Landscape Ecology vol, 11 no. 4 pp 197–202 (1996)
SPB Academic Publishing bv, Amsterdam

A note on contagion indices for landscape analysis

Kurt H. Riitters[1]*, Robert V. O'Neill[2], James D. Wickham[1] and K. Bruce Jones[3]
[1]Tennessee Valley Authority, Historic Forestry Building, 17 Ridgeway Road, Norris, TN 37828, USA; [2]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA; [3]United States Environmental Protection Agency, Las Vegas, NV 89193, USA

Keywords: Spatial pattern, image texture, information index, computation, statistics

Abstract

The landscape contagion index measures the degree of clumping of attributes on raster maps. The index is computed from the frequencies by which different pairs of attributes occur as adjacent pixels on a map. Because there are subtle differences in the way the attribute adjacencies may be tabulated, the standard index formula may not always apply, and published index values may not be comparable. This paper derives formulas for the contagion index that apply for different ways of tabulating attribute adjacencies – with and without preserving the order of pixels in pairs, and by using two different ways of determining pixel adjacency. When the order of pixels in pairs is preserved, the standard formula is obtained. When the order is not preserved, a new formula is obtained because the number of possible attribute adjacency states is smaller. Estimated contagion is also smaller when each pixel pair is counted twice (instead of once) because double-counting pixel adjacencies makes the attribute adjacency matrix symmetric across the main diagonal.

Introduction

To landscape ecologists, "contagion" refers to the degree to which mapped attributes are clumped into patches of the same attribute class. Classical definitions of the contagion index (O'Neill *et al.* 1988; Li and Reynolds 1993) consider the observed pairing of attributes on raster maps in comparison to the case of random pairing when attribute frequencies are equal. The contagion index is used often in the literature despite the limitations of using single-valued indices to represent complicated map patterns. One reason is that the index seems to be an effective summary of overall clumpiness on maps (Turner 1989). Another reason may be that in typical map collections, the contagion index is highly correlated with indices of attribute diversity and dominance (Riitters *et al.* 1995; Cain *et al.* in press) and thus may be a surrogate for those important pattern indices (O'Neill *et al.* in press).

Modifications of the contagion index have been suggested to capture more of the information about pattern complexity that is necessarily lost when using a single-valued index. Gardner and O'Neill (1991) derived attribute class-specific contagion estimators for different null hypotheses of pattern. Pastor and Broschart (1990) proposed the electivity index to measure the contagion of each pair of attribute classes separately. Cressie (1993, Chapter 7) describes other multi-parameter statistical models that are used to model spatial adjacency on raster maps. In general, the analysis can be made more realistic and the sensitivity to real pattern differences can be improved by adding additional parameters.

Sometimes, however, it is desirable to use single-valued indices. For example, multivariate state spaces with dimensions corresponding to contagion, fractal dimension, and patch shape categorize landscapes by simultaneously considering several aspects of pattern (O'Neill *et al.* in press). In these cases, a premium is placed on economy of parameters for any one dimension, and the ideal case is that of single-valued indices.

---

*Corresponding author.

To calculte contagion, the pairing of attributes is usually first summarized in an "attribute adjacency" or "co-occurrence" matrix which shows the frequency at which different pairs of attributes appear side-by-side on the map (*e.g.*, Haralick et *al.* 1973). Further analyses depend on the study objectives and approach to determine the model complexity and the number of parameters that will be used to summarize the attribute adjacency matrix.

At first glance, it appears to be a simple matter to construct a table of attribute adjacency frequencies – one simply counts the number of times that each pair of attribute classes is adjacent on the map. But several procedures are in use, including counting each pair of pixels (picture elements) twice instead of once (*e.g.*, Haralick et *al.* 1973; Musick and Grover 1991). While there are valid reasons for adopting a particular procedure, the details are not usually given when a contagion index is reported. This would not be a problem if the computing formulas were the same for these different procedures, but they are not. This makes it difficult, if not impossible, to compare published contagion values from different studies.

Contagion indices from the entropy of attribute adjacency

Consider a map made up of pixels, each with an attribute class, c_k $(k = 1...t)$. Each pixel pair (i,j) has an attribute adjacency type A_q $(q = 1...n_a)$ corresponding to the attribute class values (c_i,c_j) of the pixels. When the order of pixels in a pair is preserved, $n_a = t^2$, otherwise $n_a = (t^2 + t)/2$. For example, if the map has t = 2 attribute classes and order is preserved, then there are four types of attribute adjacencies ($A_1 = \{1,1\}$, A, $= \{1,2\}$, A, $= \{2,1\}$, A, $= \{2,2\}$). When pixel order is not preserved, $n_a = 3$, and the attribute adjacency types are $\{1,1\}$, $\{1,2\}$, and $\{2,2\}$.

Suppose there is a position operator **(O)**, a rule by which a pixel is considered to be adjacent to another pixel (Gonzalez and Woods 1992). To help distinguish among different ways of calculating contagion, define **O,** as "one pixel below or one pixel to the right", and define **O,** as "one pixel in each of the four cardinal directions". The first rule counts each immediately-adjacent pixel pair once, and the second rule counts each pair twice. It may be easier to think of these rules as counting "edges" as opposed to counting "pixel pairs".

Let n_p be the total number of pixel pairs in the map that satisfy **O,** and let F be a matrix whose element f_q is the frequency of pixel pairs satisfying **O** that are of attribute adjacency type q. The proportion $p_q = f_q / n_p$ estimates the probability that a pair of pixels satisfying **O** have attribute adjacency type q.

The entropy (E) of attribute adjacency is calculated by summing the observed proportions over the number of attribute adjacency types or "states" (equation [1]).

$$E = -\sum_{q=1}^{n_a} p_q \ln(p_q) \qquad (1)$$

The maximum possible entropy is also of interest because contagion indices are typically computed by scaling a statistic to its maximum possible value under the null hypothesis that there is no contagion (O'Neill et *al.* 1988; Li and Reynolds 1993). In general, entropy is maximum when all possible states are equally likely; that is, when all attribute adjacency types occur with probability $1/n_a$. Thus, when the pixel order is preserved, maximum entropy (E_{max}) **is** given by equation [2] (note that $n_a = t^2$).

$$E_{max}(ordered) = -\sum_{q=1}^{n_a} \frac{1}{t^2} \ln\left(\frac{1}{t^2}\right) \qquad (2)$$

$$= 2\ln(t)$$

When the pixel order is not preserved, E_{max} has the value shown by equation [3] (here, note that $n_a = (t^2 + t)/2$).

$$E_{max}(unordered) = -\sum_{q=1}^{n_a} \frac{2}{t^2+t} \ln\left(\frac{2}{t^2+t}\right) \qquad (3)$$

$$= \ln(t^2+t) - \ln(2)$$

Li and Reynolds (1993) proposed a ratio estimator of contagion that is scaled to the interval [0,1] (equation [4]). This estimator yields a zero value when the entropy of attribute adjacency is at the maximum, that is, when there is no evidence for "contagion".

$$C = 1 - \frac{E}{E_{\max}} \qquad (4)$$

Substituting E_{max} when pixel order is preserved (equation [2]), the contagion formula is equation [5].

$$C(ordered) = 1 + \frac{\sum_{q=1}^{n_a} p_q \ln(p_q)}{2\ln(t)} \qquad (5)$$

This estimator is computationally equivalent to a ratio estimator of contagion defined by equation number 23 in Li and Reynolds (1993). When pixel order is not preserved, the formula for contagion substitutes equation *[3]* into equation [4] to yield equation [6].

$$C(unordered) = 1 + \frac{\sum_{q=1}^{n_a} p_q \ln(p_q)}{\ln(t^2+t)-\ln(2)} \qquad (6)$$

Now it is possible to explore how the different rules for defining pixel adjacency affect apparent contagion. Suppose that pixel order is preserved, so that equation [5] is appropriate. In comparison to rule O_a, the "double-counting" rule O, tends to even out the proportions of different attribute adjacency types (the values of p_q). To see this, recognize that the pixel pair (c_i,c_j) is also counted as the pair (c_j,c_i), which "smooths" the attribute adjacency matrix by making it symmetric across the main diagonal. Entropy is gained whenever information (in this case, information about asymmetry) is lost, and therefore the calculated value of contagion becomes smaller. Both values of contagion are "correct", even though they are different, because both position operators (O, and O,) are justifiable. The important point is that position operators must be specified if contagion values are to be compared.

Pixel order is arbitrary under rule O,, and so rule O, may be attractive for the very reason that it obscures any ordering. The "lost information" may not be considered worth preserving; increased entropy would never be noticed if pixel pairs were always double-counted. Yet there is another way to obscure pixel order without double-counting, that is, by using rule O, but not preserving the order of pixels when each pair is counted. This approach avoids the artificial "smoothing" of the attribute adjacency matrix. It is also consistent with the motivation of an entropy calculation which requires that each "state" be distinguishable from all others.

Examples

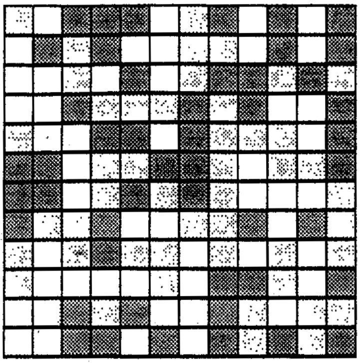
Figure 1 shows three artificial raster maps with different degrees of clumping or contagion. Each map is made up of 144 pixels and each pixel is one of three possible attribute classes ($c_k = \{1,2,3\}$). The frequency of each attribute class is 48 in each map. The frequencies and proportions of different attribute adjacency types are shown for three "cases" for each map. In case 1, rule O, (single-counting) is used and the order of pixel attributes is preserved. In case 2, rule O, is used but the order of attributes is not preserved (hence the existence of "undefined" states in this case). Rule O, (double-counting) is used in case 3. The total number of pixel pairs counted is 264 for cases 1 and 2, and 528 for case *3*.

The calculated values of entropy and contagion for the different scenarios are shown in Table 1. For each map, the contagion value **is** largest for case 1 and smallest for case 2. Case *3* yields an intermediate value which is closer to the value obtained for case 1. These examples demonstrate that, for the same map, the calculated value of contagion can vary by an order of magnitude among the different ways of counting and tabulating pixel pairs, even if the same basic formula (equation [4]) is used.

What does contagion *mean?*

Contagion usually implies that pixels having the same attribute class tend to be adjacent. The contagion indices "work" when clumping is present because they are affected by the relatively higher frequencies of the ***t*** same-class attribute pairs (c_i,c_j) where $i = j$. But the indices are equally affected by relatively higher frequencies of adjacency between ***any*** two attribute class pairs (c_i,c_j), even when $i \neq j$. This means, for example, that relatively frequent pairing of "streams" and "riparian vegetation" will increase the calculated value of contagion even in the absence of "real" clumping (of just one cover

MAP 1: LOW CONTAGION



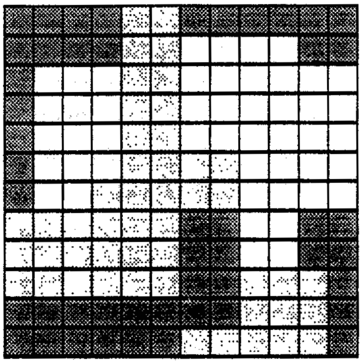
ATTRIBUTE ADJACENCY FREQUENCIES

	CASE 1				CASE 2				CASE 3		
	▓	░	□		▓	░	□		▓	░	□
▓	25	33	25	▓	25	64	56	▓	50	64	56
░	31	30	30	░	----	30	57	░	64	60	57
□	31	27	32	□	----	----	32	□	56	57	64

ATTRIBUTE ADJACENCY PROPORTIONS (X 1000)

	CASE 1				CASE 2				CASE 3		
	▓	░	□		▓	░	□		▓	░	□
▓	95	125	95	▓	95	242	212	▓	95	121	106
░	117	114	114	░	----	114	216	░	121	114	108
□	117	102	121	□	----	----	121	□	106	108	121

MAP 2: MODERATE CONTAGION



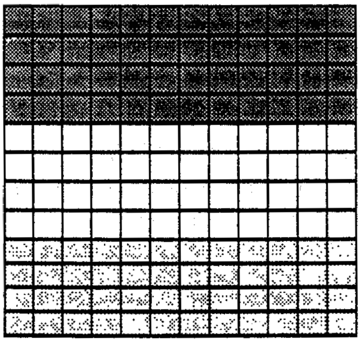
ATTRIBUTE ADJACENCY FREQUENCIES

	CASE 1				CASE 2				CASE 3		
	▓	░	□		▓	░	□		▓	░	□
▓	57	9	16	▓	57	25	21	▓	114	25	21
░	16	69	6	░	----	69	18	░	25	138	18
□	5	12	74	□	----	----	74	□	21	18	148

ATTRIBUTE ADJACENCY PROPORTIONS (X 1000)

	CASE 1				CASE 2				CASE 3		
	▓	░	□		▓	░	□		▓	░	□
▓	216	34	61	▓	216	95	80	▓	216	47	40
░	61	261	23	░	----	261	68	░	47	261	34
□	19	45	280	□	----	----	280	□	40	34	280

MAP 3: HIGH CONTAGION



ATTRIBUTE ADJACENCY FREQUENCIES

	CASE 1				CASE 2				CASE 3		
	▓	░	□		▓	░	□		▓	░	□
▓	80	0	12	▓	80	0	12	▓	160	0	12
░	0	80	0	░	----	80	12	░	0	160	12
□	0	12	80	□	----	----	80	□	12	12	160

ATTRIBUTE ADJACENCY PROPORTIONS (X 1000)

	CASE 1				CASE 2				CASE 3		
	▓	░	□		▓	░	□		▓	░	□
▓	303	0	45	▓	303	0	45	▓	303	0	23
░	0	303	0	░	----	303	45	░	0	303	23
□	0	45	303	□	----	----	303	□	23	23	303

***Fig.** 1.* Attribute adjacency frequencies and proportions for three example maps with low, moderate, and high contagion. CASE 1 counts each pair of adjacent pixels once (rule O_a, see text for explanation), and preserves their order. CASE 2 counts each pair once but does not preserve order. CASE 3 counts each pair of adjacent pixels twice (rule O,, see text). In each adjacency matrix, the left margin shows c_i and the top margin shows c_j in the pixel pair (c_i,c_j). The contagion indices calculated for these examples are given in Table 1. Note: “—” under CASE 2 means that adjacency state is undefined.

Table 1. **Landscape indices of contagion for the three example maps shown in Figure 1. The calculations are based on the number of different adjacency states, and proportions of different adjacency types as shown in Figure 1. The equation number from the text is in brackets next to each index for map 1.**

Map number	**Landscape index**	**CASE number (from Figure 1)**		
		1	**2**	**3**
1	Entropy (E)	[1] 2.19274	[1] 1.72994	[1] 2.19435
	E_{max}	[2] 2.19722	[3] 1.79176	[2] 2.19722
	Contagion	[5] 0.00204	[6] 0.03450	[5] 0.00131
2	Entropy (E)	1.79583	1.64651	1.81289
	E_{max}	2.19722	1.79176	2.19722
	Contagion	0.18268	0.08107	0.17492
3	Entropy (E)	1.36446	1.36446	1.43241
	E_{max}	2.19722	1.79176	2.19722
	Contagion	0.37901	0.23848	0.34808

type). For example, in the map with low contagion in the second example scenario (case **2,** above), the calculated contagion value is larger than expected for this reason.

The indices are also affected by variation in attribute class frequencies (Li and Reynolds **1993).** Recall that the derivations of all maximum values assumed an equality of attribute frequencies. When some attributes are relatively more common than others, then some types of attribute adjacencies are necessarily more frequent than other types. The maximum entropy (or minimum contagion) under the null model can never be realized; there must be some amount of contagion, and this will vary with attribute frequencies. This effect was demonstrated in Gustafson and Parker's **(1992)** simulation study. Cases can be made for considering this to be "real" contagion, an artifact of unequal-probability sampling, or an inappropriate application of the contagion index.

In summary, when the contagion index is small, it may be inferred that the attribute class frequencies are more or less equal, and that the frequencies of same-class adjacencies are about the same as the frequencies of different-class adjacencies. When the index is larger, it may due to a real tendency for clumping, perhaps as a result of variation in attribute class frequencies. Or it may be caused by a high frequency of adjacencies between two different classes.

If entropy (like angular second moment) measures overall image "texture" (*e.g.,* Haralick ***et* d.** **1973;** Musick and Grover **1991;** Gonzalez and Woods **1992),** then so must its contagion derivatives. An image with "coarse" texture typically displays a certain amount of clumpiness (possibly a result of unequal attribute class frequencies), whereas an image with "fine" texture does not (compare the example maps in Fig. **1).** These intuitive connections among pattern indices are supported by large empirical correlations between indices of attribute diversity, dominance, contagion, and texture (Riitters *et al.* **1995).**

Summary

Connectivity and texture are recurrent themes of spatial analysis in many fields including ecology, image processing, and statistics. As a result, there are many map-based measures of the tendency for attributes to clump or coalesce. The choice among them will depend upon circumstances, including the particular hypothesis, the format of available maps (*e.g.,* vector versus raster), the map parameters (*e.g.,* grain size and extent), and the scale of analysis (*e.g.,* pixel-level versus patch-level). The scope and depth of analysis are also important. Whereas summary indices are needed to study many aspects of pattern simultaneously (*e.g.,* fragmentation, patch compactness, and fractal dimension), more complicated models are needed to partition the detailed information from any one summary index. While no single measure can possibly be appropriate in all circumstances, it is relevant to ask whether the selected method yields the desired information.

The standard computing formula for contagion only applies when a particular method is used to reduce raster map data to an attribute adjacency matrix. Recognizing the need to compare contagion values obtained from the literature, it seems prudent to document the method used to construct the attribute adjacency table whenever a contagion index is reported. It may also be worthwhile to simply report the entire adjacency matrix, for this will enable others to fit more complicated contagion models such as those described by Gardner and O'Neill (1991), Pastor and Broschart (1990), and Cressie (1993).

An alternative, single-valued contagion estimator that is unaffected by the data reduction method is the sum of the main diagonal ("same-class") elements of the attribute adjacency matrix (Wickham and Riitters 1995). This index is defined as the proportion of all adjacencies that are same-class adjacencies. The value is not affected by adjacencies among different classes, and the same value is obtained for all of the ways that pixel pairs could be tallied.

Acknowledgemenls

The research described in this article has been funded in part by the United States Environmental Protection Agency through Interagency Agreement DW64935962-01-0 with the Tennessee Valley Authority, through Interagency Agreement DW89936104-01-0 with the Oak Ridge National Laboratory, and through Cooperative Agreement CR-819549-01-5 with the Desert Research Institute. This manuscript has been reviewed by the Environmental Protection Agency and approved for publication.

References

Cain, D.H., Riitters, K. and Orvis, K. 1996. A multi-scale analysis of landscape statistics. Landscape Ecology, in press.

Cressie, N.A.C. 1993. Statistics for Spatial Data. Revised Edition. John Wiley and Sons, New York.

Gardner, R.H. and O'Neill, R.V. 1991. Pattern, process, and predictability: the use of neutral models for landscape analysis. *In* Quantitative Methods in Landscape Ecology. Edited by M.G. Turner and R.H. Gardner. Springer-Verlag, New York.

Gonzalez, R.C. and Woods, R.E. 1992. Digital Image Processing. Addison-Wesley, Reading, MA.

Gustafson, E.J. and Parker, G.R. 1992. Relationships between landcover proportion and indices of landscape spatial pattern. Landscape Ecology 7: 101–110.

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

Li, H. and Reynolds, J.F. 1993. A new contagion index to quantify patterns of landscapes. Landscape Ecology 8: 155–162.

Musick, H.B. and Grover, H.D. 1991. Image textural measures as indices of landscape pattern. *In* Quantitative Methods in Landscape Ecology. Edited by M.G. Turner and R.H. Gardner. Springer-Verlag, New York

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. and Graham, R.L. 1988. Indices of landscape pattern. Landscape Ecology 1: 153–162.

O'Neill, R.V., Hunsaker, C.T., Timmins, S.P., Jackson, B.L., Jones, K.B., Riitters, K.H. and Wickham, J.D. 1996. Scale problems in reporting landscape pattern at the regional scale. Landscape Ecology, in press.

Pastor, J. and Broschart, M. 1990. The spatial pattern of a northern conifer-hardwood landscape. Landscape Ecology 4: 55-68.

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 10: 23–29.

Turner, M.G. 1989. Landscape ecology: the effect of pattern on process. Annu. Rev. Ecol. Syst. 20: 171–197.

Turner, M.G. and Gardner, R.H. (eds.). 1991. Quantitative Methods in Landscape Ecology. Springer-Verlag, New York.

Wickham, J.D. and Riitters, K.H. 1995. Sensitivity of landscape metrics to pixel size. Int. J. Remote Sensing 16: 3585–3594.