

# GIS data, QGIS and application in Development

Etienne Madinier<sup>1 2</sup>

<sup>1</sup>Paris School of Economics - EHESS, <sup>2</sup>ENS de Lyon

October 2022

email: [etienne.madinier@gmail.com](mailto:etienne.madinier@gmail.com)

## What to expect from this block?

- Theoretical considerations
- Examples of empirical work using GIS data
- Technical knowledge on how to use and represent those data
- Practical advice on the workflow
- NO spatial econometrics (despite being fascinating)

# Today

- 1 Why GIS?
  - Maps as a visualization tool
  - New data sources
- 2 Remote sensing information
- 3 Examples of datasets and applications
  - One dataset
  - Two examples of papers
- 4 Introduction to Qgis

# GIS: Geographic Information System

→ **georeferenced information**

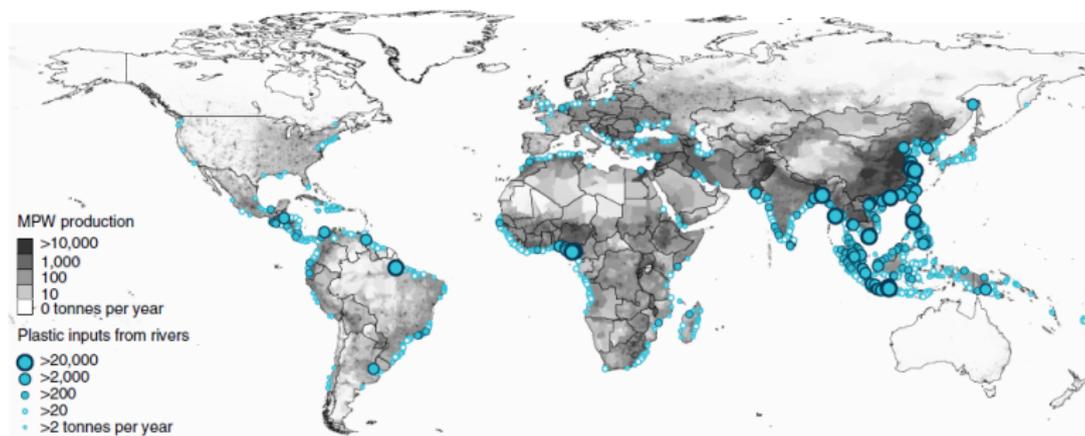
Maps take us on a journey through space and time.



**Figure:** Globe of Martin Behaim, the oldest known today, (1492)

# Why geocoded datasets?

Because maps are powerful visualization tools



**Figure 1 | Mass of river plastic flowing into oceans in tonnes per year.** River contributions are derived from individual watershed characteristics such as population density (in  $\text{inhab km}^{-2}$ ), mismanaged plastic waste (MPW) production per country (in  $\text{kg inhab}^{-1} \text{d}^{-1}$ ) and monthly averaged runoff (in  $\text{mm d}^{-1}$ ). The model is calibrated against river plastic concentration measurements from Europe, Asia, North and South America.

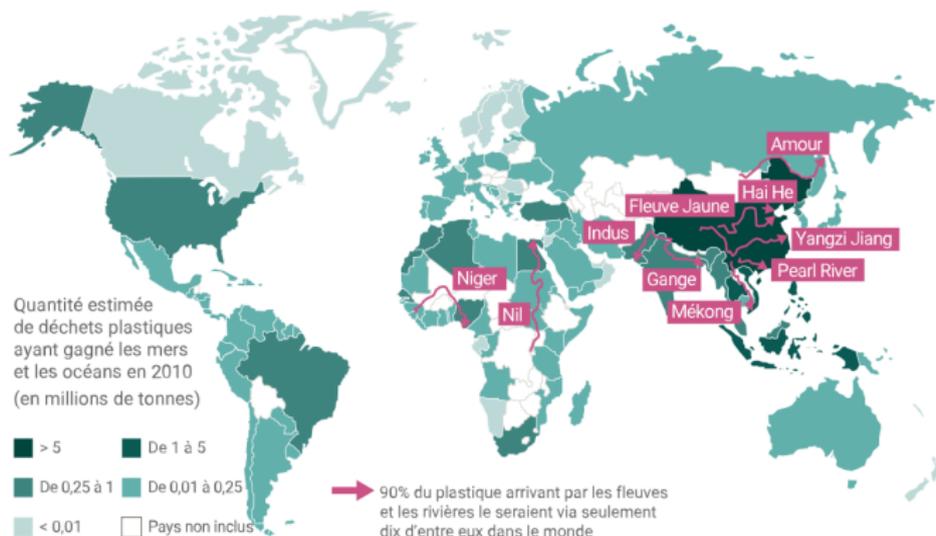
2

NATURE COMMUNICATIONS | 8:15611 | DOI: 10.1038/ncomms15611 | www.nature.com/naturecommunications

Figure:

# Why geocoded datasets?

Because maps are powerful visualization tools

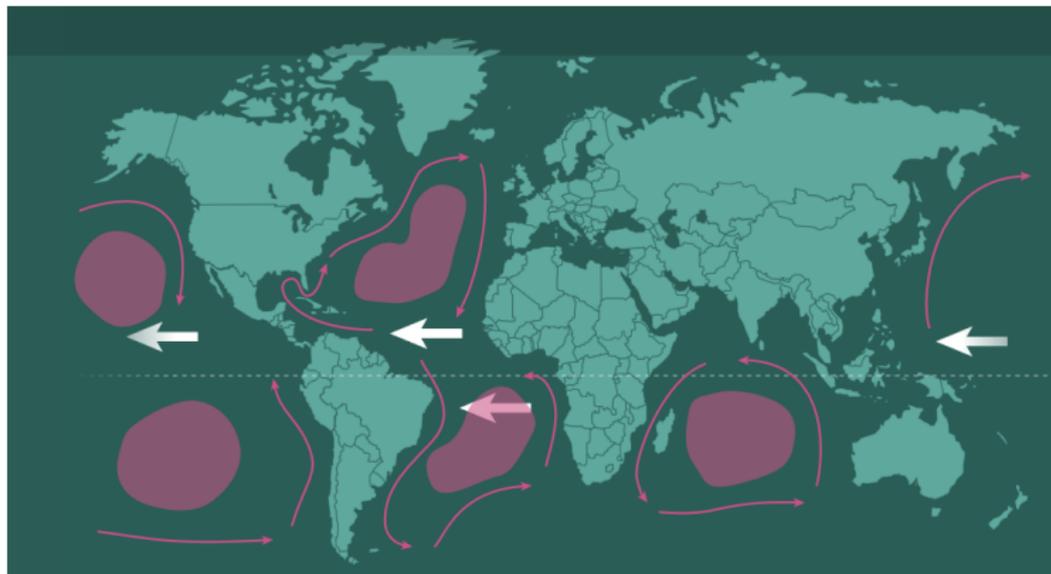


**Figure:** 5 countries and 10 rivers are generating 80% of ocean's pollution

Source: <https://labs.letemps.ch/interactive/2018/longread-homo-plasticus/>

# Why geocoded datasets?

Because maps are powerful visualization tools



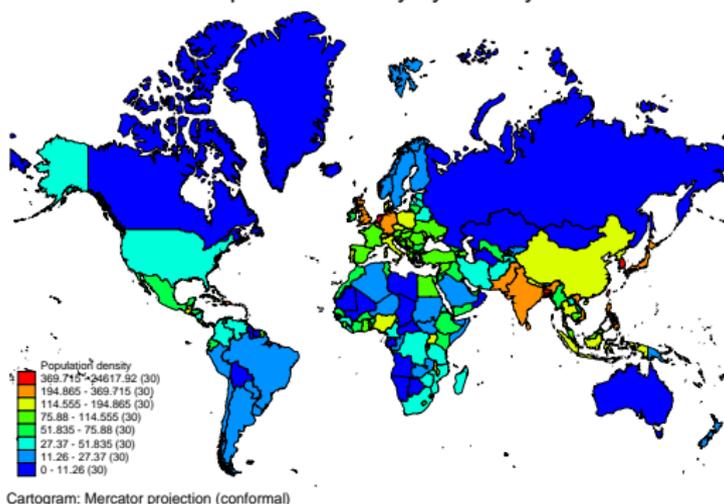
**Figure:** Accumulation zones of plastic floating waste

Source: <https://labs.letemps.ch/interactive/2018/longread-homo-plasticus/>

# Maps

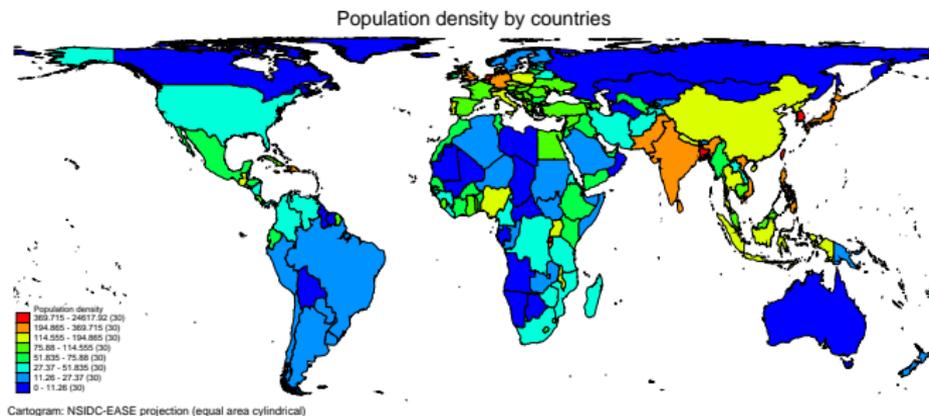
- Powerful visualization tools when used with caution...
- BUT may also produce very biased views of the world.

## Population density by country



## Mercator projection:

- larger size distortion as distance to equator increases
- inflates the importance of some countries

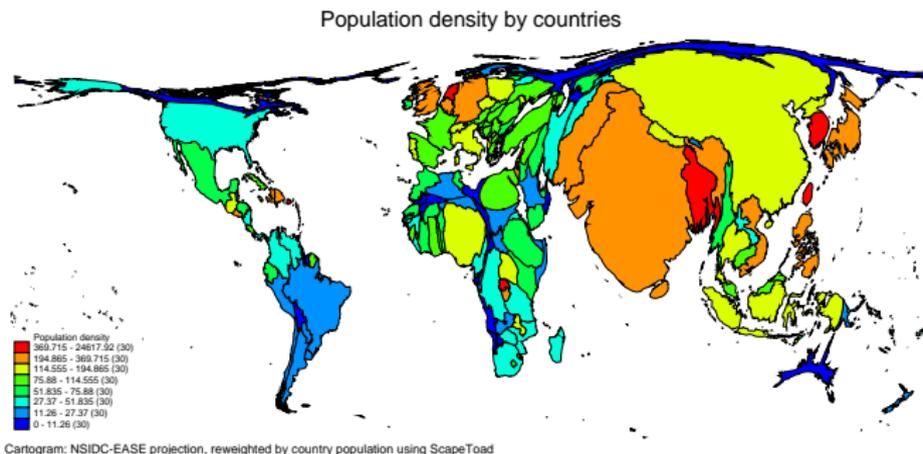


## Equal area projection (NSDIC-EASE)

- large shape distortion
- area preserving

Most of the population lives in densely populated countries.

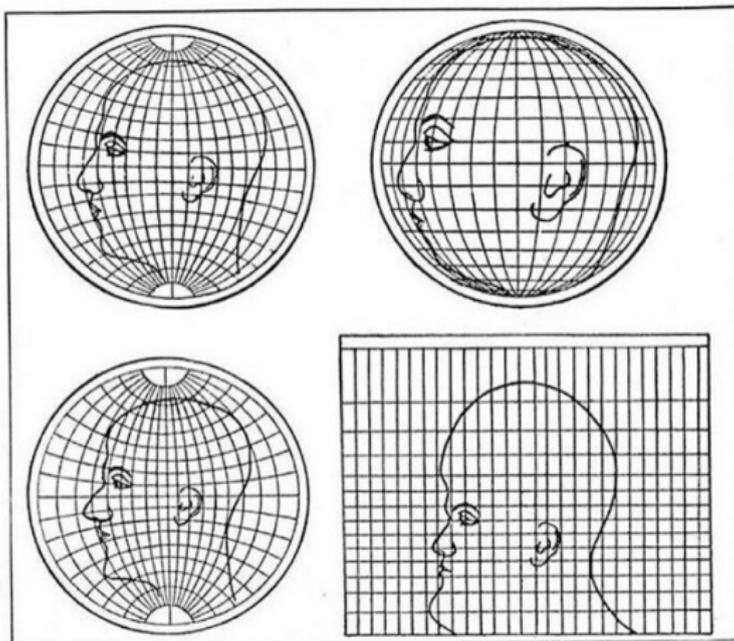
→ Are area and shape the relevant characteristics for population density?



The area of countries can represent their relative importance in terms of population

# And if you are not yet convinced...

... mind your head



*Upper left: Globular. Upper right: Orthographic. Lower left: Stereographic.  
Lower right: Mercator*

**What four commonly used projections do, as shown on a human head**

# Why geocoded datasets?

Because these data can give you new tools:

- to raise new questions
- to answer old questions in a novel and creative way

→ **Thanks to new measures**

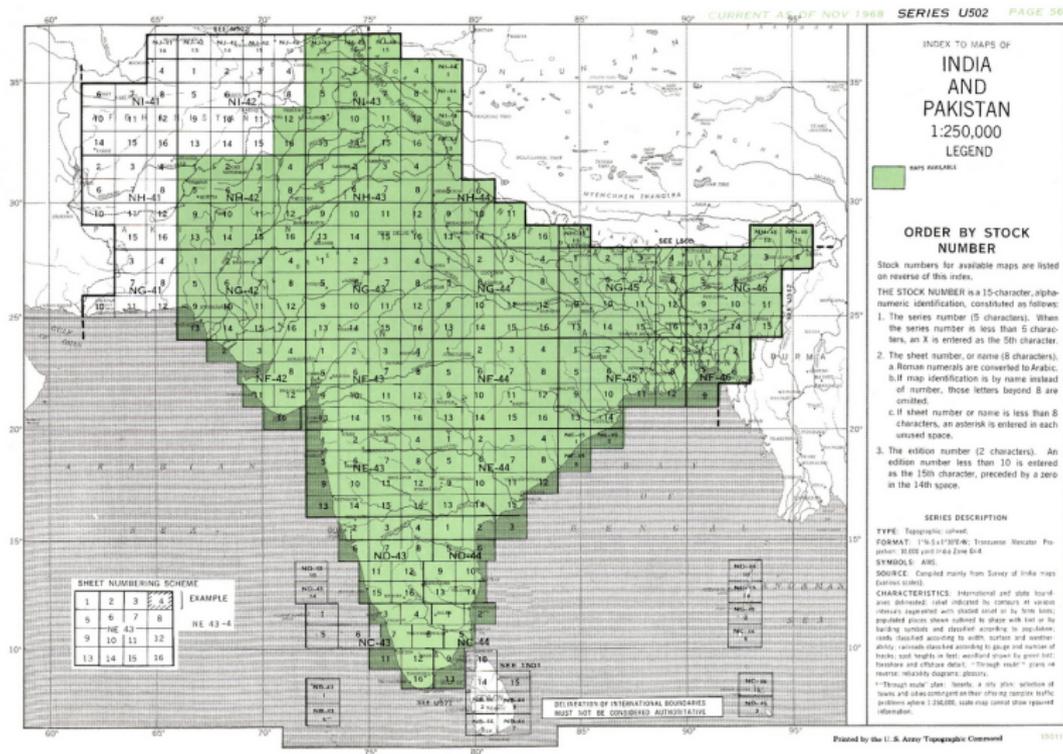
## Some data sources

- rainfall: TRMM [◀ URL](#)
- rainfall: global precipitation measurement [◀ URL](#)
- historical rainfall from University of Delaware [◀ U Delaware](#), or the University of East Anglia gridded data [◀ U East Anglia](#) or other sources [◀ GHCN](#)
- Leaf Area Index [◀ URL](#)
- NDVI [◀ URL](#)
- Land use data [◀ URL](#)
- access time to cities [◀ URL](#)
- Ethnographic atlas [◀ GREG](#) or [◀ Murdoch](#)
- altitude, ruggedness, walking time: STRM or ASTER [◀ URL](#)  
[◀ URL](#) or the most recent ALOS [◀ URL](#)

## Some data sources

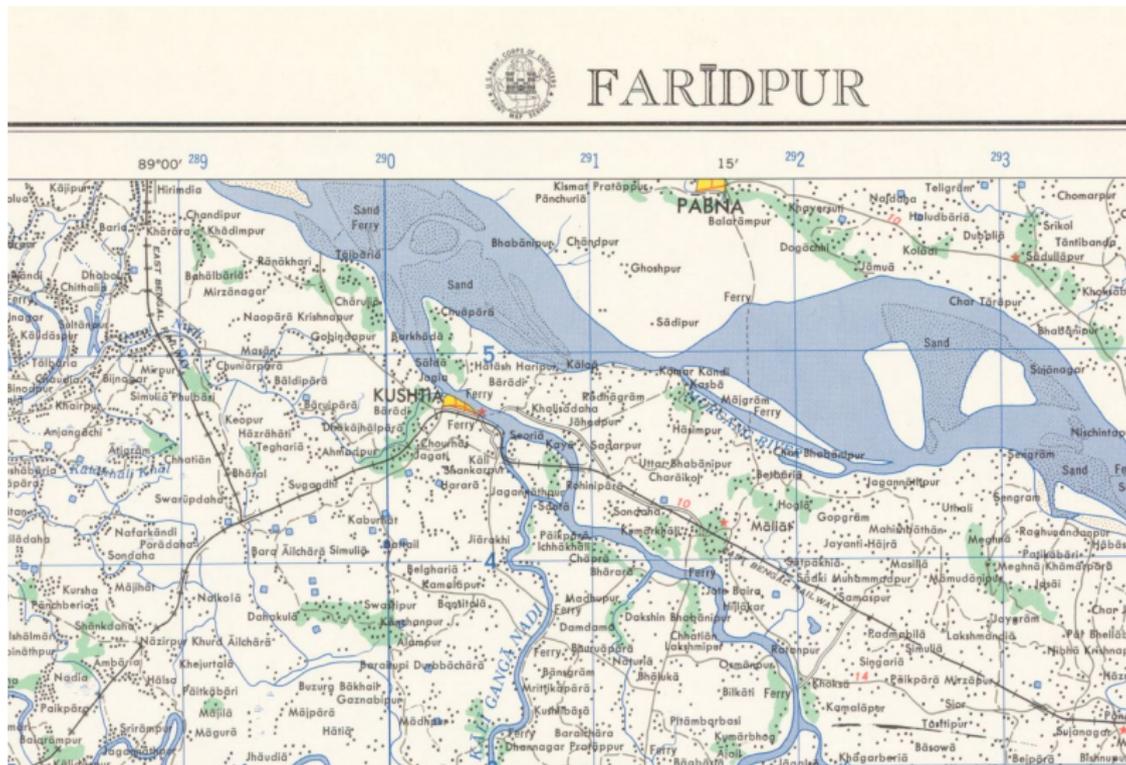
- georeferenced historical maps [◀ URL](#) [◀ URL](#)
- population density [◀ URL](#)
- Fires [◀ URL](#) [◀ URL](#)
- forest cover (Hansen et al., 2013, Science) [◀ URL](#) [◀ URL](#)
- nightlights [◀ URL](#)
- GAEZ [◀ URL](#)
- Global mapping of biomass [◀ URL](#)
- World database of protected areas [◀ URL](#)
- Geocoded aiddata [◀ URL](#)
- Nos villes vertes [◀ URL](#)
- Treepedia [◀ URL](#)
- Groundwater [◀ URL](#)
- ACLED conflict data [◀ URL](#)
- Global database of events [◀ URL](#)
- Mineral resources [◀ URL](#)

# Historical data





# Historical data



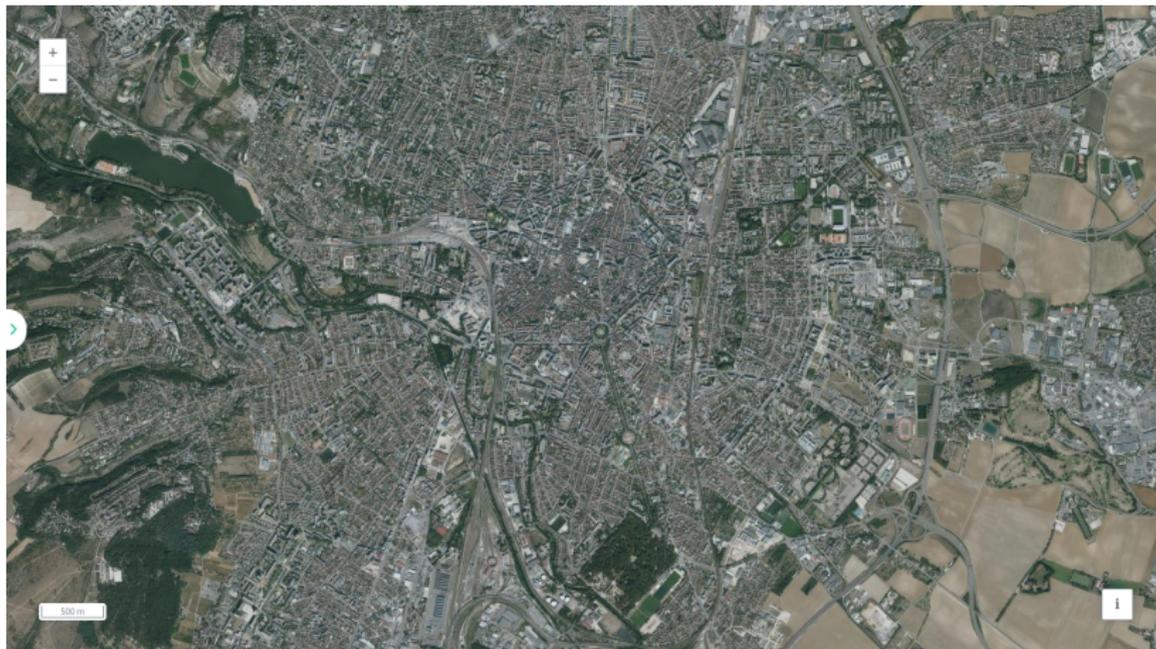
# Historical data



# Historical data

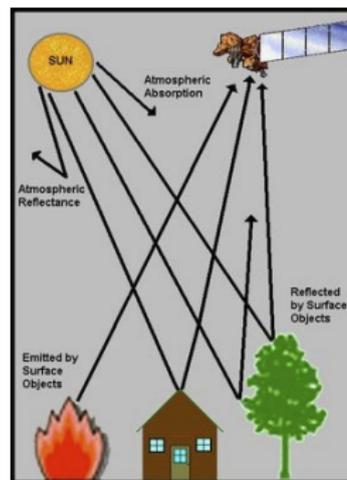
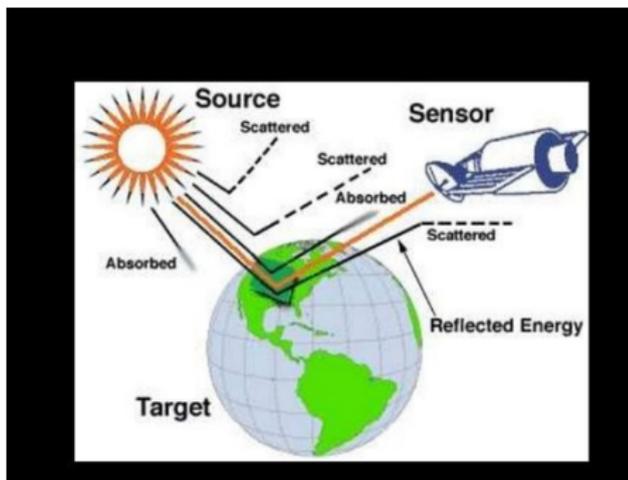


# Historical data



# Remote sensing information

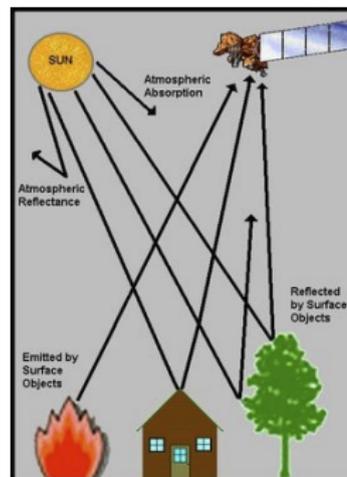
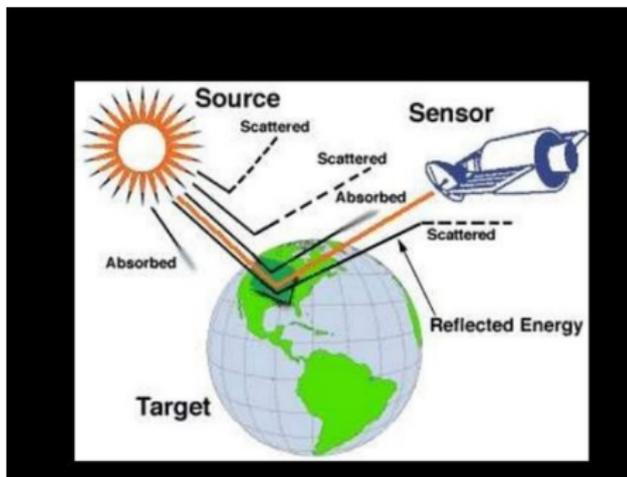
Passive sensors :



- use energy coming from another source of radiation
- very often the source of radiation is the sun and the radiation is reflected by the object of interest
- the object itself can be a source of radiation (ex.: fires → thermal remote sensing)

# Remote sensing information

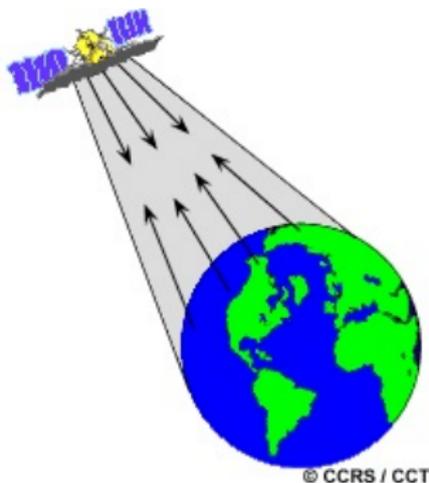
Passive sensors :



- passive sensors are very much dependent on the composition of the atmosphere (clouds, particulate matter, etc.)
- can record different frequency bands with varying precision (spectral resolution)

# Remote sensing information

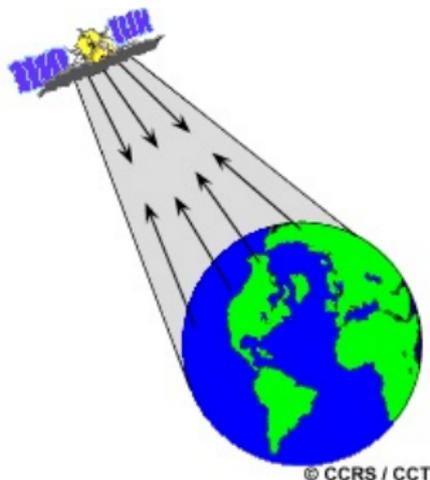
## Active sensors



- active sensors create their own electromagnetic energy
- are based on a pulse of energy sent by the sensor and backscattered by the object of interest
- active sensors are less affected by clouds and open new possibilities in terms of remote sensing

# Remote sensing information

## Active sensors



- allow to go below the canopy and measure trunks and branches
- allow to infer more properties of a surface
- noisy in hilly and mountainous area

## Remote sensing information

Information level (as defined by the NASA in 1986)

- 0 Reconstructed, unprocessed instrument and payload data at full resolution, with any and all communications artifacts (e. g., synchronization frames, communications headers, duplicate data) removed.
- 1 Reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters

## Remote sensing information

Information level (as defined by the NASA in 1986)

- 0 Reconstructed, unprocessed instrument and payload data at full resolution, with any and all communications artifacts (e. g., synchronization frames, communications headers, duplicate data) removed.
- 1 Reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters
- 2 Derived geophysical variables (e. g., ocean wave height, soil moisture, ice concentration) at the same resolution and location as Level 1 source data.
- 3 Variables mapped on uniform spacetime grid scales, usually with some completeness and consistency (e. g., missing points interpolated, complete regions mosaicked together from multiple orbits, etc.).

## Remote sensing information: using it

- Little sense for economists to use data of level 0 or 1
- Level 3 products have been produced/processed by specialists specially trained in physics, remote sensing, etc.

## Remote sensing information: using it

- Little sense for economists to use data of level 0 or 1
- Level 3 products have been produced/processed by specialists specially trained in physics, remote sensing, etc.
- Most product can be used, with “minor” processing, provided you understand what is measured and what is not.
- Good to discuss with someone trained in remote sensing

## Remote sensing information: using it

Pros and cons, from an economist's point of view

- quality is relatively constant over time and space
  - The best conditions for most remote sensing applications is an homogeneously covered, flat piece of land, on a cloud free day
  - Heterogeneity in land use, terrain slopes, and clouds tend to degrade the quality of data
  - Little effect of political borders and hard for one government to manipulate the information

## Remote sensing information: using it

Pros and cons, from an economist's point of view

- quality is relatively constant over time and space
  - The best conditions for most remote sensing applications is an homogeneously covered, flat piece of land, on a cloud free day
  - Heterogeneity in land use, terrain slopes, and clouds tend to degrade the quality of data
  - Little effect of political borders and hard for one government to manipulate the information
- data coverage gets worse, both in terms of frequency and of resolution as we go back in time
- data quality improves a lot over the years

## Remote sensing information: using it

Pros and cons, from an economist's point of view

- Trade-off between frequency and precision
  - remote sensors with high frequency tend to have lower precision. Very high resolution sensors do cover smaller parts of the Earth and therefore have smaller frequency over the same point
  - very high resolution images are costly, especially over large areas and for repeated time periods
  - The higher the resolution, the larger the memory and computational requirements.

## Remote sensing information: using it

Pros and cons, from an economist's point of view

- Trade-off between frequency and precision
  - remote sensors with high frequency tend to have lower precision. Very high resolution sensors do cover smaller parts of the Earth and therefore have smaller frequency over the same point
  - very high resolution images are costly, especially over large areas and for repeated time periods
  - The higher the resolution, the larger the memory and computational requirements.
  - There is no “one size fits all” solution
  - think twice about your question and what's the best way to answer it

*... In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography.*

purportedly from Suárez Miranda, Travels of Prudent Men,  
Book Four, Ch. XLV, Lérida, 1658

written in 1946 by Jose Luis Borges

## Remote sensing information: using it

- **very often, in applied economics, the constraints on temporal and spatial resolution do come from the socio-economic data**

## Remote sensing information: using it

- **very often, in applied economics, the constraints on temporal and spatial resolution do come from the socio-economic data**
- Economists are not the first ones to use remote sensing, its by-products, and associated applications
  - Many things that you want to do have already been programmed
  - Economists are not yet using the full power of these data → still scope for innovation
  - Do not be overly impressed by the technical difficulties described in some papers/talks

# Remote sensing information: using it CORRECTLY

- Some measures are meaningful when they are developed
- night-time light is correlated with urbanization, population density, GDP, poverty



Richer, more densely populated urban areas with high GDP and low poverty shine more during the night

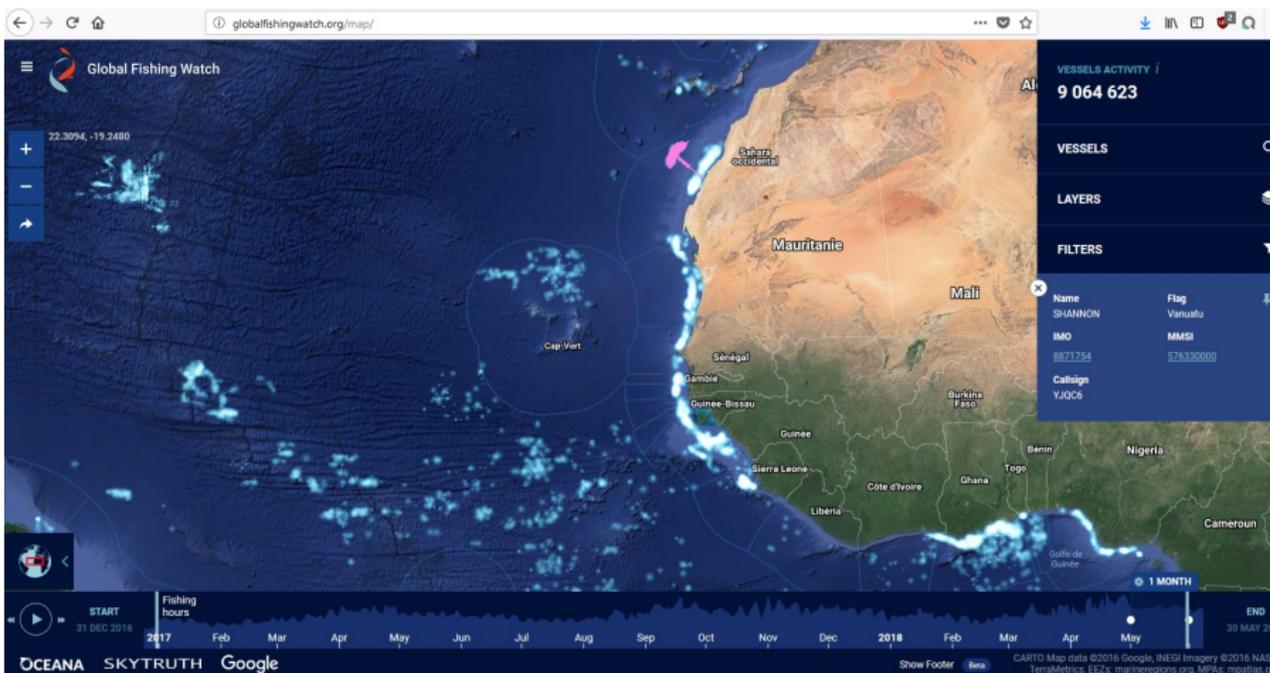




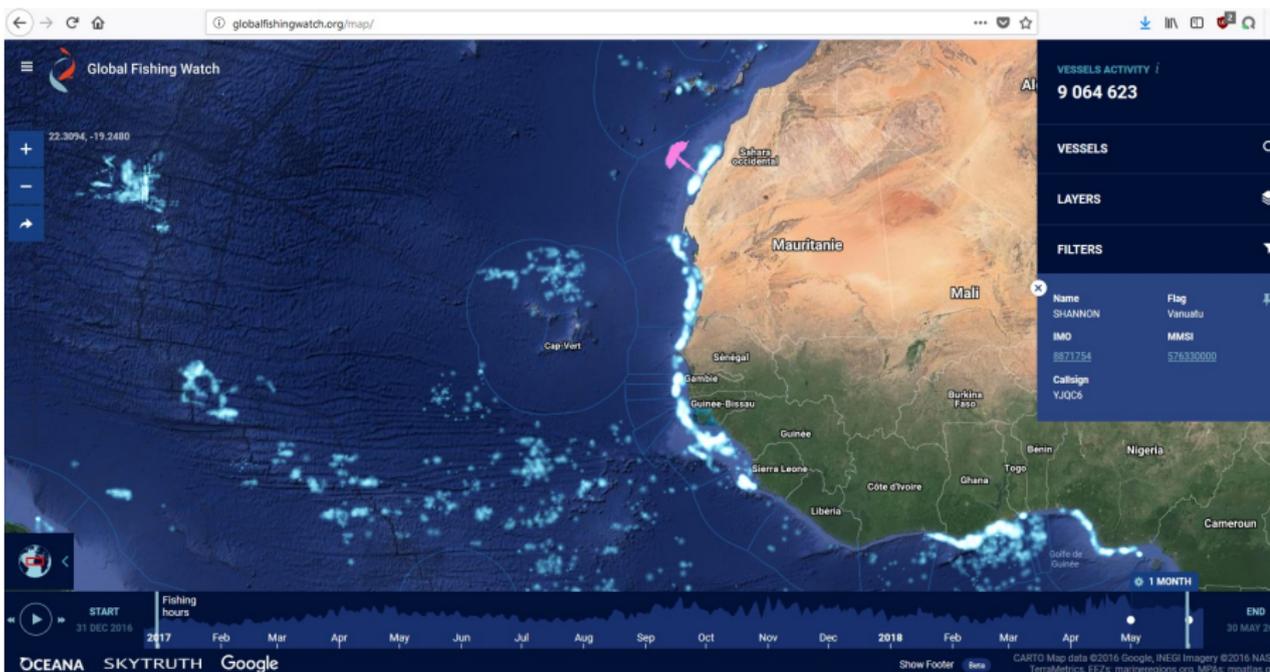
## Remote sensing information: using it CORRECTLY

- Measuring something is not neutral...
- ... especially if we draw policy recommendations, channel aid, rank countries based on these metrics
- The quality of a metric can evolve over the time

# Remote sensing information: using it CORRECTLY



# Remote sensing information: using it CORRECTLY



- Would this boat behave in the same way knowing that Global Fishing Watch tracks its behaviour?

# Tropical Rainfall Measurement Mission

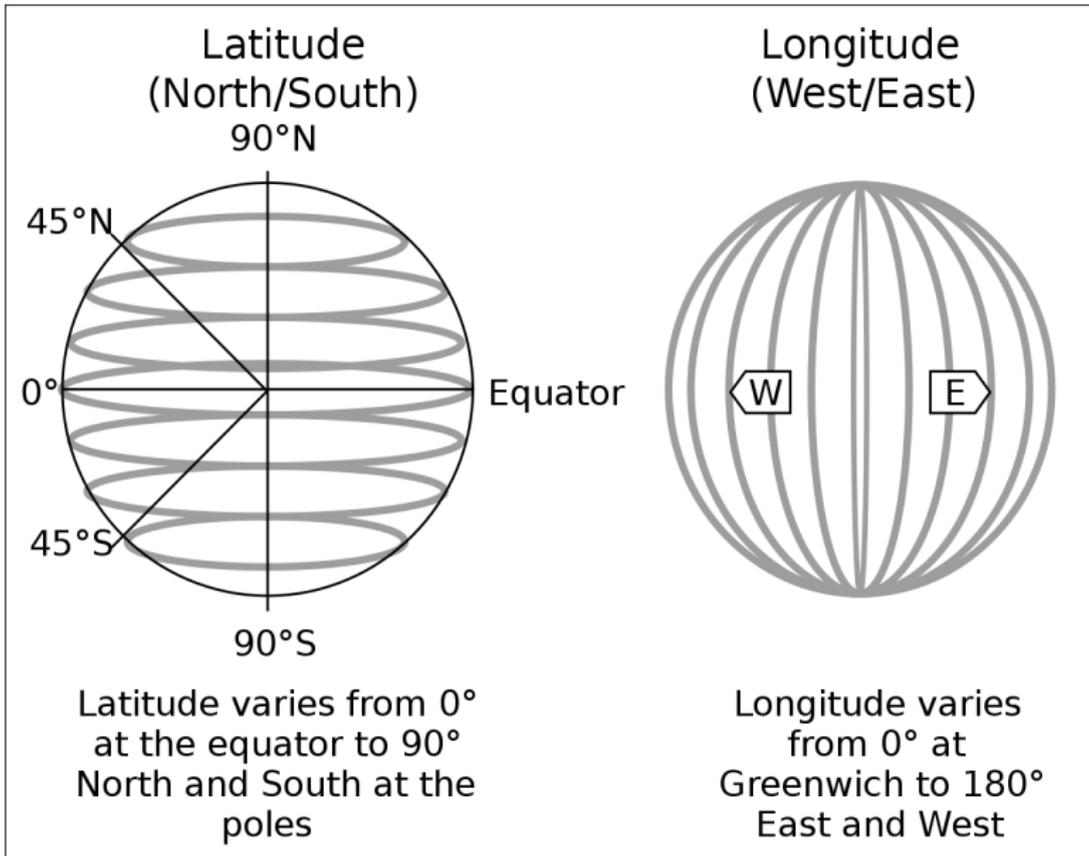
A photograph of a tropical landscape during a rainstorm. In the foreground, two palm trees with green fronds are leaning to the left. Behind them, a paved road curves through a lush, green forest. The background is heavily obscured by rain, creating a misty and grey atmosphere. The overall scene conveys a sense of a tropical environment experiencing heavy rainfall.

# Tropical Rainfall Measurement Mission

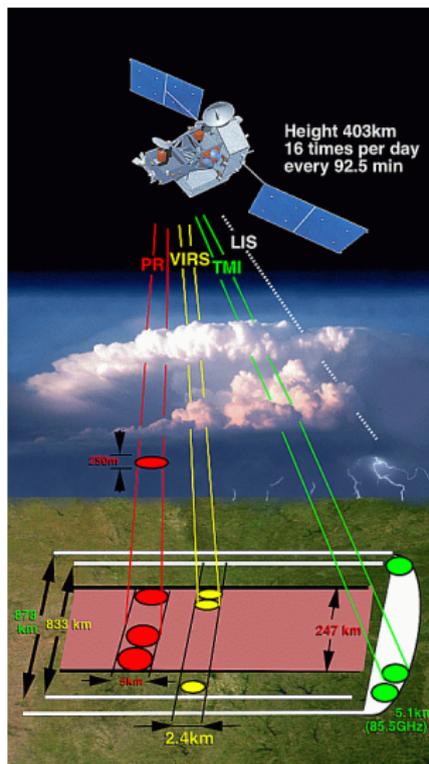
- Joint mission of NASA and JAXA operating between December 1997 and April 2015
- goal: to study the climate
- several on-board instruments, including precipitation radar and lightning imaging sensor

# Tropical Rainfall Measurement Mission

- Joint mission of NASA and JAXA operating between December 1997 and April 2015
- goal: to study the climate
- several on-board instruments, including precipitation radar and lightning imaging sensor
- coverage: between  $35^{\circ}N$  and  $35^{\circ}S$  but rainfall data available between  $50^{\circ}N$  and  $50^{\circ}S$
- circular orbit at 350km and then 403km
- period : 96 minutes
- spatial resolution : 2.4 to 5.1 km
- resolution of rainfall data:  $0.25^{\circ} \times 0.25^{\circ}$  grid. About 25km at the equator.



# Tropical Rainfall Measurement Mission



# Tropical Rainfall Measurement Mission

- Monthly data on precipitation rate
- Best-estimate precipitation rate and root-mean-square (RMS) precipitation-error estimates
- Combines 3-hourly merged high-quality/IR estimates with monthly data the monthly accumulated Global Precipitation Climatology Centre (GPCC) rain gauge analysis
- The 3-hourly merged high quality/IR estimates are summed for the calendar month, and then the rain gauge data are used to apply a large-scale bias adjustment to the multi-satellite estimates, almost exclusively over land.

More details: see *Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F. Stocker, D.B. Wolff, 2007: The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. J. Hydrometeor., 8 (1), 38-55.*

# Tropical Rainfall Measurement Mission

Main issues with the data (from an economist perspective):

- The resolution might be too coarse for some applications

*Madesdam et al., 2013, Do political protests matter: evidence from the tea party movement, Quarterly Journal of Economics*

- They exploit the variation in rainfall on April 15, 2009 to instrument protest intensity against taxes
- They find that good weather at this initial, coordinating event had significant consequences for the subsequent local strength of the movement, increased public support for Tea Party positions, and led to more Republican votes in the 2010 midterm elections
- They also look at voting behavior in the Congress and find a more conservative voting pattern of incumbents

# Tropical Rainfall Measurement Mission

- Only useful for countries “sufficiently close to the equator”
- Short time span
  - hard to address “climate change” in a proper way
  - for historical data, use the University of Delaware gridded data [◀ U Delaware](#), the University of East Anglia gridded data [◀ U East Anglia](#) or other sources [◀ GHCN](#)
  - use Global Precipitation Measurement for more recent data [◀ GPM](#)

# Social Insurance and Conflict: Evidence from India

Thiemo Fetzer (mimeo)

- Show that NREGA program in India reduces violence when a negative shock occurs
- Data at the district level
- Main variables:
  - conflict intensity at the district level computed using text analysis in 28,638 newspaper clippings collected by the South Asia Terrorism Portal (SATP)
  - NREGA participation data derived from the so-called Monthly Progress Reports (MPR) from before 2011 and from the Management Information System (MIS) from 2011 onwards
  - agricultural production and wages
  - rainfall data (using TRMM)

# Social Insurance and Conflict: Evidence from India

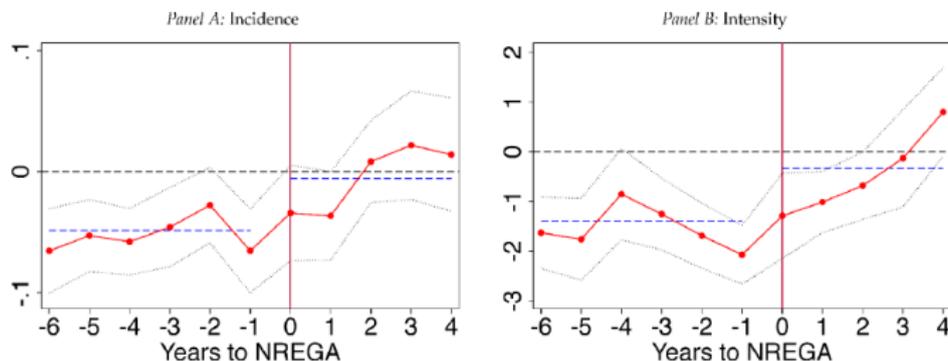
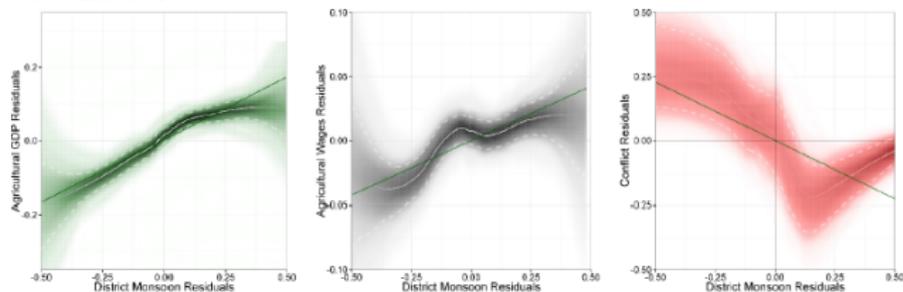


Figure 2: Effect of Monsoon Rain on Conflict Over Time. The vertical line indicates the NREGA introduction date. The blue dashed lines indicate the coefficients obtained from a simple regression interaction lagged Monsoon rainfall with the NREGA treatment indicator. The red line are each point estimates of the relationship between lagged Monsoon rainfall and conflict. 95% confidence bands are indicated as dotted black lines.

# Social Insurance and Conflict: Evidence from India

Panel A: Before NREGA



Panel B: After NREGA

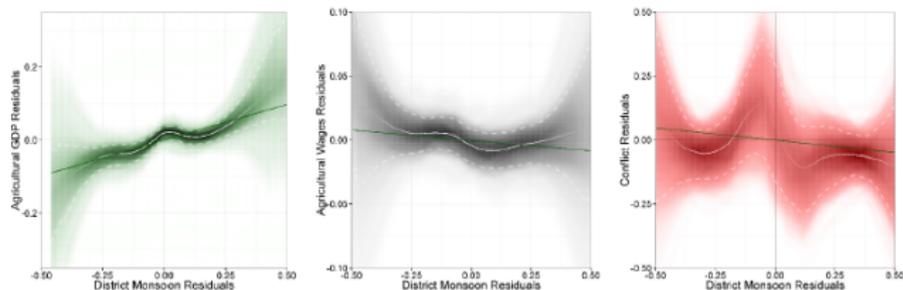


Figure 3: Non-Parametric Watercolor Regressions as in Hsiang et al. (2013): Effect of Monsoon Rain on agricultural output per Capita, Wages and Conflict Before and After Introduction of NREGA. 95% confidence bands are indicated as dashed lines. The color shading is related to the overall density of Monsoon rainfall realizations along the horizontal axis and to the density of fitted values from loess regressions along the vertical axis.

# Managing the Impact of Climate Change on Migration: Evidence from Mexico

by Isabelle Chort and Maëlys de la Rupelle

- lower than average precipitations increase undocumented migration, especially from the most agricultural states
- disaster funds (Fonden) are found to mitigate the effect of climate shocks on migration by lowering the undocumented migration response to abnormally low precipitations during the dry season
- An increase in the state-level share of an agricultural cash-transfer program (PROCAMPO) to non-irrigated plots in the ejido sector decreases documented migration after a similar shock

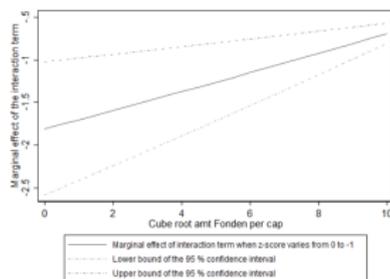
Main data: (at the state level)

- migration flows to the US by state of origin and destination, including illegal migration
- impact of hurricanes (Historical Hurricane Track tool developed by the U.S. National Oceanic and Atmospheric Administration (NOAA))
- TRMM data
- State level data on PROCAMPO and Fonden

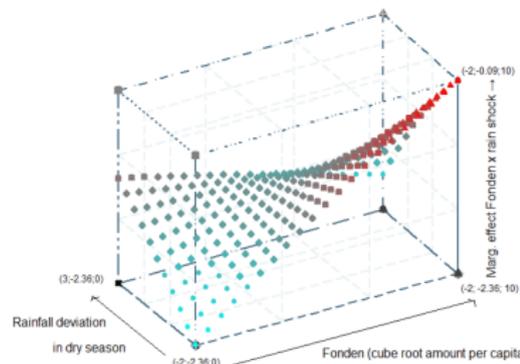
# TRMM in economics

Figure 1: Marginal effect of the interaction between Fonden (cube root amount per capita) and a negative variation in the z-score of dry season rainfall on the migration rate (per 10000 population).

(a) when the initial value of the rainfall z-score is zero.



(b) For different initial values of the rainfall z-score.

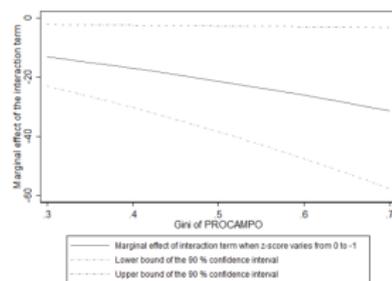


Coordinates in parentheses : (z-score; marginal effect of the interaction z-score X Fonden amount on the migration rate; Fonden cube root amount per capita).

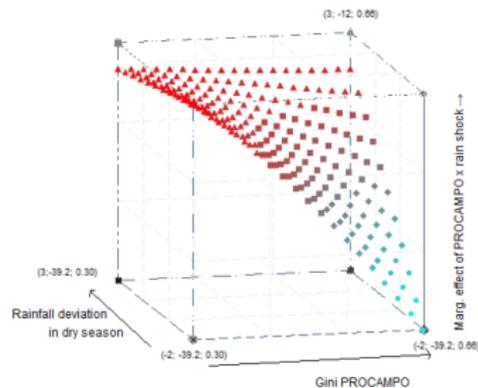
# TRMM in economics

Figure 2: Marginal effect of the interaction between PROCAMPO Gini and a negative variation in the z-score of dry season rainfall on the migration rate (per 10000 population).

(a) when the initial value of the rainfall z-score is zero.



(b) For different initial values of the rainfall z-score.

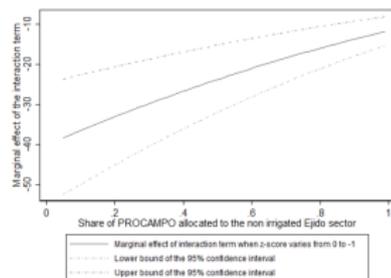


Coordinates in parentheses : (z-score; marginal effect of the interaction z-score X PROCAMPO gini on the migration rate; PROCAMPO gini).

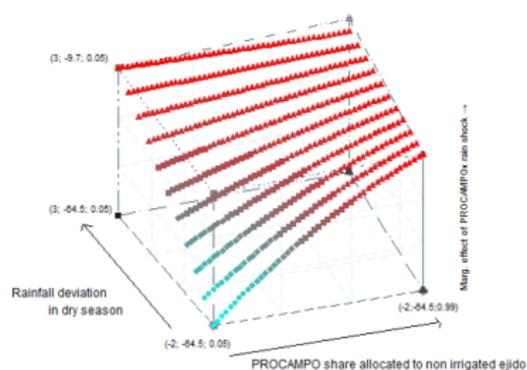
# TRMM in economics

Figure 3: Marginal effect of the interaction between the PROCAMPO share allocated to the non irrigated *ejido* sector and a negative variation in the z-score of dry season rainfall on the migration rate (per 10000 population).

(a) when the initial value of the rainfall z-score is zero.



(b) For different initial values of the rainfall z-score.



Coordinates in parentheses : (z-score; marginal effect of the interaction z-score X PROCAMPO share going to non irrigated *ejidos*; PROCAMPO share going to non irrigated *ejidos*).

## Rain as a universal instrument?

**Weather has been widely used as an instrumental variable in social science to predict many different variables.**

Instrumental variable (IV) analysis assumes the instrument only affects the dependent variable via its relationship with the independent variable. Other possible causal routes from the IV to the dependent variable are exclusion-restriction violations and invalidate the instrument.

The use of weather to instrument different independent variables represents strong *prima facie* evidence of exclusion violations for all studies using weather IVs.

Jonathan Mellon (2022) *Rain, Rain, Go Away: 192 Potential Exclusion-Restriction Violations for Studies Using Weather as an Instrumental Variable* (among 288 studies)

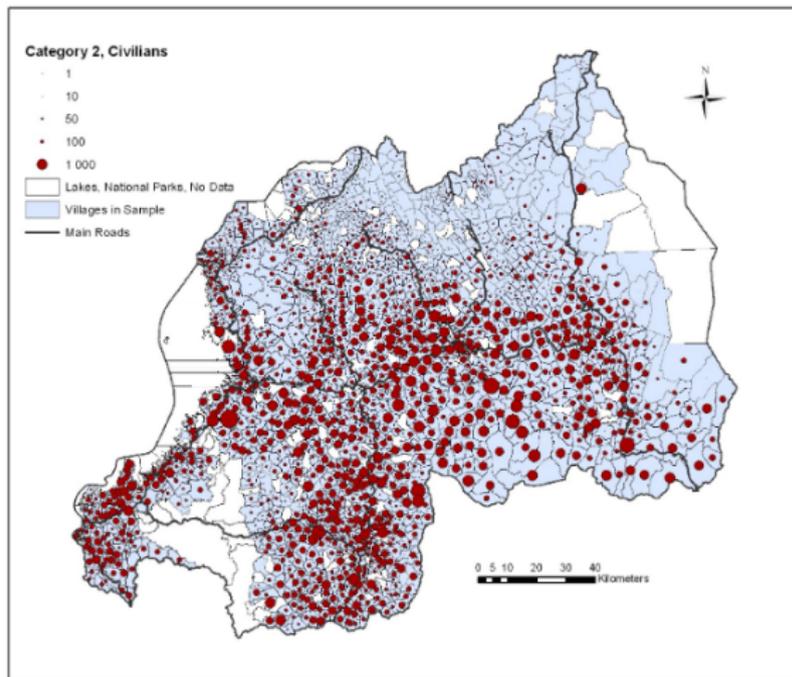
# Mobilizing the Masses for Genocide

Thorsten Rogall 2021, American Economic Review, 111 (1): 41-72.

- Do political elites use armed groups to foster civilian participation in genocidal violence?
- Are armed groups employed strategically? How do they mobilize civilians?
- Investigated in the context of the Rwandan Genocide
- Exploit exogenous variation in armed groups' transport costs induced by weather fluctuations: the shortest distance of each village to the main road interacted with rainfall along the dirt tracks between main road and village
- 1 additional armed-group member resulted in 7.3 more civilian perpetrators
- armed-group leaders employed their men strategically and invoked civilians' obedience.

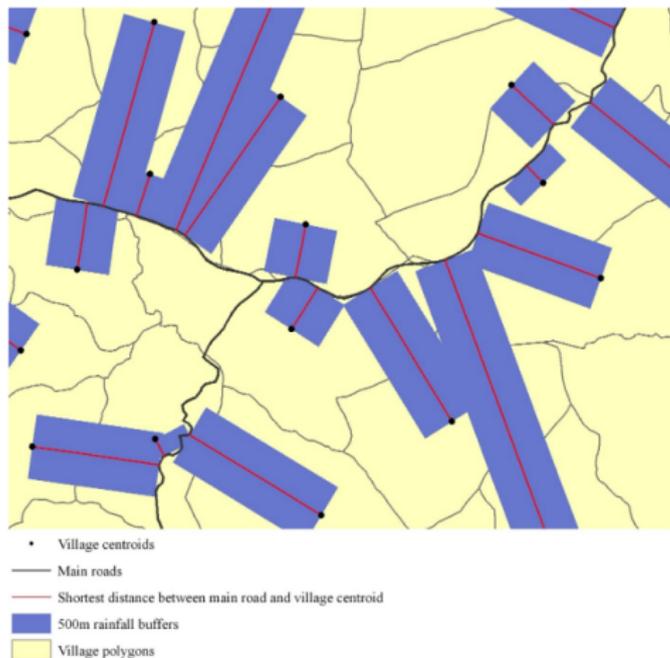
# Mobilizing the Masses for Genocide

**Figure A.3: Civilian Violence (# Prosecutions)**



# Mobilizing the Masses for Genocide

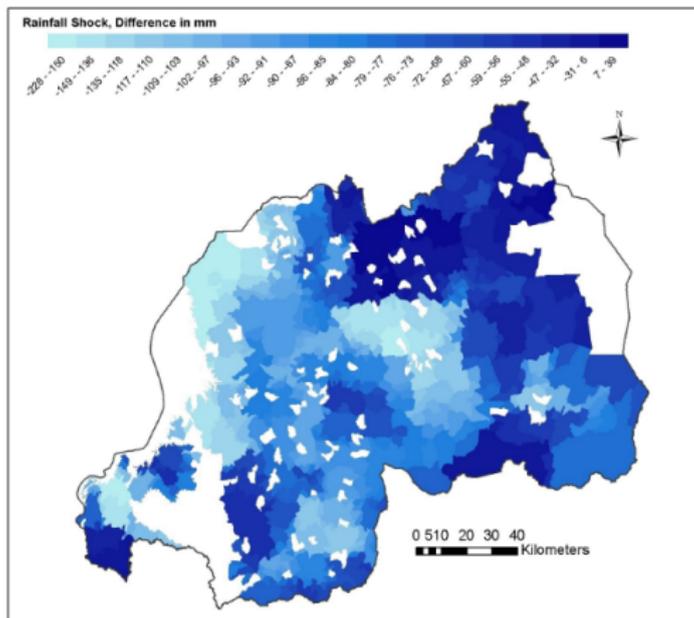
Figure A.6: Construction of the Instrument in ArcGIS



Instrument: Interaction of the length of the red line and amount of rain falling on the area of the blue rectangle during the period of the genocide.

# Mobilizing the Masses for Genocide

Figure A.8: Rainfall



Notes: This map shows rainfall along the way between main road and village during the period of the genocide in 1994 for each village, subtracting rainfall between main road and village during the 100 calendar days of the genocide of an average year (years 1984-1993). White areas are either national parks, Lake Kivu, or villages not in the sample.

# Mobilizing the Masses for Genocide

TABLE 1—SUMMARY STATISTICS (DATA AT THE VILLAGE LEVEL)\*

	Mean	SD
<i>Panel A. Endogenous variables</i>		
Number of prosecuted militiamen	51.76	70.51
Number of prosecuted civilians	290.25	286.43
<i>Panel B. Exogenous variables</i>		
Rainfall between village and main road, genocide period, 1994	122.70	35.94
Rainfall between village and main road, genocide period, 10-year average	206.18	37.78
Rainfall between village and main road, whole year, 10-year average	962.76	180.15
Rainfall between village and main road, harvest season, 1994	22.42	10.15
Rainfall between village and main city, genocide period, 1994	123.86	34.00
Rainfall at closest point on main road, genocide period, 1994	122.47	35.08
Rainfall in village, genocide period, 1994	122.68	35.62
Rainfall in village, genocide period, 10-year average	204.99	38.86
Rainfall in village, growing season, 1994	243.89	69.61
Rainfall in village, growing season, 10-year average	621.09	117.51
Rainfall in village, whole year, 10-year average	960.68	182.70
Distance to the main road	6.71	5.77
Distance to Kigali	62.65	30.00
Distance to Nyanza	64.36	30.74
Distance to main city	22.78	14.69
Distance to the border	22.60	13.93
1991 population, '000	4.88	2.48
1991 population density	494.71	850.80
Number of days with RPF presence	42.47	43.12
Tutsi rebels (RPF)	0.05	0.23
Mass grave in village	0.05	0.21
Village under pre-colonial kingdom before 1770	0.39	0.49
Average Saturday village rainfall, 1990–1994	2.43	0.60
RTL M radio coverage in village, share with reception	0.18	0.23
Inter-ethnic family measure	0.01	0.01
Tutsi minority share	0.10	0.13

# Mobilizing the Masses for Genocide

## First stage:

$$\log(\text{Militia Size}_i) = \alpha + \beta(\log(\text{Distance To Main Road}_i) \times \log(\text{Rain}_i)) + \gamma X_i + \epsilon_i$$

## Second stage:

$$\log(\text{Civilian perpetrators}_i) = \alpha + \beta(\log(\text{MilitiaSize}_i) \times \log(\text{Rain}_i)) + \gamma X_i + \epsilon_i$$

# Mobilizing the Masses for Genocide

TABLE 3—FIRST STAGE, REDUCED FORM, AND MAIN EFFECTS

<i>Panel A</i>	Militiamen, log			Civilian perpetrators, log		
	First stage			Reduced form		
	(1)	(2)	(3)	(4)	(5)	(6)
Distance × rainfall along buffer, 1994	-0.357 (0.116)	-0.460 (0.117)	-0.509 (0.115)	-0.480 (0.126)	-0.573 (0.125)	-0.661 (0.141)
Standard controls	Yes	Yes	Yes	Yes	Yes	Yes
Growing season controls	No	Yes	Yes	No	Yes	Yes
Additional controls	No	No	Yes	No	No	Yes
Province effects	Yes	Yes	Yes	Yes	Yes	Yes
<i>F</i> -statistic	9.50	15.54	19.54	14.45	20.93	21.91
<i>R</i> <sup>2</sup>	0.46	0.48	0.50	0.53	0.54	0.58
Observations	1,433	1,433	1,432	1,433	1,433	1,432

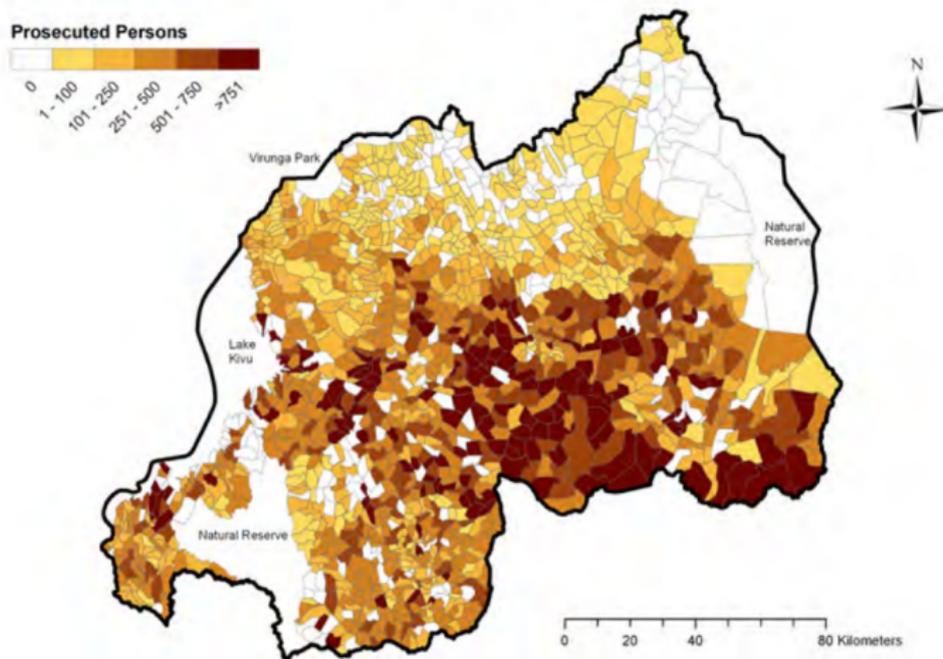
  

<i>Panel B</i>	Civilian perpetrators, log					
	OLS			IV/2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Militiamen, log	0.649 (0.065)	0.647 (0.066)	0.626 (0.051)	1.345 (0.369)	1.245 (0.241)	1.299 (0.259)
Standard controls	Yes	Yes	Yes	Yes	Yes	Yes
Growing season controls	No	Yes	Yes	No	Yes	Yes
Additional controls	No	No	Yes	No	No	Yes
Province effects	Yes	Yes	Yes	Yes	Yes	Yes
<i>R</i> <sup>2</sup>	0.72	0.72	0.74	—	—	—
Observations	1,433	1,433	1,432	1,433	1,433	1,432

# Propaganda and Conflict: Evidence from the Rwandan Genocide

David Yanagizawa-Drott (2014)

## Total Violence



# Propaganda and Conflict: Evidence from the Rwandan Genocide

David Yanagizawa-Drott (2014)

## Radio Coverage, RTLM

Share of village with reception



0

0.01 - 0.05

0.05 - 0.1

0.1 - 0.2

0.2 - 0.3

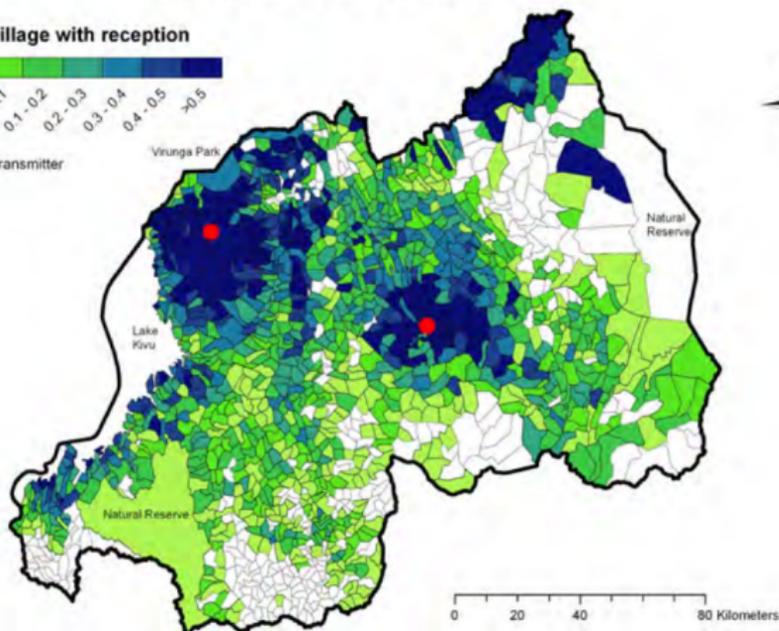
0.3 - 0.4

0.4 - 0.5

>0.5

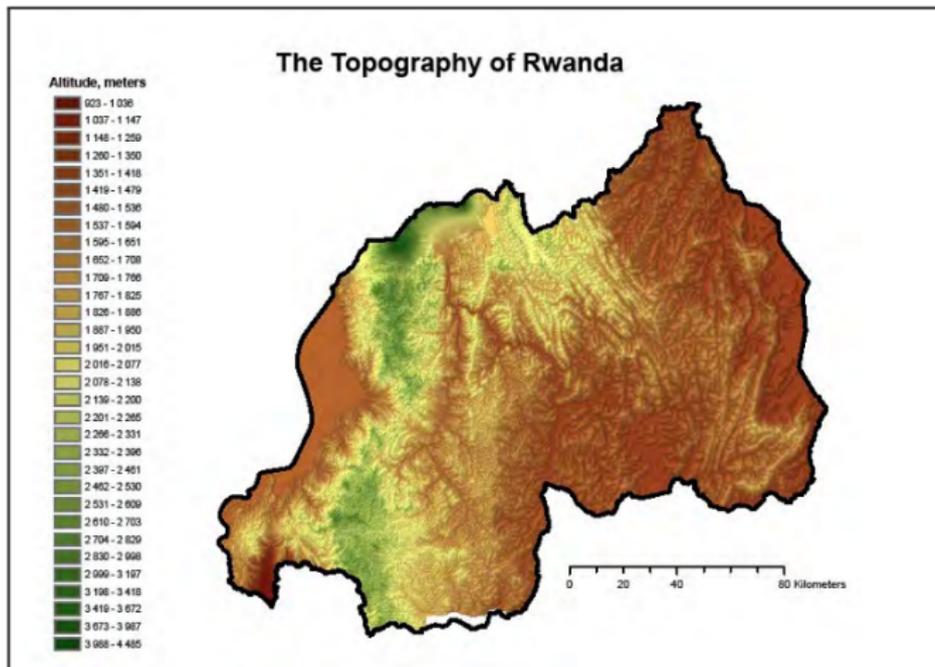


RTLM transmitter



# Propaganda and Conflict: Evidence from the Rwandan Genocide

David Yanagizawa-Drott (2014)

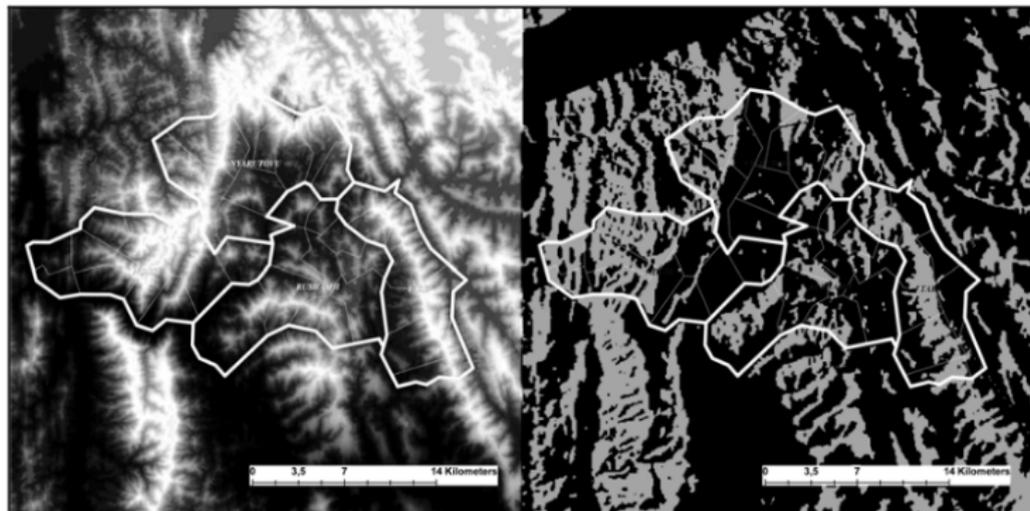


**Appendix Figure 1. The Topography of Rwanda**

Source: Shuttle Radar Topography Mission

# Propaganda and Conflict: Evidence from the Rwandan Genocide

David Yanagizawa-Drott (2014)



**Appendix Figure 2. Predicted Radio Coverage, 4 communes example**

This left picture shows the height of ground, where brighter marks higher altitude. The right picture shows the empirical radio coverage, where grey marks radio coverage. The signal comes from the Mount Muhe transmitter located 30 km to the west (outside the figure). The figures show that within each commune (boundaries in thick white lines), villages (boundaries in thin white lines) to the east of hilltops have low radio coverage due the hilltops in the line-of-sight to the transmitter. Source: SRTM 90m topography data, author's calculations of radio coverage in ArcGIS.

# Propaganda and Conflict: Evidence from the Rwandan Genocide

David Yanagizawa-Drott (2014)

**Table 2. Exogeneity Check**

	Population in 1991, log (1)	Population Density in 1991, log (2)	Distance to Major Town, log (3)	Distance to Major Road, log (4)	Distance to the Border, log (5)	North Sloping (6)	East Sloping (7)	South Sloping (8)	West Sloping (9)	Radio Coverage in Nearby Villages (<10 km) (10)	Radio Coverage in Nearby Villages (10-20 km) (11)
Radio Coverage in Village	-0.049 (0.071)	0.196 (0.145)	0.092 (0.086)	-0.238 (0.154)	0.082 (0.189)	0.113 (0.087)	-0.008 (0.099)	0.020 (0.089)	-0.125 (0.109)	0.029 (0.018)	-0.009 (0.020)
Observations	1065	1065	1065	1065	1065	1065	1065	1065	1065	1065	1065
R-squared	0.460	0.426	0.908	0.705	0.921	0.150	0.138	0.145	0.162	0.957	0.952
Propagation Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Commune FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: The radio propagation controls are: latitude, longitude, and second-order polynomials in village mean altitude, village altitude variance, and distance to the nearest RTLM transmitter. Standard errors in parentheses, adjusted for spatial correlation (Conley, 1999). Significance levels at \*10%, \*\*5%, \*\*\*1%.

# Propaganda and Conflict: Evidence from the Rwandan Genocide

David Yanagizawa-Drott (2014)

**Table 3. Main Effects**

	Dependent Variable: Log(Prosecuted Persons)								
	Total Violence			Collective Violence			Individual Violence		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Radio Coverage in Village	0.507** (0.226)	0.526** (0.242)	0.484** (0.235)	0.582** (0.239)	0.559*** (0.216)	0.544*** (0.206)	0.450* (0.233)	0.465* (0.252)	0.418* (0.246)
Population in 1991, log			0.590*** (0.131)			0.589*** (0.171)			0.624*** (0.150)
Population Density in 1991, log			-0.014 (0.070)			0.004 (0.101)			-0.015 (0.069)
Distance to Major Town, log			0.068 (0.150)			-0.233 (0.149)			0.113 (0.152)
Distance to Major Road, log			-0.196** (0.076)			-0.245*** (0.090)			-0.193** (0.075)
Distance to the Border, log			0.171* (0.103)			0.030 (0.126)			0.186* (0.103)
East Sloping, dummy			0.017 (0.070)			0.098 (0.092)			0.014 (0.084)
North Sloping, dummy			0.065 (0.068)			0.041 (0.092)			0.079 (0.068)
South Sloping, dummy			-0.013 (0.074)			-0.028 (0.101)			-0.012 (0.077)
Observations	1065	1065	1065	1065	1065	1065	1065	1065	1065
R-squared	0.63	0.64	0.66	0.52	0.53	0.55	0.62	0.63	0.65
Commune FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Propagation Controls	N	Y	Y	N	Y	Y	N	Y	Y

Notes: *Collective Violence* is crime category 1 prosecutions against organizers, leaders, army and militia; *Individual Violence* is crime category 2 prosecutions for homicides, attempted homicides and serious violence. *Total Violence* is the sum of collective and individual violence. *Radio Coverage* is the share of the village area that has RTLTM reception. The radio propagation controls are: latitude, longitude, and second-order polynomials in village mean altitude, village altitude variance, distance to the nearest RTLTM transmitter. Standard errors in parentheses, adjusted for spatial correlation (Conley, 1999). Significance levels at \*10%, \*\*5%, \*\*\*1%.

## Comparing the two results

- Yanagizawa-Drott (2014) finds that jamming the RTLH hate radio station during the genocide would have cut the number of deaths by about 10 percent.
- Thorsten Rogall (2021) finds that stopping the army and militia would have resulted in about 9 times that effect.
- Radio helped the local militia to coordinate their actions.

## Introduction to Qgis

This example is based on real work where we instrument Internet availability through physical disruptions of 3G antennas in Israel and Palestine. We aim to study the effect social media posting in Israel and Palestine on traditional media reporting of the conflict.

*Hatte, Sophie and Madinier, Etienne and Zhuravskaya, Ekaterina, Reading Twitter in the Newsroom: How Social Media Affects Traditional-Media Reporting of Conflicts (May 2021). CEPR Discussion Paper No. DP16167 Available at SSRN: <https://ssrn.com/abstract=3886588>*

## Relevance of the internet outage: Gaza Prayer Bot

- Islamic Prayer Bot from Gaza (@IPT\_Gaza) is an automatic service informing Muslims of prayer times
- Tweet example: “5:07 Now #Fajr Prayer Time for #Gaza #Palestin”

Dependent Variable:	Prayer bot did not tweet (when scheduled)					
	(1)	(2)	(3)	(4)	(5)	(6)
Lightning	0.051 <sup>a</sup> (0.017)	0.042 <sup>b</sup> (0.020)				
Paltel no visible traffic			0.059 <sup>b</sup> (0.025)	0.116 <sup>a</sup> (0.027)		
Internet Outage					0.076 <sup>a</sup> (0.014)	0.063 <sup>a</sup> (0.016)
Year, MoY, DoW, HoD FEs		✓		✓		✓
Rain and wind controls		✓				✓
Observations	7841	7841	5365	5365	7841	7841

⇒ A unit of analysis: date × hour of prayer; Dep. var.: @IPT\_Gaza did not tweet

⇒ The instrument correlates with the likelihood of prayer bot going silent

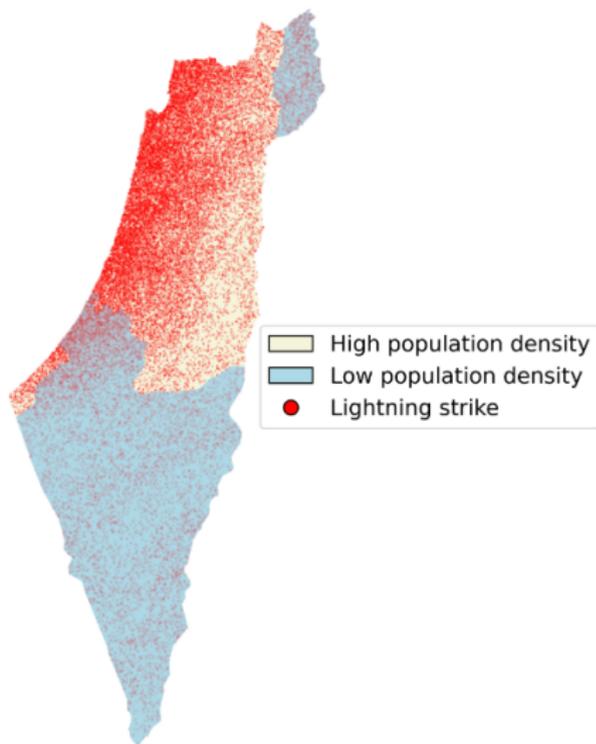
## Combining basic operations & zonal statistics

- Load the data ( Layer → Add Layer):
  - Administrative data are vector layers
  - Population density is a raster layer
  - Lightnings are stored according to their latitude and longitude in a csv table (delimited text)
- Merge (union) the four shapefiles into two separate shapefiles, according to their administrative level
  - Processing toolbox → Union
- Clip the raster data on population density
  - Processing toolbox → Clip raster by mask layer
- Colorize raster data
  - Right click on the layer → Properties → Symbology

## Combining basic operations & zonal statistics

- Sample and colorize lightning data
  - Processing toolbox → Sample raster
  - Right click on the layer → Properties → Symbology
- Produce level 1 zonal statistics: sum of population density
  - Processing toolbox → zonal statistics
  - Right click on the layer → Properties → Symbology
- Produce level 1 spatial join and export
  - Processing toolbox → Join attributes by location
  - Right click on the layer → Export → Save features as

Using what you have seen and the print layout (Project → New print layout), try to replicate this map :



- I could answer 95% of your technical questions on Qgis by  
*“Let me google that for you”*

- I could answer 95% of your technical questions on Qgis by  
*“Let me google that for you”*
- Other 4% by  
*“I did not properly understand what you wanted to do. Maybe you could rephrase. And then let me google that for you”*

- I could answer 95% of your technical questions on Qgis by  
*“Let me google that for you”*
- Other 4% by  
*“I did not properly understand what you wanted to do. Maybe you could rephrase. And then let me google that for you”*
- And the remaining 1% will be  
*“I have no idea”*

- I could answer 95% of your technical questions on Qgis by  
*“Let me google that for you”*
- Other 4% by  
*“I did not properly understand what you wanted to do. Maybe you could rephrase. And then let me google that for you”*
- And the remaining 1% will be  
*“I have no idea”*
- ⇒ This is actually good news

Most useful websites:

- Qgis Documentation: <https://qgis.org/en/docs/index.html>
- GIS Stackexchange: <https://gis.stackexchange.com/>
- You preferred research engine

## Next session

- Vector vs. rasters
- Projections systems
- Measurement errors
- Dataset guidelines
- Coding in ???
- Group project