Landscape Ecology vol. 8 no. 3 pp 155-162 (1993)
SPB Academic Publishing **bv**, The Hague

A new contagion index to quantify spatial patterns of landscapes

Habin Li and James **F.** Reynolds
Department of Botany, Duke University, Durham, North Carolina 27706 ***USA***

Keywords: Contagion index, spatial pattern, probability, information index, landscape ecology

Abstract

A contagion index was proposed by O'Neill *et al.* (1988) to quantify spatial patterns of landscapes. However, this index is insensitive to changes in spatial pattern. We present a new contagion index that corrects an error in the mathematical formulation of the original contagion index. The error is identified mathematically. The contagion indices (both original and new) are then evaluated against simulated landscapes.

Introduction

A robust landscape index should quantify two distinct components of landscape diversity: composition and configuration (Li 1989). Composition refers to both the total number of patch types and their relative proportions in the landscape, whereas configuration refers to the spatial pattern of patches in the landscape. If an index only depicts composition its usefulness is limited because many existing indices do just that (e.g., Pielou 1975). And spatial configuration and its effects on landscape function are key topics of landscape ecology (Risser *et al.* 1984; Forman and Godron 1986).

A contagion index **(D_2)** was proposed by O'Neill *et al.* (1988; Also Turner and Ruscher 1988; Turner 1989; Turner *et al.* 1990a; Turner 1990b; Graham *et al.* 1991) to characterize landscape pattern:

$$D_2 = EE_{max} - EE \qquad (1)$$

where

$$EE_{max} = 2\, n \ln(n) \qquad (2)$$

$$EE = -\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} \ln(P_{ij}) \qquad (3)$$

where n is the total number of patch types in a landscape mosaic, P_{ij} is the probability of patch type i being adjacent to patch type j, and EE_{max} is the maximum of EE if there is an equal probability of any land-use type being adjacent to a randomly chosen point in the landscape (O'Neill *et al.* 1988). Note that EE is positive and has the same form as the information index (Pielou 1975). According to O'Neill *et al.* (1988), contagion measures the extent to which landscape elements are aggregated or clumped; higher values of contagion may result from landscapes with a few large, contiguous patches, whereas lower values generally characterize landscapes with many small patches.

In this paper, we show that an error exists in the definition of EE_{max} (Eq. 2), which is a critical problem because it affects the sensitivity of D_2 to changes in landscape configuration. We introduce an alternative contagion index that corrects this error and tends to distinguish between landscape composition and configuration.

Estimates of P_{ij}

We define a landscape mosaic as the distribution of n patch types among N pixels. Two methods can be used to estimate P_{ij} based on the definition that P_{ij} is the probability of patch type i being adjacent to patch type j. First, we assume that P_{ij} is the conditional probability that, given a pixel is of patch type i, one of its neighboring pixels belongs to patch type j, i.e.:

$$P_{ij} = P_{j/i} = N_{ij}/N_i \tag{4}$$

where $P_{j/i}$ is the conditional probability, N_{ij} the number of adjacencies (joins) between pixels of patch types i and j, and N_i the total number of adjacencies between pixels of patch type i and all patch types (including patch i itself). Thus, we have

$$\sum_{j=1}^{n} N_{ij} = N_i$$

and, therefore

$$\sum_{j=1}^{n} P_{ij} = 1 \tag{5}$$

O'Neill ***et al.*** (1988) did not provide information on how to estimate P_{ij}, but an equation similar to Eq. 4 was used to estimate P_{ij} (Robert O'Neill, pers. comm.).

Second, we can assume that P_{ij} is the probability that two randomly chosen adjacent pixels belong to type i and j, respectively, i.e.:

$$P_{ij} = P_i P_{j/i} \tag{6}$$

where P_i is the probability that a randomly chosen pixel belongs to patch type i (estimated by the proportion of patch type i), and $P_{j/i}$ is the same as in Eq. 4. Here,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} = 1 \tag{7}$$

(see Appendix for details). P_{ij} is similar to the probability used by information indices, especially Shannon's conditional entropy (Shannon and Weaver 1949) and Pielou's index of mosaic's spatial diversity (Pielou 1975).

The erroneous maximum term (EE_{max})

The critical error in the D_2 (Eq. 1) is that the maximum of EE does not equal EE_{max} defined by Eq. 2. Here, we present two proofs: one being an intuitive argument based on the property of an information index, and the other being a more formal mathematical proof.

The ***intuitive argument.*** According to information theory, EE should have its greatest value when each P_{ij} is equal to a constant (e.g., Shannon and Weaver 1949; Pielou 1975). Therefore, we can assume that EE is maximum when $P_{ij} = p$ (a constant). Substituting p for P_{ij} in Eq. **3,** we have

$$EE_{max} = -\sum_{i=1}^{n} \sum_{j=1}^{n} p \ln(p)$$

hence,

$$EE_{max} = -n^2 [p \ln(p)] \tag{8}$$

The p value is dependent on how P_{ij} is estimated. First, we assume that P_{ij} is defined by Eq. 4. Based on the same property of information theory mentioned above, we assume that all patch types are of even proportion, that is, $P_i = 1/n$ for all i, where P_i is the probability that a randomly chosen pixel belongs to type i. Thus, the necessary and sufficient condition for EE to be at its maximum value is

$$P_{ij} = p = 1/n \tag{9}$$

Substituting $1/n$ for p in Eq. **8,** we have

$$EE_{max} = -n^2[(1/n) \ln(1/n)]$$

hence,

$$EE_{max} \quad = n \ln(n) \tag{10}$$

Second, we assume that P_{ij} is defined by Eq. 6. Under the assumptions of the even proportion of each patch type and the completely random pattern, $P_{j/i}$ is independent of P_i and equal to l/n. Now, the necessary and sufficient condition for EE to be at its maximum value is

$$P_{ij} = p = 1/n^2 \tag{11}$$

Substituting $1/n^2$ for p in Eq. **8,** we have

$$EE_{max} \quad = - n^2[(1/n^2) \ln(1/n^2)]$$

and, hence

$$EE_{max} \quad = 2 \ln(n) \tag{12}$$

Eq. 10 differs from Eq. 2 by a factor of 2, and Eq. 12 by a factor of n. Thus, we conclude that EE_{max} defined in Eq. 2 is not the maximum of EE. Notice that our conclusion is not affected by the definition of P_{ij} in Eq. 3. The above reasoning is easily understandable, but lacks generality due to requirement of many assumptions.

The mathematical ***proof.*** Assuming that P_{ij} is expressed by Eq. 6, we give a formal mathematical proof that the maximum of EE is given by Eq. 12. Jensen's inequality is expressed as (Hardy *et al.* 1952):

$$\sum_{i=1}^{n} q_i f(x_i) \geq f(\sum_{i=1}^{n} q_i x_i) \tag{13}$$

where f is a convex, monotonic function in the range [a,b], x_i can be any value in the range [a,b], and q_i is positive with the condition

$$\sum_{i=1}^{n} q_i = 1$$

In Eq. 13, the equality holds only if the x_i's are equal for all i. Based upon the algebraic fact that any single summations can be re-arranged and then expressed as double summation (Li 1966), Jensen's inequality can be expanded and rewritten as:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} f(x_{ij}) \geq f(\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_{ij}) \tag{14}$$

given

$$\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} = 1 \tag{15}$$

Without any loss of generality, we can define

$$x_{ij} = P_{ij} \tag{16}$$

$$f(x_{ij}) = x_{ij} \ln(x_{ij}) \tag{17}$$

and

$$q_{ij} = 1/n^2 \tag{18}$$

Notice that the function in Eq. 17 is a convex, monotonic function in the range (0,1]. Substituting Eq. 7 and Eqs. 15–18 for the corresponding terms in Eq. 14 (i.e., the expanded Jensen's inequality), we have

$$- \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} \ln(P_{ij}) \leq 2 \ln(n) \tag{19}$$

(see Appendix for details). Notice that the left side term in Eq. 19 is EE. Therefore, we come to the same conclusion that the maximum value of EE is not expresses by Eq. 2, but by Eq. 12, provided that P_{ij} is given in Eq. 6. Similar proof can be given when we assume that P_{ij} is given in Eq. 4, and the result should confirm Eq. 10.

Our conclusion **is** supported by the work of Gatrell (1977). He proposes a complexity (or uncertainty) index, H_{map}:

$$H_{map} = - \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} \ln(P_{ij}) \tag{20}$$

where n is the number of colors, and P_{ij} is the joint probability of colors (or patches) i and j (Gatrell 1977). Gatrell's P_{ij} in Eq. 20 is similar to P_{ij} in

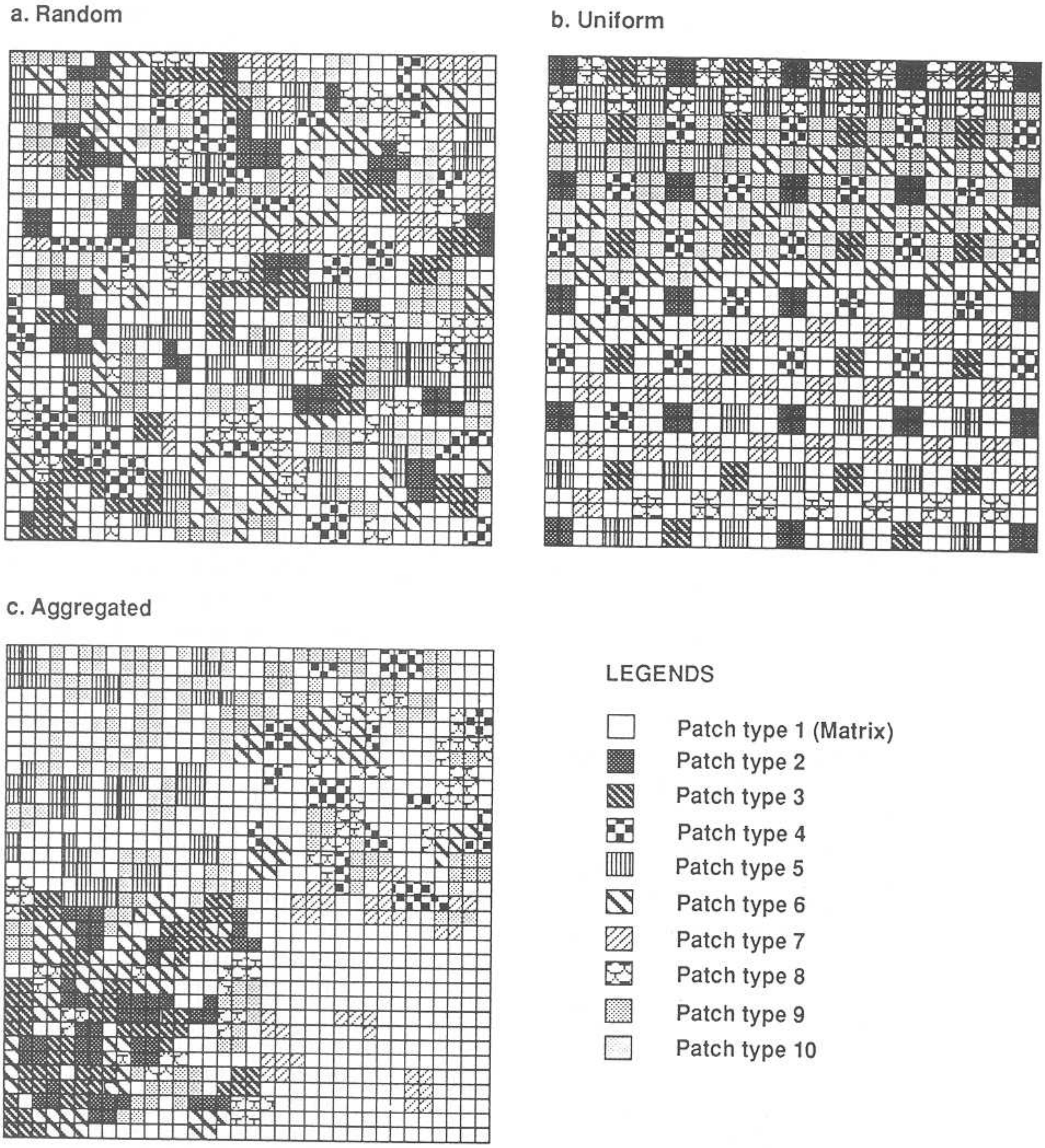

Fig. 1. Examples of the simulated landscape maps with 10 patch types for spatial configurations of (a) random (a random distribution of patches with randomly determined patch shapes, i.e, restricted random), (b) uniform (a uniform distribution of patches with strictly square patch shape), and (c) aggregated (an aggregated distribution of patches with regular patch shape). Note that a simulated map with 9 patch types can be obtained by deleting all pixels of patch type 10 and reassigning them to the matrix (i.e., patch type 1). Please refer to Li (1989) for more details about the simulation models.

Eq. 6, and H_{map} is similar to EE in Eq. 3. He has shown, using binary maps as examples, that the maximum of H_{map} is constrained by the marginal probabilities and differs with different sets of P_i's. He suggests that P_{ij} can be calculated by assuming that P_i's are statistically independent, that is, $P_{ij} = P_iP_j$. He also points out without proof that the maximum of H_{map} is "2 ln(n)", provided that $P_i = 1/n$.

An alternative contagion index

We propose a new contagion index:

$$RC = 1 - EE/EE_{max} \qquad (21)$$

where RC is the relative contagion, EE is defined in Eq. 3, and EE, is the maximum of EE defined by Eqs. 10 or 12, depending on the definition of P_{ij}. Eq. 21 has been used by authors who originally proposed D, (Robert O'Neill, pers. comm.). If P_{ij} is given by Eq. 4, then

$$RC1 = 1 + \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} \ln(P_{ij}) / n \ln(n) \qquad (22)$$

If P_{ij} is given by Eq. 6, then

$$RC2 = 1 + \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} \ln(P_{ij}) / 2 \ln(n) \qquad (23)$$

All terms on the right side of Eqs. 21, 22, and 23 are the same as defined above. Note that the correct maximum term (EE,) is used in the new relative contagion index given by Eq. 22 or 23 and that Eq. 1 has no upper limit, whereas contagion as defined in Eq. 21 is a relative index, ranging from 0 to 1. RC is in essence a function of an evenness index, EE/EE_{max}; thus, it has all the advantages and disadvantages of an evenness index (see Hurlbert 1971).

EE (Eq. 3) can be used alone as an index of spatial pattern (Robert O'Neill, pers. comm.). There is a strict, linear, inverse relationship between EE and RC; thus, EE has all the spatial information that RC has. The only other difference between the two is that RC is an evenness index, and is scaled to compare landscapes with different values of n (Turner *et al.* 1989).

Evaluation of contagion indices using synthetic landscapes

To evaluate and compare the performance of RC and D„ we generated a series of landscape maps with different components of composition and configuration. The maps were composed of the landscape matrix (empty = patch type 1) and one or more other patch types (patch types 2, 3, and up to 10); see Fig. 1). Following the algorithm described in Li (1989; Li *et al.* 1992), for each of three configurations ("random", "uniform", and "aggregated") nine maps were created by introducing one patch type to the landscape matrix at a time; each patch type (except patch type 1) covered 8% of the total area (see Fig. 1). For example, the first map had two patch types, 1 and 2 (covering 92% and 8% respectively); the second three types, 1, 2, and 3 (covering 84%, 8%, and 8% respectively); etc. Values of the three contagion indices as a function of different components of landscape composition and configuration is illustrated in Fig. 2.

D, fails to distinguish differences in the three spatial configurations (Fig. 2a). In theory, contagion has a spatial component in EE (Eq. 3), thus, changes in contiguity are reflected only by changes in EE. The failure of D, is because the inflated EE, (a function of the number of patch types) greatly outweighs EE (a function of the spatial configuration). This is also evidenced by the linear increase of D, with increase of the number of patch types. This increase of D_2 contradicts the definition of the contagion index, i.e., contagion should decrease as many small patches of different types are introduced to the landscape and as the landscape matrix is broken into many smaller pieces. These results suggest that the D, is not a good index of spatial configuration because it primarily reflects landscape composition (e.g., number of patch types) and is insensitive to spatial configuration.

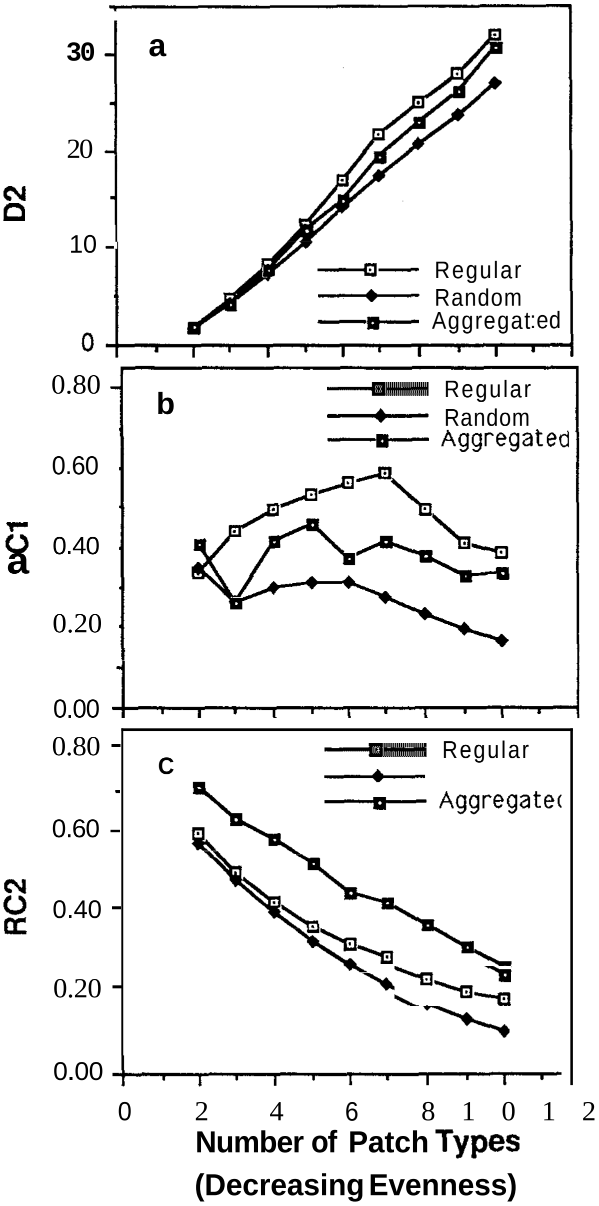

Fig. 2. Relationships between the contagion indices and the two controlled variables: spatial pattern and number of patch types. D_2, RCI, and RCI are the original contagion index (Eq. 1) and the new, relative contagion indices (Eqs. 22 and 23), respectively. See the text for details. Notice that there is a decreasing gradient of evenness of the proportions of patch types (from 0.29 to 0.09) along the x axis.

RC1 varies with different spatial configurations (Fig. 2b) but does not change much with the increase in the number of patch types, due to the method used to estimate P_{ij} (Eq. 4). RC1 is a good index if the objective is to compare spatial configuration of landscapes regardless of how many patch types they have; RC1 may not be appropriate to characterize the overall contiguity of landscapes because contagion should decrease with increase in the number of patch types, as discussed above. In contrast, RC2 distinguishes the three different spatial configurations and is sensitive to the increase in the number of patch types (Fig. 2c). Hence, RC2 is the only index that is quantifies both components of landscape diversity.

Acknowledgments

The authors wish to thank Dr. Robert O'Neill for many helpful discussions, Dr. Robert Gardner for his permission to use his percolation program, and Dr. Jialin Chen for reassuring that Eq. 14 holds. Drs. Janet Franklin, David Hilbert, Stuart Hurlbert, Robert O'Neill, and Peter Salamon reviewed earlier drafts of the manuscript and their comments are greatly appreciated. Funding for this research was provided by the US Department of Energy, Office of Health and Environmental Research to the DOE R4D program, contract no. DE-FG03-84ER60250. Some of the ideas presented have originated when the senior author was working with Research Work Unit 4356 at the USDA Forest Service, Pacific Northwest Research Station, under National Science Foundation, grant BSR85-14325.

References

Forman, R.T.T. and Godron, M. 1986. Landscape Ecology. John Wiley & Sons, New York.

Gatrell, A.C. 1977. Complexity and redundancy in binary maps. Geogra. Analysis 9: 29–41.

Graham, R.L., Hunsaker, C.T., O'Neill, R.V. and Jackson, **B.** 1991. Ecological risk assessment at the regional scale. Ecol. Appl. 1: 196–206.

Hardy, **G.H.,** Littlewood, J.E. and Polya, G. 1952. Inequalities. Second edition. Cambridge University Press, Cambridge.

Li, **H.** 1989. Spatio-temporal pattern analysis of managed forest landscapes: a simulation approach. Dissertation, Oregon State University, Corvallis, Oregon.

Li, H., Franklin, J.F., Swanson, F.J. and Spies, T.A. 1992. Developing alternative forest cutting patterns: a simulation approach. Landscape Ecology (in press).
Li, J.C.R. 1966. Numerical mathematics. Edwards Brothers, Ann Arbor, Michigan.
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. Landsc. Ecol. 1(3): 153–162.
Pielou, E.C. 1975. Ecological diversity. John Wiley & Sons, New York.
Risser, P.G., Karr, J.R. and Forman, R.T.T. 1984. Landscape ecology: directions and approaches. Illinois Natural History Survey Special Publ. No. 2. Illinois Natural History Survey, Champaign, Illinois.
Shannon, C.E. and Weaver, W. 1949. The mathematical theory of communication. University of Illinois Press, Urbana, Illinois.
Turner, M.G. 1989. Landscape ecology: The effect of pattern on process. Ann. Rew. Ecol. Syst. 20: 171–197.
Turner, m.G, 1990a. Spatial and temporal analysis of landscape pattern. Landsc. Ecol. 3(3/4): 153–162.
Turner, M.G. 1990b. Spatial and temporal analysis of landscape pattern. Landsc. Ecol. 3(3/4): 153–162.
Turner, M.G. 1990b. Landscape changes in nine rural counties of Georgia. Photogram. Engine. and Rem. Sens. 56: 379–386.
Turner, M.G. and Ruscher, C.L. 1988. Changes in landscape patterns in Georgia, USA. Landsc. Ecol. 1: 241–251.
Turner, M.G., O'Neill, R.V., Krummel, J.R., Gardner, R.H. and Milne, B.T. 1989. Effects of changing spatial scale on analysis of landscape pattern. Landsc. Ecol. 4(1): 21–30.

Appendix

1. Derivation of Eq. 7: $\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} = 1$

Assume that A_i is the total area (or total number of pixels) of patch type i and A the total area of the landscape. Then, the probability that a randomly chosen pixel belongs to patch type i can be estimated by the proportion of patch type i:

$$P_i = A_i / A \qquad (A.1)$$

Substituting Eqs. A.1 and 4 for P_i and $P_{j/i}$ in Eq. 6, we have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} P_i P_{j/i}$$

hence,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} (A_i / A)(N_{ij} / N_i)$$

and therefore

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} = \sum_{i=1}^{n} (A_i / A) \sum_{j=1}^{n} (N_{ij} / N_i) \qquad (A.2)$$

Note that, according to our definitions above, we have

$$\sum_{i=1}^{n} A_i = A \qquad (A.3)$$

and

$$\sum_{j=1}^{n} N_{ij} = N_i \qquad (A.4)$$

Substituting Eqs. A.3 and A.4 in Eq. A.2, we have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} = 1.$$

2. Derivation of Eq. 19: $-\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} \ln(P_{ij}) \leq 2\ln(n)$

We start with reconstructing the left and right sides of Jensen's inequality (Eq. 14) separately. Substituting Eqs. 16, 17, and 18 for the left side terms in Eq. 14, we obtain

$$\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} f(x_{ij}) = \sum_{i=1}^{n} \sum_{j=1}^{n} (1/n^2) P_{ij} \ln(P_{ij})$$

hence,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} f(x_{ij}) = (1/n^2) \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} \ln(P_{ij}) \qquad (A.5)$$

Substituting Eqs. 11, 15 and 17 for the right side terms in Eq. 14, we have

$$f(\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_{ij}) = \{\sum_{i=1}^{n} \sum_{j=1}^{n} (1/n^2) P_{ij} \{\ln [\sum_{i=1}^{n} \sum_{j=1}^{n} (1/n^2) P_{ij}]\}$$

hence,

$$f(\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_{ij}) = \{(1/n^2) \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij}\} \{\ln [(1/n^2) \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij}]\}$$

and therefore

$$f\left(\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij}\, x_{ij}\right) = (1/n^2)\ln(1/n^2) \qquad \text{(A.6)}$$

Replacing both sides in Jensen's inequality (i.e., Eq. 14) with the results from Eqs. A.5 and A.6, we obtain

$$(1/n^2)\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij}\ln(P_{ij}) \geq (1/n^2)\ln(1/n^2)$$

hence,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij}\ln(P_{ij}) \geq \ln(1/n^2)$$

and therefore

$$-\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij}\ln(P_{ij}) \leq 2\ln(n).$$