Using environmental data from remote sensing in demographic analysis: An introduction

EDSD, 12-15 Dec 2022

Ankit Sikarwar & Valérie Golaz

sikarwar.ankit-kumar@ined.fr

valerie.golaz@ined.fr

ined
INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES

www.ined.fr

Day 1 / Session 1

Goals, outline, expectations

Introduction:

- What is remotely sensed data?
- Ways and purposes in using remotely sensed data in demographic analysis?
- Issues in processing RS data in GIS for demographic studies

Goals

- Share experience on environmental data use in demography
- Provide an initiation to QGIS, the most commonly used GIS freeware with key concepts and tools/processes for mapping indicators
- Discuss the main sources of remotely sensed environmental data in use, their strength and limitations
- Experience the extraction of environmental data, at any scale of interest, from these sources.
- Match demographic and environmental data in analysis.

Outline

Day 1 Introduction

- The why and how of using remotely sensed data in demographic analysis
- Introduction to QGIS

Day 2 First steps in QGIS

- From indicators to maps
- Locating and importing environmental data

Day 3 Geoprocessing of remotely sensed and other gridded data

- Reading and analysing data
- Producing environmental indicators matching relevant spatial units for the analysis

Day 4 Producing indicators at individual/cluster level

- Conceptualisation
- Conclusion

Our expectations

* Active **participation**

* **Projects** in small groups in interaction / cooperation, following the steps of our 4 days of class. One country or geographical area of interest + one indicator at least.

-A **2p max doc** that will include:

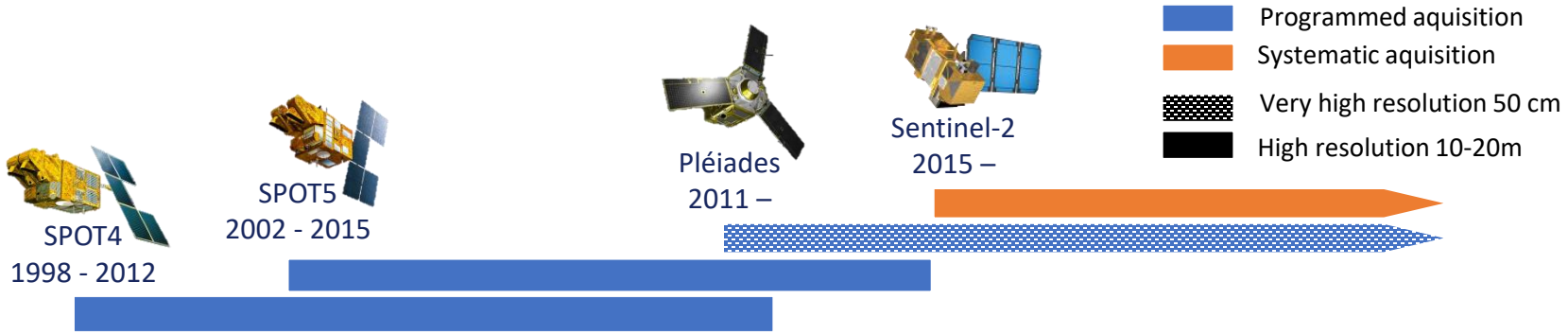
*A brief introduction: what you are going to do and for what purpose, with what idea. You need to present here the data your are going to use, and to explain at what level you need it.

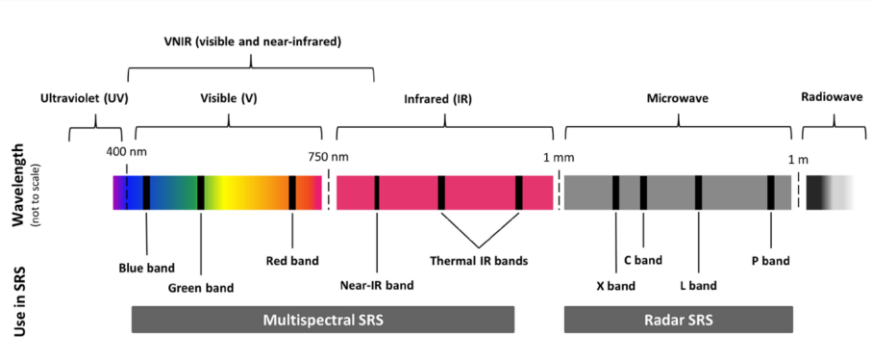
*The various steps to prepare your environmental indicators at the desired level. You may present a map of the raw indicator you are going to work on and end with a map of the processed indicator

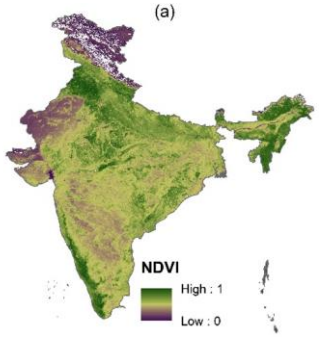
* A conclusion about the indicator

-An **excel file** including the values of the indicator(s) for your units of analysis

What is remotely sensed data?



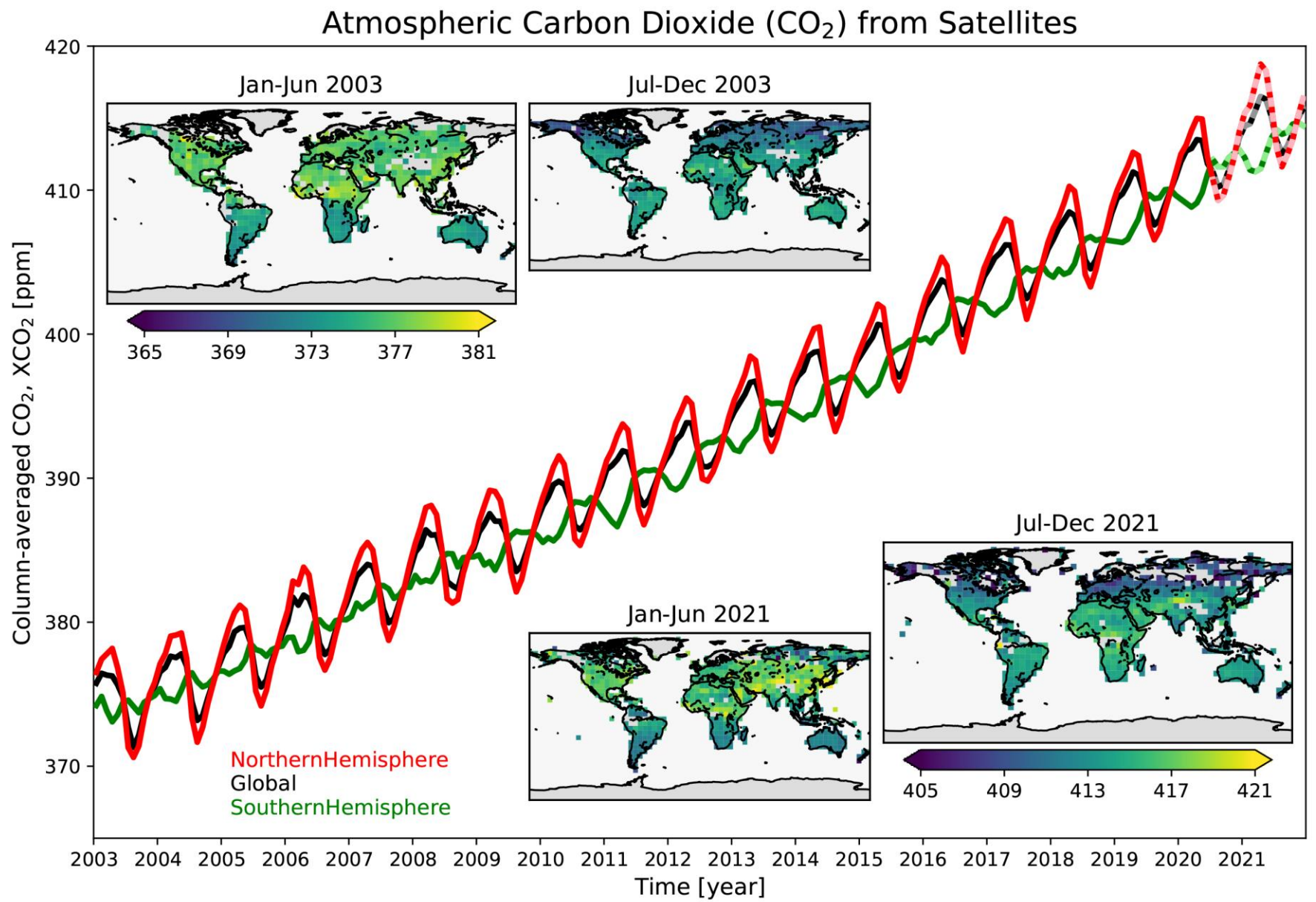




Optical and Radar Signal => Analysis => Gridded Indicator

Specific satellites / capture programmes

for specific data



Data:01.2003-06.2020:XCO2_OBS4MIPS(v4.3); after 06.2020:CAMS(NRT) - Satellites:SCIAMACHY/ENVISAT+GOSAT+OCO-2 - Credit:C3S/CCI/CAMS/Univ.Bremen/SRON v:09.02.2022

Copernicus Climate Change Service
Climate Indicators | 2021

PROGRAMME OF THE EUROPEAN UNION

Copernicus Europe's eyes on Earth

IMPLEMENTED BY ECMWF

The ‘analysis’ blackbox

- Interpreting signal through statistical procedures
- With an error factor
- Checks and adjustments: local captors / ground truth

Ways and purposes in using remotely sensed data in demographic analysis

Observing the visible print of human activities brings out social information

(lights, urbanisation, cleared land, road network…)

Physical indicators can indirectly refer us to social questions

Biomass => deforestation

Vegetation => harvest

Water => floods

Can observing and characterising the environment and environmental change help us understand socio-demographic processes ?

What research questions ?

For instance – Repeated droughts and change of agricultural practices? Decrease in migration ? Change in human fertility?

Or – Population densification and change in marriage patterns? In health ?

Or – Distance to a large water body and persistence of water-borne diseases? Of low levels of education ?

What demographic sources, at what geographical level ?

- The most accurate spatially : Individuals can be geolocalised

Census but is it accessible? Tax information, other localised administrative data, hdss data

- Clusters in a survey can be geolocalised

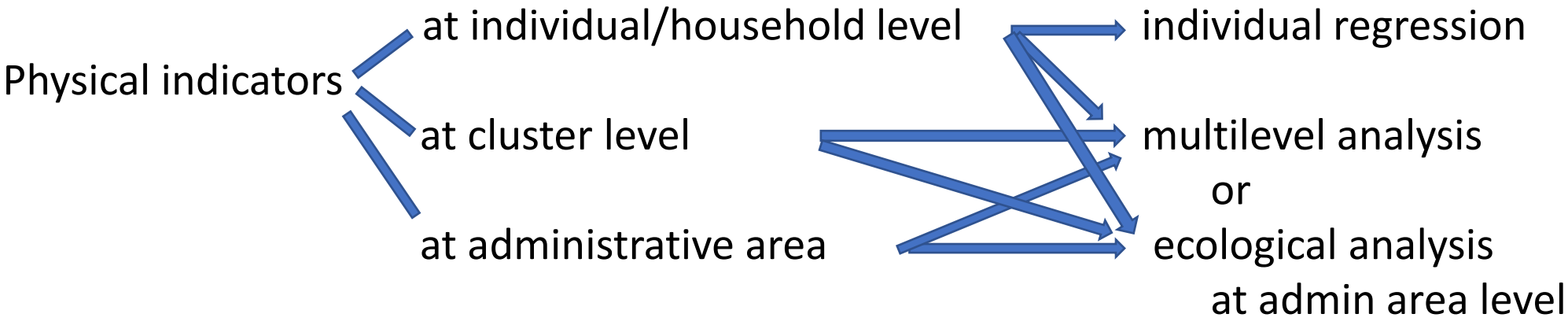
Some survey data like DHS

- Individuals belong to a given administrative area

Census data, administrative data, register data

What types of analysis ?

- Descriptive : looking for correlations between environmental and demographic data. Visually ? With classification techniques?
- Regression : because physical elements are never the only ones in the balance // in order to assess the respective roles of a set of possible factors, including those derived from the environment.




to at which level the sample is representative

First example of individual level analysis (local data)

World Development 109 (2018) 187–196

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Examining rural Sahelian out-migration in the context of climate change: An analysis of the linkages between rainfall and out-migration in two Malian villages from 1981 to 2009

Kathryn Grace [a,*], Véronique Hertrich [b], Djeneba Singare [c], Greg Husak [d]

[a] Department of Geography, Environment and Society, University of Minnesota, United States
[b] French Institute for Demography (INED), France
[c] Sahel women's empowerment and demographic dividend project (SWEDD project- Mali), Mali
[d] Climate Hazards Group, University of California, Santa Barbara, United States

Data

- 2 villages in Mali
- Demographic data: SLAM data base (1987-1989 to 2009-2010, updated every 5y) covering 1980-2010
- Climate data: CHIRPS database [Climate Hazards Group InfraRed Precipitation with Station] for the location of the 2 villages, over 30 y.

  =>Rainfall index combining 4 measures: total seasonal rainfall, rainfall distribution and onset, quality of the past year

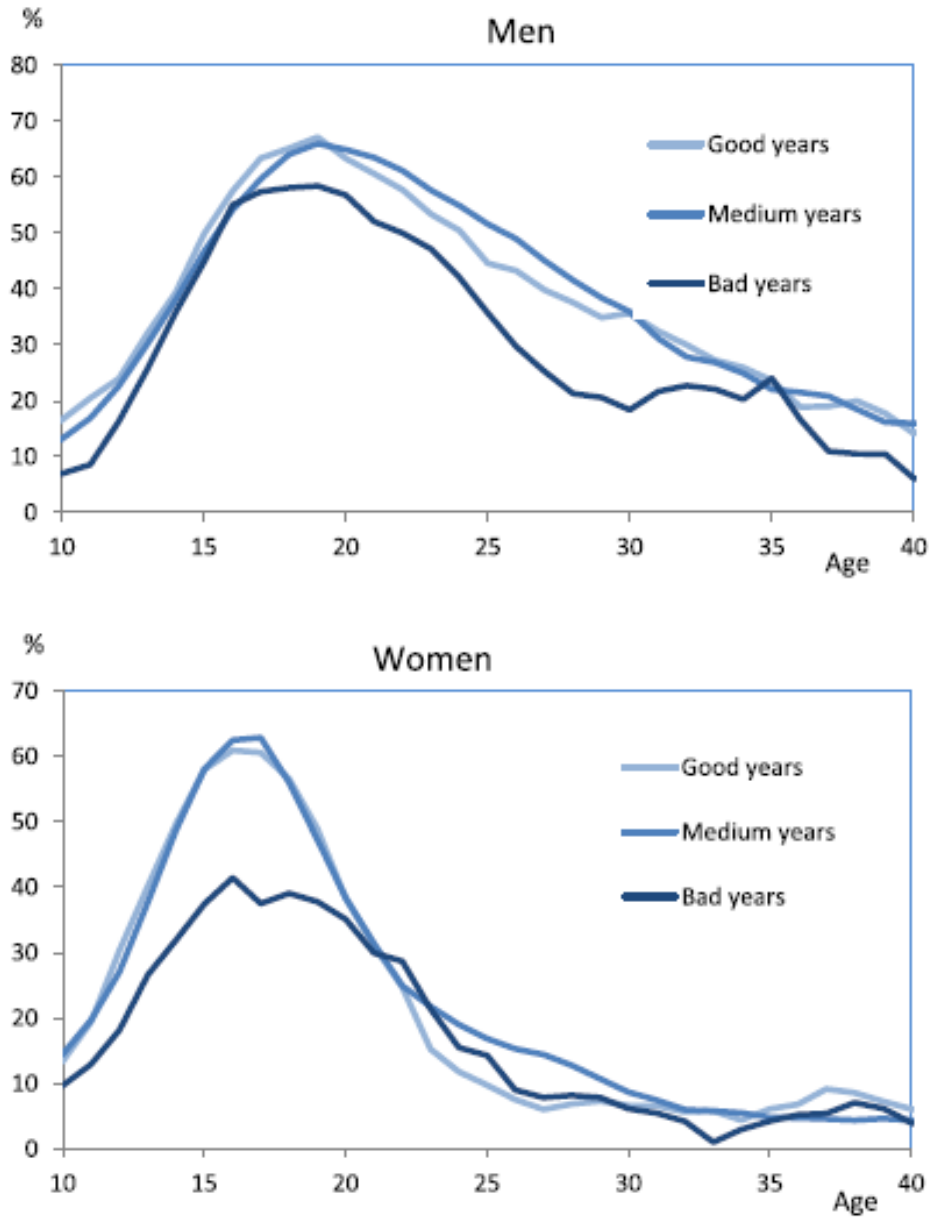


Fig. 4. Out-migration according the index of rainfall. By sex and age. (1980–2009).

Main result

⇒Negative effect of bad years on out-migration

(confirmed by regression analysis at individual level)

… second example (DHS)

Food Policy 46 (2014) 56–65

Contents lists available at ScienceDirect

Food Policy

journal homepage: www.elsevier.com/locate/foodpol

Examining the link between food prices and food insecurity: A multi-level analysis of maize price and birthweight in Kenya

Kathryn Grace [a,*], Molly Brown [b], Amy McNally [c]

[a] University of Utah, Department of Geography, 260 S. Central Campus Dr., Salt Lake City, UT 84112-9155, United States
[b] Goddard Space Center – NASA, United States
[c] University of California, Santa Barbara, United States

Data

- Kenya, DHS national survey clusters (2003 and 2008-09)
- 2 DHS Survey data on birthweight at the last birth, as well as explanatory variables
- Livelihood zone (FEWS.NET) and Price data FAO/FEWS NET
- NDVI data as a proxy of agricultural productivity (from the MODIS instrument on the NASA Terra satellite, 2000-) at 250m resolution

Main result

A logistic multilevel model (individual level/livelihood zone level) for low birth weights

=> High significance of Price x NDVI

Table 4
Interpreting the interaction between maize price and NDVI and the relationship to LBW.

	Price	
	Positive	Negative
NDVI		
Positive	Good own production, high price Model suggests high chance of LBW	Good own production, low price Model suggests low chance of LBW
Negative	Poor own production, high price Model suggests low chance of LBW	Poor own production, low price Model suggests high chance of LBW

2 syntheses of household data /remote sensing data links

Popul Environ (2014) 36:48–72
DOI 10.1007/s11111-013-0201-0

ORIGINAL PAPER

Using satellite remote sensing and household survey data to assess human health and nutrition response to environmental change

Molly E. Brown · Kathryn Grace · Gerald Shively · Kiersten B. Johnson · Mark Carroll

Population and Environment (2019) 41:209–234
https://doi.org/10.1007/s11111-019-00326-5

ORIGINAL PAPER

Check for updates

People and Pixels 20 years later: the current data landscape and research trends blending population and environmental data

Tracy A. Kugler[1] · Kathryn Grace[2] · David J. Wrathall[3] · Alex de Sherbinin[4] · David Van Riper[1] · Christoph Aubrecht[5] · Douglas Comer[6] · Susana B. Adamo[4] · Guido Cervone[7] · Ryan Engstrom[8] · Carolynne Hultquist[7] · Andrea E. Gaughan[9] · Catherine Linard[10] · Emilio Moran[11] · Forrest Stevens[9] · Andrew J. Tatem[12] · Beth Tellman[13] · Jamon Van Den Hoek[3]

Ecological analyses

SCIENCE ADVANCES | RESEARCH ARTICLE

CORONAVIRUS

Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis

X. Wu[1]*, R. C. Nethery[1]*, M. B. Sabath[1], D. Braun[1,2], F. Dominici[1†]

ELSEVIER

Environmental Research

Volume 217, 15 January 2023, 114906

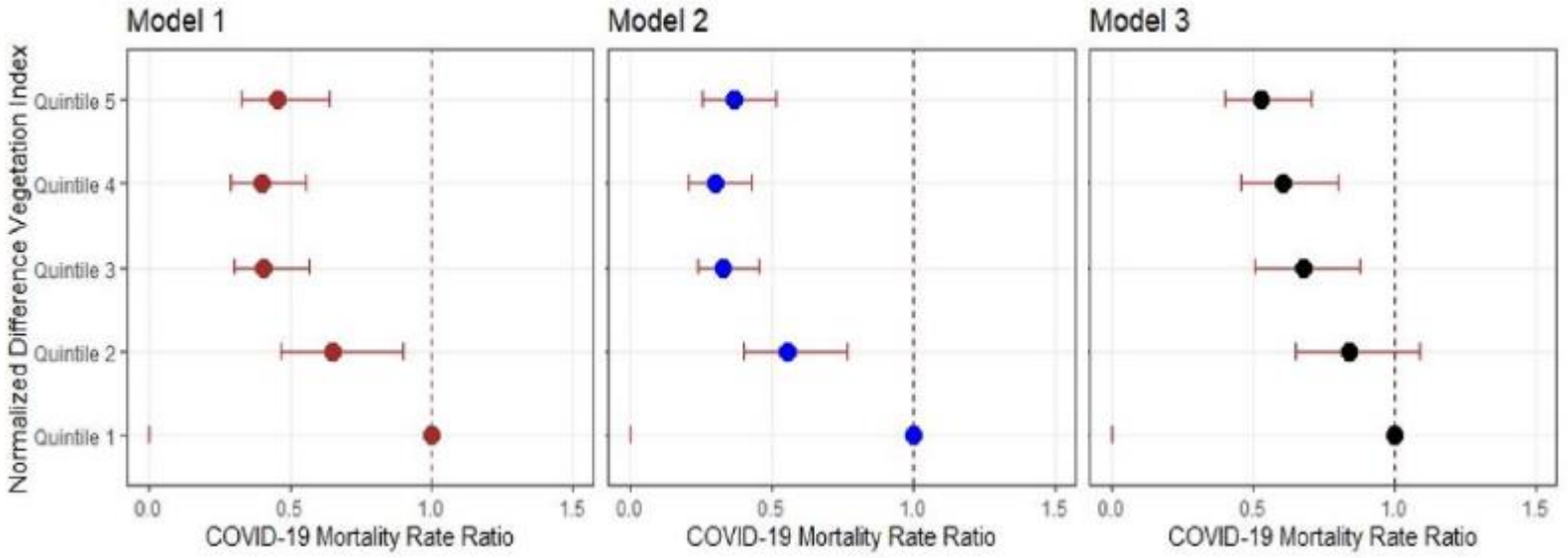
Association of greenness with COVID-19 deaths in India: An ecological study at district level

Ankit Sikarwar [a], Ritu Rani [a, b], Géraldine Duthé [a], Valérie Golaz [a, c]

Data

- India, district level
- Demographic data : Covid death counts 2020-2021=> MRR
- Environmental data : NDVI district averages for Jan-March 2019, 1km, in quintiles

Results



Another example using Worldpop data

Articles

Mapping physical access to health care for older adults in sub-Saharan Africa and implications for the COVID-19 response: a cross-sectional analysis

Pascal Geldsetzer, Marcel Reinmuth*, Paul O Ouma, Sven Lautenbach, Emelda A Okiro, Till Bärnighausen, Alexander Zipf*

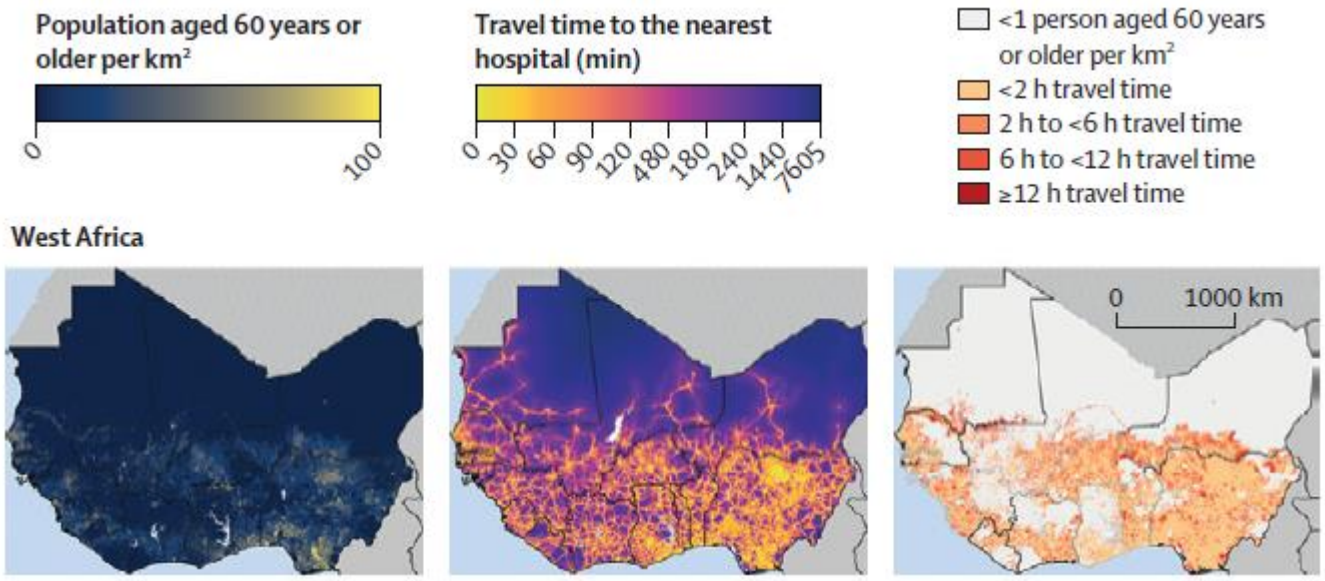
Summary

Background Severe acute respiratory syndrome coronavirus 2, the virus causing COVID-19, is rapidly spreading across sub-Saharan Africa. Hospital-based care for COVID-19 is often needed, particularly among older adults.

Lancet Healthy Longev 2020; 1: e32–42

Data

- Africa, 1kmx1km
- Demographic data : Worldpop
- Health facilities geolocations



Issues with remotely sensed data in demographic studies

- Time resolution
  - Seasonal variability
  - Time lag
- Spatial resolution
  - Large variations depending on the source
  - Towards a relevant perimeter and indexes, depending on the scale of analysis
- International databases vs direct processing ?
  - Technical difficulties in direct GIS processing (cloudy or no data areas?)
  - Keeping an eye on metadata in international databases

Using environmental data from remote sensing in demographic analysis: An introduction

EDSD, 12-15 Dec 2022

Ankit Sikarwar & Valérie Golaz

ined
INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES

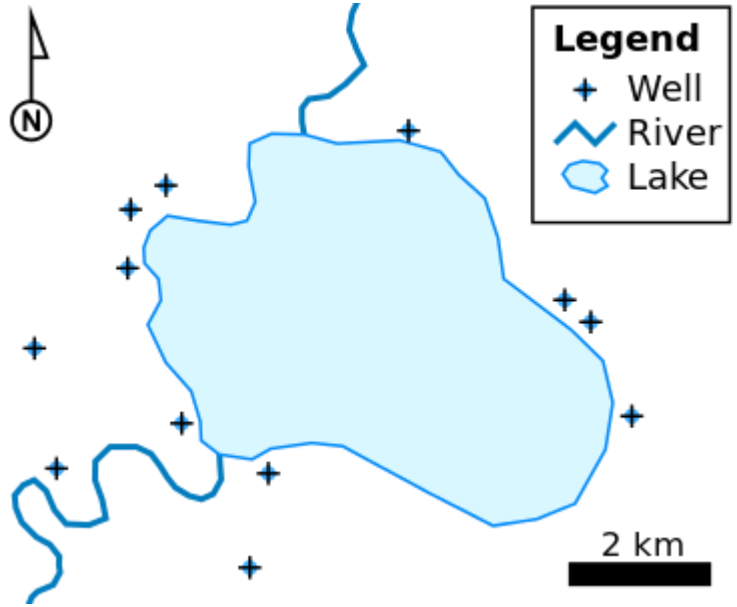
www.ined.fr

Day 1 / Session 2

- Important vocabulary and concepts: shapefiles, vector, raster, projections, ...
- Lay-out, principles, tools.
- Finding and importing data and shapefiles

Shapefiles

- geospatial vector data format for geographic information system (GIS) software
- It is developed and regulated by Esri
- spatially describe vector features: points, lines, and polygons, representing, for example, water wells, rivers, and lakes
- Each item usually has attributes that describe it, such as name or temperature.
- (.shp format)

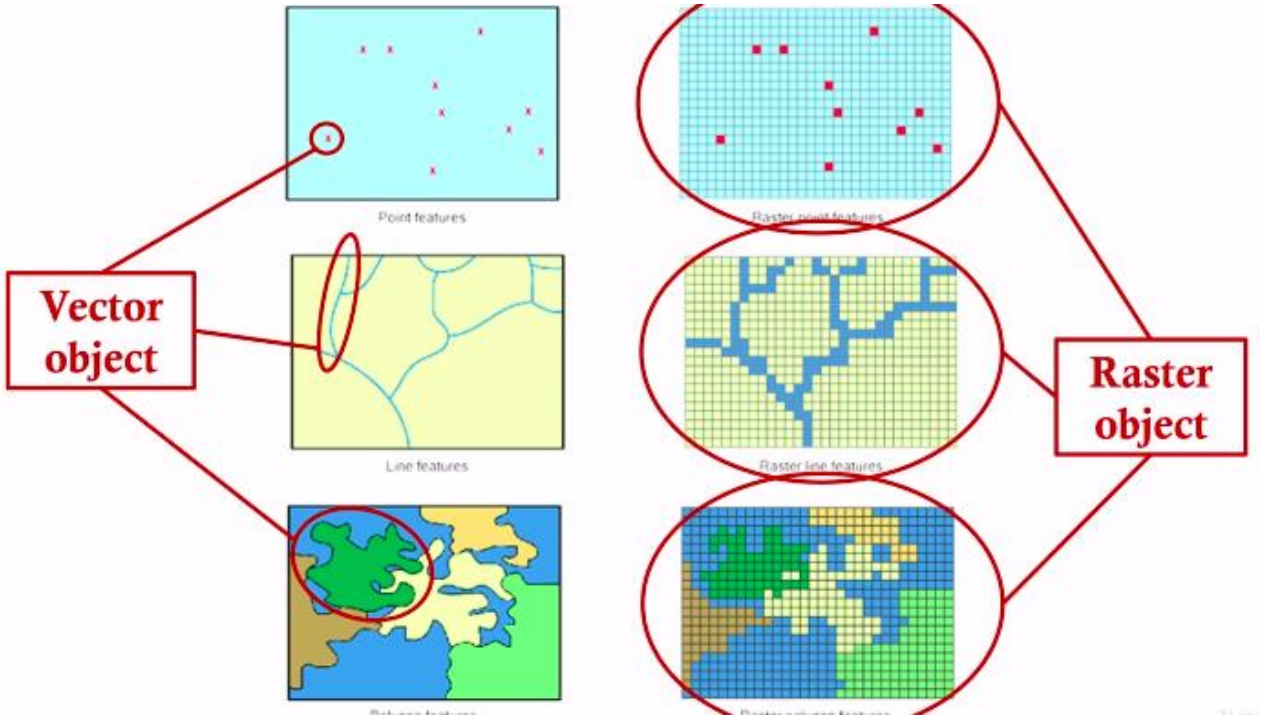


Source: wikipedia.org

Types of geospatial data

Vector data

- not made up of a grid of pixels.
- vector graphics are comprised of vertices and paths.
- points, lines, and polygons (areas).



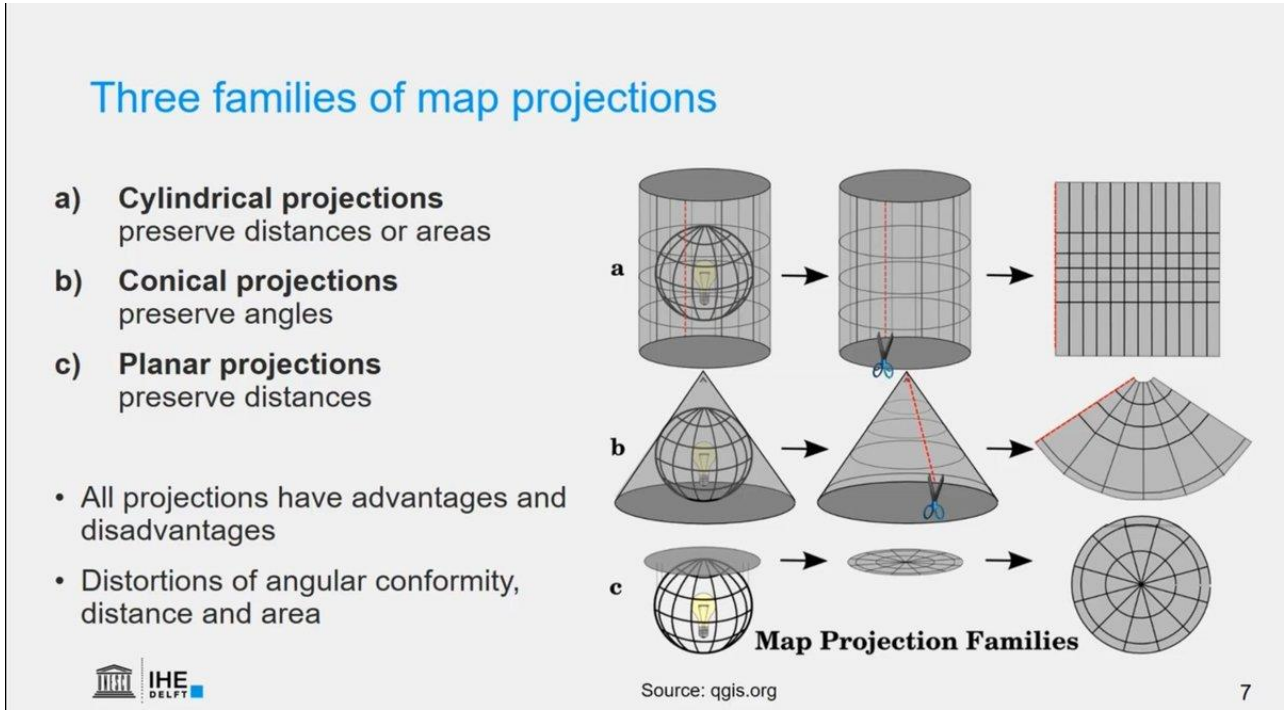
Source: gisgeography.com

Raster data

- made up of pixels (also referred to as grid cells).
- usually regularly spaced and square but they don't have to be.
- often look pixelated because each pixel has its own value or class.
- Eg. Each pixel value in a satellite image has a red, green, and blue value. Alternatively, each value in an elevation map represents a specific height. It could represent anything from rainfall to land cover.

Map projection

- Earth's surface in two dimensions is always with distortion.
- Thus, different map projections: to minimize the errors in shape and area.



ILLUSTRATIONS OF RELATIVE DISTORTIONS.

A striking illustration of the distortion and exaggerations inherent in various systems of projection is given in figures 42–45. In figure 42 we have shown a man's head drawn with some degree of care on a globular projection of a hemisphere. The other three figures have the outline of the head plotted, maintaining the latitude and longitude the same as they are found in the globular projection. The distortions and exaggerations are due solely to those that are found in the projection in question.

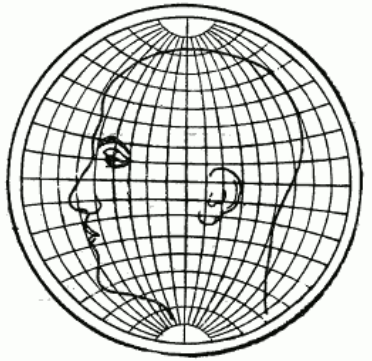


FIG. 42.—Man's head drawn on globular projection.

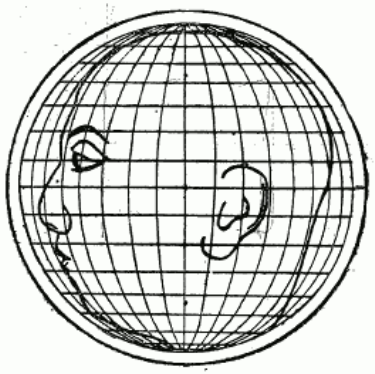


FIG. 43.—Man's head plotted on orthographic projection.

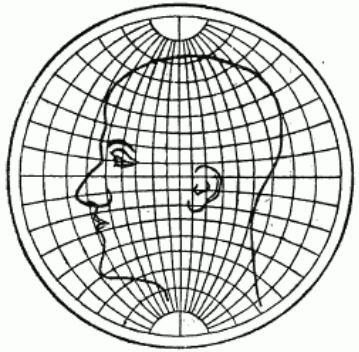


FIG. 44.—Man's head plotted on stereographic projection.

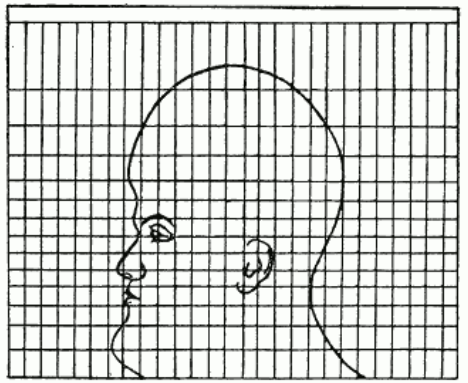


FIG. 45.—Man's head plotted on Mercator projection.

This does not mean that the globular projection is the best of the four, because the symmetrical figure might be drawn on any one of them and then plotted on the others. By this method we see shown in a striking way the relative differences in distortion of the various systems. The principle could be extended to any number of projections that might be desired, but the four figures given serve to illustrate the method.

Source: Charles Henry Deetz and Oscar Sherman Adams, 1921

QGIS: Lay-out, principles, tools

- Hands-on..
- Open QGIS - 3.22.13 (long term and stable version) – available in the shared materials

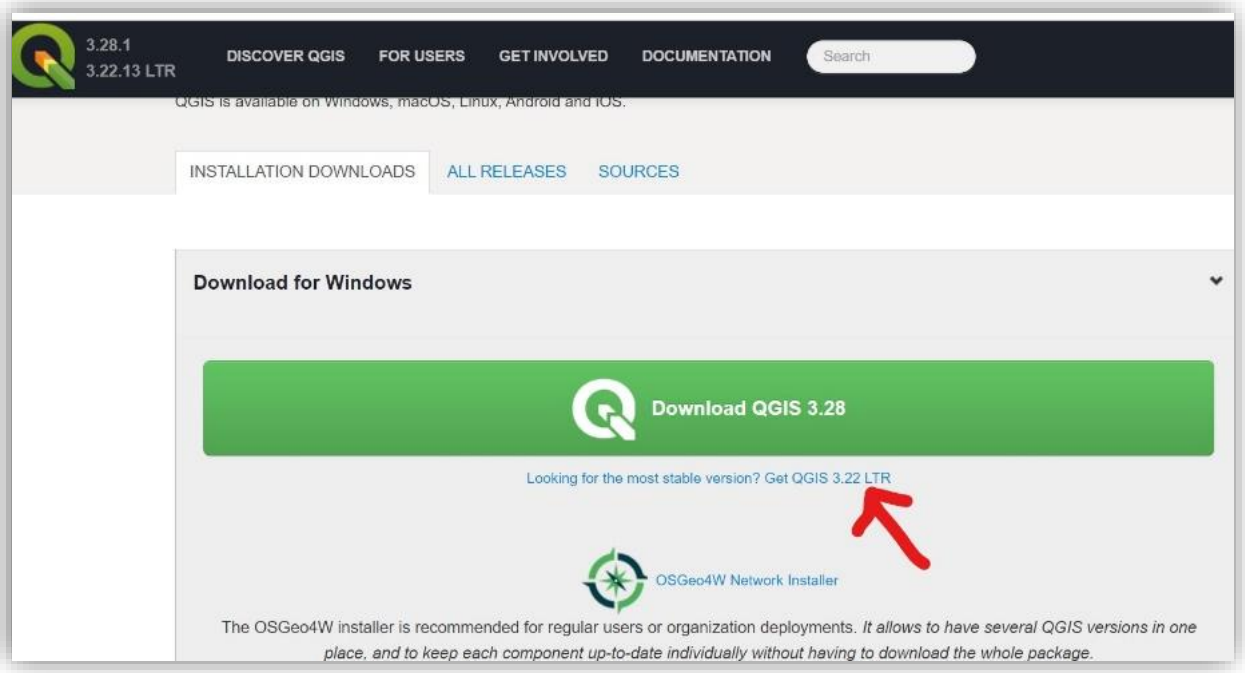
Downloading QGIS 3.22 (stable and long-term version)

1. Go to: https://www.qgis.org/en/site/forusers/download.html#

2. Select your system (Windows/macOS/Linux..)

3. Click on the blue text for long-term release. QGIS will be downloaded in your desktop.

4. Check the online USER MANUAL for QGIS: https://docs.qgis.org/3.22/en/docs/user_manual/



Adding data into QGIS

Finding and importing data and shapefiles:Hands-on session in QGIS

- Add vector layer (Shape files)
  - Point, polygon, line… (from the material folder)
  - https://www.diva-gis.org/gdata (shapefiles for all countries)
- Add raster layer (e.g., the environmental data)

- Add your data file excel (using ‘add vector layer’ tool/dragging the file directly to QGIS)

(In case if doesn’t work: Add delimited text layer (data file, e.g., our excel data table saved as CSV) (from the material folder))

- Checking and changing projection