Training Workshop provided by the IUSSP Panel on Early Career Perspectives

IUSSP Webinar Series

Extracting Spatial Data on Environment for Research

Trainer
Ankit Sikarwar
French Institute for Demographic Studies (INED) Paris, France
sikarwar.ankit-kumar@ined.fr

25 January 2023
12:00-15:00 UTC

Webinar outline

Part 1:

- Concepts and background of spatial data on environment
- Need for open access spatial data
- Applications to demographic research
- Sources to access spatial data on environment

Part 2:

- Hands-on session in QGIS
  (From importing data to extracting the indicators at desired spatial scale)
- Common errors while analysing geospatial data
- Discussion/Q&A

Concept and background of spatial data on environment

What we mean by spatial data?

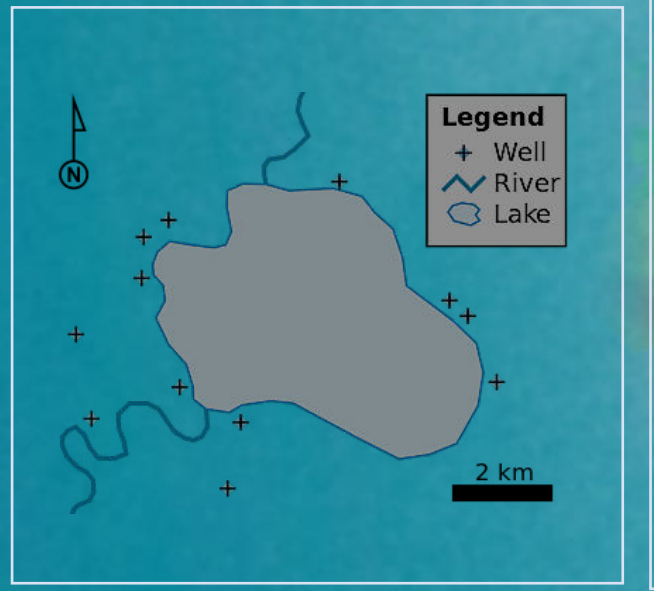
- any data directly or indirectly references a specific geographical area or location

Spatial data on Environment

- Indicators of environment for a particular location
- Rainfall, temperature, air pollution, land-use/cover, night time light,…..
- Stemming from ground data and/or earth observations

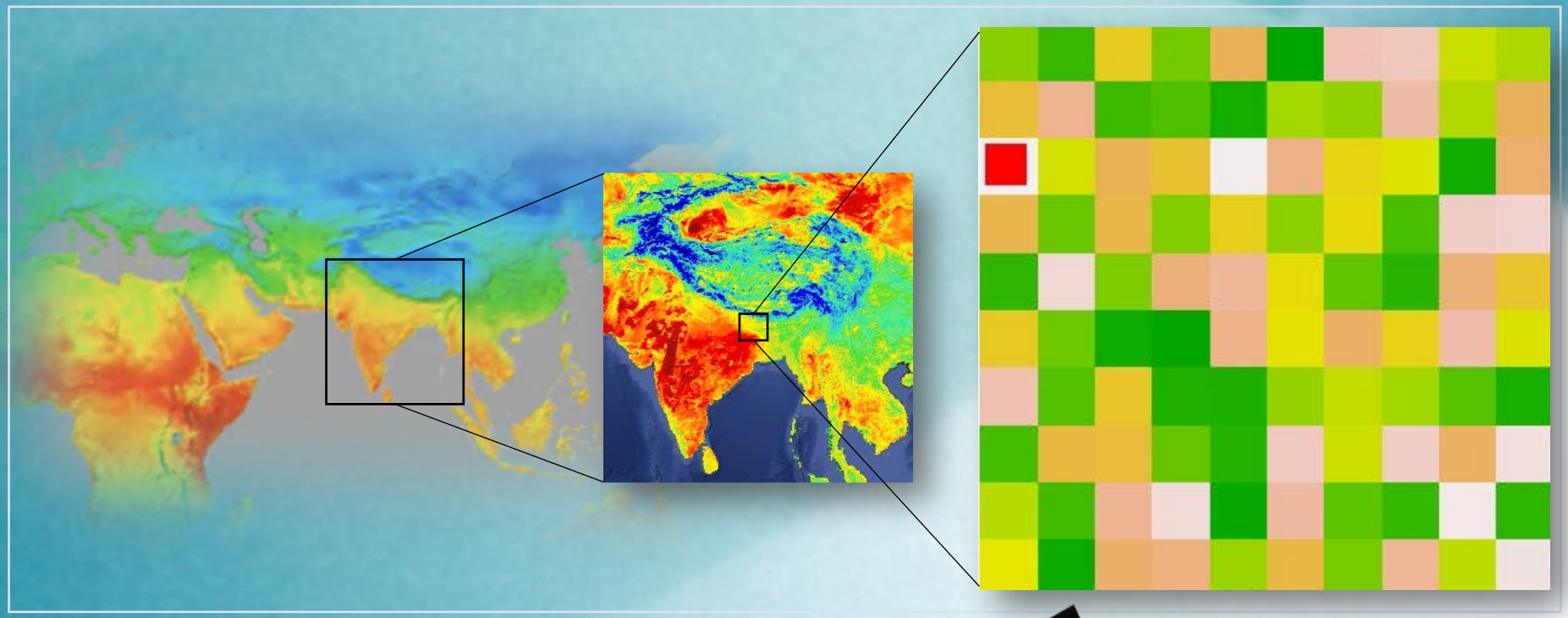
Form of spatial data on environment

Vector data



Source: wikipedia.org

Raster data

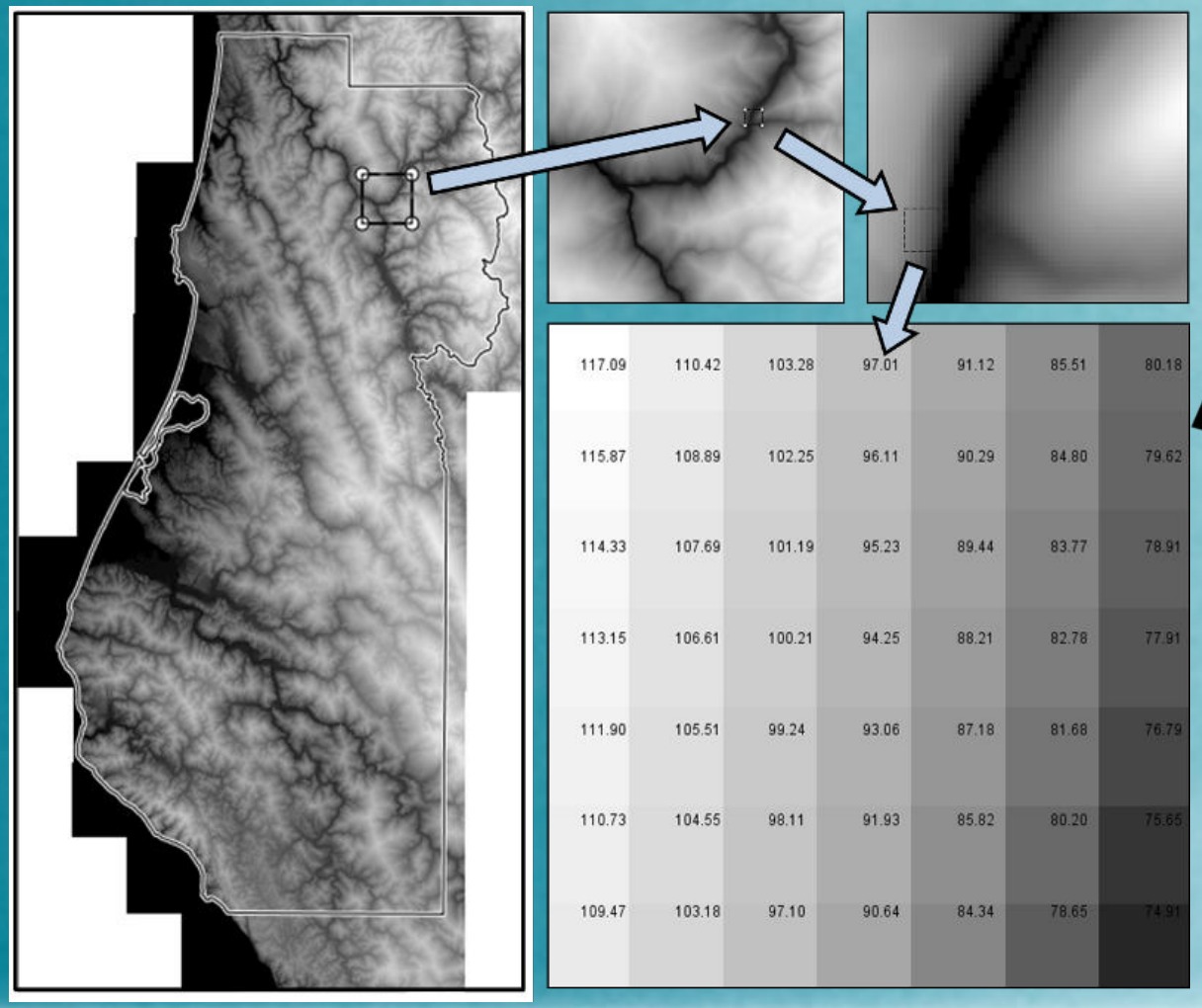


Images from ESA's climate initiative

Gridded surface of pixels (of specific size) presenting unique indicator value

Form of spatial data on environment

Raster data

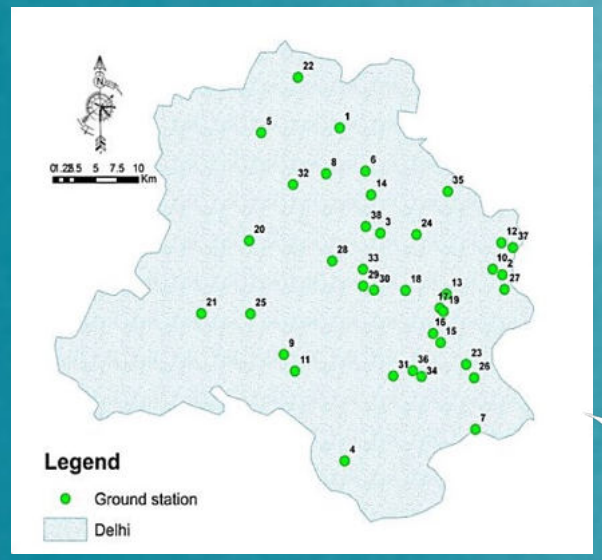


Gridded surface of pixels (of specific size) presenting unique indicator value

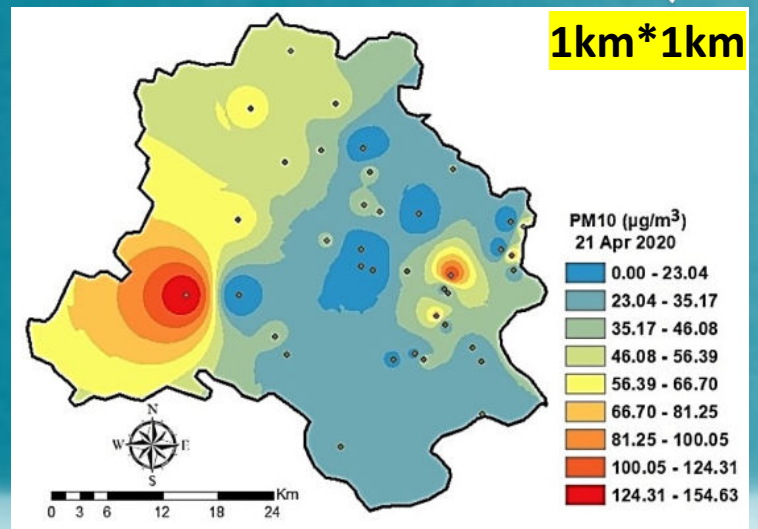
Source: https://gsp.humboldt.edu/

Concept and background of spatial data on environment

Form of spatial data on environment



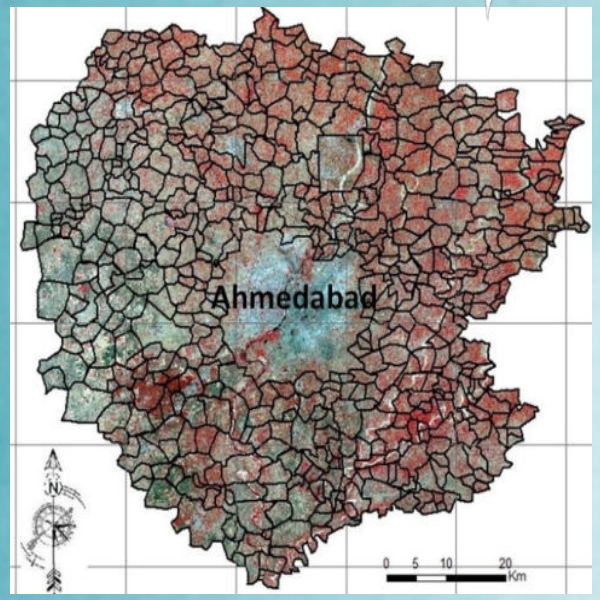
Raster generated **from vector** data using interpolation



Source: Presenter's analysis

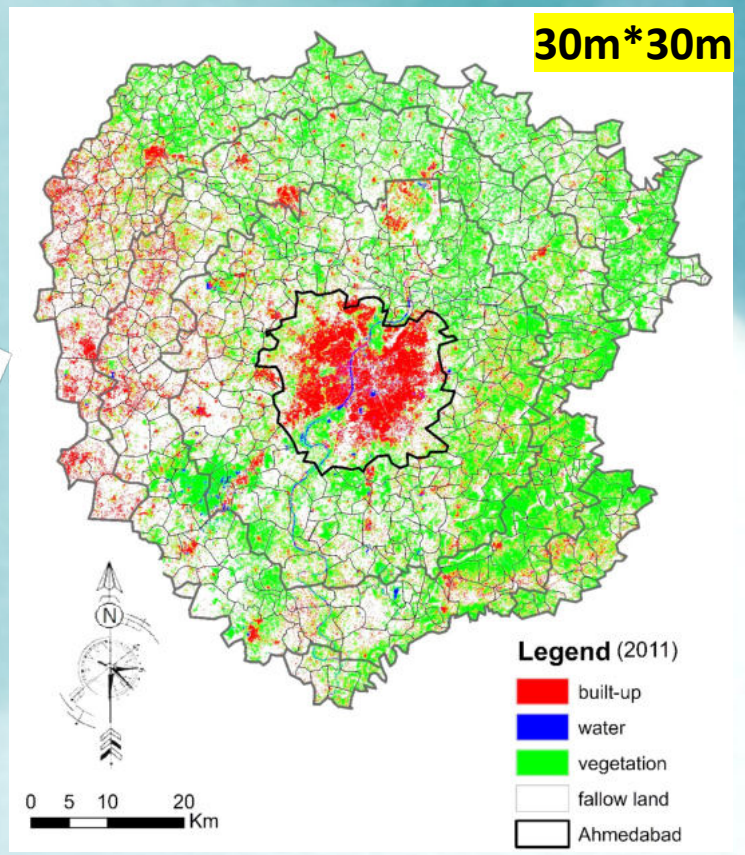
Number of ground monitoring stations

Raster re-generated **from raster** data using Supervised classification



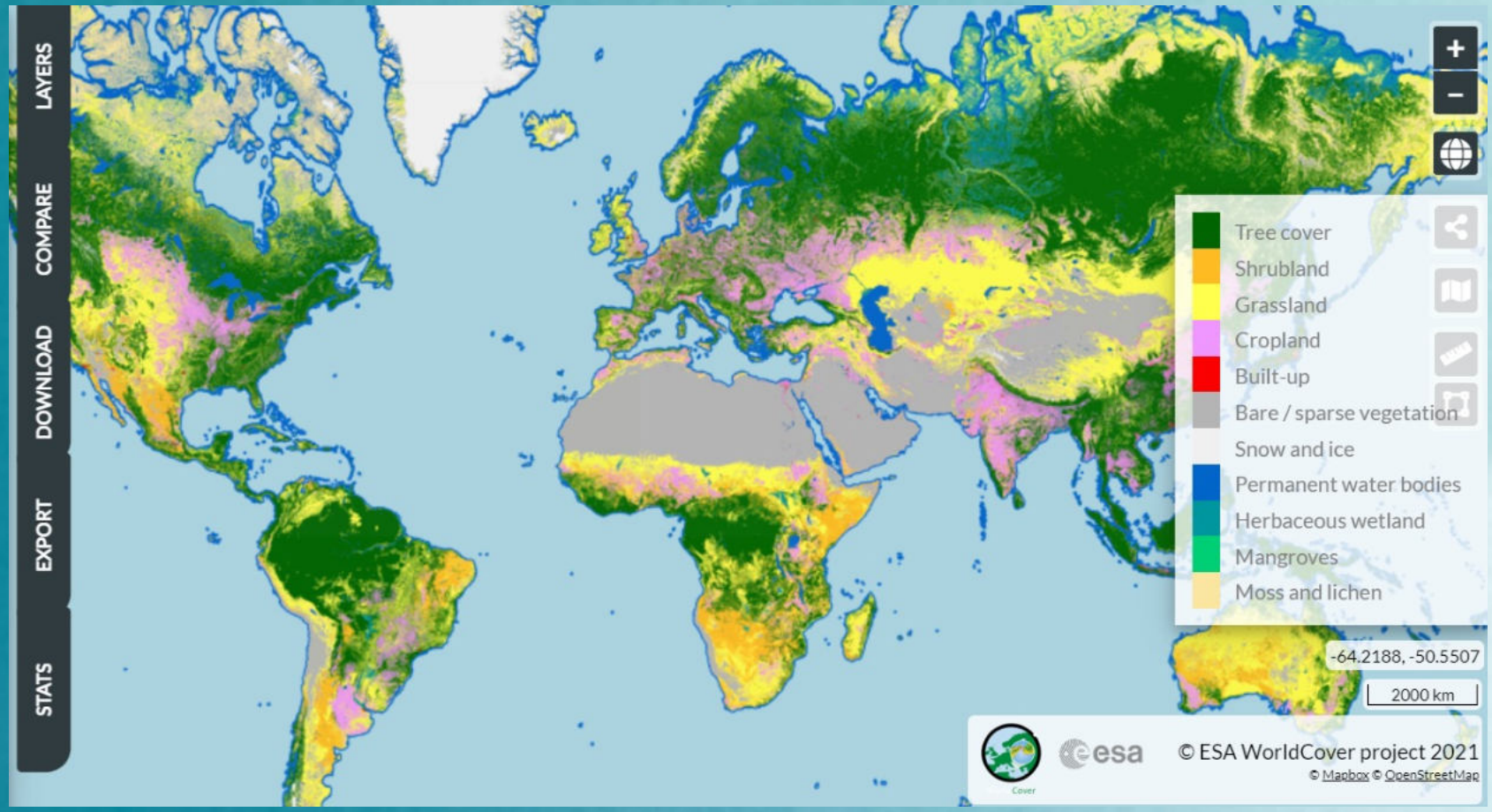
LANDSAT TM-5 imagery

Source: Presenter's analysis



Accuracy assessment

Form of spatial data on environment

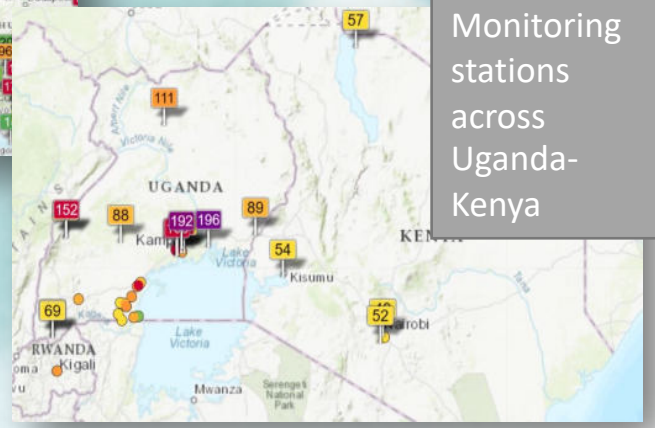


Ready to use open-access remotely sensed data pre-processed by experts

Need for open access spatial data from earth observations

- Intensifying and uncertain environmental parameters
- Need to be measured with timely and geographically rich data
- Insufficient number of ground monitoring stations in the global south, where most of the world population reside
- Fine-scale data provides a better understanding of how each variable impacts individuals, communities, populations, etc.
- Research and policy implications by identifying areas with greater environmental challenge and higher population at risk

Air Quality Monitoring stations across Europe



Air Quality Monitoring stations across Uganda-Kenya

Source: https://waqi.info/

Need for open access spatial data from earth observations

Source: Avtar et al. (2020) https://doi.org/10.1007/s10661-019-7996-9

- Direct and indirect use of earth observations and spatial data in monitoring and attainment of multiple targets of the Sustainable Development Goals
- Instrumental in countries where census and survey data on such indicators are not up-to-date and frequent.

SUSTAINABLE DEVELOPMENT GOALS

IUSSP UIESP

Applications to demographic research

- To explore complex connections of Population-Development-Environment
- Research questions: Where, when, how?

  - ❑ Climate change – migration patterns/socio-economic conditions/ fertility?
  - ❑ Air pollution – mortality and morbidity
  - ❑ Land use change – sustainability
  - ❑ Proximity connections – distance to water body and water-borne diseases

- Gridded population data solving the issues of spatial and temporal scales

Applications to demographic research

What type of analysis?

- Depends on research interest, objectives and availability of demographic data.

- Descriptive analyses: looking for correlations between environmental and demographic data - Visually / With classification techniques

Pixel level analysis

- Regression analyses: 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.

At individual/household level

At cluster level

At administrative area level

Applications to demographic research

Example of Individual level analysis

- 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)

ELSEVIER

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

Applications to demographic research

Example of cluster level analysis

- 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

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

Applications to demographic research

Example of area level (ecological) analysis

- India, district level
- Demographic data : Covid death counts 2020-2021 till second wave (crowd source data)
- Environmental data : Greenness-NDVI, district averages for Jan-March 2019, 1km, in quintiles (ISRO satellite data)

ELSEVIER

Contents lists available at ScienceDirect

Environmental Research

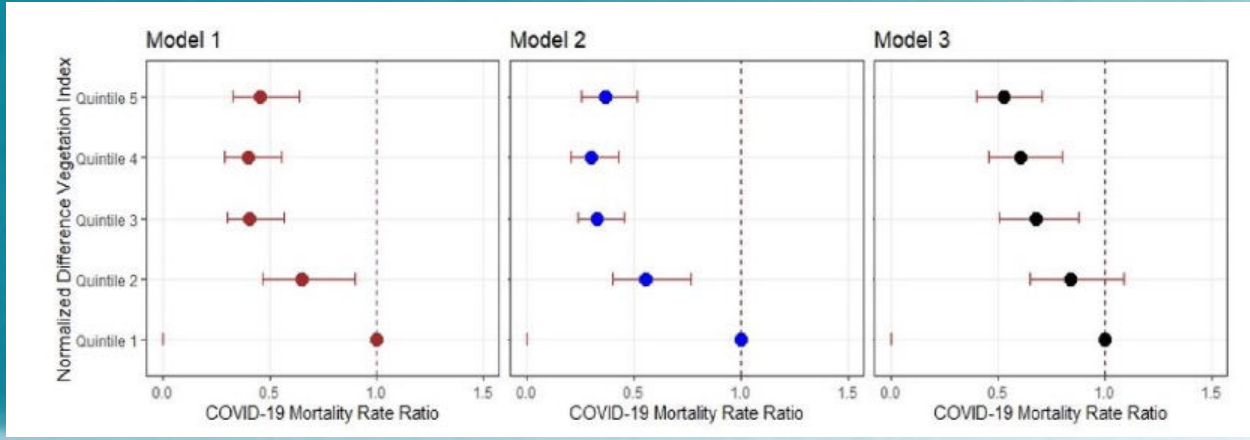
journal homepage: www.elsevier.com/locate/envres

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]

[a] French Institute for Demographic Studies (INED), Aubervilliers-Paris, France
[b] International Institute for Population Sciences, Mumbai, India
[c] Aix-Marseille University, IRD, LPED, Marseille, France

Results



Applications to demographic research

Example of pixel level analysis

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

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

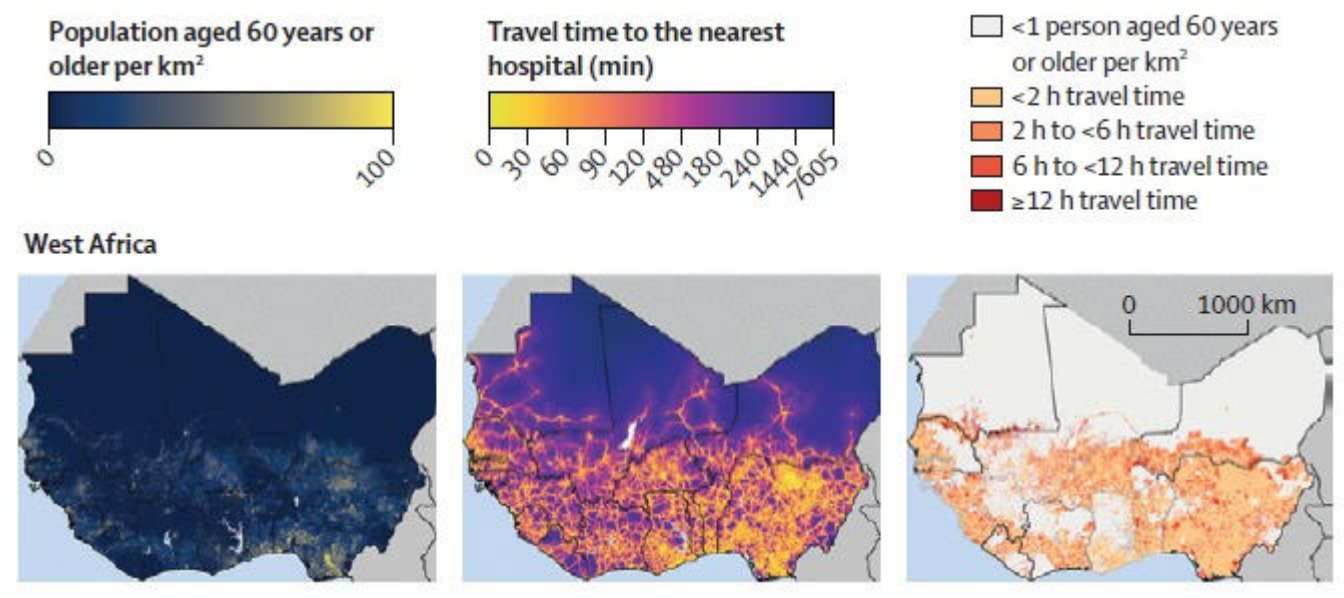
Pascal Geldsetzer, Marcel Reinmuth*, Paul O Ouma, Sven Lautenbach, Emelda A Okiro, Till Bärnighausen, Alexander Zipf*

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

Results



Sources to access spatial data on environment

Weather/Climate data

CHIRPS: Climate Hazards center InfraRed Precipitation with Station data
(Rainfall Estimates from Rain Gauge and Satellite Observations)
"Since 1999, USGS and CHC scientists—supported by funding from USAID, NASA, and NOAA—have developed techniques for producing rainfall maps, especially in areas where surface data is sparse."

CRU-TS: the UK Climate Research Unit's time series datasets
(exclusively based on weather stations observation)
The last version, released 26 May 2022, covers the period 1901-2021
Coverage: All land areas (excluding Antarctica) at 0.5° resolution
Variables: temperature • volume of hydrological precipitation • vapour pressure • wet days • cloud cover

Demonstration:

Worldclim (temperature and precipitation data at various resolution)

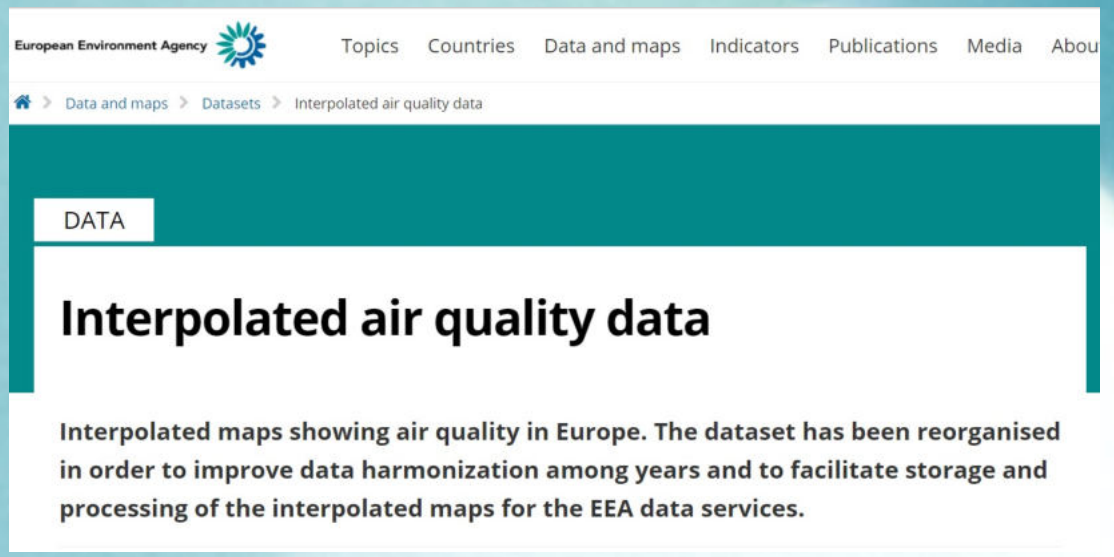
https://www.worldclim.org/data/worldclim21.html

Sources to access spatial data on environment

Atmospheric data

Gridded data on air-quality indicators

Data based on ground observations



Demonstration:

Atmospheric Composition Analysis Group (PM2.5, NO2..)

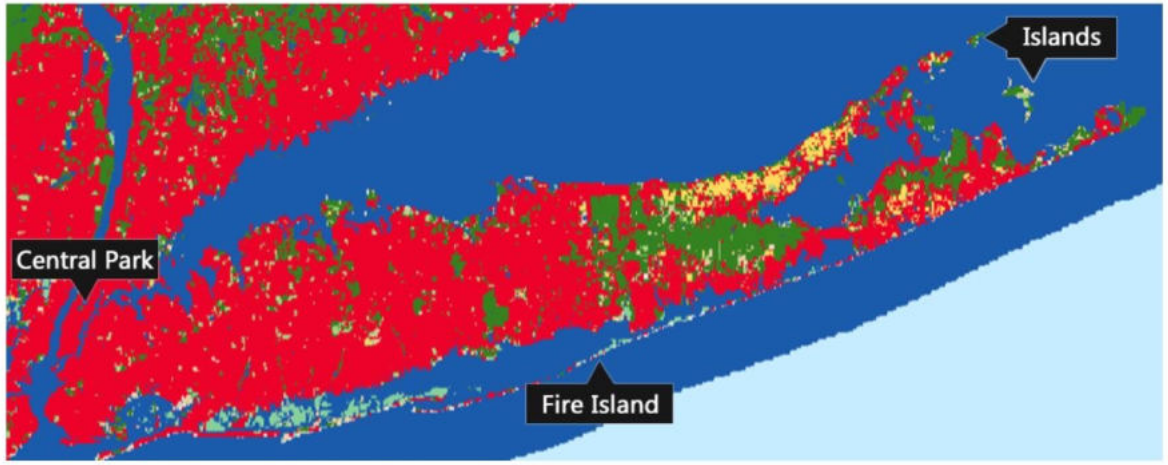
https://sites.wustl.edu/acag/datasets/

Sources to access spatial data on environment

Land use / Land cover data

- Pre-processed data sets:
- Multiple sources of data (varying in spatial resolution, temporal scales, methods, and categories)

1. Esri Land Cover 10m



2. Global Land Survey (GLS)

More details and other sources:
https://gisgeography.com/free-global-land-cover-land-use-data/

Demonstration:

ESRI living Atlas (ready to use LULC data)

https://livingatlas.arcgis.com/landcover/

Sources to access spatial data on environment

Satellite imageries

MODIS : the Moderate Resolution Imaging Spectroradiometer on NASA Terra (and Aqua) satellite(s) (temperature data retrieved at the moment)
« viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands - These data will improve our understanding of global dynamics and processes occurring on the land, and in the lower atmosphere.

SENTINEL : The Copernicus Program is an ambitious initiative headed by the European Commission in partnership with the European Space Agency (ESA). The Sentinels include all-weather radar images from Sentinel-1A and -1B, high-resolution optical images from Sentinel 2A and 2B, as well as ocean and land data suitable for environmental and climate monitoring from Sentinel 3.

How to access them ?
Copernicus - Google Earth Engine –Earth Explorer

Demonstration:

NASA: earth explorer (satellite images)
Platform to access satellite images for specific times and regions...
https://earthexplorer.usgs.gov/

Sources to access spatial data on environment

Global gridded population data

From **GPW** (gridded population of the world) in 1995 [1kmx1km, little modelling (areal weighting/Protected areas and water bodies)]

… to **GRUMP** from 2011 (a little more modelling, using night lights)

… to **Worldpop** today [100m x 100m, more modelling (machine learning / Roads, land cover, built structures, cities or urban areas, night time lights, infrastructure, environmental, protected areas, and water bodies)]

Demonstration:

PopGrid: a compilation of different gridded data for demographic indicators https://www.popgrid.org/
https://worldpop.org

Hands-on session in QGIS

Objective

- Estimating climate data for the administrative units (cantons) of France

- Importing remotely sensed data
- Reading data
- Data editing
- Geo-processing
- Extraction of indicators at desired spatial scale

	NAME_4	VARNAME_4	TYPE_4	ENGTYPE_4	_tmaxcount	_tmaxsum
1	Bischwiller	NULL	Canton	Cantons	12	214.233812332...
2	Haguenau	NULL	Canton	Cantons	20	348.496332168...
3	Niederbronn-les...	NULL	Canton	Cantons	16	263.496637344...
4	Molsheim	NULL	Canton	Cantons	17	270.597736358...
5	Rosheim	NULL	Canton	Cantons	8	112.717980384...
6	Saales	NULL	Canton	Cantons	7	96.0789451599...
7	Schirmeck	NULL	Canton	Cantons	11	147.440326690...
8	Wasselonne	NULL	Canton	Cantons	12	187.742876052...
9	Barr	NULL	Canton	Cantons	10	153.845879554...
10	Benfeld	NULL	Canton	Cantons	7	124.459865570...

Side by side things!

Map projection

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

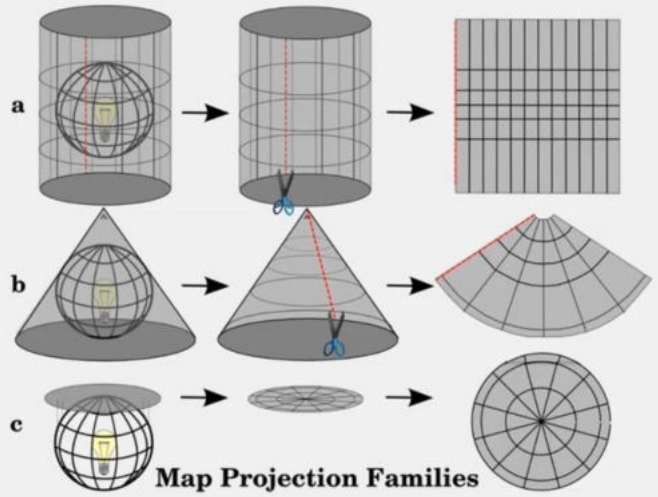
Three families of map projections

a) **Cylindrical projections**
preserve distances or areas

b) **Conical projections**
preserve angles

c) **Planar projections**
preserve distances

- All projections have advantages and disadvantages
- Distortions of angular conformity, distance and area



Map Projection Families

Source: qgis.org

7

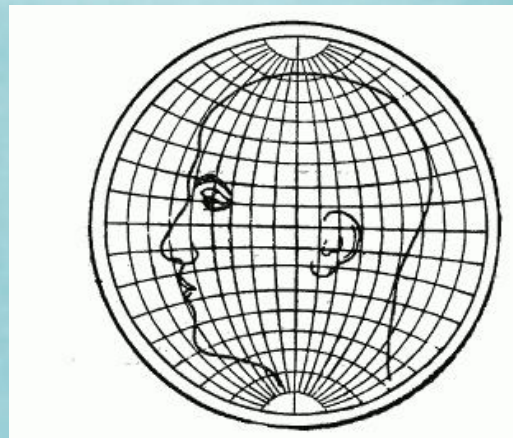


Fig. 42.—Man's head drawn on globular projection.

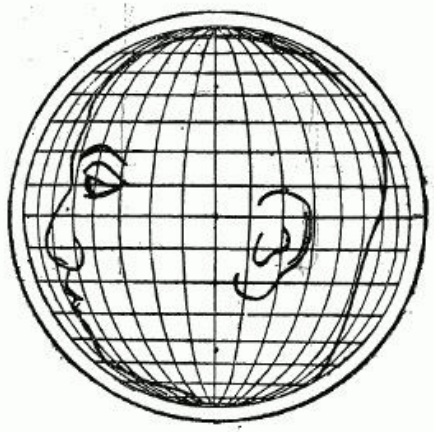


Fig. 43.—Man's head plotted on orthographic projection.

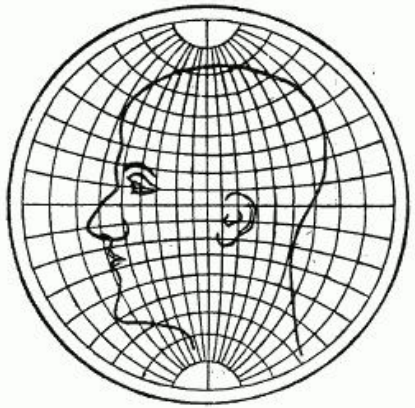


Fig. 44.—Man's head plotted on stereographic projection.

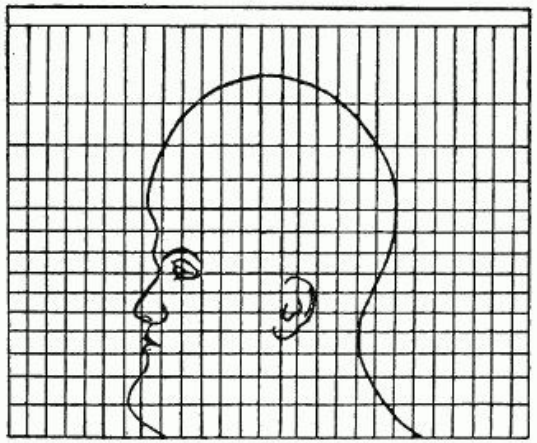
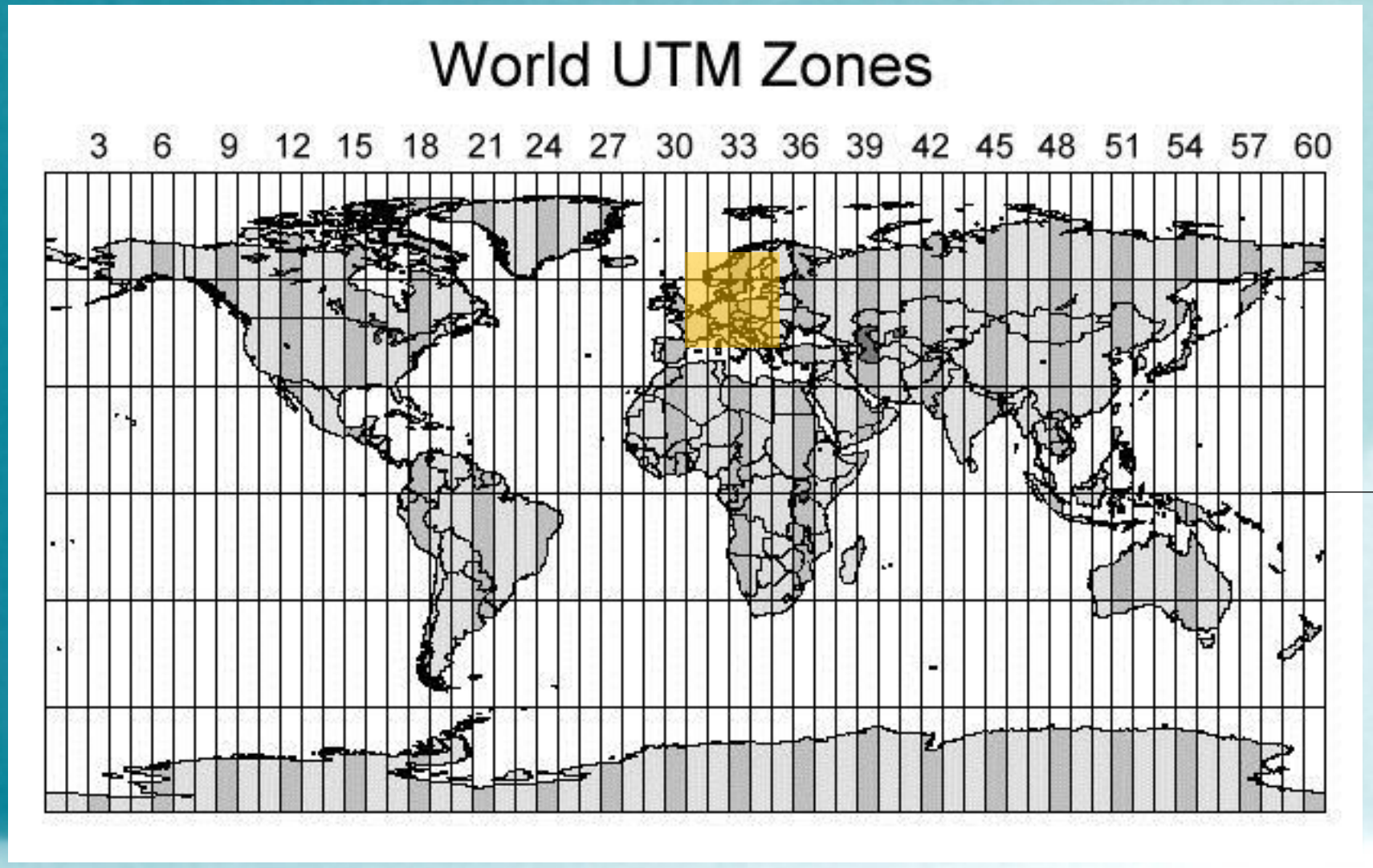


Fig. 45.—Man's head plotted on Mercator projection.

Source: Charles Henry Deetz and Oscar Sherman Adams, 1921

Side by side things!

Map projection



For this area of interest (Part of Europe)
30-35 N zone is ideal

https://www.xmswiki.com/wiki/UTM_Coordinate_System

Side by side things!

Raster calculator is important to perform calculations over your raster.

For instance,

to convert temperatures in degrees Fahrenheit to Celsius

(Raster layer - 32) x .5556

to calculate NDVI (Normalized difference vegetation index)
from multispectral satellite images

(NIR - R) / (NIR + R)

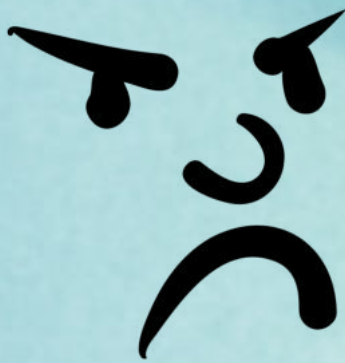
Also to apply weights or to make criteria based selection of indicators

Land surface temperature > 40 degree Celsius

Avoid common errors during raster processing

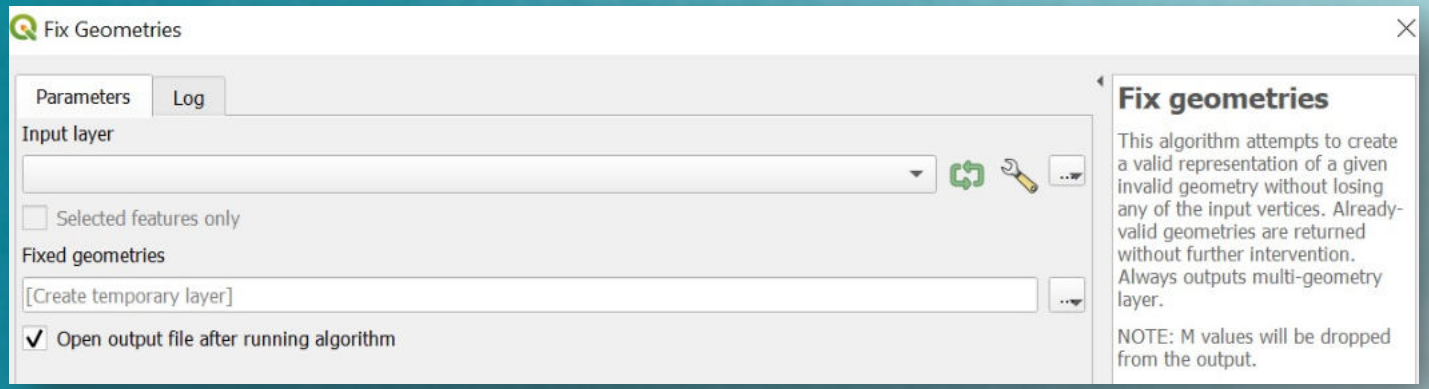
Right choice of projection

- For the objectives related to distance (proximity, buffer, area, etc.) projection unit must be in meters/km/mile.. (not as degrees)
- Idea is UTM projection.
- Match the projections of all layers (vector/raster)



Don't
PANIC !
Errors are common in
RS and GIS

Fix geometry of shapefile before processing



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

IUSSP
UIESP

Acknowledgements

- IUSSP team: Nico van Nimwegen, Paul Monet, Mary Ellen Zuppan, and Shireen Jejeebhoy for organizing this webinar and providing necessary intellectual support.
- Big thanks to Valérie Golaz from INED, who has helped in conceptualizing and design the course.
- My team members of the Early Career Perspective Panel at IUSSP
- Enthusiastic participants from various parts of world
- Léo Lipovac and Basile Rousse (INED) for extending help during the webinar.
- French Institute for Demographic Studies (INED), France.

Thank you!

Ankit Sikarwar (INED)

Sikarwar.ankit-kumar@ined.fr / anks.sik@gmail.com