DELCLAUX F. and DEPRAETERE C., 2001. Quantitative approach for the determination of hydrological landscapes. Proceedings of River Basin Manageme Conference, Cardiff, UK, WITPress, 283–292

Quantitative approach for the determination of hydrological landscapes

F. Delclaux[1] & C. Depraetere[2]
[1]*UMR HydroSciences Montpellier, IRD, France*
[2]*Hydrologie, IRD Cotonou, Bénin*

Abstract

Due to spatial heterogeneity of watersheds hydrologists segment the landscape in homogeneous zones. Then, two points are explored: the homogeneity validation and the role of data quality. The concept of homogeneous units is validated in terms of texture metrics. These indices are calculated with a co–occurence matrix whose elements define the probability of pixel adjacency within a raster map. Data and computational tools are handled by the GRASS GIS. First, the index sensitivity is tested with two types of data: (1) linear and sinusoidal virtual landscapes, (2) 314 West African watersheds characterised by the Mean Quadratic Curvature parameter of the relief. The inter–comparison shows a great dispersion in the heterogeneity description. Moreover, the study highlights the correlation between metrics through a principal components method and clustering analysis: the nine texture metrics are gathered in three sets, each one represented by one index: contagion, IDM, contrast. The second part of our work is dedicated to the spatial error propagation and its consequence on a watershed classification. A stochastic Monte–Carlo simulation applied on watershed maps reveals an increasing of heterogeneity.

1 Introduction

On one hand, more and more spatial data are involved in distributed hydrological modelling. But capturing and evaluating spatial heterogeneity is not obvious to achieve. On the other hand, the increasing use of Geographic Information System (GIS) has promoted handling of spatial databases and development of specific techniques. Consequently, hydrologists can easily segment the landscape into homogeneous units according to scale, process and

land property criteria [1], [2], [3], [4], [5]. Nevertheless, as the homogeneous level of delineated units is not validated as such, it is necessary to address this point in terms of homogeneity quantification.

The second point our work focuses on is error propagation: as handling spatial data always introduces error, the question is to assess the sensitivity of the homogeneity model to the data uncertainty.

1.1 Homogeneity quantification

Quantifying homogeneity in landscapes has already been addressed by ecologists. As they are concerned by relationships between spatial patterns and ecological dynamics, they developed a large variety of metrics describing the landscape spatial characteristics [6], [7], [8], [9]. Among them, the texture indices are dedicated to homogeneity quantification. They are based on the processing of a co–occurrence probability matrix p(i,j), with the dimensions (mxm) where m is the total number of attributes in the zone concerned. p(i,j) is the probability the pixel of attribute i is adjacent to the pixel of attribute j. The texture metrics are summarised in Table 1 according to the definitions given by Baker [10]. We added three no–spatial heterogeneity indicators: standard deviation, M and MS.

Table 1: Heterogeneity indices. p(i,j) is the co–occurence probability matrix. i, j and k are pixel attributes: i and j range between 1 and m, with m attribute number. k ranges between 1 and n with n total number of pixels. S is the surface of the study area.

Heterogeneity index (HI)	Symbol	Definition
Angular Second Moment	ASM	$\Sigma_i \Sigma_j \; p(i,j)^2$
Inverse Difference Moment	IDM	$\Sigma_i \Sigma_j \; [\, p(i,j) / (1+(i-j)^2) \,]$
Contrast	CON	$\Sigma_i \Sigma_j \; [\, (i-j)^2 \; p(i,j) \,]$
Entropy	ENT	$- \Sigma_i \Sigma_j \; [\, p(i,j) \; \ln(p(i,j)) \,]$
Contagion	CONTA	$2 \ln(m) - ENT$
Shannon	SHA	$ENT/2 \ln(m)$
Standard deviation	STDDEV	$[\, \Sigma_\kappa (k - ((\Sigma_\kappa k)/n)/ (n-1) \,]^{1/2}$
Number of attributes	M	m
Area related number of attributes	MS	$m / \ln(S)$

1.2 Spatial error propagation

Except few cases where the spatial operator is quite simple, Heuvelink [11] discussed two alternative methods for describing error propagation: (1) Taylor series method when the spatial operator can be linearised in the form of Taylor

series, (2) Monte–Carlo simulation, in which mean and standard deviation of outputs are computed 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.

2 Data and methods

2.1 Homogeneity index

First, a set of virtual landscapes has been realised for evaluating index sensitivity. These maps are characterised by: (1) one North–South variable, (2) simple shapes including linear, disturbed linear and sinusoidal structures, (3) 720x720 pixel size, (4) identical spatial statistical properties of the variable: average=100 and standard deviation=20. Indices are calculated for these configurations. They are then compared together and their change is studied with respect to these landscape structures.

The second dataset is made up of watersheds. 314 basins, for which the surface ranges from 10^2 to 1.5 10^6 km^2, have been extracted from the USGS 30 second DEM (GTOPO30) within a [18W:4N–18E:21N] geographic window. A relief descriptor parameter, the Mean Quadratic Curvature (MQC), has been selected for characterising the watersheds. MQC, which quantifies the slope gradients in the neighbourhood of a point, is important for hydrologists as it is a conditioning factor for water processes and the result of the process action itself: it thus reflects the strong interdependence between water and relief. The computational algorithm is based on a derivation of altitude using 2nd–order Taylor series according to Depraetere [12]. The indices are normalised, and the watersheds are classified and compared according three categories, ranging from *homogeneous* to *heterogeneous*. Afterwards, the whole set of metrics is divided into three groups through a principal component analysis and hierarchical clustering method, each group being characterised by the more representative index.

2.2 Error propagation

A Monte–Carlo stochastic simulation is applied to a set of random DEM fields according the following scheme:

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

where Z, ε, MQC are random fields of elevation, error and MQC, HI random variable of a heterogeneity index relative to the ith watershed. Z_0 is the structural component of Z.

The characteristics of the random field $\varepsilon(x,y)$ are necessary for generating realisations of the elevation surface [13]. Unfortunately, the only available information given in the GTOPO30 package is a map of elevation Root Mean Square Error (RMSE) for each square degree. DEM realisations are thus completed in a simplified way in order to match the original mean Z_0 , the elevation RMSE and relief curvature. This method runs as follows: first,

disturbed elevations are assigned to a set of spatial random points; second, random DEM is generated by interpolating with inverse distance squared weighting algorithm; then, MQC fields are computed. Finally, one sample of each texture index per watershed is obtained from which statistical characteristics are extracted. Due to size of one map (20 Mbytes) and the overall computational time, the simulation number has been limited to 100.

All the data and tools are managed in the GRASS Public Domain GIS.

3 Results and discussion

3.1 Homogeneity index for virtual landscapes

A representation of linear virtual landscapes is given in Fig. 1. The profile lines are made up of one or several plates. Furthermore, adding perturbation to normal landscapes allows high local heterogeneity to be simulated.

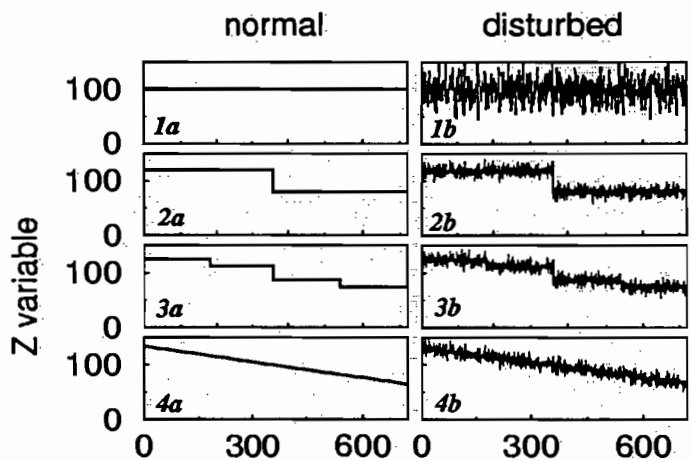

Figure 1: Profile lines of Z variable for normal (1a–4a) and disturbed (1b–4b) landscapes. Spatial average and standard deviation (except for 1a) are set to 100 and 20. Disturbed maps are generated using a gaussian noise.

As shown in Fig. 2, heterogeneity measures are quite different depending on the landscape structure. IDM and ASM are similar, but ASM is more sensitive than IDM to global heterogeneity increase. The two fall to zero in the case of disturbed landscapes. Concerning contrast (CON), it sharply increases in case of large local gradients. In general, Entropy (ENT), Shannon (SHA) and contagion (CONTA) are not sensitive to local heterogeneity.

Lastly, the results obtained in the case of sinusoidal landscapes agree with the above observed trends.

3.2 Homogeneity index for watershed landscapes

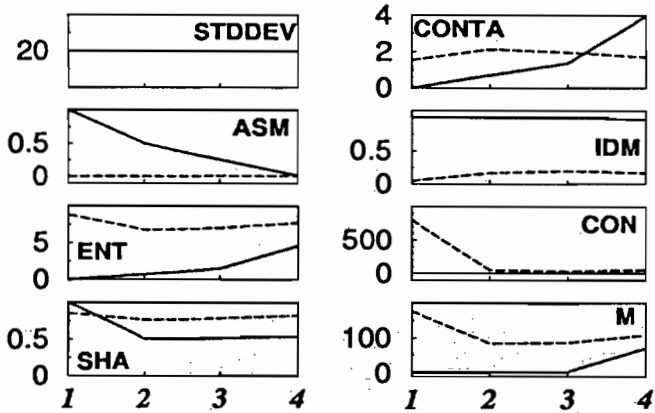

Figure 2: Texture index evolution with virtual landscapes. Landscape numbers (1–4 according to Fig. 1) are on x axis. Solid lines show index evolution for normal landscapes and dashed lines for disturbed landscapes.

The variety of heterogeneity descriptors is illustrated by referring to Table 2, where it appears that the index distributions are very different, in terms of either number of watersheds, either watershed surface.

Table 2: Classification of watershed heterogeneity in homogeneous, medium and heterogeneous classes. Each class is characterised by its cumulative (in million km^2) and percentage (%) surface and by its number of watersheds (n w).

	homogeneous		medium		heterogeneous	
	S (%S)	n w	S (%S)	n w	S (%S)	n w
STDDEV	5.86 (93.7)	293	0.39 (6.2)	16	0.01 (0.1)	5
CONTA	3.77 (60.3)	19	2.15 (34.4)	121	0.33 (5.3)	174
ASM	0.89 (14.2)	80	1.04 (16.7)	47	4.32 (69.1)	187
IDM	2.35 (37.6)	158	3.59 (57.5)	71	0.31 (4.9)	85
ENT	1.67 (26.6)	129	2.12 (34)	83	2.46 (39.4)	102
CONT	5.72 (91.5)	296	0.53 (8.5)	17	0 (0)	1
SHA	1.54 (24.7)	96	4.2 (67.3)	100	0.5 (8)	118
M	2.4 (38.4)	290	1.42 (22.7)	20	2.43 (38.9)	4
MS	2.16 (34.6)	270	1.66 (26.5)	40	2.43 (38.9)	4

As it has been observed in the case of virtual landscapes, some metrics appeared to have similar properties. A principal component analysis has then been applied on the correlation matrix whose results are shown on Table 3. We

have retained the two first components as: (1) they explain 84% of the variance, (2) the third eigenvalue is less than 1. These results have been completed by a hierarchical clustering analysis which allows finally to describe the set of metrics with three groups represented by IDM, contagion and contrast.

Table 3: Results of principal components factor analysis.

	Component 1	Component 2	Component 3
eigenvalue	5.22	2.33	0.73
% variance	58	26	8
STDDEV	0.6313203	0.3455804	0.52931093
CONTA	–0.2122502	0.8603646	–0.30167281
ASM	–0.8950385	0.225571	0.25248168
IDM	–0.9144161	0.370243	0.07643413
ENT	0.9608923	–0.220767	–0.10987004
CON	0.7431288	0.2901959	0.4426902
SHA	0.8697442	–0.4581152	–0.14048888
M	0.5789931	0.7365266	–0.22031033
MS	0.7529322	0.6287034	–0.10090742

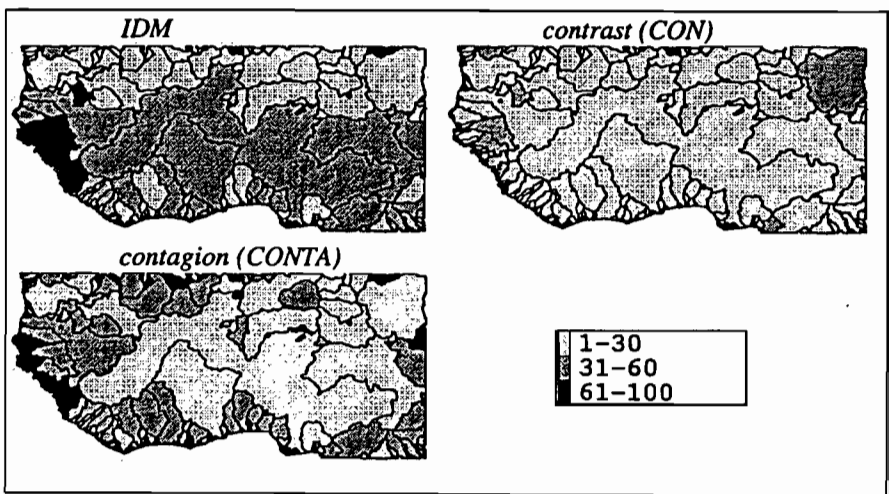

Figure 3: Classified maps of watershed heterogeneity indices. For each map, the Mean Quadratic Curvature texture index has been normalised between 1 and 100, and classified in three classes ranging from homogeneous [1–30] to heterogeneous [61–100].

Fig. 3 shows the corresponding classified maps for these 3 indices, and Table 4 summarises their intersection. In fact, according to the high values in intermediate column, few watersheds belong to the same classes. This is mainly due to the difference of heterogeneity underlying concepts: IDM characterises a global homogeneity, contrast deals with the local gradients whereas contagion expresses a patch clustering degree.

Table 4: Intersection of IDM, contagion and contrast maps. The three first columns contain cumulative (and percentage) surface and number relative to watershed belonging to the same classes. Last column concerns remaining watersheds.

	homogeneous	medium	heterogeneous	intermediate
S (%s)	1.06 (16.9)	0.04 (0.7)	0.001 (0.02)	5.15 (82.4)
n w	15	3	1	295

3.3 Error propagation for watershed landscapes

The effect of error clearly appears on Fig. 4a and 4b where the curvature map reproduces the low uncertainty elevation areas: the lower the error, the rougher the relief and the higher the curvature. It should be noted that these areas are not visible on the elevation map.

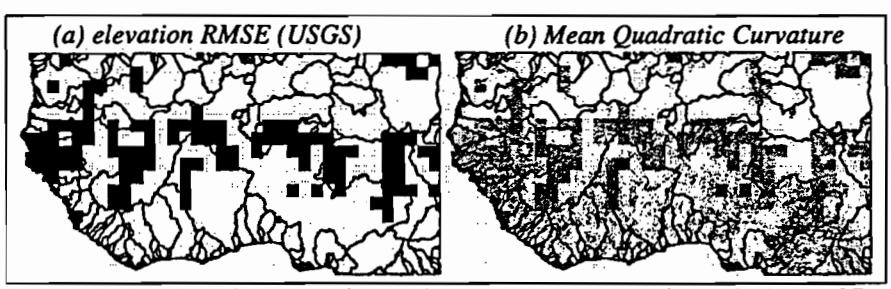

Figure 4: (a) elevation RMSE map by square degrees: 18m (black) or 97m (white) (b) MQC map: high values (dark) are associated, either with high natural curvature, either with low RMSE. Max. MQC value is 1080.

Starting from the RMSE map, a set of 100 random DEM and MQC fields have been simulated through a Monte–Carlo method. Then, for each watershed and texture index, a sample of 100 values have been obtained. Finally, the mean has been calculated and compared to the original values. As illustrated by Fig. 5 for the three previous selected indicators, the overall trend is an heterogeneity increase: IDM and entropy are more sensitive than contagion. This increase can be more or less important. For example, ASM homogeneity index dramatically decreases to 0. On the opposite, non–spatial metrics such as standard deviation, M and MS are less influenced by error. Actually, due to its

definition, MQC magnifies elevation perturbations. The co–occurence matrix then propagates theses errors, adding to the initial value a pseudo heterogeneity.

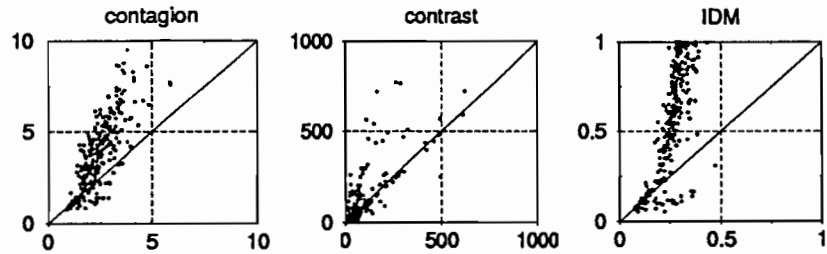

Figure 5: Mean of random values (x axis) versus initial values (y axis) for the 3 most significant indices. Each point represents one of the 314 watersheds.

4 Conclusion

This paper suggests to hydrologists a quantitative method for validating homogeneity assumptions with regards to their spatial data. The meaning and sensitivity of nine homogeneity indices have been assessed, first with virtual landscapes for which the spatial structure is known, second with African watersheds characterised by the relief curvature. Comparing these indices induces very different results in terms of homogeneity quantification. This is quite obvious as the surface and number of classified homogeneous watersheds are compared. Some criteria are then given to select the suitable index according to the study case: numerical or nominal data, local or global scale heterogeneity, sensitivity to perturbations. This qualitative approach has been completed by a principal component analysis which shows that two axes are satisfactory for explaining the variance: the nine metrics have been grouped in three classes, each one being represented by one index: IDM for global homogeneity, contagion for patch clustering level and contrast for local heterogeneity.

The second problem that our study addresses is the effect of data uncertainty. A stochastic Monte–Carlo model was implemented for generating elevation and curvature random fields, and corresponding index samples for each watershed. The final result showed that in the areas where the original elevation error was important, the heterogeneity values increase significantly independently of the index definition.

In the future, some points would need further developments: getting better index selection criteria, testing metrics with another watershed parameter, improving random elevation fields algorithm and modelling heterogeneity increase.

Acknowledgements

We gratefully acknowledge Mrs H. Lubès and Mr. E. Elguero, research engineers in IRD, for their constructive comments. This study has been funded by the French *Programme National de Recherche en Hydrologie* (PNRH), grant #234.

References

[1] Esteves, M., Cartographie d'unités hydrologiques homogènes et modélisation hydrologique: exemple de l'expérience Hapex–Sahel. *Actes des Xèmes Journées Hydrologiques*, eds. IRD, pp. 463–473, 1994.

[2] Viné, P., Apport de la télédétection à l'étude des zones contributives aux écoulements. Cas de la Mare d'Oursi. *Proc. of Int. Workshop on Remote Sensing and Water Resources*, eds. CEMAGREF/IRD/FAO, 1995.

[3] Loyer, J.Y., Moriaud, S. & Descroix, L., Unités de paysage pour l'hydrologie au Nord du Mexique. *Proc. of Int. Workshop on Remote Sensing and Water Resources*, eds. CEMAGREF/IRD/FAO, 1995.

[4] Flugel, W.A., 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**, pp. 423–436, 1994.

[5] Jeton, A.E. & Smith, J.L., Development of watershed models for two Sierra Nevada basins using a Geographic Information System. *Water Resource Bulletin*, **29(6)**, pp. 923–932, 1993.

[6] Lavers, C.P. & Haines–Young, R., Equilibrium landscapes and their aftermath: spatial heterogeneity and the role of new technology (Chapter 5). *Landscape Ecology and GIS*, ed. R Haines–Young, DR Green & S Cousins, Taylor & Francis: London, pp. 57–74, 1993.

[7] Delcros, P., *Ecologie du paysage et dynamique végétale post–culturale en zone de montagne*, eds. CEMAGREF, Collection Gestion des Territoires, 334p., 1994.

[8] Musick, H.B. & Grover, H.D., Image textural measures as indices of landscape pattern (Chapter 4). *Quantitative methods in Landscape Ecology: The analysis and Interpretation of Landscape Heterogeneity*, Ecological Studies 82, ed. M.G. Turner & R.H. Gardner, Springer–Verlag: New–York, pp. 77–103, 1991.

[9] Riitters, K.H., O'Neill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B. & Jackson B.L., A factor analysis of landscape pattern and structure metrics. *Landscape Ecology*, **10(1)**, pp. 23–39, 1995.

[10] Baker, W.L. & Cai, Y., The r.le programs for multiscale analysis of landscape structure using the GRASS geographical information system. *Landscape Ecology*, **7(4)**, pp. 291–302, 1992.

[11] Heuvelink, G.B.M., Propagation of error in spatial modelling with GIS. (Chapter 14). *Geographical Information Systems: principles and technical issues*, ed. P.A. Longley, M.F. Goochild, D.J. Maguire & D.W. Rhind, Wiley & Sons: New York, pp. 207–217, 1999.

[12] Depraetere, C., *Demiurge: Chaîne de Production et de Traitement de Modèles Numériques de Terrain*, eds. IRD, Collection Logorstom, 2000.
[13] Ehlschlager, C.R. & Shortridge, A., Modelling elevation uncertainty in geographical analysis. *Proc. of the Int. Symp. on Spatial Data Handling*, **9B**, pp. 15–25, 1996.