1. Introduction

LandEx-USA (Landscape Explorer USA) is a GeoWeb-based tool for exploration of patterns of land cover classes in the National Land Cover Dataset 2006. (www.mrlc.gov/nlcd2006.php). Standard GIS tools can perform queries for spatial extent of single class of land cover, but it is sometimes more interesting to search for spatial extent of characteristic patterns of land cover classes rather than a single class. This is exactly what LandEx does by performing query by pattern similarity (QBPS) calculations.

LandEx user interface works in an internet browser environment and is based on JavaScript libraries: ExtJS with GeoExt and OpenLayers. Through this interface a user can access all functionalities expected from a modern GeoWeb web page.

LandEx has been developed in the Space Informatics Lab (sil.uc.edu) at the University of Cincinnati and is available at sil.uc.edu/landex. References for LandEx are:


2. Opening Screen

2.1 - The main window
The center of the screen shows National Land Cover Dataset 2006 (NLCD2006). Different colors indicate 16 different land cover classes. The NLCD2006 has a resolution of 30 meters/cell; there are approximately $10^{10}$ cells in the dataset.

2.2 - Map layers
This panel shows all layers available to LandEx. The layers are divided into three categories:

- **Boundaries**: these auxiliary layers help user to find proper location

- **Similarity layers**: are user generated, each is available for 20 minutes after generation

- **Base layers**: include land cover and hillshade. Hillshade is for reference only; it’s not used in similarity calculation.
2. Openning Screen

2.3 - Scale and navigation

Map navigation is provided by panning and zooming. In order to pan the map press the left mouse button and drag the map. Use mouse wheel to zoom in and out. Zooming into a selected rectangle is accomplished by pressing Shift button and using mouse to indicate an approximate area into which to zoom in.

Navigation icon at the top-left part of the screen has a map scale slider (A), zoom-out to maximum extent button (B) and help button (C) that brings up this document.

2.4 - Opacity

Opacity slider is located at the lower-left part of the screen in the Map Layers section. It acts on the selected layer. Opacity is mostly used for visual comparison of local and reference patterns (see section 3.4).

2.5- Similarity calculation

Similarity button is initially located at upper-right part of the screen. It initiates search for local regions having patterns similar to a user-selected example. Clicking on this button brings up a region of interest (ROI) square that a user needs to position at selected location (see section 3.1)
2. Opening Screen

2.6 - Legends

When LandEx opens the only legend visible is the land cover (NLCD2006) legend that shows 16 different classes. A more detailed description of every class is given at www.mrlc.gov/nlcd06_leg.php

The land cover legend is always present.

Once a similarity map is calculated, a similarity legend appears in the Legends section of the screen. Calculation of subsequent similarity maps does not bring new legends as every similarity map has the same legend. The values are from 100% indicating perfect similarity to 0% indicating no similarity whatsoever.
3. Getting similarity map

3.1 - Navigate to region of interest

This illustrates how to get a similarity map using a region of interest (ROI). An example ROI, located in the southwest region of Ohio, is used. Starting from the map of the entire USA (A) use navigational tools to zoom into the southwest corner of Ohio (B).

Click on the “Similarity” icon located in the upper-right corner of the screen.

“Similarity” icon expands showing available options; simultaneously, a pink square (C) appears in the middle of the screen. This pink square indicates selection of ROI. It can be repositioned using a mouse. Using navigational tools preserves the presence and location of ROI. Position ROI so it covers a site exhibiting a pattern of land cover to be used as reference for US-wide search. Click “Start calculation” button (D) to start the search. The search will take few seconds.

There are two choices for the size of ROI (E); 500 x 500 px is the default. ROI can be dismissed by clicking on “Similarity” button (F) before starting calculations. Resulting similarity map can be smoothed using options (G); the default is “none”. Clicking button (H) zooms to the ROI.
3. Getting similarity map

3.2 - Examine resultant similarity map

The ROI selected in the southwest Ohio example (A) consists of characteristic pattern dominated by deciduous forest and pasture/hay classes of land cover. The same area is also shown in uncategorized satellite image (B).

This rendition of ROI shows map of land cover wrapped around hillshade. LandEx uses such rendition in order to provide additional information about a site. However, NO terrain information is used in a search, only the information about the pattern of land cover is utilized.

Google map image of ROI helps to relate a characteristic pattern of land cover to visual impression of the land. In this version of LandEx satellite image is not provided.
Once similarity calculation finishes, the screen is updated: A new layer - a similarity map” - is displayed (A) and added to available layers (B). Similarity layer has its unique ID number. A user can create many similarity layers, but each will expire after 20 minutes from its creation.

Also added is a view of the ROI (a reference pattern of land cover) (C), so it can be visually compared to any other pattern indicated as similar on the similarity map. Clicking on ROI zooms the main map to its location.

Finally, a legend to similarity map (D) is added to legend section. Increasingly red colors indicate increasing similarity of local pattern to the reference pattern. Most people perceive patterns with >90% similarity as highly similar and those with similarities in the range between 80% and 90% as similar.

In this particular example, a broad region on the two sides of Ohio - West Virginia boarder and extending to Pennsylvania contains pattern of land cover similar to the selected reference.
3. Getting similarity map

3.3 - Examine similarity map on global scale

Zooming out reveals the similarity map over the entire United States. The entire map can be downloaded for further analysis using a download button (A). Map is saved in geotiff format preserving color legend; the size of the file is 3.7MB/13.6MB depending on size of ROI (and the granularity of the map). The values in geotiff are the similarity values multiplied by 100 in order to make them integers.

Zooming into the Missouri area reveals existence of several hot spots characterized by high similarity to the reference pattern taken from southwest Ohio. Clicking on the map brings a pop-up showing a similarity value and a land cover class.
3. Getting similarity map

3.4 - Visual inspection of pattern search results

Zooming into one of the hot spots (A) of pattern similarity in the Missouri area. The specific cell underlying the letter (A) carries a similarity value that relates to comparison between the ROI (C) and a local area shown by dashed square and centered on (A).

In order to see the actual pattern of the local area centered on (A), the similarity layer must be "selected" (in Map layers section); opacity slider (B) can then be used to uncover the land cover map and enabling direct visual comparison of patterns in (C) and (A).
4. How similarity is calculated?

4.1 - Introduction to similarity measure

A similarity measure between patterns of land cover in two different sites is the core ingredient of LandEx. It is important to note that all such measures are heuristic by nature and the goal is to use a measure that minimizes a disparity between a notion of alikeness between two sites as perceived by a human analyst and a degree of alikeness between the sites as calculated by a similarity algorithm. LandEx uses an original similarity algorithm published in IEEE Geoscience and Remote Sensing Letters. The algorithm is based on information theory notion of mutual information; details can be found in the aforementioned paper, here an explanation based on a specific example is given.

Following steps are preformed:

- connected components labeling algorithm is first used to find clumps in both sites.

- every cell in both sites inherits two attributes of a clump to which it belongs: class and size (measured in number of pixels).

- three two-dimensional (class, size) histograms are constructed from cell attributes; two from cells restricted to individual sites, respectively, and the third from cells combined from the two sites. Histograms are normalized to add up to 1, so they represent a probability distribution functions.

- informational entropy is calculated for each histogram; the three values of entropies are combined into a single measure corresponding to a mutual information. By construction mutual information (as calculated in LandEx) has a range between 0 (if the two sites are identical) and 1 (if the two sites are dominated by two different classes).

- the similarity measure is taken as \(100(1 - \text{mutual information})\); these values are reported in similarity map produced by LandEx.
4. How similarity is calculated?

4.2 - Clumping

Reference site

Local site #1

Local site #2

1131 clumps

968 clumps

1290 clumps

Example: Reference site and two local sites; each site has size of 150 x150 cells. Top row: land cover pattern. Middle row: clumping (each clump is indicated by a random color). Bottom row: 2-D histograms of cells in each site. Clump sizes are categorized in bins based on the powers of two. High peaks in histograms indicate combination of land cover classes and clump sizes that contain large number of cells.
4. How similarity is calculated?

4.3 - Mutual information: example 1

Calculation of similarity between the reference site and the local site #1 uses three histograms (see above). Visually, these two sites appear fairly similar which is reflected in fairly similar histograms. This also means that a histogram constructed from cells combined from both sites is also fairly similar to the two histograms calculated from individual sites. Entropy is a value that encapsulates a character of a histogram; more distributed (flatter) histograms have larger values of entropy, and more concentrated histograms have smaller values of entropy.

Mutual information (MI) is an information gain; in this case, how much information can be gained (on average) about pattern if only a single site (rather than both sites) is considered. This is calculated as:

\[ MI = H_{R+\#1} - 1/2 (H_R + H_{\#1}) = 0.216778 \]

Thus, on average an entropy of class/clump size distribution is reduced by about 0.22 if calculated from only one of the two sites instead of both. Note that MI=0 if both sites are identical, thus MI serves as a measure of informational “distance” between the two histograms (and thus, the sites). A similarity, defined as 100%(1-MI), can be interpreted as a percentage of information shared by the two histograms (sites). In this example a similarity is about 78.3%
4. How similarity is calculated?

4.4 - Mutual information: example 2

Calculation of similarity between the reference site and the local site #2 uses three histograms (see above). Visually, these two sites appear dissimilar which is reflected in dissimilar histograms. Mutual information (MI) is:

\[
MI = H_{R+#1} - 1/2 (H_R + H_{#1}) = 0.472996
\]

Thus, on average an entropy of class/clump size distribution is reduced by about 0.47 if calculated from only one of the two sites instead of both. A similarity, defined as \(100\%(1-\text{MI})\), is about 52.7\% The two sites are not similar which confirms a visual impression.
LandEx calculates similarity map by imposing an overlapping grid of windows (data blocks) on the entire extent of NLCD2006. Two grids are available.

The coarser grid has 500 x 500 cells (15 km x 15 km) windows that overlap by 100 cells (3 km). In this option the reference site is also 15 km x 15 km The resolution of similarity map resulting from imposing the coarser grid is 3 km. The value of similarity attached to a cell in the coarser grid refers to comparison between 15 km x 15 km reference site to the 15 km x 15 km local site centered on that cell.

The finer grid has 150 x 150 cells (4.5 km x 4.5 km) windows that overlap by 50 cells (1.5 km). In this option the reference site is also 4.5 km x 4.5 km The resolution of similarity map resulting from imposing the finer grid is 1.5 km. The value of similarity attached to a cell in the finer grid refers to comparison between 4.5 km x 4.5 km reference site to the 4.5 km x 4.5 km local site centered on that cell.

LandEx offers several smoothing methods to obtain more diffused similarity maps. These methods work by taking into account similarity values for all local sites overlapping with a given map cell.
6. LandEx development team

**DR. TOMASZ STEPINSKI**  
SPACE INFORMATICS LAB, UNIVERSITY OF CINCINNATI  
215 Braunstein Hall  
Cincinnati, OH 45221, USA  
stepintz@uc.edu

**DR. PAWEL NETZEL**  
SPACE INFORMATICS LAB, UNIVERSITY OF CINCINNATI  
219 Braunstein Hall  
Cincinnati, OH 45221, USA  
netzelpl@ucmail.uc.edu

**DR. JAROSLAW JASIEWICZ**  
GEOECOLOGY AND GEOINFORMATION INSTITUTE ADAM MICKIEWICZ UNIVERSITY  
Dziegielowa 27  
60-680 Poznan, Poland  
jarekj@amu.edu.pl

**JACEK NIESTEROWICZ**  
SPACE INFORMATICS LAB, UNIVERSITY OF CINCINNATI  
219 Braunstein Hall  
Cincinnati, OH 45221, USA  
niestejk@mail.uc.edu