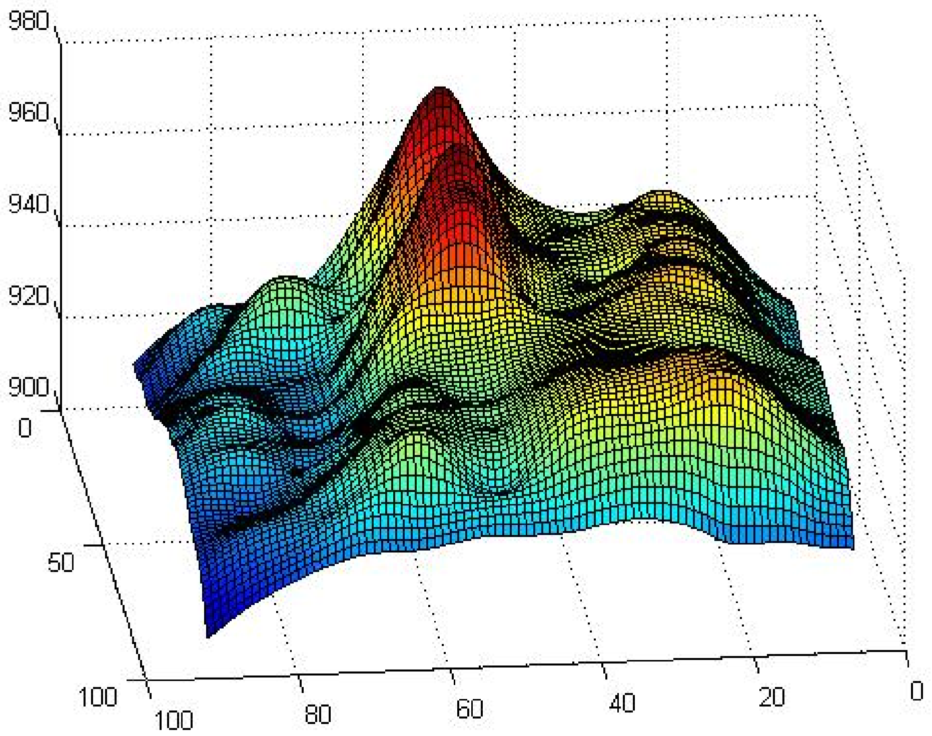


**Fig. 7 Terrain No. 9**

## 4. Results and Discussion

**Table 1 Results for methods Approach$_{11}$ and Approach$_{12}$**

| Terrain No. | Area ($10^6$ $m^2$) | | | | |
|---|---|---|---|---|---|
| | Actual | Appraoch$_{11}$ | $\frac{\text{Mag.(Actual - Approach}_{11})}{\text{Actual}}$ *100% | Approach$_{12}$ | $\frac{\text{Mag.(Actual - Approach}_{12})}{\text{Actual}}$ *100% |
| 1 | 4.0011094 | 4.0002511 | 0.02145 | 4.0002517 | 0.02144 |
| 2 | 4.1178202 | 4.0810097 | 0.89393 | 4.0814457 | 0.88334 |
| 3 | 4.0469591 | 4.0288906 | 0.44647 | 4.0288286 | 0.44800 |
| 4 | 4.3208904 | 4.2017661 | 2.75694 | 4.2019745 | 2.75212 |
| 5 | 4.4147348 | 4.2727965 | 3.21510 | 4.2733029 | 3.20363 |
| 6 | 7.4814986 | 6.6119335 | 11.62287 | 6.6177374 | 11.54530 |
| 7 | 4.2095931 | 4.1482530 | 1.45715 | 4.1500898 | 1.41352 |
| 8 | 4.3932822 | 4.2694438 | 2.81881 | 4.2716723 | 2.76809 |
| 9 | 5.2152814 | 5.2093933 | 0.11290 | 5.2094309 | 0.11218 |

Table 1 shows the results for the method suggested in [2] and [4] (Approach$_{11}$) and the same method after increasing the resolution of the DEM by two times by interpolation (Approach$_{12}$). In this study we did not investigate the effect of different interpolation techniques and in all the cases cubic interpolation was used. Comparing the two methods we see that in all the cases the accuracy has increased. Within the same method, however, the accuracy is not even. For flat surface as represented by the terrain in fig. 1, the accuracy is maximum; accuracy is minimum for terrain no. 6 which happens to be a surface with slope more than 60°. Fig. 7

represents a mountainous area, however, over the terrain the variation of the height is only about 100 meter and variation is slow it has resulted in comparatively low error in the calculated area. Steepness of the variation of the elevation describes the increase in the error in going from terrain nos. 2 to 3 to 7 (Figs. 4, 5, 6).

**Table 2 Results for methods Approach$_{21}$ and Approach$_{22}$**

| Terrain No. | Area ($10^6$ $m^2$) | | | | |
|---|---|---|---|---|---|
| | Actual | Appraoch$_{31}$ | $\frac{\text{Mag.(Actual - Approach}_{21})}{\text{Actual}}$ *100% | Appraoch$_{32}$ | $\frac{\text{Mag.(Actual - Approach}_{22})}{\text{Actual}}$ *100% |
| 1 | 4.0011094 | 4.0002119 | 0.02243 | 4.0002376 | 0.02179 |
| 2 | 4.1178202 | 4.0728074 | 1.09312 | 4.0786087 | 0.95224 |
| 3 | 4.0469591 | 4.0248419 | 0.54652 | 4.0274161 | 0.48291 |
| 4 | 4.3208904 | 4.1835811 | 3.17780 | 4.1956256 | 2.89905 |
| 5 | 4.4147348 | 4.2647015 | 3.39847 | 4.2702899 | 3.27188 |
| 6 | 7.4814986 | 6.5390123 | 12.59756 | 6.5915505 | 11.89532 |
| 7 | 4.2095931 | 4.1246656 | 2.01748 | 4.1425531 | 1.59255 |
| 8 | 4.3932822 | 4.2390668 | 3.51026 | 4.2619203 | 2.99006 |
| 9 | 5.2152814 | 5.2082202 | 0.13539 | 5.2090439 | 0.11960 |

Referring to table 2, the improvement in accuracy of area calculation by increasing the DEM resolution by interpolation is apparent.

**Table 3 Results for methods Approach$_{31}$ and Approach$_{32}$**

| Terrain No. | Area ($10^6$ $m^2$) | | | | |
|---|---|---|---|---|---|
| | Actual | Appraoch$_{31}$ | $\frac{\text{Mag.(Actual - Approach}_{31})}{\text{Actual}}$ *100% | Appraoch$_{32}$ | $\frac{\text{Mag.(Actual - Approach}_{32})}{\text{Actual}}$ *100% |
| 1 | 4.0011094 | 4.0002487 | 0.02151 | 4.0002511 | 0.02145 |
| 2 | 4.1178202 | 4.0813260 | 0.88625 | 4.0815920 | 0.87979 |
| 3 | 4.0469591 | 4.0288196 | 0.44823 | 4.0288448 | 0.44760 |
| 4 | 4.3208904 | 4.2009979 | 2.77472 | 4.2018595 | 2.75478 |
| 5 | 4.4147348 | 4.2721033 | 3.23081 | 4.2731322 | 3.20750 |
| 6 | 7.4814986 | 6.6068207 | 11.69121 | 6.6161740 | 11.56619 |
| 7 | 4.2095931 | 4.1468209 | 1.49117 | 4.1497952 | 1.42052 |
| 8 | 4.3932822 | 4.2645380 | 2.93048 | 4.2704234 | 2.79651 |
| 9 | 5.2152814 | 5.2091817 | 0.11696 | 5.2093755 | 0.11324 |

Refering to table 3, the improvement in accuracy by improving the DEM resolution by interpolation is apparent.

Comparing all the methods we see that dividing the pixel into only two parts for area calculation (Appraoch$_{21}$ and Approach$_{22}$) perform the worst. Between eight and four triangle methods, four triangle method performs slightly better in terrains with less undulations and eight triangle method performs better in steep terrains. It is logical as in steep terrain, averaging distorts the true elevation but four triangle method uses averaging of four pixel elevation to calculate the elevation at the corner of the center pixel but in the case of eight triangle method it is done with only two pixels at a time. Based upon this, it is apparent using eight triangle method with interpolation in the highly undulating areas, and using four triangle method with interpolation in relatively flat areas will give better results in terms of accurate surface area. Fig. 8 shows the steps for calculating the surface area and fusing the area with satellite imagery. The source code for interpolation based area calculation which can be used from GRASS-GIS will be released very soon.

Time series images of NOAA/AVHRR of year 2001 covering almost all parts of Nepal downloaded from [7] were classified into three classes. 90m SRTM-DEM covering the same area was downloaded from [6]. Using the method outlined here the area was calculated for each class. The forest class occupied 41813 pixels, however, the surface are was calculated to be 46820 km$^2$. Similarly, for water/snow class came 724 pixels but the area was found to be 7400 km$^2$.

Third class that included all other things possessed 102821 pixels; however, it had an area of 113320 km$^2$. The increase in area is natural as Nepal is a mountainous country.

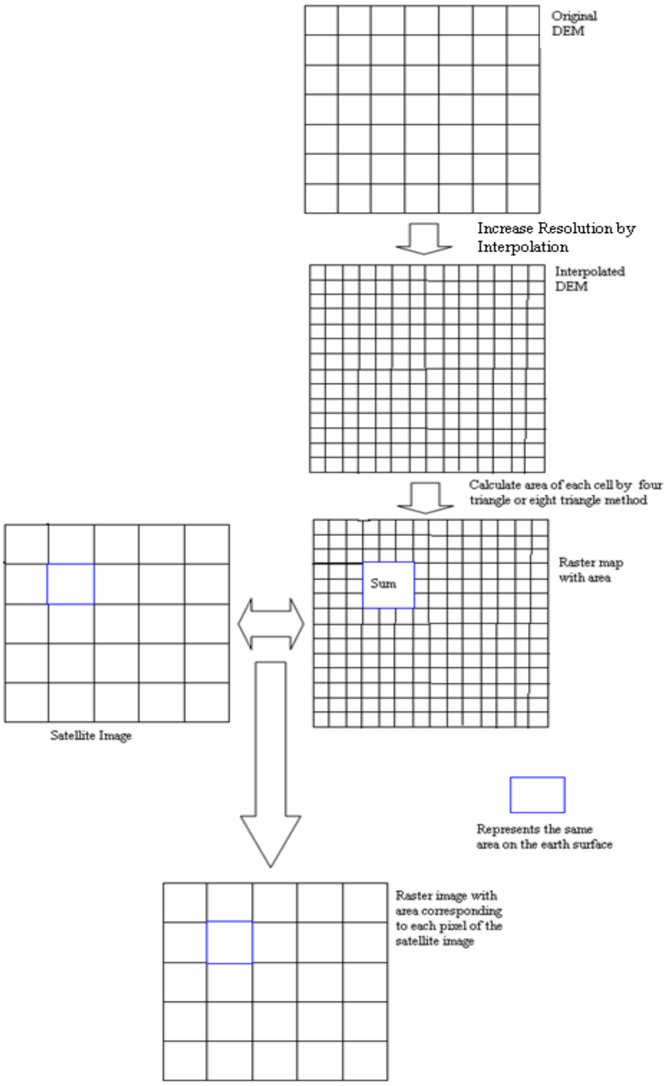


**Fig. 8 DEM-Satellite Image Fusion: Calculating the surface area of satellite imageries**

Though the area calculated using the proposed method is more near to the actual area than the previously suggested techniques, a quantitative analysis of the improvement in accuracy has not yet been performed. Moreover, the accuracy is bound to vary with the landform. A transform based approach might be better for finding the exact area or finding the limit of accuracy that can be achieved

## 6. Conclusion

Existing methods for calculating the surface area from grid DEM were assessed for their accuracy using synthetic terrains and synthetic DEMs. Accuracy assessment was done by comparing the calculated area with the 'actual area'. It was found that increasing the resolution of the DEM by interpolation before calculating the surface area produces more accurate results. Moreover, in the plain area four triangle method was found to be more accurate and in the steep terrains eight triangle method produced better results. Based upon this, a procedure for calculating the surface are for satellite imageries using DEM was outlined and was used to calculate the areas of different classes in a classification image produced from NOAA/AVHRR.

## References


[1] Ines, A. V. M., and K. Honda, 2004. On quantifying agricultural and water management practices from low spatial resolution RS data using genetic algorithms: A numerical study for mixed-pixel environment, *Advances in Water Resources*, 28(8): 856-870.

[2] Jenness, J. S., 2004. Calculating landscape surface area from digital elevation models, *Wildlife Society Bulletin 2004*, 32(3): 829-839.

[3] **URL:** Geoimage Pty Ltd. Australian SRTM data Available at: www.geoimage.com.au/geoweb/dfs/products/SRTM_Australia.pdf

[4] **URL:** Kundu, S. K., and B. Pradhan, 2003. Surface area processing in GIS. Available at: www.gisdevelopment.net/technology/gis/ma03191pf.htm

[5] **URL:** Olsen, J, 2004. Realtime procedural terrain generation. Available at: www.oddlabs.com/download/terrain_generation.pdf

[6] **URL:** International Center for Tropical Agriculture (CIAT). CGIAR-CSI SRTM 90 m database. Available at: http://srtm.csi.cgiar.org

[7] **URL:** WebPaNDA-AVHRR Data Service Center. Composite imagery for Asia. Available at: http://webpanda.iis.u-tokyo.ac.jp/WebPaNDA/index_j.php