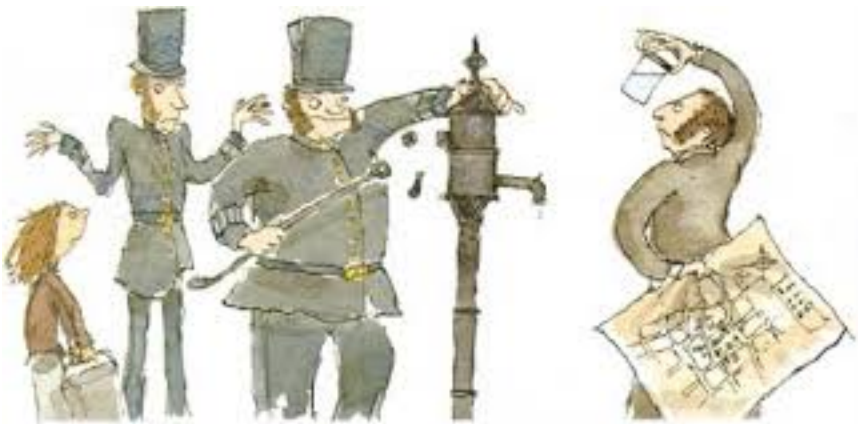
Intro to Spatial Analytics
and
Remote Sensing applications for urban and spatial economics

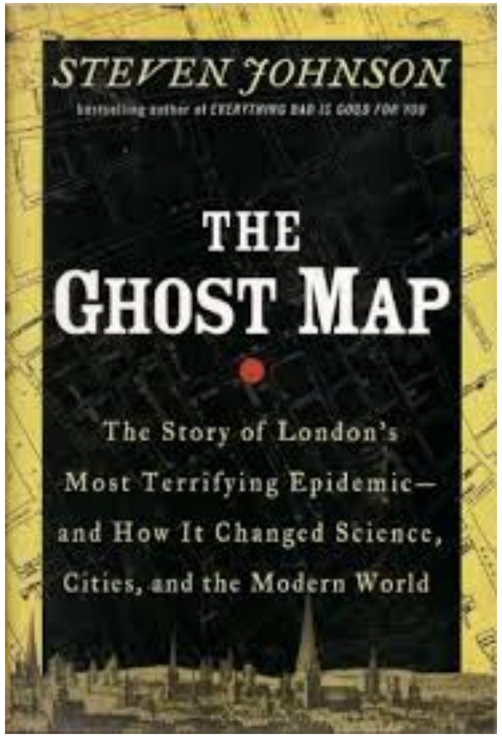
Esteban López Ochoa
Profesor Asistente, Escuela de Negocios
Director ejecutivo - Centro de Economía y Política Regional (CEPR)
Director académico – Diplomado en Big Data para Políticas Públicas

Part 1: Intro to Spatial Analytics

Who is John Snow?

Who is John Snow?





First Law of Geography

"everything is
related to
everything else,
but near things
are more related
than distant
things."

Waldo Tobbler

Motivation

- Substantive
  - From Adam Smith invisible hand to social networks
    - Individual vs. socio-spatial interaction
    - Peer effects, contagion, imitation, trends
  - Spatial externalities
    - Costs/Benefits from activities that impact in a different location
    - Spatial spillovers:
      - Good: neighbor playing the piano
      - Bad: neighbor with COVID-19 and not quarantining
    - Spatial multipliers
      - Investment in a park → housing prices
  - Spatial Mismatch/Disparities
  - Spatial Context

Motivation

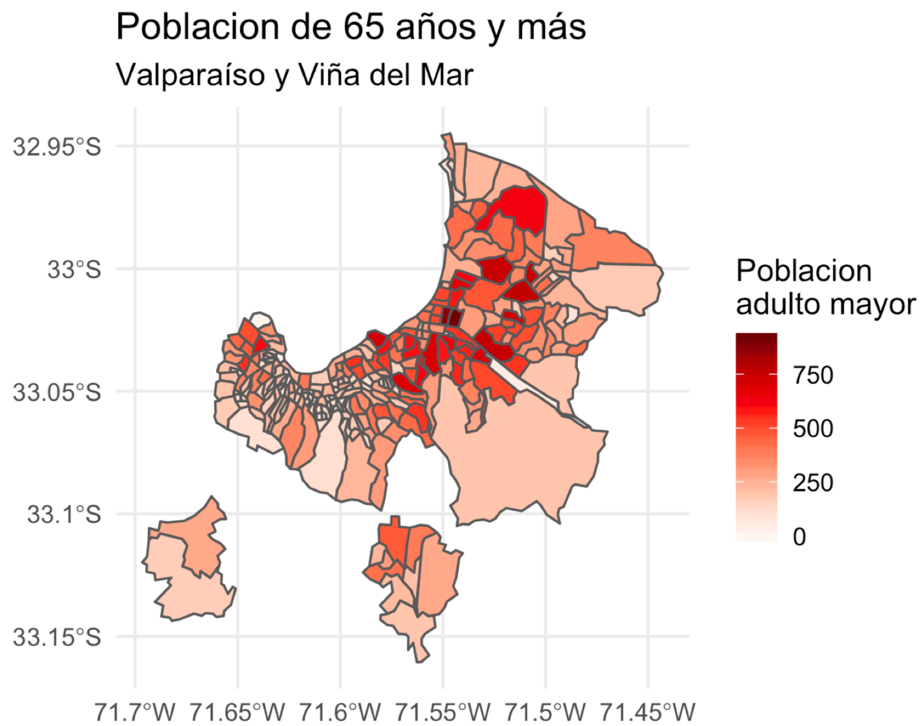
- Practical
  - Data: geo-located observations
    - Private sources
    - Public sources
    - Self reported vs. Web scraped
  - Spatial mismatch between data and social processes
    - Labor markets vs. Cities and Counties → Labor Market areas
    - Epidemic spread vs. Data collection in hospitals
  - Neighborhood Effects
  - Spatial interpolation
  - Change of support problem
    - Data at different spatial levels that don't overlap
      - Ej: Parents' income aggregated at the school level used to predict municipalities' income level

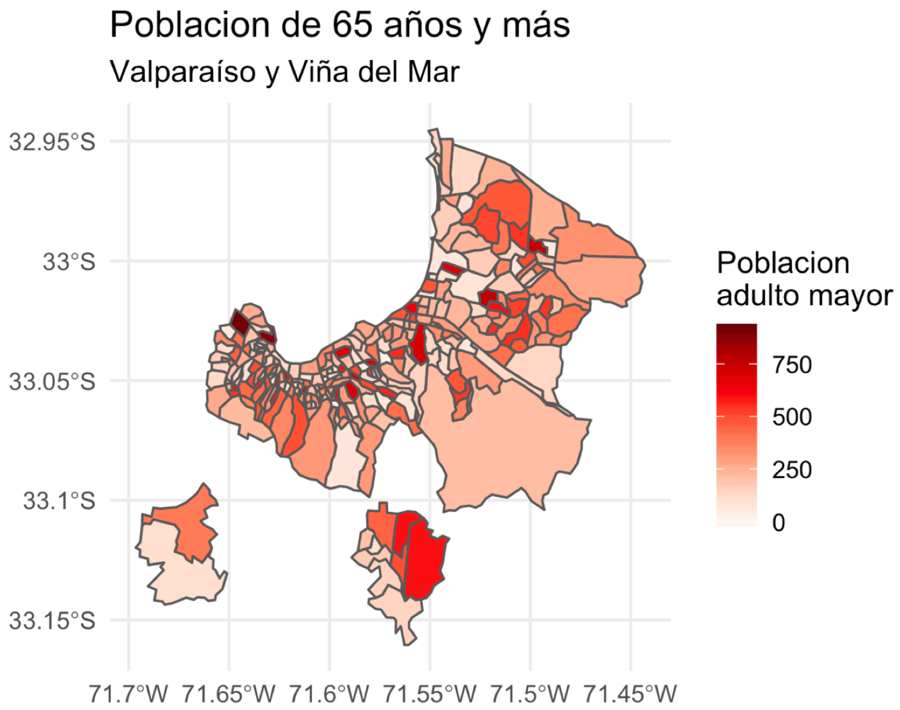
What is Spatial Analytics?

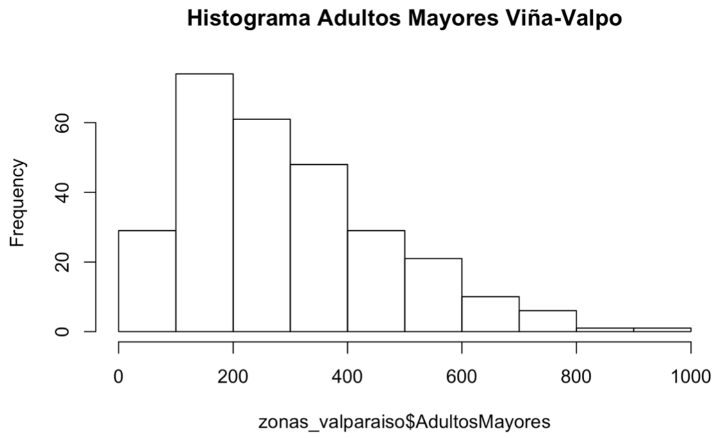
- More than just mapping
  - Added value in the explanation
  - Combination of methods, theory, data manipulation
  - Knowledge discovery
  - “from data, to information, to knowledge, to wisdom”
  - Correlation is not causation, and this applies also to spatial analysis.

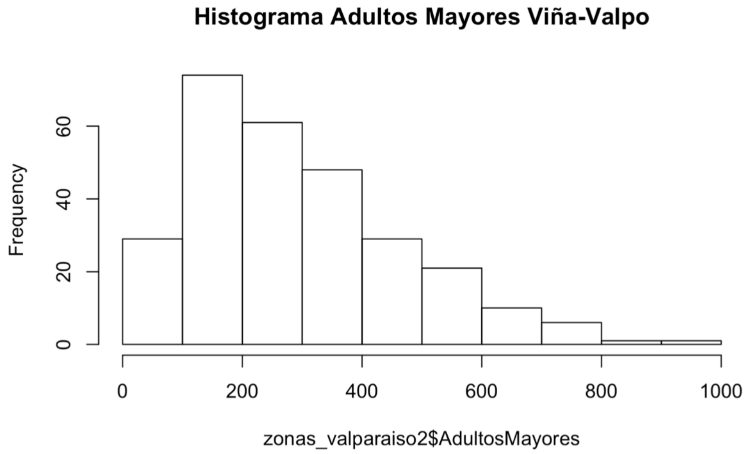
Questions of Spatial Analytics

- Where do things happen
  - Patterns, clusters, hot spots, disparities,..
- Why do things happen
  - Location decisions
- How, things that happen, affect other things (spillovers) and how context affect what happens (interaction)
- Where should things be happening/be located
  - Optimization


Poblacion de 65 años y más
Valparaíso y Viña del Mar
32.95°S
33°S
33.05°S
33.1°S
33.15°S
71.7°W 71.65°W 71.6°W 71.55°W 71.5°W 71.45°W
Poblacion
adulto mayor
750
500
250
0


Poblacion de 65 años y más
Valparaíso y Viña del Mar
32.95°S
33°S
33.05°S
33.1°S
33.15°S
71.7°W 71.65°W 71.6°W 71.55°W 71.5°W 71.45°W
Poblacion
adulto mayor
750
500
250
0


Histograma Adultos Mayores Viña-Valpo
Frequency
0
20
40
60
0 200 400 600 800 1000
zonas_valparaiso$AdultosMayores


Histograma Adultos Mayores Viña-Valpo
Frequency
0
20
40
60
0 200 400 600 800 1000
zonas_valparaiso2$AdultosMayores

Part 2: Applications

Outline

1. Economic impacts of land use changes and forestry expansion

2. Using ML for predicting forest fires from forest plantations

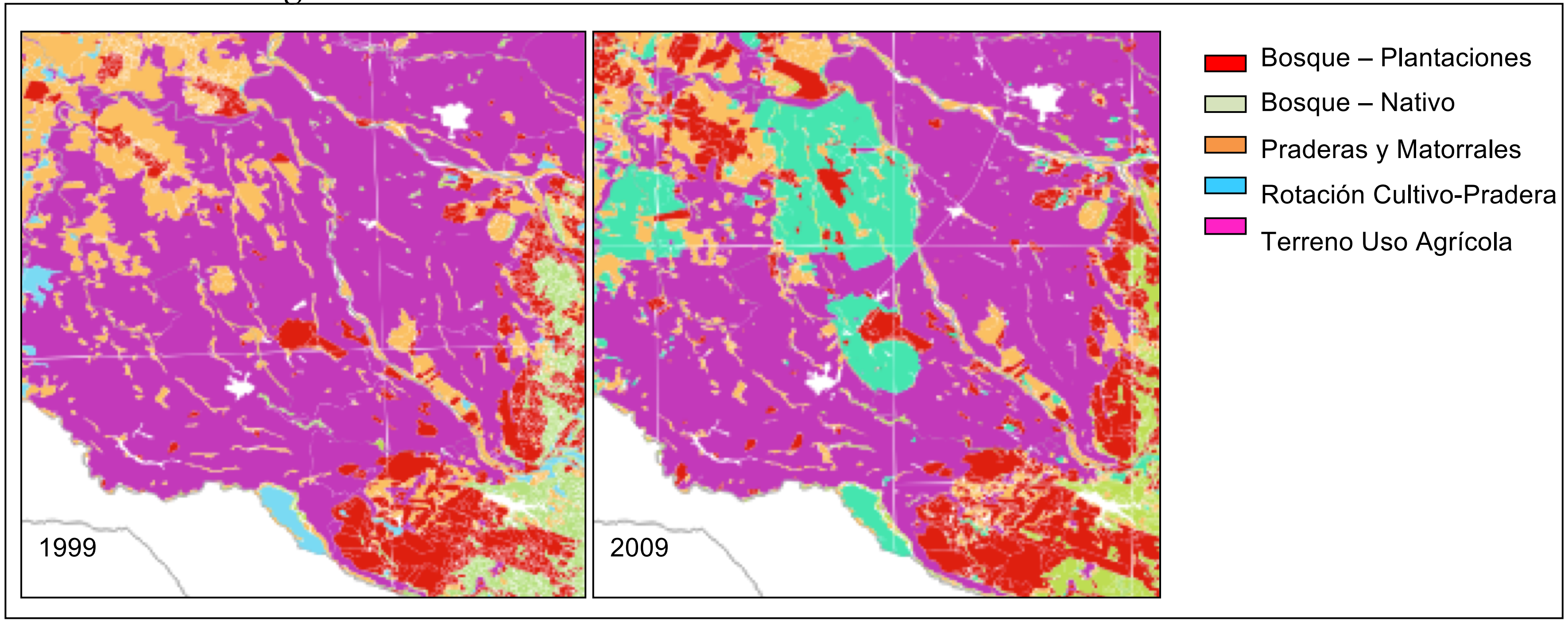
3. Satellite imagery for estimating the impact of green areas in housing prices

4. Effects of school disntances on Educational Outcomes

5. Other
    1. Bonus – Patrones espaciales de la Ley de Protección del Empleo (Boletin CEPR)
    2. ML on predicting Deforestation patterns (Ivan Flores)

1 Spatial patterns of forestry expansión matter

Figura 2. Ejemplo del cambio de uso de suelo potencialmente causado por la expansión Forestal en la Región del Maule 1999-2009



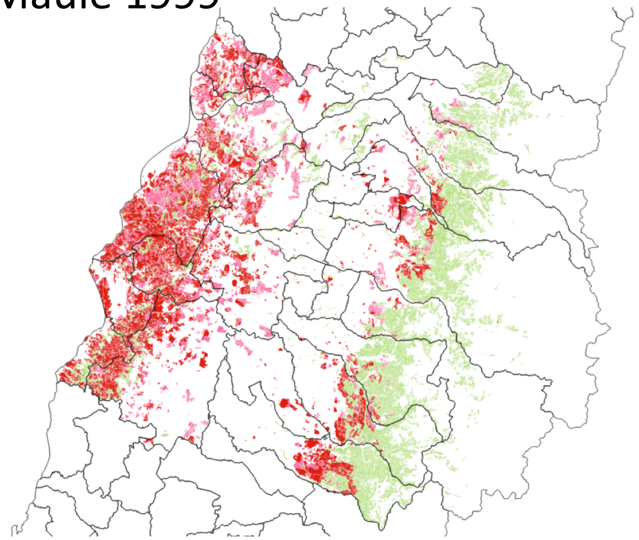
Fuente: Elaboración propia con datos del Sistema de Información Territorial (SIT) CONAF
Nota: Zoom de la Figura 1. Como se aprecia, en 2009 el área de plantación de bosques (en rojo) que tiene a su alrededor áreas de Rotación Cultivo-Pradera (en cian), es nueva. Estas dos áreas en 1999 estaban clasificadas como Praderas y Matorrales y Terreno de Uso Agrícola respectivamente. La otras dos áreas que pasaron de Terrenos Agrícolas a Rotación Cultivo-Pradera en el panel del 2009, también contienen asentamientos de plantaciones de bosques. Estos cambios pudieron deberse a la escases de agua producto de la instalación de plantaciones de bosques de especies exóticas de alto consumo de agua.

Forestry expansion has occurred at the expense of agriculture and pasture land-uses

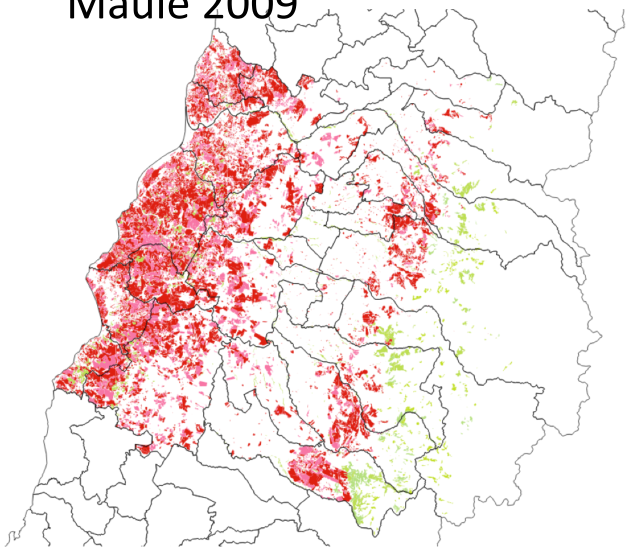
Percentage		(1) Agriculture	(2) Pasture	(3) Wetland	(4) Native forest
	Forestry	-0.388	-0.609	0.000	0.003
		(0.034)***	(0.036)***	(0.004)	(0.014)
	Year	-0.024	-0.231	-0.041	-0.138
		(0.271)	(0.289)	(0.029)	(0.111)
	Constant	29.362	35.804	0.854	21.824
		(0.660)***	(0.702)***	(0.070)***	(0.269)***
	Observations	232	232	232	232

*** p<0.001; ** p<0.005; *p<0.010

Maule 1999



Maule 2009



Forestry expansion reduces:

- employment, particularly in non-forestry agricultural activities,
- reduces labor earnings and the probability of finding formal jobs

Urbano y Rural	(1) ln_income
forestry_sd	-0.213***
	(0.076)
Mean(y)	12.395
N	25657

Urbano	(2) ln_income
forestry_sd	-0.340***
	(0.085)
Mean(y)	12.080
N	9395

Rural	(2) ln_income
forestry_sd	-0.145**
	(0.072)
Mean(y)	12.578
N	16262

Standard errors in parentheses
="* p<0.1, ** p<0.05, *** p<0.01"

2 Predicciones de incendios forestales

Variable dependiente:

- Incendios forestales (dicotómica)

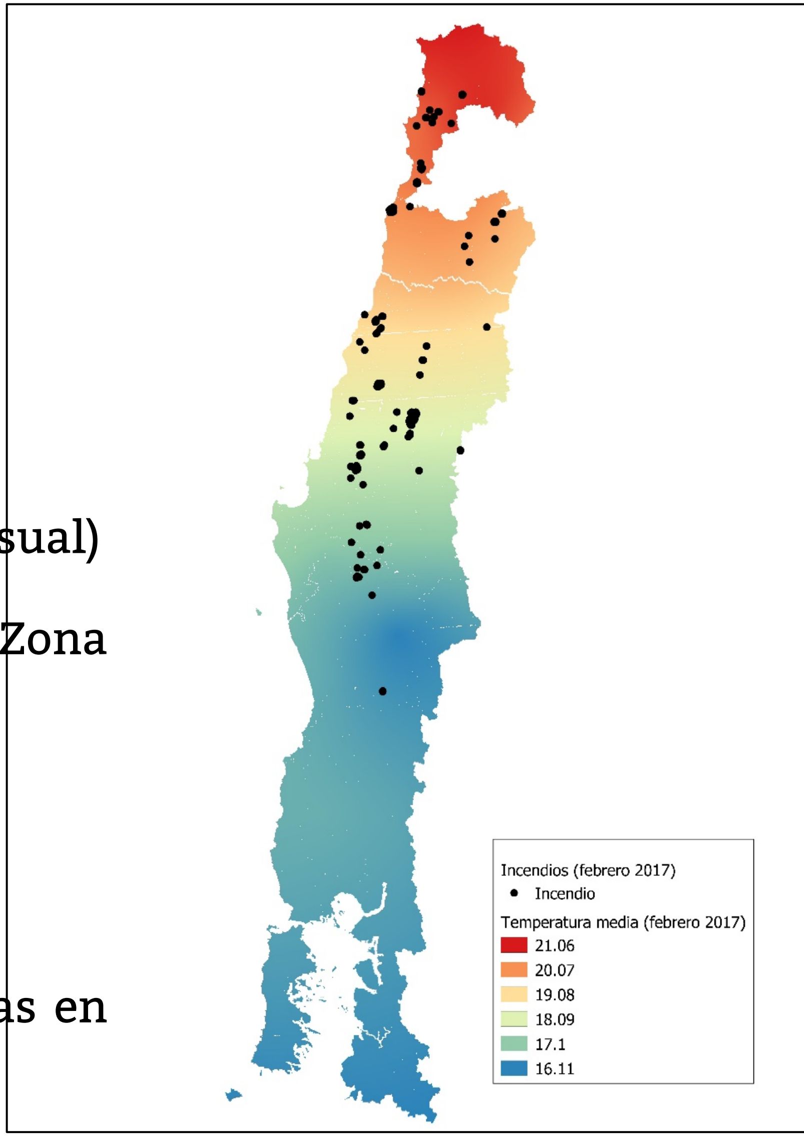
Variables independientes:

- Meteorológicas (precipitaciones mensuales y temperatura media mensual)
- Geográficas (Altitud, Pendiente, Rugosidad, TRI)
- Uso de suelos (Áreas protegidas, Áreas forestales, Bosque nativo, Zona agrícola, Pasturage, Humedales)
- Dummys de mes (mes_1, mes_2, mes_3)

Random forest:

- Número de árboles a estimar: 100
- Training set: muestra aleatoria del 70% de observaciones
- Testing set: muestra aleatoria del 30% de observaciones (no incluidas en training set)

Temperatura media e incendios en febrero de 2017



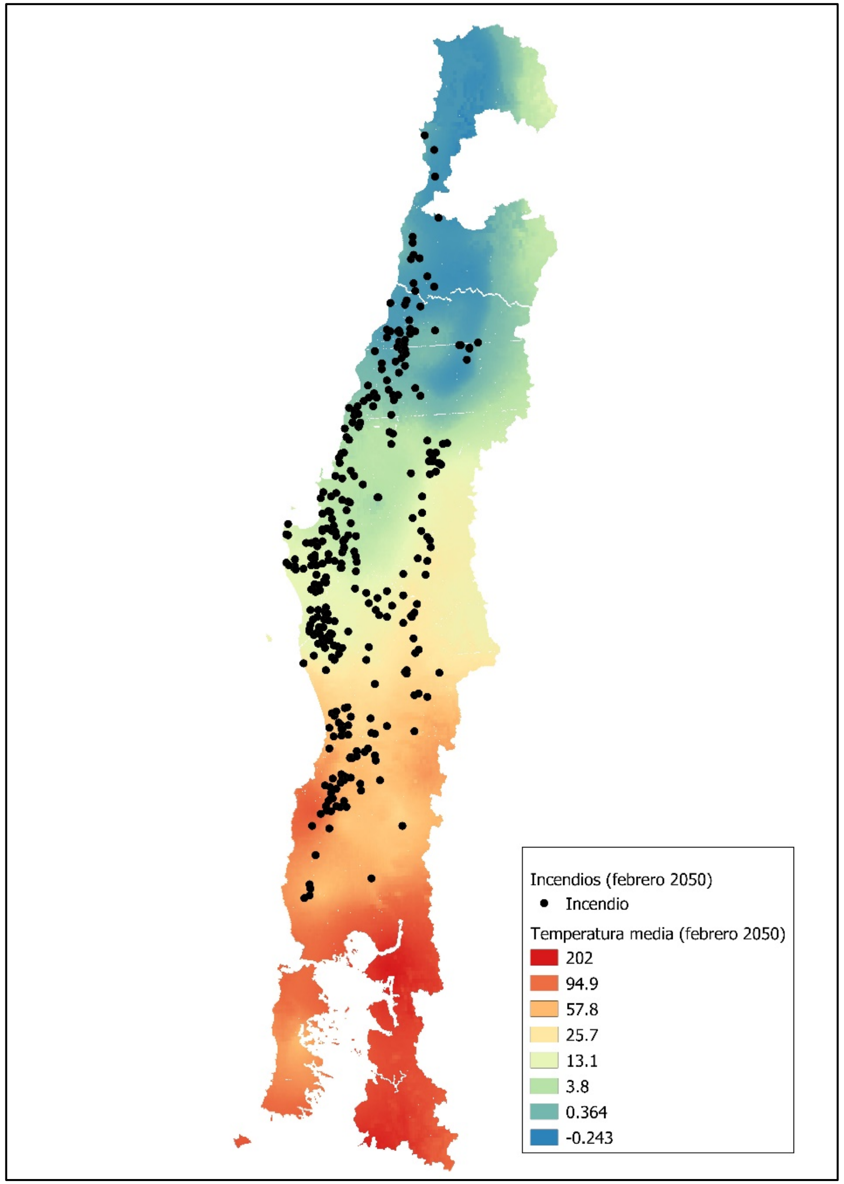
Incendios forestales y cambio climático (preliminar)

Para predecir la ocurrencia de incendios en zonas forestales en 2050 en un contexto de cambio climático, se realizó lo siguiente:

- Se descargaron y procesaron los datos históricos de clima de 1970-2000 de WorldClim
- Se hizo match con la base de escenarios climáticos y se calcularon las variaciones en temperatura y precipitaciones para el escenario más drástico (rcp85): var_temp_rcp85 = (temp_rcp85/temp_1970_2000).
- Posteriormente se agregaron las variaciones a la base con datos de estaciones meteorológicas y se calculan las posibles temperaturas de chile en 2050 (temp_2017 * var_temp_rcp85).

Con esta base (y agregándole variables geográficas y de uso de suelo) se prueba el modelo previamente estimado.

Incendios 2050 y Forestales



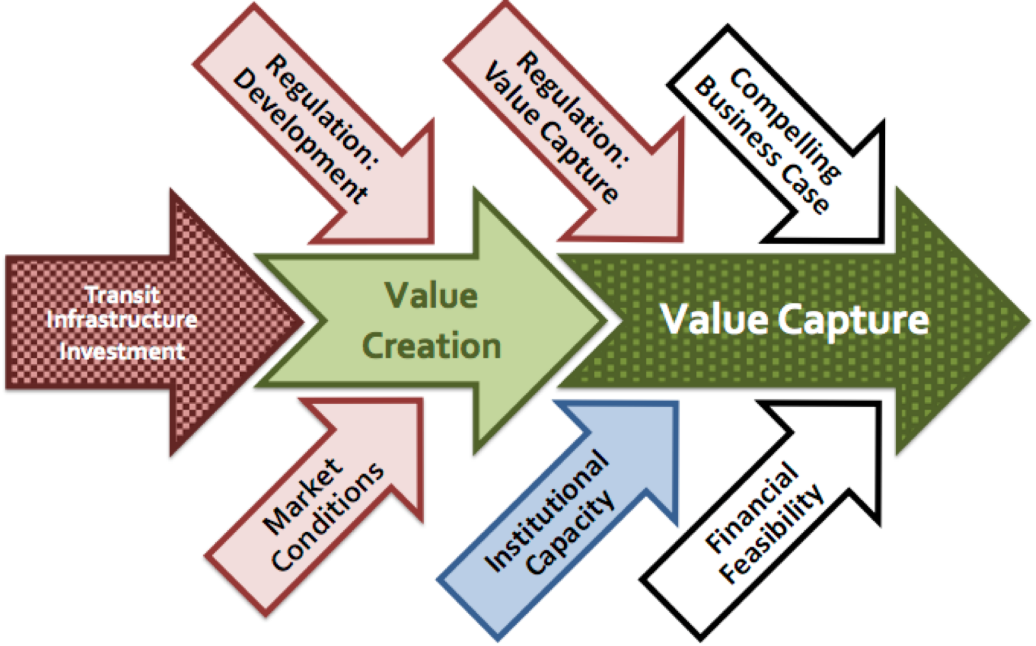
3

Main Question

How much should we ‘capture’ in property taxes from public investment in green areas?

Main Goal

Revealing how much people is willing to pay can help in creating a plan and advocating for urban renewal



- Non Parametric
  - Locally Weighted Regression (LWR)

$$\sum_{i=1}^{n}(y_i-\alpha-\beta'(x_i-x))^2K(\psi_i)$$

$$Z_i=\binom{1}{x_i} \qquad \theta=\binom{\alpha}{\beta}$$

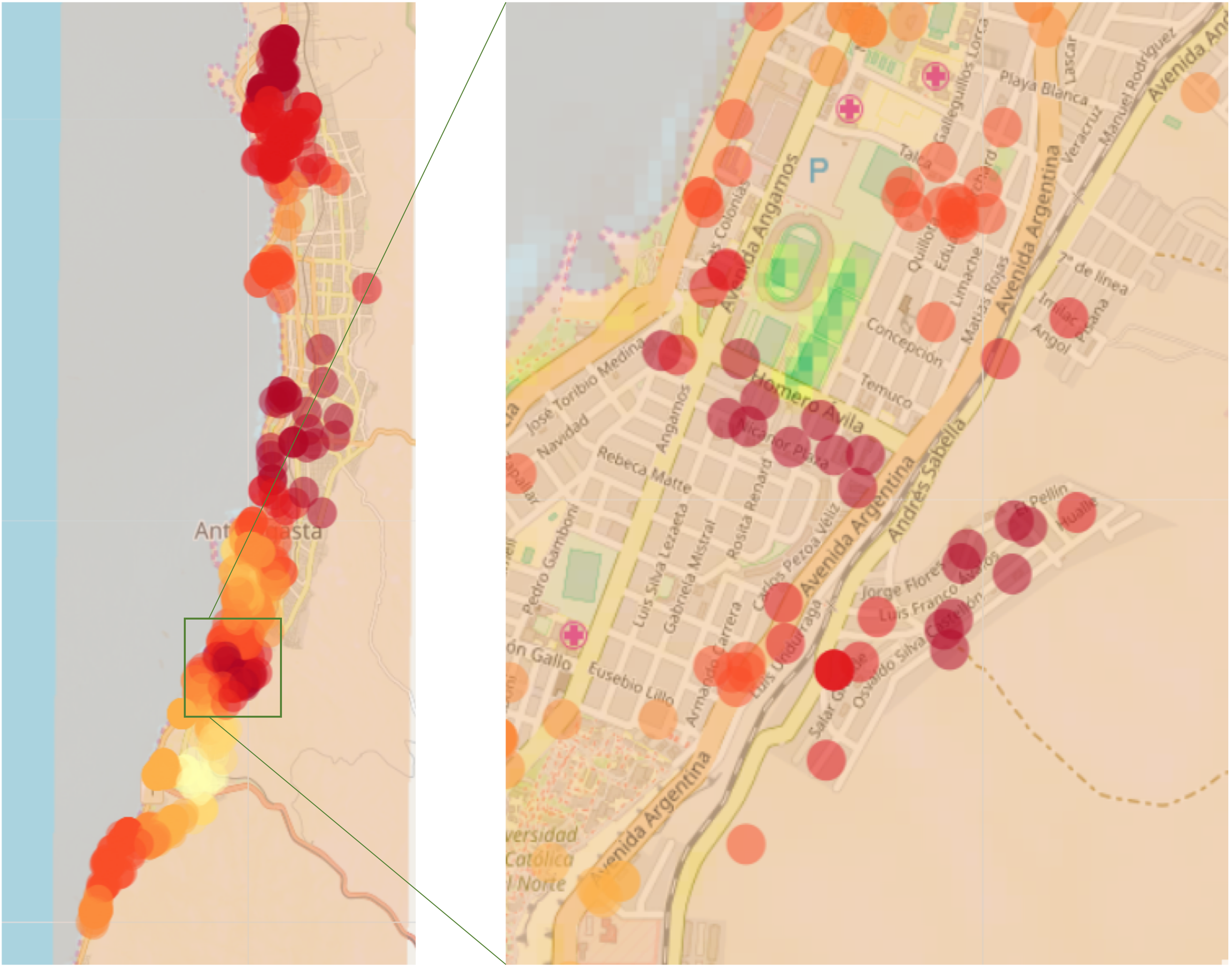
$$\hat{\theta}(x)=(\sum_{i=1}^{n}K(\psi_i)Z_iZ_i')^{-1}\sum_{i=1}^{n}K(\psi_i)Z_i'y_i$$

A regression of wy on wz, where

$$w_i=K\left(\frac{x_i-x}{h}\right)^{1/2}$$

Each point is a housing unit colored by the willingness to pay for closeness to green areas

More red colors indicate higher willingness to pay

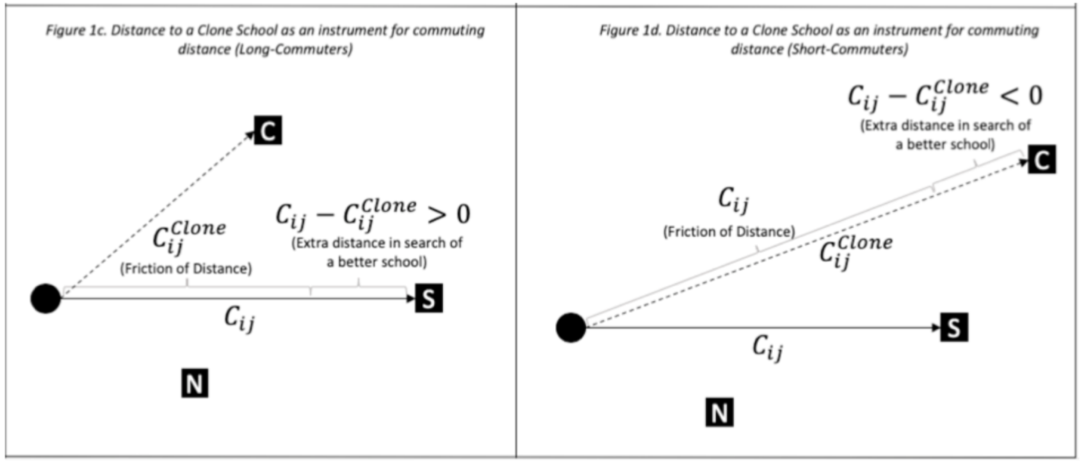


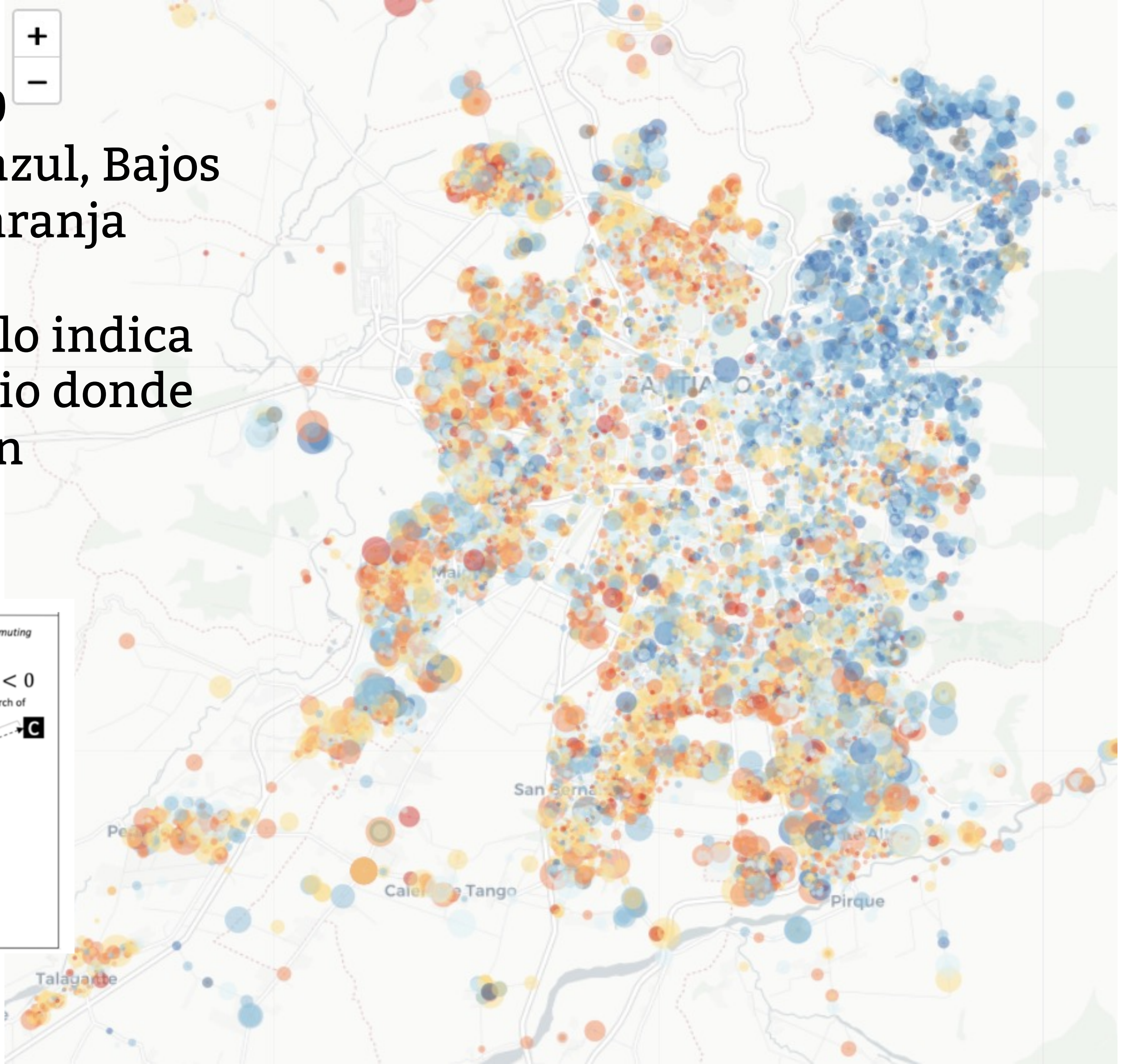
4

PSU 2010

Altos puntajes en azul, Bajos puntajes en naranja

Tamaño del círculo indica distancia al colegio donde estudiaron





BOLETÍN LABORAL CEPR

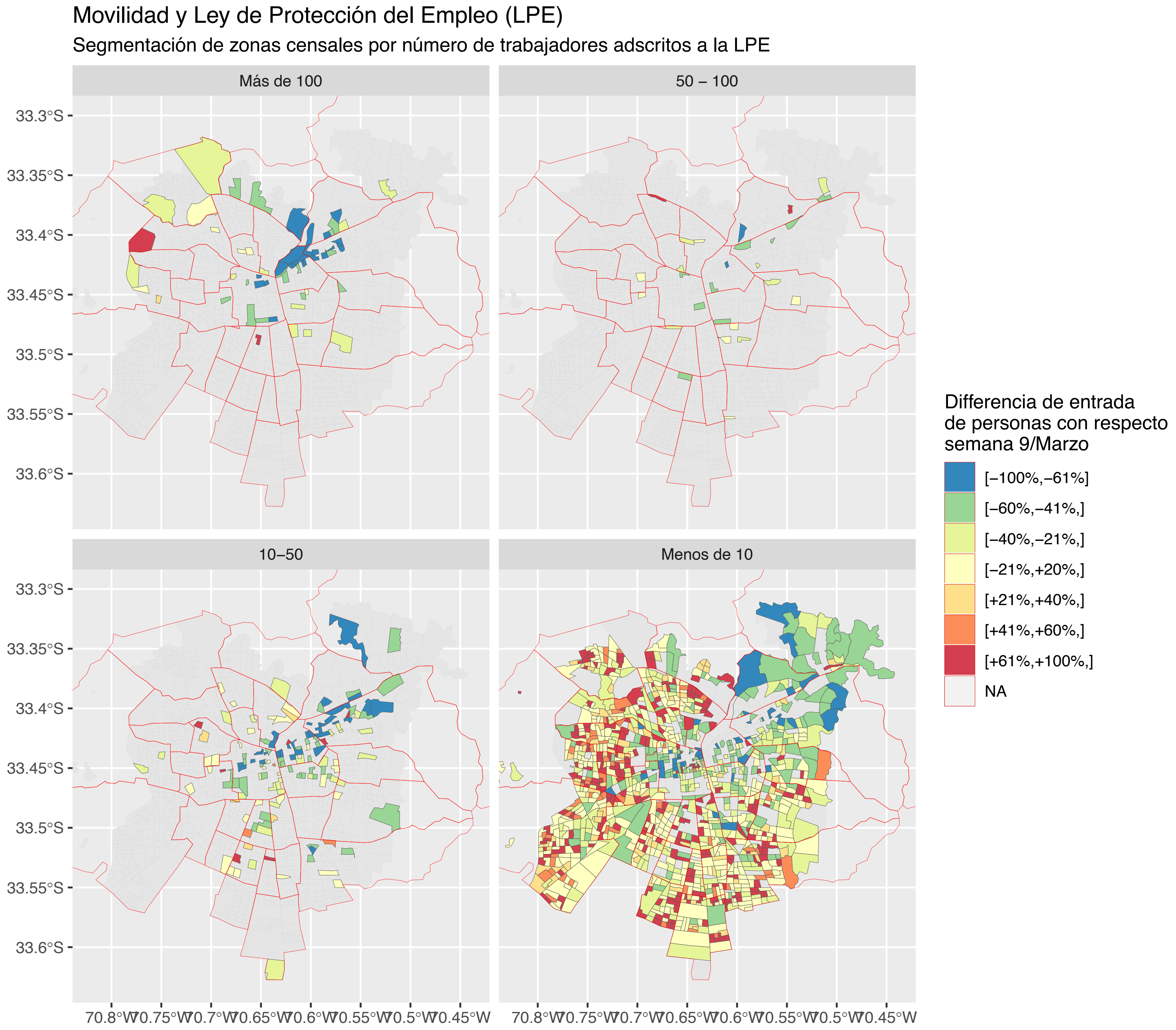
REGIÓN METROPOLITANA

Bonus

La mayoría de las personas cubiertas por la LPE están en areas que hasta el 6/07 tenían baja movilidad y que son de mayores ingresos.

Potencial pérdida de empleo en las áreas de residencia de mayor poder adqusitivo

Desafíos importantes de movilidad y organización de la reapertura



Elaboración: CEPR–UAI

Fuente datos: Registro Público Ley 21.227 – Protección del Empleo, Dirección del Trabajo, SII, Chilemapas, ISIC & Entel

4

Testing the accuracy of machine learning methods to predict deforestation spatial patterns using Monte Carlo simulations

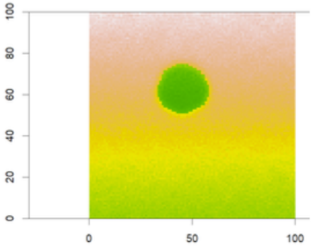
Ivan Flores

Department of Agricultural and Consumer Economics
University of Illinois at Urbana-Champaign

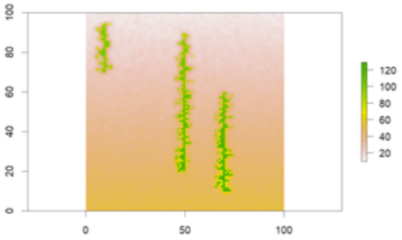
June 07, 2020

Methods & Data

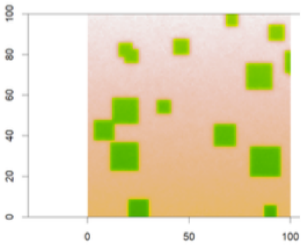
- Variation in lags (0, 0.5, 0.9) and W definition (Queen, Rook)
- Simulated data



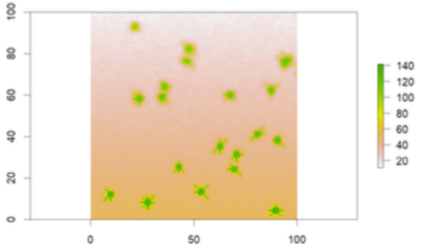
(a) Radial



(b) Fish-bone



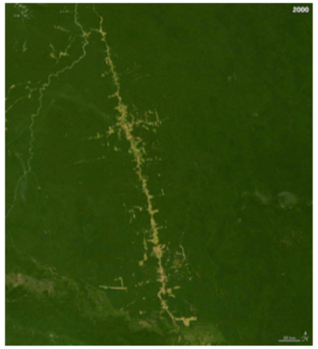
(c) Geometric



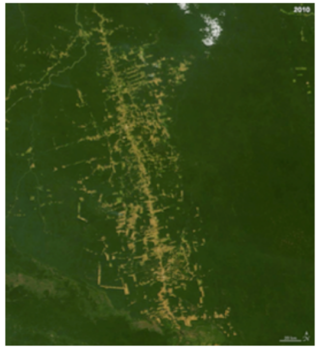
(d) Diffuse

Spatial patterns of human activities

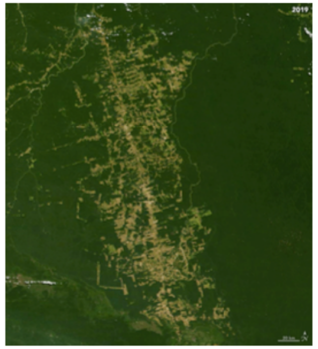
- Fish-bone
  - Perpendicular smaller roads spawn from a main road simulating the structure of a fish spine
  - Related to new colonizations and planned movements as well as early stages of illegal activities such as logging and mining (Solinge, 2014)



(a) 2000



(b) 2010



(c) 2019

Figure 3: BR-163, State of Pará, Brazil

Practical example

Fork: https://github.com/estebanlp/Intro2SpatialAnalytics.git