

## USE OF SEMANTIC CONCEPTS FOR THE ANALYSIS OF STRUCTURAL HETEROGENEITY IN RURAL LANDSCAPES

Díaz Varela, E. R.<sup>1</sup>; Roces Díaz, J. V.<sup>2</sup>; Cardin Pedrosa, M.<sup>1</sup>; Álvarez López, C. J.<sup>1</sup>

<sup>1</sup> Universidad de Santiago de Compostela, <sup>2</sup> Universidad de Oviedo

The analysis of landscape heterogeneity is a useful tool in classification and mapping processes, and delineation of landscape planning units. Habitually, definition of the level of landscape heterogeneity and its spatial distribution at a number of scales use diversity indices derived from Information Theory (Shannon, Simpson...). Despite their undeniable usefulness, such indices are focused in the analysis of spatial distribution of quantitative (or "syntactical") information, which allows just for a limited comprehension of results. Nevertheless, possibilities exist for the integration of qualitative elements (relative to "semantic" information) in the analysis, which would permit an enhancement of the analysis' possibilities.

In this work the potential of heterogeneity analysis is defined through the use of two types of heterogeneity indices derived from Information Theory: those analyzing the syntactical information contributed by land cover categorical maps, and those allowing for the inclusion of semantic analysis criteria. Results show the advantages in the application of the latter, related to the improvement of the analysis of the qualitative aspects of landscape heterogeneity.

**Keywords:** *Heterogeneity analysis; Rural landscape; Rural development; GIS; Information Theory; Semantic Spaces*

## APLICACIÓN DE CONCEPTOS SEMÁNTICOS PARA EL ANÁLISIS DE LA HETEROGENEIDAD ESTRUCTURAL DEL PAISAJE RURAL

El análisis de la heterogeneidad territorial es una herramienta de gran utilidad en los procesos de clasificación, cartografía y delimitación de unidades de planificación en ordenación del paisaje. Habitualmente, la definición del nivel de heterogeneidad paisajística y su distribución espacial a distintas escalas recurre al empleo de índices de diversidad derivados de la teoría de la información (Shannon, Simpson...). Pese a su innegable utilidad, dichos índices se centran en el análisis de la distribución espacial de información cuantitativa (o "sintáctica"), lo que permite una comprensión limitada de los resultados. Sin embargo, existe la posibilidad de integrar elementos cualitativos (relativos a información "semántica") en el análisis, que permitirían una ampliación de sus posibilidades.

En este trabajo se definen las posibilidades de análisis a través del empleo de dos tipos de índices de heterogeneidad procedentes de la teoría de la información: aquellos que analizan la información sintáctica aportada por mapas categóricos de cobertura del suelo, y aquellos que permiten criterios de análisis semántico. Los resultados permiten ver las ventajas de aplicación de estos últimos, relativa a la mejora del análisis de los aspectos cualitativos de la heterogeneidad territorial.

**Palabras clave:** *Análisis de la heterogeneidad; Paisaje rural; Desarrollo rural; SIG; Teoría de la Información; Espacios Semánticos.*

## 1. Introducción

Un elemento importante en los procesos actuales de clasificación y cartografía del paisaje es el análisis de la heterogeneidad estructural (Dramstad et al., 1996; Botequilha & Ahern, 2002; Ahern, 2005). Esta nos permite la definición de unidades de paisaje con una elevada precisión (Díaz-Varela, 2009a), de forma que la cartografía pueda ser empleada en procesos de planificación, gestión y conservación. Con tal fin, habitualmente se recurre al empleo de índices espacialmente explícitos, aplicados sobre cartografía digital empleando Sistemas de Información Geográfica. Dichos índices permiten la definición cuantitativa de los niveles de heterogeneidad paisajística y cómo ésta se distribuye espacialmente a distintas escalas. Como tales, suponen una gran ayuda para la comprensión de la complejidad estructural de los paisajes rurales y de los agroecosistemas que los componen.

Los índices derivados de la teoría de la información han demostrado extensivamente su utilidad en el análisis espacial (O'Neill et al., 1988; Magurran, 1988; Riitters et al., 1995; Fjellstad et al., 2001). El empleo de estos índices procede de la analogía de los ecosistemas que componen el territorio como canales que "proyectan" información (por ejemplo, especies o taxones) hacia el futuro (Margalef, 1993). La extensión desde un análisis por especies a un análisis territorial (Pielou, 1975; Magurran, 1988) ha sido facilitada por el uso conjunto de modelos de representación del paisaje como el de Matriz-Mancha-Corredor (Forman & Godron, 1986; Forman, 1995), y de los Sistemas de Información Geográfica (Haines-Young et al., 1993).

Entre los índices de información más empleados en el análisis del paisaje se encuentran el índice de Shannon-Wiener y el de Simpson. El índice de Shannon-Wiener se calcula siguiendo la expresión:

$$H_1 = - \sum_{i=1}^n p_i \log p_i = H' \quad (1)$$

Donde  $p_i$  es, en su aplicación a los paisajes, la frecuencia relativa de las  $n_i$  clases de uso o cobertura del suelo.

El índice de Simpson se calcula mediante la expresión:

$$H_2 = \sum_{i=1}^n p_i^2 = S \quad (2)$$

Ambos índices están relacionados matemáticamente, pudiendo considerarse un caso especial de la función estadística general empleada en teoría de la información: la entropía generalizada de orden  $\alpha$  (Pielou, 1975; Ricotta, 2000), que se calcularía mediante la expresión:

$$H_\alpha = \frac{\log \sum_{i=1}^n p_i^\alpha}{1-\alpha} \quad (3)$$

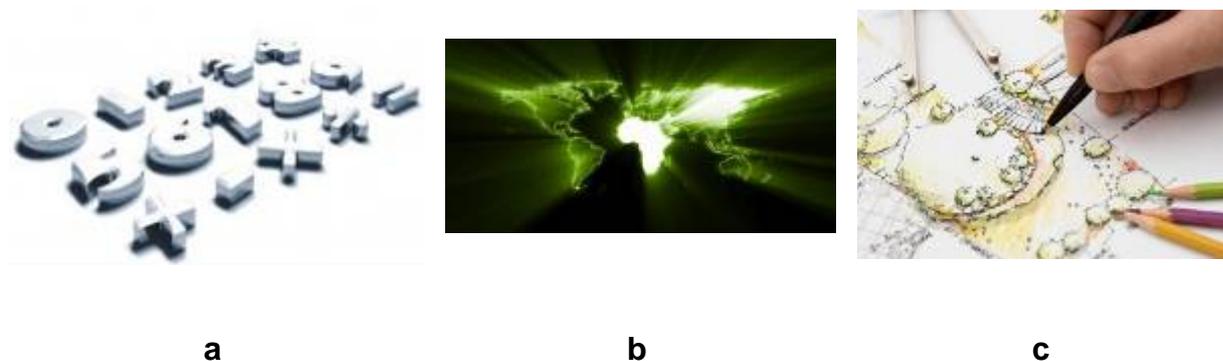
En la que  $\alpha$  es un número real arbitrario que respresenta el orden de la entropía, de forma que  $\alpha=1$  nos permite encontrar la expresión de Shannon-Wiener, y  $\alpha=2$ , la de Simpson (Pielou, 1975; Ricotta, 2000). Trabajos anteriores de los autores han permitido desvelar algunos elementos interesantes de la aplicación de estos índices al análisis del paisaje, tales como su consistencia a través de cambios en la escala (Díaz-Varela et al., 2009b), o la capacidad de modelos logísticos para predecir comportamientos multi-escalares por parte del índice de Shannon-Wiener (Díaz-Varela et al., 2009c).

Pese a su innegable aplicabilidad, se han descrito algunas limitaciones en su utilización, derivadas de su tendencia a ocultar el componente de normalización (“evenness”) de la heterogeneidad (Odum, 1969; Yue et al., 2007), y la insuficiencia ocasional del número de manchas analizadas para la obtención de resultados significativos (Yue et al., 2007; Steinhardt et al, 1999). No obstante, estas limitaciones se superan con relativa facilidad cuando el objetivo es la obtención de una expresión cuantitativa de la variación espacial de la heterogeneidad. Resulta más complejo superar la limitación de estos índices en cuanto a que informan solamente de los aspectos cuantitativos de la heterogeneidad.

La razón para estas últimas limitaciones la debemos buscar en las diferentes aproximaciones semióticas hacia el concepto de información (Jantsch, 1980; Naveh & Liebermann, 1993). Así, los índices de información se orientan hacia la medición de la información *sintáctica*: la cantidad de información (en *bits*) que transmite un mensaje (en el contexto original del índice), o en la que se puede traducir una comunidad ecológica o ecosistema. La información en este caso no posee un significado inherente: ésta es la razón por la que el grado de heterogeneidad medido por el índice de Shannon-Wiener es el mismo para dos paisajes en los que los valores de  $p_i$  y  $n_i$  son idénticos, pero en los que las tipologías analizadas (p.ej., de coberturas de suelo, ecotopos, usos, etc.) pueden variar mucho en composición, calidad, etc. En consecuencia, los resultados de los índices tienen una baja capacidad de aplicación práctica.

Por el contrario, la información *semántica* hace referencia al significado real de la información. Ésta supone un tipo cualitativo de información, con capacidad real para cambiar nuestra percepción y guiar nuestras acciones (Nonaka, 1988). Si bien su medición puede resultar difícil, existe como veremos la posibilidad de integrar elementos cualitativos en los índices espacialmente explícitos, que permitirían una ampliación de sus posibilidades de análisis (ver Fig.1).

**Figura 1: Tipos de información**



a. Información *sintáctica*: es la información almacenada en forma cuantitativa, sin significado inherente. b. información *semántica*, que permite una comprensión de la información y su forma de ordenarse. c. información *pragmática*, desarrollada a partir de la ubicación de la información en un contexto, y que permite la participación en el mismo. Ver texto para detalles.

Un último tipo de información, la *pragmática*, constituiría la información necesaria para ser manejada en el proceso de planificación y gestión de los paisajes: se trata de la información que, al apoyarse en sistemas de retroalimentación (es decir, la novedad aportada por la reacción del receptor de la información ante la misma) permite acciones reales puestas en práctica sobre el terreno a través de la planificación (Naveh, 2007).

La aproximación científica y técnica hacia el paisaje demanda una perspectiva integradora de los tres tipos de información mencionados. Efectivamente, en los últimos años se ha reconocido la necesidad de extender las orientaciones convencionales basadas en el análisis de patrones y procesos, a lo que ya se conoce como *Design in Science Paradigm*, o el “Paradigma del Diseño en Ciencia” (Nassauer & Opdam, 2008; Musacchio, 2009, Opdam., 2010; Musacchio, 2011), integrador del diseño (y por tanto, necesitado de la información pragmática para su desarrollo).

## 2. Objetivos

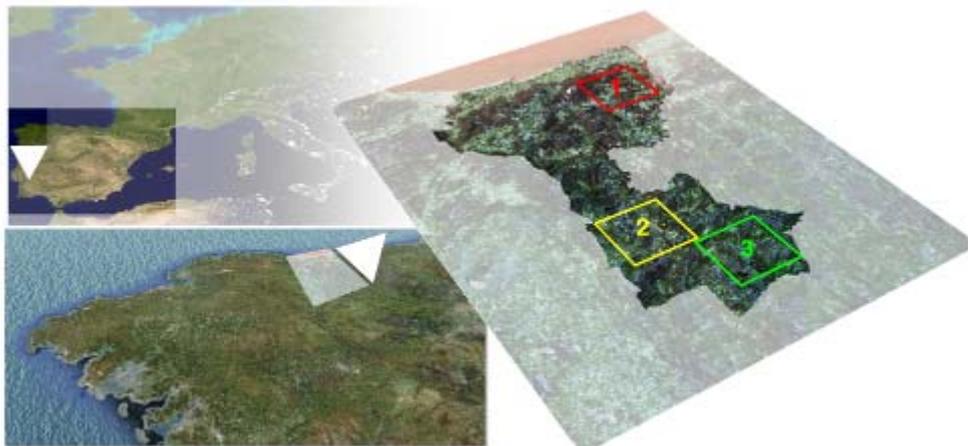
Como primera aproximación a las mencionadas necesidades de integración de aspectos cualitativos en el análisis, en este estudio hemos testado las posibilidades de un índice de aplicación novedosa en el campo del análisis del paisaje. Se trata del Índice de Entropía Cuadrática (Rao, 2010; Ricotta & Szeidl, 2006), combinado con el cálculo de distancias de disimilitud basadas en atributos semánticos (Alqvist & Shortridge, 2009) para su aplicación a cálculos espacialmente explícitos y a escalas múltiples a través de Sistemas de Información Geográfica. El objetivo fundamental es poner en práctica su utilización en el mencionado ámbito, y compararlo con otros índices derivados de la teoría de información.

## 3. Material y métodos

### 3.1. Zona de estudio

El estudio se ha desarrollado en la comarca de “A Mariña Oriental” ( $43^{\circ}26'05''\text{N}-7^{\circ}10'17''\text{W}$ ), situada al noreste de la comunidad autónoma de Galicia (España).

**Figura 2: Zona de estudio y áreas de trabajo**



Zona de estudio (Comarca de A Mariña Oriental), situada en el NE de la Comunidad Autónoma de Galicia. Se destaca en tres colores diferentes la ubicación de las tres áreas de trabajo.

Dentro de la misma, se han seleccionado tres áreas de trabajo (ver Fig.2) de dimensiones 5000x5000 metros, representando tres tipos de paisaje diferentes dentro de la misma zona, de forma que:

Área 1: Paisaje homogéneo, mostrando elevada intensificación en la actividad forestal (predominio de masas de *Eucalyptus Globulus*) y agrícola (praderas y cultivos de maíz de orientación forrajera).

Área 2: Paisaje heterogéneo, mostrando un mosaico que combina manchas forestales multiespecíficas, manchas seminaturales de bosque autóctono, zonas de agricultura tradicional y manchas de matorral (brezal-tojal).

Área 3: Paisaje homogéneo seminatural, formado mayoritariamente por manchas de bosque seminatural multiespecífico, matorral, y agricultura tradicional.

### 3.2. Cartografía

Se ha partido de la cartografía de coberturas del suelo de obtenida en estudios previos (Marey-Pérez, 2002; Díaz-Varela, 2005) por fotointerpretación analógica de fotografías aéreas escala aprox. 1:18.000. La leyenda se ha reclasificado siguiendo criterios de clasificación jerárquica del European Nature Information System (EUNIS) (Agencia Ambiental Europea, 2013), obteniéndose un total de 20 clases de cobertura del suelo. Se ha empleado el software ArcGIS 9.3® para el procesado y manejo de la información cartográfica.

### 3.3. Índices

Se ha aplicado a cada una de las zonas el cálculo de dos índices. Por una parte, el índice de Simpson (ver Ecuación 1), y el Índice de Entropía Cuadrática (Rao, 2010; Ricotta & Szeidl, 2006), cuyo cálculo responde a la expresión:

$$Q = \sum_{i=1}^s \sum_{j=1}^s p_i p_j d_{ij} \quad (4)$$

En donde  $p_i$  y  $p_j$  son, respectivamente, la proporción de las coberturas de suelo  $i$  y  $j$ ; y  $d_{ij}$  la disimilitud existente entre ambas clases de cobertura del suelo. Dicha disimilitud se ha estimado a través de la distancia semántica existente entre *espacios conceptuales* (Alqvist & Shortridge, 2009) establecidos para cada clase de cobertura del suelo, obtenidas a partir de un índice de distancia  $d_{ij}$  que responde a la expresión:

$$d_{ij} = \sqrt{\frac{|U|}{\sum_k W_{Bk} (P_{ik} - P_{jk})}} \quad (5)$$

En donde  $P_{ij}$  y  $P_{ik}$  son, respectivamente, los espacios conceptuales respectivos de las coberturas de suelo  $i$  y  $j$ . Los espacios conceptuales se han establecido definiendo para cada clase de cobertura del suelo los valores numéricos de una serie de variables (ver Tabla 1).  $W_{Bk}$  es un coeficiente de ponderación establecido para cada diferencia entre pares de variables. Y  $U$ , es el conjunto de los espacios conceptuales considerados.

**Tabla 1. Variables definidas para los espacios conceptuales**

Variable	Unidad de medida
Cobertura de vegetación	%
Caducifolias	%

---

Coníferas	%
Matorral	%
Herbáceas	%
Altura media	metros
Grado de intervención humana	%
Índice de forma	%
Tasa de reemplazo (turnover)	años
Presencia de agua	%
Persistencia de agua	%
Elementos culturales	%
Cobertura por especies	%

---

El cálculo de los diferentes dij ha permitido la creación de una matriz de disimilitud, a partir de la cual se han obtenido los valores para el cálculo del índice de Entropía Cuadrática.

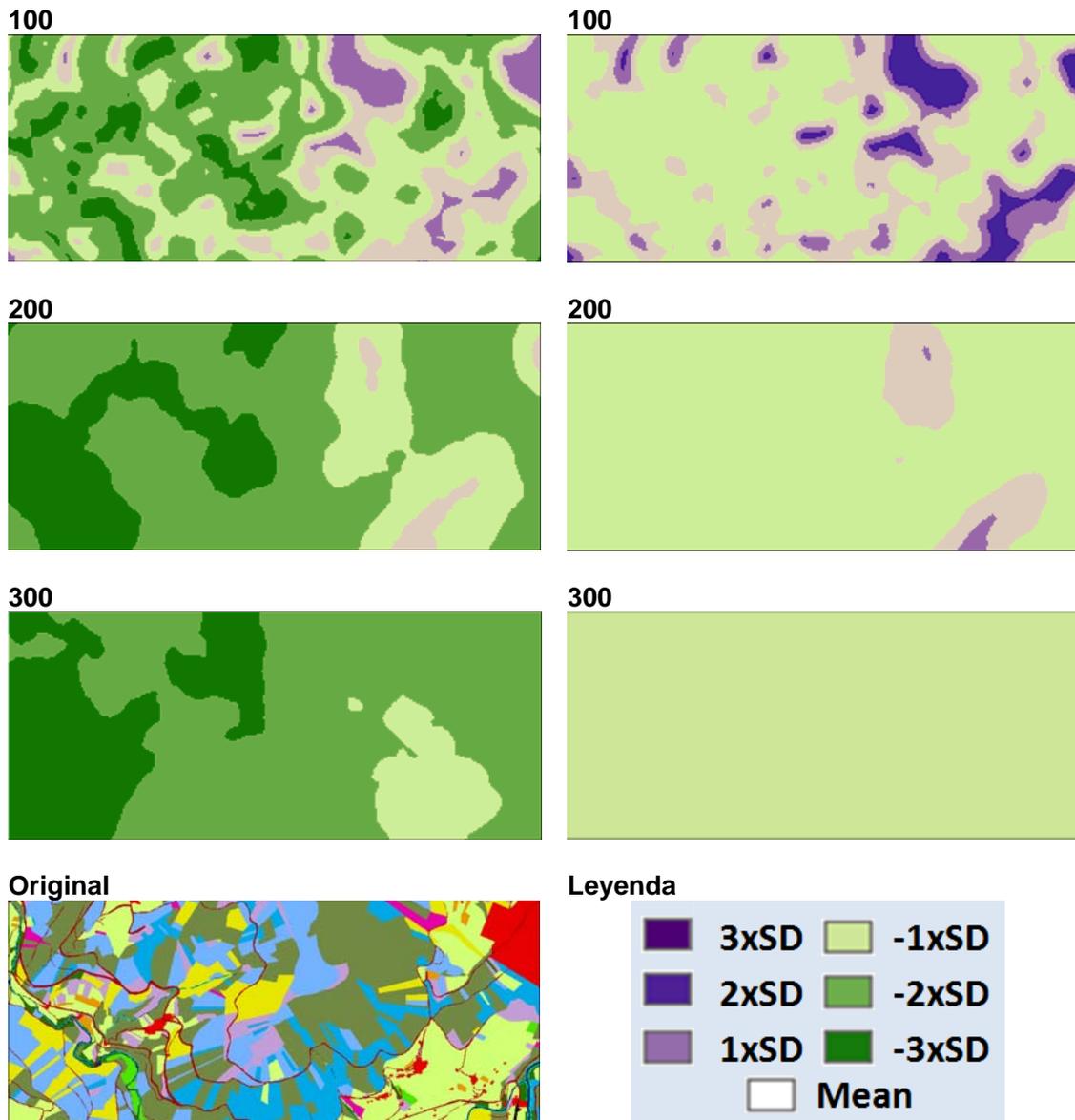
Para el cálculo de los índices ha empleado el software Fragstats (McGarigal et al., 2002), a través de la opción de ventana móvil circular, estableciéndose un rango de tamaños de ventana de 100 a 1000 m de radio, obteniéndose mapas a intervalos de 100m. De esta forma, el análisis se ha podido realizar a múltiples escalas.

#### **4. Resultados y discusión**

A partir del cálculo realizado se han generado un total de 10 mapas raster de valores continuos para cada una de las tres zonas y cada uno de los índices (siendo por tanto un total final de 60 mapas). Un ejemplo comparativo de los resultados puede encontrarse en la Fig. 3. Los resultados permiten observar cómo el Índice de Entropía Cuadrática actúa de una forma similar al índice de Simpson, del que se puede considerar un caso especial para el que se han integrado valores cualitativos (ver ecuaciones 2 y 4). No obstante, la integración de criterios semánticos asegura que la distribución espacial del índice muestra variaciones en la heterogeneidad basadas en las distancias establecidas entre variables que tienen un significado claro (esto es, basadas en criterios semánticos).

Por otra parte, la generación de mapas mediante tamaños de ventana variables permite la detección de diferencias en la heterogeneidad a escalas múltiples, lo cual resulta de ayuda para la identificación de dominios de escala espacial, que difieren en el comportamiento de la heterogeneidad para un área geográfica dada.

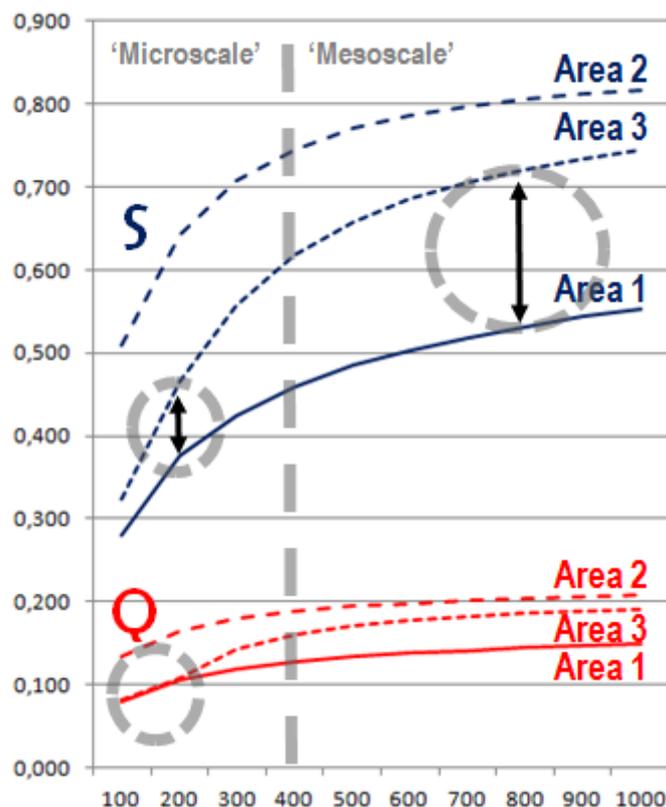
Figura 3: Mapas de resultados en los cálculos de índices



Resultados del cálculo del Índice de Entropía Cuadrática (izda.) y del Índice de Simpson (dcha.), para tres escalas diferentes, correspondientes a las ventanas de cálculo de radio 100, 300 y 1000 metros. Obsérvese el efecto progresivo de promediado relativo al incremento del área de cálculo. Se muestran abajo el mapa original y la leyenda, que muestra el alejamiento de la media de los datos en unidades de desviación típica.

Por otra parte, el análisis del comportamiento de los valores medios de los dos índices en función de la escala adoptada para el cálculo (Fig. 4) permite ver la diferencia en la discriminación entre niveles de heterogeneidad entre los tres ejemplos. Los dos índices distinguen de forma clara las diferencias entre paisajes homogéneos y heterogéneos. Asimismo, se definen claramente dos dominios de escala, asociados a un cambio en la pendiente de los valores de los índices, que se verifica en torno a la ventana de radio 400 metros, y que se pueden definir como dominios de “microescala” y de “mesoescala”.

Figura 4: Comportamiento multiescalar de los índices



Se muestra el comportamiento a escalas múltiples de los valores medios del Índice de Entropía Cuadrática (Q) y del Índice de Simpson (S). El incremento en los valores con la escala sufre un importante cambio en el valor de la pendiente en torno a la ventana de 400 m de radio, permitiendo definir dos dominios: de microescala y de mesoescala. Se observan asimismo diferencias claras en la detección de diferencias entre áreas de paisajes homogéneos y heterogéneos.

Tal como muestra la figura, las diferencias en heterogeneidad son mínimas entre los diferentes tipos de paisaje cuando se comparan en el dominio de microescala, acrecentándose las diferencias con el incremento en la escala de análisis. Esto demuestra la necesidad de establecer enfoques multiescalares en los análisis de paisaje orientados a la planificación, debido a la posibilidad de que existan escalas a las que diferencias críticas en la composición y configuración de los elementos del paisaje puedan no ser apreciadas.

## 5. Conclusiones

En este trabajo, se han integrado criterios semánticos en el cálculo del Índice de Entropía Cuadrática, y se ha aplicado el mismo al análisis de mapas categóricos establecidos como modelos de representación de paisajes rurales. La capacidad de dicho índice para la detección de diferencias en la heterogeneidad entre paisajes con diferentes características, la posibilidad de interpretar dichas diferencias en base a los criterios semánticos empleados y el estudio de la variación detectada a diferentes escalas permite señalar a la integración de elementos cualitativos en combinación con perspectivas multiescalares en el análisis de la heterogeneidad de los paisajes puede considerarse como un avance en la aplicación de índices de paisaje.

## 6. Referencias

- Agencia Ambiental Europea (2013). European Nature Information System (EUNIS). Disponible en <http://eunis.eea.europa.eu/> (ultimo acceso 17/04/2013)
- Ahern, J. (2005). Theories, methods and strategies for sustainable landscape planning. En, B.Tress, G. Tress, G. Fry, & P. Opdam, (Eds.): From landscape research to landscape planning: Aspects of integration, education and application. (Pages 119-131). Dordrecht: Springer.
- Alqvist, O., Shortridge, A. (2010). Spatial and semantic dimensions of landscape heterogeneity. *Landscape Ecology*, 25, 573-590.
- Botequilha Leitao, A., Ahern, J. (2002). Applying landscape ecological concepts and metrics in sustainable landscape planning. *Landscape and Urban Planning*, 59, 65-93.
- Diaz Varela, E.R. (2005). Planificación ecológica multiescala con Sistemas de Información Geográfica para la sostenibilidad de los paisajes. Aplicación a la comarca de A Mariña Oriental (NE de Galicia-España). Tesis Doctoral. Edición digital (CD), Santiago de Compostela: Servicio de Publicaciones de la Universidad de Santiago.
- Diaz-Varela, E., Álvarez-López, C.J., Marey-Pérez, M. (2009a). Multiscale delineation of landscape planning units based on spatial variation of land-use patterns in Galicia, NW Spain. *Landscape and Ecological Engineering*, 5, 1-10
- Díaz-Varela, E.R., Marey-Pérez, M.F., Álvarez-Álvarez, P. (2009c). Use of simulated and real data to identify heterogeneity domains in scale-divergent landscapes. *Forest Ecology and Management*, 258, 2490-2500.
- Dramstad, W.E., Olson, J.D., Forman, R.T.T. (1996). *Landscape Ecology Principles in Landscape Architecture and Land Use Planning*. Washington: Island Press.
- Fjellstad, W.J., Dramstad, W.E., Strand, G.-H., Fry, G.L.A. (2001). Heterogeneity as a measure of spatial pattern for monitoring agricultural landscapes. *Norsk Geografisk Tidsskrift*, 55, 71–76.
- Forman, R.T.T. (1995). *Land mosaics: The ecology of landscapes and regions*. Cambridge: Cambridge University Press.
- Forman, R.T.T., Godron, M., (1986). *Landscape Ecology*. New York: John Wiley and Sons.
- Haines-Young, R., Green, D.R., Cousins, S. (1993). *Landscape ecology and geographic information systems*. Londo: Taylor & Francis.
- Jantsch, E. 1980. *The self-organizing universe. : Scientific and human implications of the emerging paradigm of evolution*. Oxford: Pergamon Press.
- Magurran, A.E. (1988). *Ecological Diversity and its Measurement*. London: Croom Helm,
- Marey-Pérez, M. F. (2002). Tenencia de la Tierra en Galicia. Modelo para la caracterización de Propietarios Forestales (Land Tenure in Galicia. A Model for Characterization of Forest Owners), Tesis Doctoral. Edición digital (CD), Santiago de Compostela: Servicio de Publicaciones de la Universidad de Santiago.
- Margalef, R. (1993). *Teoría de los Sistemas Ecológicos*. Barcelona: Edicions Universitat Barcelona.
- McGarigal, K., Cushman, S. A., Neel, M. C., & Ene, E. 2002. FRAGSTATS: Spatial Pattern Analysis. Disponible en [www.umass.edu/landeco/research/fragstats/fragstats.html](http://www.umass.edu/landeco/research/fragstats/fragstats.html). (ultimo acceso 17/04/2013)

- Musacchio, L.R. (2009). The scientific basis for the design of landscape sustainability: A conceptual framework for translational landscape research and practice of designed landscapes and the six Es of landscape sustainability. *Landscape Ecology*, 24, 993-1013.
- Musacchio, L.R. (2010). The grand challenge to operationalize landscape sustainability and the design-in-science paradigm. *Landscape Ecology*, 26, 1-5.
- Nassauer, J.I., & Opdam, P. (2008). Design in science: extending the landscape ecology paradigm. *Landscape Ecology*, 23, 633-644.
- Naveh, Z. (2007). *Transdisciplinary Challenges in Landscape Ecology and Restoration Ecology*. Dordrecht: Springer.
- Naveh, Z. Liebermann, A.S. 1993. *Landscape Ecology: Theory and Application*. Dordrecht: Springer.
- Nonaka, I.(1988). Creating organizational order out Of chaos: Self-renewal in Japanese firms. *California Management Review*, 30, 57-73.
- O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H., Graham, R.L. (1988). Indices of landscape pattern. *Landscape Ecology*, 1, 153-162.
- Odum, E. (1969). The strategy of ecosystem development. *Science*, 164, 262-270.
- Opdam, P. (2010). Learning science from practice. *Landscape Ecology*, 25, 821-823.
- Pielou, E.C. (1975). *Ecological Diversity*. New York: John Wiley & Sons.
- Rao, C.R. (2010). Quadratic Entropy and Analysis of Diversity. *Sankhya: The Indian Journal of Statistics*, 72-A, Part 1, 70-80.
- Ricotta, C. (2000). From theoretical ecology to statistical physics and back: selfsimilar landscape metrics as a synthesis of ecological diversity and geometrical complexity. *Ecological Modelling*, 125, 245-253.
- Ricotta, C., Szeidl, L. (2006). Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, 70, 237-243.
- 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. (1995). A factor analysis of landscape pattern and structure metrics. *Landscape Ecology*, 10, 23-39.
- Steinhardt, U., Herzog, F., Lausch, A., Müller, E., Lehmann, S. (1999). Hemeroby index for landscape monitoring and evaluation. In: Pykh, Y.A., Hyatt, D.E., Lenz, R.J. (Eds.), *Environmental Indices-System Analysis Approach*. Oxford: EOLSS Publishing. Pp. 237-254.
- Yue, T.-X., Ma, S.-N., Wu, S.-X., Zhan, J.-Y. (2007). Comparative analyses of the scaling diversity index and its applicability. *International Journal of Remote Sensing*, 28, 1611-1623.