Echelon analysis

W.L. Myers & G.P. Patil

Volume 2, pp 583–586

in

Encyclopedia of Environmetrics
(ISBN 0471 899976)

Edited by

Abdel H. El-Shaarawi and Walter W. Piegorsch

Echelon analysis

Quantitative spatial data are important inputs for many environmental models that determine future implications of current resource use, policies, and interventions. End products of applying such models are often mappings of indices for level of potential environmental impact (*see* **Impact, environmental**), which then become guides to allocation of economic and technical resources for amelioration.

Errors in quantitative spatial data will propagate through environmental models and find expression in the resulting impact indices. However, the consequences of such errors for decision making may well depend upon where the errors occur. There may be relatively little confusion introduced by moderate errors occurring in a vicinity that otherwise has consistently high values of a variable. In contrast, errors compound confusion in areas that are highly variable. Errors can also substantially distort the apparent state of areas that otherwise have consistently low values of a variable.

It is therefore desirable to have a systematic means of determining spatial organization in mappings of quantitative variables, both for input variables to environmental models and for indexes of potential impact generated by the models. Modern computer capabilities for visualization of surfaces are helpful in this regard, but their interpretation is subjective. *Echelons* present an innovative alternative for objectively determining quantitative spatial structure for direct mapping, either with or without computer-assisted visualization [1–9]. Thus, they can facilitate analysis of errors associated with environmental models that take quantitative layers as input, or produce quantitative output layers, or both.

Echelons of Spatial Variation

The spatial variables for echelon analysis are essentially topographies, whether real or virtual. Such terrain information is typically formated for processing in a **geographic information system** (GIS) as a digital elevation model (DEM). This comprises a raster in which an 'elevation' value is specified for the center of each cell. Echelons divide the (virtual) terrain into structural entities consisting of peaks, foundations of peaks, foundations of foundations, and so on, in an organizational recursion. Saddles determine the divisions between entities. Each entity is assigned an echelon number for identification purposes. The peaks constitute one series of structural entities, being numbered in decreasing order of summit elevation. The foundations constitute a second series of entities that are likewise numbered in order of decreasing top level, starting with the next number after that assigned to the lowest peak. Consider, for example, the terrain depicted in profile with division as seen in Figure 1.

The numbered entities thus determined are called echelons. Echelons are determined directly by organizational complexity in the spatial variable, and not by either absolute 'elevation' or steepness.

Echelons form extended families of terrain entities having a genealogy similar to that of an extended human family, except that each echelon has only one parent. In the case of echelons, an entity that rises from another is more aptly termed an 'ascendant', rather than a 'descendant'. Likewise, a 'parent' entity is termed a 'founder'. The echelon relations determine a family tree as illustrated in Figure 2. This is a 'scaled tree' in the sense that the height of each vertical edge corresponds to the height of the echelon above its founder (parent). The cumulated height above the root is the height of the terrain. The number of 'ancestors' for an echelon is a local measure of regional complexity.

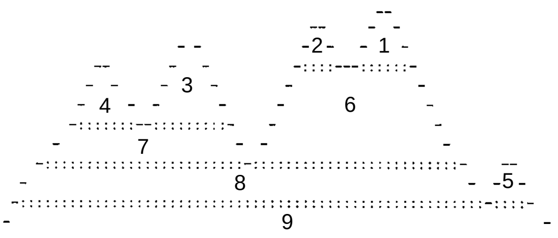

Figure 1

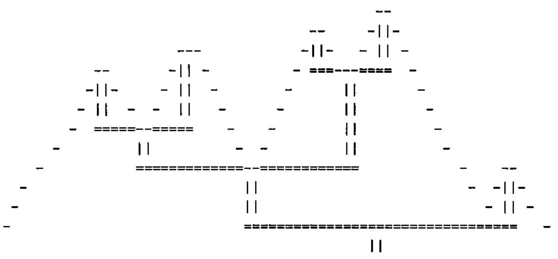

Figure 2

The echelons also comprise a structural hierarchy of organizational orders. The orders of the hierarchy are assigned and numbered in the same manner as for a network of streams and tributaries [8]. Thus, peaks are akin to unbranched tributaries, and have order 1. A foundation for two or more order 1 entities is of order 2. Likewise, a foundation for two or more order 2 entities is of order 3. A low order entity (see 5 in Figure 1) sharing a foundation with a higher order entity does not increase the order of the foundation. This is not unlike the case of an unbranched tributary entering a higher order stream.

Echelon Characteristics

A suite of form attributes can be determined for each echelon, including area extent of the basal slice and vertical projection above its founder. Some form attributes may depend upon an interval scale of measure for the vertical dimension, but the echelon decomposition only requires an ordinal scale of measurement. A standard table of echelon characteristics contains a record (row) with 10 fields for each echelon, including echelon identification number, order, founder, maximum level, minimum level, relief, cells, progeny, ancestors, and setting within the tree. The table is associated with an echelon map file giving the 'level' value and echelon identification number for each cell. Echelons thus formalize the structural complexity of the (surface) variable without incurring any loss of information with respect to (surface) level.

Echelon Trees

Since most echelon trees are much too complicated for visual study as dendrograms, characterization and comparison of echelon trees is done through analytical processes such as pruning. Consider, for example, a pruning process on an echelon tree that recursively partitions the tree into an inner set of limbs and an outer set of boughs. The first stage of pruning traces all terminal (order 1) echelons down to the root while counting the echelons that comprise each path. The subset of terminals thus identified as having the maximum path is then retraced to determine which echelons are shared by all such paths. The mutual echelon path working upward from the root becomes the first (main or trunk) limb of the echelon tree. All other components of the tree are pruned away at the trunk forming a set of subtrees called boughs. This partitions the tree into a set of limb echelons and another set of bough echelons.

In the second stage of pruning, the process is conducted separately on each of the boughs. Each bough thus yields, in turn, a limb and a set of additional boughs. This repartitions the tree into a larger set of limb echelons and a reduced set of bough echelons. Further stages of pruning will eventually convert the boughs entirely to limbs, so that the bough fraction will become zero. Some boughs will continue to yield residual boughs longer than others, depending upon depth of structure in the respective subtrees.

Echelon Profiles

Plotting the limb fraction (as a percentage) vs. the stage number in the pruning process will yield a divergence profile of surface organization. Simple trees will have the profile climb rapidly to 100.

A scope profile plots the percentage of cells in limbs vs. stage number. Since echelons vary in number of cells (even for the same order), such a profile captures the scope for different degrees of complexity on the surface. A bunching profile plots boughs as a percentage of order 1 nodes against stage (pruning cycle) number. This profile is particularly sensitive to depth of structure and its consistency among the branches. A stacking profile plots the percentage of order 1 nodes in limbs against stage number. This profile is indicative of propensity for echelon siblings to be of the same order.

Echelons may also be determined after filtering the (surface) variable in several ways to remove 'jaggedness' that has high spatial frequency. The degree of change in the echelon structure due to filtering is indicative of the data error instability and the amount of 'noise' expressed on the surface.

Echelon Research

The current stage of development for echelon technology provides a mature descriptive capability for characterizing quantitative spatial variables. A major question concerning quantitative spatial variables with respect to many applications is whether there are substantial sectors of the surface having particularly high or particularly low values relative to the mean level. These are the 'uplands' and 'lowlands' of the virtual surface. Currently the manager or investigator is obliged to resort to subjective examination of visualizations on maps and/or computer displays in an attempt to gain insights on possible ‘focal’ areas.

In the domain of echelons, candidate focal areas may be conceptualized as principal families; the sectors that they occupy can be viewed as principalities. The information needed for determining principal families resides in the echelon table and tree representation. Once the principal families are identified, the sectors that they occupy can be extracted by exploiting the linkage between the echelon map and the echelon table. Analytical and computational strategies must be formulated for segregating the principal families from what are typically hundreds of upper-level echelon families.

Probabilities based on a null model using a planar random process allow the user to specify a criterion for areas of potential concern to be extracted computationally. In other words, an echelon family would be seen as a candidate for focus if the probability of its extent receiving observed amounts is less than the criterion under a random distribution of quantity over area.

Since echelon determination is computationally intensive, there is further advantage in extracting principal families from partially determined echelons. This scenario would terminate the top-down progression of echelon determination for an area when the probability of observing encountered values under planar randomization exceeds the criterion level. The echelon table would then consist of a series of subtrees, with a subtree for each principal family.

There may also be environmental and/or economic importance in especially ‘rough’ or irregular areas of high spatial variability. A promising strategy for this purpose is analytical specification of tree models representing specific types of spatial roughness. Binary trees and ‘vines’ are special forms that provide points of departure for this component of the work. Roughness develops through bifurcation in a binary tree. A ‘vine’ is a ‘noise’ tree in which each node has only one ‘fertile’ element that undergoes further branching. Corresponding computational capability must then be configured for traversing a tree and excising those sequences of nodes that are consistent with the branching model of interest.

Filtering strategies can be explored for the purpose of assessing robustness of spatial structure to errors in the variable. Extraction of principal families and principalities after filtering severity corresponding to assessed error will indicate whether or not principalities are locationally stable and therefore robust to errors.

Environmental Applications

A further line of research involves comparative study of spatial complexity as expressed by a suite of indicators for different aspects of ecosystem health (*see* **Ecosystem monitoring**). A basis for approaching this problem lies in the echelon capability for local specification of regionalized complexity. This specification can take several forms such as echelon order, number of ancestors, and precedence in terms of sequential echelon determination. Each such specification yields a synthetic image band. These pseudo-image bands can be assembled as synthetic multiband complexity image datasets for the region in question. Hypercluster compression of the synthetic multiband image data will extract prevailing patterns of complexity among the several indicators of ecosystem health. Spatial patterns of joint complexity and/or simplicity among health indicators for stressed **ecosystems** can provide a new diagnostic tool. The complexity clusters will occur in spatial patches that can, in turn, be modeled by hierarchical Markov transition matrices (*see* **Conditional entropy profiles**).

References

[1] Johnson, G.D., Myers, W.L., Patil, G.P. & Walrath, D. (1998). Multiscale analysis of the spatial distribution of breeding bird species richness using the echelon approach, in *Assessment of Biodiversity for Improved Forest Planning*, P. Bachmann, M. Kohl & R. Paivinen, eds, Kluwer, Boston, pp. 135–150.

[2] Kurihara, K., Myers, W.L. & Patil, G.P. (2000). Echelon analysis of the relationship between population and land cover pattern based on remote sensing data, *Community Ecology* **1**, 103–122.

[3] Myers, W.L., Patil, G.P. & Joly, K. (1997). Echelon approach to areas of concern in synoptic regional monitoring, *Environmental and Ecological Statistics* **4**, 131–152.

[4] Myers, W.L., Patil, G.P. & Taillie, C. (1995). Comparative paradigms for biodiversity assessment (Invited paper at the IUFRO Symposium in Chiang-mai, Thailand), in *Measuring and Monitoring Biodiversity in Tropical and Temperate Forests*, T.J. Boyle & B. Boontawee, eds, CIFOR, Bogor, Indonesia, pp. 67–85.

[5] Myers, W.L., Patil, G.P. & Taillie, C. (1999). Conceptualizing pattern analysis of spectral change relative to ecosystem health, *Ecosystem Health* **5**, 285–293.

[6] Patil, G.P. & Taillie, C. (1999). A Markov model for hierarchically scaled landscape patterns, *Bulletin of the International Statistical Institute* **58**, 89–92.

[7] Ramakomud, A. (1998). Change detection using hyper-clustered data: the spatial averaging approach, MSc. Thesis, Pennsylvania State University, University Park.

[8] Rodriguez-Iturbe, I. & Rinaldo, A. (1997). *Fractal River Basins: Chance and Self-Organization*, Cambridge University Press, Cambridge.

[9] Smits, P.C. & Myers, W.L. (2000). Echelon approach to characterize and understand spatial structures of change in multi-temporal remote-sensing imagery, *IEEE Transactions, Geoscience and Remote Sensing* **38**, 2299–2309.

(*See also* **Spatial analysis in ecology**; **Tree-structured methods**)

W.L. MYERS & G.P. PATIL