A distance dependent contagion function for vector-based data

Authors: Habib Ramezani * and Sören Holm

* H. Ramezani (Corresponding author).

Department of Forest Resource Management, Swedish University of Agriculture Science, SLU, SE-901 83 Umeå, Sweden

E-mail: Habib.Ramezani@slu.se , Ramezani.habib@gmail.com

Phone: +46 90 786 81 51/ Fax: +46 90 77 81 16

Abstract

Landscape pattern is of primary interest to landscape ecologists and landscape metrics are used to quantify landscape pattern. Metrics are commonly defined and calculated on raster-based land cover maps. One metric is the contagion, existing in several versions, e.g., unconditional and conditional, used as a measure of fragmentation. However, mapped data is sometimes in vector-based format or there may be no mapped data but only a point sample. In this study a definition of contagion for such cases is investigated. The metric is an extension of the usual contagion, based on pairs of points at varying distances and gives a function of the distance. In this study the extended contagion is calculated for vector-based delineated real landscapes and for simulated ones. Both unconditional and conditional contagions are studied using two classification systems. The unconditional contagion function was decreasing and convex, with upper and lower limits highly correlated to the Shannon diversity index, thus carrying only area proportion information. The spatial information lies in the speed by which the function converges to the lower limit; using a proxy function this can be expressed by a single parameter *b*, with high values for fragmented landscapes. No proxy function was found for the conditional contagion, for which only qualitative information was found. The extended contagion is applicable both in patch mosaic models of landscapes and in gradient-based models, where landscape characteristics change continuously without distinct borders between patches. The extended contagion can be useful in sample based surveys where there no map of the entire landscape is available.

Key words: Landscape pattern analysis; Landscape metrics; contagion; vector-based; point sampling

1. Introduction

Landscape pattern is of primary interest to landscape ecologists, because it is assumed that landscape pattern can significantly affect ecological processes (Turner, 1989) such as biodiversity and population dynamics (Forman, 1995; Schumaker, 1996; Wiens et al., 1997). Thus, landscape metrics, as predictor variables, can help us better understand pattern-process relationships (Bebber et al., 2005; Hernandez-Stefanoni, 2005). The metrics can also be used to detect differences between various landscapes and changes in a given landscape over time. The metrics typically are defined in terms of landscape elements such as the number, area, and edge length of patches (O'Neill et al., 1988; Turner, 1990; Hunsaker et al., 1994).

A variety of landscape metrics have been developed to capture both composition and configuration aspects of landscape structure (McGarigal and Marks, 1995; Gustafson, 1998). Composition refers to the number of land cover types and their proportions within landscapes whereas configuration refers to the spatial distribution of land cover types. An example of a configuration metric is the contagion, which was first proposed by O'Neill et al. (1988) to measure the degree of clumping of patches. This metric was proposed by United States Environmental Protection Agency (1994) as an effective indicator in landscape pattern analysis. Indirectly, the contagion can provide information on landscape fragmentation (Hargis et al., 1998), and fragmentation is important for many ecological processes (Fahrig, 2003). Furthermore, the contagion is highly correlated with metrics of diversity, dominance, and patch richness (Riitters et al., 1995; Cain et al., 1997; Frohn, 1998).

Landscape metrics in general, and the contagion metric in particular, are defined and commonly calculated on raster-based land cover/use maps and a frequently used computer software is FRAGSTATS (McGarigal and Marks, 1995). However, in some environmental monitoring programs, such as the Norwegian 3Q (NIJOS, 2001) and the National Inventory of Landscapes in Sweden (NILS)(Ståhl et al., 2010) the maps are vector-based (i.e., from aerial photographs). In addition, metrics like contagion are sensitive to pixel size, so that with increasing pixel size contagion value decreases since the number of within-patch edges decreases faster than between-patch edges (Wickham and Riitters, 1995; Ricotta et al., 2003). Some errors can be introduced by converting vector data to raster data, for instance, small patches may disappeared (Lunetta et al., 1991; Wade et al., 2003; Jenness, 2004).

In this study, vector-based data set was used where vector-based means that landscapes can be delineated into polygons (homogenous area), each with a uniquely defined cover type (class). Photo-interpreted polygons are assumed to provide an accurate description of existing land cover classes. However to use survey data for *estimating* the contagion metric defined here, there is no need to really perform a delineation.

Since vector format data sets are important sources for many environmental monitoring programs it would be useful to develop new metrics or to redefine currently used metrics to meet this data format. Whereas a few studies have been conducted on vector-based data sets to calculate metrics such as Shannon's diversity and edge density (Corona et al., 2004; Ramezani and Holm, 2009; Ramezani et al., 2010) and contagion (Wickham et al., 1996), more attention is needed in this area.

The purpose of this paper was to develop a new definition of contagion for vector-based data and derive properties of it for a variety of landscapes. In general the definition was guided by two goals 1) the aim to avoid the resolution problem in raster-based data, and 2) to find a definition that should admit estimation of the contagion metric from a point sample in a non-delineated landscape, for which only raw, non-delineated remotely sensed data is available.

2. Material and methods

2.1. The raster-based contagion metric (C)

Contagion (C) is as a measure of clumping of classes within a landscape. There are several alternative raster-based definitions (Riitters et al., 1996), of which two are considered here, one called conditional contagion, C_u and the other unconditional, C_c (Li and Reynolds, 1993). The definition of C_u is

$$C_u = 1 + \frac{\sum_{i}^{s} \sum_{j}^{s} p_{ij} \cdot \ln(p_{ij})}{2\ln(s)} \qquad (1)$$

where p_{ij} is the proportion of all pairs of adjacent pixels that belongs to the land cover type i and j. The definition of C_c is

$$C_c = 1 + \frac{\sum_i^s \sum_j^s p_{j/i} \cdot \ln(p_{j/i})}{s \ln(s)} \qquad (2)$$

where $p_{j/i}$ is the proportion of pixels of class *j* adjacent to pixels that belong to the class *i* and *s* is the number of land cover types considered. Usually two pixels are defined as adjacent if they are neighbours in any of the four principal directions but other definitions are possible (Turner et al., 2001). Both definitions are normalized to give values between 0 and 1, with low values indicating fragmented landscapes with adjacency types in roughly equal proportion (conditionally or unconditionally). Calculation of the two metrics from raster data is simple from an adjacency matrix that shows the frequency of all adjacency types on the raster-based land cover map (Haralick et al., 1973).

2.2. A vector-based contagion metric

A contagion metric for vector data was proposed by Wickham et al. (1996), essentially using Eq. (1) but with a different normalization. The definition was based on the proportions of edge length between all possible adjacent classes to total edge length within the landscape, implying that the main diagonal of the adjacency matrix (within-class edge) has been ignored. It is, at least theoretically, possible to estimate edge lengths from a point sample (Ramezani et al., 2010). However, in this study we propose another metric which is better adapted for point sampling and which is also an extension of the raster-based definition (Li and Reynolds, 1993).

For any given distance *d* we define $p_{ij}(d)$ as the probability that two randomly chosen points at distance *d* belongs to the classes *i* and *j* (in that order). Thus $p_{ij}(d)$ is a function of the distance *d* and we also have $p_{ij}(d) = p_{ji}(d)$. The unconditional contagion function $C_u(d)$ is, for each given distance *d*, defined exactly as C_u according to Eq. (1) thus $C_u(d)$ can be written as

$$C_u(d) = 1 + \frac{\sum_{i=1}^s \sum_{j=1}^s p_{ij}(d) \ln(p_{ij}(d))}{2 \ln(s)} \qquad (3)$$

The conditional probability that the "second point" in a pair of a randomly chosen points at distance *d* belongs to class *j*, given that the "first" point belongs to class *i* equals $p_{j/i}(d) = p_{ij}(d) / p_i(d)$, where $p_i(d) = \sum_{j=1}^s p_{ij}(d)$. The conditional probability function $p_{j/i}(d)$ gives a conditional contagion function $C_c(d)$ according to Eq. (2), defined as

$$C_c(d) = 1 + \frac{\sum_{i=1}^s \sum_{j=1}^s p_{j/i}(d) \ln(p_{j/i}(d))}{s \ln(s)} \qquad (4)$$

The functions are independent of any mapping resolution, except for restrictions caused by the minimum mapping unit. The interpretation is similar to that of the raster-based contagion but is extended to "fragmentation at distance *d*". From the definitions of the functions it is also clear that they can be estimated by point sampling. However, this issue is not pursued further in this paper.

2.3. The study

In this study both an analytical and an empirical investigation of the properties of the two contagion functions (Eq. 3 and 4) were performed. This was accomplished by calculating the values of the functions (for a number of distances) and interpreting the metrics values in terms of landscape pattern properties. This study was conducted for real and already manually photo-interpreted delineated landscapes and also for simulated ones. The simulated landscapes were used in order to assess the behaviour of the metrics in extreme cases or cases not covered by the real ones. An underlying idea of the study was that describing landscape pattern properties by continuous functions would reveal more about them than describing them by simple numbers as in the raster-based case. A comparison between raster-based (by FRAGSTATS) (McGarigal and Marks, 1995) and vector-based contagion was also performed.

2.4. Material

The study was conducted on data from the National Inventory of Landscapes in Sweden (NILS)(Ståhl et al., 2010), which is a major environmental monitoring program run by the Swedish Environmental Protection Agency. A 25 km^2 quadrate is used in order to capture the broad landscape context. Within a 1 km^2 centrally located quadrate, a detailed delineation of polygons (homogenous areas) is manually made. To obtain a genuine sample of landscapes for our study, we used data from 50 randomly selected quadrates across entire Sweden.

The aerial photographs in which interpretations were made were colour infrared and had a ground resolution of 0.4 m. Polygon delineation was made using the interpretation program Summit Evolution from DAT/EM and ArcGIS from ESRI. For the purpose of the present study, the NILS variables were used together with two different classification systems (7 and 20 classes, see Table 1 for more details) in order to produce land cover maps. The survey was conducted on systems of classification with seven and twenty classes. The classes of the two systems are given in Table 1.

Table1. Classes according to the two different classification systems (7 and 20 classes)

Seven classes	Twenty classes
1- Forest	1-1- Coniferous-Dense
	1-2- Coniferous-Sparse
	1-3- Deciduous-Dense
	1-4- Deciduous-Sparse
	1-5- Mixed-Forest- Dense
	1-6- Mixed-Forest- Sparse
2- Urban	2-1- Housing-Areas
	2-2- Urban-Green-Areas
	2-3- Urban-Forest
3- Cultivated fields	3-1- Crop fields
	3-2- Grassland
4- Wetlands	4-1- Bog
	4-2- Fen
	4-3- Mixed-Wetland
5- Water	5-1- Open-Water 5-2- Water-Vegetation
6- Pasture	6-1- Open- Pasture
	6-2- Pasture-Sparse-Trees
	6-3- Wooded-Pasture
7- Other land	7- 1- Other land

Simulated landscapes with four classes were also created through a raster-based approach of maximum size 512 by 512 squares, with the possibility to build pixels of different sizes (1, 2, 4, …; e.g., pixel size 16 results in $32 \cdot 32 = 1024$ pixels). Several methods were used to simulate landscapes, 1) by creating more or less regular patterns (like chess-boards or strips) with classes randomly or not assigned to pixels, 2) by assigning classes randomly in different proportions to pixels, with constant or non-constant intensity over the landscape, and 3) by assigning classes with probabilities depending on the classes of neighboring squares. The third method was applied in two ways, either directly by assigning classes in succession after a randomly chosen class in a corner of the landscape or by a version of Gibbs sampler. The Gibbs sampler method is as follows: A. Start with any landscape (e.g., with random classes for all pixels). B. Choose one pixel at random. C. A new (could be the same as before) class of the pixel chosen is assigned with a probability depending on the classes of the four neighbors in the main directions. D. Repeat B and C a very large number of times. By choosing the probability matrix in different ways landscapes where classes are repelling or attracting can be created. The simulated landscapes were seen as landscapes delineated into polygons consisting of adjacent pixels of the same class.

2.5. Calculations

To calculate the exact contagion value, even for a single distance d, for a given delineated landscape, is too complicated to be done in practice (in principle it would require calculation of a large number of quadruple integrals with pairs of polygons as domains of integration). Instead we have to rely on Monte Carlo simulations and/or numerical methods. The simplest method is just to apply the definition and take a very large number of pairs of points at distance *d* randomly in the mapped landscape, considering the boundary problem that occurs when the second point falls outside the map. If this happens either a new "first" point should be taken or the pair should be weighted by its inclusion intensity; and the first alternative is the simplest. This "new first point" alternative was applied for the simulated landscapes since the cover class of each point was easily determined by its pixel. For each distance and replicate (see below) 3000-5000 pairs were simulated.

For the mapped real landscape the total computer time to determine the polygon belongings of all points was relatively long. A polygon of the real landscapes could have up to about 800 sides. The average number of polygons was 26 and the average numbers of sides per polygon was 79 for the seven class system and for the twenty class system the figures were 58 and 61, with large variation between landscapes. For this reason an alternative method was applied, allowing the "second" point intensity ("probability") to be calculated exactly. For each map and each polygon (of say class *i*) a sample of "first" points was laid out systematically, with random start, and with a number of points depending on the polygon area (lower and upper bounds were 5 and 290 points per polygon). With the given point as a centre and the distance *d* as radius a circle is defined on which the second point must fall. The lengths of the circumference of the circle within all polygons were determined. The mean of all such lengths over the systematic sample estimates the "local" p_{ij} s for the given polygon and the final "global" p_{ij} is determined by area weighting over all polygons of class *i*.

For both methods of calculation the sample simulations were replicated independently 20 times and the mean value of the contagion was used as a final value, estimated with high precision as judged by the standard error (estimated from the replicates). The relation $p_{ij} = p_{ji}$ was used for the estimation.

For the real landscapes the contagion functions were estimated for the nine distances 2, 5, 10, 20, 30, 60, 100, 150 and 250 meters. The samples were taken independently for different distances. For the unconditional contagion the highest and average standard errors observed from the 450 estimates were 0.0019 and 0.0001. In 81 % of the cases the standard error was less than 0.0011. For the conditional contagion the highest and average standard errors observed were 0.0011 and 0.0003. Hence, we conclude that the precision is high, well within the second decimal of the estimated value.

For the simulated landscapes up to 14 distances, from 0.1 up to 200 length units were used (the sides were 512 length units long) and the values calculated had standard errors of the same size as those for the real landscapes.

3. Results

In this study, the properties of the contagion function for different point distances were investigated. There are both theoretical and empirical results. Some mathematical results are followed by a comparison with empirical values. Finally, the behaviour of the contagion functions are presented and compared to the raster-based analogue.

Mathematically derived properties

(i) If we use the relation $p_{ij}(d) = p_{j/i}(d) \cdot p_i(d)$ we get

$$\sum_{i=1}^{s}\sum_{j=1}^{s} p_{ij}(d)\ln(p_{ij}(d)) = \sum_{i=1}^{s} p_i(d)\ln(p_i(d)) + \sum_{i=1}^{s} p_i(d)\sum_{j=1}^{s} p_{j/i}(d)\ln(p_{j/i})$$

and thus

$$C_u(d) = 1 + \frac{\sum_{i=1}^{s} p_i(d)\ln(p_i(d))}{2\ln(s)} + \frac{\sum_{i=1}^{s} p_i(d)\sum_{j=1}^{s} p_{j/i}(d)\ln(p_{j/i}(d))}{2\ln(s)}$$

while $C_c(d)$ can be written

$$C_c(d) = 1 + \frac{\sum_{i=1}^{s}(1/s)\sum_{j=1}^{s} p_{j/i}(d)\ln(p_{j/i}(d))}{\ln(s)}$$

The third term in $C_u(d)$ corresponds to the second in $C_c(d)$, but for $C_u(d)$ the classes with small areas, implying low values of $p_i(d)$, have smaller impact on the value than they have on $C_c(d)$ where all classes have the same weight. The second term in $C_u(d)$ is independent of *j* and is related to the Shannon diversity index

$$H = -\frac{\sum_{i=1}^{s} a_i \ln(a_i)}{\ln(s)}$$

where a_i is the area proportion of class *i*. Since $p_i(d)$ is the probability that the "first" point in a pair of randomly chosen points belong to class *i* it seems to follow that $p_i(d) = a_i$. However, due to boundary effects this is only approximately true, but at least for small distances *d* it is a good approximation.

(ii) When the distance *d* tends to 0, the two points in a pair tend to fall into the same polygon and the probability that the polygon belongs to class *i* converges to its area proportion a_i. Thus when *d* tends to 0 then $p_{ij}(d)$ tends to a_i for $j = i$ and to 0 for $j \neq i$. Since the function $t \cdot \ln(t)$ is continuous and tends to 0 when *t* does so, we can deduce that

$$C_u(d) \to 1 + \frac{\sum_{i=1}^{s} a_i \ln(a_i)}{2\ln(s)} = 1 - H/2 \qquad \text{when } d \to 0$$

where *H* is the Shannon diversity index.

In the conditional case $p_{j/i}(d)$ tends to 0 when *d* tends to 0 if $j \neq i$ and otherwise to 1, so $C_c(d) \to 1$ when $d \to 0$

(iii) for each of the two function and small distances *d* we have the approximation

$(C(d)-C(0))/d \approx \alpha \cdot \ln(d)+\beta$ with different values for the constants α and β for the two contagion types (for a proof, see Appendix 1). This implies that the (right hand) derivatives of the contagion functions at $d=0$ are (negatively) infinite, i.e. the vertical axis is tangent to both functions.
(iv) for landscapes that are "stationary" (no trends) it is likely that $p_{ij}(d)$ for large distances d is close to the product $p_i(d)\cdot p_j(d)$ (due to long distance independence). When this approximation holds we get, by inserting into the definitions and some algebra, the result that

$$C_u(d) \approx 1+\frac{\sum_{i=1}^{s} p_i(d)\ln(p_i(d))}{\ln(s)} \qquad \text{for large } d$$

The second term is, as mentioned above, related to and likely close to the Shannon diversity index *H*. In the conditional case the approximation depends on the number of classes actually present in the landscape, denoted *r*. We obtain, for large *d*,

$$C_c(d) \approx 1+\frac{r}{s}\cdot\frac{\sum_{i=1}^{s} p_i(d)\ln(p_i(d))}{\ln(s)}$$

Comparison between theoretical and empirical values

The items (ii) and (iv) have been compared to the estimated values of the contagion functions for the real landscapes, considering 2 meters as close to 0 and 250 meters as close to infinity (for the 1 by 1 km landscape). The results are given in Table 2.

Table 2. Comparison between empirical and theoretical values for conditional and unconditional contagion for small and large distances *d* for the two classification systems. Figures show mean values of absolute values of the differences for the 50 maps. Numbers in parentheses show the number of positive differences of the expression given within the absolute sign. *H* is the Shannon diversity index, *s* the number of classes (7 or 20) and *r* the number of classes actually present. Extrapolation means that the empirical values for 2 and 5 meters were used for a linear extrapolation to $d = 0$.

		7 classes	20 classes
		$\lvert 1-H/2-C_u(d)\rvert$	
Short distance ($d \to 0$)	$d = 0.05$	0.0006 (50)	0.0007 (50)
	$d = 2$	0.0170 (50)	0.0201 (50)
	Extrapolation	0.0048 (47)	0.0048 (50)
		$\lvert 1-C_c(d)\rvert$	
Short distance ($d \to 0$)	$d = 0.05$	0.0016 (50)	0.0010 (50)
	$d = 2$	0.0486 (50)	0.0319 (50)
	Extrapolation	0.0160 (49)	0.0093 (50)
		$\lvert C_u(d)-(1-H)\rvert$	
Long distance ($d \to \infty$)	$d = 250$	0.0137 (47)	0.0152 (47)
		$\lvert C_c(d)-(1-r\cdot H/s)\rvert$	
Long distance ($d \to \infty$)	$d = 250$	0.0227 (34)	0.0230 (40)

In general the values indicate that the 2 and 250 meter distances are small and large enough to catch the values of the contagion functions for small and large distances. One exception is perhaps the short distance for the conditional contagion, for which the decrease is relatively large close to $d = 0$. However, the good agreement for the extrapolation shows that the infinite derivative at $d = 0$ (item (iii) above) should not be a serious problem for approximations or interpolations of the functions even for short distances.

A proxy function for the unconditional contagion function

For all the 50 landscapes studied the unconditional contagion function was a convex and decreasing function of distance (within the range studied). It was found that the contagion function could be described by $C_u(d) \approx f(d)$, where

$$f(d) = c + a \cdot e^{-b \cdot d} \qquad (5)$$

The parameters a, b and c were estimated by nonlinear regression (SAS® NLIN, version 9.2) for each of the 50 landscapes and both classification systems, without any restrictions on *a* and *c*. The fit was good; the average estimated standard deviation around the function was 0.0051 and 0.0063 for the seven and twenty classes systems. By letting *d* tend to 0 and to infinity in $f(d)$ we would expect (from items (ii) and (iv) and from Table 2) that $c+a$ should be close to $1-H/2$ and *c* close to $1-H$ and thus *a* close to $H/2$. The estimated value of *c* exceeded $1-H$ on an average by 0.017 (4 %) and the value of *a* fell short of $H/2$ by 0.028 (12 %) for the seven class system and by similar but slightly larger figures for the twenty classes system. These over- and underestimations are well in line with the corresponding values of the contagion functions (see Table 2). In Figure 1 two examples are used to illustrate the outcome for the seven classes system, one for a landscape with average standard deviation around the function and the other the landscape with the highest standard deviation.

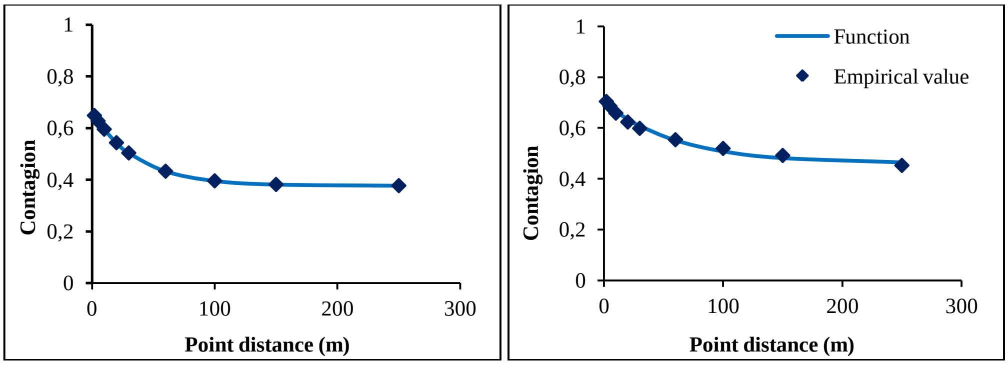


Figure 1. Two examples of unconditional contagion functions with average (left) and largest (right) standard deviation around the function among the 50 real landscapes using the 7 classes system.

Results for the conditional contagion function

For the conditional contagion no simple and unique form of the function was detected and hence no proxy function could be derived. However, the conditional contagion carried qualitative information about the landscape pattern (see the discussion section). In appendix 2 some landscapes with their unconditional and conditional contagion is provided.

Comparison between raster-based and vector-based contagion

The raster-based contagion was, for the seven classes system, compared to the vector-based for a couple of distances. The pixel size for the raster-based case was chosen equal to the distance *d* and the calculations made by FRAGSTATS (McGarigal and Marks, 1995). The two values were very close to each other. In Figure 2 the two measures are plotted against each other. The same normalization was used, i.e., the calculation was based on the number of classes present within the landscape.

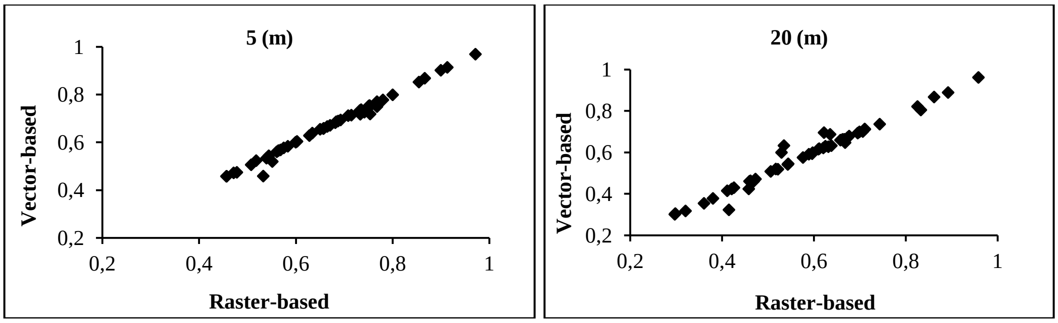


Figure 2. Plots of raster-based versus vector-based contagion values using the seven classes system. The pixel sizes (side of the squares) were the same as the distance between points for the vector-based contagion.

However, the average distance between two randomly chosen points in adjacent pixels exceeds the pixel size by about 9 %, so distance and raster size are not completely equivalent and for this reason the comparison is not perfect.

4. Discussion

In this study, a vector-based definition of the contagion (C) metric is developed. It is an extension of the raster-based definition and gives a contagion metric that depends on distance, i.e., a contagion function. The properties of the contagion function for different point distances and both under unconditional and conditional definitions are investigated.

In a raster environment, contagion is sensitive to pixel size and its value can be increased when decreasing pixel size (Ricotta et al., 2003; Li et al., 2005). PPU (patch-per-unit area) (Frohn, 1998) and UNMIX (independent-resolution) (Ricotta et al., 2003) are two alternative metrics which are insensitive to pixel size but they may fail to capture the configuration aspect of landscapes. Indeed, PPU is equivalent to patch density, a composition metric, (Wu et al., 2002) and UNMIX does not utilizes all information (Ricotta et al., 2003).

The empirical study of the unconditional contagion function revealed that the function could be approximated as $c + a \cdot \exp(-b \cdot d)$ where d is the distance between points. According to the theoretical and empirical findings both c and a are strongly related to the Shannon index and thus to the area proportions of the classes. Hence, we can make two conclusions: 1) the unconditional contagion cannot be interpreted without considering the area proportions (the Shannon index), and 2) the parameter b carries most, if not all, information about the spatial distribution or fragmentation. The larger the parameter b the faster the contagion function tends to its lower bound showing fragmentation for short distances. To illustrate the two conclusions above, the 50 real landscapes were classified into nine categories based on three categories of Shannon diversity (H) and three of the value of the parameter b. (The estimated values for the parameter b ranged from 0.0072 to 0.0656 for seven classes system and from 0.0014 to 0.0850 for twenty classes system). Maps of the four categories with low and high values of H and b are given in Figure 3.

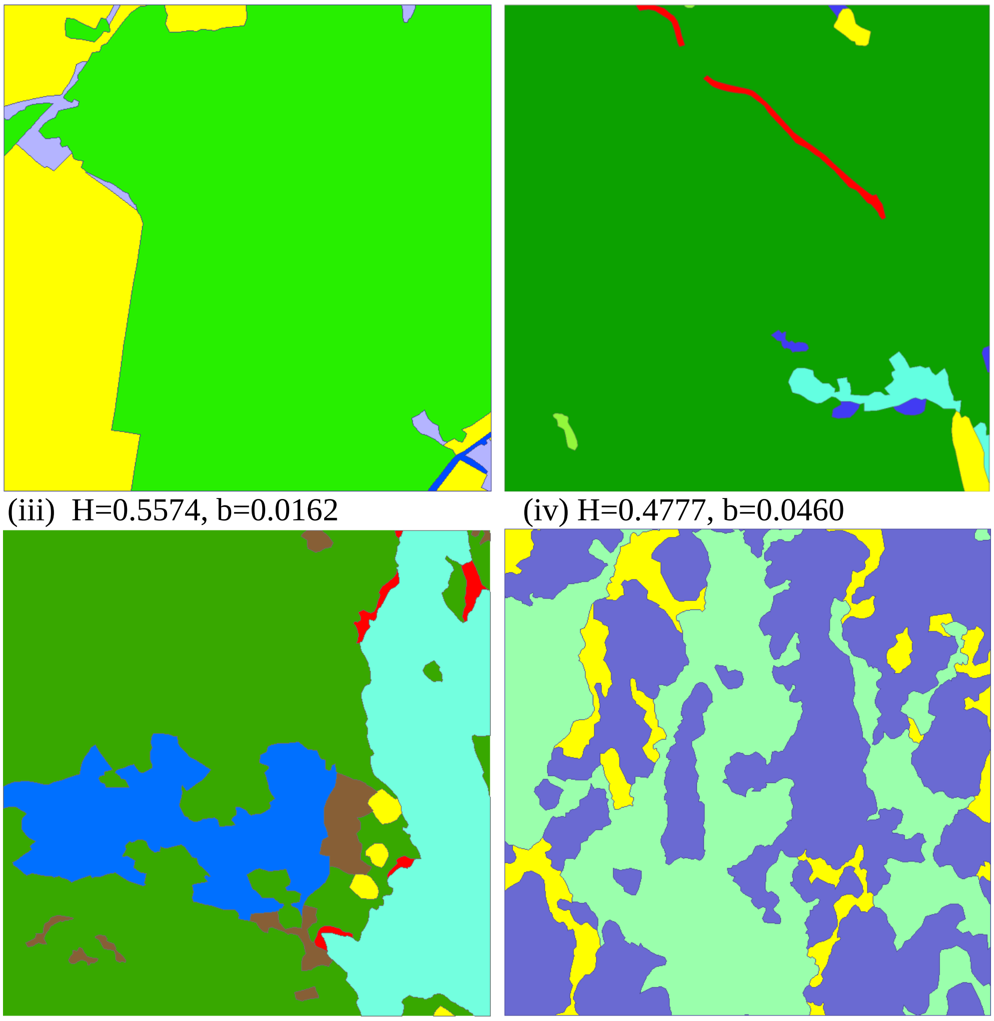


Figure 3. Four landscapes illustrating the necessity to consider the Shannon diversity index when interpreting the unconditional contagion function as a measure of fragmentation (seven classes system). Category (i) has low Shannon and low *b*, (ii) has low Shannon and high *b*, (iii) has high Shannon and low *b*, and (iv) has high Shannon and high *b*.

In Figure 3, the landscapes to the right (high values of *b*) comprise smaller and more scattered polygons within a class than the landscapes to the left. However, landscape (iii) might visually be considered as more fragmented than landscape (ii) but that is an effect of the area proportions. In landscape (iii) the polygons of the non-dominant classes are in general larger and more compact than in landscape (ii).

By visual inspection of the 50 landscapes it was found that high values of *b* were found in landscapes where many patches of small classes were embedded in the dominant classes (as in figure (ii) above). Low values were consequently found when dominating classes contained few patches of smaller classes. By its definition, the unconditional contagion characterizes especially the pattern of the large classes.

Simulated landscapes were used in order to assess the contagion function in extreme cases. Four of these simulated landscapes and some of their properties are shown in Figure 4 and Table 3. There are four classes within all the simulated landscapes. In simulated landscape SL_1 the pattern is completely random; SL_2 is simulated by Gibbs sampler with the classes 3 and 4 highly repellent to class 1;

SL_3 has classes simulated conditionally on the classes of the west and north neighbours, creating a trend, and with classes 3 and 4 repellent to class 1; and SL_4 has two dominating classes dividing the landscape in two halves, with some random fragments of two other classes. The area proportions of the classes are equal or almost so for SL_1 and SL_3. In general the fit of the proxy function for the unconditional contagion is fairly good even for these extreme simulated landscapes.

Table 3. Unconditional contagion and proxy function parameters for four simulated landscapes. For a comparison with the real landscapes with side length 1 km the values of *b* should be multiplied by 0.512.

Simulated landscape (SL)	*c*	*a*	*b*	*H*	*SE*
1	0.0025	0.4955	0.2054	1	0.0062
2	0.1423	0.3956	0.0948	0.9000	0.0088
3	0.0034	0.4507	0.0424	0.9940	0.0166
4	0.2024	0.3601	0.0220	0.8471	0.0165

c, *a*, and *b* are contagion function parameters, *H* is Shannon diversity, *SE* is standard deviation around the function.

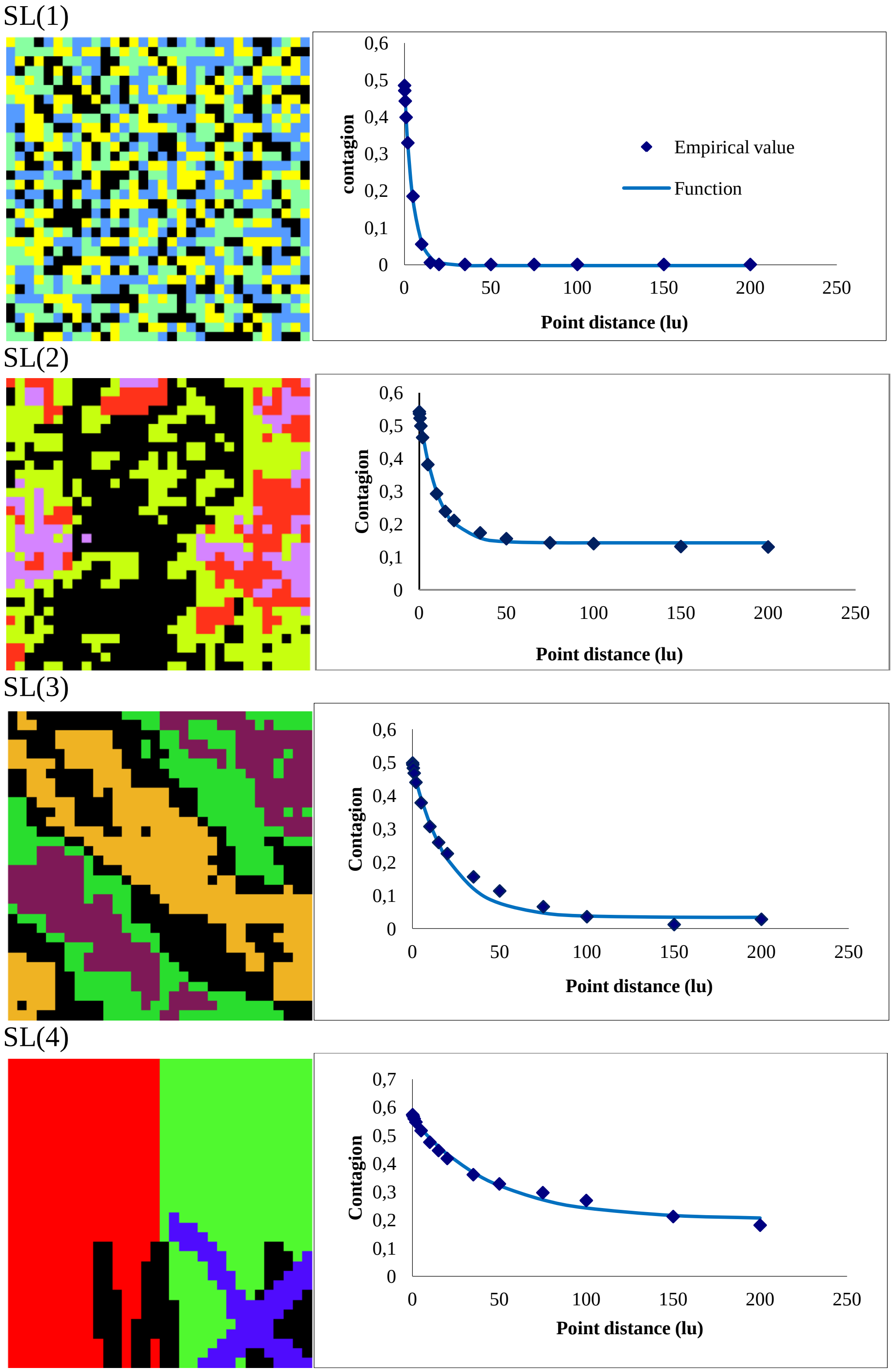


Figure 4. Illustration of four simulated landscapes and their unconditional contagion values and proxy functions. Side length of landscapes is 512 length units.

For the conditional contagion no simple proxy function was found, but the contagion function itself carries some qualitative information. The functions were classified into three categories. While one category, the most common, had a behavior similar to the unconditional contagion, two others showed different behavior. In Fig.5 the conditional contagion for a typical example of each of the two other categories is shown. For landscape (I) the function has a minimum value at a distance of about 30 m while for landscape (II) it decreases monotonically. For landscape I there is one dominating class, say class *j*, and three very small and fragmented ones. For a small class, say class *i*, we have $p_{i/i}(d) = 1$ at very short distances; for moderate distances both $p_{i/i}(d)$ and $p_{j/i}(d)$ are between 0 and 1, while $p_{j/i}(d)$ tends to 1 at longer distances. This implies that the contribution from the small class *i* to the numerator in the definition of the conditional contagion equals 0 at $d = 0$, is negative for moderate distances and approaches 0 again for long distances. Due to this, and that the contributions from all classes are weighted equally, the conditional contagion will look like that of landscape (I) if the landscape contains several small (or oblong) and fragmented classes. The smaller the polygons of the small classes the shorter the distance to the minimum value. In landscape (II) there is no dominating class and the polygons in general have a compact form. For about one third of the remaining 48 landscapes the conditional contagion looks like either that of landscape (I) or (II). The rest of them are intermediate, some with a minimum value at longer distances than 30 m.

Functions of type (I) tend to have lower value of the Shannon index than those of type (II). This is also clear from property (iv) in the result section. Due to this there is a certain connection between the conditional and unconditional contagion for long distances. It was also found that the parameter *b* of the proxy function of the unconditional contagion is somewhat smaller for landscapes of type II than for the other. Otherwise the two kinds of contagion seem to carry different information.

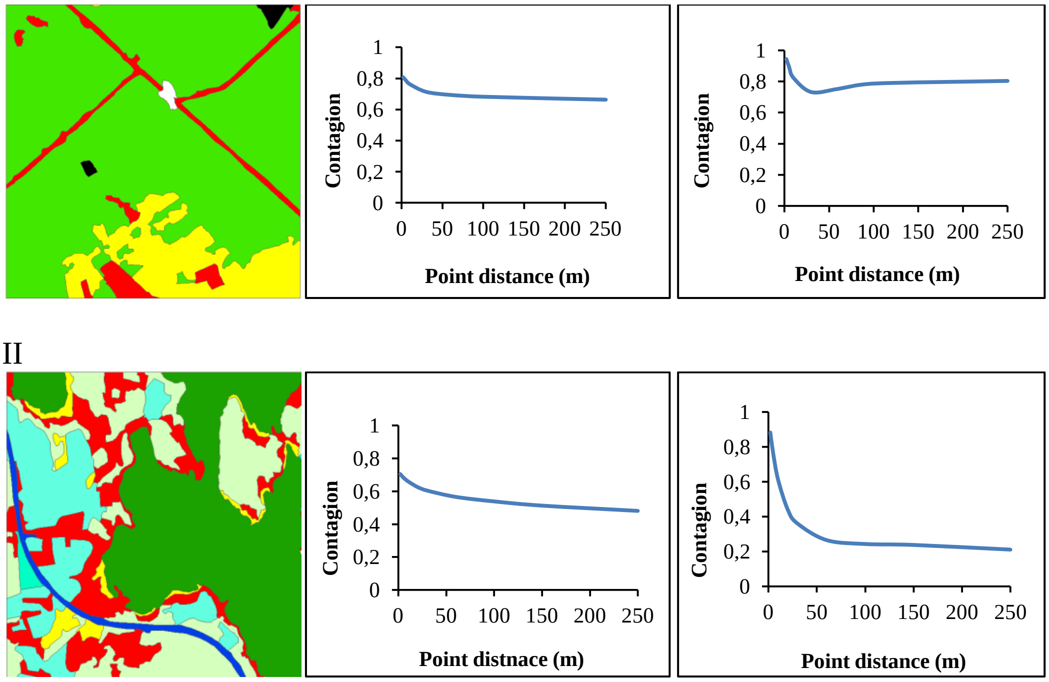


Figure 5. Comparison of contagion behavior at given landscape under unconditional and conditional definitions with 7 classification system

It is clear that a projection like contagion of a landscape onto a function (or a number) never can describe everything about the landscape. Nevertheless the two contagion functions tell us something.

The unconditional contagion emphasis the properties of the patches of the large classes; to what extent they contain nested patches of smaller classes or not. The conditional contagion tells us more about the fragmentation or clumping of the small classes and sometimes their polygons sizes.

The proposed contagion definition appears also to be a basis for sampling-based estimation of the contagion. If a non-delineated map is available it is possible to estimate the probabilities $p_{ij}(d)$ and $p_{j/i}(d)$ from a sample of point pairs at distance *d* in the landscape, where the $p_{ij}(d)$ is estimated by the relative frequency of points in classes *i* and *j*. The estimators $\hat{p}_{ij}(d)$ and $\hat{p}_{j/i}(d)$ are then inserted into the defining expressions (3) and (4) to obtain estimators of $\hat{C}_u(d)$ and $\hat{C}_c(d)$ of the contagion functions.

The comparison between the raster-based and vector-based values (see fig.5) indicates that almost the same continuous contagion functions could be obtained by rasterizing at “all” sizes as well. This requires mapped data and it is difficult to see how to estimate true rasterized maps from point sample data. However, the vector-based definition allows a sample-based estimation (at least in theory) and this can indirectly be applied with some approximation to a thought rasterized case.

Sample-based assessment of landscape metrics is recognized as an alternative to traditional wall-to-wall mapping in terms of cost-efficiency, and metrics can be derived without land cover/use map of the entire landscape (Corona et al., 2004; Ramezani et al., 2010). It would therefore be of interest to consider statistical properties of a contagion estimator for different designs such as systematic and random sampling for different patterns. Further, point sampling appears to be applicable in a gradient-based landscape model (McGarigal and Cushman, 2005), where landscape characteristics changes continuously and no distinct border is assumed between patches.

Acknowledgments: we are grateful to Prof. Göran Ståhl for constructive comments at SLU in Umeå, Sweden.

5. References

Bebber, D.P., Cole, W.G., Thomas, S.C., Balsillie, D., and Duinker, P., (2005). Effects of retention harvests on structure of old-growth Pinus strobus L. stands in Ontario: Forest Ecology and Management, v. 205, p. 91-103.

Cain, D.H., Riitters, K., and Orvis, K., (1997). A multi-scale analysis of landscape statistics: Landscape Ecology, v. 12, p. 199-212.

Corona, P., Chirici, G., and Travaglini, D., (2004). Forest ecotone survey by line intersect sampling: Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, v. 34, p. 1776-1783.

Fahrig, L., (2003). Effects of habitat fragmentation on biodiversity: Annual Review of Ecology Evolution and Systematics, v. 34, p. 487-515.

Forman, R.T.T., (1995). Land mosaics: the ecology of landscapes and regions: Cambridge, Cambridge University Press, 632 p.

Frohn, R.C., (1998). Remote sensing for landscape ecology: new metric indicators for monitoring, modeling, and assessment of ecosystems: Boca Raton, Lewis Publishers, 99 , 4 pl.- (i färg) p.

Gustafson, J.E., (1998). Quantifying landscape spatial pattern: What is the state of the art?: Ecosystems, v. 1, p. 143-156.

Haralick, R.M., Shanmugam, K., and Dinstein, I., (1973). Textural features for image classification: IEEE Transactions on Systems, Man, and Cybernetics, v. 3, p. 610-621.

Hargis, C.D., Bissonette, J.A., and David, J.L., (1998). The behavior of landscape metrics commonly used in the study of habitat fragmentation: Landscape Ecology, v. 13, p. 167-186.

Hernandez-Stefanoni, J.L., (2005). Relationships between landscape patterns and species richness of trees, shrubs and vines in a tropical forest: Plant Ecology, v. 179, p. 53-65.

Hunsaker, C.T., O'Neill, R.V., Jackson, B.L., Timmins, S.P., Levine, D.A., and Norton, D.J., (1994). Sampling to characterize landscape pattern: Landscape Ecology, v. 9, p. 207-226.

Jenness, J.S., (2004). Calculating landscape surface area from digital elevation models: Wildlife Society Bulletin, v. 32, p. 829-839.

Li, H., and Reynolds, J., (1993). A new contagion index to quantify spatial patterns of landscapes: Landscape Ecology, v. 8, p. 155-162.

Li, X.Z., He, H.S., Bu, R.C., Wen, Q.C., Chang, Y., Hu, Y.M., and Li, Y.H., (2005). The adequacy of different landscape metrics for various landscape patterns: Pattern Recognition, v. 38, p. 2626-2638.

Lunetta, R.S., Congalton, R.G., Fenstermaker, L.K., Jensen, J.R., Mcgwire, K.C., and Tinney, L.R., (1991). Remote-Sensing and Geographic Information-System Data Integration - Error Sources and Research Issues: Photogrammetric Engineering and Remote Sensing, v. 57, p. 677-687.

McGarigal, K., and Cushman, S.A., (2005). The gradient concept of landscape structure, *in* Wiens, J., and Moss, M., eds., Issues and perspectives in landscape ecology: Cambrideg, Cambrideg University press.

McGarigal, K., and Marks, E.J., (1995). FRAGSTATS: Spatial pattern analysis program for quantifying landscape pattern. General Technical Report 351. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.
NIJOS, (2001). Norwegian 3Q Monitoring Program: Norwegian institute of land inventory.
O'Neill, R.V., Krumme, J.R., Gardner, H.R., Sugihara, G., Jackson, B., DeAngelist, D.L., Milne, B.T., Turner, M., Zygmunt, B., Christensen, S.W., Dale, V.H., and Graham, L.R., (1988). Indices of landscape pattern: Landscape Ecology v. 1, p. 153-162.
Ramezani, H., and Holm, S., (2009). Sample based estimation of landscape metrics: accuracy of line intersect sampling for estimating edge density and Shannon's diversity . Environmental and Ecological Statistics (in press).
Ramezani, H., Holm, S., Allard, A., and Ståhl, G., (2010). Monitoring landscape metrics by point sampling: accuracy in estimating Shannon's diversity and edge density: Environmental Monitoring and Assessment v. 164, p. 403-421.
Ricotta, C., Corona, P., and Marchetti, M., (2003). Beware of contagion!: Landscape and Urban Planning, v. 62, p. 173-177.
Riitters, K.H., O'Neill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B., and Jackson, B.L., (1995). A factor-analysis of landscape pattern and structure metrics: Landscape Ecology, v. 10, p. 23-39.
Riitters , K.H., O'Neill , R.V., Wickham , J.D., and Bruce Jones , K., (1996). A note on contagion indices for landscape analysis
Landscape Ecology, v. 11, p. 197-202.
Schumaker, N.H., (1996). Using landscape indices to predict habitat connectivity: Ecology, v. 77, p. 1210-1225.
Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sture Sundquist, S., Christensen, P., Gallegos Torell , Å., Högström, M., Lagerqvist, K., Marklund, L., Nilsson, B., and Inghe, O., (2010). National Inventory of Landscapes in Sweden (NILS) - Scope, design, and experiences from establishing a multi-scale biodiversity monitoring system: Environmental Monitoring and Assessment (in press).
Turner, M.G., (1989). Landscape ecology: The effect of pattern on process: Annual Review of Ecology and Systematics, v. 20, p. 171-97.
—, (1990). Spatial and temporal analysis of landscape patterns: Landscape Ecology, v. 4, p. 21-30.
Turner, M.G., Gardner, R.H., and O'Neill, R.V., (2001). Landscape ecology in theory and practice : pattern and process: New York, Springer, xii, 401 p.
United States Environmental Protection Agency, (1994). Landscape Monitoring and Assessment Research Plan. EPA 620/R-94/009, Office of Research and Development, Washington, DC.
Wade, T.G., Wickham, J.D., Nash, M.S., Neale, A.C., Riitters, K.H., and Jones, K.B., (2003). A comparison of vector and raster GIS methods for calculating landscape metrics used in environmental assessments: Photogrammetric Engineering and Remote Sensing, v. 69, p. 1399-1405.
Wickham, J.D., and Riitters, K.H., (1995). Sensitivity of landscape metrics to pixel size: International Journal of Remote Sensing, v. 16, p. 3585-3594.

Wickham, J.D., Riitters, K.H., ONeill, R.V., Jones, K.B., and Wade, T.G., (1996). Landscape 'contagion' in raster and vector environments: International Journal of Geographical Information Systems, v. 10, p. 891-899.
Wiens, J.A., Schooley, R.L., and Weeks, R.D., (1997). Patchy landscapes and animal movements: do beetles percolate?: OIKOS, v. 78, p. 257-264.
Wu, J., Shen, W.J., Sun, W.Z., and Tueller, P.T., (2002). Empirical patterns of the effects of changing scale on landscape metrics: Landscape Ecology, v. 17, p. 761-782.

Appendix 1. The behavior of the contagion functions for distances close to 0.
For very short distances *d* the second point of a random pair tends to fall into the same polygon as the first point. For the second point to fall in a different polygon the first point must fall within a distance *d* from an edge. If so, consider the circle with the first point as center and radius *d*. The probability that the second point belongs to an adjacent polygon equals the proportion of the circumference of the circle that is outside the polygon of the first point. The expected value of that proportion equals $1/\pi$ (derived as the integral of $2\cdot\arccos(t/d)$ with t from 0 to 1, divided by 2π). Thus if the first point belongs to a class *i* polygon, the probability that the second point belongs to a class *j* polygon approximately equals

$$p_{j/i}(d) \approx d\cdot e_{ij}/(a_i\cdot\pi) \qquad \text{(A1)}$$

where e_{ij} is the edge length between the classes and $a_i \approx p_i(d)$ is the area of the class *i*. The numerator equals the area close to the edge. Effects of boundaries and polygon corners are neglected.

By summing Eq. (A1) over *j* we get

$$p_{i/i}(d) \approx 1 - d\cdot e_i/(a_i\cdot\pi) \qquad \text{(A2)}$$

where e_i is the total edge length of class *i* polygons.
Inserting this in the definition of the conditional contagion function, using the approximation $\ln(1+x)\approx x$ for small *x* and neglecting terms of second order of *d* we obtain (after some simplification)

$$\sum_{i=1}^{s}\sum_{j}^{s} p_{j/i}\cdot\ln(p_{j/i}) \approx \frac{d\ln(d)}{\pi}\sum_{i=1}^{s} e_i/a_i + \frac{d}{\pi}\sum_{i=1}^{s}\sum_{i=1}^{s} e_{ij}^{'}\ln(e_{ij}^{'}) - \frac{d(1+\ln(\pi))}{\pi}\sum_{i=1}^{s} e_i/a_i$$
(A3)

where $e_{ij}^{'} = e_{ij}/a_i$ and $e_{ii} = 0$

Hence, for the conditional function, we obtain the difference ratio

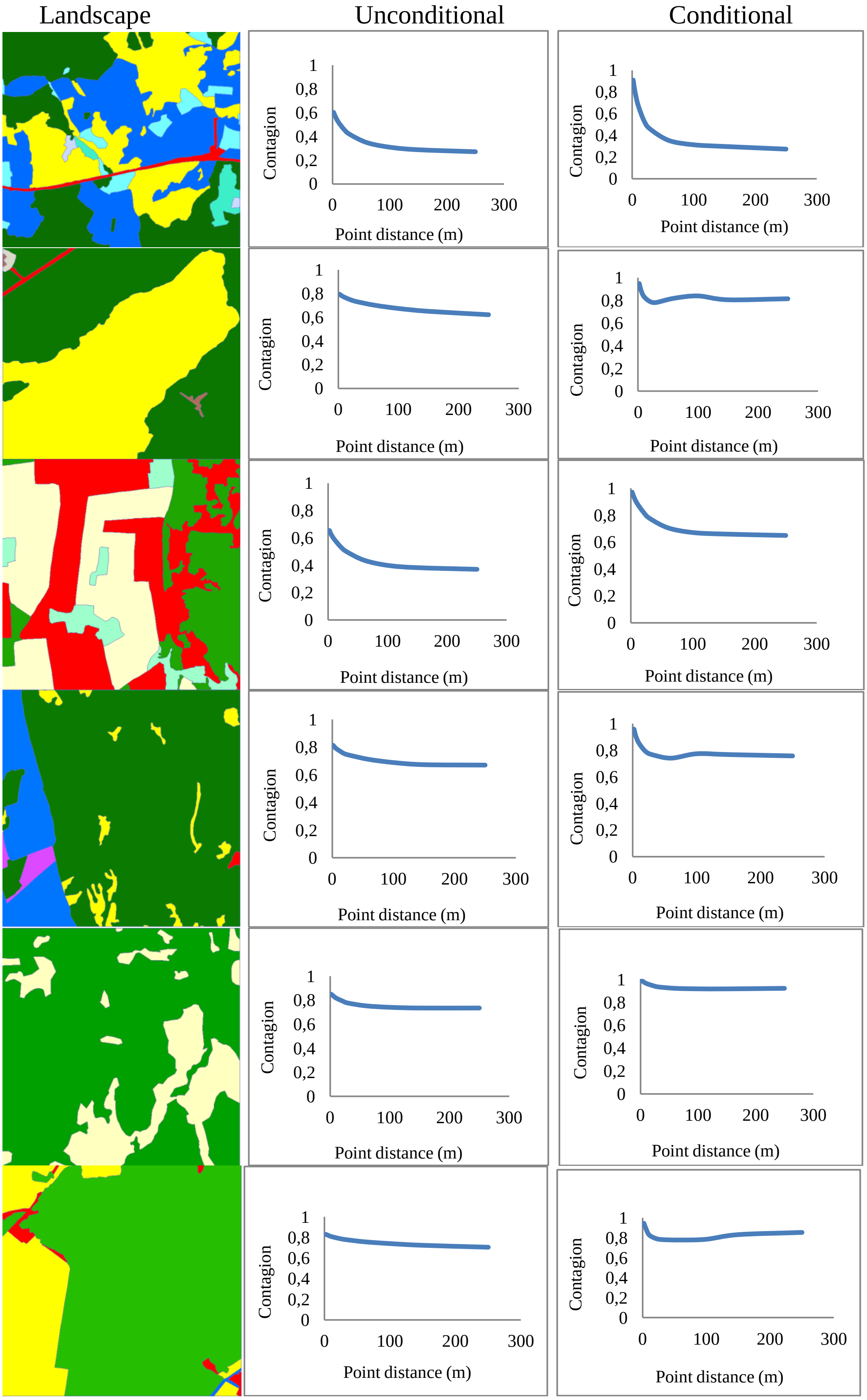
$$\left(C_c(d) - C_c(0)\right)/d \approx \alpha\cdot\ln(d) + \beta \qquad \text{(A4)}$$

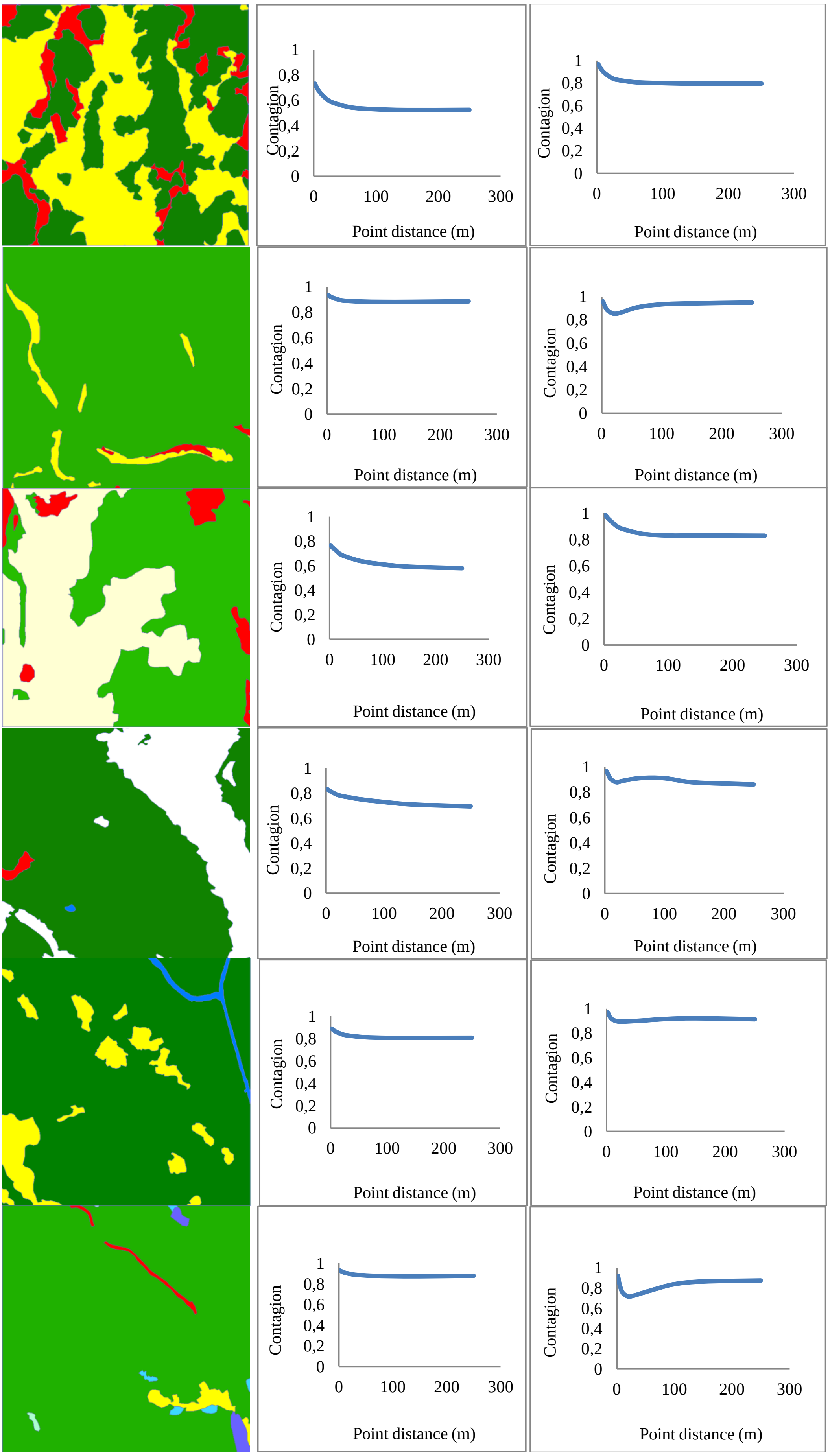
where $\alpha = \frac{1}{\pi s\ln(s)}\sum_{i=1}^{s} e_i/a_i$ and β are constants.
From (A4) it follows that $C_c(d)$ has no finite (right hand) derivative at $d = 0$

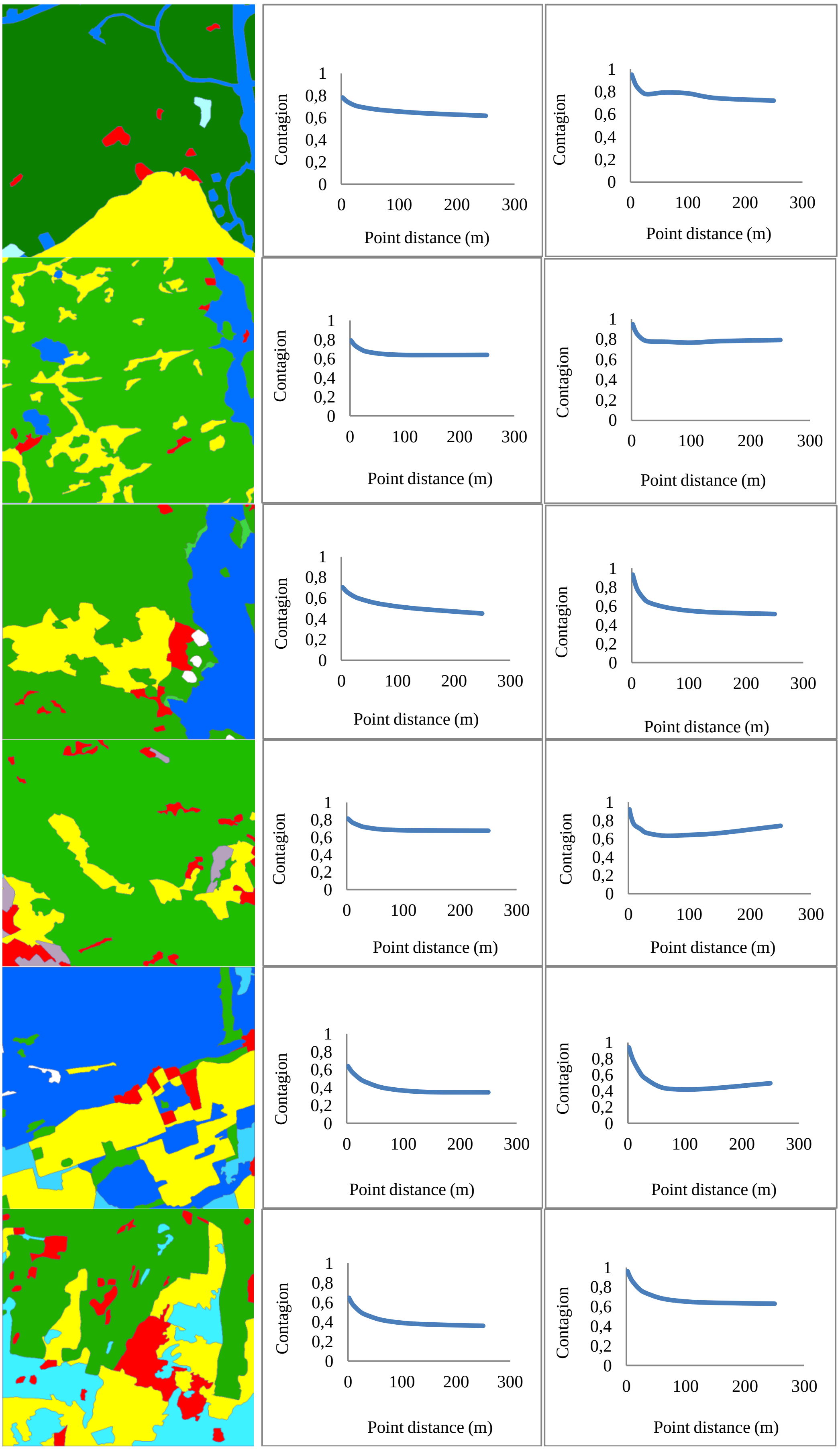
since $\ln(d)$ tends to negative infinity when d tends to 0.

For the unconditional contagion function the same argument gives the same kind of expression for the difference ratio, but with $\alpha = \frac{1}{2\pi \ln(s)} \sum_{i=1}^{s} e_i / a_i$ and a different value for the constant β

Appendix 2. Example of 18 landscapes and their unconditional and conditional contagion




Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)


Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)
Contagion
1
0,8
0,6
0,4
0,2
0
0
100
200
300
Point distance (m)