GeoPAT 2.0 – Software for Pattern-Based Spatial and Temporal Analysis of Large Earth Science Datasets

Tomasz Stepinski, Jaroslaw Jasiewicz, Pawel Netzel, Jacek Niesterowicz

Space Informatics Lab
http://sil.uc.edu
University of Cincinnati

ESTF 2017
What is GeoPAT?

GeoPAT (Geospatial Pattern Analysis Toolbox) is a standalone suite of modules written in C and dedicated to analysis of large Earth Science datasets in their entirety using spatial and/or temporal patterns.

Global scale, high resolution spatial datasets are available but are mostly used in small pieces for local studies. GeoPAT enables studying them in their entirety.

High resolution spatial data at local scale tells nothing about global properties.

To comprehend a global character, entire high resolution spatial data needs to be analyzed.
GeoPAT’s core idea

Tessellate global spatial data into grid of square blocks of original cells (pixels). This transforms data from its original form (huge number of cells each having simple content) to a new form (much smaller number of supercells/blocks with complex content).

![Image of original complex cell containing a pattern of original variable.]

GeoPAT provides means for succinct description of such patterns and for calculation of similarity between patterns.

This enables spatial analysis such as clustering, segmentation, and search to be performed on the grid of complex cells (local patterns).

Working with data compressed to necessary minimum GeoPAT can analyze large data using modest resources.
GeoPAT 2 versus GeoPAT 1

GeoPAT 2 is a completely new software (not an update) based on an idea first implemented in GeoPAT 1.

<table>
<thead>
<tr>
<th>ELEMENT</th>
<th>GeoPAT 1</th>
<th>GeoPAT 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>System/Environment</td>
<td>GRASS GIS/ Linux</td>
<td>STANDALONE Linux/Windows</td>
</tr>
<tr>
<td>Input data format</td>
<td>Categorical raster</td>
<td>Any numerical structure</td>
</tr>
<tr>
<td>Type of patterns</td>
<td>spatial</td>
<td>spatial, temporal, spatio-temporal</td>
</tr>
<tr>
<td>Type of analysis</td>
<td>search, change, clustering</td>
<td>Search, change, clustering, segmentation</td>
</tr>
</tbody>
</table>
GeoPAT architecture
What datasets can GeoPAT 2 handle and what analyses it can perform?

DATASETS

1. Land cover (CCI-LC, GlobCover, NLCD, EOSD-Canada)
2. Topography (SRTM, NED)
3. High resolution (1 m/pixel or less) image or LIDAR
4. Global climate (WorldClim)
5. Global phenology (EVI)

ANALYSES

1. Spatial search
2. Change detection
3. Clustering/classification
4. Segmentation/classification
Using GeoPAT for search
Using GeoPAT to search for spatial patterns similar to a query: U.S. land cover

Dataset: National Land Cover Dataset (NLCD) res 30m/cell

You can perform your own search on NLCD at sil.uc.edu/webapps/landex_usa/
Using GeoPAT to search for spatial patterns similar to a query: global topography

Dataset: Shuttle Radar Topography Mission (SRTM) res 90m/cell

Original data: topographic map
Converted data: geomorphons map

Saudi Arabia Empty Quarter
Search of global topography continue

Examples computed online by TerraEx (http://sil.uc.edu/webapps/terraex/) (beta)
Using GeoPAT to search for temporal patterns similar to a query: global climate

Dataset: WorldClim res 4km/cell

You can perform your own search on global climate at sil.uc.edu/webapps/climateex
Using GeoPAT to calculate change
Using GeoPAT to calculate change: U.S. land cover

Dataset: NLCD 2001 and 2006 res 30m/cell

You can explore this map of change in more details at sil.uc.edu/webapps/dataeye_usa/
Using GeoPAT to calculate change: Global climate

Dataset: WorldCilm 2000 and 2070 res 4 km /cell

Climate change 2000 -2070

You can explore global climate change at more details at sil.uc.edu/webapps/climateex
Using GeoPAT for clustering
Using GeoPAT to find regions with similar spatial pattern via clustering: U.S. land cover

Dataset: NLCD res 30 m /cell

Clustering all local patterns of NLCD into nine clusters yields a generalization of US land cover.
Using GeoPAT to find regions with similar spatial pattern via **clustering**: Global climate

Dataset: WorldClim res 4 km /cell

Global climates clustered using a concept of climate as time series (temporal pattern) and utilizing the Dynamic Time Warping (DTW) as similarity function.

Standard, Koppen-Geiger climate classification
Using GeoPAT for segmentation

Illustrating pattern-based segmentation
Pattern-based segmentation – perhaps the most useful analysis using GeoPAT

Segmentation on the basis of homogeneity of spatial patterns

Segmenting land cover (NLCD) pattern. Each segment is characterized by homogeneous pattern of land cover categories.

Segmentation on the basis of homogeneity of temporal patterns

Segmenting climate (WorldClim). Each segment is characterized by homogeneous climate.
Using GeoPAT to **segment** and classify land cover patterns worldwide

**Dataset:** Land cover (CCI-LC) res 300 m /cell

Calculations for entire world, western Europe is shown for details

~100,000 segments delineated worldwide using 9 km as local scale of a pattern

Segments worldwide classified to ~600 categories based on pattern
Using GeoPAT to segment and classify urban structure types

Dataset: LIDAR res 12 point/m²

LIDAR Classification of LIDAR cells
Pattern Segmentation (GeoPAT)
Classification of segments

22000 X 21000 raster based on LIDAR measurement in a city of Poznan is used to derive urban structure types (USTs)
Using GeoPAT to **segment** climates worldwide

Dataset: WorldClim 4 km/cell

~24,100 segments labeled by random colors, area of each segment has homogeneous climate.

**Important result:** Climate is much more inhomogeneous in equatorial regions than suggested by classical Koppen-Geiger classification. This inhomogeneity is due to large differences in amount and timing of precipitation.
Using GeoPAT to **segment** phenology

Dataset: EVI time series 4 km/cell

~134,100 segments labeled by random colors, area of each segment has homogeneous phenology *(annual progression of EVI)*.

**Important result:** First regionalization of phenology worldwide. Could be compared to regionalization of climate *(next slide)*.
Climate versus phenology

The redder the color the more inhomogeneous is the segment

Inhomogeneity of phenology in climate segments

This region has relatively homogeneous climate but varies in phenology

Amazon region is characterized by a single phenology but many different climates (vegetation annual progression insensitive to differences in precipitation)

Inhomogeneity of climate in phenology segments
Summary

GeoPAT is an idea-driven rather than data-driven software.

GeoPAT fills a niche devoted to analysis of large spatial datasets in their entirety.

GeoPAT achieves its goal by starting from radical compression of original data leaving only information necessary to complete its tasks.

GeoPAT does not require supercomputer or cloud computing because core calculations are performed on a grid orders of magnitude smaller than a grid of original data. It works well on a moderate server.

GeoPAT is an open and free software.

Binaries for Linux and Windows are available for download.
Examples of public data that could be used by GeoPAT

1. Land cover (NLCD, GLC30, GlobCover, CORINE, EOSD)
2. Topography (NED, SRTM)
3. Croplands (USDA CropScape)
4. Urban Structure Types (UST) (National Map, LIDAR)
5. Global climate (WorldClim)
6. Phenology data (EVI)

Examples of problems that could be addressed using GeoPAT

1. Ecology (delineation of ecoregions, climate versus phenology)
2. Forestry (conservation planning and management)
3. Agriculture (identify spatio-temporal patterns of usage of U.S. croplands)
4. Climate (climate change)