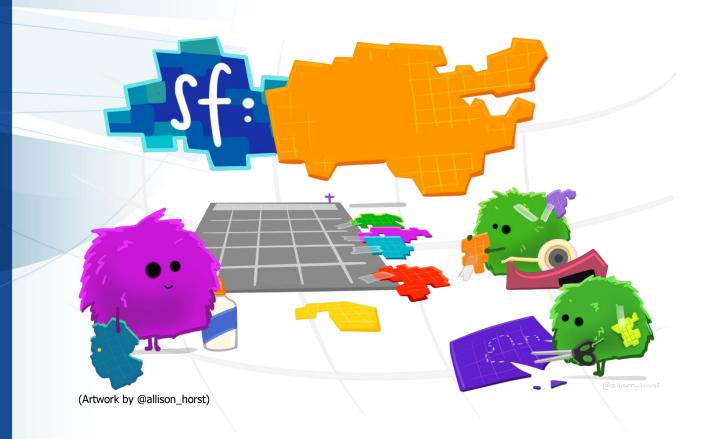
Spatial data and analysis in R ESM – CDAT course series



Course plan

Introduction talk (14:00 – 14:45pm, 22th October 2020)

Self study tutorials

(online self paced, link at end of slides)

All slides and materials online

https://tinyurl.com/CDATSpatR



What is GIS?

A geographic information system (GIS) is a system designed to capture, store, manipulate, analyze, manage, and present types of geographical data.









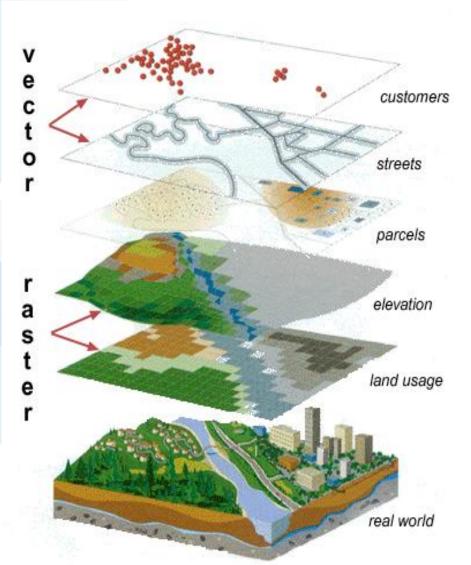








Types of spatial data



Format:

*.shp, *.gpkg, *.gpx, *.kml, ...

Format:

*.tif, *.vrt, *.hdr, *.asc, ...



Vector

Advantage: Accuracy, more visually pleasing

Disadvantage: Space-inefficient. Every vertex needs to be stored. Algorithms computational intensive.

Raster

Advantage: Geogr. Position associated with data, easier for analysis

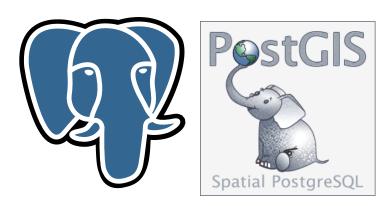
Disadvantage: Resolution dependent on cellsize. Lack of attributes, MAUP



Spatial databases







PostgreSQL & PostGIS

Stop using shapefiles!



```
Multifile system (.shp,.shx,.dbf,.prj, ... )
Limited to 2GB (4GB)
Attribute names limitation
Etc...

http://switchfromshapefile.org/
```

Solution → **OGC Geopackages**



Geographic projections

Spatial data requires a projection

Choosing an appropriate geographic projection is important

- Meter or degree based?
- Tradeoff between shape, area or distance distortion
- Aesthetics vs accuracy







MERCATOR PROJECTION; YOU JUST WISH IT WEREN' SQUARE. THE EARTH'S NOT A SQUARE, IT'S A ORCLE YOU LIKE CIRCLES, TODAY IS GONNA BE A GOOD DAY





THAT YOU WEAR EVERYWHERE. YOU LIKE COFFEE AND ENJOY THE BEATLES, YOU THINK THE ROBINSON IS GOGGLES. WHICH YOU USE TO VIEW ROTATING MODELS THE BEST-LOOKING PROTECTION, HANDS DOWN OF BETTER 3D GOGGLES YOU TYPE IN DURRAK





FASY FNOUGHTO YOU YOU LIKE FASY SOLUTIONS YO 1998. BUT YOU'VE BEEN A WIT FAN SINCE LONG BEFORE THINK WE WOULDN'T HAVE SO MANY PRORIEKS IF WE THAT FLECT MORNAY PROPER TO CONCRESS INSTEAD AMED OUT, AND ARE THINKING OF SLITTCHING TO THE OF PALITICIANS. YOU THINK AIRLINES SHOULD TIKE RIM KAURAYSKIY YN I ONCE LEFT A PARTY IN DISGIKT WHEN FOOD FROM THE RESTAURANTS NAME THE GATES AND QUEST SHOULED UP WEARING SHOES WITH THES. YOUR SERVE THAT ON BOARD. YOU CHANGE YOUR CAR'S OIL



YOUTHINK THIS ONE IS FINE. YOU LIKE H YOU'VE HEARD BAD THINGS ABOUT GALL-PETERS. MAP TO LATTTUDE AND LONGITUDE. THE OTHER YOU'RE CONFLICT-AVERSE AND BUY ORGANIC. YOU PROTECTIONS OVERCOMPLICATE THINGS. YOU WANT HE LISE A RECENTLY-INVENTED SET OF GENDER-NEUTRAL TO STOP ASKING ABOUT MAPS SO YOU CAN ENDOY DINNE PRONOUNS AND THINK THAT WHAT THE WORLD



THE 1909 CAHILL MAP IT'S BASED - ... YOU HAVE A FRAMED REPRODUCTION AT HOME?! WHOA. ... LISTEN, FORGET







PERFORESTIONS ARE YOUT DOING ANYTHING TONKY-IT



Source: https://xkcd.com/977

Why use R for spatial analyses?

- ✓ Open source
- ✓ Efficiency ('Don't repeat yourself')
- ✓ Cross system availability (Win,*Nix, MacO\$
- Extendable and rich functionality
- Clean coding (also for 'tidy' data concept)
- ✓ Parallel computing support
- ✓ Integration of C, C++ code for speed



Many spatial packages



Environmental Modelling & Software

Volume 133, November 2020, 104799



Position Paper

Harmonise and integrate heterogeneous areal data with the R package arealDB

Steffen Ehrmann ^{a, d} $\stackrel{\triangle}{\sim}$ $\stackrel{\boxtimes}{\sim}$, Ralf Seppelt ^{b, c}, Carsten Meyer ^{a, c, d} $\stackrel{\triangle}{\sim}$ $\stackrel{\boxtimes}{\sim}$

Source: https://doi.org/10.1016/j.envsoft.2020.104799





geemap: A Python package for interactive mapping with Google Earth Engine

Qiusheng Wu¹

1 Department of Geography, University of Tennessee, Knoxville, TN 37996, United States

DOI: 10.21105/joss.02305



One comprehensive list of spatial packages

https://cran.rproject.org/web/views/Spatial.ht ml



Why/When not to use R for spatial analyses?

- R can be slow
- Many (sp) packages not memory efficient
- Often little support
- ❖ No GUI
- Greater proficiency in other languages



(Personal opinion)

The diversity of open source GIS solutions is both its greatest strength and weakness



The backbone of most open-source GIS

OGG®

Making location count.

Standards like WMS, KML, GML, SFC



C/C++ libraries



R as a GIS

Main packages: 'Sp', 'raster' and 'rgdal' still go-to functions to use

Problem: Each have their own object-based model, often inefficient code



TIDY DATA is a standard way of mapping the meaning of a dataset to its structure.

-HADLEY WICKHAM

In tidy data:

- each variable forms a column
- each observation forms a row
- each cell is a single measurement

(each co	lumn a v	ariable	
	id	id name		
	1	floof	gray	R each row
	2	max	black	()
	3	cat	orange	Mobservatio
	4	donut	gray	2//
	5	merlin	black	4/
	6	panda	calico	1

Wickham, H. (2014). Tidy Data. Journal of Statistical Software 59 (10). DOI: 10.18637/jss.v059.i10

Artwork by @allison_horst



Tidy data and simple features

Simple Features are a set of OGC standards how spatial (vector) data is to be stored



Source: https://r-spatial.github.io/sf/articles/sf1.html

Simple feature collection with 100 features and 6 fields

geometry type: MULTIPOLYGON

dimension: XY

bbox: xmin: -84.32385 ymin: 33.88199 xmax: -75.45698 ymax: 36.58965

epsg (SRID): 4267

proj4string: +proj=longlat +datum=NAD27 +no defs

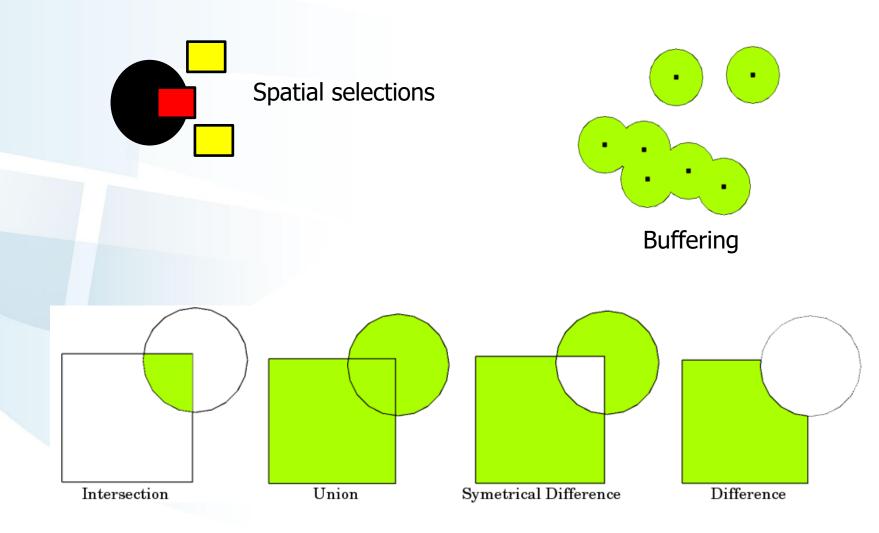
precision: double (default; no precision model)

First 3 features:

$\pi\pi$	٠.	TISC 3	i Ca cui	C3.					
##		BIR74	SID74	NWBIR74	BIR79	SID79	NWBIF	R79	geom
##	1	1091	1	10	1364	0		19	MULTIPOLYGON(((-81.47275543
##	2	487	0	10	542	3	1	12	MULTIPOLYGON(((-81.23989105
##	3	3188	5	208	3616	6	/ 2	260	MULTIPOLYGON(((-80.45634460
						Simple fea	ture	Simp	Simple feature geometry (sfg)



Spatial analyses – vector data





Source: QGIS documentation

Example code in R using 'sf'

```
World %>%
    st_transform(crs = 54009) %>%
    st_buffer(1000) %>%
    st_intersects(hotspots) %>%
    group_by(hotspot_name) %>%
        summarise(
        geometry = st_union(geometry),
        area = st_area(geometry)
```



Spatial analyses – raster data

Examples

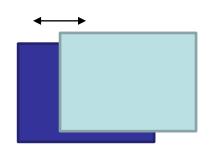
- Aggregations, disaggregations
- Region growth, reclassifications
- Band arithmetic (NDVI etc)
- Terrain analyses (Slope, Aspect, Curvature)

₹



Common spatial tasks I do in R

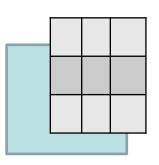
Aligning raster input data



Extracting zonal statistics

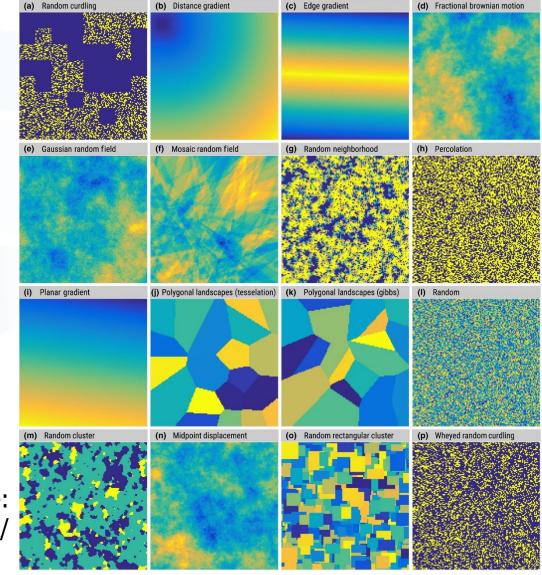


Joining spatial and non-spatial data





Examples: Neutral landscapes



Source: https://ropensci.github.io/NLMR/



Example: Landscape metrics

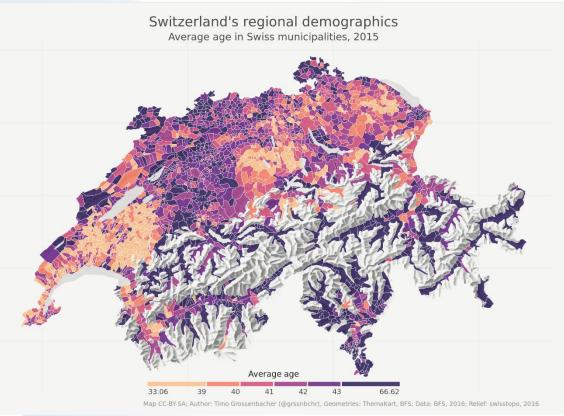
```
# list all available metrics
list_lsm()
#> # A tibble: 132 x 5
     metric name
                                                              level function name
#>
                                           type
#> <chr> <chr>
                                                            <chr> <chr>
                                           <chr>>
#> 1 area patch area
                                           area and edge met... patch lsm p area
                                           core area metric patch lsm p cai
#> 2 cai core area index
   3 circle related circumscribing circle
                                           shape metric
                                                              patch 1sm p circle
                                           shape metric
                                                              patch 1sm p contig
   4 contig contiguity index
   5 core core area
                                           core area metric patch 1sm p core
   6 enn euclidean nearest neighbor dis... aggregation metric patch lsm p enn
  7 frac fractal dimension index
                                           shape metric
                                                              patch 1sm p frac
   8 gyrate radius of gyration
                                           area and edge met... patch lsm p gyrate
  9 ncore number of core areas
                                           core area metric patch lsm_p_ncore
#> 10 para perimeter-area ratio
                                           shape metric
                                                              patch 1sm p para
#> # ... with 122 more rows
```

Source:

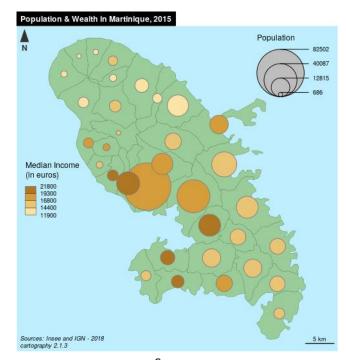
https://r-spatialecology.github.io/landscapemetrics/



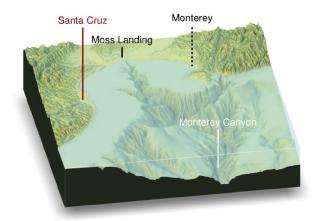
Making maps in R



Source: https://timogrossenbacher.ch/2016/12/beautiful-thematic-maps-with-ggplot2-only/



Source: http://riatelab.github.io/cartography



Source: https://github.com/tylermorganwall/rayshader

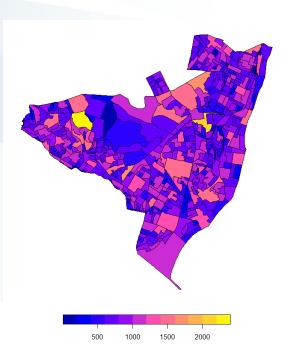


Spatial statistics

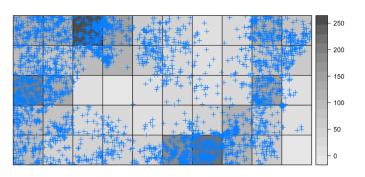
(Not covered in the online course!)

Geographically weighted regressions, Kriging, Network analyses, Spatial clustering, Machine learning

. . .



A lot of data has spatial structure!





What to do if things don't work

Too slow

→ Check memory requirements, consider tiling

No Function

→ Check external tools. Is there a wrapper?

Visualization

→ Use QGIS for quick queries





Free online books and materials

Welcome

This is the online home of *Geocomputation with R*, a book on geographic data analysis, visualization and modeling.

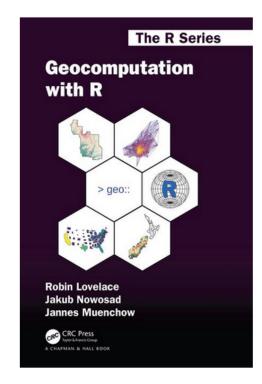
Note: This book has now been published by CRC Press in the R Series. You can buy the book from CRC Press, Wordery, or Amazon.

Inspired by **bookdown** and the Free and Open Source Software for Geospatial (FOSS4G) movement, this book is open source. This ensures its contents are reproducible and publicly accessible for people worldwide.

The online version of the book is hosted at geocompr.robinlovelace.net and kept up-to-date by GitHub Actions, which provides information on its 'build status' as follows:

Render-Book-from-master passing

This version of the book was built on GH Actions on 2020-10-06.





Online course materials

Spatial data and analysis in R

Starting page

Lecture

Installing packages

Contents ▼

Resources

Code **▼**

Spatial data and analysis in R

Martin Jung

Ecosystems Services and Management

International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

Apart from the lecture slides this online self-learning course aims to provide you with basic knowledge about spatial datasets in R, how to load and analyse them. In many instances R might not be the fastest tool one can use for these kinds of analyses, but it certainly is the fastest in terms of time spent in code development. Here we will use R as a wrapper to load in external tools. This course assumes that users already have basic knowledge of R.

I generally tried to avoid replicating things that are already openly available online through other resources. Thus, if you are interested in more or other training materials regarding spatial data and analyses in R, check out the resources link at the top with more examples and free self-learning tutorials.



https://tinyurl.com/CDATSpatR



Thank you for your attention.

Good success with your spatial analyses!

