An Evaluation of the Accuracy of DEM-derived Altitude and Slope Values

Sara Byne, M. Sc..

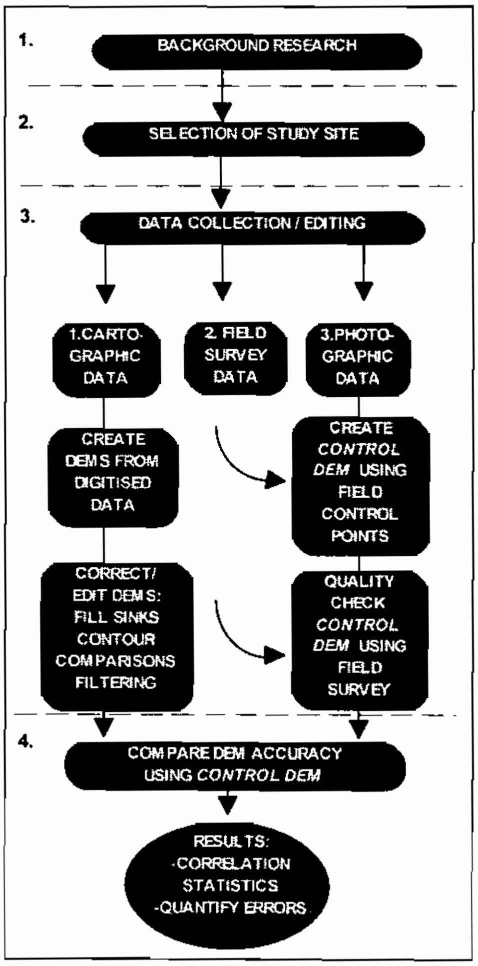
40066

ABSTRACT

This project was set up to investigate the level of accuracy which can be expected for slope and altitude values derived from low-cost Digital Elevation Models (DEMs). Eight gridded DEMs were generated from digitized contour maps at a range of scales:- 1:10,000 1:25,000 1:100,000 1:200,000 - and using a range of contour intervals 25m, 50m and 100m. A Control DEM was then produced using large scale aerial photographs (1:28,000) which were registered for auto-extraction of z-values using Helava software and accuracy tested using 91 differentially measured GPS ground control points. The DEM showed an vertical RMSE of 4.26m which is well within the accuracy standards for a 'level one' DEM as stipulated by the USGS. The altitude and slope readings derived from each of the eight test models were then compared to the values derived from the Control to assess the relationship between the cost of production of a DEM and the accuracy of the results. The relationship between cellsize and slope correlation was also examined. Several recommendations have been made regarding optimal production methods for a DEM based on application needs.

1. PROJECT STRUCTURE

As illustrated in Figure 1.1 the project was divided into four phases. Each phase deals with a particular stage of the research, development and assessment of nine Digital Elevation Models (DEMs). Eight of these models were produced using a range of topographic maps as source data. The ninth DEM was produced using highly accurate large scale stereo photography. It is considered to represent 'true' altitude and was developed to act as a control for the study. This model was itself quality controlled using a large number of GPS ground control points which were collected in the field.



2. BACKGROUND AND OBJECTIVES

A Digital Elevation Model (DEM) is a three dimensional computerized model of the earth's surface which is used to store topographic attributes in digital form. These models have been developed within the field of Geographical Information Systems (GIS) and are a valuable source of data for agricultural research. The information which they provide can be used as input to a wide range of projects such as soil erosion modelling (Vertessy et al, 1990), crop suitability (Bradley, 1994), drainage basin monitoring and flood control (Rosenthal, 1995), hydrological run-off modelling (MacMillan et al,1994; Chieng & Luo, 1993), land classification (Jones, 1993; Dikau, 1989), viewshed analysis (Lee, 1991; Smart et al 1991) and pollution dispersion modelling (Woodrow, 1993). They are also being used in the field of Remote Sensing (RS) to aid geometric and radiometric correction of satellite images (Conese et al, 1993).

It is widely acknowledged that the products of a GIS will reflect and, in some cases, augment, any errors which are present in the source data, (Goodchild & Gopal, 1989). Likewise, the accuracy of a DEM and its products is dependent on the quality of the altitude data from which the surface has been generated.

A large number of data sources can be used to generate and store DEMs. These range from analytical photogrammetry and softcopy photogrammetry of stereo satellite imaging and aerial photographs (Day & Muller, 1988; Toutin & Beaudoin, 1995; Welch & Papacharalmpos, 1992) to the digitization of topographic maps (Eklundh & Martensson, 1995) and ground surveys (McLaren & Kennie, 1989;81). The range of data sources for deriving DEM surfaces is diversifying rapidly. Likewise, the need to investigate and assess the reliability of each one is becoming more important. Accuracy assessments of DEMs and their products have begun to appear in scientific journals (Brown & Barra, 1994; Bolstad & Stowe, 1994; Adkins & Merry, 1994; Fisher, 1990; Skidmore, 1989) and the United States Geological Survey published a set of accuracy specification for their own DEM products in 1990 (USGS, 1990). However, there are still no international standards to which a DEM should conform to lend legitimacy to its resulting data, and many of the accuracy assessments which have been performed are specific to a single data source product in a particular type of landscape. Current research in this field is unable to provide the *CIAT Hillsides program*[*] with the information required to include validated DEMs in their agricultural models, as findings are non-transferable or inapplicable to the data and complex landscapes encountered in the Hillsides areas.

An additional problem for the CIAT research program is the severe lack of information regarding the *costs* associated with producing a DEM and their relationship with the level of accuracy which can be expected. It is generally understood that the data sources for producing DEMs which are believed to be more accurate are also more expensive (Rhind, 92) although little work has been conducted to investigate this relationship and to quantify the level of performance of data in the low-cost categories.

This project was set up in an attempt to provide CIAT with an understanding of the accuracy and costs of DEMs which are produced for use in the hillsides regions, and to answer the following questions; Which method of low-cost DEM production is most appropriate for the work in CIAT?; How accurate will each data set be?; What is the cost of production for each of these data sets? The evaluation and validation of the models used in research studies conducted by CIAT is part of an ongoing strategy to

---

[*] Centro Internacional de Agricultura Tropical, Cali, Colombia. Member of the CGIAR network of agricultural research centres.

develop procedures for quantifying and refining the precision of input data sets to spatial analysis projects.

A number of specific objectives were outlined at the beginning of the project. These were as follows:-

- To determine the level of accuracy of slope and altitude values derived from DEMs which have been produced using different scales of cartographic data.
- To examine the relationship between increased accuracy of a DEM and corresponding increases in the cost of production.
- To examine the relationship between cell resolution and the correlation between 'true' and derived slope values.
- To examine the occurrences of high error and document where they are, develop reasons as to why they are occurring and work towards reducing them.
- Finally, it is hoped to increase the *users* awareness of the level of accuracy which can be expected of DEMs which are being used in CIAT projects and to ensure that future DEMs will be produced according to agreed standards.

3. STUDY AREA

In order to fulfill the Hillsides program goals it is imperative that accurate and up-to-date information be collected on the area, to provide decision makers with the tools they need to understand the patterns and processes involved. In particular *Digital Elevation Models* (DEMs) provide aspect, altitude and slope data which can be used as input to agricultural and hydrological simulation models, which are used to understand the consequences of proposed changes to the agroecosystem of the area. In a rugged landscape such as the ones encountered in hillsides, height and aspect change dramatically over very short distances, and errors can accrue rapidly in the absence of very dense (and expensive) data sets. Thus, it is imperative that the DEMs are accuracy checked prior to their inclusion in any decision support system. This study was conducted to provide validation of the DEMs which will be used by the Hillsides program, and thus the study area was selected in an area which was considered to be typical of the *Tropical American Hillsides*.

A site was chosen which covers 30km^2 and is located within the Rio Ovejas watershed, in the Andean foothills of Southern Colombia (see figure 1.2). There are a number of reasons why this area was considered suitable for the project;

Physical Characteristics
The region has an elevation range of 1,400 - 2,200 meters and can be considered typical of many hillside areas found in the Andean regions of South America. It is dissected by two large rivers and has a complex network of tributaries. There are many v-shaped valleys with steep lower slopes and high plain areas, (See plates 1, 2 & 3). As a result it is considered to have a sufficient range of topography on which to test the representational capabilities of DEMs for the Hillsides program.

Data Availability

Both cartographic and photographic data were available for the area at a range of scales. CIAT were involved in a watershed study in the area and as a result several photo-control points had already been recorded and a permanent GPS control station had been set up nearby.

Manageability

The area had to be of manageable size in terms of time spent in the field, computer model size (number and diversity of pixels for raster storage), complexity and processing time. Thirty square kilometers was considered large enough to include a diverse range of landscape features yet small enough to be handled in an efficient manner.

Accessibility

The area is easily accessible from CIAT by car and the road network within the area is also good. However, as the roads are largely unpaved and there are steep slopes in the area, many of them become impassable after heavy rainfall. Thus, it was necessary that all fieldwork be conducted in the dry season.

Future Project Work

Finally, as it is a designated research site for the CIAT Hillsides program, the data collected for this study can be used to contribute to future CIAT projects.

Figure 2.1 The Location of the Study Are
Plate 1 Perspective Views of the Study Area
Plates 2 & 3 Photographs of the Study Area

4. DATA COLLECTION AND PREPARATION

4.1 Cartographic Data

The following section describes the procedures which were followed for digitizing data from Colombian topographic maps and interpolating the values to produce eight Digital Elevation Models for accuracy testing.

Source Data

Five topographic maps were used as cartographic input to create the test DEMs. All of these maps were produced by the Instituto Geografico Augustin Codazzi (IGAC[*]). The maps and the source data from which they were derived from are set out in Table 3.1. There are no documented accuracy specifications for Colombian maps. The maps are all horizontally referenced using a national Transverse Mercator Projection with the central meridian positioned through Bogota, the capital of Colombia. The maps are vertically referenced using Mean Sea Level in Buenaventura, a city located on the Pacific coast, 150km northwest of the study area. The contour lines were derived from aerial photographs using aero-triangulated photo- identifiable points, collected by the IGAC.

Model Name	Source Map Scale	Interval Between digitized Contours	Shreve magnitude of digitized rivers	Area of Map digitized	Total length of arcs-(Rivs &Contours digitized)	Original Contour Interval on Map	Map Sheets	Date of Publ.	Date of Source Photos.
10-25	1: 10,000	25m	>1	3,000cm^2 *(2 sheets)*	8,636 cm	25m	3431a1 3431a3	1967	N/A
10-50	1: 10,000	50m	>1	3,000cm^2 *(2 sheets)*	5,030 cm	25m	3431a1 3431a3	1967	N/A
25-50	1: 25,000	50m	>2	480cm^2	1,960 cm	50m	3431a	1967	N/A
100-50	1:100,000	50m	>3	30cm^2	335 cm	50m	343	1990	N/A
10-100	1: 10,000	100m	>3	3,000cm^2 *(2 sheets)*	2,487 cm	25m	3431a1 3431a3	1967	N/A
25-100	1: 25,000	100m	>3	480cm^2	960 cm	50m	3431a	1967	N/A
100-100	1:100,000	100m	>3	30cm^2	205 cm	50m	343	1990	N/A
200-100	1:200,000	100m	>3	7.5cm^2	85 cm	100m	2-809	1984	75/79/84

Table 3.1 Cartographic Data sources used to produce DEMs.

Model Names

The DEMs have been assigned names according to the scale of the input data source and the vertical interval between contours which were digitized to produce them. (Table 3.1, Columns 1, 2 and 3). Examples;, Model *10-25* was produced from a topographic map with a source scale of 1:*10*,000 and had a vertical contour interval of *25* meters; Model

---

[*] Colombian National Mapping Agency

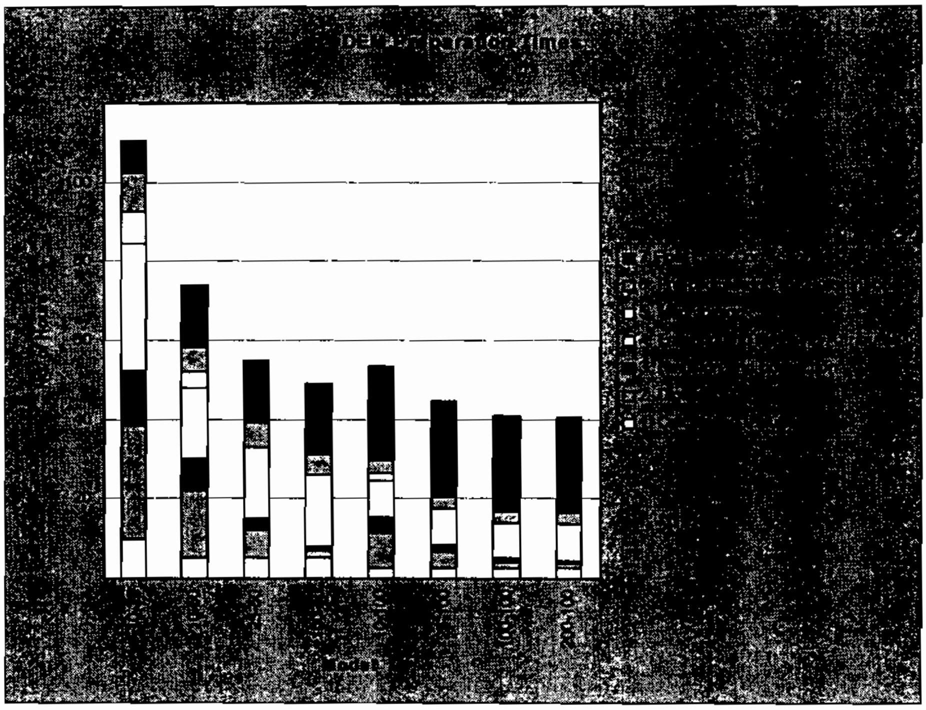
25-100 was produced from a topographic map with a source scale of 1:*25*,000 and had a vertical contour interval of *100* meters.

Digitizing

All of the maps were digitized by the CIAT Digitizing Unit, Cali, using the ESRI ArcInfo Digitizing System (ADS). The contour and river arcs for each model (as specified in Table 3.1) were digitized for the 30km^2 study area and also for a 500 meter *zone of interpolation*[*] surrounding its boundary. Some of the models were produced using data from the same map sheet but fewer contour lines were digitized. (i.e. Model 10-50 used the same source map as 10-25 but only every second contour line was digitized). The area of each map sheet used (Table 3.1,Column 5) differs significantly with scale, with some of the models requiring two map sheets to cover the study area. The total length of arcs which were digitized for each model (Column 6) provides a good indication of the relative workloads involved in preparing the source data.

Time/Costs

The amount of time which was taken to prepare each of the models is set out below (Figure 3.1). The tasks involved in preparing and correcting a DEM can be divided into seven categories.



---

[*] This refers to a strip of land surrounding the study area whose altitude values will be used in the interpolation process. When the DEM has been generated this area is then clipped away from the model leaving only the area of interest. If a zone of interpolation is *not* used there will be errors toward the edges of the DEM due to the lack of surrounding elevation values available for the interpolation computations.

Task Description	10-25	10-50	25-50	100-50	10-100	25-100	100-100	200-100
1. Calculating Values	10	5	5	5	2.5	2.5	2.5	2.5
2. Digitizing	28.5	17	7	2	9	4	1.5	1
3. Revision and correction	14	8	3	1	4	2	1	1
4. Entering and Revising Attributes	32	18	18	18	9	9	9	9
5. Map Joining	8	4	0	0	2	0	0	0
6. Pre-processing for Hutchinsons'	10	6	6	5	3	3	3	3
7. Final Error Correction	8	16	16	18	24	24	24	24
TOTAL (Hours)	**110.5**	**74**	**55**	**49**	**53.5**	**44.5**	**41**	**40.5**

Figure 3.1 Preparation Time for each of the DEMs

1. Calculating Values
This refers to the process of manually assessing and marking altitude values on the maps to ease the digitizing process.

2. Digitizing
This includes the setting up of tic points and digitizing of all contour lines, rivers and spot heights.

3. Revision & correction
The digitized maps are then compared to the originals to detect errors such as contour misplacement or missing arcs.

4. Entering and revising Attributes
Altitude information is added and error checked. The time taken for this process is mainly dependent on contour interval, rather than the scale of the map.

5. Map Joining
Edge matching is performed where more than one map sheet was used.

6. Pre-processing for Hutchinsons'
The data must then be prepared for input to Hutchinsons' interpolation algorithm (see page 24 for further detail). This involves ensuring that all streams are pointing downstream, preparing a *known sinkhole coverage*[*] where necessary, preparing an interpolation boundary and generalizing contour lines to minimize data concentration along contours. The time taken for this process increases as map scales increase.

7. Final Error Correction
Finally, there are errors which only become apparent after producing and accuracy checking an initial DEM. Contour comparisons and slope profiles can reveal rivers which do not follow valley floors, incorrectly coded contours, missing lakes, and areas where supplemental altitude data is required. There tend to be more errors of this kind in coverages which have been digitized from smaller scale data, due to the close proximity

---

[*] A coverage which contains information about natural depressions or waterholes in the landscape. It is used to ensure that they are preserved in the final DEM. All other depressions are treated as data errors and removed.

of contours on the map and complexity of the data. Thus, there is an inverse relationship between map size and time spent on this phase; i.e. the smaller the scale the more post-processing work is necessary to maximize return from the data.

Costs

The costs displayed below (Figure 3.2) provide a useful indicator of the *worth* of each of the models when examining their level of accuracy in later sections. In calculating costs it was assumed that tasks 1-5 could be performed by a trained digitizing operator whereas tasks 6 and 7 would need to be performed by someone with a higher skill level, including a good knowledge of DEM production methods and interpolation algorithms. Thus, costs were calculated according to wage levels, assigning \$8 per hour for tasks 1-5 and \$15 per hour for tasks 6 and 7.

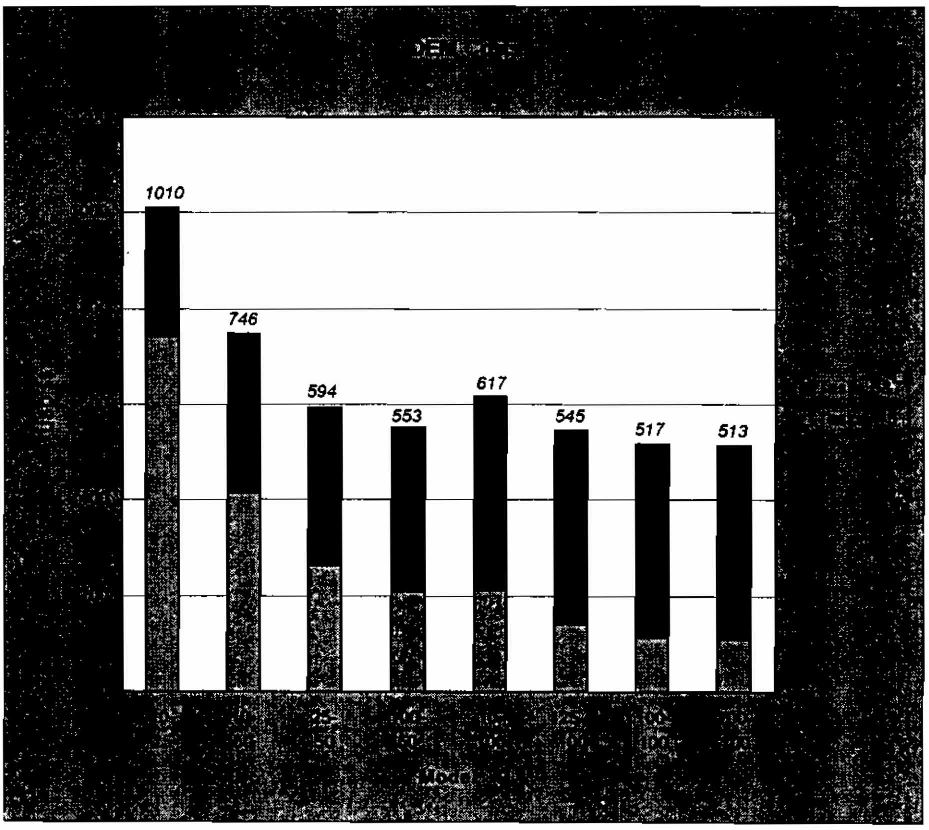


Figure 3.2 Cost in US Dollars for producing each of the DEMs

The differences in cost appear minor with regard to models 100-50, 25-100, 100-100 and 200-100. However the size of the study area is relatively small (30km^2) and minor differences at this level could lead to significant savings when working with larger areas.

The estimated costs per square kilometer are shown in Table 3.2. A degree of caution should be exercised if using this chart to calculate DEM production costs for larger areas.

Model	**10-25**	**10-50**	**25-50**	**100-50**	**10-100**	**25-100**	**100-100**	**200-100**
Cost	**$33.67**	**$24.87**	**$19.80**	**$18.43**	**$20.57**	**$18.17**	**$17.23**	**$17.10**

Table 3.2 Cost per model per square kilometer (US Dollars)

Considerations such as the *number* of map sheets that the area covers should be taken into account due to the additional tasks involved in 'stitching' multiple map sheets together. DEMs which include large tracts of flatter terrain would also be considerably cheaper due to the reduction in contours lines and increase in resolution required to represent them (see next page). It is also reasonable to assume that as the size of the area increases the cost-per-square-kilometer will decrease. In such cases large amounts of data can be handled simultaneously saving processing, checking and verification time-per-unit-area of land.

Additional costs such as hardware, software, running costs, map acquisition and training have not been included here, as they may vary greatly from one organization to another. The figures shown above should be considered to act as a good indicator of the *relative* cost of producing each DEM , *all other factors being equal.*

Although these costs may appear high they should be considered in context with the *worth* of the information. Field survey costs would be considerably higher if you were to attempt to collect altitude, slope and aspect data in a regular grid pattern over such a large area. There would also be a considerable amount of time and expense incurred in translating field survey results to digital form for analysis. DEM production does have high initial costs but once the model has been produced it is easy and cheap to manipulate and can be integrated into a wide range of agricultural applications.

4.2 DEM Production

The procedure for generating a DEM from topographic maps can be divided into three stages. Firstly, it is necessary to choose a data structure with which to represent the surface and, in the case of raster representation, to select an appropriate cell size. Next, a method of interpolation must be selected to transform the source data to a continuous elevation surface. Finally, there are a number of standard error checking procedures which should be undertaken to ensure that the DEM is an accurate reflection of the data which was used to produce it.

4.2.1 Data Structure

There are a number of different data structures which can be chosen to store the elevation data. The most commonly used are the *Triangulated Irregular Network*, whereby contiguous, planar triangles are fitted to the input data points (Peucker et al, 1978), and the *Grid* (or *Raster* based) data structure, which uses a matrix to store elevation values at regular intervals. A Grid structure was chosen for this study for a number of reasons.

Firstly, and most importantly, it is in a format which is easy to manipulate to derive secondary characteristics from the surface (i.e. slope, aspect) and the size of the data blocks can be standardized for multiple-grid comparisons. It is well suited to overlay and other spatial analysis procedures and can be linked to the many other coverages in CIAT which are also stored in this format. Finally it is stored in a compact structure using *run length encoding**, which makes it possible to open and work with a large number of grids simultaneously.

Cell Resolution

When using a Grid structure the most important consideration is the size of its pixels. This is known as the grid cell *resolution.* The resolution which is chosen will directly affect the level of generalization introduced to the data as it determines the size of each block of land which is represented by a single value in the grid. It is of particular importance when determining slope values as it affects the surface area which contributes to each slope value calculation. While it is desirable to have a cell size which is as small as possible to 'fit' the terrain closely, the size chosen must also enable efficient handling and storage of data and attempt to minimize data redundancy. A resolution of 5 meters was chosen as a *base resolution** for each of the models. This was chosen for the following reasons;

a) The Complexity of the Landscape This refers to the level of detail present in the landscape and is conceptually similar to the 'frequency' of the image in remotely sensed data. Different levels of complexity need to be represented using different cell resolutions. The level of complexity can be estimated by the steepness and distribution of slopes in any given area. To take an extreme example, one could say that the 'ideal' resolution with which to represent a 100km^2 area where the land is completely flat, would be with a single 100km^2 cell. Any smaller resolution would lead to data redundancy as a single cell would be sufficient to represent the landscape completely. As the landscape becomes more complex the optimal resolution to represent it depends on determining some measure of the complexity of its features. The study area in this project is considered to be highly complex as it contains slopes of up to 70 degrees with frequently changing aspect values. In a landscape such as this, if you were to move 5 or 10 m on a horizontal plane the difference in vertical height and slope direction could be considerable. This indicates that a cell resolution should be at, or lower than, these distances, to minimize potential errors in z-value readings.

b) The level of detail available in the source data. The density of source data for altitude and river channels is also critical when selecting a cell size which is optimal for the data. Important detail may be lost if a cell resolution is overlapping two or more contours lines in steep areas. It is desirable to retain as much of the input data

---

* A compression technique which stores long sequences of numbers as single references.

* All models were produced at a resolution of 5 meters and later aggregated to larger cell sizes to investigate the relationship between cell size and slope values.

as possible without generalizing it and losing valuable information. Although this may lead to data redundancy in flat areas (where the horizontal interval between the contours is wider) it is necessary to preserve the character of the more complex regions of the DEM.

c) The project application. The nature of the application is also of significance in selecting a cell resolution. If it is a large scale project where only a general idea of the regions with high altitude and slope are required, more can be gained by generalizing the surface (processing times will be faster and production costs lower) than by representing it with very fine detail. In the Hillsides Program at CIAT, agricultural modelling projects are typically concerned with small scale farming plots where many of the farms have less than 5 hectares of land. For modelling at these scales it is desirable to generate information at the field level rather than to have a single pixel representing several fields with different crops. A higher level of detail will be required in this case than, for example, the geometric correction of a 30 meter resolution remotely sensed image.

4.2.2. Method of Interpolation

To produce a DEM a method of interpolation must be developed to transform a set of discrete altitude data points (i.e. contour data and spot heights) to a continuous data set (i.e. a DEM). There are numerous algorithms which exist for this purpose (i.e. Kriging, Inverse Distance Weighting, Linear Regression Analysis, Splines etc.) and a choice between them should be based on the nature of the input data, the type of landscape and the data structure which has been chosen to store the DEM. Hutchinson's (1988) method of interpolation was chosen to generate the DEMs for this project, based on its proven performance in a series of accuracy tests conducted in similar terrain, in October 1993 (Rincon, 1993). This method utilizes a grid-based iterative finite-differences technique which honors input altitude data points and drainage channels according to a user defined set of 'accept/reject' tolerances. A successively smaller grid is placed over the input points and values are calculated for the cell centers using the data which falls within each cell. Following each iteration, drainage through the model is assessed and points which block flowpaths by less than a specified amount are eliminated to enable drainage. This results in a continuous surface which is *depressionless and drainable** and is thus very valuable for use in conjunction with agricultural and hydrological datasets.

Each of the models in this study was produced using Hutchinsons' algorithm (ARC/INFO TOPOGRID), using the 'accept/reject' tolerance settings shown in Table 3.3. The models were then error checked and verified as described in the following section. Where errors were detected the input data was corrected and Hutchinsons method of interpolation was re-run to produce a new model. This process was performed iteratively until the optimal model had been produced.

---

* Contains no walls or sinkholes which would prevent simulated water from flowing to the edges of the DEM.

Model Name	**Min. dist betw pts on input arcs. (Generalized using DouglasPeucker alg)**	**Specifications for Hutchinsons' Interpolation Algorithm**			**Filter Used (Post Algorithm Smoothing)**
		RMS	**Tol 1***	**Tol 2***	
10-25	5m	0.15	12.5	25	Low-Pass 3x3
10-50	5m	0.25	25	50	Low-Pass 3x3
25-50	5m	0.25	25	50	Low-Pass 3x3
100-50	5m	0.25	25	50	Low-Pass 3x3
10-100	5m	0.5	50	100	Low-Pass 3x3
25-100	5m	0.5	50	100	Low-Pass 3x3
100-100	5m	0.5	50	100	Low-Pass 3x3
200-100	5m	0.5	50	100	Low-Pass 3x3

**Tol 1 refers to the maximum height for which input data values blocking drainage are retained in the final DEM. Points blocking drainage by less than this value are eliminated to allow drainage to take place*

**Tol 2 refers to the maximum height for which data points surrounding a depression may be considered as an exit point for the water. If the surrounding heights block a water exit by more than this value, the depression will remain in the DEM.*

Table 3.3 Tolerances Set for Hutchinsons' Algorithm

4.2.3. Error Checking and Verification

Before a DEM which has been produced using Hutchinsons algorithm can be used as input to a spatial model it is imperative that the following standard error checking procedures are undertaken.

Contour Comparisons

Contours (set to half the input contour interval) were derived from the output DEMs and compared back to the originals to determine the *goodness of fit* to the original data. Where grave deviations were detected it was usually due to errors in the input coverages, (i.e. rivers pointing in the wrong directions, crossing contour lines, mis-coded altitude values etc.) or to lack of data. Where possible changes and corrections were made to input digitized data and the interpolation process was re-run.

Profiling and Histogram Analysis

Slope profiles can be useful for detecting and eliminating data errors (i.e. erroneous peaks due to mistyped values). Natural data (i.e. landscapes) are known to tend toward a normal distribution so badly skewed or misshaped frequency histograms of the data can also be used to provide a warning signal if there are serious errors in the data.

Profiles of the surface were taken at random throughout the DEMs, which, together with histogram analysis, were used to detect abnormal clustering of cells holding the same values as the input contours. This clustering was then alleviated somewhat by deliberate line generalization (i.e. thinning of data points along contour lines) using the Douglas-Peucker algorithm, to remove any redundant data input before interpolation took place, (Douglas & Peucker, 1973). The clustering described occurs as a result of the interpolation algorithm, in that many of the original input values (i.e. contour heights) are retained and unaffected by interpolation resulting in many cells holding *exactly* the same value whereas values for the intervening slopes are determined based on distance-weighting *between* the input data and thus there are more unique values. To further alleviate this unnatural 'terracing' effect, each of the DEM's were smoothed using a low pass 3x3 spatial filter.

It was decided that no further 'clean up' processes would be applied to the models in order to retain a relatively repeatable set of post-processing methods. Further editing would affect the integrity of the input data and thus bias later comparative analysis.

5. FIELD CONTROL

Two sets of ground control points (GCPs) were required to produce and quality check the Control DEM. The first set were used to georeference the photographs and to provide vertically referenced data. The second set were used to compute the vertical accuracy of the DEM. This set included over one hundred points which were selected throughout the area using a random stratified sample.

As this project was an investigation of *accuracy* it was imperative that the GCPs were accurate to within one half of one pixel resolution of the computed DEM (i.e. 2.5 meters). To achieve this level of precision, Global Positioning Systems (GPS) were used to provide field control.

In order to maximize accuracy for this project the following equipment and methods were employed;

Hardware
CIAT's GIS Leica GPS - System 200 equipment was used. This is a 'top of the range' GPS with the ability to collect data from up to nine satellites simultaneously using a dual frequency (L1 & L2), tracking system. If differential GPS methods are used (as in this project) decimeter accuracy can be expected..

GDOP
A maximum Geometric Dilution of Precision (GDOP) level of 2 was permitted. This is a computed indicator of the loss of precision due to unfavorable satellite distribution. distribution of detected satellites above the receiver at the time of measurement. a GDOP of 1 indicates optimum precision, while GDOP values of up to 5 are generally considered acceptable.

Number of Satellites
The minimum number of detected satellites allowed before taking a measurement was five. Generally, the higher the number of satellites the higher the ability to fix the location of the point and the lower the GDOP.

Differential GPS

Differential GPS methods refer to the simultaneous use of two GPS receivers, one positioned continuously over a *known point* located within a few kilometers of the second *roving receiver*. At the post-processing stage the data from both receivers is used to calibrate and eliminate error. The contribution of Differential GPS methods to the elimination of locational error is dependent on the length of the baseline (distance between the roving receiver and the base station). The shorter the baseline, the shorter the time needed at each point to resolve any locational ambiguities, and achieve a high accuracy.

Temporary Reference Station

To minimize baselines a temporary reference station was set up in the center of the area. This ensured that the maximum distance between permanent and roving receiver was under five kilometers, which is well within the recommended maximum baseline lengths for Stop-and-Go, Rapid Static and Static Surveying methods (Leica, 1992).

Measuring Times

Measuring times were set according to recommended levels used in Leica Field Tests. The temporary reference station was measured for over three hours with baselines of up to 20km between it and the 'known' points. Photo-control points were measured in Static survey mode for 30 minutes with baselines of under 5 km. DEM checkpoints were measured in Stop-and-Go mode with an initialization time of 10-11 minutes and 30-60 seconds of data at each following point. Again the baselines were never more than 5 kilometers long.

Using the methods described above a positional accuracy to within 25 centimeters was expected in the resulting data. The GPS work was organised and conducted in two field sessions. In session one the GCPs for error-checking the DEM were collected and in session two the photo-control points were collected. Three to five people and two vehicles were required for this work. One to maintain the base station. Two to drive and locate the field points, carry equipment, take fieldnotes etc.

Session 1: Set up of Temporary Reference Station & Collection of DEM checkpoints

This work was conducted over a seven day period in May/June 1995. The following procedures were adopted;

Preparation

Thirty six points within the area were generated at random using a C program. It was decided to visit each of these areas and to sample 3 or 4 points at each site.

A number of criteria were outlined for selecting the location of the temporary reference station. These included; vehicle access, freedom from obstruction (trees/power cables), high altitude (to avoid topographic interference), centrality (to minimize baseline lengths), freedom from human/animal interference and land access permission. Three potential sites were located using the aerial photographs and field knowledge of the area.

Reference station

The temporary reference station was set up in the northeastern part of the study area, near Buena Vista at *2°45'48.912N 76°30'24.82604"W (Latitude/Longitude).* Aero-triangulation was performed using known GPS stations located at *2°36'20.2254"N 76°29'9.61809"W*, *2°53'56.39144"N 76°29'9.61809"W* and *2°48'4.16066"N 76°33'19.67059"W*. The locations of the permanent and temporary reference stations relative to the study area. are shown in Figure 3.5.

Checkpoints

In all, 34 of the 36 random locations were visited (or as close as possible depending on accessibility). At each of these locations a number of points were collected situated up to 20 meters apart. In total, 126 checkpoints were collected using STOP and GO survey mode with a rapid static initialization of 11 minutes. Their locations are shown in Figure 3.6.

Session 2: Collection of Photo-Control Points

The GCP's for photo-control were collected on 29th January 1996. The temporary reference station which was set up for session one was used as a base station for this work.

Preparation

Twelve photo-control points were selected from the aerial photographs to use as reference control for the DEM. Having conducted extensive fieldwork in the area during session one this acquired knowledge of the area was invaluable in selecting suitable points and estimating travel times for field planning.

The control points were chosen according to the following criteria:
Accessibility - The quality of roads/paths had been noted during the first fieldwork session
Clarity of Photograph -each point had to be easy to recognize on *all* photographs
Location - Permanent field objects such as the corner of a house, or a concrete structure were selected rather than road junctions (where possible) as the roads in this area are unpaved and thus prone to rapid change.
Frequency- Points were chosen which occurred in all three photos (i.e. featured in both overlaps) so only 11 measurements were required to attain 7 GCPs *per* overlap. The extent of each photograph and the locations of the control points are shown in Figure 3.6.

Figure 3.5 Locations of the permanent and temporary reference stations.
Figure 3.6 Locations of GCPs and DEM checkpoints, Limits of the Orthophotographs

Before leaving for the field, enlarged graphic printouts of each photograph for *every* control point were prepared to aid correct identification in the field, and GPS field sheets were prepared (see Appendix A). A travel schedule was organised for the day, to optimize efficiency and aid radio communication in the field.

At each point, GPS data was collected using *Static Survey Mode*. The points were measured for 30 minutes each, and the maximum baseline was 4.65 kilometers. While waiting for the measurements, detailed sketches of the tripod position and photographs were taken at each location to ease identification at the post processing stage.

Post-Processing
The data from both field sessions was then processed in CIAT using SKI software, the static/kinematic post processing software for the Leica System. After the ambiguities had been resolved the calculated positions were transformed from Latitude and Longitude to the Colombian National Transverse Mercator horizontal reference system. All heights were vertically referenced to the WGS84 (for further information see page 34). Point coverages were then generated for the GCPs collected in session one (using ArcEdit), and stored for later verification of the Control DEM.

6. PHOTOGRAMMETRIC DATA

It has been stated (Maclaren & Kennie, 1989) that large scale stereo photographs are a highly accurate, although expensive, means of producing a three dimensional model of the terrain. For this study, a model was produced using stereo photographs at a scale of 1:28,000. This was assumed to represent the truest model of the terrain and was used as a control for assessing the value of the models produced using topographic maps.

6.1. Methods

A set of three stereo photographs were selected for the area according to their contrast quality, scale (1:28,000), date (1989), proportion of cloud cover and the availability of metadata. The negatives were then purchased from the *Instituto Geografico Augustin Codazzi* (IGAC) in Bogota, and sent to the USA where they were commercially scanned using a high resolution scanner (25 microns). The GCPs collected in the second fieldwork were then used to geographically reference the photographs and height values were determined using a soft-copy terrain mapping package (Helava) provided by the IGAC.

Sample points were measured at five meter intervals on the ground to correspond with the base resolution of the eight test DEMs, thus avoiding the need to resample the data points later on. Height values were then automatically extracted from the correlated stereo model using the strategy file and model parameters detailed in Appendix B. The DEMs were then checked visually and errors of 2-3 meters were corrected manually. It is acknowledged that additional errors may have been introduced to the process via canopy and cloud cover, as discussed below.

Forest Cover and Vegetation
Altitude adjustments were *not* made for forest cover or high vegetation. This is standard practice at the IGAC as in many mapping agencies around the world (Bolstad & Stowe, 1994). Due to the rapid rate of deforestation in the study area it is probable that the DEMs produced from early map sources (created over 25 years ago) will exhibit inconsistencies with the Control Model in recently cleared areas. However, short of producing a separate control model for each map source year (considered infeasible due to lack of data) there was no standard way to eliminate this problem. Suffice to say, a certain degree of inaccuracy on altitude readings can be expected in areas of (current or former) tall growth, which, in this region, was expected to particularly affect the thickly forested valley floors.

Cloud Cover
Due to the difficulties in obtaining cloud-free photographs of the area, a method for dealing with the distortions and error introduced by cloud had to be developed after the Control DEM had been produced.

In the stereo-correlation process, the height values within clouded areas were interpolated using a mean of the heights of the pixels which surrounded the clouded area. This interpolation process provides a 'best estimate' solution for a general-purpose DEM but in this project, where accuracy was 'of the essence', it was necessary to clip these areas out of the DEM and eliminate them from the analysis.

Areas which had been interpolated were identified through photo-analysis and automatically excluded where a height deviation of 15m or more was apparent between the Control DEM at a 5m resolution and a parallel model developed at a 10m resolution. A mask-grid was produced which, (as its name suggests) effectively masks out the areas where cloud or uncertainty in height values were apparent. The mask-grid was then used to eliminate the same areas from each of the cartographically derived DEMs. A second mask was created to eliminate edge areas from the computed slope values (i.e. where less than 9 of the cells in a 3x3 window contained data).

6.2. Verification Of The Control DEM

For verification of the Control DEM the GCPs collected in the first fieldwork session were used. Twenty-one of these points were dismissed as they occurred in areas which had been masked out of the DEM due to cloud or distortion. A further 15 were dismissed due to minor locational errors which were detected in the post-processing stage. Ninety points were considered reliable (to within 25 cms) and were used to check the accuracy of the Control DEM. Of these 90 points 15 were considered to be 'edge points' i.e. Those which fall on or near to the outer edges of the DEM. Edge points are valuable in assessing the quality of a photogrammetrically derived DEM as accuracy is typically reduced as you move away from the center of the photographs.

In accordance with the accuracy standards specified by the USGS[*] a *Level One* DEM (i.e. derived photogrammetrically) should have a vertical root mean square error (RMSE) of *not more than* 7 meters. The maximum RMSE permitted is 15 meters, (USGS, 1992). The RMSE is defined as:

$$\text{RMSE} \qquad \frac{\sqrt{\Sigma(Z_i - Z_t)^2}}{n}$$

where:

z_i	*interpolated DEM elevation of a test point*
z_t	*true elevation of a test point*
n	*number of test points*

In this study the computed RMSE was 4.26 meters which is well within the USGS error specifications (see Appendix C for figures) for a *level one DEM.* It is also worth noting that 83% of the checkpoints tested were within 5 meters of the DEM derived z-value, which is a good fit considering that the DEM cell resolution was 5 meters. As discussed earlier (and illustrated in figure 3.3) vertical errors of up to 4 or 5 meters *per 5 meter ground interval*, can be anticipated in a complex landscape such as this. In effect 10% of the points showed an error grteater than 50m, and 15% more than 30m. Maximum error was 244m.

6.3. Transforming The Vertical Reference System

The GPS points and consequent Control DEM both used the United States military World Geodetic System (WGS84) datum as a vertical reference system (VRS). The models produced from cartographic data were all vertically referenced to mean sea level in Buenaventura. Thus, before any comparative analysis could be performed it was necessary to transform the Control DEM to the same VRS as the cartographically-derived models. The US Defense Mapping Agency (DMA) have provided WGS84-to-local-geodetic-system *shifts* for local areas by considering one or more points within the local datum area and comparing them to the corresponding WGS84 co-ordinate (DMA, 1991). However, none of these *shifts* have been performed for the study area and the IGAC were unable to provide the required parameters. Thus, in order to provide a standard VRS to compare the models, the following transformation procedures were adopted;

I) IGAC aerotriangulated control points and their z-values for in and around the study area were read from the topographic maps, and used to generate a point coverage. The x,y and z values for these points had been used previously by the IGAC as GCPs for the stereo-photographs from which the contour lines used in this study were derived.

---

[*] United States Geological Survey

ii) The z-values for these points were then subtracted from the corresponding z-values held in the Control DEM. The WGS84 surface was consistently higher than the IGAC control points with a mean difference of 26.505 meters. A set of points which held the difference between each VRS as their z-value were then generated.

iii) A 'least squares' surface was fitted through these points using a first order polynomial regression function. The RMSE for this surface was 7.124 meters.

iv) The regression surface was then subtracted from the Control DEM to transform it to the same VRS as the models which were derived from cartographic data.

7. RESULTS

Each of the DEMs were assessed for accuracy based on their degree of similarity to the Control Model. This chapter describes the findings and conclusions which can be drawn from this assessment. Several recommendations for the production and use of DEMs are outlined.

7.1. Assumptions

The analysis and conclusions which follow are reliant on the assumption that the Control Model represents "truth" in the field. There are several factors which should be borne in mind regarding the results of this analysis.

1. The presence of noise in the control model.

Despite the automated removal of known and detected cloud areas from the Control Model it is likely that there are areas where wispy cloud or unadjusted forest cover were retained in the final output. These areas will exhibit small errors (of up to 10 meters) and may influence the reliability of the final results.

2. Bias in GCP site selection.

The GCPs which were used to error check the control DEM were collected in as random a method as possible, despite problems of accessibility. However, bias will have been introduced as a GPS can only operate in open areas. This eliminates the possibility of selecting checkpoints in forested areas or in deep gorges. Unfortunately, these are the places where maximum errors are expected in producing a DEM. As these areas were unchecked in computing the RMSE value it is probable that the Control DEM is less accurate than it appears.

3. Map Accuracy Standards

In Colombia, at the present time, there are no printed map accuracy standards such as those in the United States and other countries. Thus it is difficult to estimate the degree to which positional errors in the map contour lines may be affecting the accuracy of the DEMs. One can only assume that, without any strict controls, cartographic errors are likely to be higher than in countries with established standards. These kinds of errors will be propagated during the production process and may be responsible for some of the observed discrepancies between the Control Model and the DEMs which are based on cartographic data.

4 Results are specific to this type of landscape
The findings and conclusions drawn from this study are to a large extent dependent on the characteristics of the landscape being modelled, together with the quality of the Colombian-produced source maps used. Although they are considered to provide a useful indication of relative value of DEMs which have been produced using different scales of cartographic data, it should be borne in mind that the results may not be readily applicable to studies made in different countries with different topographic conditions.

7.2. Altitude

The models were first assessed with regard to the accuracy of their altitude values. Table 4.1 shows the descriptive statistics for each of the DEMs. After the removal of cloud from each of the models there were 991,573 cells (24.79km^2) which were evaluated. At first glance the figures in Table 4.1 reveal a high degree of similarity between the models. It is worth noting that the *range* of altitude values *narrows* for models 1-9 as the precision of the source data decreases. This is because the altitude values were determined using an averaging technique which, together with the lack of input “peak” or “pit” data in the small scale models, produces a smoothing effect resulting in 'higher lows' and 'lower highs' than those which exist on the ‘true’ altitude model.

Number	Model	MIN	MAX	MEAN	STD	RANGE
1	Control	1556.6	2188.8	1863.4	135.7	632.2
2	10-25	1558.6	2188.4	1863.1	137.3	629.8
3	10-50	1558.4	2171.9	1861.9	137.1	613.5
4	25-50	1561.5	2174.8	1861.0	135.2	613.3
5	100-50	1560.9	2172.9	1860.5	137.7	612.0
6	10-100	1593.1	2138.0	1861.4	134.0	544.9
7	25-100	1567.1	2133.4	1859.1	129.9	566.3
8	100-100	1585.8	2140.8	1859.9	134.3	555.0
9	200-100	1584.6	2151.2	1858.5	131.7	566.6

Table 4.1 Descriptive Statistics for Altitude models

A test of correlation was then performed on each pair of models to provide an *initial indication* of the relationships between ‘true’ and calculated altitude values. The correlation co-efficient (r), also known as the standardized co-variance, is a measure of the mutual relationship between two variables (Snedecor, 1967). The value range of r is from minus one to plus one. A negative value indicates that there is an *inverse* relationship between the datasets being tested. In other words, as the values on one model are increasing with regard to its mean value, the corresponding values on the model which is being compared with it will be *de*creasing, as illustrated in Figure 4.1. A value of zero would mean that the models being measured were unrelated. Finally, a positive value for r indicates a tendency for both of the models being tested to *increase* and *decrease* simultaneously. The closer r lies towards one, the stronger the similarity between the models. Perfect correlation, (i.e. $r = +1$) would mean that the *relative* topography in both models is identical, i.e. if there is a hillock or depression on the model representing 'true' altitude there will be a hillock or depression on the model which is being tested. However, it is possible that these hillocks, although identical in form, and in relation to their respective means, may exist at different altitudes. Thus, it should be understood that high values for r serve to confirm the existence of a close relationship and indicate the strength of such a relationship yet lack information as to the degree of *absolute* synonimity between the models.

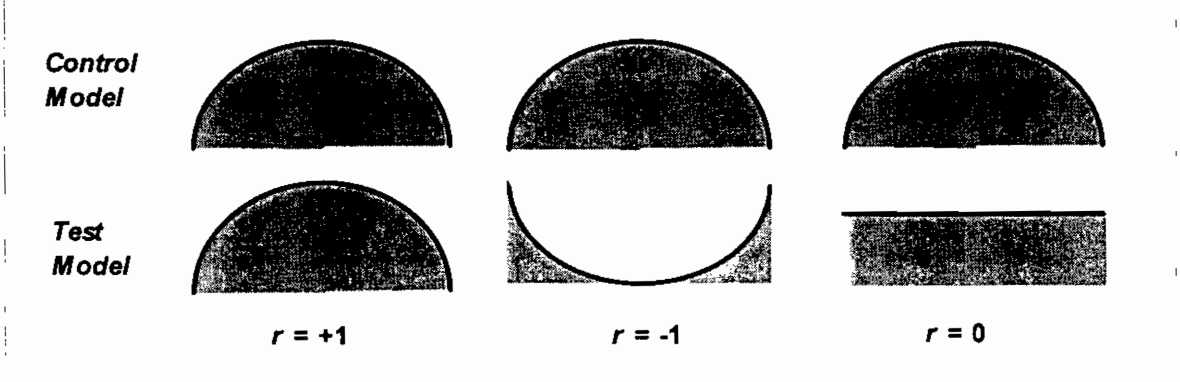


Figure 4.1 Cross-sections of landscape features illustrating significance of r value.

The correlation co-efficient is calculated in the following manner;

$$r_{ij} \quad \frac{[\sum_{k=1}^{N} (Z_{ik} - \mu_i)(Z_{jk} - \mu_j) / (N - 1)]}{\sqrt{\sigma^2_i \sigma^2_j}}$$

where:

Z	*value of a pixel in the DEM*
i,j	*the DEMs being compared*
μ	*the mean value of a DEM*
N	*the total number of DEM pixels*
k	*denotes a particular pixel*
σ^2	*the variance of a DEM*

The correlation co-efficient (r) was calculated for each pair of models using the total populations[*] (i.e. over 900,000 values). Thus, all of the r values can be considered significant at the 100% probability level. The results are displayed in the correlation matrix shown in Table 4.2 .

---

[*] All populations were first normalised using z-scores

Model Name	Control	10-25	10-50	25-50	100-50	10-100	25-100	100-100	200-100
Control	**1.0000**								
10-25	**0.9977**	1.0000							
10-50	**0.9972**	0.9989	1.0000						
25-50	**0.9952**	0.9962	0.9965	1.0000					
100-50	**0.9866**	0.9894	0.9902	0.9908	1.0000				
10-100	**0.9933**	0.9944	0.9957	0.9930	0.9879	1.0000			
25-100	**0.9886**	0.9893	0.9895	0.9920	0.9862	0.9914	1.0000		
100-100	**0.9844**	0.9863	0.9872	0.9875	0.9937	0.9894	0.9907	1.0000	
200-100	**0.9803**	0.9808	0.9816	0.9814	0.9821	0.9819	0.9821	0.9836	1.0000

Table 4-2 Correlation Matrix for Altitude Models.

From Table 4.2. it is clear that there is a strong positive relationship between each of the models and the 'true' surface, with all values for *r* lying very close to 1.0. There are a number of preliminary observations which can be drawn from this data, with respect to the *relative* strengths of the models. Firstly, the correlation co-efficients in column one for model 10-25 and model 10-50, are very similar with values of *0.9977* and *0.9972* respectively. The value for 10-100 is only slightly lower (*0.9933).* All of these models were based on the same map scale (1:10,000) yet a different number of contour lines were digitized in each case. The closeness of these figures suggest that it may be more cost-efficient to digitize every second, or indeed every fourth contour on a 1:10,000 map rather than every single contour. Secondly, the co-efficients for the degree of correlation between the Control as compared with model 25-100 (0.9886) and model 100-50 (0.9866), imply that the contour interval itself is not the *only* determining factor. In this case the closer contour interval of 50m produced a model which was actually *less* similar to 'true' altitude than digitizing with a contour interval of 100m due to the difference in base map scale which was used. Thus a larger scale with less contour lines may be preferable to a smaller scale map with more contours. This suggests that both the contour interval and the base map scale should be considered *together* before selecting the optimal data source for a DEM.

In order to examine the *absolute* differences which exist between the models, the altitude values were then compared to 'true' altitude on a cell-by-cell basis. The differences were then reclassed into error categories and the results were charted, as displayed in Figure 4.2. For the purposes of this study a height difference of up to 7 meters was considered 'good', bearing in mind the surface area of each cell ($25m^2$) and the probability for height alteration within each individual window (as illustrated in Figure 3.3). A difference of up to 15m[*] in height, although not desirable, was considered acceptable for the type of modelling undertaken for the *Hillsides* program. Thus, the percentage of each model which falls within classes 1 and 2 is termed "*usable*" and is hereafter regarded as an indicator of the value of each of the models.

---

[*] This allows a two-cell horizontal placement error which is considered realistic given the scales of the input data sources.

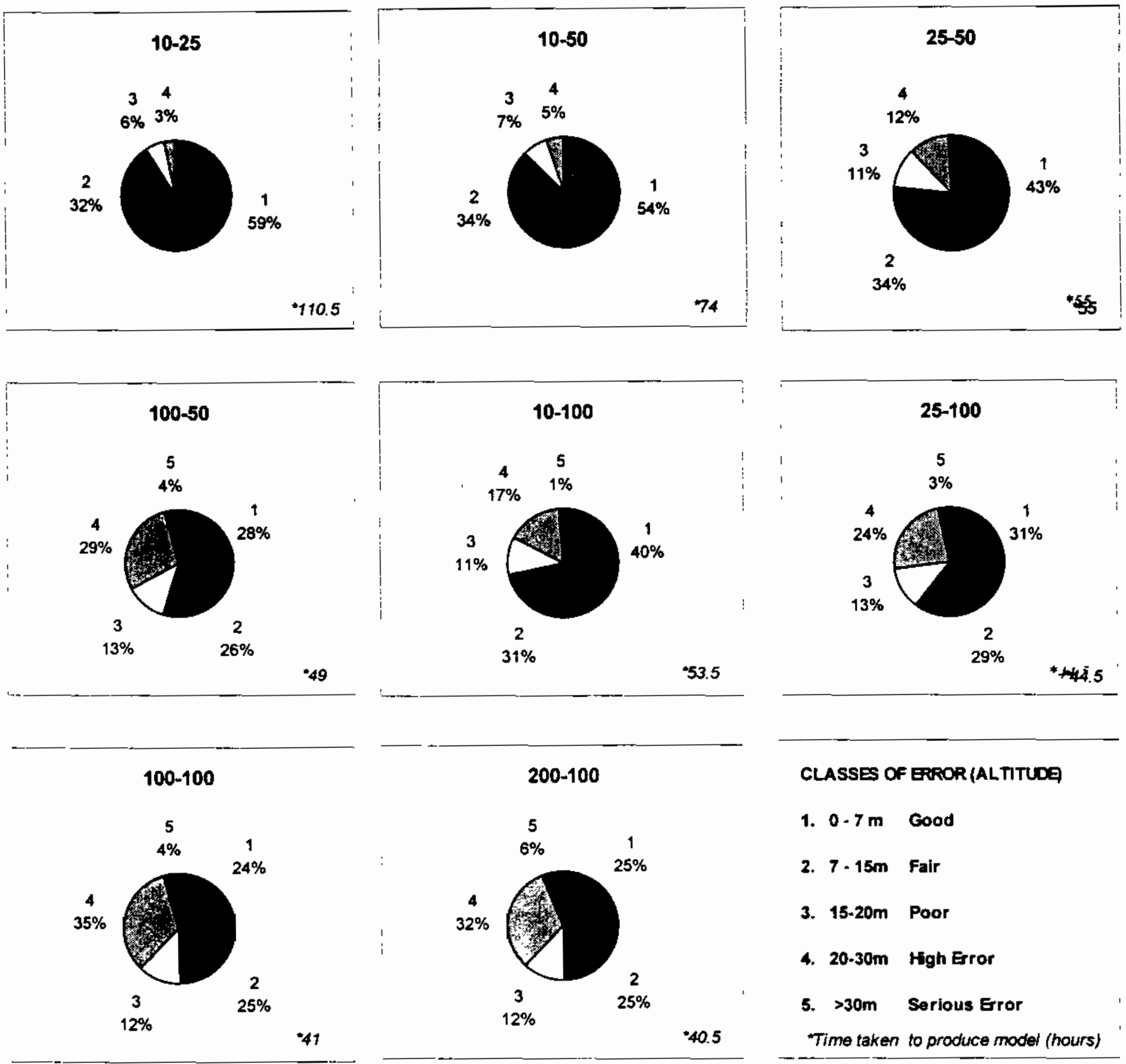


Fig 4.2 Difference between DEM altitude values and 'true' altitude. (all 991,573 cells have been classified)

In the charts shown above it is worth noting that in model 10-25, 91% of the cells are considered *usable*, as opposed to only 50% in model 200-100, at the other end of the scale. However, these figures should be considered with regard to the time taken to produce each of the models (shown in italics). Models 100-100 (49%) and 200-100 (50%) show very poor results, yet, if considered with regard to the short time taken to produce them, (almost one third of the time - and consequent expense of model 10-25), then they may be considered a reasonable option for producing a *general* representation of the terrain. Model 100-50 has only 54% of *usable* terrain which is not much better than these models and it is worth noting that this model actually took longer to produce than model 25-100 which shows better results (60%).

If we set a cut-off level of the *usability* of a model at 68% (i.e. over 68% of the cells must be within 15 meters of 'true' altitude otherwise the model cannot be used), there are only four models which are worthy of further discussion. These are models 10-25, 10-50, 25-50 and 10-100. It is significant that the list includes all three models which relied on the 1:10,000 scale map as a base data source, as this implies that even if a contour interval is closer on a smaller scale map, the level of detail which is lost regarding line positioning will have knock-on effects in the process of producing a DEM. In other words, a line which is misplaced by just one or two millimeters on a 1:100,000 scale map will affect the positioning of a whole hillside slope when a DEM has been produced. One millimeter on a 1:100,000-scale map represents 100 meters on the ground. On a 1:10,000-scale map it represents only 10 meters. Thus, the larger the base map scale, the less severe these kind of errors are going to be in terms of their effects on landscape positioning within the DEM.

Of the four models listed above, the differences which occur between models 10-25 and 10-50 are so minimal that the second model is preferable to the first in terms of time and costs saved in the production process (it takes less than 75% of the time taken to produce model 10-25 and is $250 cheaper). Models 10-100 and 25-50 also show very similar levels of accuracy, with 71% and 77% of each model (respectively) falling within an acceptable level of accuracy. As they take approximately the same time to produce, and model 10-100 costs only $1.27 more per square kilometer, a decision between them may depend on other factors, such as their ability to represent slope accurately (see next section). Other factors affecting the decision may be the amount of time spent on each individual phase of the production process. For example, the total times may be similar but it is clear from Figure 3.1 that the amount of time spent in the digitizing laboratory (tasks 1-4) are lower for model 10-100. This means that a higher proportion of the processing time for this model is spent on the tasks which require a higher skill level. It may be desirable to distribute the labor such that the more highly skilled personnel will have more time for other work.

Clearly an informed decision should be based on an understanding of the time and money which is available, together with the level of accuracy required. Model 10-50 appears to provide a good compromise at the top end of the field, and model 25-50 could be used to produce a less accurate, although *usable* model at three-quarters of the cost. The models which were produced using 1:100,000 and 1:200,000 source scale data have relatively low standards of accuracy and their use should be avoided if intending to conduct research at a local level, (i.e. for use in the CIAT Hillsides Program)

7.3. Slope

Calculating Slope

There are a number of algorithms which can be used for deriving slope values from a DEM. Most of these operate by fitting a function to four or more of the eight elevation values which surround the central cell (Evans, 1980). A 3x3 analysis window or *kernel* is passed over the grid and slope values are assigned to each cell based on the relationship between its height value and the heights of its immediate neighbors. Calculation methods include determination of the slope of the steepest fall or rise (Goetz, cited in Theobald, 1992), calculating directional finite differences within the window (with and without weighted kernels), (Sharpnack and Akin, 1969; Horn 1981) and using multiple linear regression to fit a surface to the data points (Skidmore, 1989).

Using any of these algorithms it is possible to produce a set of slope values which may be evaluated for accuracy based on the degree of correlation between *calculated* and *true* slope values. However the concept of 'true' slope is, in itself, problematic. In order to calculate a 'true' slope value one must break a continuous phenomenon up into discrete units. Consequently, the slope values which are calculated cannot, and should not, be considered independently of these units (usually known as *slope baseline* or *run length*) which have been used to compute it. In a DEM, slope values should be considered together with the size and number of *cells* which were used to compute them. The area which the slope value represents is a function of these variables. Thus, rather than stating a slope value as a single figure it is more correct to include a reference to the surface area which it represents.

Hodgeson (1995) studied a number of slope algorithms in order to determine the size of the area which each computed value represents. It was determined that Horn's algorithm (Horn, 1981), (which has been used in this study) computes a value which represents a surface area of twice the cellsize of the DEM. Horn's algorithm uses a third order finite differences method whereby all eight neighbors are included and a weighting system is added for cells which are adjacent (i.e. shorter baseline) to the central cell (see below). This method was assessed for accuracy by Andrew Skidmore (1989) and compared favorably with several other common methods for calculating slope.

Z_{-+}	Z_{0+}	Z_{++}
Z_{-0}	Z_{00}	Z_{+0}
Z_{--}	Z_{0-}	Z_{+-}

3x3 Filter

$$p_w \quad [(Z_{++} + 2\,Z_{+0} + Z_{+-}) - (Z_{-+} + 2\,Z_{-0} + Z_{--})]\ /\ 8\phi x$$
$$q_w \quad [(Z_{++} + 2\,Z_{0+} + Z_{-+}) - (Z_{+-} + 2\,Z_{0-} + Z_{--})]\ /\ 8\phi y$$

Slope x for $Z_{00} \Rightarrow \quad \tan x = \sqrt{}(p_w^2 + q_w^2)$

where: $8\phi x$ is the grid cell resolution from west to east
$8\phi y$ is the grid cell resolution from south to north

Slope values were calculated using the method shown above for each one of the models. Values which had neighboring cells without data values (i.e. beside cloud cover or at the edges of the DEM) were then eliminated. There were 979,808 cells (24.50km^2) which were used in the analysis.

Deviation in Slope Values

Table 4.3 shows the descriptive statistics for each of the slope models. Although the range of values increases as the accuracy of the input data decreases, it should be noted that the mean slope value is lower. As the standard deviation remains relatively stable this suggests that there are higher incidences of lower slope values in the models with lower data density. In other words, it appears as if the slope values are being underestimated, when represented by models produced from smaller scale data.

Number	Model Name	MIN	MAX	MEAN	STD	RANGE
1	Control	0.0	73.5	20.0	10.5	73.5
2	10-25	0.0	76.9	19.4	11.2	76.9
3	10-50	0.0	71.1	17.7	10.5	71.1
4	25-50	0.0	80.6	17.7	10.8	80.6
5	100-50	0.0	82.9	16.1	10.7	82.9
6	10-100	0.0	71.3	16.3	10.2	71.3
7	25-100	0.0	73.0	15.6	10.4	73.0
8	100-100	0.0	72.9	15.1	10.3	72.9
9	200-100	0.0	82.7	14.9	10.7	82.7

Table 4.3 Descriptive Statistics for Slope models - Slope values represent a $10m^2$ *surface area*

To examine this observation in more detail, nine samples of 475 cells each were taken from the normalized populations of the models. First, Analysis of Variance (the F-ratio test) as described in Ebdon (1981), was used to establish whether a significant difference existed between the populations. The following hypotheses were set:

H_o: The samples were taken from identical populations.
H_1: *At least one* of the samples was taken from a population with a significantly different distribution to the other samples.

An F-Value of 18.09 was computed and the Null Hypothesis was rejected at the 0.0001 probability level.

The Tukey-Kramer Studentised Range (HSD[*]) Test was then employed to examine the character and extent of the differences which existed between each of the populations (Kramer, 1956). This is a multiple comparison means test which is similar to the t-test but the probability of conclusive errors is lower as the MEER[*] is controlled, assuming equal sample sizes (®SAS, 1988). A pair of sample means are considered to be significantly different if:

$$|\mu_i - \mu_j| / s \sqrt{(1/n_i + 1/n_j) / 2} \leq q(\alpha;k,v)$$

where $q(\alpha;k,v)$ *is the* α*-level critical value of a studentised range distribution of k independent normal random variables with v degrees of freedom*

---

[*] Honestly Significant Difference Test

[*] Maximum Experimentwise Error Rate. Probability that the null hypothesis will be falsely rejected.

Three ' Tukey' Groupings were identified and are listed in Figure 4.3. The models which are marked with the same, single letter are not considered to be significantly different.

Alpha 0.05 df 4266 MSE 1.815246
Critical Value of Studentized Range 4.389
Minimum Significant Difference 0.2713

Tukey Grouping	√μ	N	Model	Ctrl - μ
A	4.20242	475	10-25	-0.2
A	4.18432	475	Control	0.0
B A	3.96079	475	10-50	1.8
B A	3.95840	475	25-50	1.8
B C	3.76228	475	100-50	3.4
B C	3.69269	475	10-100	3.9
C	3.56389	475	25-100	4.8
C	3.54960	475	100-100	4.9
C	3.54269	475	200-100	5.0

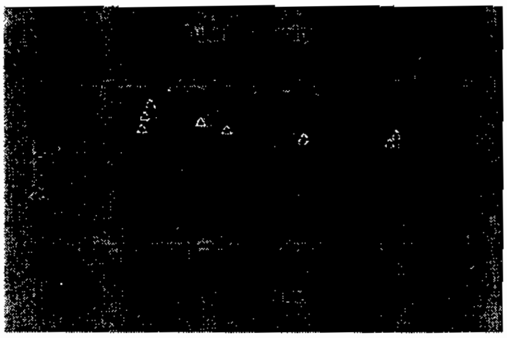


Figure 4.3 Tukeys Studentized (HSD) Range Test for DEM Slope populations

Using the groupings which were identified we can say, with 95% confidence, that the mean slope values computed by models 25-100, 100-100 and 200-100 differ significantly from the 'true' slope values, and are underestimating the slope values by an average of five degrees (Column 5). The slope values calculated by Models 10-50, 25-50, 100-50 and 10-200 are between two and four degrees *less* than the *'true'* slope values, the latter two models tending towards the larger errors. The only model which exhibits no significant mean deviation from the 'true' slope values is model 10-25.

These findings confirm the assumption that models derived from small-scale map data (i.e. 1:100,000 or 1:200,000) tend to underestimate slope values. It can be explained by the lack of topographic detail available in the source data which leads to larger tracts of interpolated land values. Due to the nature of most interpolation algorithms used these values will be placed along an unrealistically smooth scale between known data points. Although these findings are specific to the methods of interpolation and type of landscape which was modelled for this study, it is clear that a *general* correction factor should be applied to models derived from small-scale data. The correction factors shown in Table 4.4 can be used as a 'rule of thumb' compensation for generalized slopes in studies where only small scale maps are available.

Model	10-25	10-50	25-50	100-50	10-100	25-100	100-100	200-100
CF (Degrees)	0	+2	+2	+3	+4	+5	+5	+5

Table 4.4 Correction Factors for underestimated slope values (Cell resolution = 5 meters)

Correlation co-efficients were then calculated for each of the slope grids as compared to the Control model using the methodology described in the previous section (page 38). The low correlation values initially calculated (Table 4.5, Column three), reflect the inherent structural differences between the Control DEM and the models which were derived through interpolation. As the control model was never interpolated, any roughness (i.e. small undulations) occurring in the terrain have been maintained. As illustrated below, (Figure 4.4) it does not take a particularly high deviation in local altitude values to produce a high deviation in slope values.

1995	*1997*	*1996*
1996	*1997*	*1997*
1997	*1998*	*1997*

Altitude values in a section of the Control Model (Minor undulations are apparent in the surface)

1993	*1995*	*1997*
1994	*1997*	*1999*
1996	*1998*	*1999*

Altitude values in a Section of model G10-25 (Smooth surface)

Cell resolution = 5 meters

2	*2*	*-1*
2	*0*	*-2*
1	*0*	*-2*

Difference in Altitude Values Control Model - G10-25

Slope Value for central cells:
Control = 8 degrees
G10-25 = 26 degrees

→ **Difference in calculated slope:**
18 degrees

(*Slopes calculated using Horn's algorithm. Example taken from actual data - mapex 1062937.5, 793802.4, 1062967.6, 793823.1*)

Figure 4.4 Slope values calculated using 3X3 kernels. Minor altitude differences lead to high deviation in calculated slope values.

To decrease the roughness of the terrain which is visible in the Control Model and alleviate minor errors in elevation values, the surface was smoothed using a 3x3 low-pass filter. The figures shown in Column Four of Table 4.5, show the correlation co-efficients after this initial smoothing of the data. Although some improvement is visible, these figures are still not high enough to produce a trustworthy estimate of slope values.

Number	Model Name	Control model (raw data) *r*	Control model (after smoothing) *r*	Improvement in *r*
1	Control	1.00000	1.00000	-
2	10-25	0.41353	0.47412	0.06059
3	10-50	0.39581	0.45263	0.05682
4	25-50	0.34262	0.39042	0.04780
5	100-50	0.26454	0.30137	0.03683
6	10-100	0.27210	0.30867	0.03657
7	25-100	0.27334	0.30945	0.03611
8	100-100	0.22452	0.25464	0.03012
9	200-100	0.24281	0.27580	0.03299

Table 4.5 Correlation co-efficient for slope models (as compared to true slope)

The models could be smoothed repeatedly to iron out the surface undulations and thus increase the correlation between *'true'* and calculated slope. However, the resulting slopes, although more similar, would be highly generalized, and would represent a surface area far larger than twice the cell size due to the averaging of neighboring cells to achieve the smoothed effect. Thus it was considered undesirable to smooth the data further and alternative explanations for the high internal deviation between the DEMs were examined.

Potential Sources of Error

Various hypotheses were proposed to determine where the incidences of high deviation from *'true'* slope were occurring in order to examine them in a spatial context. Areas which were considered to be error-prone were identified, isolated and analysed separately. These included; *(a)* areas where thick forest cover existed, *(b)* areas with very steep slopes and *(c)* specific, identifiable terrain features, i.e. ridges and valley floors.

Firstly, higher errors were expected in forested areas. It is standard practice that no adjustments are made for canopy cover when deriving altitude data from stereo-photographs, (Bolstad and Stowe, 1994). This leads to inaccurate vertical height readings where trees or high vegetation is present, which were expected to lead to inaccurate slope calculations. Areas covered by thick vegetation were identified on the aerial photographs and isolated from the 'clear' areas for analysis. In Table 4.6 the differences between the representation of *'true'* slope in each type of area are shown. Although a slight improvement in slope calculations is apparent in the 'clear' areas the margin is too slim for this to be considered an overriding factor. Its lack of influence was possibly due to the fact that large forested areas tend to reflect the topography at ground level, having relatively uniform growth patterns within each lot. Thus this factor is more likely to affect the accuracy of altitude data than slope values, with the exception of the forest boundaries where steep slopes (from tree-tops to field level) may explain the slight difference in correlation values which was encountered.

Secondly, it was proposed that higher incidences of error may be occurring in areas with steeper slopes, as documented by Paul Bolstad in a study conducted in the Appalachians, Virginia (Bolstad et al, 1994). The study area was divided into those regions with steep slopes (over 20 degrees - 37% slope) and those with gentle to moderate slopes. There was a slight improvement in correlation between 'true' and calculated slope in the areas of gentle to moderate slopes (Figure 4.6), although, again, it was not considered a significant factor in itself to explain the low overall performance.

Finally, the cells in the Control Model were reclassified according to their position within the terrain, to examine whether errors could be linked to particular characteristics of the landscape. The area is characterized by steep, gorge-like valleys with abrupt directional changes at the base of the slopes, uniform less steep mid-slopes, and upper ridges with smooth directional changes, gentler slopes and limited vegetation. It was anticipated that higher errors would be found in the base of the valleys than higher up the slopes. To identify the valleys, water flow paths were first determined through the region. Areas with high accumulations of water were identified as valley floors which were then expanded by 25 meters on either side to include the lower slope areas. Ridges were identified using the same technique, having first inverted the DEM.

	(a) Forest Cover		*(b) Steepness of Slopes*		*(c) Terrain position*		
Categories	***Cover***	***No Cover***	***<20°***	***≥20°***	***Ridge***	***Mid-slope***	***Valley Floor***
Corr. to 'true' Slope (r)	*0.39025*	*0.50071*	*0.51376*	*0.45744*	*0.53198*	*0.51881*	*0.37501*
Count	*119,527*	*860,281*	*532,372*	*447,436*	*82,101*	*657,239*	*240.468*
Min	*-57.286*	*-61.919*	*-61.919*	*-54.615*	*-57.992*	*-61.919*	*-57.286*
Max	*54.180*	*66.254*	*19.642*	*66.254*	*41.521*	*66.254*	*47.740*
Range	*111.465*	*128.173*	*81.561*	*120.869*	*99.512*	*128.173*	*105.026*
Mean	*1.033*	*0.558*	*-3.164*	*5.114*	*0.048*	*0.670*	*0.663*
Stddev	*10.801*	*11.009*	*9.311*	*11.130*	*10.416*	*10.984*	*11.171*

Table 4.6 Potential sources of error in the calculation of slope. Specific features were isolated and comparisons were drawn between the slope values calculated by Model 10-25 and by the Control Model, in these areas.

As expected, the correlation between slope values calculated from cells in the mid-slope and ridge positions and 'true' slope were slightly higher than those for the valley floors, (Figure 4.6). However the mean and standard deviation were relatively similar, hence, as in the case of forested and steep slope areas, the improvement to be gained by eliminating these areas is minimal. Thus, although the slope deviations could, to some extent, be explained by the factors examined above, it became apparent that the cells which exhibited the severest levels of deviation were dispersed more-or-less *randomly* throughout the region.

This implied that the majority of errors were due to structural differences within the models. The high resolution of the DEMs is the most influential factor affecting the level of detail (i.e. surface undulations) in the models. As discussed previously and illustrated in Figure 3.3, in an area of this character, a small locational error could lead to a large deviation in the recorded z-value. Although desirable to minimize the cell size to maintain terrain complexity for altitude values, the local errors which result from this level of detail contribute heavily to the deviation of slope values. The only way to increase slope accuracy in this case, is to increase the size of the area which the slopes represent, thereby avoiding local error and producing more generalized slope values, (i.e. perhaps calculating the slope for a larger section of a field (i.e. 15 or 20 meters wide) rather than breaking it up into 5 meter units).

The relationship between cell resolution and slope correlation was examined to determine whether an optimal cell resolution could be recommended for each of the models, which would reflect the quality of the input data. Figure 4.5 illustrates the effect that increasing the cellsize has on the correlation of slope values.

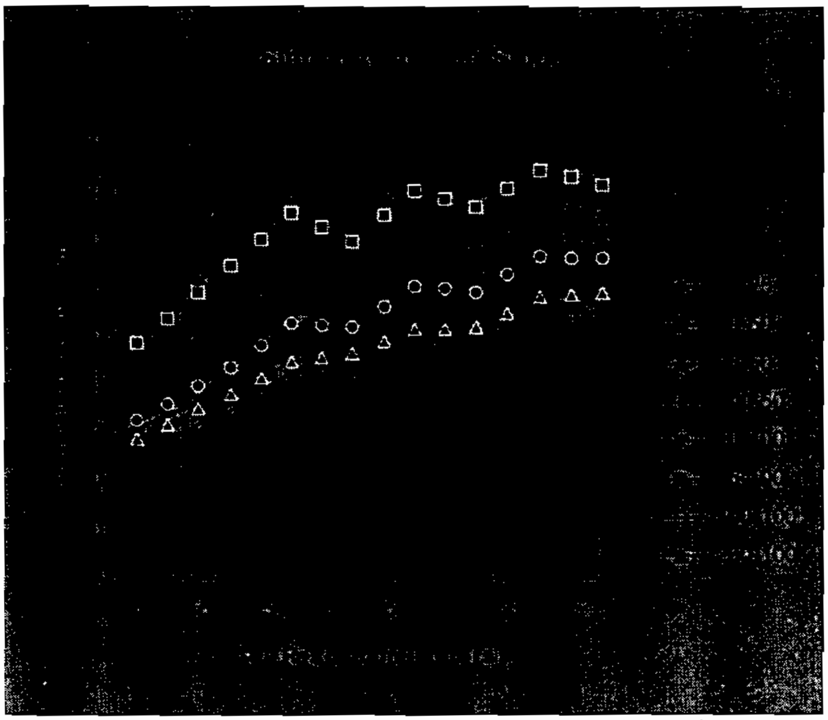


Figure 4.5 The Relationship between Cell resolution and Correlation to True Slope.

It is clear from figure 4.5 that the degree of correlation increases sharply as the cellsize is increased up to a *sillpoint* after which the rate of change increases slowly tending towards one. It seems safe to assert that this sillpoint is *the lowest* one can go with the cell size to produce a reasonable definition of slope. Below this level there is unwanted noise present in the data. It is clear from this graph that a different cell size suits a different level of input data.

It is worth noting that in Figure 4.5 the models with a wider contour interval have consistently lower correlation with 'true' slope despite their higher degree of correlation to 'true' altitude (see previous section). This confirms the importance of a narrow contour interval to maintain accurate intra-cell relationships. Clearly a DEM which has been interpolated from contour data at a 100m vertical interval does not contain sufficient detail to be accurately represented with a 5 meter cell interval. To attempt to do so would be lending a false sense of accuracy to the computed results. Therefore it is important to determine the cellsize which is best suited to the level of detail which is available in the source data.

To do this the correlation co-efficients were first compared with the internal cell deviations to determine a means of defining a level at which the r value represents a usable model. Figure 4.6 demonstrates the relationship between the correlation value and the proportion of the model which falls within 6 degrees of 'true' slope[*] . Thus the regression line can be used to determine the quality of a model at any resolution, with

---

[*] Six degrees was chosen as the cut-off level for the accurate representation of slope. Beyond this level errors were considered to be too high for use in agricultural databases.

regard to the proportion which is considered usable. If we set a cut off point of 68% (i.e. over 68% of the model must be within 6 degrees of 'true' slope), a model will only be accepted with a correlation co-efficient of 0.66 or more [i.e.. 68 = (85.7)(0.66) + 11.228.]

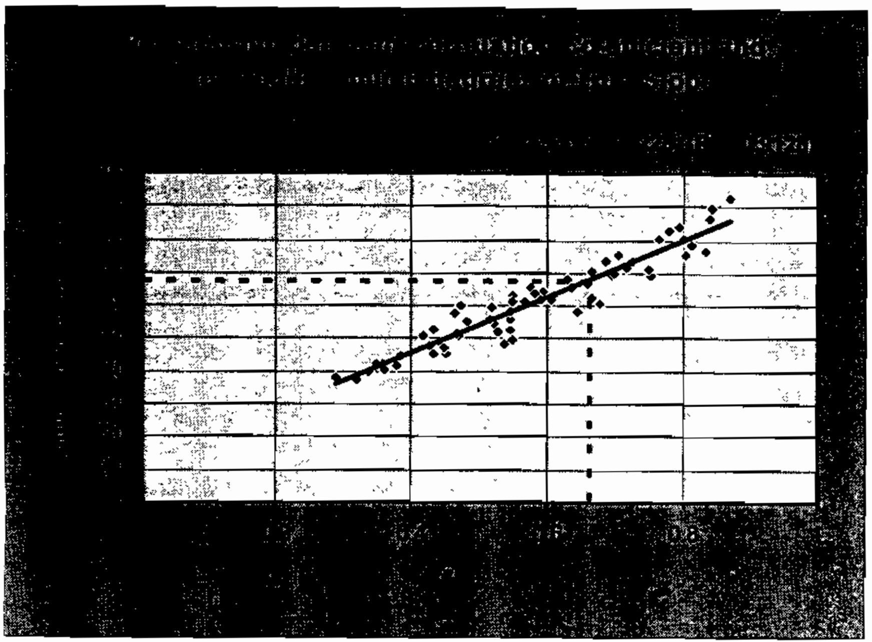


Figure 4.6 Cut-off point for determining model usability

Using this method we can refer back to figure 4.5 to determine *at what resolution* each of the models can be considered 'usable'. Thus, in this case (with a cut-off of $r = 0.66$) we can say that models 10-25 and 10-50 are best represented with a cellsize of 20 meters, 25-50 at 30 meters, and 10-100 and 25-100 at 70 meters. Models 100-50,100-100 and 100-200 aren't considered usable with a cellsize of less than 80 meters. Using these two charts it is possible to select a cellsize and model that suits the level of accuracy which is required. However, it should be stressed again, that increasing the resolution results in larger and larger slope surfaces which are being represented. Before using any of the derived slope grids it is necessary to decide whether these generalized slopes are good enough for the application. In the case of a Hillsides research project a cellsize of 30 meters is probably the maximum level of precision which you would want to work at. In this case, only models 10-25, 10-50 and 25-50 are considered *usable.* The other models may provide a good general representation and a relative indicator of high/low slope areas but cannot be relied upon for large scale modelling purposes.

8. CONCLUSIONS

The following conclusions have been drawn with regard to deriving altitude and slope values from DEMs produced using cartographic data sources.

Data Sources
It is possible to save a substantial amount of time and expense in producing a DEM from topographic maps by enlarging the interval between digitized contours. This research indicates that it would be cost-effective to digitize every *n*th contour on a map rather than every single line *as long as the new interval is not more than 25 meters wider than the original interval*, when modelling altitude *or* slope. This interval can be increased to *50 meters wider than the original contour interval* in cases where you only wish to model altitude.

With regard to slope determination, the contour interval has more influence than the map scale in providing realistic slope representations. Models with a wide contour interval were less accurate in modelling slope at acceptable resolutions (under 40m), regardless of the scale of the source map.

Cartographic data sources at a scale of 1:100,000 or 1:200,000 with a contour interval of 100 meters or more *do not provide sufficient detail* to accurately represent slope in the Hillside areas. If, however, these are the only source which is available in a designated research area it may be possible to *estimate* 'true' slopes by applying a correction factor to the calculated slope values. Small scale maps such as these have been proven to consistently underestimate slope by an average of 5 degrees. However, *even if* a correction factor is applied it is probable that there will be high incidences of error and it is recommended this method only be employed as a last resort.

Cell Resolution
The optimal cell resolution of a DEM is not necessarily 'the smaller the better' if intending to derive slope values in a complex landscape. A resolution which is too small incorporates too much noise (i.e. local deviations) in the data and can lead to erroneous slope calculations.

In selecting a cell size to model slope one must take the type of landscape *and* the level of detail in the source data into account. With a base map scale of 1:100,000 the best resolution you can use is upwards of 70 meters despite the desirability of modelling this kind of landscape at a smaller resolution. If the base map scale is unable to represent slope accurately at the desired resolution (i.e. up to 30 meters) then it should be considered unsuitable as an input data source for the DEM.

Filtering
Although low-pass filters can be used to "iron out" local errors in the surface they should not be used too liberally as such methods are ultimately affecting the slope values in the

same way as increasing the cell size. Namely, slopes derived from multi-filtered data will be representing a surface area which is several times larger than the cell itself.

Geo-Correction

With regard to using DEMs derived from topographic maps as a source for geo-correcting remotely sensed images, it appears that only scales of 1:10,000 and 1:25,000 (less reliable) would be usable. As the slopes derived using Horns' algorithm represent up to twice the cellsize, a resolution of 15 meters would be considered desirable to provide slope input to a 30 meter resolution satellite image. Unless using a different algorithm (i.e. Fleming & Hoffers (1979) represents only 1.6 times the cellsize), stereo-photos or other methods should be considered as a more suitable data source from which to derive the DEM.

Comment

Finally and most importantly, this information goes some way towards providing users with an *awareness* of the level of accuracy which can be expected of a DEM. It is clear that the scale of the source data from which it has been produced is fundamental to the dependability of the model. Care should be taken to obtain specifications regarding production methods and error checking procedures and their results before attempting to incorporate a DEM *or any of its derived products* into a spatial modelling system.

9. SUGGESTIONS FOR FURTHER WORK

Measuring Landscape Complexity

One of the most challenging aspects of this study was the complex nature of the terrain and the difficulties it presented with regard to choosing the optimal cell size and input data density. In theory, if it were possible to categorize the level of landscape complexity in a given area, recommendations and guidelines could be developed to assist the non-expert in choosing the appropriate data source and cell resolution to represent it in digital form. It is possible that a classification scheme could be developed, using indicators such as the spatial frequency of the data via Fourier Transforms (Mather, 1987), the character of cell-to-cell aspect changes, slope frequency distribution, and the standard deviation of the elevations. Previous work by Evans (1972) and Pike (1988) has touched on these themes but, in the light of recent technological advances there is a need to bring them together in todays context and examine the implications which a landscape classification scheme could bring to the field of Digital Elevation Modelling.

Other Data Sources

This study concentrated on investigating the ability of DEMs to represent the type of landscapes typically found in the Hillsides of Latin America. Its findings were specific to these areas and the resources available for this particular CIAT project. However, there are numerous other sources of data and methods for producing DEMs which may produce more accurate results. From this study it is clear that 1:100,000 and 1:200,000 scale topographic maps are very unreliable sources of data for slope and altitude mapping in the Hillsides areas. Larger scale maps should also be handled with a certain degree of caution. It is acknowledged that satellite image data is considerably more expensive but it is possible that the expense could be justified by the increase in accuracy and availability of cloud-free data at very large scales. Work needs to be done to investigate these sources to provide a critical analysis of *all* data sources for producing DEMs.

Sources of Error
A number of potential sources of error in DEM values were discussed in this project, (trees, steep slopes, landscape position and cloud cover). It was determined that they contributed to the deviation in slope values to a certain degree, but, as their *individual* influence was minimal, it was not considered appropriate to examine their effects in greater detail. There is limited research into quantifying the effects of these errors on DEMs. Further study into identifying and possibly reducing these kind of errors would be useful for developing methods which work towards reducing or compensating for their effects on the accuracy of a DEM. Work towards standardizing the methods for accuracy checking and verifying DEMs would also be useful.

REFERENCES

ADKINS, Kirk; Merry, Carolyn (1994) 'Accuracy Assessment of Elevation Data Sets Using the Global Positioning System', *Photogrammetric Engineering & Remote Sensing*, Vol. 60, No. 2, February, pp. 195-202.

AUGUST, Peter; Michaud, Joanne; Labash, Charles; Smith, Christopher (1994) 'GPS for Environmental Applications: Accuracy and Precision of Locational Data', *Photogrammetric Engineering & Remote Sensing*, Vol.60, No. 1, January 1994, pp41-45.

BOLSTAD, Paul; Stowe, T (1994) 'An Evaluation of DEM Accuracy: Elevation', Slope and Aspect, *Photogrammetric Engineering & Remote Sensing*, Vol.60, No.11, November 1994, pp.1327-1332.

BRADLEY, R; Jones, R. J.; Siddons, P (1994) A Land Information System for Environmental Risk Assessment, *Mapping Awareness,* 8(8), pp.20-23

BROWN, Daniel; Bara, Thaddeus, (1994) 'Recognition and Reduction of Systematic Error in Elevation and Derivative Surfaces from 7.5- Minute DEMs', *Photogrammetric Engineering and Remote Sensing*, Vol.60, No.2, February 1994, pp.189-194.

BURROUGHS (1988), Patrick *Principles of GIS for Land Resources Assessment*, Clarendon Press, Oxford.

CEPAL (1990) Magnitud de la Pobreza en America Latina en los Anos Ochenta, *Comision Economica Para America Latina y el Caribe*, Santiago, Chile.

CHIENG, S;Luo, J; (1993) Application of GIS to Peak Flow Estimation, in Applications of Advanced Information Technology, Washington Conference, 18th -19th June, USA.

CIAT (1994), *Hillsides Program Annual Report 1993-1994*, Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

CONESE, C; Gilabert, M.A.; Maselli, F; Bottai, L (1993) Topographic Normalization of TM Scenes through the Use of an Atmospheric Correction Method and Digital Terrain Models, *Photogrammetric Engineering and Remote Sensing*, Vol.59, No.12, December 1993, pp.1745-1753.

DAY, T; Muller, A (1988) 'Quality Assessment of Digital Elevation Models Produced by Automatic Stereomatchers from SPOT Image Pairs', *Photogrammetric Record* 12(72), pp.797-808,October 1988.

DIKAU, Richard (1989) Digital Relief Models in Landform Analysis, in *Three Dimensional Applications in Geographic Information Systems,* Raper, J (ed), Taylor & Francis, UK.
DMA (1991) *Department of Defense World Geodetic System 1984, Its Definitions and Relationships with Local Geodetic Systems,* Second Edition, Defense Mapping Agency Combat Support Centre, Colorado, USA.

DOUGLAS, D; Peucker, T (1973) 'Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature', *Canadian Cartographer,* 10, No. 2, December.

EBDON, D (1981) *Statistics in Geography: A Practical Approach*, Basil Blackwell Publisher, Oxford, UK.

EDWARDS, Kathleen; Davis, Philip (1994) 'The Use of Intensity-Hue- Saturation Transformation for Producing Color Shaded Relief Images', *Photogrammetric Engineering & Remote Sensing*, Vol. 60, No. 11, November, pp. 1369-1374.

EKLUNDH, L; Martensson, U (1995) 'Rapid generation of Digital Elevation Models from Topographic Maps', *International Journal of Geographical Information Systems*, Vol.9 No.3 pp.329-340.

EVANS, S; (1980) An Integrated System of Terrain Analysis and Slope Mapping, *Zeitschrift fur Geomorphologie,* NF, Suppl Bd. 36, 274-295.

EVANS, S; (1972) General Geomorphometry, derivatives of altitude and descriptive statistics, in *Spatial Analysis in Geomorphology,* Chorley, R.J (ed) pp 17-90, Methuen & Co., UK.

GOODCHILD, Michael; Gopal, Sucharita (eds) (1989) *Accuracy of Spatial Databases*, Taylor & Francis Ltd, UK.

HELPKE, C (1995) 'State-of-the-Art Digital Photogrammetric Workstations for Topographic Applications', *Photogrammetric Engineering & Remote Sensing*, Vol. 61, No.1, January, pp.49-56.

HODGSON, Michael; (1995) What Cell Size Does the Computed Slope/Aspect Angle Represent? *in Photogrammetric Engineering & Remote Sensing*, Vol.61, No.5 May, pp.513-517.

HORN, B.K. (1991) Hill Shading and the Reflectance Map, *Proceedings of the IEEE,* Vol 69 No1, January, pp. 14-47.

HUTCHINSON, M.F.; (1989) A New Procedure for Gridding Elevation and Stream Line Data with Automatic Removal of Spurious Pits, *Journal of Hydrology*, 106 (1989) 211 - 232.

JONES, P (1993) *Hillsides Definition and Classification*, Centro Internacional de Agricultura Tropical, CIAT, Cali, Colombia.

KING, Douglas; Walsh, Pat; Cluffreda, Frank (1994) Airborne Digital Frame Camera Imaging for Elevation Determination, *Photogrammetric Engineering & Remote Sensing*, Vol.60, No. 11, November, pp.1321-1326.

KRAMER, C. Y, (1956) Extension of Multiple Range Tests to Group Means with unequal Numbers of Replications, *Biometrics*, Vol 12 No. 3, pp307-311.

LEE, J (1991) Analysis of Visibility Sites on Topographic Surfaces, *International Journal of Geographical Information Systems,* Vol 5, No. 4, pp413-429.

LEE, Y.C. (1985) Comparison of Planimetric and Height Accuracy of Digital Maps, *Surveying and Mapping*, Vol. 45, pp. 333-340, December.

LEICA, (1992) *Guidelines to Static and Rapid Static GPS Surveying*, October, Leica AG, Switzerland.

MacMILLAN,R.A.; Furley, P.A.; Healey, R.G. (1994) Using Hydrological Models and Geographic Information Systems to Assist with the Management of Surface Water in Agricultural Landscapes, in *Landscape Ecology and GIS,* Haines-Young, R, Green, D and Cousins, S (eds), Taylor & Francis, UK.

MATHER, Paul (1987) *Computer Processing of Remotely Sensed Images*, Wiley & Sons, UK.

McLAREN, R; Kennie, T. (1989) Visualisation of Digital Terrain Models: Techniques and Applications, in *Three Dimensional Applications in Geographical Information Systems*, Raper, J (ed), Taylor & Francis, London, UK.

PACHIKO, D; Ashby, J; Sanint, L.R; (1994), Natural Resource and Agricultural Prospects for the Hillsides of Latin America, Paper prepared for the IFPRI 2020 Vision workshop, Washington, DC, November 7-10, CIAT, Cali, Colombia.

PEUCKER, T; Fowler, RJ; Little, JJ;Mark, DM; (1978) The Triangulated Irregular Network. Proceedings of the DEM Symposium, St. Louis, May 9-11 pp516-540.

PIKE, (1988) The Geometric Structure: Quantifying landslide Terrain Types from DEMs, *Mathematical Geology,* 20(5), pp. 491-511.

RHIND, D (1992) Data Access, Charging and Copyright and their Implications for Geographical Information Systems, in *International Journal of Geographical Information Systems,* , Vol 6, No.1 pp.13-30.

RINCON, Mauricio (1995) *Validacion de Modelos de Elevacion Digital*, Unpublished Masters Thesis, Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

ROSENTHAL, W.D;Srinivasan, R.; Arnold, J.G.; Alternative River Management Using a Linked GIS Hydrology Model, *Transactions of the ASAE,* Vol. 38(3):783-790.

SAS (1988) *Users Guide*, Release 6.03 Edition, SAS Institute Inc., NC, USA.

SHARPNACK,D; AKIN,G;(1969) An Algorithm for Computing Slope and Aspect from Elevations, *Photogrammetric Engineering and Remote Sensing*, , 35, 247.

SKIDMORE, Andrew (1990) Terrain position as mapped from a gridded digital elevation model, *International Journal of Geographical Information Systems*, Vol.4, No.1, 33-49

SKIDMORE, Andrew (1989) A Comparison of Techniques for Calculating Gradient and Aspect from a gridded digital elevation model, *International Journal of Geographical Information Systems*, 1989, VOL. 3, NO. 4, pp. 323-334.Taylor & Francis, UK.

SMART, J; Mason, M; Corrie, G; Assessing the Visual Impact of Development Plans, in *GIS Applications in Natural Resources*, Heit, M;Shortreid, A (eds), pp 295-303, Canada.

SNEDECOR, G.W.; (1967) Cochran, W.G. *Statistical Methods*, Iowa State University F]Press, Iowa.

THEOBALD, David (1992) Accuracy and bias issues in surface representation, Ch. 9 in *Accuracy of Spatial Databases*, M. Goodchild (ed), Taylor & Francis.

TOUTIN, Thierry; Beaudoin, Marc (1995) Real-Time Extraction of Planimetric and Altimetric Features from Digital Stereo SPOT Data Using a Digital Video Plotter, *Photogrammetric Engineering & Remote Sensing*, Vol. 61, No. 1, January, pp. 63-68.

U.S. GEOLOGICAL SURVEY (1992) Standards for Digital Elevation Models, *National Mapping Program, Technical Instructions,* (Available at ftp site: nmdpow9.er.usgs.gov)

USERY, E. L. (1993) Virtual Stereo Display Techniques for Three Dimensional Geographic Data *Photogrammetric Engineering & Remote Sensing*, Vol.59 No.12, December, pp. 1737-1744.

VERTESSY, R; Wilson, C.J.; (1990) Predicting Erosion Hazard Areas using Digital Terrain Analysis, in *Research Needs and Applications to Reduce Erosion and Sedimentation in Tropical Highlands,* Proceedings of the PH Symposium, June, No. 192.

WELCH, R; Papacharalampos, D (1992) Three-Dimensional Terrain Visualization on Personal Computers: An application with Stereo SIR-B Images *Photogrammetric Engineering & Remote Sensing*, Vol. 58, No. 1, January, pp. 71-75.

WOODROW, D (1993); Environmental Applications of GIS in the Water Industry, *MAGISE*, 7(6), pp 22-23.

ZHILIN, Li (1991) Effects of Check Points on the Reliability of DTM Accuracy Estimates Obtained from Experimental Tests, *Photogrammetric Engineering & Remote Sensing*, Vol.57, No.10, October, pp.1333-1340