CIAT

Centro Internacional de Agricultura Tropical
International Center for Tropical Agriculture

An ArcView® toolkit

for cross-scale raster image analysis

Andy Nelson, Grégoire Leclerc, Ron Knapp.
July 1999

CGIAR Consultative Group on International Agricultural Research

This toolkit was built as a contribution to Activities 2.1, 2.2 amd 2.9 of the PE3 project Methodologies for integrating data across geographic scales in a data rich environment: Examples from Honduras. Pis: Ron Knapp and Grégoire Leclerc. These activites aim to:

Activity 2.1

Compare and contrast a subset of socio-economic and biophysical variables from output 1 in terms of frequency distribution, redundancy, noise and extreme values at different scales of aggregation and disaggregation.

Activity 2.2

Perform data reduction and low dimensional representation of multiple scales for site sampling and hypothesis generation.

Activity 2.9

Test significance of spatial structure in hillsides agroecosystems characterization.

The final report for this project contains the background, justification and results for these outputs. This document describes the spatial analysis techniques that can be applied with this toolkit, and some example applications.

The interface

The cross-scale models can be accessed through ArcView® GIS, a graphical user interface from ESRI. The toolkit is packaged as an *extension* that can be pluged-in to the GIS software. Figure 1 is the raster analysis interface runnin g in unix. It is started from the ArcView® View menu.

Although the interface is written in Avenue® the scripting language of ArcView® the backbone of the model is written in ANSI C, and so the entire toolkit is extememly portable and has been tested on Windows 95, NT and Sun stations running Solaris 2.5.1. The models were written in C for speed, since the ArcView® primatives for raster data manipulation are painfully slow.

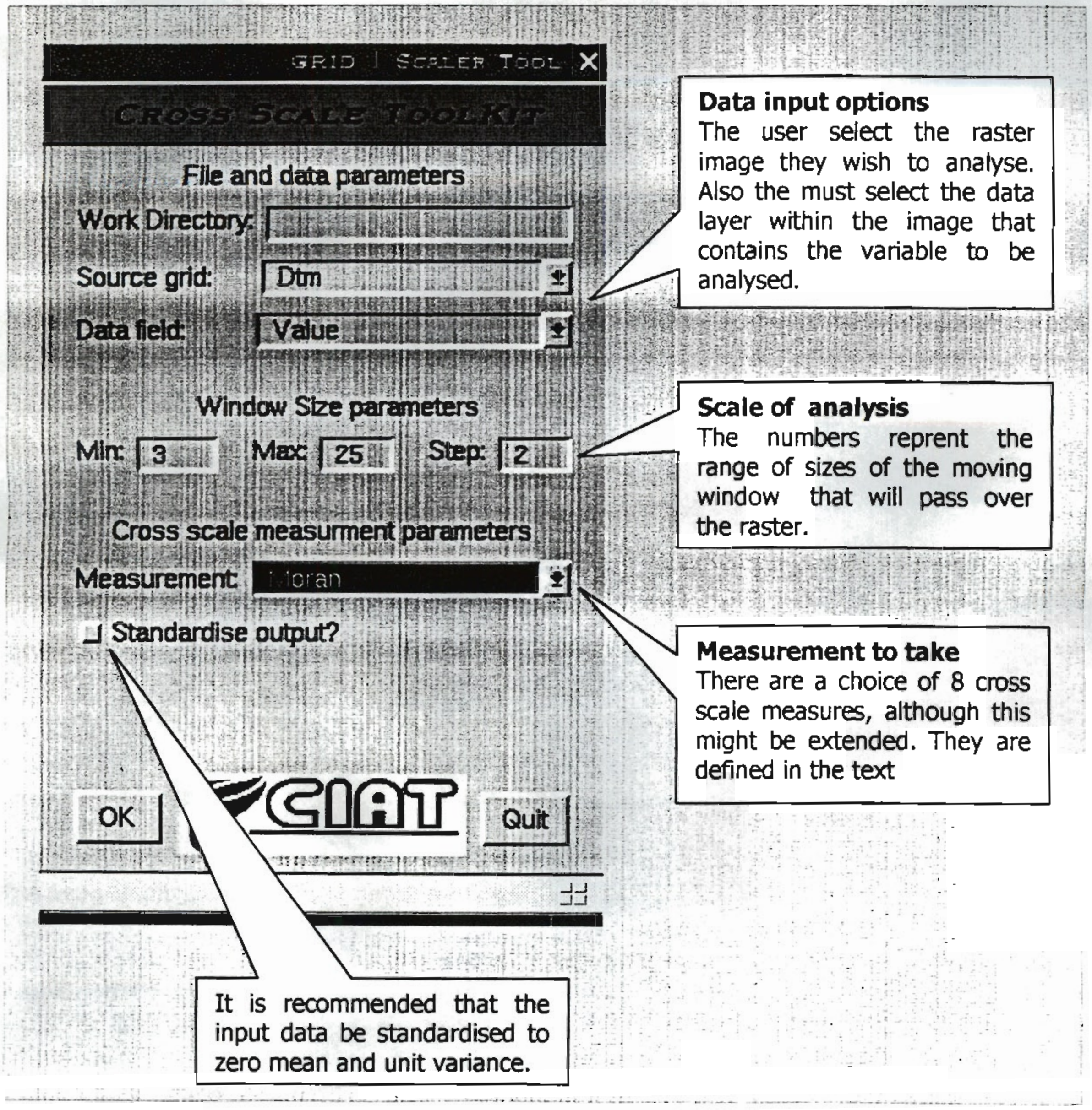


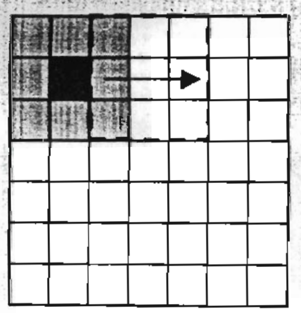
Figure 1. *The raster cross-scale interface.*

The concept of a spatial neighbourhood

The techniques for characterisation of rasters are all constrained by the resolution of the image. The information derived using these techniques is relevant only to the scale implied by the resolution of the image. Since this scale is often arbitrarily defined and not necessarily related to the scale of characterisation required, derived results may not always be appropriate. Indeed, most of the acknowledged problems of characterisation result from either variation at a finer scale than the image ('noisy data') or variations on a coarser scale ('flat regions'). What is required are characterisation techniques that are somewhat independent of the resolution of the database used to store topographic information. From a spatial perspective, it would seem ludicrous to only consider surface variation at a fixed scale when an assessment of an entire landscape is desired. Our own judgements both scientifically and 'intuitively' rely on an appreciation of landscape at a variety of scales simultaneously. Variation that occurs with scale is in itself a useful landscape diagnostic. It is for these reasons, that this tool was produced, to consider how scale can be incorporated into the automatic characterisation of raster images.

Moving window analysis

The neighbourhood is defined by the size (in cells) of the moving window centered on the selected cell. All cells within the window are considered. The greater the size of the window the greater the number of cells considered. The window floats over the raster centering on each cell in turn, Figure 2.



The black cell is the centre cell of the moving window, with the gray cells representing the extent of the window. Each cell in the window contributes to the measuement being taken. The window moves one cell to the left, with the new window extents indicated by a dotted line and another measurement is taken. This continues until each cell has been considered. Edge effects are taken into account at corners and limits of the raster image

Figure 2. *The moving window analysis*

The window can be square or circular in shape, although square is slightly more computaitionally efficient. Also distance weighting can be included so that the window becomes a spatially sensitive filter, in that ccells further away form the centre cell contribute less to the measure beinbg taken. This distance weighting is an appreciation of Toblers first law of geography

> ***Everything is related to everything else, but close things are more related than those farther away.***

The cross-scale functions

The six functions that are curently in the model are described in this section

Local Mean / Mode

The output is assigned the mean or modal value of it's local spatial neighbourhood

$$P_{mean} = \frac{1}{n}\sum_{i=0}^{n} P_i$$

Local Spatial Lag

The output is assigned the spatial lag value of it's local spatial neighbourhood

$$P_{lag} = P_0 - \left(\frac{1}{n}\sum_{i=1}^{n} P_i \right)$$

Local Variance / Standard Deviation

The output is assigned the variance or standard deviation value of it's local spatial neighbourhood

$$P_{var} = \frac{1}{n}\sum_{i=0}^{n} (P_i - P_{mean})^2$$

$$P_{\sigma} = \sqrt{\frac{1}{n}\sum_{i=0}^{n} (P_i - P_{mean})^2}$$

Local Semivariance

The output is assigned the semi variance value of it's local spatial neighbourhood

$$P_{\gamma} = \frac{1}{2n}\sum_{i=1}^{n} (P_i - P_0)^2$$

Local Moran index of spatial autocorrelation

The output is assigned the moran index value of it's local spatial neighbourhood

$$P_I = \frac{1}{n \cdot P_{\text{var}}} \sum_{i=1}^{n} (P_0 - P_{mean})(P_i - P_{mean})$$

Local morphometric feature extraction

The output is assigned 1 of 6 possible morphometric features based on it's local spatial neighbourhood. The most widely used set of morphometric characteristics, is the subdivision of all points on a surface into one of pits, peaks, channels, ridges, passes and planes. The names of these features suggest a geomorphological interpretation, but they may be unambiguously described in terms of rates of change of three orthogonal components (see Table). Note that the components x and y are not necessarily parallel to the axes of the lattice, but are in the direction of maximum and minimum profile convexity.

Feature	*Derivative Expression*	*Description*
Peak	$\frac{\delta^2 z}{\delta x^2} > 0, \frac{\delta^2 z}{\delta y^2} > 0$	Point that lies on a local convexity in all directions (all neighbours lower).
Ridge	$\frac{\delta^2 z}{\delta x^2} > 0, \frac{\delta^2 z}{\delta y^2} = 0$	Point that lies on a local convexity that is orthogonal to a line with no convexity/concavity.
Pass	$\frac{\delta^2 z}{\delta x^2} > 0, \frac{\delta^2 z}{\delta y^2} < 0$	Point that lies on a local convexity that is orthogonal to a local concavity.
Plane	$\frac{\delta^2 z}{\delta x^2} = 0, \frac{\delta^2 z}{\delta y^2} = 0$	Points that do not lie on any surface concavity or convexity
Channel	$\frac{\delta^2 z}{\delta x^2} < 0, \frac{\delta^2 z}{\delta y^2} = 0$	Point that lies in a local concavity that is orthogonal to a line with no concavity/convexity.
Pit	$\frac{\delta^2 z}{\delta x^2} < 0, \frac{\delta^2 z}{\delta y^2} < 0$	Point that lies in a local concavity in all directions (all neighbours higher).

Standardising local values

To standardise a data set, i.e. transform it so that it has a mean value of 0 and a standard deviation of 1, we apply the following equation to all values of P.

$$P_i' = \frac{P_i - P_{global\_mean}}{P_{global\_\sigma}}$$

Further visualisation and analysis

Summary grids

When the model has been run, a new grid will be created for each window size. This can be a large number of grids, maybe 10 or more. It is possible to make two summary grids of all this information, a mean/median grid and a standard deviation/entropy grid.

There are likely to be contexts when a single measure or classification is required, so it is useful to at least quantify the variability of measurements. For (the ratio scale) parameters, the mean and standard deviation are calculated. For the categorical feature classification, the mode and scaled entropy are calculated,

$$E = -\sum_{i=1}^{n1} p(f_i).\log[p(f_i)]$$

$$E_{mean} = -\log(1/n)$$

$$E_s = \frac{E}{E_{mean}}$$

where $p(f_i)$ is the proportion of measurements classified as feature type i and n is the number of feature types (6 in this study).

This may be used to distinguish locations that are consistently classified as the same feature ($Es = 0$) from locations that have a high degree of scale dependency in their classification ($Es = 1$). While this form of interrogation is useful in an interactive visualisation context, it is difficult to identify the spatial pattern of entropy across the image. To visualise both the spatial distribution and scale dependency of measures it is necessary to create several raster layers containing morphometric feature classification at a range of scales. Using simple map algebra, these maps can be combined to produce two new images, one of the modal classification, the other of entropy Thus it is possible to produce a feature membership map and a classification uncertainty map. These may be combined into a single hue-intensity image.

Data interogation

When the grids are displayed in the view, a further tool is available to interogate the variable in one location across scale. By clicking on the grid, a graph appears, charting the change in the variable or its corresponding measure across scale.

Examples

The Yorito watershed, Yoro, Honduras:

10m resolution Digital Terrain Model

Area 112km^2

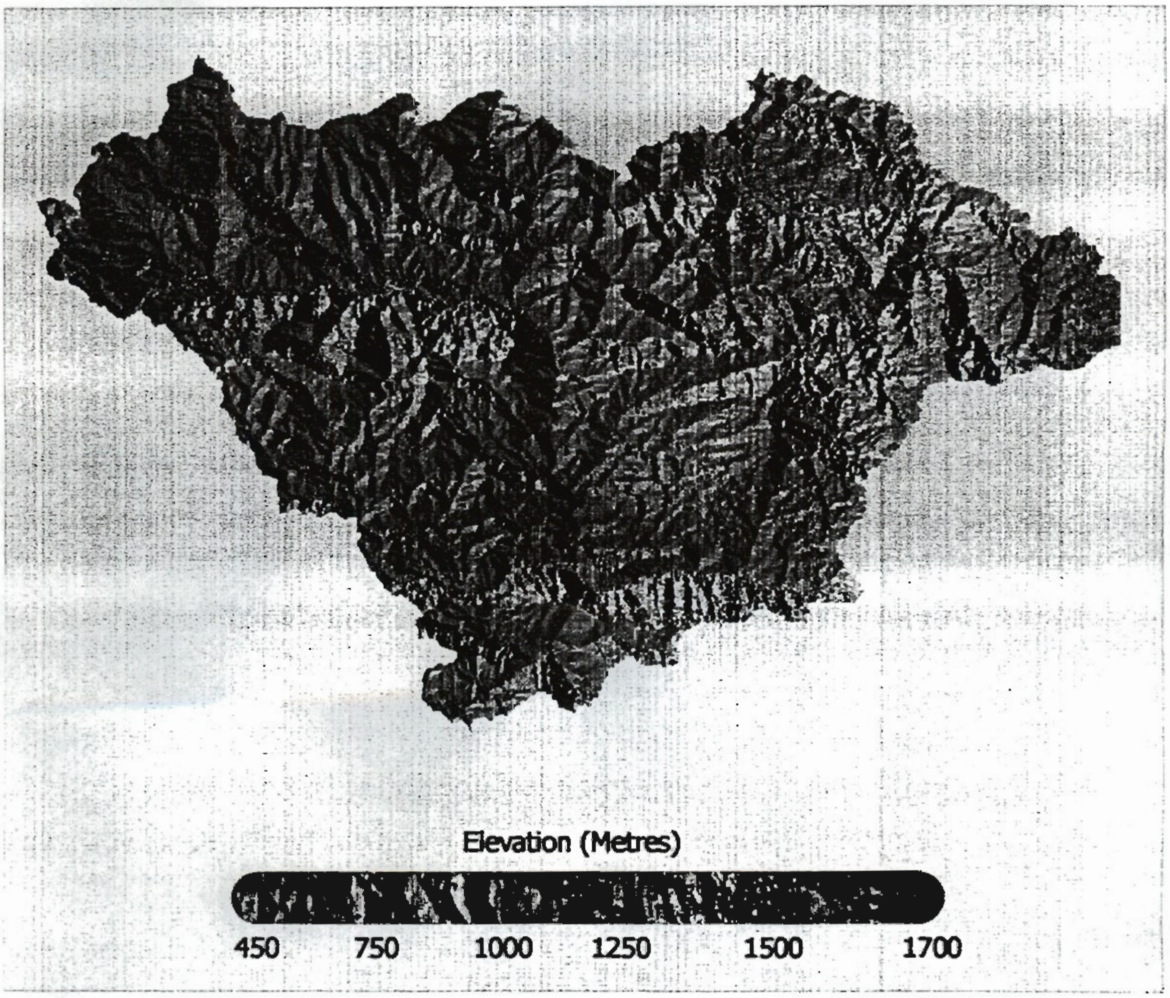
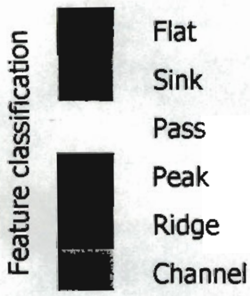


Figure 3 Shaded relief image of the 10m resolution digital terrain model of the Yorito watershed

Figure 4. Feature extraction with a 50m filter



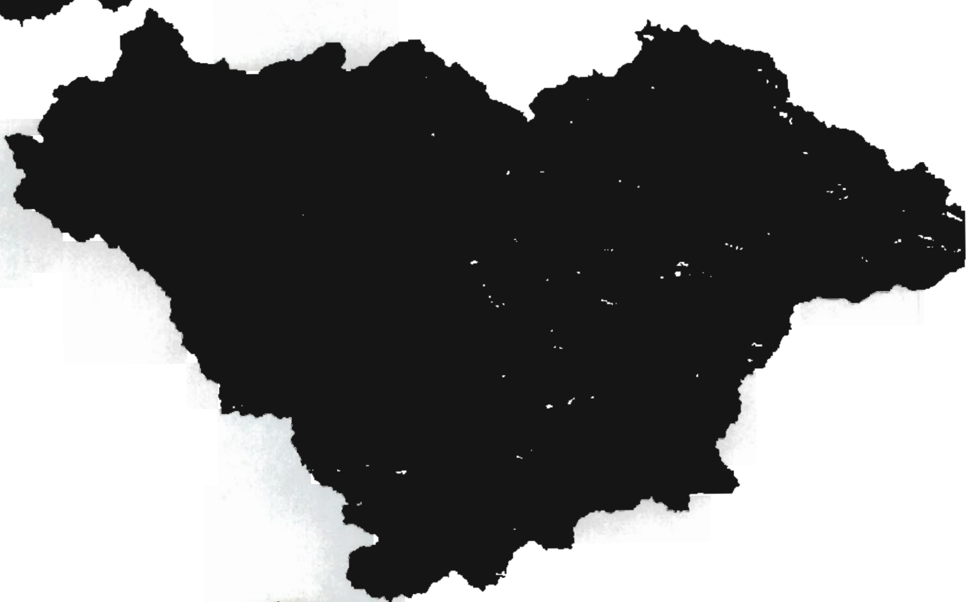


Figure 5 Feature extraction with a 200m filter

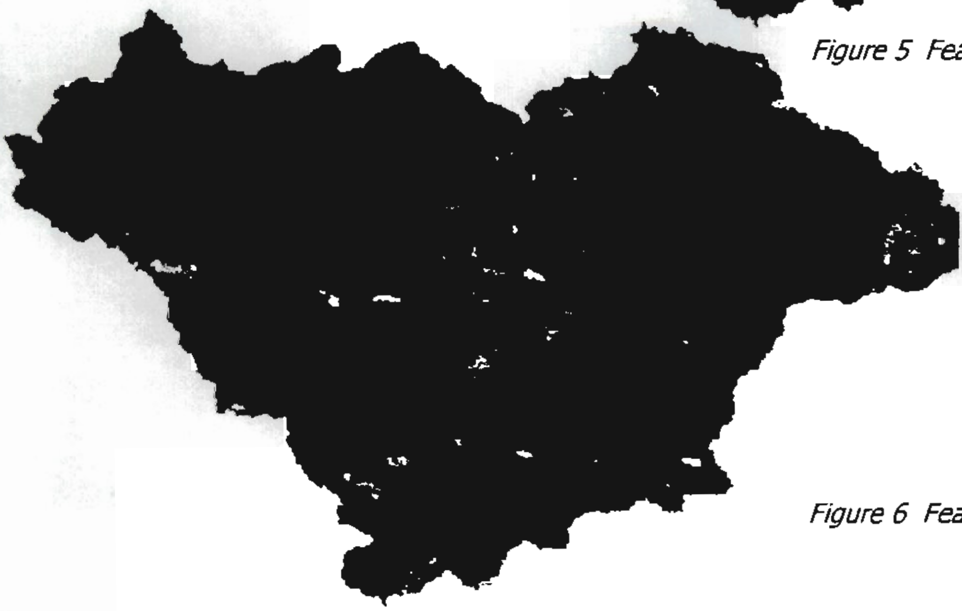


Figure 6 Feature extraction with a 500m filter

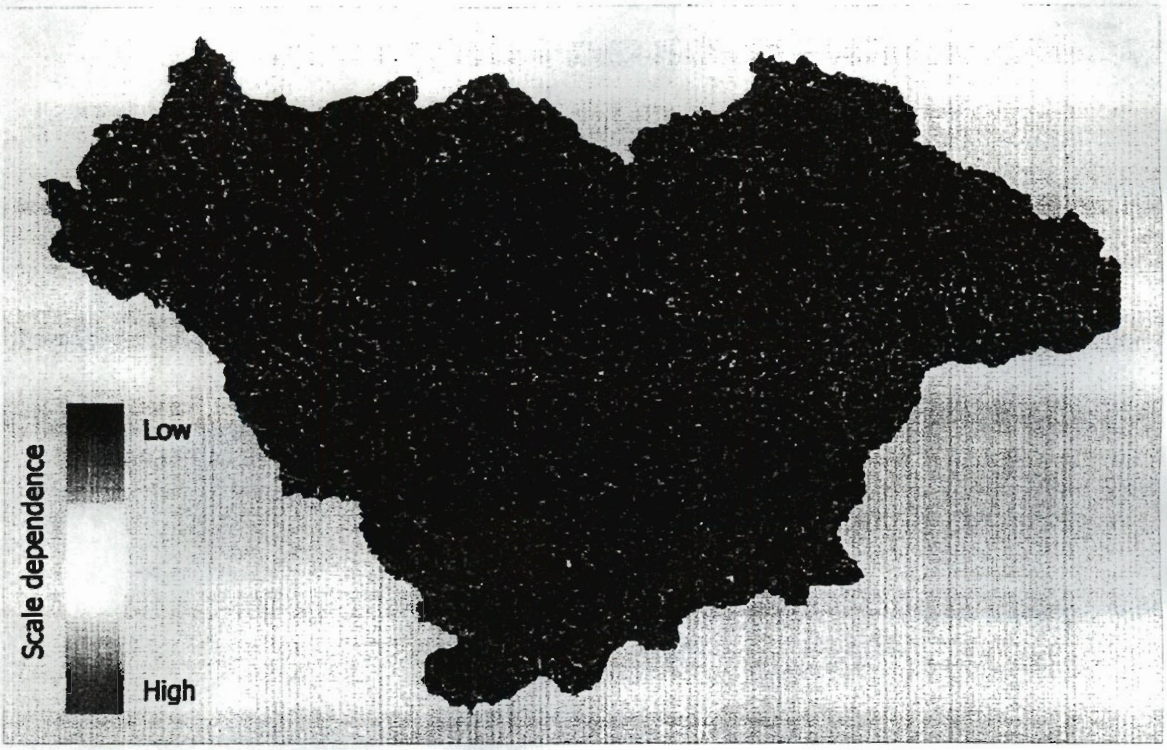


Figure 7. Entropy of the feature classification across scale. Dark blue colurs indicate scale independnce, where as Red indicates that the feature classicfiation is likely to change with scale.

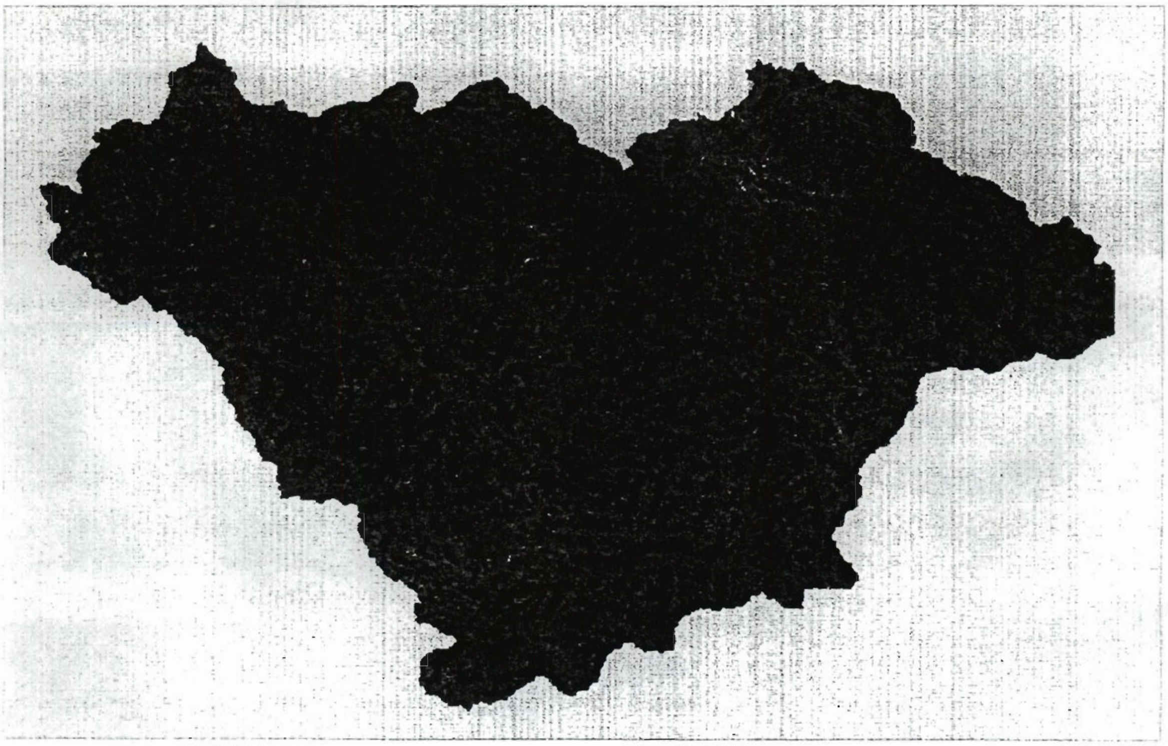


Figure 8. Hue Intensity blend of entropy and modal classification. This shows the most likely feature classicfication, by colour and the confidence of the classification by intensity, the brighter the colur the lower the confidence.