THE ESTIMATE OF TOPOGRAPHICAL VARIABLES FOR SOIL EROSION MODELLING THROUGH GEOPROCESSING

Márcio de Morisson VALERIANO*, **Gilberto José GARCIA****
*Agronomic Institute of São Paulo, Brazil
valerian@barao.iac.br
**São Paulo State University, Brazil
gilberto@rc.unesp.br

Working Group IV/2

KEY WORDS: Agriculture, Soil conservation, Watersheds, GIS, Raster, DTM/DEM/DSM, Model-based processing, Information extraction.

ABSTRACT:

The technical requirements for geoprocessing of topographical data were studied, from digitizing until the obtention of the topographical factor of the Universal Soil Loss Equation (USLE), through calculation of slope (angle) and length. The development of the methods were based on data from manual cartographic survey and geographical information system (GIS) results of São Joaquim creek watershed (Pirassununga, SP, Brazil). A preliminary study, on the spatial behavior of the variables, their relations to local relief and to the topographical factor sensibility supported the selection of geoprocessing optimal procedures and parameters to achieve the maximum quality of the generated digital terrain models (DTM), and therefore, of all GIS results. Tests with the slope calculation indicated the need of high resolution, and the importance of post-processing (smoothing) and numerical fit. A method to obtain slope length was developed using spatial analysis functions of GIS. Results of the anisotropic cost analysis were correlated to measureed length, requiring linear fit. Errors for digital estimates prevailed in steep relief areas.

RÉSUMÉ:

Les impératifs techniques pour geoprocessing des données topographiques ont été étudiés, de numériser jusqu' à l'obtention du facteur topographique de l'équation universelle de perte de sol (EUPS), par le calcul de la pente (angle) et de la longueur. L'élaboration des méthodes ont été basées sur des données d'étude cartographique manuelle et des résultats de système d'information géographiques (SIG) de ligne de partage de crique de São Joaquim (Pirassununga, SP, Brésil). Une étude préliminaire, sur le comportement spatial des variables, de leurs relations au soulagement local et à la sensibilité topographique de facteur a supporté la sélection des procédures et des paramètres optimaux de traitement pour obtenir la qualité maximum des modèles digitaux de terrain (MDT) produits, et donc, de tous les résultats de SIG. Les essais avec le calcul de pente ont indiqué le besoin de la haute résolution, et l'importance du post-traitement (lisser) et de l'ajustement numérique. Une méthode pour obtenir la longueur de pente a eté developped en utilisant des fonctions d'analyse spatiales de SIG. L'analyse du coût anisotrope ont été corrélés avec la longueur measureed, exigeant l'ajustement linéaire. Les erreurs pour des évaluations digitales ont régné dans des zones de soulagement raide.

1. INTRODUCTION

GIS (Geographical Information Systems) capabilities have been used to integrate environmental data towards erosion modelling in watersheds. Raster-based data related to rain, soil, topography and vegetation properties are overlaid to perform erosion model calculations spatially. This technique results, as applied directly, in a soil loss prediction map. Through the inverse application, it can be used to calculate adequate managing practices (Paez, 1992), as well as the adequate land-use spatial distribution in order to cause an acceptable erosion level.

Topographical data are of fundamental importance in determining levels of erosion (Zachar, 1981). The factors of the relief applied directly on most erosion models are slope inclination and length. Ocasionally, slope form and aspect, as the elevation itself, may be included in more complex calculus. The USLE-Universal Soil Loss Equation (Wischmeier & Smith, 1978) topographical factor (LS) takes into account the slope angle (S) and length (L), according to the following equation, as adapted to Brazilian conditions by Bertoni & Lombardi (1992):

$$LS = 0{,}00984\, L^{0.63}\, S^{1.18} \qquad (1)$$

Despite the importance of the topographical data, there is a relative lack of studies concerning the obtention of the related raster layers. GIS modelling of USLE has often been performed using slope automatically obtained from the DTM-Digital Terrain Model and slope length measured graphically on maps. The accuracy of these layers is seldom studied, and their obtention methods often are not specifically developed with real data support. This study was conducted to determine optimal procedures for generating automatically slope angle raster layer and to develop a GIS alternative for the obtention of slope length layer.

2. MATERIAL AND METHODS

2.1 Data sources

The study area comprises gentle terrain, in general, with rugged terrain in the northwest region, throughout the 3,142ha watershed of São Joaquim creek (Pirassununga, São Paulo, Brazil). Its elevation ranged from 600m to 820m asl., with slope angles ranging mostly between 2% and 20%. Slope length reaches 1,500m at maximum, in the southern sheds.

The topographic data were extracted from the 1:10,000 IGC sheets covering the area of the São Joaquim creek watershed (*Córrego São Joaquim*, *Bairro Barrocão* and *Bairro da Saúde* quadrangles). The contour lines are drawn at 5m intervals. The original map was enlarged by a factor of 2 to improve the accuracy of the topographical measurements and for easy digitizing.

2.2 Processing

DTM. The contour lines were digitized (with ILWIS) through a purposive sampling scheme (Gao, 1995), which allows the construction of accurate DTM. Purposive scheme consists in sampling elevations only on contour lines, in opposition to a sistematic scheme. The digitized contour lines were transferred to a raster layer (grid) of 10m resolution, in order to sample 1,000 spatially randomic points. A geostatistical analysis (with VarWin) was carried out with the resulting sample file (x,y,z structure) to obtain the optimal krigging parameters for interpolation. These parameters (sill, range, model type and nugget effect) were applied in krigging interpolation to generate raster elevation input layers into grids of 20m, 40m, 100m and 200m resolution. The file containing all digitized elevations was exported to Surfer, which performed the interpolations.

Slope. The input DTM layers were operated by GIS slope angle determination module under different conditions, such as resolution and smoothing procedures. To test resolution effects, slope module calculations were applied to construct the 20m, 40m, 100m and 200m grid DTM. Chosen the best resolution, smoothing (3x3 pixel mean filter) was tested on the DTM before operating the slope module and on the resulting slope layer itself. The result sets were compared to control data through regression analysis.

Length. Assuming the conceptual similarity between the anisotropic cost analysis concept (Eastman, 1995) and the slope length geographical distribution, this spatial analysis was used to estimate slope length, as topographical elements could be regarded as cost elements. Anisotropic cost analysis may be described as a spread function, which is one of the connectivity functions among GIS spatial analysis (Valenzuela, 1991). Spread functions model phenomena intensity as related to a measurement of distance, evaluating the movement of a resource through an heterogeneous surface. The anisotropic cost analysis was performed using Idrisi, and the input images were processed so as to extract topographic features from the DTM. The five anisotropic cost elements, direction, forces, source, friction and angular function were substituted respectively by the topographical elements aspect, slope, hilltops, unit distance and an angular function. The first two elements are simple to obtain and easily understood, but special developments were required to the rest. Direction can be taken as the aspect image directly, or after some pre-processing. Forces image has to vary between 0 and 1, so that slope shall be normalized by its maximum.

The hilltops correspond to the origin (or source) of the modeled movement. The hilltop image was achieved using directional gradient filters as to simulate the derivative of elevation, in four direction axes. After normalizing the 1st derivative images (by dividing by its module), a second application of the same filters enhanced the hilltops as well as the channels, in each direction, with gray levels respectively –2 and 2. The final hilltop image was generated overlaying the 2nd derivative images, for the different directions, and selecting the suitable gray level to generate a boolean image of the hilltops.

The unit distance is the flow pathway distance, on the surface, of runoff through each pixel. It is thus a function of the pixel dimensions (resolution), slope angle and direction. Unit-distance was calculated spatially through image overlay and scalar operations, after trignometric functions applied on the aspect and slope images.

The angular function is the only scalar element of the anaysis. The aspect image values express the direction of maximum force but not the only direction of positive forces. It is since necessary to describe the decay pattern of the forces, from the aspect direction (maximum) to the tansverse (nule), which is the angular function. The angular functions tested were cosine (to the 1st, 2nd, 10th and 100th power) and abruptly cut at 80°, 60°, 45° and 22.5°.

Regression analyses with control data of slope length allowed the selection among the tested angular functions, as well as was done with the many options and input parameters during the development of the four images. The development of the whole analysis followed a cyclic flux of tests, programmed with a macro routine, until no improvement was observed. Finally, all results were fit by regression equations, evaluated through their coeficients. After stabilishing all steps in optimal options and parameters, the entire sequence was programmed in a macro file (Idrisi macro language), which calculates L, S and LS factor by Equation 1. The results presented in this study are limited to L and S data, because of the regional utility of LS calculated by Equation 1.

2.3 Control data

The variables were measured graphically on map in order to achieve a reliable data set for the statistical analyses. The measurements were made systematically in a regular grid of 200m spacing, in a total of 786 samples. Samples were picked on the same positions of this grid also from the digital processing result images, so as to build the files of the corresponding pairs (processed/measured) for the regression analyses. Field observations were collected in selected points and representative sheds, and the use of this data was only to support discussion comments.

Slope control data was calculated in every grid dot by combining the measured distance between adjacent contour lines (horizontal distance), the map scale and the contour interval (vertical distance). On hilltops, the vertical distances were calculated as the difference between the upper contour and the peak elevation. When the sample dot was centered exactly in channels, the horizontal distance was measured between contour lines in the channel direction. One can make two important considerations about this measuremenbt method: there is an assumption that altitude varies linearly between the contour lines, untrue for the reality and for the constructed DTM; the area corresponding to each measurement is not constant, but an inverse function of the slope angle, with plain samples being represented by much more area than steeper ones.

Length control data were calculated as the hypotenuse having the vertical and horizontal distances as cathetus. The distances were measured from each grid dot to its slope origin, at the top, through the hypothetical pathway of the water. For the simplest case, the measurement was made upward from the sampling point, directed in a straight line 90° from the local contour, until it reached the top of the ridge. However, sampling points often fell in curved sheds so that the distance measurement have to be interrupted were the direction indicates no more upward movement. This cut must be placed where the measuring line and the surrounding contours turn paralell.

3. RESULTS

3.1 Slope

DTM resolution affected drastically the results, limiting maximum slope systematically as pixel size increased (Figure 1). Additionally, weak resolution caused more data dispersion for the lower slope values. Indeed, regression with control data showed the decreasing coeficients (r^2) 67.3%, 62.6%, 50.4% and 24.1% for slopes determined from DTM of 20m, 40m, 100m and 200m, respectively.

Smoothing was found to affect inversely when applied on the DTM or on the resulting slope (Table 1). It was shown to be a interesting technique to smooth the resulting slope image sucessively until the fourth application, when regression coeficients stabilize at its maximum. In other hand, DTM smoothing previous to slope analysis showed similar effect to the generalization caused by coarse resolution, decreasing drastically the regression coeficients.

Smoothing	None	1 filterings	2 filterings	3 filterings	4 filterings
on DTM	67.3%	48.5%	47.6%	37.3%	29.7%
on slope	67.3%	71.5%	72.6%	72.9%	73.2%

Table 1. Regression coeficients (r^2) of GIS calculated slope affected by smoothing

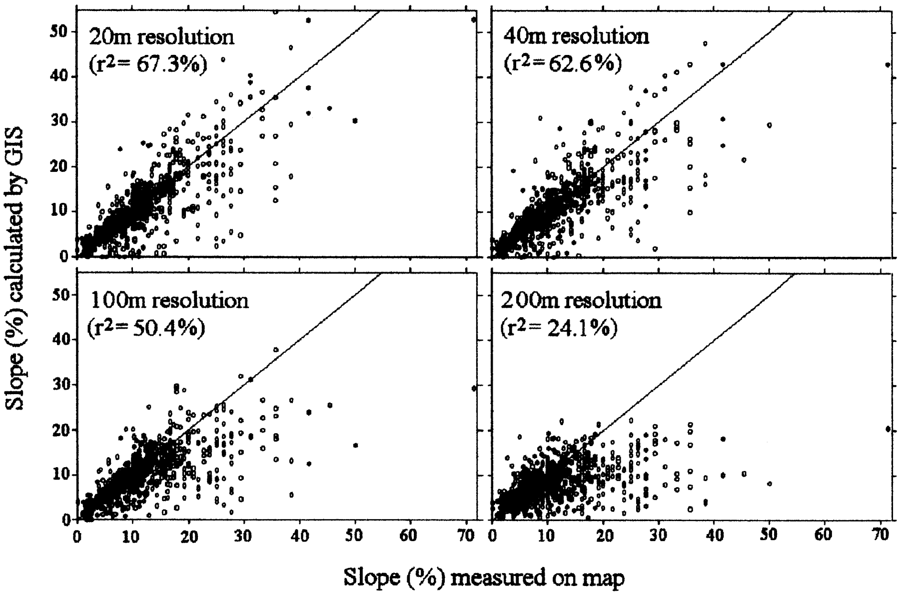


Figure 1. Dispersion of slope calculated by GIS under different DTM resolutions relative to control data measured graphically on map

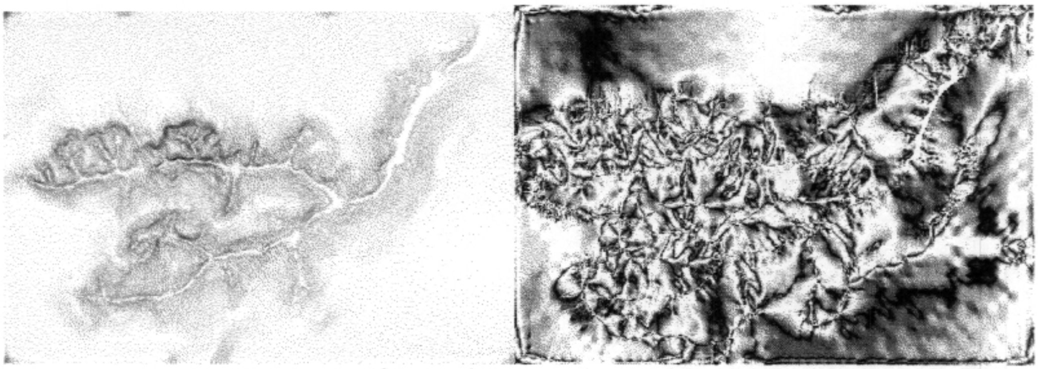
These results may be explainable by the spatial properties of terrain slope. It has been observed that high slope values are outstanding exceptions, dispersely distributed, among large areas of gentle terrain. Processing techniques causing generalization of the relief therefore tend to obscure these small areas, "flatenning" higher slopes. When applied on the slope result itself, smoothing caused a spread on areas of high values. The increase on the regression coeficients corresponded to a slight dispersion decrease (not shown), as effect of the mean filter on the calculus undesirable sensitivity. The mitigation of positional errors may be another mechanism of diminishing data dispersion, though the specific tests necessary to evaluate these effects are not possible with the availlable data of this research.

3.2 Length

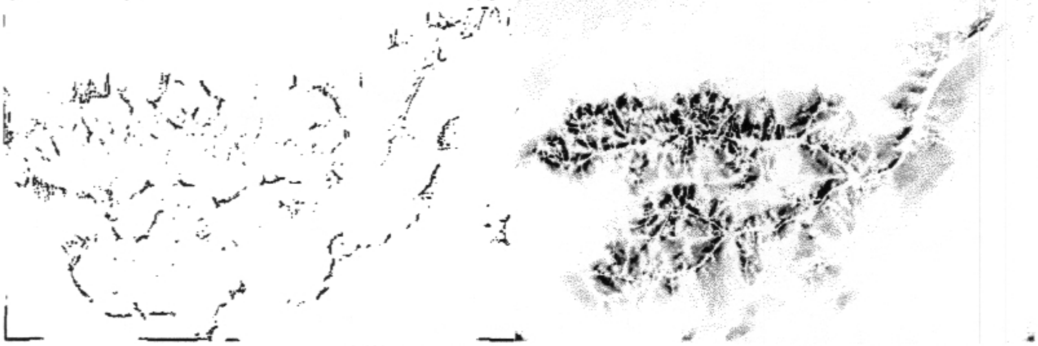
The four images prepared for the anisotropic cost analysis showed distinct appearance (Figure 2), though all descended from the same DTM. The slope, aspect and unit distance were presented with an inverted grayscale palette, darker for higher levels and white for zero. The slope image may give the impression of the actual appearance of the watershed, due to the fact that illumination is also distributed according to the slope angle. The brightness variation according to the horizontal exposure explains the metallic-paper appearance of the direction image, due to the behavior of specular reflexion. Unit-distance image combines properties of both slope and direction images, since it is calculated through trignometric functions of these. Hilltop image is presented with a 0-white/1-black palette for its boolean nature.

It was expected that the analysis would result in values near the length control data, but, as the parameters were changed, the resulted cost magnitude varied largely. They required always a factor between 2 and 10 to match the control data magnitude inside the watwershed, but they had extremely high values outside, increasing towards the image's edge. The application of a mask to remove this meaningless part of the image turned out to be important for the use and visualization of the resulting cost spatial distribution, which was in conformity to the slope length variation through the sheds. There was observed agreement between cost results and length control data, indeed, and the correlation coeficients varied mostly interesting, during the tests.

Best agreement was achieved with the images presented, which were selected in the cyclic test squence that evaluated each obtention procedure (and input parameter) through the achieved correlation coeficient. The procedures for the obtention of these images shall include some choices and specific operations that improve correlation, but there are also some that must be avoided. Cost correlated slightly better with length when forces image (normalized slope) was smoothed twice. Major improvements were observed when the direction image was obtained from the DTM smoothed 6 times. Angular functions were shown to strongly modify the spatial distribution of the resulting cost. The angular function $\cos^2$ was the best from the other tested functions. Correlation decreases with direction image smoothing or when the forces distribution is modified by taking square root sucessively. For each tested option, different fit equations were generated by the regression analyses, varying both intercept and linear coeficient.



forces (slope) direction (aspect)



hilltops (source) friction (unit distance)

Figure 2. Input images for the anisotropic cost analysis

3.3 Fitting the results

The developed procedures for the estimate of topographic variables, summarized in Figure 3, were codified into a macro program file, and applied definitely on the DTM. The fit equation applied on the last step varies as a function of the procedures, but it can be said that for slope, all equations had intercept around zero and linear coeficient around the unity. The only exception was the effect of working with progressively lower resolutions. The equations found by the analysis systematically compensated the reduction of maximum values by increasing linear coeficient.

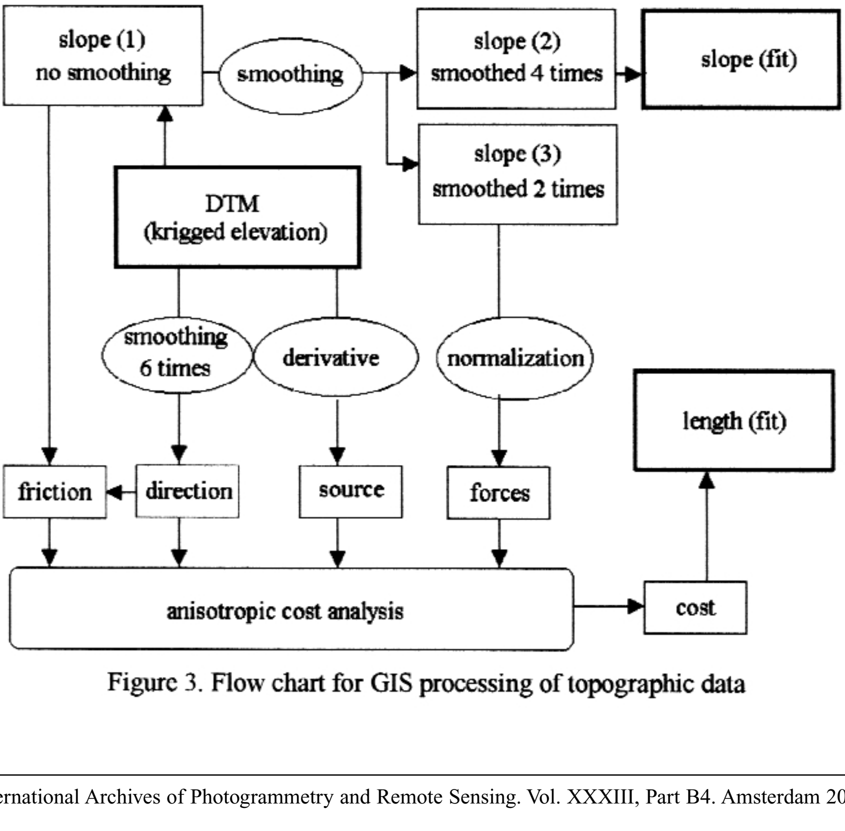


Figure 3. Flow chart for GIS processing of topographic data

With the programmed procedures, GIS estimated slope (r^2=73.2%) better than length (r^2=64.2%). The simplified fit equation (intercept=0) for the applied procedures are presented in the dispersion diagrams of Figure 4. Although length estimates presented more dispersion, it was found to be usefull through a wider range than slope estimates, which presented smaller dispersion, but only for low values. The few high values of slope were poorly estimated.

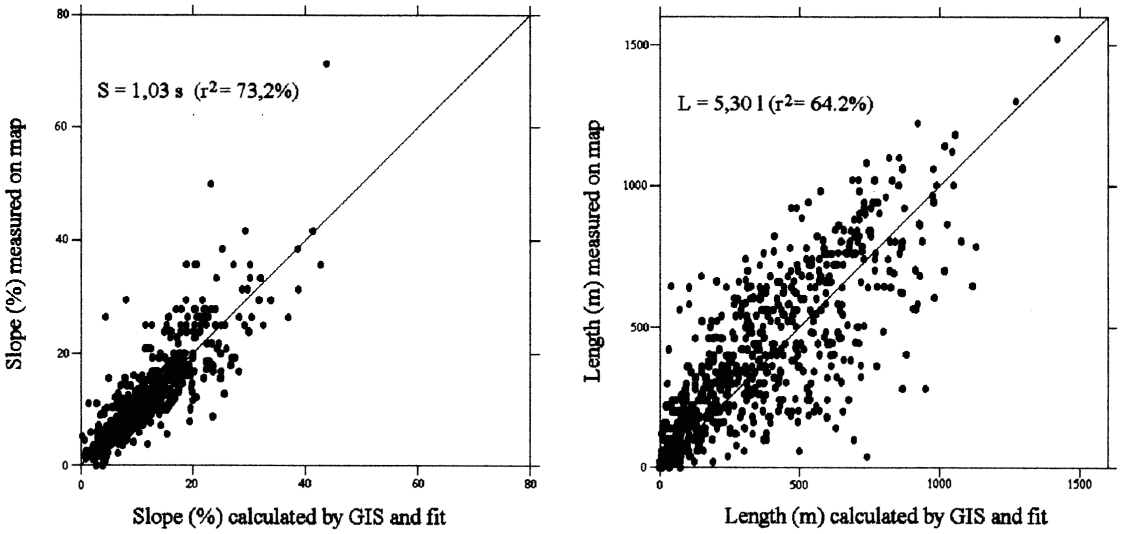


Figure 4. Fitting GIS results (s and l) for the estimate of slope (S) and length (L)

From the magnitude variation of its results, due to the sensitivity of the calculus, it is advisable to fit the anisotropic cost values using regression analysis with a cartographic-measured length data set. The fit equation for length presented at Figure 4 is valid for the exact set of procedures and parameters applied in this study, what may hardly be repeated. A set of length measurements on transects of representative sheds (paired with the corresponding anisotropic costs) is indicated to support the regression analysis. The regression equation found must be applied on the anisotropic cost image so as to fit cost values to slope length distance.

3.4 Spatial considerations

The relations between GIS-estimated and real topographic variables can be visually checked by looking simultaneously to the relief appearance (Figure 2, at the top left) and the digittal results as referenced in a geographical structure (Figure 5). But, the numerical similarities and diferences are better observed comparing the respective images to the control data presented in corresponding palettes and geographical basis. It is necessary only to ignore the resolution difference between the image sets.

The spatial distribution of GIS and the corresponding control data agreed relatively, while the two variables were distinctly distributed from each other. As expected from the regression analysis, slope mapping by GIS could be said to work perfectly in areas of gentle terrain, showing occasional differences in steeper areas, where slope slightly tended to be underestimated.

Depite the first-look geographical similarity between GIS and control data, length estimation was less satisfactory. The intrinsec connectivity of its spatial distribution was perfectly modeled, for the classes succed through shed transects in numerical order. However, the proportion between the level classes were strongly biased, with an apparent increase of the lower lengths classes on areas of lower slope and higher classes for steeper areas, when estimated by GIS. Neither multiple regression including slope, nor non-linear slope correction factors improved the correlation achieved through the exposed fit (linear) equation for the estimative of length by GIS.

The estimative of the topographical variables by GIS calculations with the described methods, though optimized, has shown marked mathematical limitations, from the point of view of data precision. It was found, however, that uncertainities are not exclusive to GIS estimations of length. Manual measurements are also imprecise because of the impossible to guess how much can a surface water flow curve its pathway. The measurement criteria is regarded to represent reality as weak as the empirically selecteed angular function ditribution.

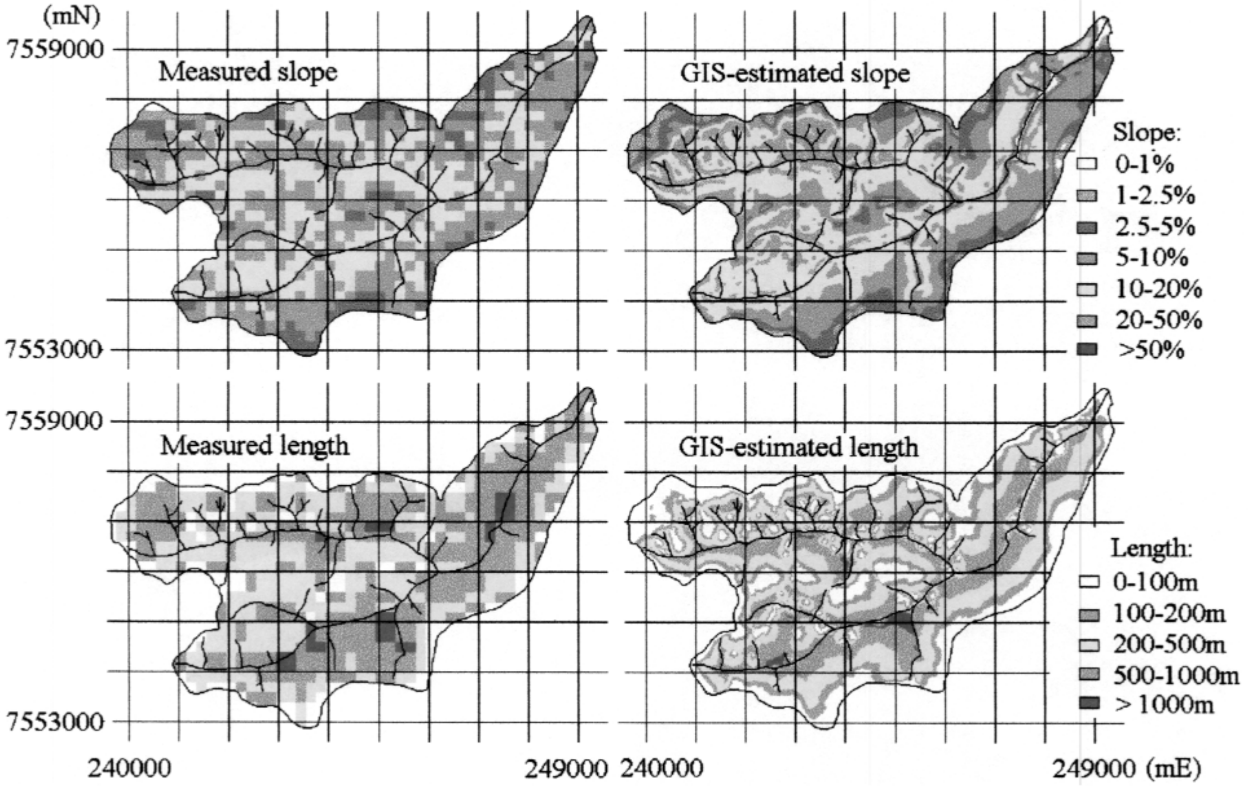


Figure 5. Measured and GIS-estimated results for slope (S) and length (L)

Nevertheless, the developed techniques allow mapping these variables with a higher resolution, since its computational nature, so as to generate geometrically detailed description of its distribution. The main importance of such approach is its operational convenience, as compared to the manual alternative, in addition to all the benefits of working with digital spatial databases. For slope, field data was strongly biased from cartographic measurements, much more than were these from GIS estimatives. The opposite occurred with length, noting that this variable field data could be better estimated by GIS than slope field data.

3.5 Errors

In order to verify the areas of higher deviations, the standard residue (Neter & Wassermann, 1974) was calculated for the estimative of both variables. Standard residue is the estimative deviation (error) normalized by its real value. The images were operated as to map standard residues, which were classified according to the legend of Figure 6.

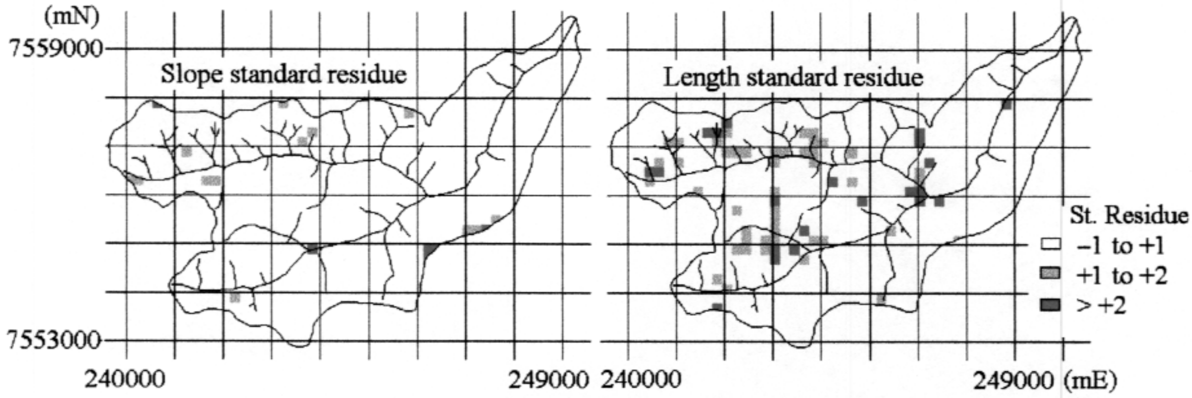


Figure 6. Spatial distribution of standard residues for the estimatives of slope(S) and length (S)

All major deviations were positive. Slope showed less residues than length, and the major deviations occurred in areas recognized as hilltops on the watershed limits and river channels. High residues were more frequent at the northwest portion of the watershed, were shorter and steeper sheds prevail. Curiously, the major standard residues for length and for slope did not occur in the same place (except once), despite the concentration of both in the same region. This suggests that the limitations to precise GIS estimating of topographical variables are different to slope and length.

The regional differences in the distribution of residues as related to the relief characteristics may encourage a research toward a stratified fit analysis. The regression equation was less influenced by data from the northwest region (with its particular characteristics) than by data from the rest of the watershed, because of its smaller suurface proportion. Further tests, on the classification of the relief oriented by the sheds' characteristics may be recommended, if one is interested in improving GIS estimatives of topogaphical variables.

4. CONCLUSIONS

This study was conducted to develop the obtention of topographical variables from Digital Terrain Model (DTM) so as to support erosion modelling through geoprocessing. The results led to the following conclusions:

Slope was adequately estimated by geoprocessing techniques when taking into account methodological cares. Higher resolution of DTM allows better estimative by Geographical Information Systems (GIS) of slope and, therefore, length. The superior limit of slope variation is reduced systematically with coarser resolution, decreasing correlation. It was found advisable not to work with resolution much coarser than 20m-40m. Slope also requires high resolution for its adequate description, since its spatial variability demmands detail. Field level observations showed topography mapping (and further, geoprocessing) to be generalized enough to hide the detailed variability of slope.

Otherwise, length presented more continuity, and may be represented (or stored) under lower resolution, but it also must be obtained under higher resolutions. The operational advantage allows the use of high resolution mapping, compensating partially the estimative numerical precision errors. Field observations were in satisfactory agreement with the geoprocessing results for length as well.

Length was conveniently estimated through geoprocessing, by means of connectivity functions, namely the anisotropic cost analysis. GIS-processing of DTM was capable of generating the images required to feed the analysis. Anisotropic cost was correlated to length, though magnitude must be fit by regression analysis using measured control data. Further efforts would be required in order to improve the use of GIS connectivity functions to estimate length and other related phenomena.

The estimates of slope and length were affected by the relief type. Gentle terrain had the variables better estimated than steeper areas. The errors were more pronounced and frequent around sheds with short length and high slope.

REFERENCES

BERTONI, J. & LOMBARDI NETO, F. Conservação do Solo. 3rd ed. São Paulo, Editora Ícone, 1992. 395p.

EASTMAN, J. R. Idrisi for Windows – User's Guide. Clark University. Worcester, MA, 1995.

GAO, J. Comparison of sampling schemes in constructing DTMs from topographic maps. ITC Journal, 1995-1. pp 18-22

NETER, J & WASSERMANN, W. Applied linear statistical models: regression, analysis of variance and experimental designs. Homewood, Richard. D Irwin, 1974. 842p.

PAEZ, M. L. Diseño de Practicas de Conservacion con la Ecuacion Universal de Perdida de Suelo. CIDIAT, Serie Suelos y Clima, SC-64. Mérida, Venezuela. 2nd ed., 1992. 130p

VALENZUELA, C. R. Data analysis and models (in: BELWARD, A. S,; VALENZUELA, C. R. Remote Sensing and Geographical Information Systems for Resource Management in Developing Countries). Euro Courses: Remote Sensing, vol 1. Kluwer Academic Publishers, the Netherlands, 1991. Chapter 18, 335-348.

WISCHMEIER, W. H. e SMITH, D. D. Predicting Rainfall Erosion Losses – a Guide to Conservation Planning. USDA/ARS, Washington, D. C., 1978. 58p. (Agricultural Handbook nº 537).

ZACHAR, D. Soil Erosion (Developments in Soil Science, n.o. 10), Publishers Elsevier/Veda, Amsterdam, 548p. 1982.