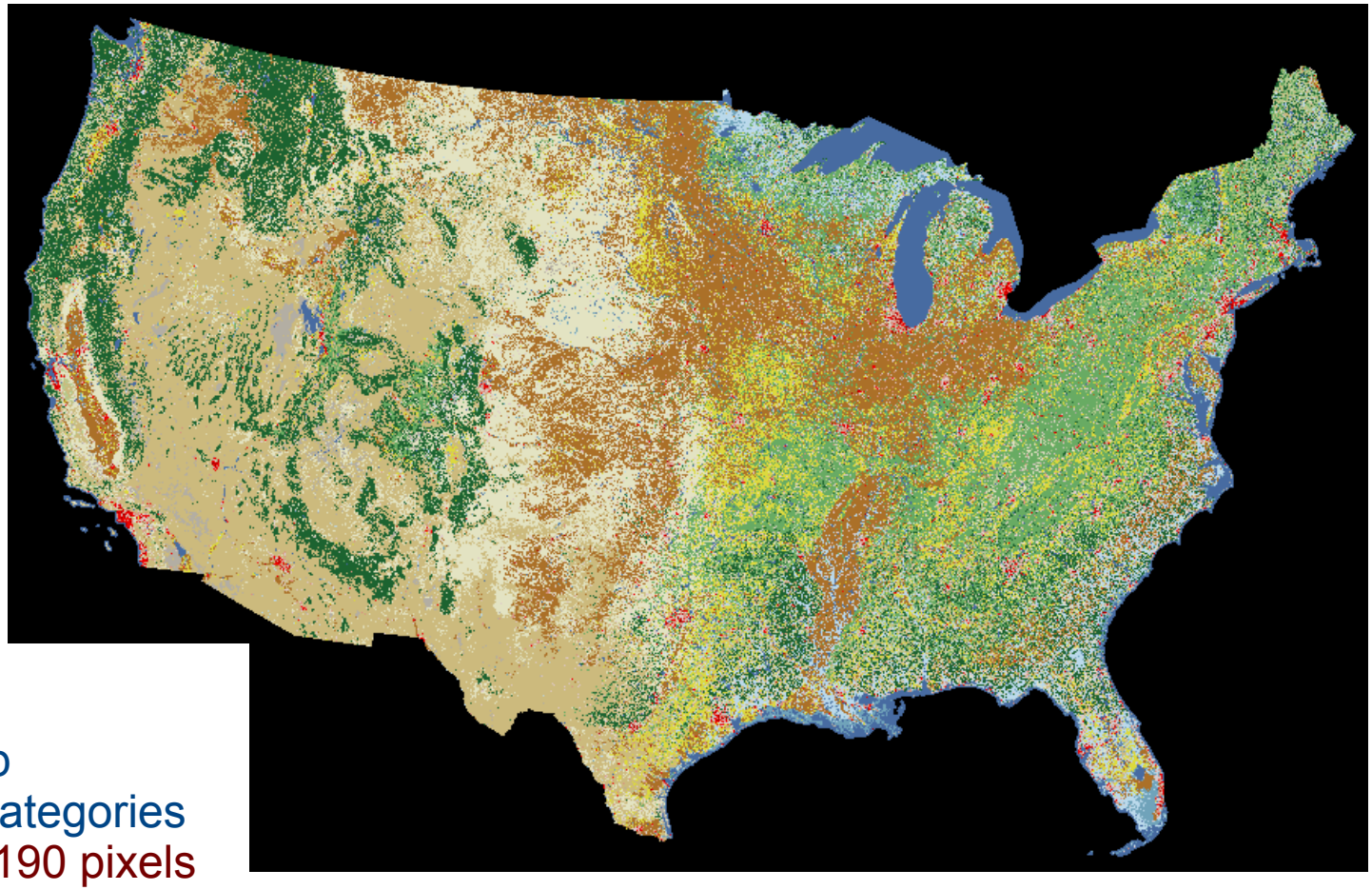
GeoPAT – Pattern-based GIS Software for understanding content of large Earth science datasets

Tomasz Stepinski, Jaroslaw Jasiewicz, Pawel Netzel, Jacek Niesterowicz

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

ESTF 2016

Example of a dataset we are interested in



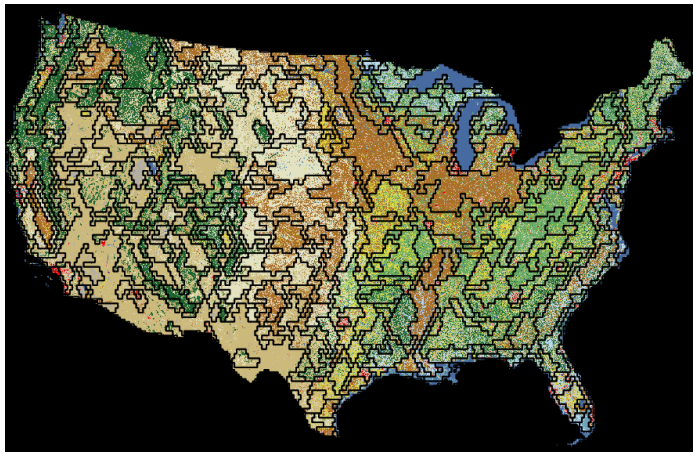
NLCD2011
categorical map
16 land cover categories
104,242 x 161,190 pixels

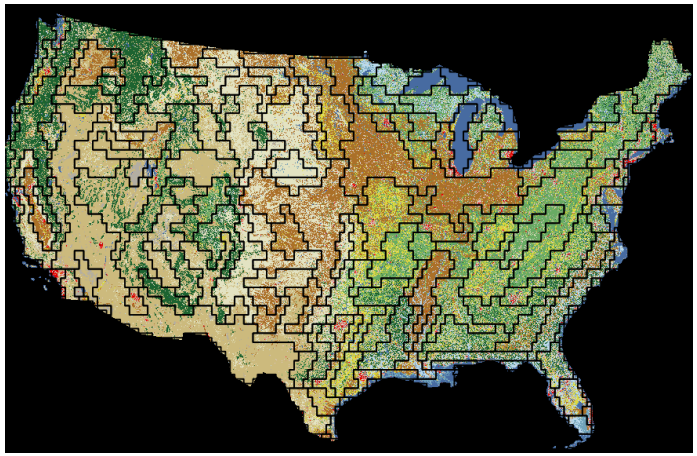
How to analyze such dataset in its entirely?
What relevant information can we extract from such dataset?
What are other, similar datasets?
What are the applications for such analysis?

Some answers…

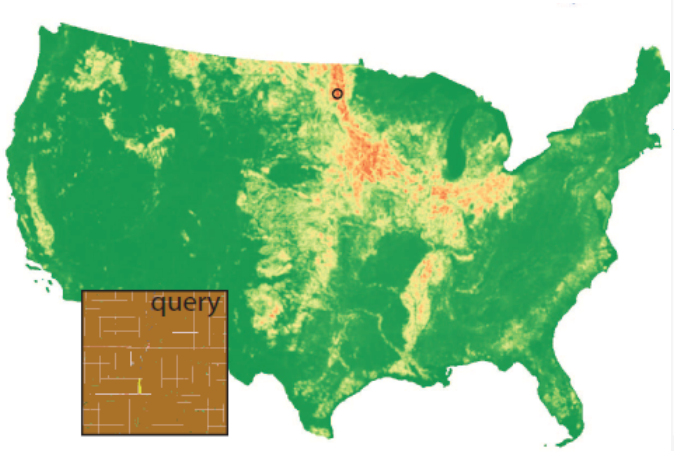
Tools relevant to analyzing NLCD in its entirety:

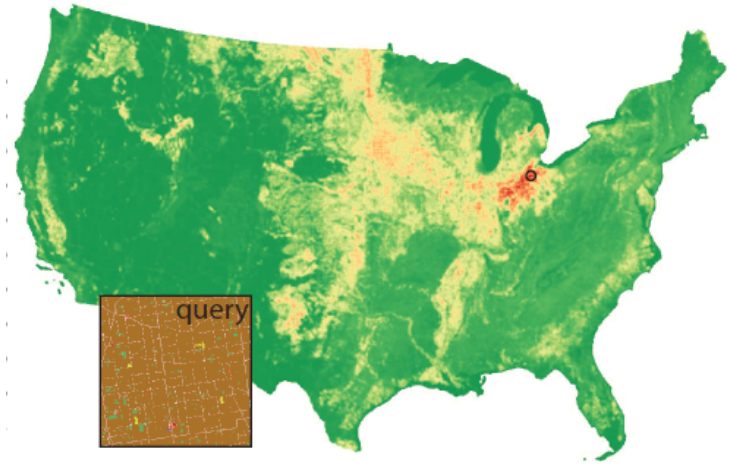
Segmentation into regions of homogeneous land cover patterns



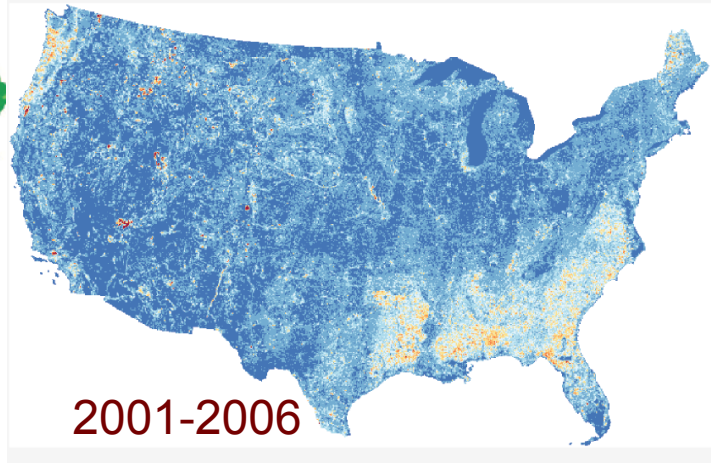


Spatial search

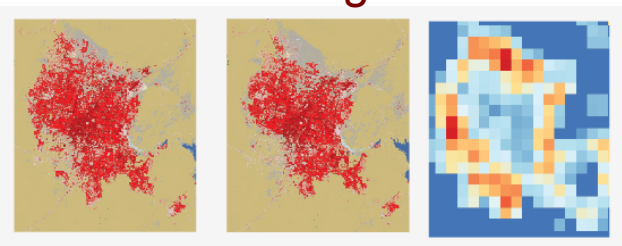




Temporal change

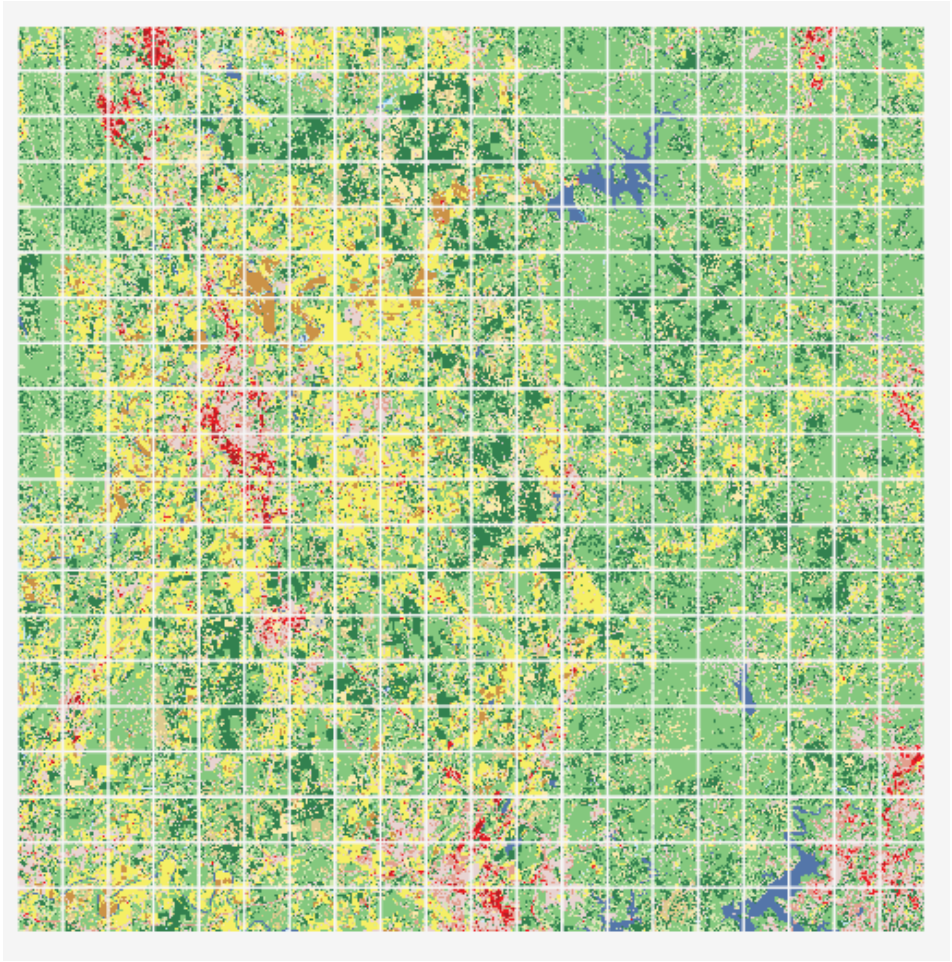


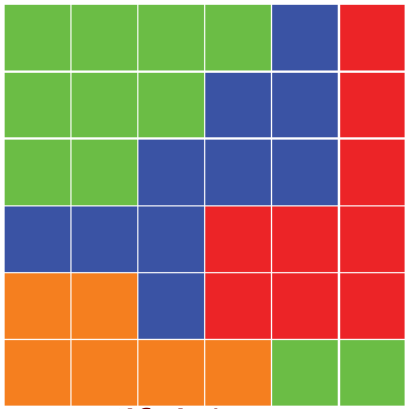
Las Vegas



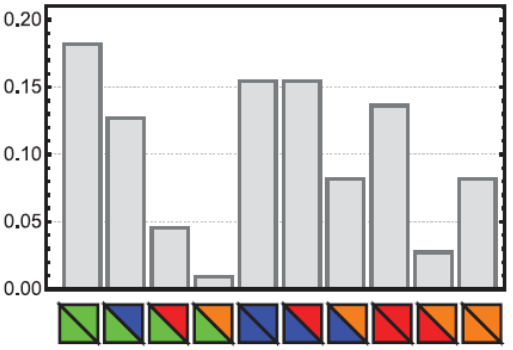
2001 2006 change

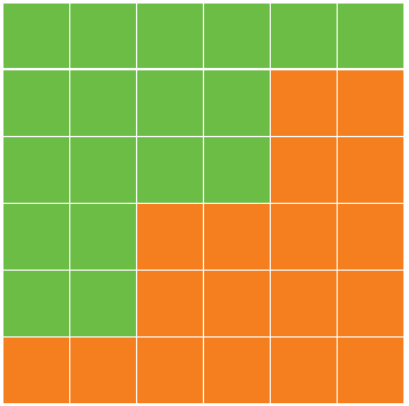
Basic concepts



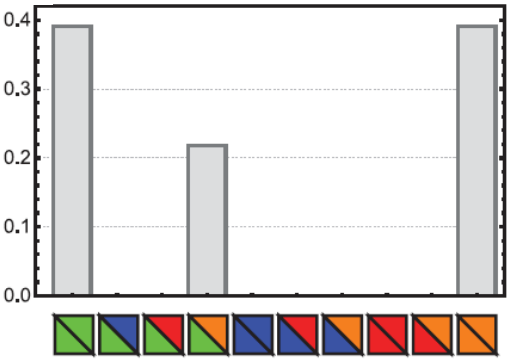


motifel 1





motifel 2



histograms of features

Grid of motifels

Motifel – elementary unit of analysis - a square block of pixels representing local landscape

Motifel is represented by histogram of features (co-occurrence, decomposition)
Distance between motifels is a distance between their histograms (Jensen-Shannon Divergence

Difference between our method and object-based image analysis (OBIA)

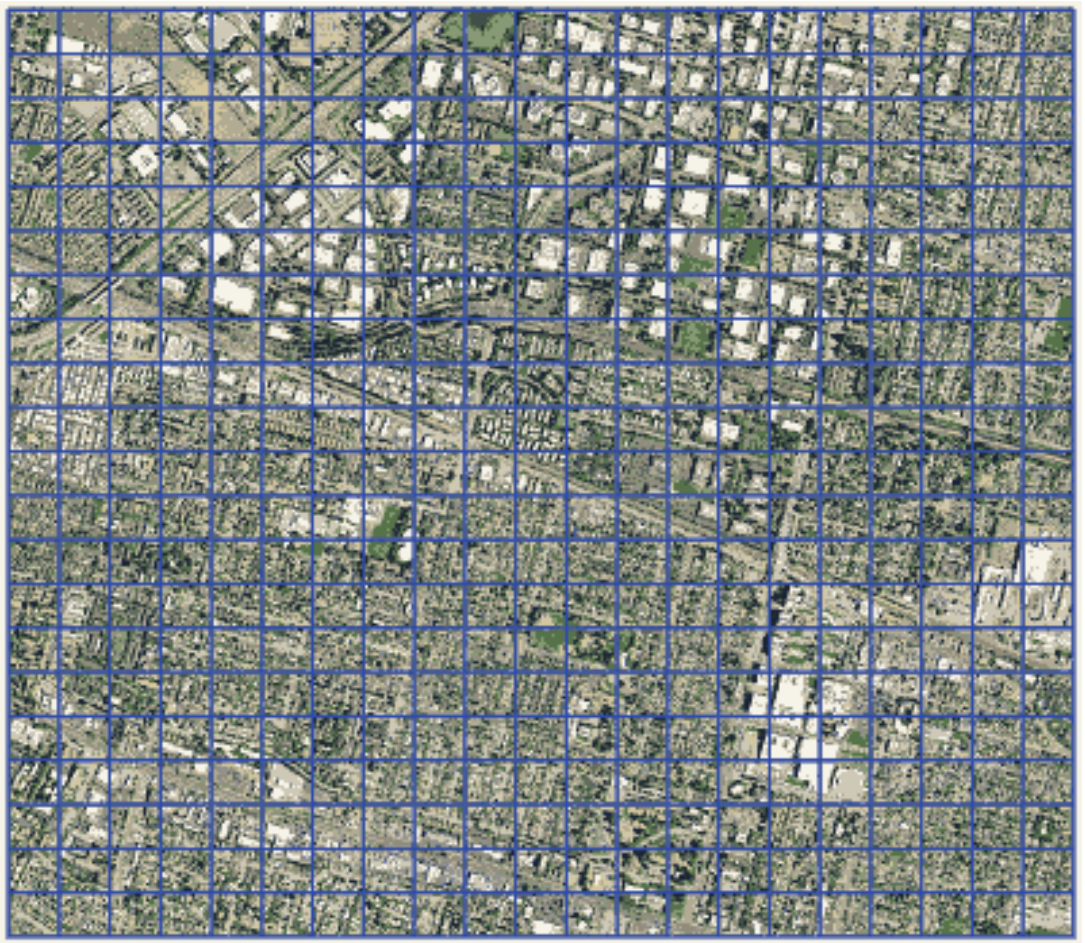
object-based image analysis

works best for small scenes
identifies individual objects
well-eveloped

irregular objects having simple, homogeneous content

complex object-based image analysis

works best for large scenes
identifies types of neighborhoods
our project



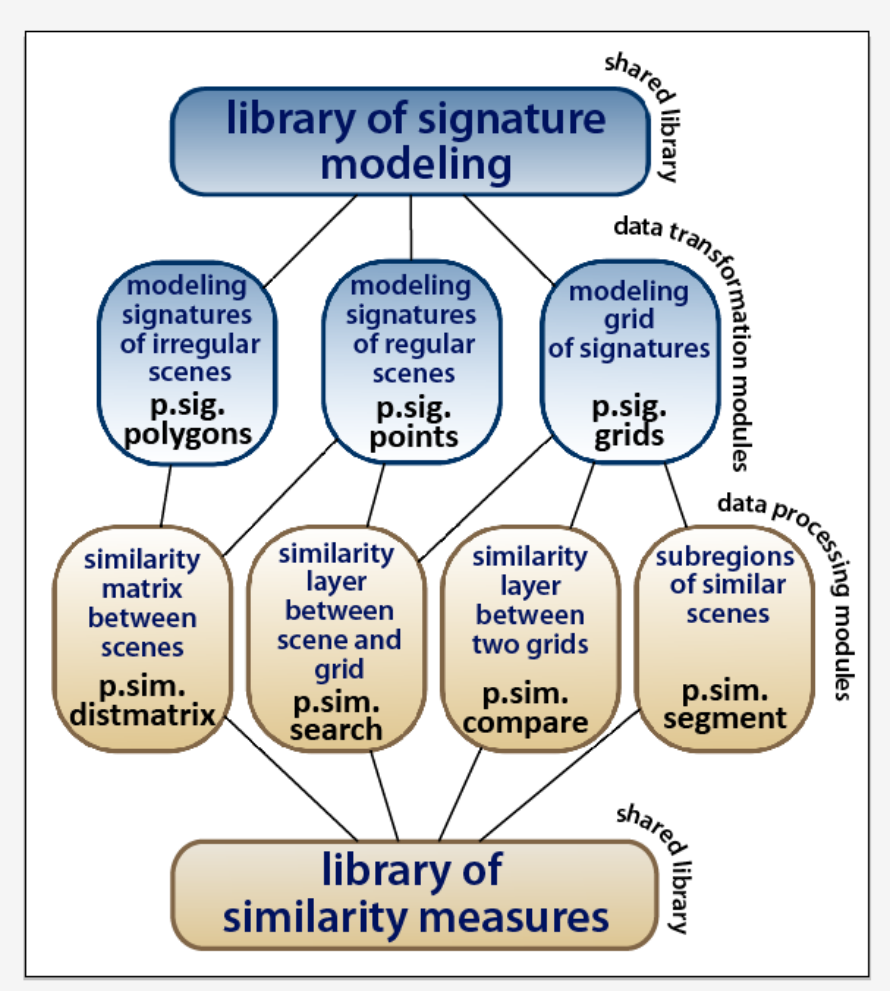
regular objects having complex, inhomogeneous content

GeoPAT

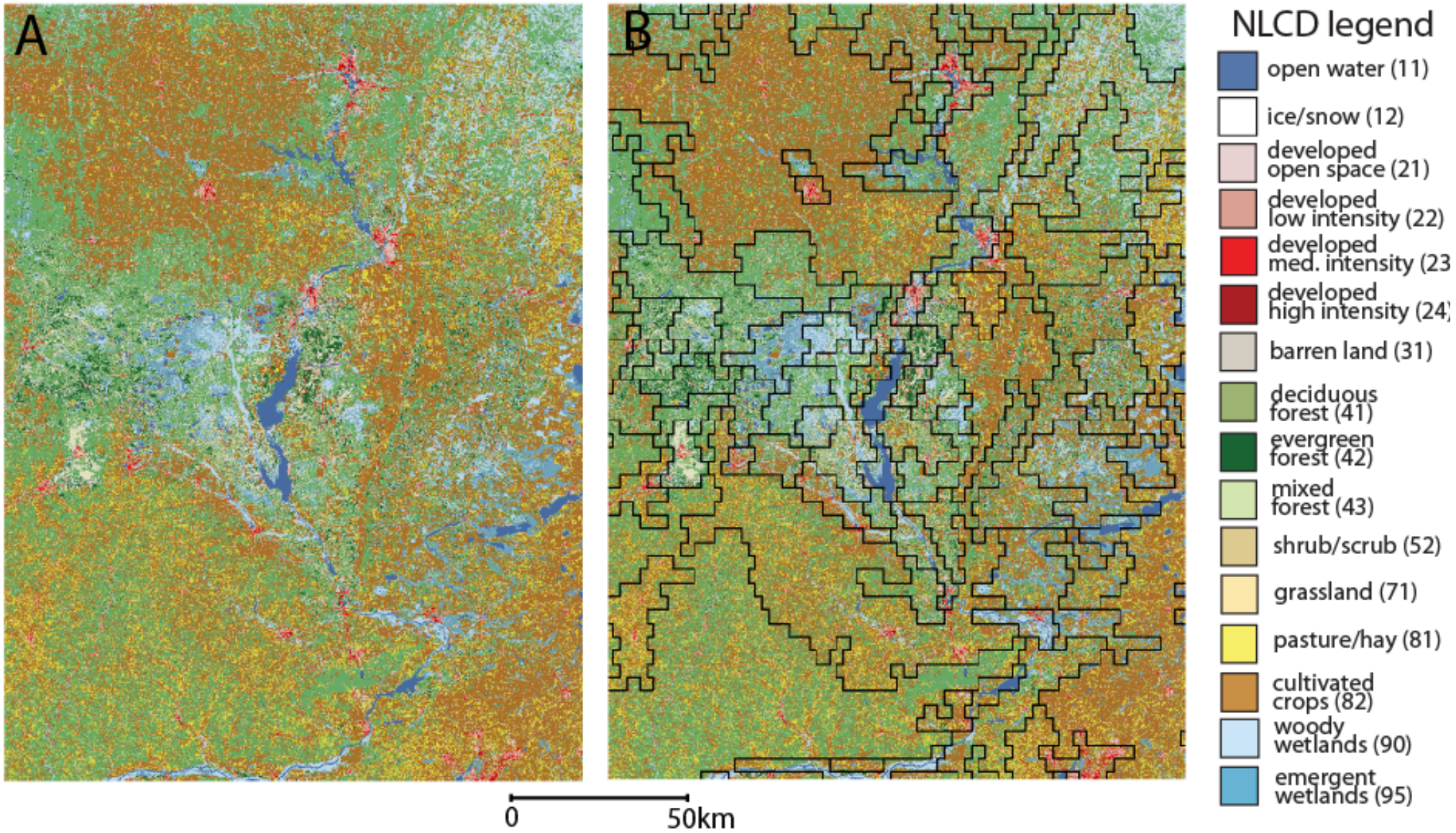
Geospatial Pattern Analysis Toolbox (GeoPAT) is a collection of GRASS GIS modules, written in C, for carrying out pattern-based geospatial analysis of large categorical spatial datasets.

We are developing GeoPAT 2, a stand alone version of our software, which does not require GRASS or Linux.

The new version has expanded functionalities, features brand new segmentation module, and will be optimized to work with Xeon Phi co-processor.



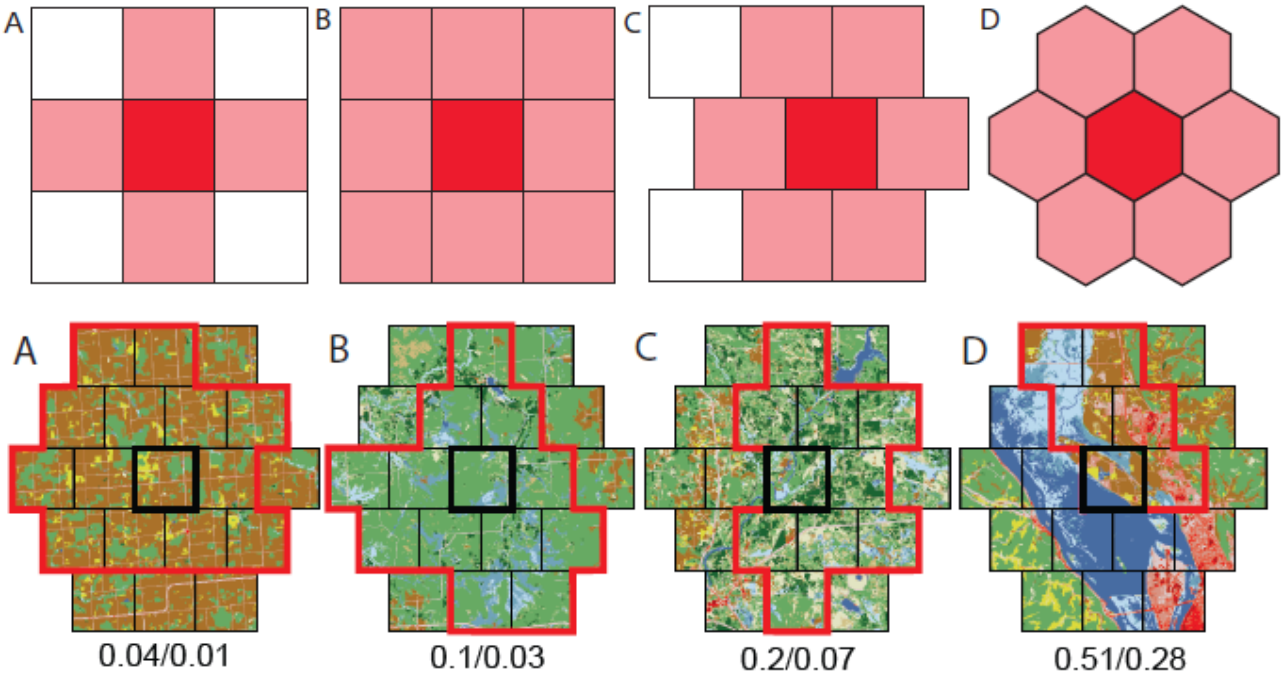
Segmentation 1 of 8



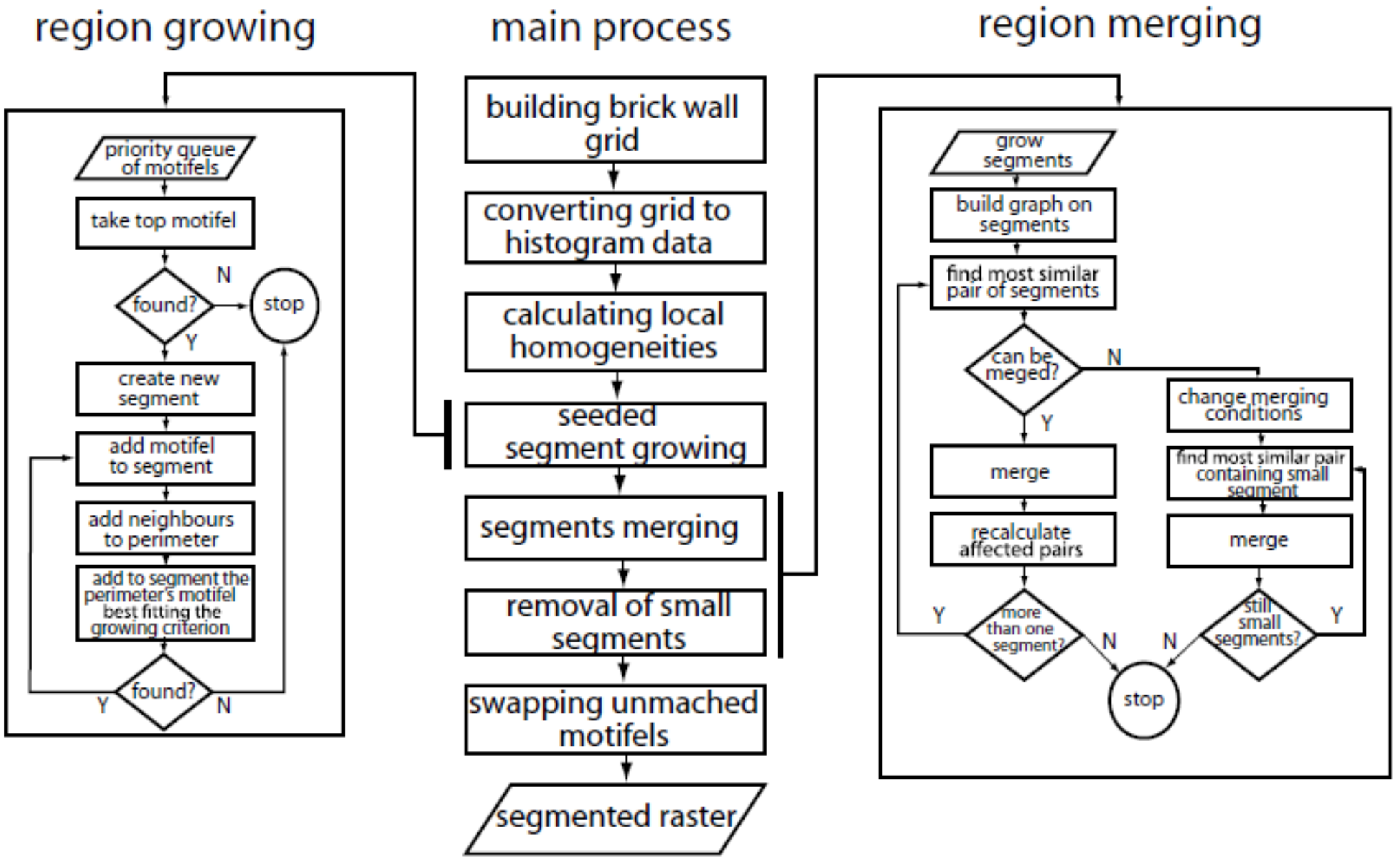
Segmentation 2 of 8

GeoPAT segmentation algorithm is based on the principle of seeded region growing (SRG) but introduces a number of novel features in order to make the SRG applicable to the task of pattern-based segmentation of categorical rasters:

1. Motifels instead of pixels (new representation and distance)
2. Brick-wall topology of grid
3. Novel method of constructing a priority queue for potential seeds and adoption of locally determined growth-stopping criteria for regions growing from these seeds



Segmentation 3 of 8

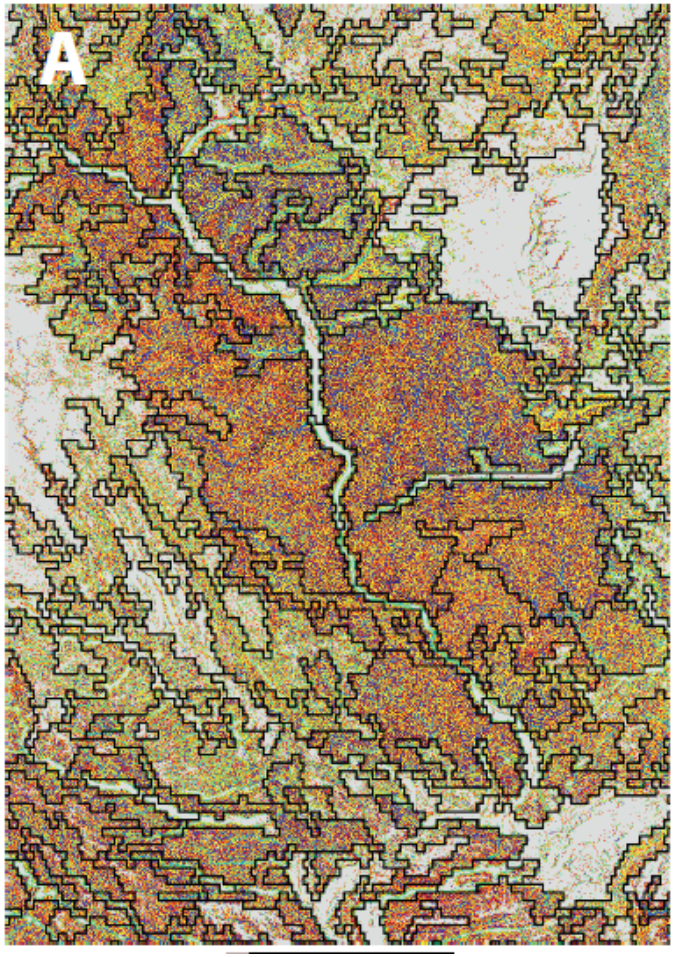


Segmentation 4 of 8

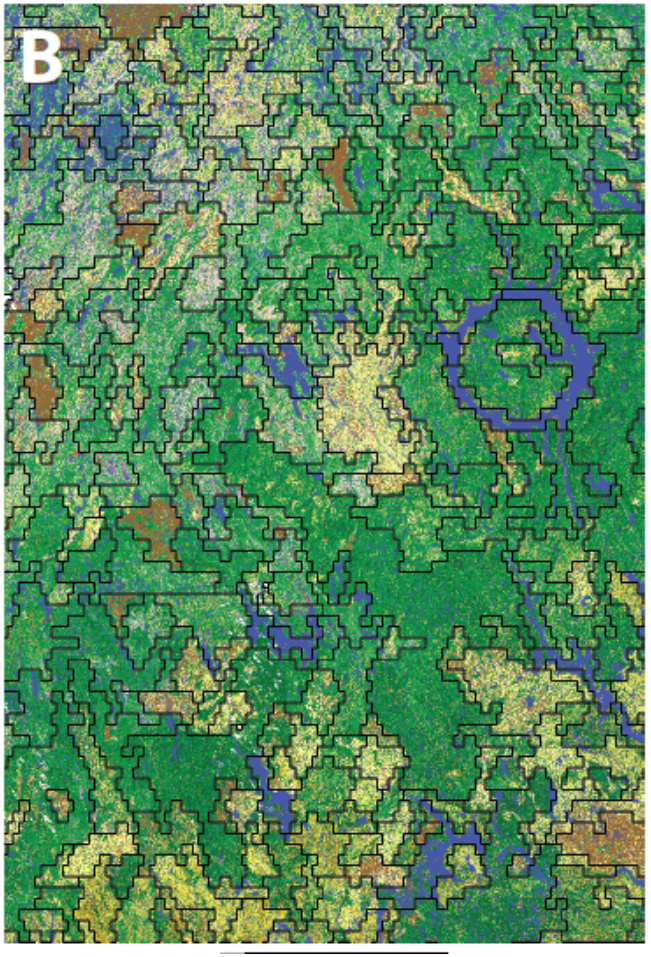
Dataset	Data size	Motifel	# segments	Total time
NLCD (exam.)	16896×16192	64	3135	3m 34s
		128	605	58s
		256	155	33s
		512	63	33s
NLCD2011 U.S.	104242×161190	128	24097	1h 48m
		256	6130	33m 11s
		512	1799	12m 3s
Topo. U.S.	104242×161190	128	26041	2h 11m
		512	4041	13m 24s
EOSD Canada	224400×130800	400	3399	33m 17s
LA image	41600×50200	200	2894	1h 23m

All calculations were performed on computer with Intel 3.4GHz, 4-core processor and 16 GB of memory running the Linux operating system.

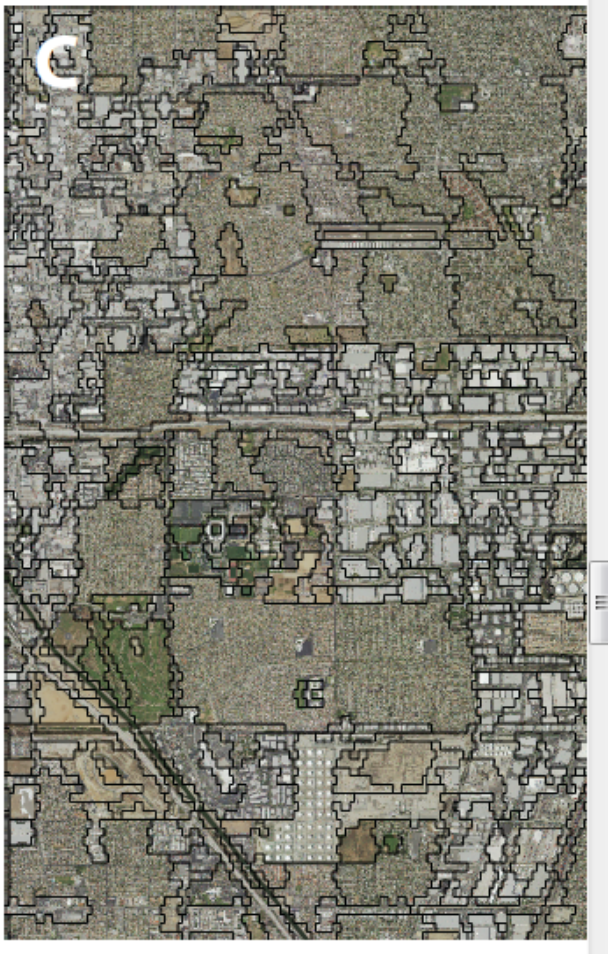
Segmentation 5 of 8



US topography 30m

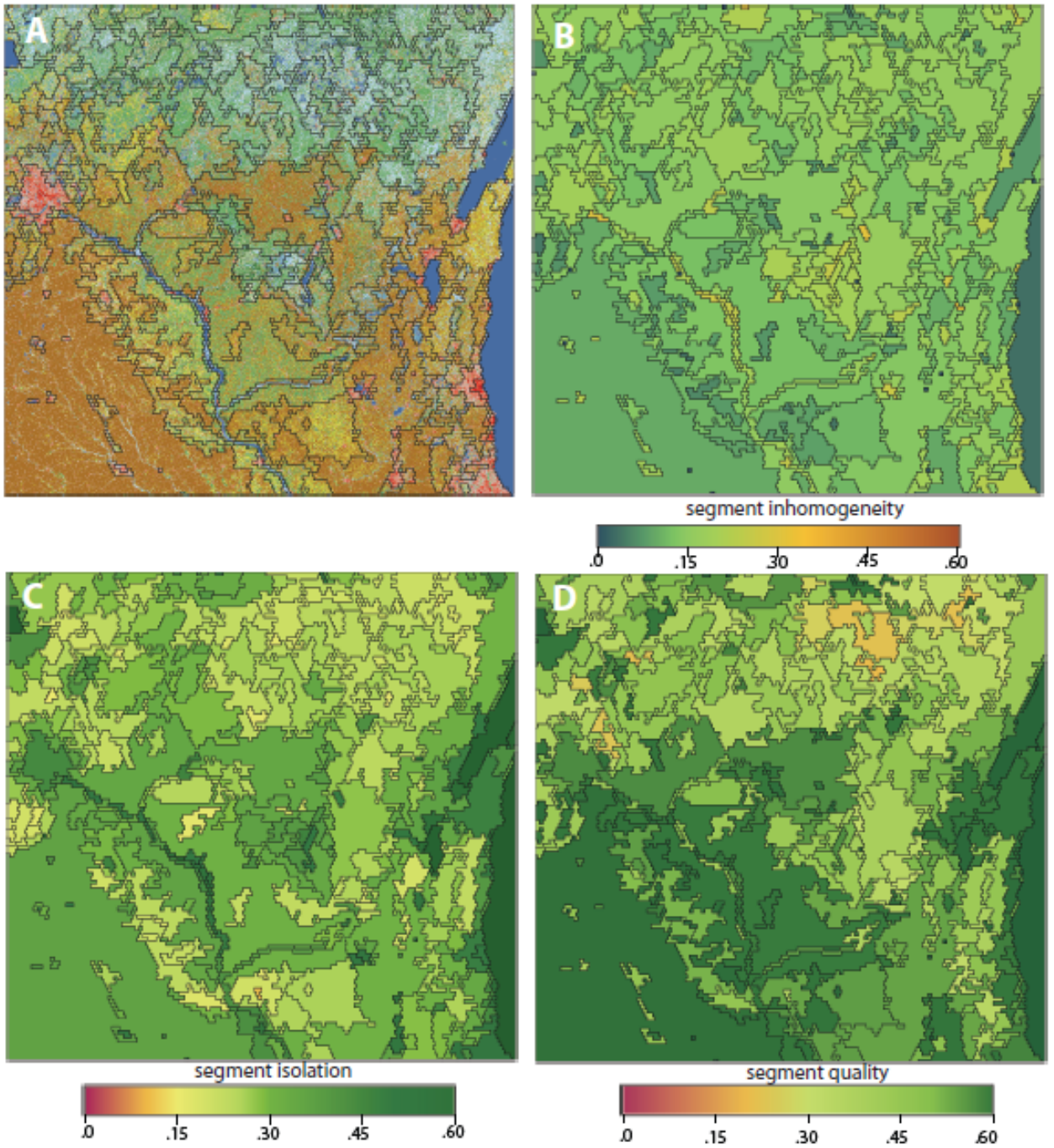


Canadian forest 30m



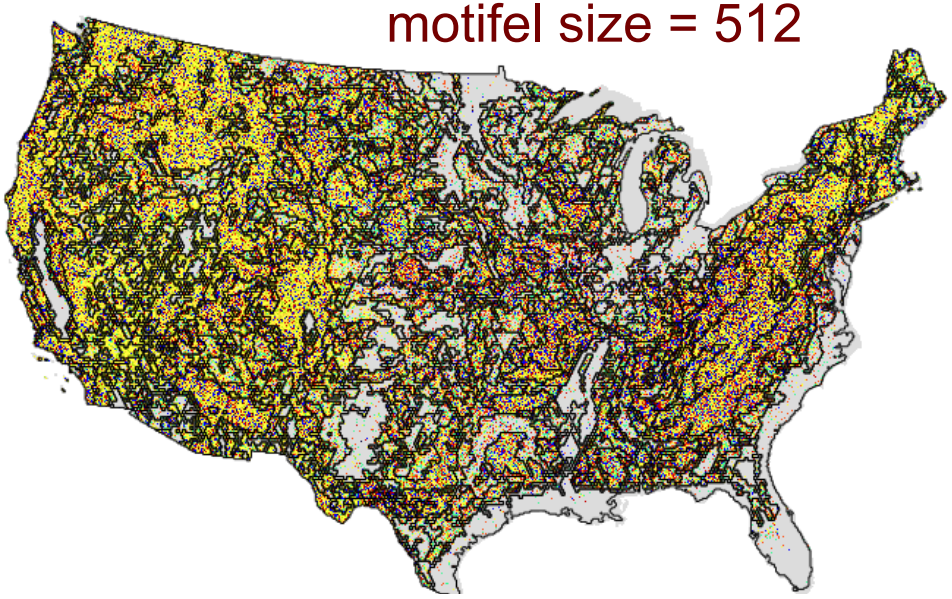
Los Angeles image 1m

Segmentation 6 of 8

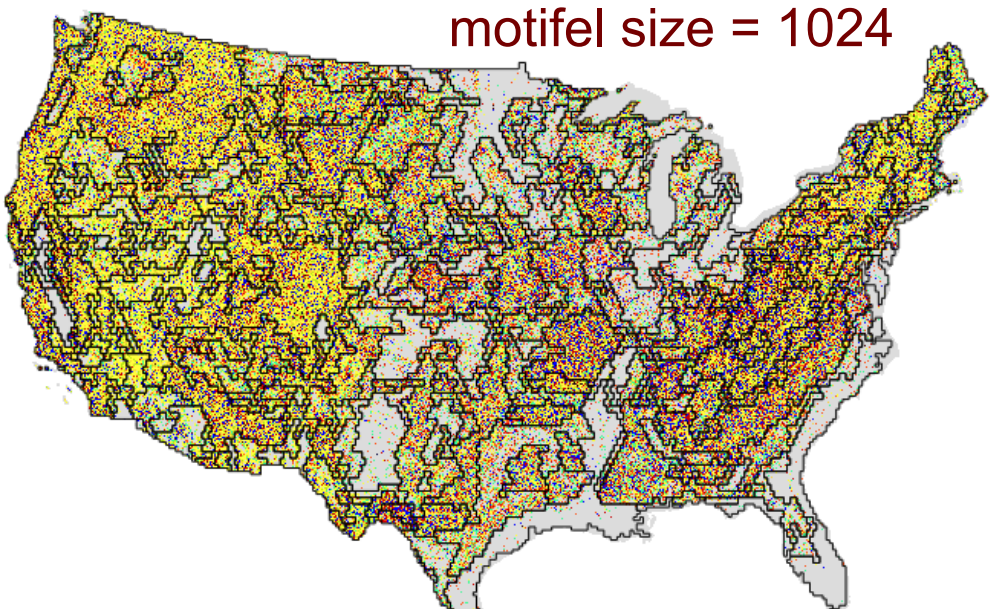


Segmentation 7 of 8

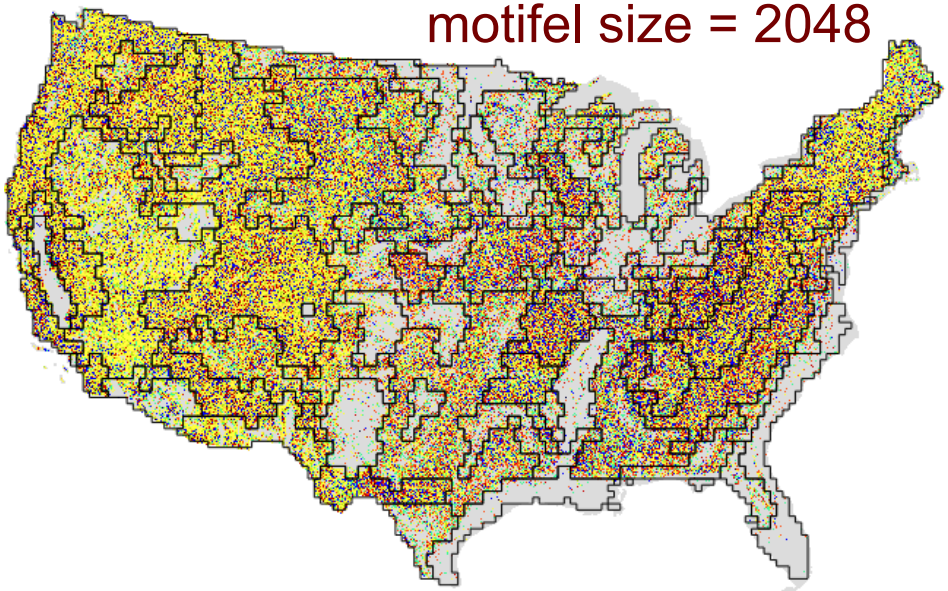
motifel size = 512



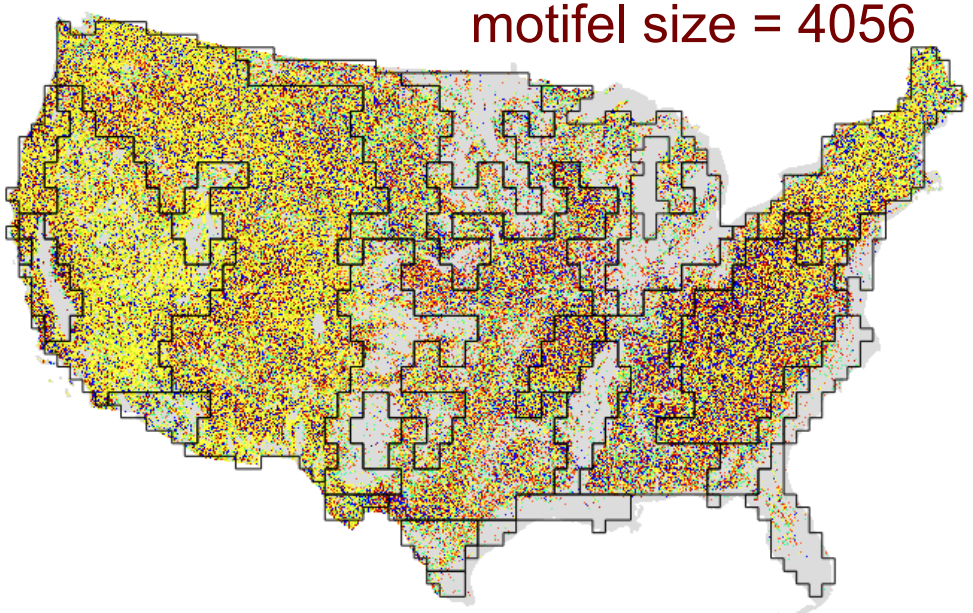
motifel size = 1024



motifel size = 2048



motifel size = 4056



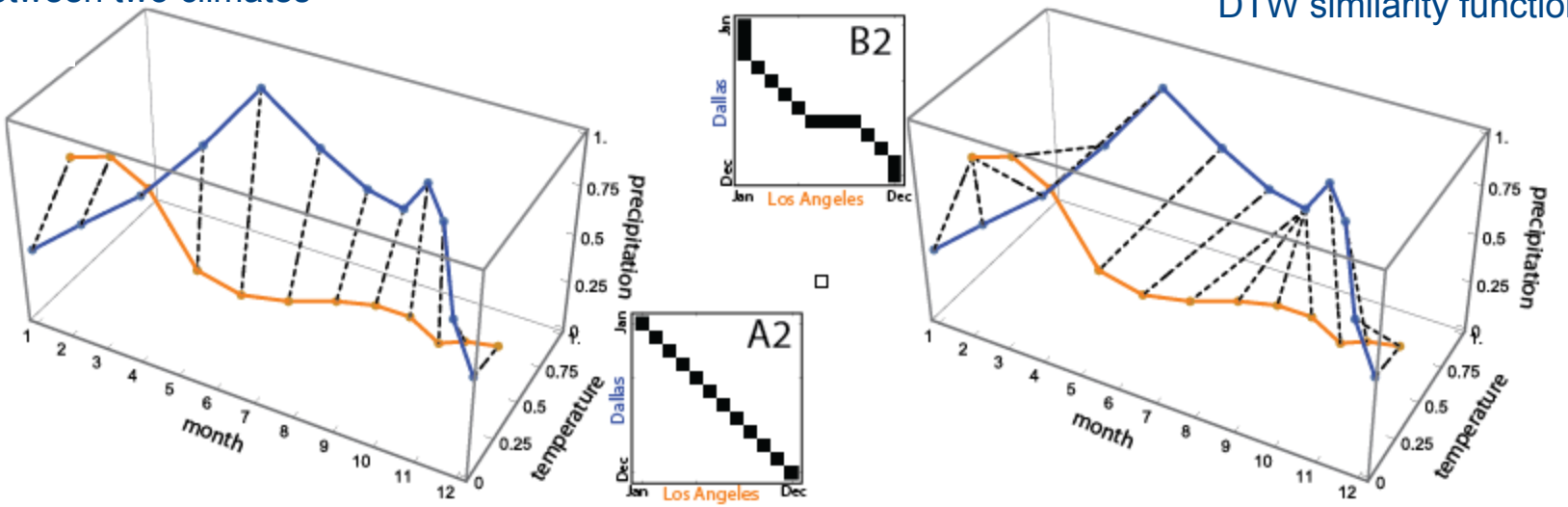
Multi-scale segmentation by changing the size of motifel

Segmentation 8 of 8

Climate, Dynamic Time Warping

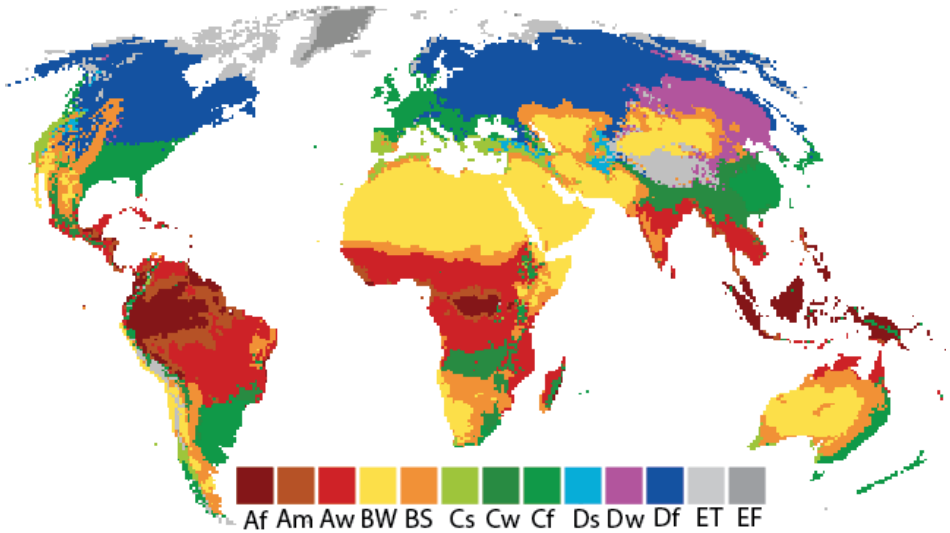
Standard, Euclidean similarity function between two climates

DTW similarity function

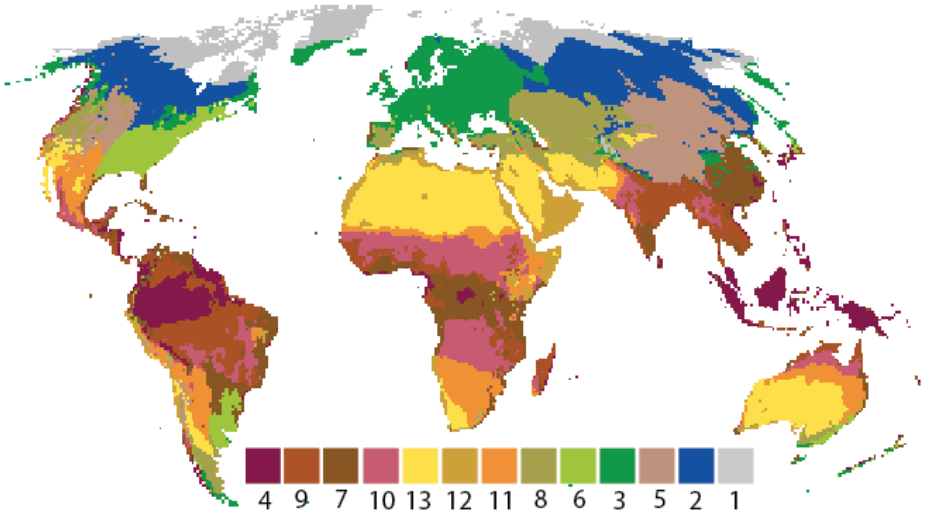


We calculate regionalization of world climates using a concept of climate as time series (temporal pattern) and utilizing the Dynamic Time Warping (DTW) as similarity function

Standard climate classification: Koppen-Geiger



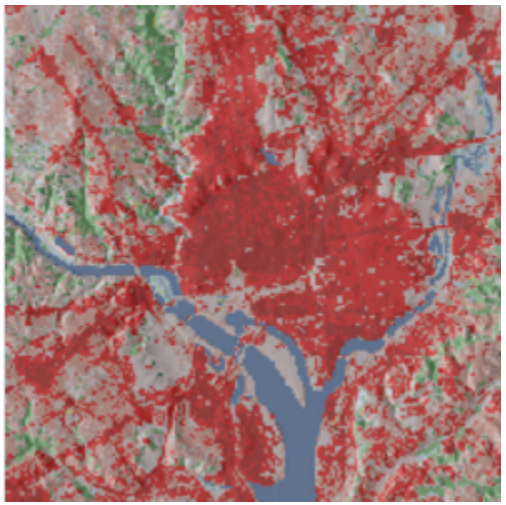
Climate classification based on temporal patterns



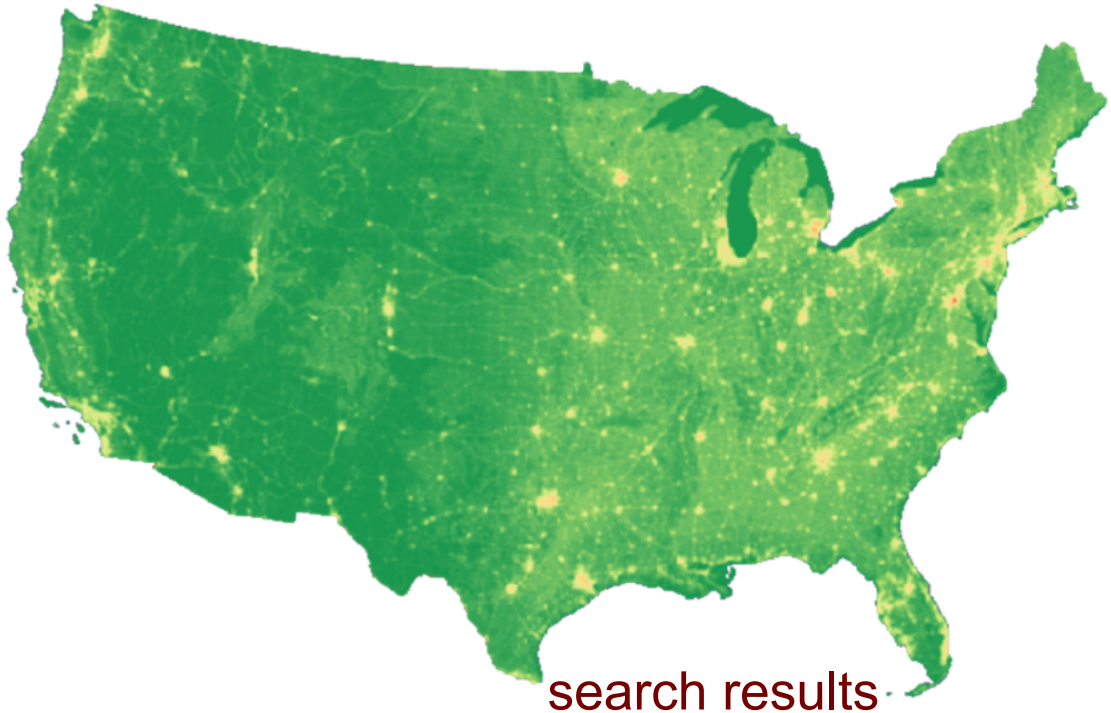
Search 1 of 5

GeoPAT is used to implement pattern-based searches as GeoWeb applications:

1. LandEx – search for similar patterns of land cover over the U.S. sil.uc.edu/webapps/landex_usa/
2. ClimateEx – search for similar climates over the world. sil.uc.edu/webapps/climatex/
3. TerraEx – search for similar landforms over the world (beta) sil.uc.edu/webapps/terraex/

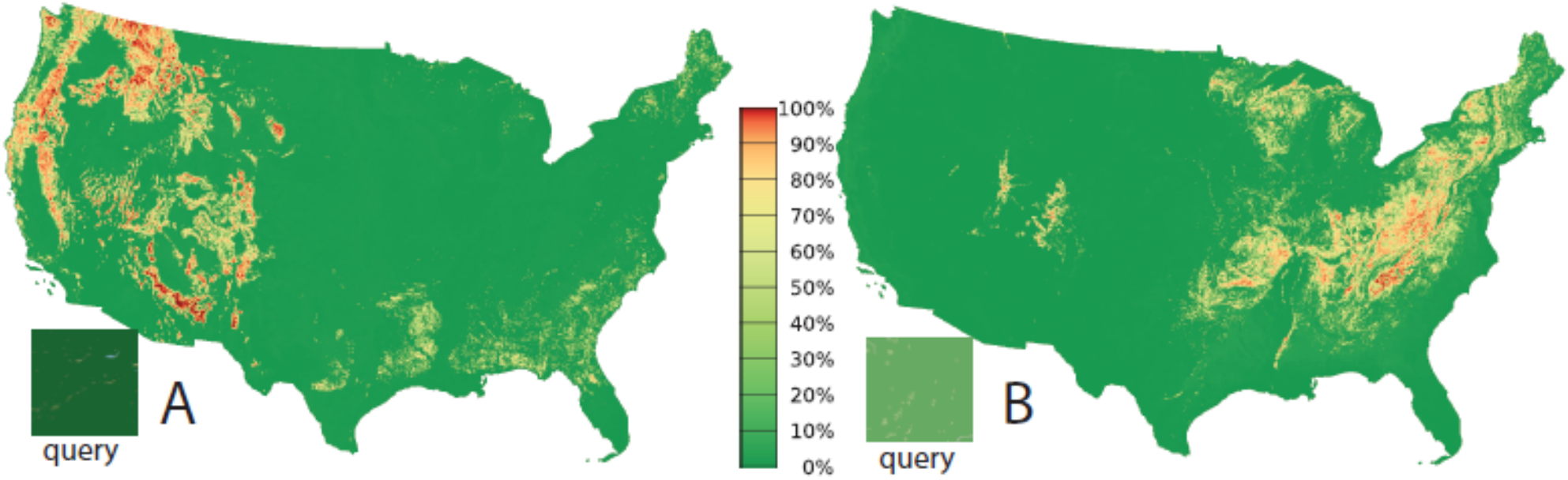


query – Washington DC



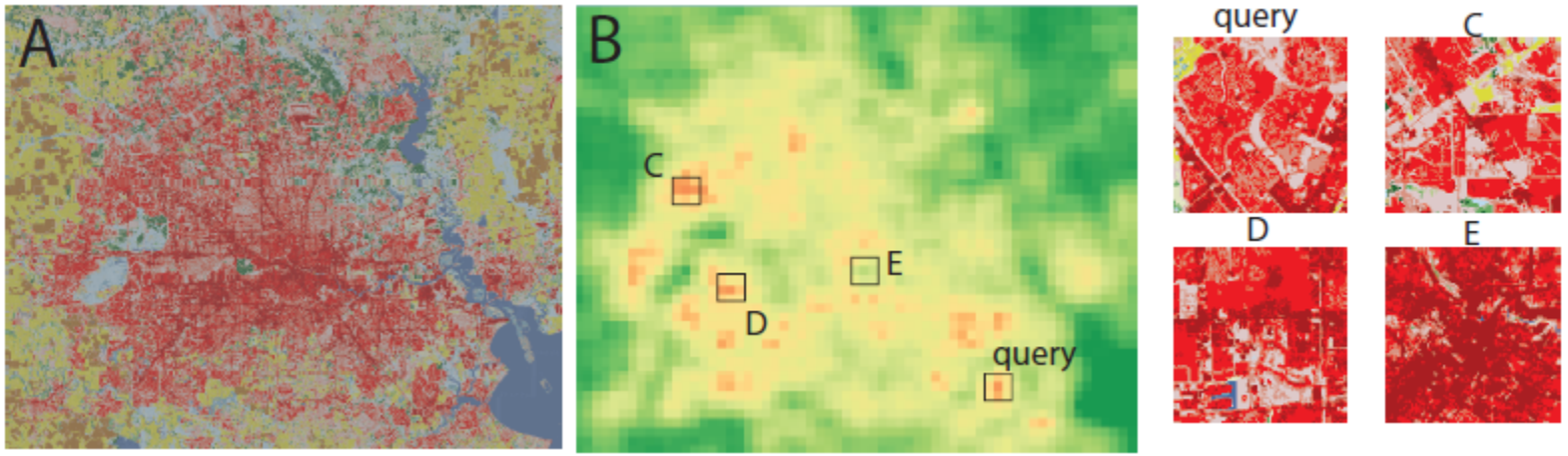
search results

Search 2 of 5



consolidated evergreen forest

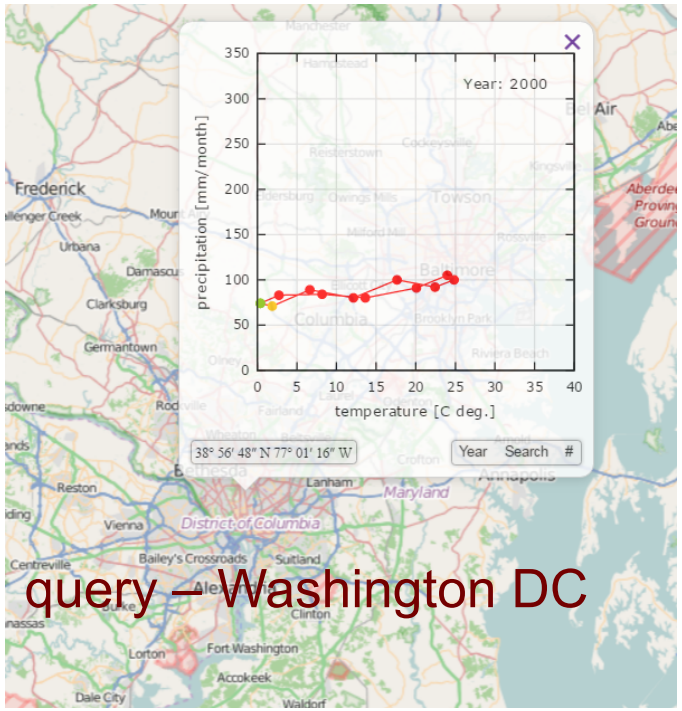
consolidated deciduous forest



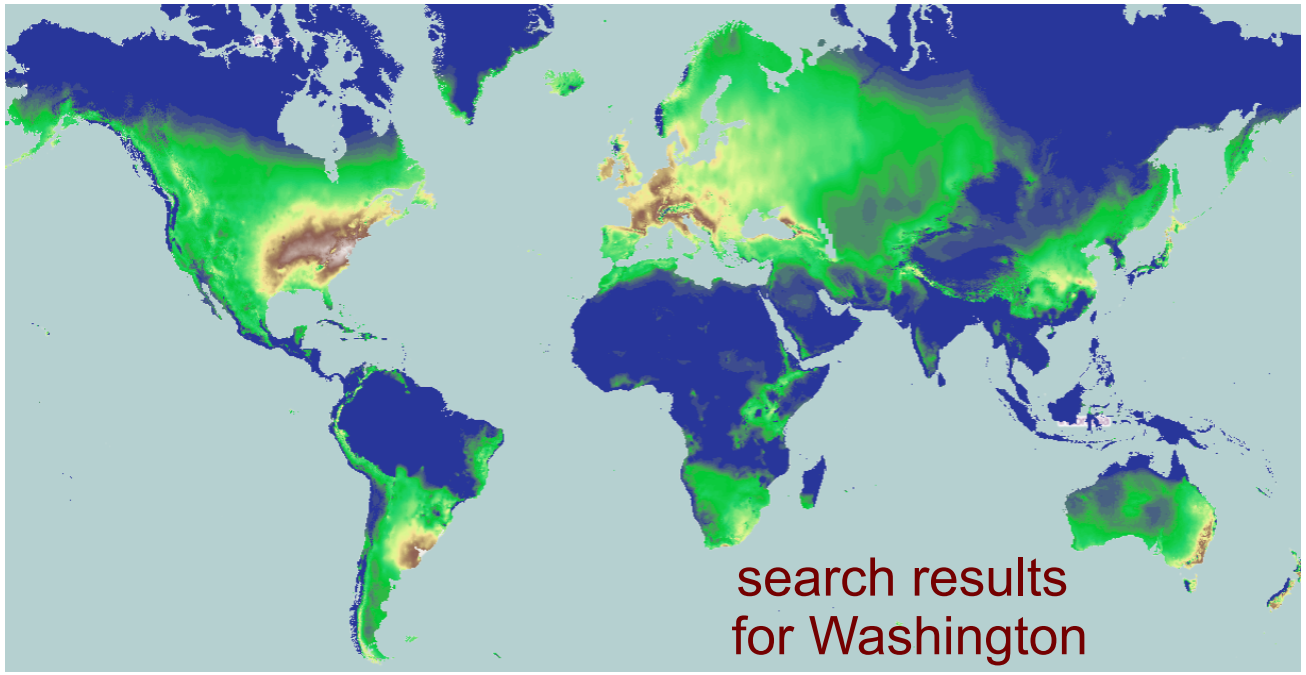
Search of similar urban landscapes in Houston, TX

all examples computed online by LandEx (sil.uc.edu)

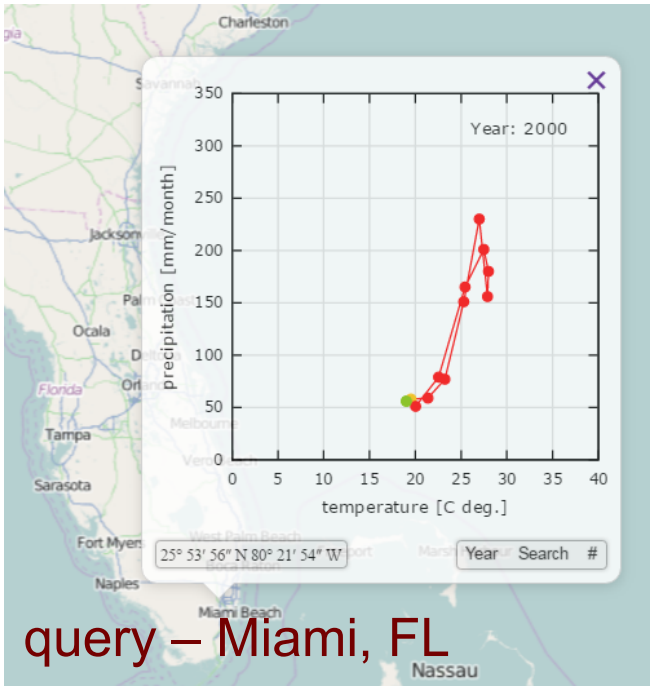
Search 3 of 5



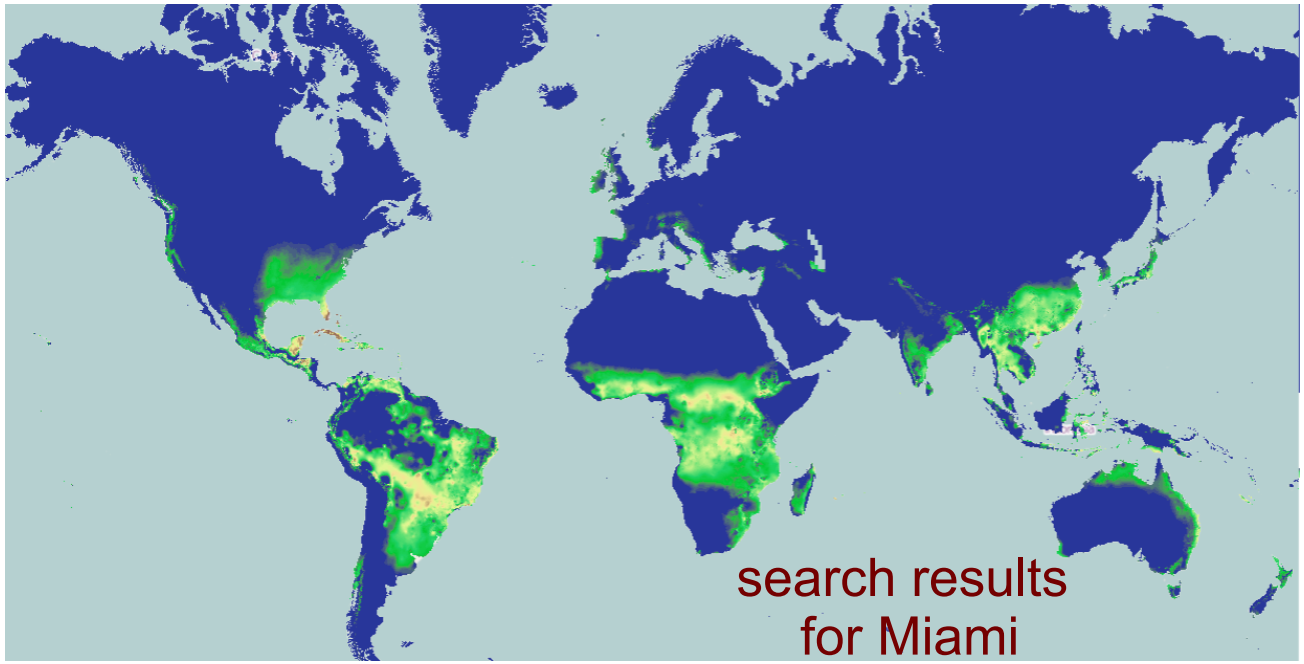
query – Washington DC



search results for Washington



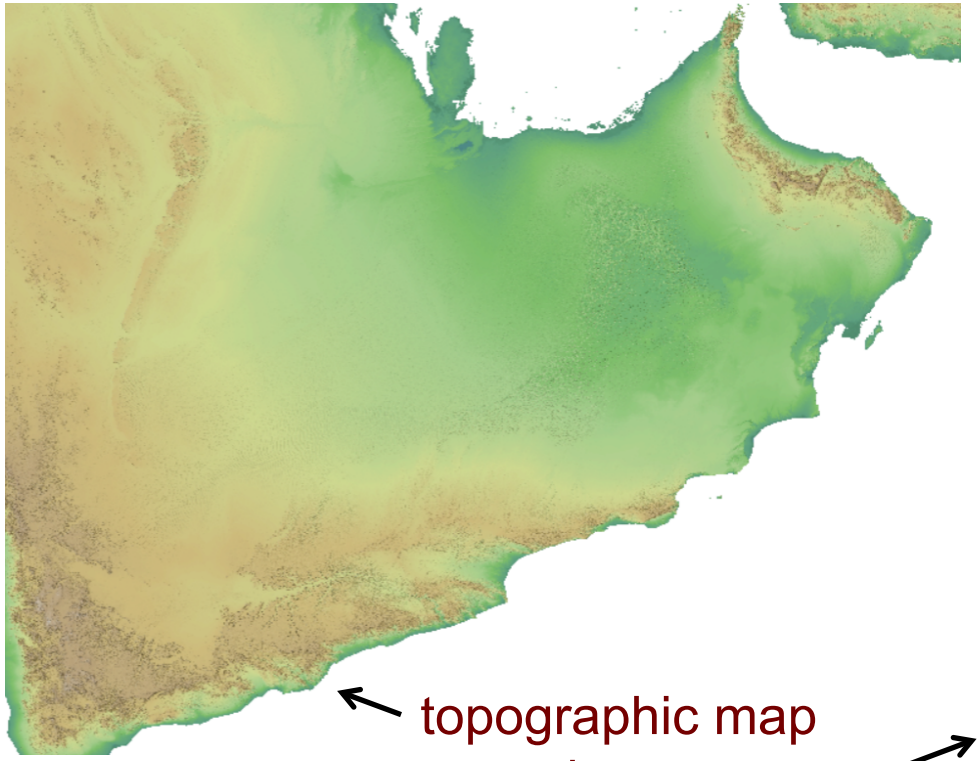
query – Miami, FL



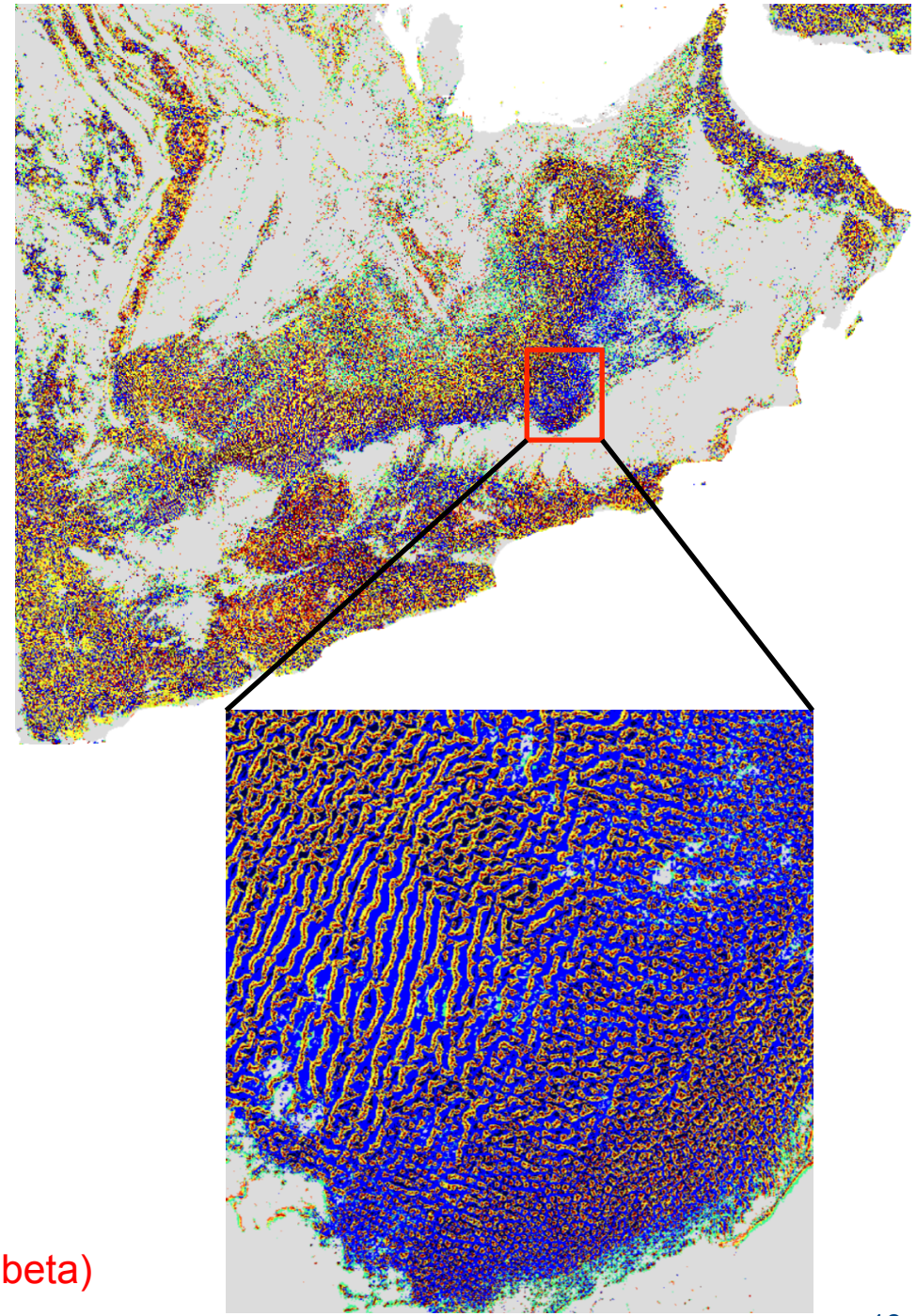
search results for Miami

all examples computed online by ClimateEx

Search 4 of 5



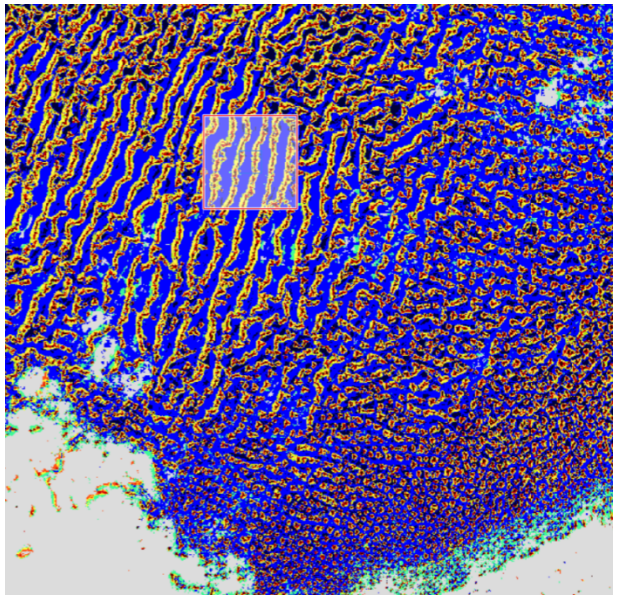
topographic map
geomorphons map

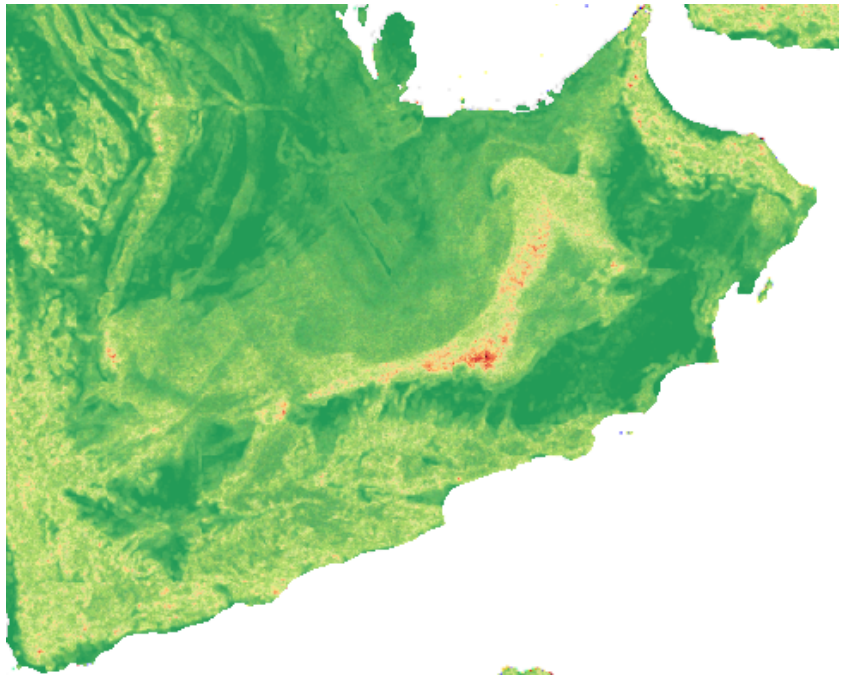


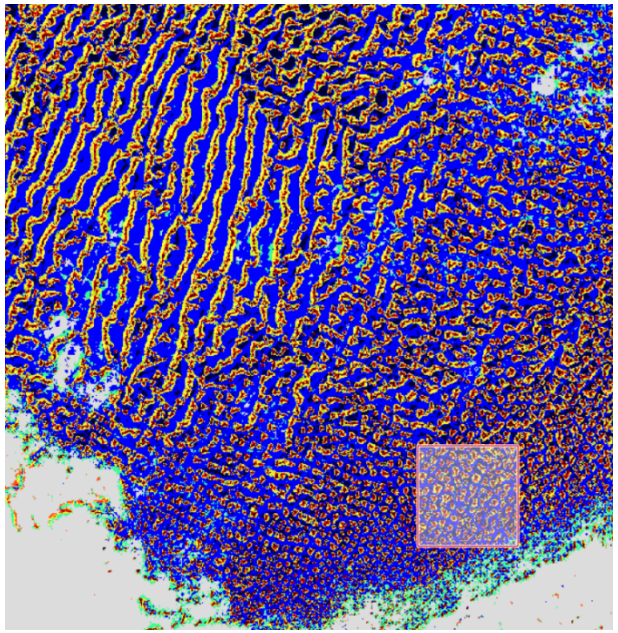
Saudi Arabia Empty Quarter

all examples computed online by TerraEx (sil.uc.edu) (beta)

Search 5 of 5

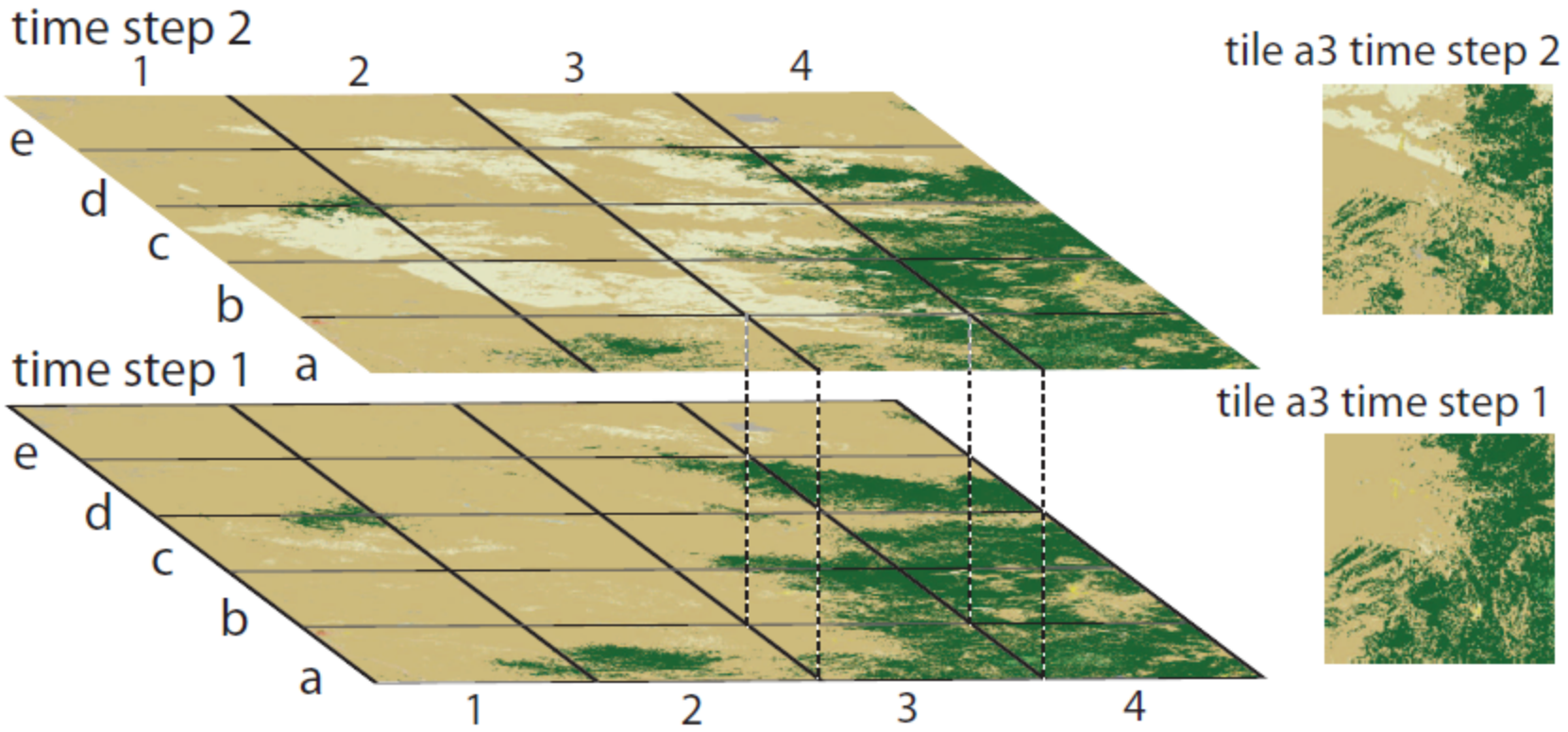






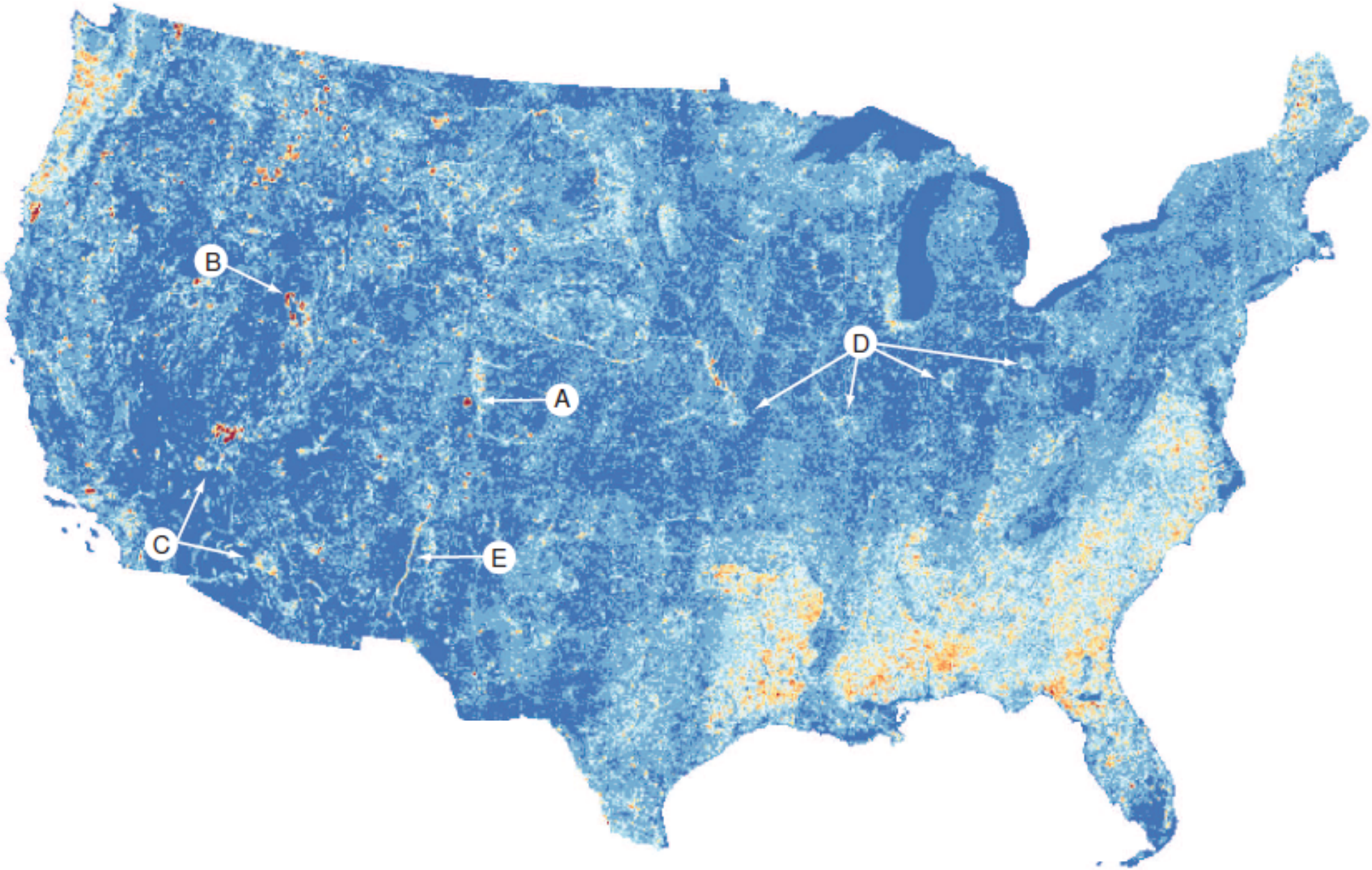
all examples computed online by TerraEx (sil.uc.edu) (beta)

Change 1 of 5



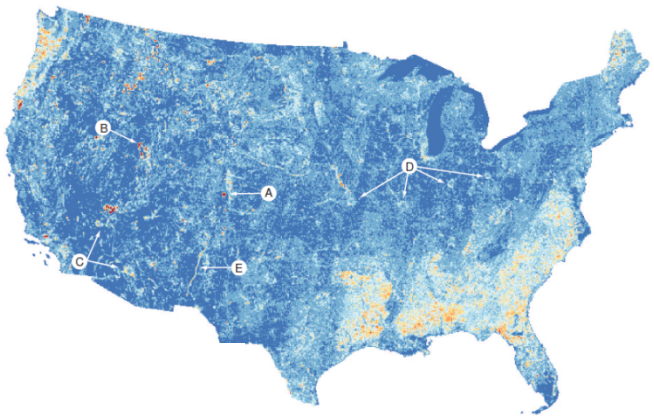
A degree of dissimilarity between the same locations (motifels) at two different times is a measure of change in pattern.

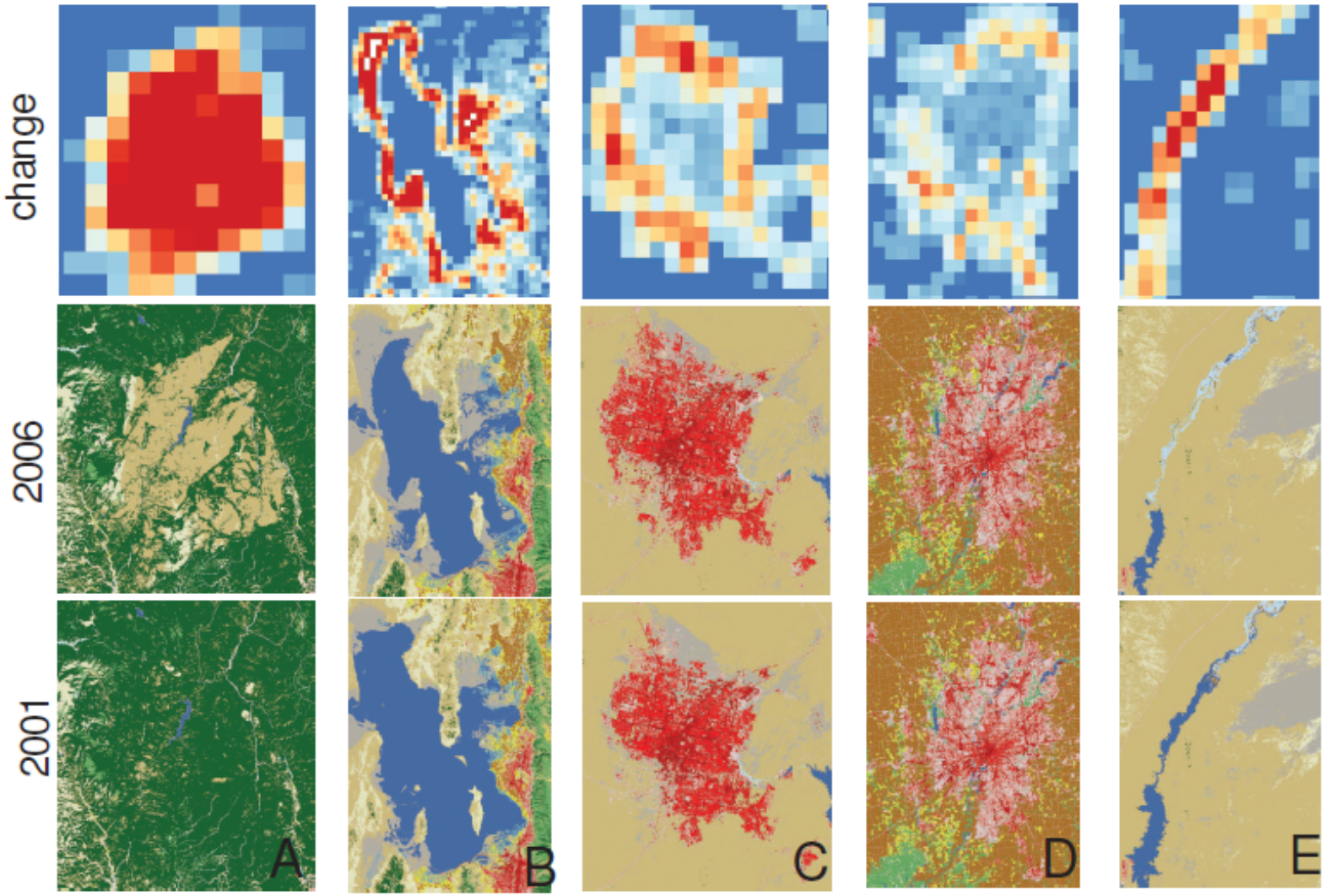
Change 2 of 5



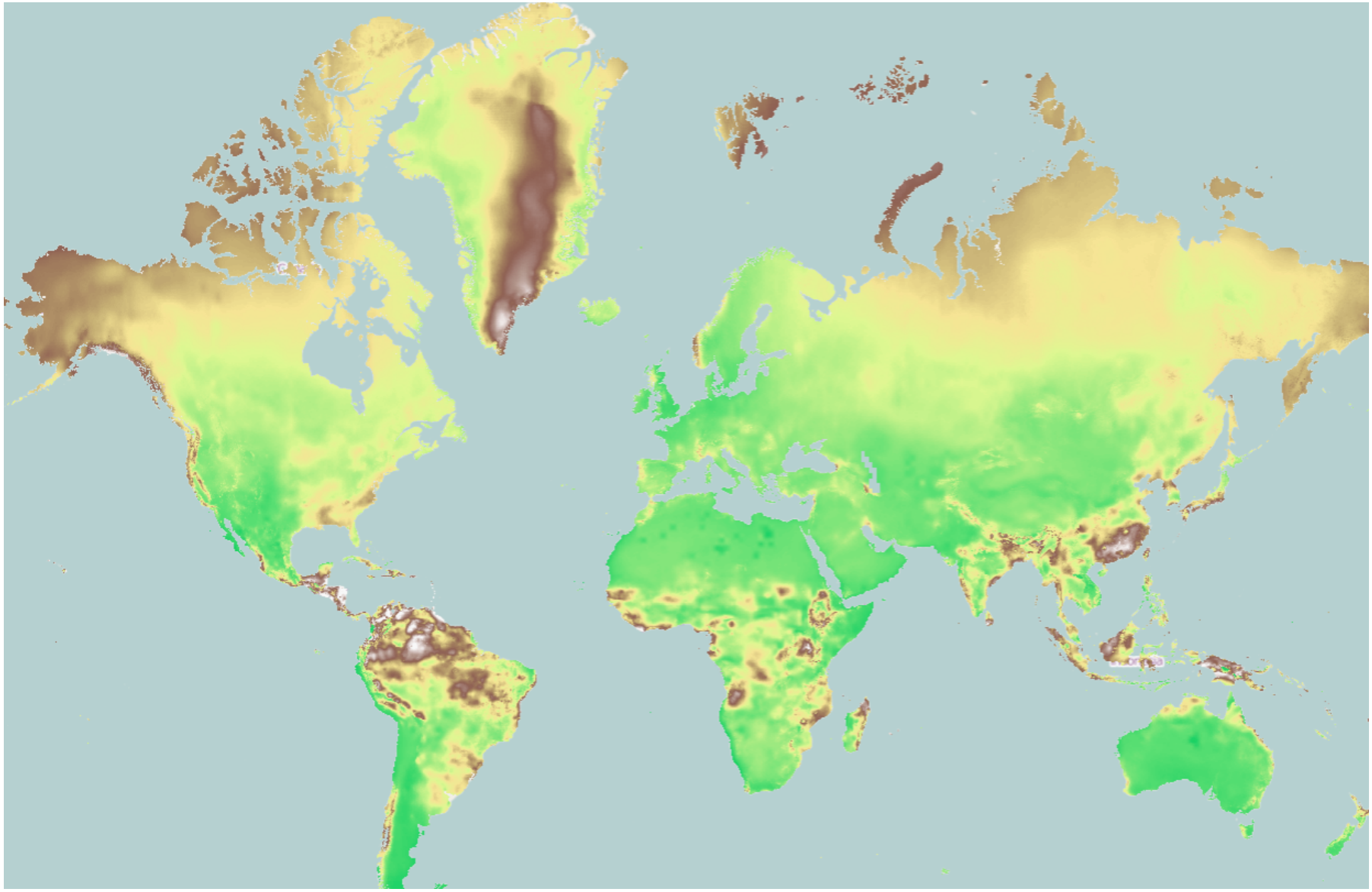
Change in land cover patterns over conterminous U.S. 2001-2006

Change 3 of 5





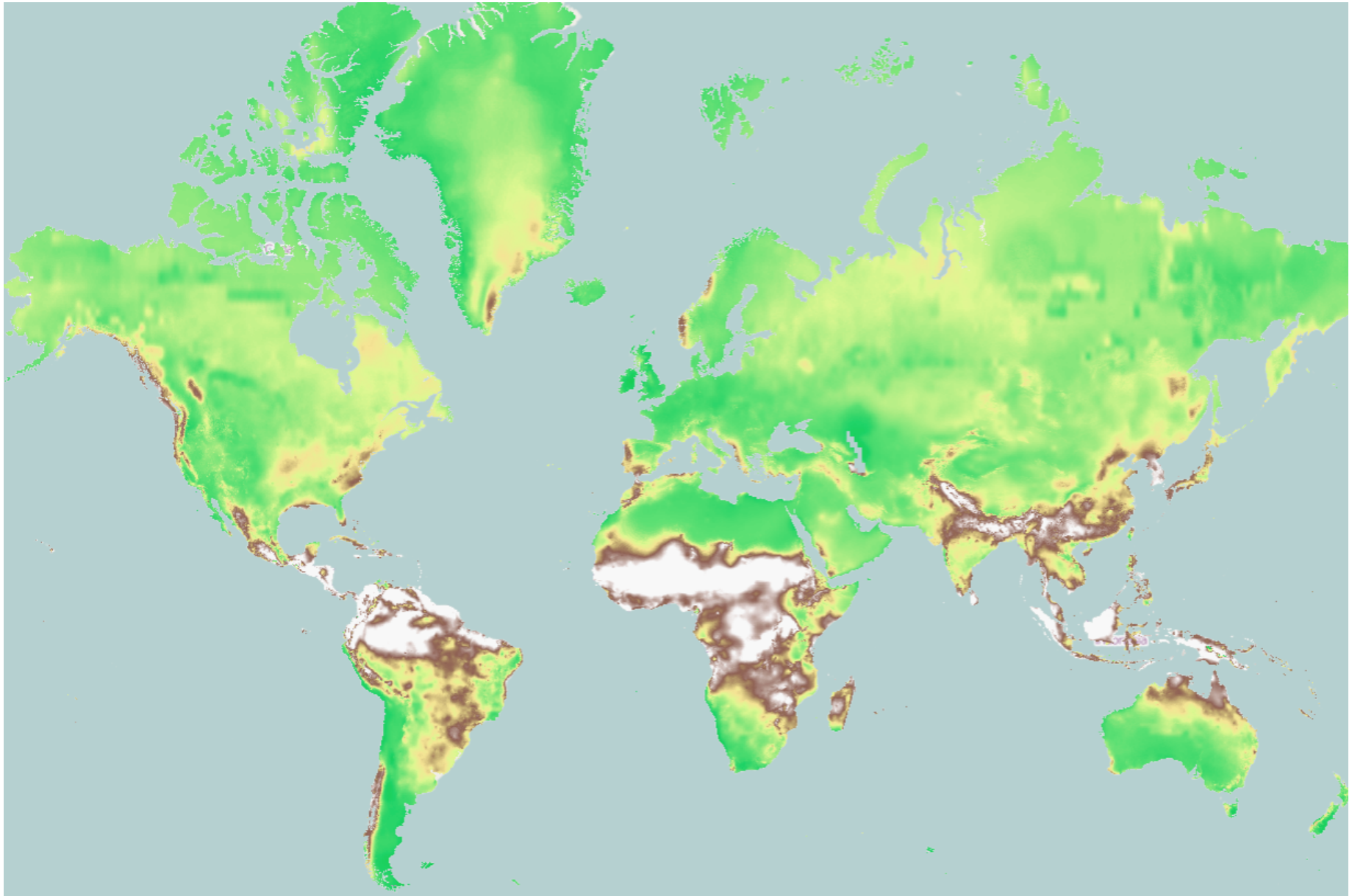
Change 4 of 5



Climate change 2000 -2070

all examples computed online by ClimateEx (sil.uc.edu)

Search 5 of 5



Climate change -6000 -2000

all examples computed online by ClimateEx (sil.uc.edu)

Stand-alone version of GeoPAT

Present version of GeoPAT requires GRASS and (in most cases) Linux because GRASS needs to be compiled together with GeoPAT.

GeoPAT 2 is GRASS-free. All computations can be performed in terminal mode and exported to GIS software of choice (ArcGIS, QGIS, etc)

GeoPAT 2 will be provided as source code and as executable for Linux and Windows.

We have already converted several GeoPAT modules GRASS environment to stand-alone and demonstrated to work on Windows.

GeoPAT 2 will be open and free software

Summary

GeoPAT fills a niche devoted to analysis of large rasters (starting with GeoPAT 2 they don't need to be categorical).

Public data:

1. Land cover (NLCD, GLC30, GlobCover, CORINE, EOSD)
2. Topography (NED, SRTM)
3. Croplands (USDA CropScape)
4. Urban Structure Types (UST) (National Map)

Applications:

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

Roadmap for the next 12 months

GeoPAT:

1. Finish development and release GeoPAT 2 for GRASS
2. Finish development and release the stand-alone GeoPAT 2
3. Finish development and release GeoPAT 2 optimized for Xeon Phi

Online applications:

1. Finish TerraEx

Testing GeoPAT on public data

1. Calculate and validate regionalization of US into physiographic units
2. Calculate and validate regionalization of US into landscape types
3. Validate utility of GeoPAT for delineation of UST in American cities