Utah State University

DigitalCommons@USU

Articles   Publications

11-16-2010

The Landscape Similarity Toolbox: New Tools for Optimizing the Location of Control Sites in Experimental Studies

Thomas E. Dilts
University of Nevada, Reno

Jian Yang
University of Nevada, Reno

Peter J. Weisberg
University of Nevada, Reno

Follow this and additional works at: https://digitalcommons.usu.edu/sagestep_articles

Part of the Plant Sciences Commons

Recommended Citation

Dilts, T.E., Yang, J. and Weisberg, P.J. (2010), The Landscape Similarity Toolbox: new tools for optimizing the location of control sites in experimental studies. Ecography, 33: 1097-1101. doi:10.1111/j.1600-0587.2010.06577.x

UtahState University
MERRILL-CAZIER LIBRARY
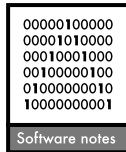
The Landscape Similarity Toolbox: new tools for optimizing the location of control sites in experimental studies

Thomas E. Dilts, Jian Yang and Peter J. Weisberg

T. E. Dilts (tdilts@cabnr.unr.edu), J. Yang and P. J. Weisberg, Great Basin Landscape Ecology Lab, Dept of Natural Resources and Environmental Science, Univ. of Nevada Reno, 1000 Valley Road, Reno, NV 89512, USA.

Large-scale manipulative experiments are critically important for linking ecological theory with land management at a relevant spatial scale. Statistically powerful inferential approaches such as the before-after-control-impact design involve pairing a small number of treatment sites with control sites of analogous ecological structure and landscape context. Pairing treatment and control sites that are as analogous as possible is an important step to ensuring that differences are due to a treatment effect. The Landscape Similarity Toolbox provides tools for optimizing the location of potential control sites based upon the spatial characteristics of the treatment site.

There is increasing recognition of the need for large-scale experimental studies in ecology, to provide an appropriate scope of inference for understanding episodic and spatially heterogeneous phenomena (Carpenter 1996, Grace et al. 1997). However, robust experimental design for large-scale studies can be problematic for two primary reasons: lack of true replication and lack of true randomization (Hobbs 2003). High cost and logistical limitations when experimental treatments are conducted over large areas generally lead to a low number of replicates. The potential for bias also increases with scale (Hobbs 2003) because "all landscapes are different" with respect to various ecological processes including those related to or confounding treatment effects. Hence random assignment of large replicate areal units to treatments becomes infeasible (Hargrove and Pickering 1992).

Faced with the challenges of implementing inferentially strong, landscape-level experiments, some have proposed that the before-after-control-impact (BACI) design (Stewart-Oaten et al. 1986) provides the greatest inferential power, particularly where the treatment is expressed as a one-time perturbation such as road construction, modification or removal (Roedenbeck et al. 2007) or forest harvesting impacts (Bennett and Adams 2004). The BACI design compares treated sites with untreated control sites (i.e. control-impact main effect) both before and after treatment occurs (i.e. before-after main effect). A significant effect of the treatment is indicated when the difference in the response variable between treatment and control sites is greater after treatment occurs (i.e. a significant interaction term between the two main effects). Such an approach requires a high level of ecological similarity between control

and treatment sites; so that "before-after" treatment effects are not masked by "control-impact" effects, where differences in the response variable between the treatment and the control may be due to initial site characteristics rather than the actual treatment effect (Underwood 1992). Thus, the placement of control and treatment sites is critical to the overall viability of the experiment.

Despite the great importance in identifying suitable control sites there is a lack of objective methods to facilitate this process. Commonly used methods, such as field visits by experts and comparison using maps or aerial photographs, lack repeatability among observers. Furthermore it may be infeasible to visit all candidate field sites. Existing GIS methods, such as Boolean logic (Burrough and McDonnell 1998), are commonly used to narrow the search for suitable control areas, but tend to be simplistic. For example, GIS-based methods that employ Boolean logic can only tell us whether a site is suitable according to some pre-defined criteria. If the criteria are too strict it is possible that no candidate sites meet all of the criteria. Alternatively, if the criteria are too relaxed then there may be many potential sites that the researcher needs to choose from. This paper describes a more advanced GIS-based approach (the Landscape Similarity Toolbox) that incorporates many types of spatial information to maximize the similarity between paired control and treatment sites.

The Landscape Similarity Toolbox incorporates three types of spatial variation: compositional, configurational, and continuous. All three types of spatial variation are calculated at a local level using moving window analysis and then mapped throughout the entire study region. Compositional variation refers to the frequency distribution of cells

among categories without regard to their spatial location. Configurational variation provides information about the arrangement of individual cells or groups of contiguous cells referred to as patches. Much of the discipline of landscape ecology has focused on describing the shape and arrangement of patches which can be characterized using patch-based metrics (Turner 1989). Configurational variation at the cell or pixel level has received less attention than at the patch level, although one might argue that for two sites to be maximally similar they should have precisely the same arrangement of cells. Continuous patterns of variation most appropriately describe environmental gradients such as soil, climate, and topography that vary continuously rather than abruptly across the landscape (McGarigal et al. 2009).

The Landscape Similarity Toolbox methodology allows the researcher to retain critical spatial information concerning the arrangement of cells and patches within the treatment and to compare that arrangement to all potential control sites. The method is flexible in that it allows the researcher to incorporate the level of spatial information that most closely matches the landscape attributes requiring experimental control. Moving window (neighborhood) analysis within a GIS environment is used to generate per-pixel maps of similarity between the treatment and all potential control areas that can be used to augment the control site selection process. The Landscape Similarity Toolbox consists of python scripts that can be run stand-alone or within ArcGIS software and is freely available for download at <http://purl.oclc.org/similaritytools>. All tools employ a moving window methodology in which a window of similar size and dimensions to the treatment area is generated (Fig. 1). This window is referred to as the "template". The surrounding area from which potential control sites can be located is known as the "region". Moving window analysis is used to compare the treatment window to all moving windows (potential control sites) within the region to generate maps of similarity. The approach is flexible in that it can generate similarity maps for any spatial dataset provided by the user, and it allows the user to determine how to combine those maps to provide an overall measure of site similarity. Users may opt to combine various maps of similarity outside of the Toolbox using a number of existing methods, such as principal components analysis or weighted overlay.

The Toolbox includes tools for Data Preparation and Data Analysis of both categorical and continuous data (Fig. 2). The compositional similarity tool measures similarity between the treatment and all possible control sites without regard to the relative or absolute position of those cells within the treatment or the moving window. It first creates a value table that contains each class and the number of cells present in the treatment as well as each window. Then the absolute difference in the count of each class in the treatment is weighted by the number of cells in each class to generate a dissimilarity map.

$$d(x) = \sum_{i=1}^{m} \frac{1}{p_{i,T}} |N_{i,T} - N_{i,Mx}|$$

where $d(x)$ is dissimilarity for cell x, $P_{i,T}$ is the proportion of class i in the treatment window T, $N_{i,T}$ is abundance (i.e.
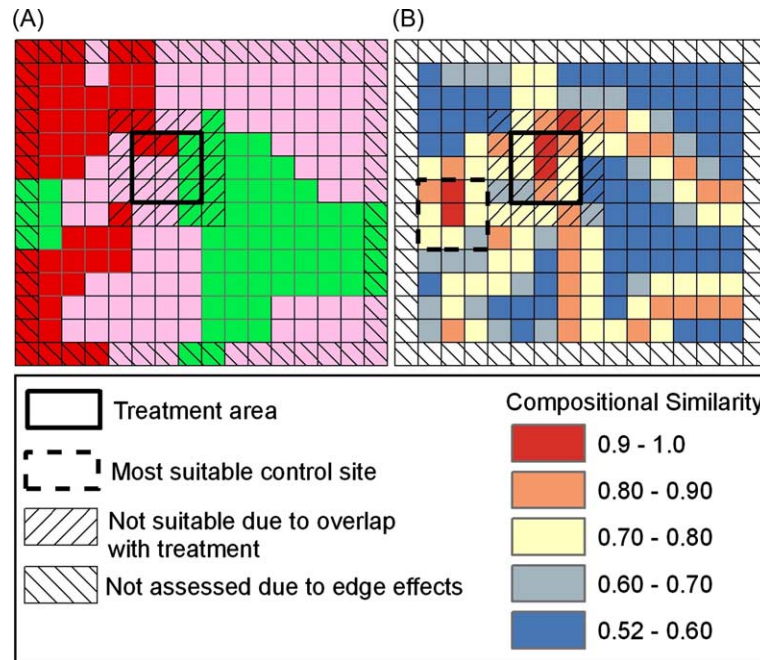

Figure 1. Moving window procedure used for conducting similarity analysis. (A) Shows a three category raster map (colored using red, green, and pink) with the treatment area outlined with the bold box. Areas with hatching represent potential sites that cannot be included because of boundary issues (hatching at 225°) and overlap with the treatment area (hatching at 45°). (B) Shows the results of the compositional similarity analysis. Cells with hotter colors represent potential sites that are more similar to the treatment. Each nine cell neighborhood is represented by the value in the center cell. Values represent the proportion of cells that are the same as the treatment without regard to position within the nine cell neighborhood. Cells with edge effects have been removed from the illustration on the right while cells within and overlapping with the treatment area have been included for visual reference. The value of the center cell in the nine cell treatment area is 1.00.
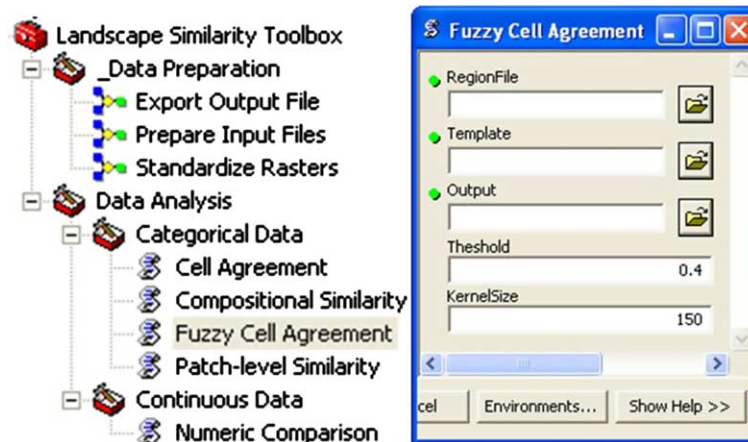
Figure 2. The Landscape Similarity Toolbox within ArcGIS. The Landscape Similarity Toolbox is divided into a set of tools designed for preparing data for analysis and for conducting analysis. The data analysis tools are split into separate toolsets for continuous data and categorical data. The dialog for the fuzzy cell agreement tool is shown to illustrate the types of inputs and parameters required. The data analysis tools are written in Python and can be run stand-alone in a free Python development environment.

number of pixels) of class i in the treatment window T, $N_{i,Mx}$ is abundance of class i in the moving window Mx that is centerred on cell x, and m is total number of classes.

The patch-level similarity tool measures configurational similarity at the patch level. A patch is defined as a group of contiguous cells of the same class surrounded by cells of a different class (Forman 1995). Patches can have properties that describe their shape as well as their position relative to other patches. Patches can be derived from imagery using object-based image analysis and object-based image analysis could be used as a tool to consider multiple scales of data in a hierarchical fashion (Burnett and Blaschke 2003). This tool utilizes output raster datasets derived from a free software known as FRAGSTATS (McGarigal et al. 2002). Patch metrics are summarized in FRAGSTATS at the class level for the treatment; moving window analysis is then used to calculate those same metrics for each possible control site. The patch-level similarity tool compares metric values derived from the treatment to the values for each window.

$$d(x) = |C_{Mx} - C_T|$$

where C_{Mx} is the class-level metric value for the moving window Mx, C_T is the class-level metric value for the treatment window T. Both C_{Mx} and C_T are calculated by FRAGSTATS.

Both the cell agreement and the fuzzy cell agreement measure configurational similarity at the cell level rather than the patch level. Measures of cell-level configurational similarity may offer advantages over measures of patch-level similarity because they can simultaneously incorporate information about both overlap and structure (Hagen-Zanker 2006). The cell agreement tool compares the specific arrangement of cells within the treatment with the moving window and determines which cells match in terms of their relative postions within the window. The tool totals the number of matches between the treatment and the window and compares this value to the total number of cells. The cell agreement tool can be viewed as the strictest of all of the tools for comparing similarity.

$$d(x) = 1 - \frac{1}{N_{Mx}} \sum_{z \in Mx} I_A(z); \;\; I_A(z) = \begin{cases} 1; z \in A \\ 0; z \notin A \end{cases};$$
$$A = \{z : z \in Mx \text{ and } V_z = V_{T(z)}\}$$

where $I_A(z)$ is the indicator function (i.e. $I_A(z)$ equals one if cell z belongs to set A and zero otherwise), A is the subset of non-background cells in the moving window that have the same class as the corresponding cells in the treatment window, V_z is the value (i.e. class) of cell z, $T(z)$ is the cell in the treatment window that spatially corresponds to cell z in the moving window, Mx is the set of all non-background cells in the moving window, and N_{Mx} is the total number of such cells in the moving window.

Precise cell agreement becomes unlikely at the landscape level. Therefore, the fuzzy cell agreement tool uses fuzzy set theory to relax the assumption of an exact match in terms of relative position. Fuzziness of location is incorporated by allowing partial values when cells of the same category are within a specified distance (Hagen 2003). The partial values are assigned based upon a user-specified distance using a linear decay function.

$$d(x) = 1 - \frac{1}{N_{Mx}} \sum_{x \in Mx} S_{fuzzy}(x);$$
$$S_{fuzzy}(x) = \begin{cases} 1 - \dfrac{\min(dist(x,y)|y \in Y)}{dist_{max}} & , Y \neq \varnothing; \\ 0 & , Y = \varnothing \end{cases}$$
$$Y = \{y : y \in Mx \text{ and } V_y = V_{T(x)} \text{ and } dist(x,y) \leq dist_{max}\}$$

where Y is the subset of cells in the moving window Mx that are within the maximum search distance ($dist_{max}$) of the focal cell x and match the class of the corresponding cell $T(x)$ in the treatment window, and $dist(x,y)$ is the distance between cell x and cell y.

The numeric comparison tool allows for similarity comparisons using continuous rather than categorical data. The tool first identifies the row and column position for each cell in the treatment as well as the corresponding
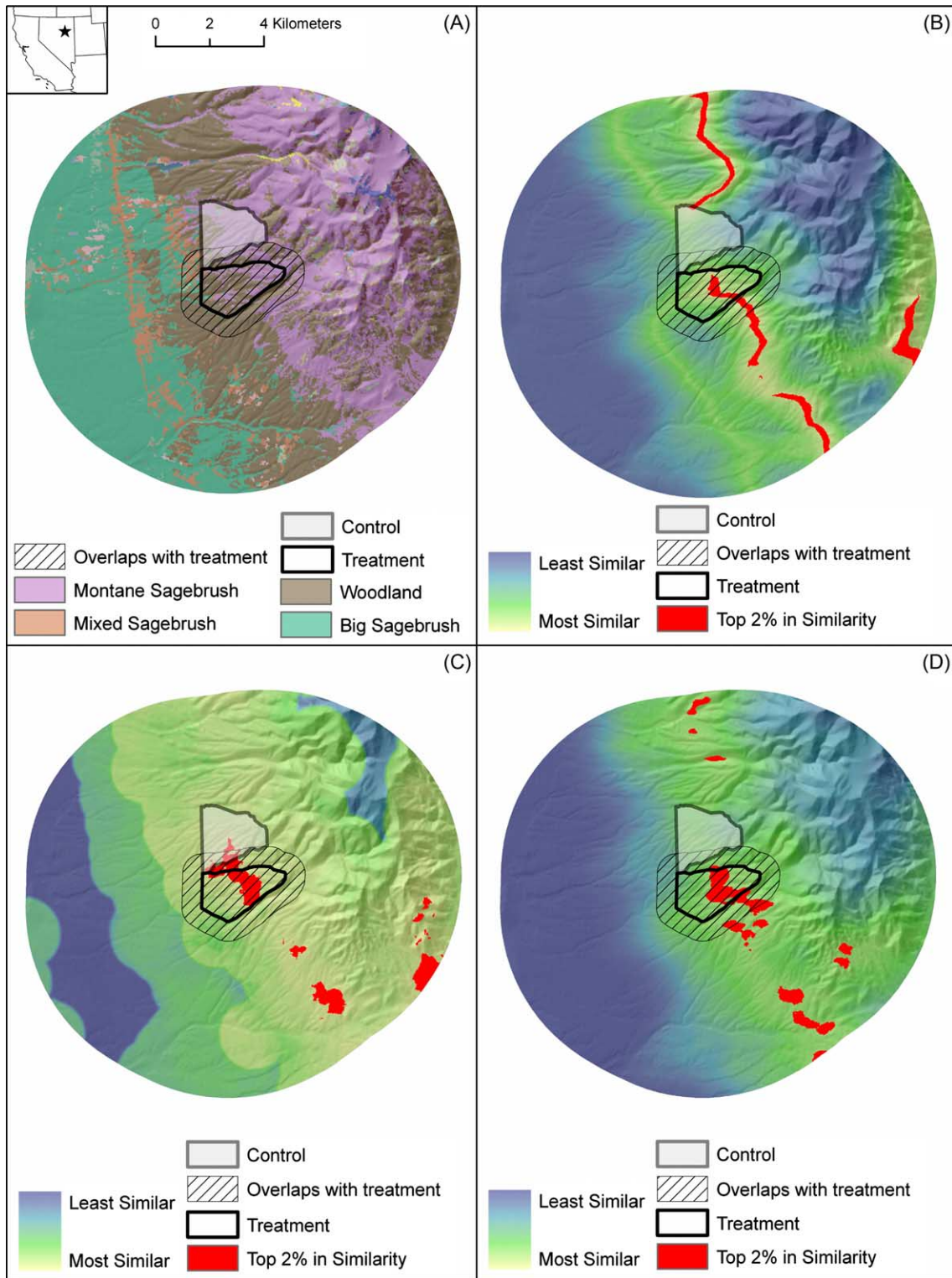
Figure 3. Maps showing the output of the Similarity Toolbox for the South Ruby Mountain site in Nevada, USA. Values for each cell on the map represent an area the same shape, size, and dimensions as the treatment. Control and treatment sites were identified by scientists and managers without using the Landscape Similarity Toolbox. (A) Vegetation map with underlying topographic relief, (B) compositional similarity for vegetation with the top 2% of most similar sites highlighted, (C) patch-level configuration similarity for vegetation with the top 2% of most similar sites highlighted, (D) cell-level configuration similarity using the fuzzy cell agreement method for vegetation with the top 2% of most similar sites highlighted. Results from the cell agreement tool and the continuous data tool are not shown.

row and column position of each cell within each moving window. It then takes the absolute value of the difference between the value in the treatment and the corresponding value in the moving window. Finally, the absolute value of the differences is summarized for each window and is written to the center cell to produce a map of similarity.

$$d(x) = \frac{1}{N_B} \sum_{z \in B} |V_z - V_{T(z)}|;$$
$$B = \{z : z \in Mx \text{ and } T(x) \in T\}$$

where Mx and T comprise the set of all non-background cells in the moving window and the treatment window respectively, and B is the subset of non-background cells in the moving window whose corresponding cells in the treatment window are also non-background cells.

We tested the Landscape Similarity approach by comparing the outputs of each of the tools using a Sagebrush Steppe Treatment Evaluation Project (SageSTEP) treatment site (McIver et al. 2005) and the Southwest ReGAP vegetation map (Lowry et al. 2007). Maps of vegetation similarity were produced using three methods: compositional similarity, patch-level configurational similarity, and the fuzzy cell agreement version of cell-level configurational similarity (Fig. 3). The compositional similarity map selected areas along ecotone boundaries where the proportion of each class was the most similar. In contrast, both the patch-level and the cell-level configurational similarity approaches produced relatively similar maps. However, the fuzzy cell-level configurational similarity approach identified those areas where patch configurations were more clustered compared to the patch-level configurational similarity approach.

Although we have developed the Landscape Similarity Toolbox to assist with the selections of control sites for large-scale manipulative experiments, other types of studies could benefit from improved pairing of treatment and control sites. ''Natural experiments'' in which the treatment is not controlled by the researcher, but rather results from a natural disturbance event, could benefit from improved pairing of treatment and control sites. Quantitative approaches to identifying control sites can be used for smaller-scale studies as well as larger ones by incorporating landscape context. Landscape context can easily be incorporated into the Landscape Similarity framework by expanding the template beyond the boundary of the treatment area to incorporate neighboring areas. The Landscape Similarity Toolbox represents an improvement in the identification of potential control sites that are maximally similar to a pre-defined treatment area. The Toolbox builds upon existing software (e.g. FRAGSTATS) by comparing metrics derived from a treatment to all potential control sites in the surrounding region, producing maps that indicate the degree of similarity between the treatment and potential control sites. Use of the Toolbox enforces clear definition of landscape similarity in the context of sources of potentially confounding, spatial variation for large-scale ecological studies.

To cite the Landscape Similarity Toolbox or acknowledge its use, cite this Software note as follows, substituting the version of the application that you used for ''Version 1.0'':

Dilts, T. E., Yang, J. and Weisberg, P. J. 2010. The Landscape Similarity Toolbox: new tools for optimizing the location of control sites in experimental studies. – Ecography 33: 1097–1101 (Version 1.0).

References

Bennett, L. T. and Adams, M. A. 2004. Assessment of ecological effects due to forest harvesting: approaches and statistical issues. – J. Appl. Ecol. 41: 589–598.

Burnett, C. and Blaschke, T. 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. – Ecol. Model. 168: 233–249.

Burrough, P. A. and McDonnell, R. A. 1998. Principles of GIS. – Oxford Univ. Press.

Carpenter, S. R. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. – Ecology 77: 677–680.

Forman, R. T. T. 1995. Land mosaics: the ecology of landscapes and regions. – Cambridge Univ. Press.

Grace, J. et al. 1997. Tackling large-scale problems by scaling-up. – In: Van Gardingen, P. R. et al. (eds), Scaling-up: from cell to landscape. Cambridge Univ. Press, pp. 7–16.

Hagen, A. 2003. Fuzzy set approach to assessing similarity of categorical maps. – Int. J. Geogr. Inform. Syst. 17: 235–249.

Hagen-Zanker, A. 2006. Map comparison methods that simultaneously address overlap and structure. – J. Geogr. Inform. Syst. 8: 165–185.

Hargrove, W. W. and Pickering, J. 1992. Pseudoreplication a sine qua non for regional ecology. – Landscape Ecol. 6: 251–258.

Hobbs, N. T. 2003. Challenges and opportunities in integrating ecological knowledge across scales. – For. Ecol. Manage. 181: 223–238.

Lowry, J. H. Jr et al. 2007. Land cover classification and mapping. – In: Prior-Magee, J. S. et al. (eds), Southwest regional gap analysis final report. U.S. Geological Survey, Gap Analysis Program, Moscow, ID, chapter 2.

McGarigal, K. et al. 2002. FRAGSTATS: spatial pattern analysis program for categorical maps. – Computer software program, Univ. of Massachusetts, Amherst, <www.umass.edu/landeco/research/fragstats/fragstats.html>.

McGarigal, K. et al. 2009. Surface metrics: an alternative to patch metrics for the quantification of landscape structure. – Landscape Ecol. 24: 433–450.

McIver, J. H. et al. 2005. A regional experiment to evaluate effects of fire and fire surrogate treatments in the sagebrush biome. – <www.sagestep.org/pdfs/SageSTEP_proposal.pdf>.

Roedenbeck, I. A. et al. 2007. The Rauischolzhausen agenda for road ecology. – Ecol. Soc. 12: 11, <www.ecologyandsociety.org/vol12/iss1/art11>.

Stewart-Oaten, A. et al. 1986. Environmental impact assessment: ''pseudoreplication'' in time? – Ecology 67: 929–940.

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

Underwood, A. J. 1992. Beyond BACI – the detection of environmental impacts on populations in the real, but variable, world. – J. Exp. Mar. Biol. Ecol. 161: 145–178.