**Introduction**

**GeoPAT 2** is an open-source software written in C and dedicated to pattern-based spatial and temporal analysis.  
Four main types of analysis available in **GeoPAT 2** are (i) search, (ii) change detection, (iii) segmentation, and (iv) clustering.  
However, additional applications are also possible, including extracting information about spatial patterns.

Global landscape diversity (based on Shannon entropy of land cover categories in a 9 by 9 km local window). It shows where only one land cover category dominates, and where there is a large number of lc categories. [#sil](https://twitter.com/hashtag/sil?src=hash&ref_src=twsrc%5Etfw) [#gis](https://twitter.com/hashtag/gis?src=hash&ref_src=twsrc%5Etfw) [#landscapeecology](https://twitter.com/hashtag/landscapeecology?src=hash&ref_src=twsrc%5Etfw) [pic.twitter.com/E7pVkIwwY7](https://t.co/E7pVkIwwY7)

— Jakub Nowosad (@jakub\_nowosad) [January 15, 2019](https://twitter.com/jakub_nowosad/status/1085174957678780417?ref_src=twsrc%5Etfw)

In the above case, I was interested in extracting values of Shannon entropy of land cover categories for local landscapes on a global scale.  
It had prompted a comment from [Steffen Ehrmann](https://twitter.com/DerEhrmann):

The LC data I have, but it would be awesome to see the code you use in GeoPAT2, would be a good excuse to learn it!

— Steffen Ehrmann (@DerEhrmann) [January 16, 2019](https://twitter.com/DerEhrmann/status/1085651345552490496?ref_src=twsrc%5Etfw)

Therefore, this blog post describes how to calculate Shannon entropy of land cover categories using **GeoPAT 2** and R.

**Setup**

Installation instructions for **GeoPAT 2** can be found at <https://github.com/Nowosad/geopat2#Installation> and for R at <https://cloud.r-project.org/>.  
Additionally, a few R packages should be installed.

pkgs = c(

"rgeopat2", # helper functions for GeoPAT 2 in R

"sf", # vector data classes

"raster", # raster data classes

"dplyr", # data manipulation

"tmap", # spatial visualisation

"landscapemetrics" # contains example dataset

)

to\_install = !pkgs %in% installed.packages()

if(any(to\_install)) {

install.packages(pkgs[to\_install])

}

Now you have the necessary packages, the next step is to attach the ones we will use.

library(rgeopat2)

library(sf)

library(raster)

library(dplyr)

library(tmap)

library(landscapemetrics)

**Input data**

We will use the augusta\_nlcd dataset from the **landscapemetics** package as an example.  
It is a raster representing land cover categories for an area of about 270 km^2^ west from Augusta, Georga.

data("augusta\_nlcd")

augusta\_nlcd = deratify(augusta\_nlcd, "NLCD.2011.Land.Cover.Class")

dir.create("data")

writeRaster(augusta\_nlcd, "data/augusta\_nlcd.tif", overwrite = TRUE)

nlcd\_colors = c("#000000", "#00F900", "#476BA0", "#D1DDF9", "#DDC9C9", "#D89382",

"#ED0000", "#AA0000", "#B2ADA3", "#68AA63", "#1C6330",

"#B5C98E", "#dcca8f",

"#fde9aa", "#DBD83C", "#AA7028", "#BAD8EA", "#64b3d5")

p1 <- tm\_shape(augusta\_nlcd) +

tm\_raster("NLCD.2011.Land.Cover.Class", palette = nlcd\_colors) +

tm\_layout(legend.outside = TRUE)

p1

**Calculate Shannon entropy**

The gpat\_gridhis module of **GeoPAT 2** reads the input raster data from a GeoTIFF file and creates an output based on selected size and signature.  
Firstly, we need to choose the size of the local landscapes.  
The input data has a resolution of 30 meters, and we decided on local landscapes of 600 by 600 meters (0.36 km^2^) by setting both size (-z) and shift (-f) to 20 (map resolution of 30 \* 20 = 600 meters).  
Secondly, we need to pick a proper signature – in this case, it is ent (Shannon **ent**ropy).  
For this signature, we also must disable any value normalization (-n 'none').

system("gpat\_gridhis -i data/augusta\_nlcd.tif -o data/augusta\_ent.grd -z 20 -f 20 -s 'ent' -n 'none'")

The output is a binary file and we can easily convert it to text with the gpat\_grid2txt module.

system("gpat\_grid2txt -i data/augusta\_ent.grd -o data/augusta\_ent.txt")

**Create a vector grid**

The output text file can be next processed in any software.  
However, we will use R here, as there is an R package for working with **GeoPAT 2** outputs called **rgeopat2**.  
One of its functions, gpat\_create\_grid(), creates a spatial vector object recreating the grid used in the previous **GeoPAT 2** calculations.

augusta\_grid = gpat\_create\_grid("data/augusta\_ent.grd.hdr")

This new object contains a number of regular square polygons (aka local landscapes), each having a size of 600 by 600 meters.

p2 = tm\_shape(augusta\_nlcd) +

tm\_raster("NLCD.2011.Land.Cover.Class", palette = nlcd\_colors) +

tm\_shape(augusta\_grid) +

tm\_borders() +

tm\_layout(legend.outside = TRUE)

p2

**Read data to R and connect it to the grid**

The **rgeopat2** also has a helper function, gpat\_read\_txt(), for reading text outputs from **GeoPAT 2**.

augusta\_ent = gpat\_read\_txt("data/augusta\_ent.txt", signature = "ent")

The new object augusta\_ent is a data frame with five columns describing each local landscape – (i) value of Shannon entropy, (ii) number of land cover categories, (iii) area with values in cells^2^, (iv) column number, (v) row number.

head(augusta\_ent)

## Shannon\_entropy number\_of\_categories object\_size col row

## 1 1.174216 5 400 1 1

## 2 2.005147 6 400 2 1

## 3 1.485236 5 400 3 1

## 4 2.200218 10 400 4 1

## 5 2.226918 8 400 5 1

## 6 1.918720 6 400 6 1

Now, we can combine the spatial object with the data frame.

augusta\_grid = bind\_cols(augusta\_grid, augusta\_ent)

The output, augusta\_grid is a set of polygons, where each is represented by the value of Shannon entropy.  
It allows distinguishing the local landscapes with only one or two land cover categories (low values of Shannon entropy) and the ones with many land cover categories (high values of Shannon entropy).

p3 = tm\_shape(augusta\_nlcd) +

tm\_raster(legend.show = FALSE, palette = nlcd\_colors) +

tm\_shape(augusta\_grid) +

tm\_polygons("Shannon\_entropy") +

tm\_layout(legend.outside = TRUE)

p3

**Summary**

This post has shown how to use **GeoPAT 2** to extract certain metric (Shannon entropy), and how to connect the result with its spatial representation for a relatively small area in Georgia, USA.

tmap\_arrange(p1, p2, p3, ncol = 1)

However, **GeoPAT 2** was written to handle large spatial rasters, including ones on continental and global scales.  
Therefore, I encourage you to try it on your own study area – regardless of its size.  
You can also read more about additional applications of **GeoPAT 2** and see more examples at <https://github.com/Nowosad/geopat2>.