Image Anal Stereol 2001;20(Suppl 1):590-595

METHODOLOGICAL APPROACH FOR HYDROLOGICAL LANDSCAPE DEFINITION: APPLICATION OF TEXTURE MEASURES ON WEST AFRICAN WATERSHEDS

FRANCOIS DELCLAUX[1] AND CHRISTIAN DEPRAETERE[2]
[1]UMR HydroSciences, Maison des Sciences de l'Eau - IRD, BP 5045, 34032 Montpellier Cedex, France, [2]IRD Bénin, 01 BP 4414, Recette Principale, Cotonou, Benin
e-mail: delclaux@mpl.ird.fr, ird@bow.intnet.bj

ABSTRACT

A method for estimating the homogeneity level of hydrological spatial units is presented here. The approach is based on (i) the quantification of homogeneity by using specific indices, (ii) the influence of the spatial data error on homogeneity. Texture metrics, based on the co-occurrence probability matrix and dedicated to homogeneity quantification, are investigated. The sensitivity of these texture metrics, considered as homogeneity indices, is assessed in two types of landscapes: virtual landscapes characterised by elevation, and watersheds characterised by the relief curvature. The calculation of homogeneity indices reveals considerable variation in the way the spatial homogeneity is perceived. Three groups of indices are then formed by applying the principal components method and clustering analysis, each group being characterised by one index: contagion, contrast and Angular Second Moment or Inverse Difference Moment. In the second part of this paper, the influence of the data quality is analysed. The spatial error propagation is modelled using a stochastic simulation in the watershed dataset framework: a set of random elevation surfaces is generated using a Monte-Carlo method from which the corresponding curvature fields are derived. Next, homogeneity index samples are obtained for each watershed. The result shows a global decrease in homogeneity depending on the value of the elevation error and the type of index.

Keywords: error propagation, homogeneity, landscape, relief, texture, watershed.

INTRODUCTION

By observing the environment, the hydrologist can determine the relevant processes involved. This qualitative approach, completed by data acquisition and modelling, will help to understand the behaviour of the hydrological system in response to external variables. As information is rarely available on every point of the study area, simplifying assumptions are often made about spatial variability: the landscape is considered as an assemblage of homogeneous sub-units with respect to their description and hydrological dynamics (Esteves, 1994; Flugel, 1994). Nevertheless, these approaches face two main difficulties : the validation of homogeneity assumption and the effect of the spatial error.

The first point we consider is the question of homogeneity validation, and consequently, homogeneity quantification. This problem has been addressed by landscape ecologists who are concerned with relationships between ecological processes and spatial patterns (Musick and Grover, 1991; Riitters *et al.*, 1995). A landscape can be described in two ways (Sanderson and Ustin, 1998). First, pattern-oriented indices consider a landscape as an organised structure composed of patches and describe their spatial and statistical. The second way is concerned with pixel-oriented landscapes for which no hypothesis is made about specific structures. In this case, a wide variety of metrics based on pixel interrelation have been developed: some of them are directly concerned with texture and are considered as homogeneity quantifiers, i.e. homogeneity indices.

We used the programs developed by Baker and Cai (1992) for several reasons: (i) they operate on nominal as well as on numerical raster data, (ii) they focus on several different texture metrics, (iii) they are implemented in the GRASS GIS (Byars *et al.*, 1999) which handles our raster maps. These calculations involve processing a co-occurrence probability matrix, p(i,j), where p(i,j) is the probability for a pixel of attribute i to be adjacent to a pixel of attribute j. The dimensions of the matrix are (MxM) where M is the total number of attributes in the concerned zone. Six texture metrics were identified:

Angular Second Moment:	$ASM = \Sigma_i \Sigma_j\ p(i,j)^2$	(1)
Inverse Difference Moment:	$IDM = \Sigma_i \Sigma_j\ [\ p(i,j) / (1+(i-j)^2)\ ]$	(2)
Contrast:	$CON = \Sigma_i \Sigma_j\ [\ (i-j)^2\ p(i,j)\ ]$	(3)
Entropy:	$ENT = - \Sigma_i \Sigma_j\ [\ p(i,j)\ \ln(p(i,j))\ ]$	(4)
Contagion:	$CONTA = 2 \ln(M) - ENT$	(5)
Shannon:	$SHA = ENT/2 \ln(M)$	(6)

ASM and IDM are normalised homogeneity measures: the higher the values, the more homogeneous the landscape. On the other hand, contrast dramatically decreases with homogeneity. Concerning entropy, it increases with heterogeneity and is maximal when all the pixels i are as far apart from one another as possible. Shannon is the normalised value of entropy and contagion, which is the deviation between maximum entropy and entropy, increases with the tendency of a landscape to be clustered in large patches. Furthermore, we enlarged this review by adding three non-spatial indices: STDDEV (standard deviation), M, and MS (M divided by the surface logarithm). MS is a priori introduced insofar as M varies with respect to large surface changes.

The second purpose of our study is concerned with spatial error propagation. The raster maps stored in a spatial database have been acquired, classified, estimated, and in all these steps, errors on pixel values are introduced. Consequently, each spatial operation between maps, such as derivation, logical or numerical calculation and neighbourhood analysis, generates and propagates errors to the output. For example, a digital elevation model (DEM), which stores elevation data in the form of an altitude grid, can be used to derive a slope map. In this case, an error associated to the altitude will be propagated by the slope operator. Except for a few cases where the spatial operator is quite simple, Heuvelink (1999) discussed two alternative methods for modelling error propagation: (i) the Taylor series method when the spatial operator can be linearised in the form of Taylor series: the statistical properties of output are then estimated by an analytical expression, (ii) Monte-Carlo method, in which mean and standard deviation of output are calculated from a set of input random fields. Due to the complexity of the spatial operators involved in our study, we applied the Monte Carlo method.

MATERIAL AND METHODS

HOMOGENEITY INDEX

The first dataset consists of virtual landscapes characterised by elevation raster maps with identical statistical properties (mean = 100, standard deviation = 20). Three kinds of landscapes are created: (i) linear landscapes, i.e. made up of homogeneous areas (cf. Fig. 1a), (ii) disturbed linear landscapes which are the previous ones disturbed by a Gaussian noise (cf. Fig. 1b), (iii) undulating landscapes where elevation is a sinusoidal function. The homogeneity indices are calculated for these configurations. They are compared with each other and their behaviour is studied with respect to the landscape structures.

The second dataset is made up of 314 West African watersheds. The study area coordinates are [18W:4N-18E:21N]. A raster map is composed of [2040 × 4320] squared pixels whose size is set to 1km^2. The pixel value is the Mean Quadratic Curvature (MQC) which quantifies the slope changes in the neighbourhood of a point. This relief descriptor is important for hydrologists since it is a conditioning factor for hydrological processes and it is the result of the process itself, thus reflecting the strong interrelation between water and relief. Mathematically, MQC is equivalent to the 2nd derivative of altitude. The computational algorithm is based on a derivation of altitude using 2nd-order Taylor series on a [3 × 3] pixels window (Depraetere, 2000). The nine homogeneity indices of MQC are then computed for each watershed. They are normalised with an interpolation between their extreme values and range from 0 to 100 as a function of increasing heterogeneity. Next, the watersheds are classified and compared according five categories, ranging from *very homogeneous* to *very heterogeneous*. Afterwards, the whole set of indices is divided into three groups by using a principal components analysis and the hierarchical clustering method.

(a) linear and *(b) disturbed linear landscapes* *(c) MQC landscape*

Fig. 1. *1a. and 1b. Examples of virtual elevation landscapes. 1c. MQC raster map. The pixel values range from 0 to 1080: high values (dark) are associated, either with a natural high curvature (for example: mount Cameroon) or with low altitude error (further comments are in Fig. 4.). The black lines delineate the watersheds.*

ERROR PROPAGATION

First, a Monte-Carlo stochastic simulation is applied for generating a set of random DEM and MQC fields according the following scheme:

$$Z(x,y) = Z_0(x,y) + \varepsilon(x,y) \Rightarrow MQC(x,y),$$

where Z, ε, MQC are random elevation, error and curvature fields, and Z_0 is the structural component of Z. Theoretically, the spatial autocorrelative characteristics of $\varepsilon(x,y)$ are necessary to generate realisations of the elevation surface. Unfortunately, the only available information is a map of elevation Root Mean Square Error (RMSE). We then simulated DEM and MQC realisations in a simplified way: (i) assigning disturbed elevations to a set of randomly located points: the elevation disturbance distribution is Gaussian with a mean of zero and a standard deviation equivalent to RMSE, (ii) generating DEM with interpolation using inverse distance squared weighting algorithm, (iii) computing MQC field from DEM.

Second, one sample of each homogeneity index is obtained for each watershed from which statistical characteristics are extracted. Due to the size of one map (20 Mb) and the computation time required for each run, the simulation number was set to 100.

RESULTS AND DISCUSSION

HOMOGENEITY INDEX

Some of the results for virtual landscapes have already been discussed (Delclaux *et al.*, 2000): (i) ASM is more sensitive than IDM to spatial heterogeneity; if the landscape is disturbed by a Gaussian noise, they fall to zero; (ii) contrast sharply increases in the presence of large local variations; (iii) Shannon, entropy and contagion are more sensitive to a global homogeneity rather to the local gradients.

Concerning the watershed dataset, Fig. 2 shows the difficulty in estimating the level of homogeneity of a basin. For example, the spatial distributions of watersheds obtained with contrast and ASM reveal two different geomorphological structures. Moreover, crossing these eight maps reveals that only eighteen watersheds, i.e. 1.1% of the surface, belong to the same class (*very homogenous* class). Nevertheless, some visual trends can be observed: IDM, Shannon and entropy indices give similar results. On the opposite, contrast and contagion distributions are quite different.

STDDEV *CON* *ASM*

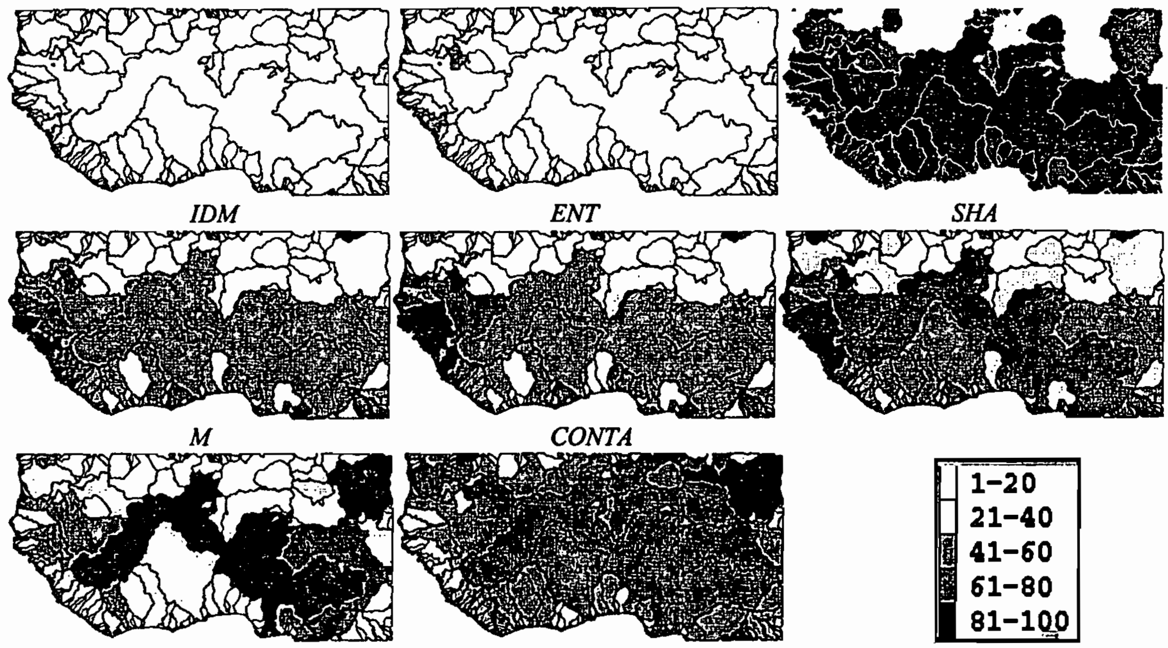

Fig. 2. *West African watershed classification in term of MQC homogeneity index. Indices are normalised, and the five classes range from very homogenous [1-20] to very heterogeneous [81-100]. Classified MS map is not represented as it is identical to M map.*

In order to reduce the number of indices, we processed the 314 index values through the principal components method applied to the correlation matrix: 84% of the variance is explained by the two first axes and the three first eigenvalues are equal to 5.22, 2.33 and 0.73. Consequently, we considered two components for which the variable correlation values are plotted in Fig. 3a. This result has been completed with a hierarchical clustering analysis (cf. Fig. 3b) which allows indices to be gathered in three groups, each one represented by one index: [1] contagion; [2] ASM for (ASM, IDM, Shannon, entropy); [3] contrast for (contrast, standard deviation, M, MS). In the last group, contrast has been retained as it is the only spatially dependant index. Lastly, it appeared that MS index did not contain more relevant information than M index.

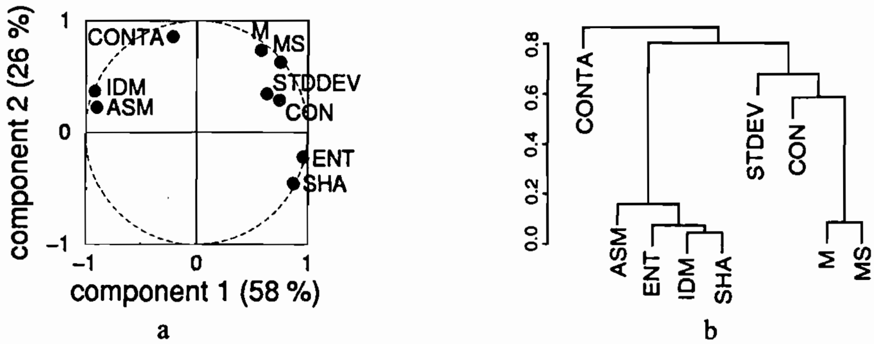

Fig. 3. *a: Index correlation values versus two first components. b: Resulting tree from hierarchical clustering analysis.*

ERROR PROPAGATION

The curvature map in Fig. 1c highlights the influence of altitude error as shown in Fig. 4. DEM error thus cannot be neglected and has to be taken into account in modelling MQC homogeneity. The assumption is that homogeneity is made up of two parts: the first one is natural as expressing the real world homogeneity and is quantified by the indices mentioned above; the second one is related to the data quality as decreasing function of the spatial error.

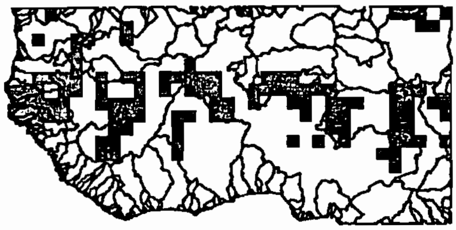

Fig. 4. *Spatial variability of elevation error: square degrees with 18m (black) or 97m (white) RMSE. A part of MQC high values (grey areas in Fig. 1c.) are due to low RMSE: the more accurate the elevation, the rougher the relief, and the higher the curvature.*

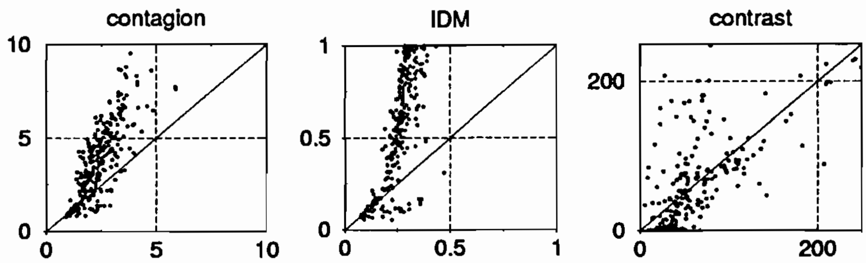

Fig. 5. *Mean of sample values (x axis) versus initial values (y axis) for contagion, IDM and contrast indices. Points above the line for contagion and IDM, and points below the line for contrast correspond to an increase in heterogeneity.*

Starting from the index samples generated with the Monte-Carlo simulation, the comparisons between the mean and original values are plotted on Fig. 5. They concern the three previous representative indices, except for ASM which has been replaced by IDM. The trend is an increase in heterogeneity for contagion and IDM. Concerning ASM, it appears to be very sensitive as all the values dramatically decrease to 0, corroborating the results obtained for the virtual landscape dataset. Moreover the behaviour of contrast is more complex since homogeneity falls for low values, and rises for high values: the same trend has been found for the indices belonging to group [3].

CONCLUSIONS

Hydrologists can calculate the watershed homogeneity for a given parameter, but they still need to select the most relevant quantifier of homogeneity. Our study presents a review of nine available textural and statistical indices and proposes a classification in three groups in order to reduce the number: the first group is represented by contagion and measures the level of aggregation. The second group is characterised by ASM or IDM for global homogeneity: IDM cannot be used for nominal data, but it is more convenient to low homogeneity than ASM. The third group is represented by contrast which is very sensitive to local gradients.

The second part of our study highlights the influence of data accuracy. A stochastic simulation shows that taking into account the spatial error in the original data leads to a global increase in heterogeneity. Nevertheless, the rise has to be modulated for some indices which are dedicated to local heterogeneity such as contrast. Finally, homogeneity indices seem to integrate both the real physical homogeneity and a pseudo-heterogeneity due to the inaccuracy of the data. The problem is now to model and to link these homogeneity components as a function of data error characteristics, statistical distribution of index values and mathematical properties of each index.

ACKNOWLEDGEMENTS

This work has been funded by the French *Programme National de Recherche en Hydrologie* (PNRH), grant #234. The authors also wish to thank Hélène Lubès, research engineer in IRD, for her constructive comments.

REFERENCES

Baker WL, Cai Y (1992). The r.le programs for multiscale analysis of landscape structure using the GRASS geographical information system. Landscape Ecology 7(4):291-302.

Byars B, Neteler M, Clamons S, Cherry S (1999). Geographic Resources Analysis and Support System GRASS 4.2 Fact Sheet. Baylor University. (http://www.geog.uni-hannover.de/grass).

Delclaux F, Depraetere C, Laborde JP (2000). Détermination de paysages hydrologiques de référence en Afrique de l'Ouest. Journées du Programme National de Recherche en Hydrologie, Toulouse.

Depraetere C (2000). Demiurge: Chaîne de Production et de Traitement de Modèles Numériques de Terrain. In: IRD, eds. Collection Logorstom.

Esteves M (1994). Cartographie d'unités hydrologiques homogènes et modélisation hydrologique, exemple de l'expérience Hapex-Sahel. In: IRD, eds. Montpellier: Actes des Xèmes Journées Hydrologiques, 463-73.

Flugel WA (1994). Delineating hydrological response units by GIS analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the river Brol, Germany. Hydrological Processes, 9:423-36.

Heuvelink GBM (1999). Propagation of error in spatial modelling with GIS. In: Longley PA, Goochild MF, Maguire DJ, Rhind DW, eds. Geographical Information Systems: principles and technical issues. Ch. 14. New York: Wiley & Sons, 207-17.

Musick HB, Grover HD (1991). Image textural measures as indices of landscape pattern. In: Turner MG, Gardner RH, eds. Quantitative methods in Landscape Ecology: The analysis and Interpretation of Landscape Heterogeneity. Ch. 4. Ecological Studies 82. New-York: Springer-Verlag, 77-103.

Riitters KH, O'Neill RV, Hunsaker CT, Wickham JD, Yankee DH, Timmins SP, Jones KB, Jackson BL (1995). A factor analysis of landscape pattern and structure metrics. Landscape Ecology 10(1):23-39.

Sanderson EW, Ustin SL (1998). Evaluation of landscape structure using AVIRIS quicklooks and ancillary data. Proc of the 7th Earth Sc Airborne Workshop, JPL, Pasadena.