Data Science in Economic History: Spatial Analysis and Mapping

Timur Öztürk

Lecture 1: Introduction

April 29, 2025

Overview

Introduction

- This course provides an overview of spatial analysis and mapping in R with applications in economic history.

- **First part**:
  - Introduction to the basic toolkit in geospatial research (as used in economics).
  - Topics: Handling of vector and raster data, processing of spatial data, map making.

- **Second part**:
  - Applies the tools introduced in the first part.
  - Replicates spatial analyses of papers in economic history and long-run development.

Textbooks

The following two online textbooks cover much of the tools discussed in this course:

1. Taro Mieno (2022). R as GIS for Economists. (short: GISEcon)
   `https://tmieno2.github.io/R-as-GIS-for-Economists/`
2. Robin Lovelace, Jakub Nowosad, Jannes Muenchow (2023). Geocomputation with R. 2nd edition. CRC Press. (short: LNM)
   `https://geocompr.robinlovelace.net/`

Geographic Information Systems (GIS)

- GIS is a computer-based system designed to capture, store, manipulate, analyze, and present spatial data.
- Spatial data is any data that has a location or geographic component, such as GPS coordinates, addresses, or zip codes.
- GIS allows users to visualize spatial data on maps, and to perform spatial analyses and modeling.
- Many of you might have some first experience with using GIS software emphasizing graphical user interfaces (QGIS, ArcGIS).

R as GIS

- This course focuses on programming GIS tasks using R rather than manually implementing them using graphical interfaces.

- Why? Two key advantages of command-line interfaces:
  1. Reproducibility[1] is essential for scientific research.
  2. Manual operations are not scalable.

- Python is a valuable alternative to R.

[1] "A process in which the same results can be generated by others using publicly accessible code." (LNM)

Overview

Why GIS for economists & economic historians?

There are at least two reasons for the increasing popularity of GIS (e.g., Kudamatsu, 2018).

1. **Data**: GIS allows economists to observe the previously unobserved.
2. **Identification**: Geography as a source of exogenous variation.

GIS as data source: Satellite images[2]

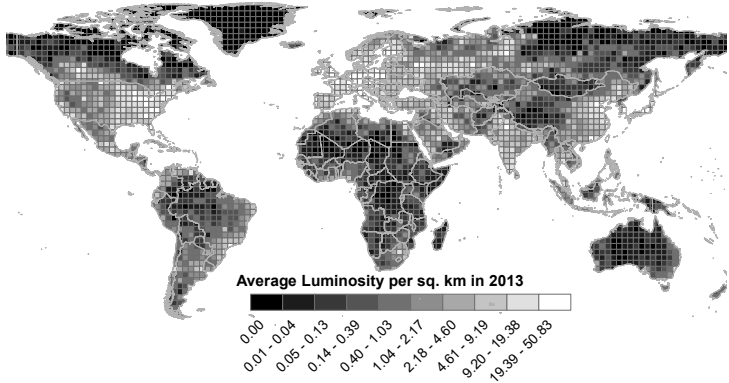
Satellite images can close gaps in the official statistics due to illegality (e.g. deforestation in India) or lack of state capacity (e.g. GDP growth).

Source: Michalopoulos and Papaioannou (2018). *Spatial Patterns of Development*. Fig. 1.

[2] See Michalopoulos and Papaioannou (2018) for a survey.

GIS as data source: Satellite images (ctd.)

Satellite images are also not bound to administrative borders and can be aggregated as needed.



Source: Michalopoulos and Papaioannou (2018). *Spatial Patterns of Development*. Fig. 7.

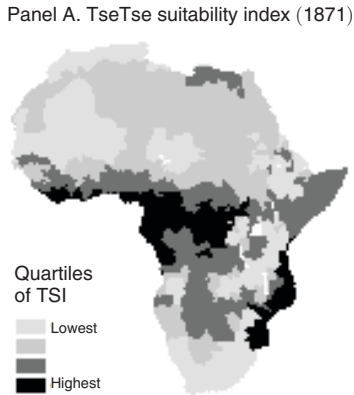
GIS as data source: Old maps

Using GIS, scanned old maps can be used to construct spatial historical data often retained by no other written source.

Source: Magnus, O. (1572) *Map of the Sea*. Rome: Antoine Lafréry. `https://www.loc.gov/item/2021668418/`.

GIS for identification: Examples

Climate as a source of (exogeneous?) variation (e.g. TseTse fly suitability index used in Alsan, 2015).



Source: Alsan, M. (2015). *The Effect of the TseTse Fly on African Development*. Fig. 4, Panel A.

GIS for identification: Examples (ctd.)

Geographic features as instrumental variables (e.g. distance to important destinations of slave trades in Nunn, 2008).

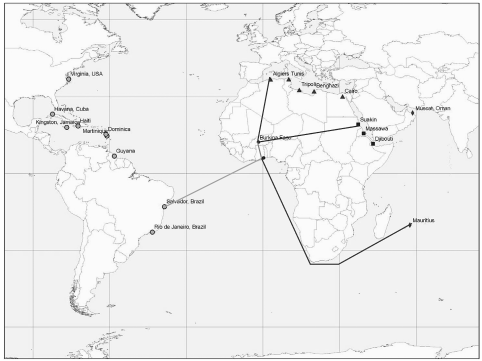


FIGURE V
Example Showing the Distance Instruments for Burkina Faso

Source: Nunn, N. (2008). *The Long-Term Effects of Africa's Slave Trades*. Fig. V.

GIS for identification: Examples (ctd.)

Regression discontinuities using (historical) borders (e.g. borders of Peru's forced mining labor system in Dell, 2008).

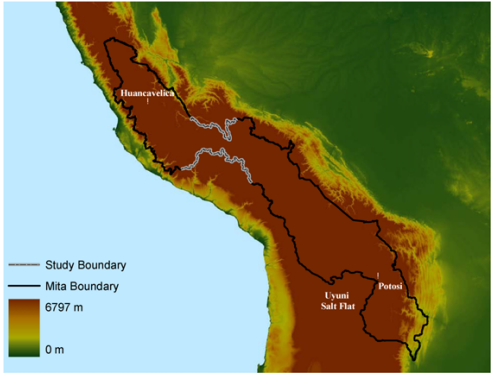


FIGURE 1.—The *mita* boundary is in black and the study boundary in light gray. Districts falling inside the contiguous area formed by the *mita* boundary contributed to the *mita*. Elevation is shown in the background.

Source: Dell, M. (2010). *The Persistent Effects of Peru's Mining Mita.* Fig. 1.

Overview

Demonstration I: Nunn (2008)

- Studies the effect of Africa's slave trades 1400-1900 on per capita GDP today.
- Uses an IV approach, instrumenting slaves taken from a country with the distance to the nearest slave trade centre in the Americas.
- Let's see how to construct the instrumental variable using R as GIS...

Instrumental variable

Distance from the point on the coast that is closest to the country's centroid to the closest major market of the Atlantic slave trade.

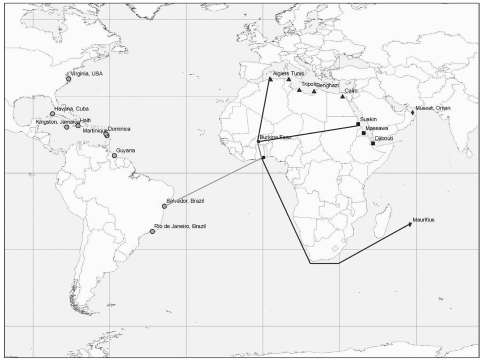


FIGURE V
Example Showing the Distance Instruments for Burkina Faso

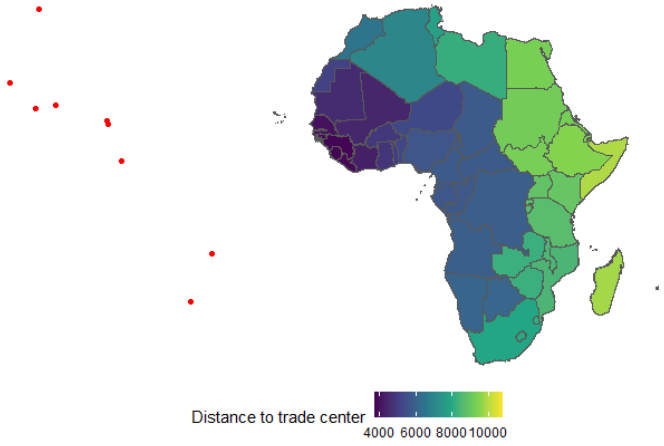
Source: Nunn, N. (2008). *The Long-Term Effects of Africa's Slave Trades*. Fig. V.

Determine each country's centroid [3]

[3] R-script: `chapter-1-6-R-as-GIS-Nunn-2008.R`

Determine point on the coast closest to centroid 🖵

Calculate distance to closest major market

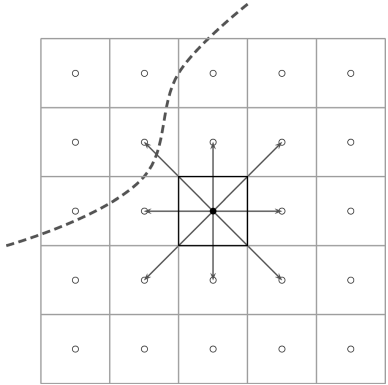


Demonstration II: Nunn and Puga (2012)

- Rugged terrain is tough to farm, costly to traverse, and often inhospitable to live in–and thus impairs economic development.
- Study shows that in Africa, countries with a rugged landscape tend to perform better than flatter ones.
- In Africa, rugged terrain decreased the exposure to the slave trades and thus their negative long-run effects on economic development.
- Let's see how to construct a ruggedness index using R as GIS...

Measuring ruggedness

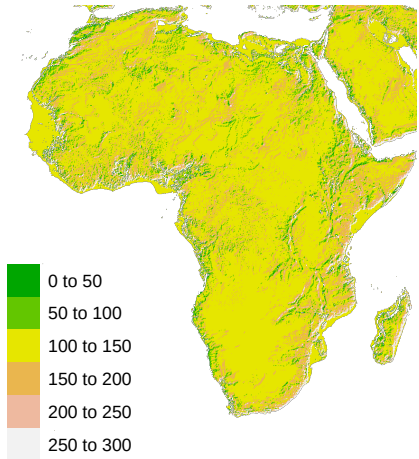
Square root of the sum of the squared differences in elevation between the central point and the eight adjacent points.[4]



Source: Nunn, N. and D. Puga (2012). *Ruggedness: The Blessing of Bad Geography in Africa*. Fig. 1.

[4] Meant to capture small-scale terrain irregularities, such as caverns, caves, and cliff walls.

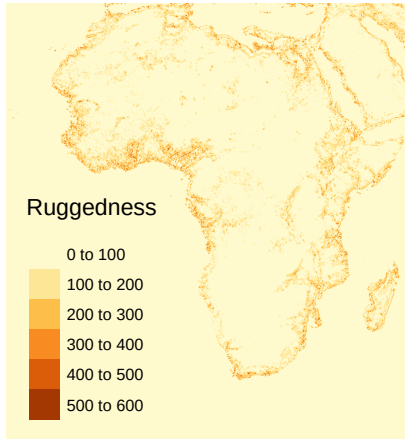
Load and plot raster data on elevation [5]



[5] R-script: `chapter-1-7-R-as-GIS-Nunn-2012.R`

Let R calculate the ruggedness index 🖵

- Define function for terrain ruggedness index (TRI).
- Loop through raster cells and calculate TRI (see map).
- Averaging across grid cells on land gives a country's index.



Overview

Seminar programme: Part 1 (preliminary)

Date	**Topic**	Core reading	**Lecturer**
18.4.	Introduction	GISEcon Ch. 1	SB
25.4.	Getting started with R		TÖ
2.5.	Handling vector data	LNM Ch. 2.2; GISEcon Ch. 2	SB
9.5.	Processing and manipulating vector data	LNM Ch. 3.2, 4.2, 5.2; GISEcon Ch. 3	SB
16.5.	Handling raster data	LNM Ch. 2.3; GISEcon Ch. 4	SB
23.5.	Processing and manipulating raster data	LNM Ch. 3.3, 4.3, 5.3; GISEcon Ch. 5	SB
30.5.	No lecture (Pentecost)		
6.6.	Raster-vector interactions	LNM Ch. 6	RF
13.6.	Making maps with R	LNM Ch. 9	RF

General information

- Slides and R-scripts will be made available on **e-learning**.
- **Target groups**:
  - Master's students in History and Economics, Economics and related degree programmes
  - Advanced undergraduate students with a profound knowledge of econometrics and R
- **Prerequisites**:
  - Students should be familiar with basic econometric methods.
  - Prior exposure to R is recommended (although not strictly necessary).

Grading

1. Two take-home **problem sets**: 20%
   - Problem sets will apply methods taught in the course to real-world data using R.
   - Due on 16.05. and 20.06.2023,

2. **Term paper** (max. 5000 words): 80%
   - Term paper will analyse a pre-specified research question in R using the methods taught in the course.
   - Involves both coding and discussion of results.
   - Due on 30.9.2023

Overview

Readings (* optional; ** recommended)

* Kudamatsu, M. (2018). GIS for Credible Identification Strategies in Economics Research. *CESifo Economic Studies* 64 (2): 327–338.

* Michalopoulos, S. and E. Papaioannou (2018). Spatial Patterns of Development: A Meso Approach. *Annual Review of Economics* 10 (1): 383–410.

** Mieno, T. (2022). *R as GIS for Economists*. Chapter 1.6 and 1.7.

References

Alsan, M. (2015). The Effect of the TseTse Fly on African Development. *American Economic Review* 105 (1): 382–410.

Nunn N. (2008). The Long Term Effects of Africa's Slave Trades. *Quarterly Journal of Economics* 123(1): 139–176.

Nunn, N. and D. Puga (2012). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics* 94 (1): 20–36.