Complexity of Raster Spatial Adjacency Graphs

In a spatial adjacency graph (SAG) the graph nodes or vertices are nominal or categorical spatial entities—for example soil types, landform types, geological formations, or vegetation communities. Here we will just call them environmental units (EU). Any two nodes are connected (i.e., there exists link between them) if they are spatially contiguous. Thus, if EU types A and B at least sometimes occur adjacent to each other, they are connected, and if they never occur spatially adjacent to each other, there is no edge connecting A, B. In the SAGs analyzed by Phillips (2013), for instance, connectivity or spatial adjacency was based on whether soil types (taxa) occurred within the same mapping unit, or in mapping units with shared boundaries. In Phillips (2016) a SAG was constructed based on contiguity of mapping units themselves as represented in digital soil maps. SAGs have also been applied to landform types and coastal environments (Phillips, 2018).

SAGs are intermediate between spatially explicit graphs, where nodes represent specific locations, and structural graphs, where nodes represent system components (most state-and-transition models, for instance, can be represented as structural graphs)(Heckmann et al., 2015). In this note I address a spatially explicit form of SAGs, based on raster representation of categorical spatial units.

Assuming that each raster cell is assigned to a single category, the categories are considered connected where they occur in adjacent cells. Note that the categories (landforms, vegetation, soil types, etc.) are the nodes or vertices of the graph, not the raster cells. But rather than two environmental types being considered connected if they *ever* occur contiguously, in this case a link is defined if *every case* that they occur in contiguous cells. For example, assume there are nine environmental units or categories, X1, X2, . . . , X9. In the top of Figure 1, there are 8 links associated with X5, because there are eight contiguous cells with different EUs. In the middle there are no links, as all the cells are the same type. In the lower part of Fig. 1, six links are associated with the X5 cell in the middle.

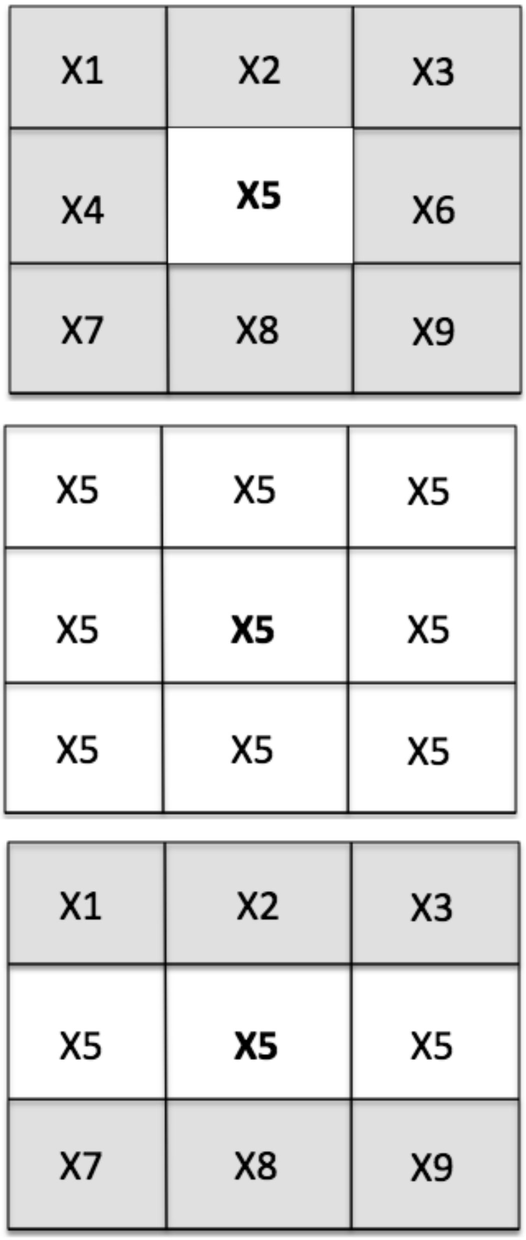

Figure 1. Explanatory diagram. Number of links (shaded) for the X5 cell in the middle is 8 for the top, 0 for the center, and 6 for the bottom example.

The method in this note based on algebraic graph theory and the analysis of graph adjacency matrices. An adjacency matrix for a network with *N* nodes (in this case *N* EUs) is an *N x N* matrix with cell entries of zero if the row and column nodes are unconnected, and nonzero otherwise. In this case cell values indicate the number of cases where cells of the row and column EU are spatially contiguous, and 0 otherwise (by convention, diagonal entries are zero).

The largest eigenvalue (λ_1) of the adjacency matrix is the graph spectral radius. Spectral radius is a key indicator of many network properties; particularly graph complexity. λ_1 is sensitive to the number of cycles in the graph (sequences of edges that begin and end at the same node), and is inversely related to critical coupling strength, a threshold at which a graph transitions from incoherent to coherent behavior (Restrepo et al., 2006, 2007). While coherence is not directly relevant to SAGs, this property reflects the fact that spectral radius is an indicator of graph complexity (see, e.g., Fath, 2007; Phillips, 2011a, 2011b).

Now we define:

N = number of EUs = number of graph nodes
A = size of domain (e.g., area) = nx
n = number of raster cells
x = minimum sample, observation, or classification unit (e.g., cell size)
N_{max} = A/x (each sample or cell is unique)
m = number of identified links
m_{max} = number of links for case of N_{max}

For a typical rectangular raster, $m_{max} = 4n$ since every link is associated with two cells. This can be worked out for other raster geometries; for a hexagonal model, for instance, $m_{max} = 3n$. This assumes, for convenience, that there is an extra layer (row or column) of pixels around the edge of the study area that are not included as part of A or n. If not, appropriate adjustments must be made for the reduced adjacency possibilities of edge pixels.

The most complex spatial pattern would occur where every cell has a different EU, such that $m = m_{max}$ and $N = N_{max}$. The upper limit of the spectral radius for a graph of a given N, m is

$$\lambda_{1,upper} = [2m(N-1)/N]^{0.5}$$

Thus we can compute the reduction in complexity associated with having adjacent cells with the same EU as

$$\lambda_{1,upper}/\lambda_{1,max} = [2m(N-1)/N]^{0.5} / [8N_{max}(N_{max}-1)/N_{max}]^{0.5}$$

Note that this implies with constant N, m, A, the ratio varies as the square root of cell size ($x^{0.5}$).

The observed spectral radius $\lambda_1 \leq \lambda_{1,upper}$. For the observed spatial pattern we can then compute the proportion of complexity associated with having $N < N_{max}$ and $m < m_{max}$ (variability of EUs, or ev), and that associated with the specific spatial pattern of adjacency (the "wiring" of the graph).

$$\zeta_{ev} = (\lambda_{1,max} - \lambda_{1,upper}) / (\lambda_{1,max} - \lambda_1)$$

$$\zeta_{wiring} = 1 - \zeta_{ev}$$

Worked example

Suppose we have a 1 km^2 study area with 10 X 10 m pixels, with a total of 10,000 cells. There are 20 EUs, and the observed m is 1000, and the spectral radius is 10.00. Thus we have $N = 20; N_{max} = n = 10{,}000; m = 1000; m_{max} = 40{,}000.$ Then we have $\lambda_{1,upper}/\lambda_{1,max} = 0.154$; $\zeta_{ev} = 0.877$; $\zeta_{wiring} = 0.123$. For this hypothetical example the maximum possible complexity for a graph with 20 nodes and 1000 links is about 15 percent of that associated with a situation where each of the pixels was a separate EU. Nearly 88 percent of the order (reduction in complexity) of the observed pattern compared to the maximum possible is due to the limited number of EUs actually observed, with the rest owing to the specific adjacency pattern.

References

Fath, B.D., 2007. Structural food web regimes. Ecol. Model. 208, 391–394.

Heckmann, T., Schwanghart, W., Phillips, J.D., 2015. Graph theory – recent developments and its application in geomorphology. Geomorphology 243, 130–146.

Phillips, J.D., 2011a. Predicting modes of spatial change from state-and-transition models. Ecol. Model. 222, 475–484.

Phillips, J.D., 2011b. The structure of ecological state transitions: amplification, synchronization, and constraints. Ecol. Complex. 8, 336–346.

Phillips, J.D., 2013a. Sources of spatial complexity in two coastal plain soil landscapes. Catena 111, 98–103.

Phillips, J.D., 2016. Identifying sources of soil landscape complexity with spatial adjacency graphs. Geoderma 267: 58-64.

Phillips, J.D., 2018. Environmental gradients and complexity in coastal landscape response to sea level rise. Catena 169: 107-118.

Restrepo, J.G., Ott, E., Hunt, B.R., 2006. Emergence of synchronization in complex networks of interacting dynamical systems. Physica D 224, 114–122.

Restrepo, J.G., Ott, E., Hunt, B.R., 2007. Approximating the largest eigenvalue of network adjacency matrices. Phys. Rev. E 76. http://dx.doi.org/10.1103/PhysRevE.76.056119.

Jonathan D. Phillips[1,2], 12 November 2020

[1]Professor Emeritus of Earth Surface Systems, University of Kentucky
[2]Adjunct Professor of Geography, Planning & Environment, East Carolina University

816 Shippoint Avenue, New Bern, NC 28560, USA
jdp@uky.edu