Ciencia de Datos Geoespaciales

IMT2118
Paula Aguirre
Clase 22 –08/06/2023

MODELOS DE ANÁLISIS ESPACIAL

	Tarea	Modelo “convencional” (datos tabulares)	Modelo para análisis espacial
No supervisado	**Clustering**	K-means Clustering jerárquico DBScan	Análisis de patrones de puntos ✅ Clustering basado en densidad✅ Autocorrelación espacial ✅ Análisis de Hot Spots (LISA) ✅ Clustering multivariados con restricción espacial✅ Segmentación Patrones espacio-temporales
Supervisado	**Regresión** Predicción de valores numéricos	Regresión lineal Regresión kNN Decision tree Gradient boosting SVM ...etc	OLS Regresión ponderada espacialmente **Kriging (empirical bayesian kriging)** Interpolación por área
	Clasificación Predicción de valores categóricos o clases	Decision tree Random forest Support vector machine Máximum likelihood ...etc	Decision tree Random forest Support vector machine Máximum likelihood ...etc

ANÁLISIS DE REGRESIÓN Y DEPENDENCIA ESPACIAL

- El análisis de regresión permite modelar y predecir algún proceso en base a su relación con una o varias variables dependientes.

- Los modelos estándar de regresión pueden ser insuficientes para modelar datos con dependencia espacial (i.e. cuando los datos presentan autocorrelación espacial).

  - Modelos de regresión asumen independencia entre observaciones: lo que ocurre en un área $\mathbf{x_i}$ no tiene ninguna relación con lo que ocurre en el área $\mathbf{x_j}$

  - Esta suposición puede ser invalidada si las dos áreas x_i, x_j son vecinas, y los datos presentan una autocorrelación espacial.

- El **análisis de regresión espacial** permite incorporar dependencias espaciales en el modelo.

ANÁLISIS DE REGRESIÓN (NO ESPACIAL)

- **Regresión lineal ➔** busca explicar la variación de una variable (*dependiente*) como una combinación lineal de una serie de otras variables (*explicativas o independientes*)

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \epsilon$$

- **Ejemplos:**

$$P = \alpha + \beta_1 NEW + \beta_2 IMD + \epsilon$$

¿Cómo explicamos el precio de una casa en función del nivel de deprivación y antigüedad del área donde está ubicada?

¿Cómo explicamos la tasa de criminalidad en una determinada zona en función de variables sociodemográficas?

ANÁLISIS DE REGRESIÓN (NO ESPACIAL)

- **Regresión lineal ➔** busca explicar la variación de una variable (*dependiente*) como una combinación lineal de una serie de otras variables (*explicativas o independientes*)

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \epsilon$$

- **Interpretación:**

  Coeficientes $\boldsymbol{\beta_1, \beta_2, \beta_3}$ entregan información respecto a cómo y en qué grado la variable respectiva X_1, X_2, X_3 determina el valor de Y

  $\boldsymbol{\beta_k}$ representa el grado de correlación entre la variable explicativa k y la variable dependiente, *mantiendo todas las otras variables constantes*

  $\boldsymbol{\alpha}$ indica el valor promedio de Y cuando todas las otras variables son 0

ANÁLISIS DE REGRESIÓN (NO ESPACIAL)

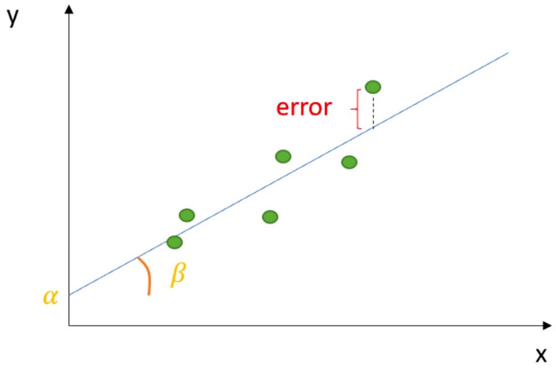
¿Cómo determinamos los valores de $\boldsymbol{\alpha}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}$?

- **Mínimos cuadrados ordinarios (OLS):** método que optimiza los parámetros de la regresión lineal de manera de **minimizar los errores cuadráticos**.

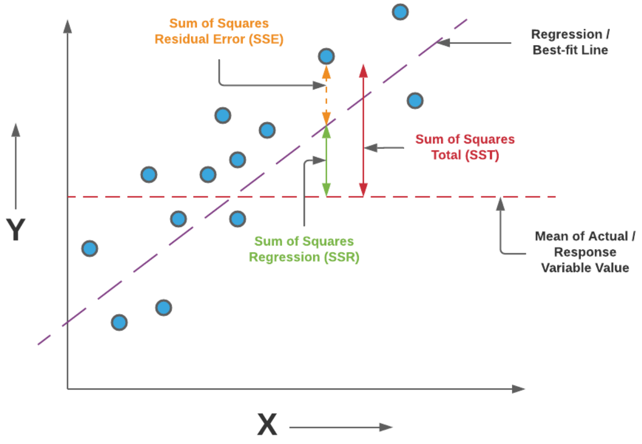
- Estimadores α, β_i son los mejores estimadores lineales no-sesgados (Best Linear Unbiased Estimators, BLUE) de los valores reales.

- ¿Qué tan bueno es el ajuste obtenido? Podemos analizarlo en base a la métrica **R²**:

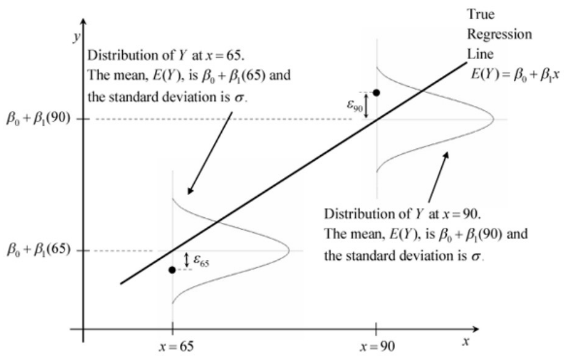
$$R^2 = \frac{SSR}{SST} = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$



$$y_i = \alpha + \beta x_i + \varepsilon_i$$



ANÁLISIS DE REGRESIÓN (NO ESPACIAL)



$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \epsilon$$

El método OLS se basa en ciertas suposiciones respecto a las observaciones:

1. **El modelo de regresión es lineal en los coeficientes y error.**

2. **El error sigue una distribución normal, con valor esperado 0** (i.e., en promedio el error es 0).

$$\epsilon \sim N(0, \sigma^2)$$

➔ Podemos definir errores estándar para los coeficientes, e intervalos de confiabilidad

3. **Ninguna variable es combinación lineal perfecta de las otras variables.**

ANÁLISIS DE REGRESIÓN (NO ESPACIAL)

El método OLS se basa en ciertas suposiciones respecto a las observaciones:

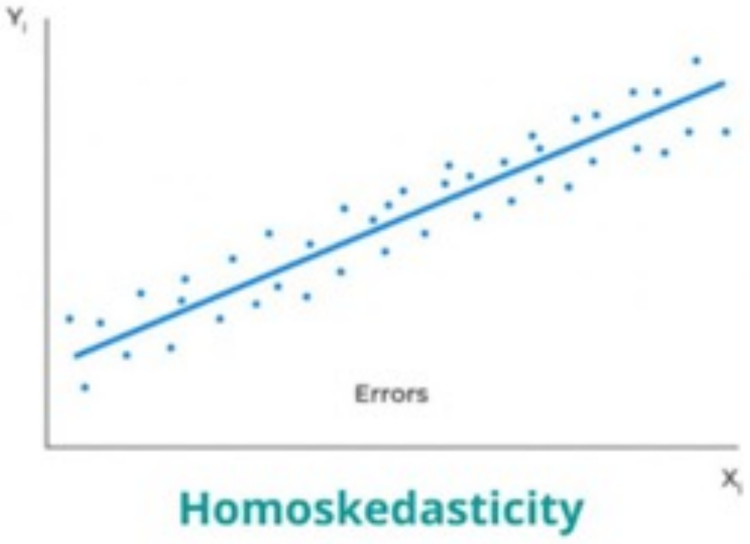
4. **No hay correlación entre las variables independientes y el error.**

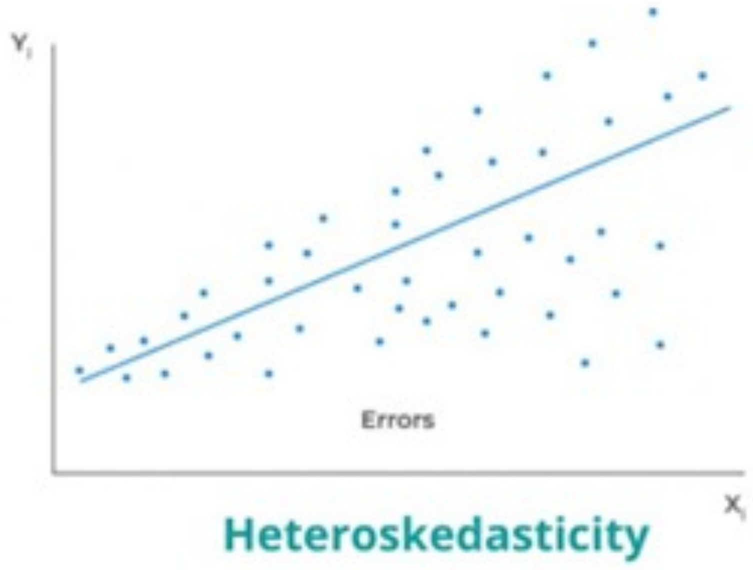
➔Si no se cumple, el error asociado a la variable X_i se traspasa al coeficiente (ej: cuando hay errores de medición en una variable)

3. **No hay correlación entre los errores (o, las observaciones son independientes entre sí).**

➔ Si el error para una observación es positivo, y eso aumenta la probabilidad de que el error en la siguiente observación sea positivo, entonces hay una correlación.

ANÁLISIS DE REGRESIÓN (NO ESPACIAL)





La varianza del error es igual para todas as observaciones.

La varianza del error varía para distintas observaciones.

REGRESIÓN ESPACIAL

- Los modelos estándar de regresión pueden ser insuficientes para modelar datos con dependencia espacial.

- Modelos de regresión asumen independencia entre observaciones: lo que ocurre en un área x_i no tiene ninguna relación con lo que ocurre en el área x_j

- Este requisito no siempre se cumple cuando modelamos un proceso espacial, donde la ubicación *sí importa*.

- O, visto de otra forma, un modelo espacial puede ayudar a explicar la estructura de errores (residuos) en modelos de regresión no-espaciales.

Ejemplos:

➔ Precios de viviendas dependen de la ubicación, e influyen en el valor de las propiedades vecinas.

➔ Proyección espacial de instalaciones de paneles fotovoltaicos (energía solar)

REGRESIÓN ESPACIAL

Identifying the determinants of housing prices in China using spatial regression and the geographical detector technique

Yang Wang [a, b], Shaojian Wang [c, *], Guangdong Li [d, **], Hongou Zhang [a, b], Lixia Jin [a, b], Yongxian Su [a, b], Kangmin Wu [a, b]

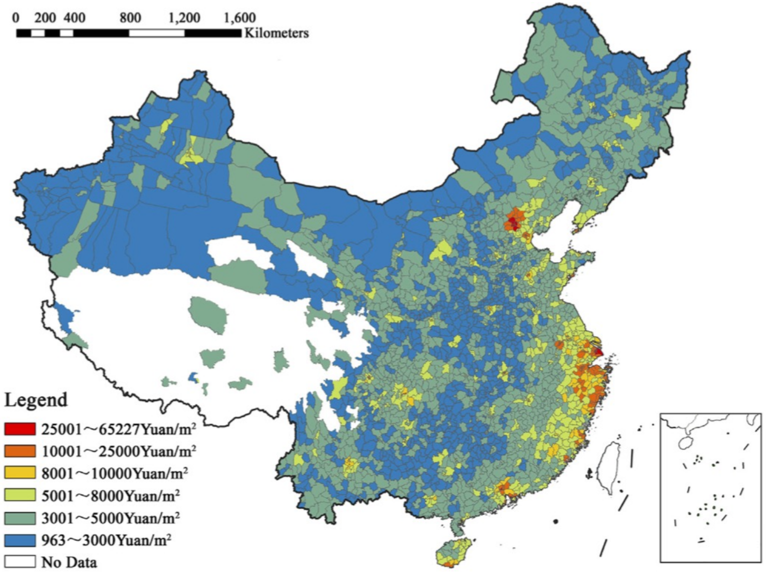


Fig. 3. The spatial pattern of housing prices in China at the county level.

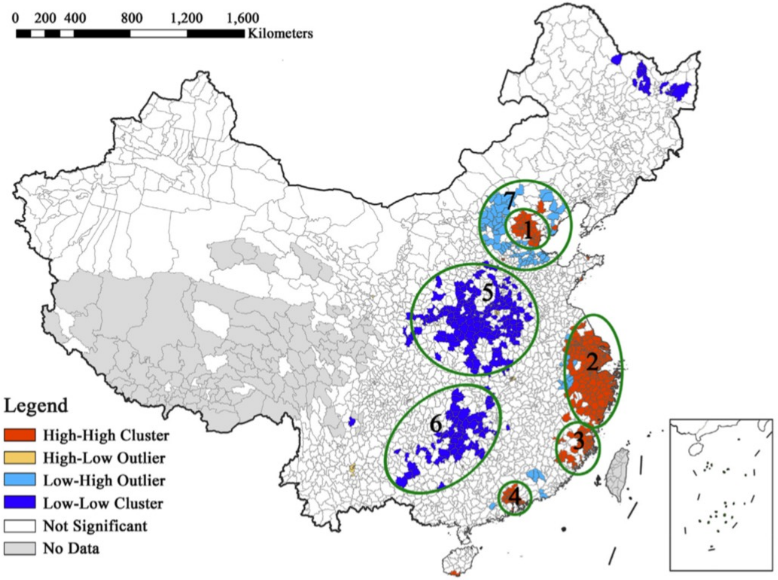


Fig. 5. Local Moran's I for China's housing prices at the county level.

REGRESIÓN ESPACIAL

Identifying the determinants of housing prices in China using spatial regression and the geographical detector technique

Yang Wang [a, b], Shaojian Wang [c, *], Guangdong Li [d, **], Hongou Zhang [a, b], Lixia Jin [a, b], Yongxian Su [a, b], Kangmin Wu [a, b]

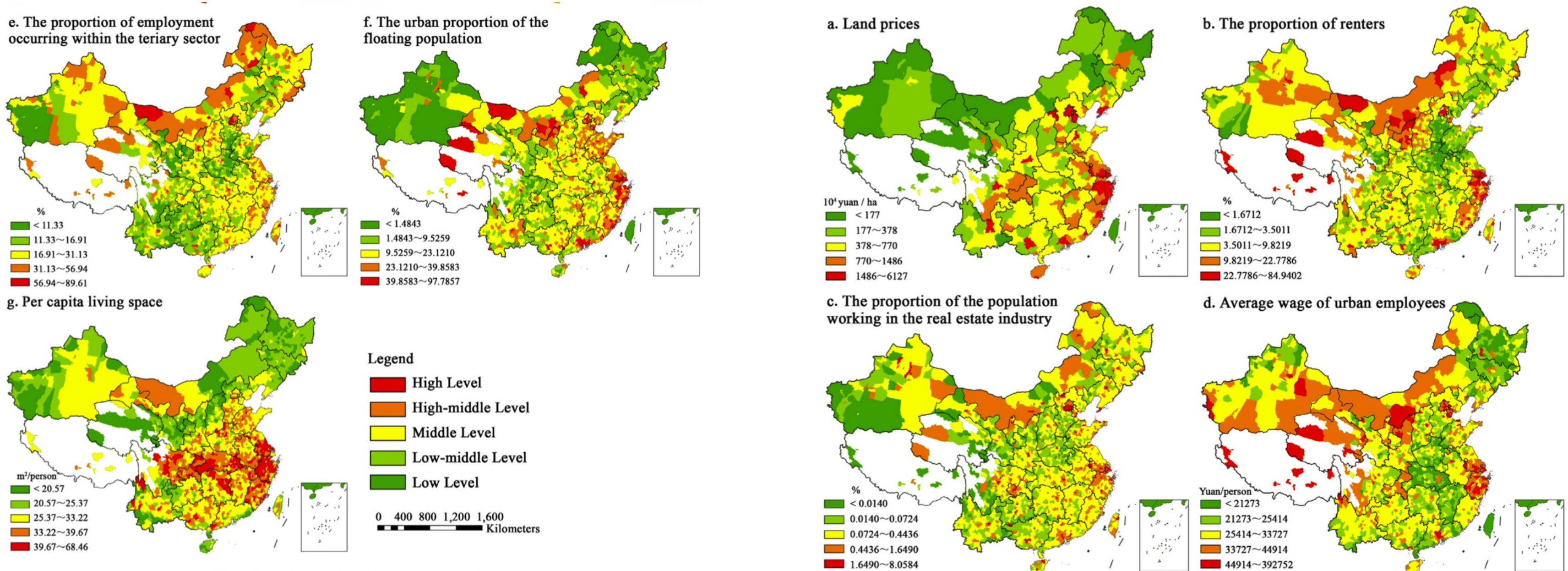


Fig. 6. Maps of seven factors of housing prices in China.

REGRESIÓN ESPACIAL

- Los modelos estándar de regresión pueden ser insuficientes para modelar datos con dependencia espacial.

- Modelos de regresión asumen independencia entre observaciones: lo que ocurre en un área x_i no tiene ninguna relación con lo que ocurre en el área x_j

- Este requisito no siempre se cumple cuando modelamos un proceso espacial, donde la ubicación *sí importa*.

- O, visto de otra forma, un modelo espacial puede ayudar a explicar la estructura de errores (residuos) en modelos de regresión no-espaciales.

Ejemplos:

➔ Precios de viviendas dependen de la ubicación, e influyen en el valor de las propiedades vecinas.
➔ Proyección espacial de instalaciones de paneles fotovoltaicos (energía solar)

REGRESIÓN ESPACIAL

Applied Energy
Volume 265, 1 May 2020, 114747

Spatial projections of solar PV installations at subnational level: Accuracy testing of regression models

Jonas Müller [a, b], Evelina Trutnevyte [a]

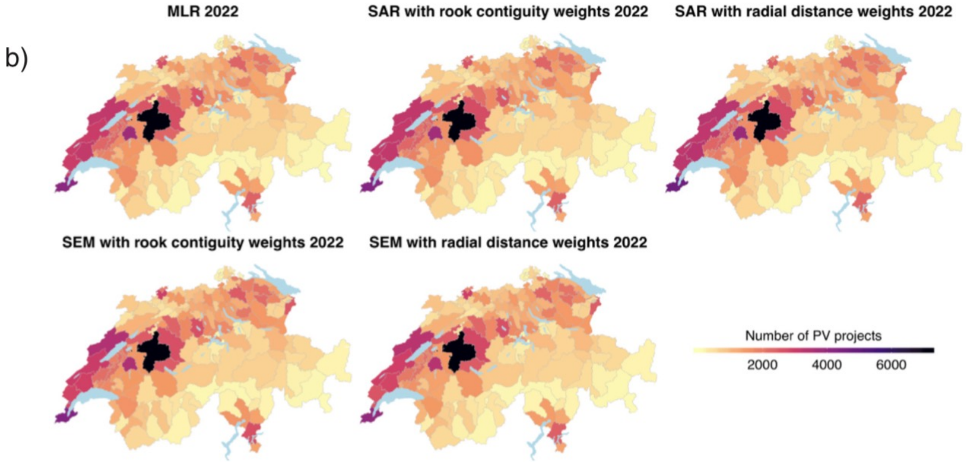
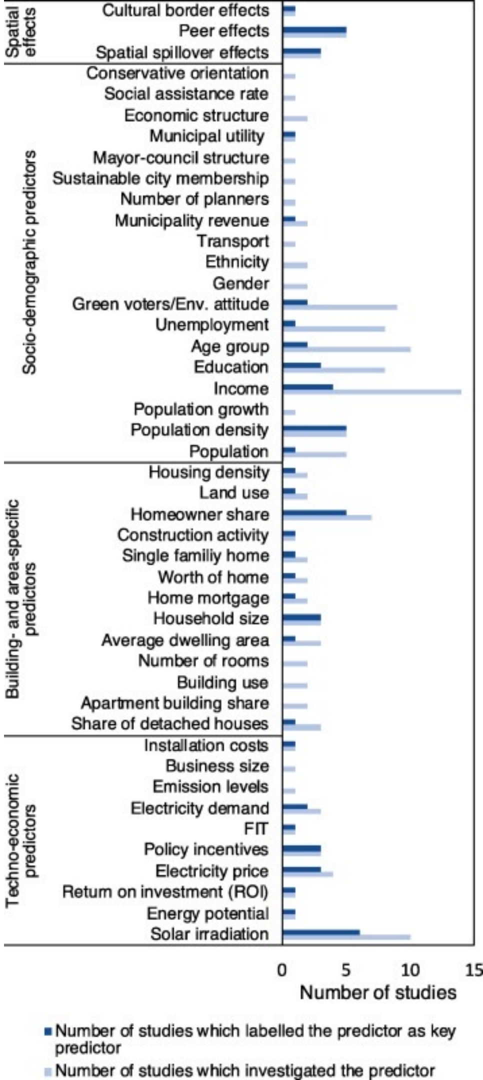


Fig. 11. 5-year-ahead spatial projections of the number of PV projects up to 2022.

https://www.sciencedirect.com/science/article/abs/pii/S0306261920302592



REGRESIÓN ESPACIAL

Ejemplos:

- Precios de viviendas dependen de la ubicación, e influyen en el valor de las propiedades vecinas.
- Proyección espacial de instalaciones de paneles fotovoltaicos (energía solar)
- Análisis de regulaciones locales a actividades productivas

REGRESIÓN ESPACIAL

Understanding Local Regulation of Fracking: A Spatial Econometric Approach

Patrick J. Walsh, Stephen Bird, and Martin D. Heintzelman

Fracking is a controversial practice but is thriving in many areas. We combine a comprehensive data set on local bans and moratoria in the state of New York with local-level census data and spatial characteristics in a spatial econometric analysis of local fracking policies. Some factors, including location in the Utica shale, proportion of registered Democrats, and education level, increase the probability of restrictions on fracking. Extent of local land development, location in highly productive petroleum areas, and number of extant oil and gas wells are among factors that have a negative impact on the likelihood of a ban or moratorium.

Key Words: fracking, local policy analysis, spatial econometrics, survival analysis

Variable
Ban
Moratorium
Ban or moratorium
Share over Marcellus
In prime Marcellus region
Share over Utica
Share over Utica sweet spot
Percent manufacturing employment
Percent arts and tourism employment
Percent natural resource and construction employment
Percent developed land area
Count of recent water wells (since 2000)
Percent open water
Percent wetlands
Ratio of Democrats to Republicans
Existing oil and gas wells
Share within priority watershed
Homes with solar systems
Civilian unemployment rate
Median household income ($1,000)
Percent over 65 years of age

DOI: 10.1017/S1068280500010261

REGRESIÓN ESPACIAL

Understanding Local Regulation of Fracking: A Spatial Econometric Approach

Patrick J. Walsh, Stephen Bird, and Martin D. Heintzelman

Fracking is a controversial practice but is thriving in many areas. We combine a comprehensive data set on local bans and moratoria in the state of New York with local-level census data and spatial characteristics in a spatial econometric analysis of local fracking policies. Some factors, including location in the Utica shale, proportion of registered Democrats, and education level, increase the probability of restrictions on fracking. Extent of local land development, location in highly productive petroleum areas, and number of extant oil and gas wells are among factors that have a negative impact on the likelihood of a ban or moratorium.

Key Words: fracking, local policy analysis, spatial econometrics, survival analysis

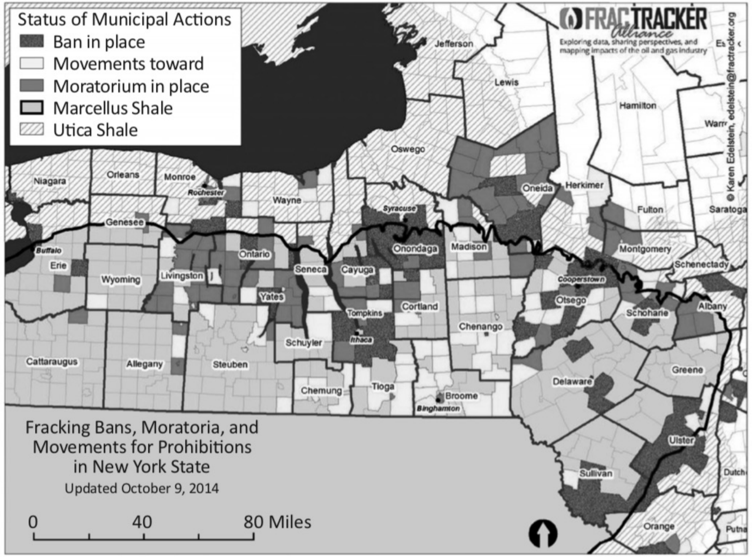
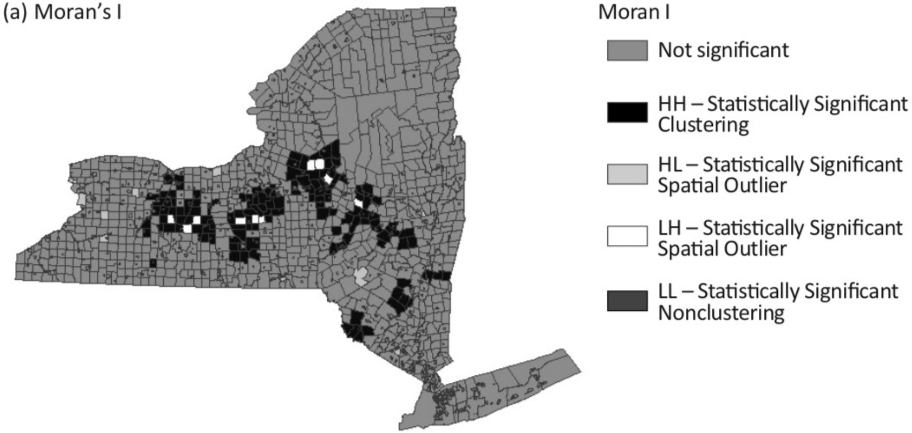


Figure 1. Locations of Bans and Moratoria in New York



DOI: 10.1017/S1068280500010261

REGRESIÓN ESPACIAL

Ejemplos:

- Precios de viviendas dependen de la ubicación, e influyen en el valor de las propiedades vecinas.
- Proyección espacial de instalaciones de paneles fotovoltaicos (energía solar)
- Análisis de regulaciones locales a actividades productivas
- Modelamiento de contaminación ambiental urbana

REGRESIÓN ESPACIAL

Original Research

Accounting for spatial effects in land use regression for urban air pollution modeling

Stefania Bertazzon [a], Markey Johnson [b], Kristin Eccles [c], Gilaad G. Kaplan [d, e]

Show more

Share Cite

https://doi.org/10.1016/j.sste.2015.06.002 Get rights and content

Highlights

- Spatial effects in LUR models are important for heterogeneous pollutants, e.g., NO_2.
- Meteorological variable (wind) improves fit; model is simpler than spatially autoregressive.
- Meteorological variable (wind) model can be specified as geographically weighted model.
- Geographically weighted models assess spatial variability of regression coefficients.
- Adding a meteorological variable (wind) addresses a gap identified in the LUR literature.

Response variable

NO_2 (ppb)

Land use variables

Local roads

Major (arterial) roads

Primary highways

Expressways

Sum of major roads, primary highways, expressways

Sum of primary highways and expressways

Traffic volume

Population density

Dwelling density

Land use: residential

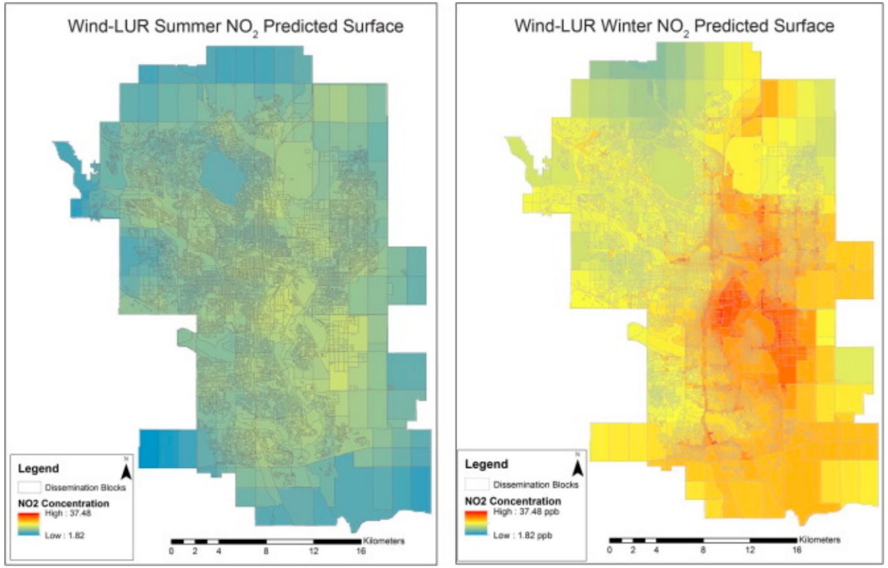
Land use: parks

Land use: institutional

Land use: commercial

Land use: industrial

The coefficients of the summer and winter wind-LUR models were used to predict NO_2 concentrations at the DB level for the entire city of Calgary (Fig. 2).



https://www.sciencedirect.com/science/article/pii/S1877584515000325#f0015

REGRESIÓN ESPACIAL

- Para determinar si existe o no independencia espacial en las observaciones a modelar, podemos aplicar pruebas de autocorrelación.

- Para determinar si un modelo de regresión lineal es una buena representación de la realidad, o si se requiere incorporar el efecto de autocorrelación espacial, podemos analizar si existe un patrón de **autocorrelación en los residuos (Ej: Índice de Moran, LISA)**

- Si la prueba de autocorrelación es significativa → necesitamos un modelo de regresión espacial, **que difiere del modelo lineal tradicional** $\boldsymbol{Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon}$

REGRESIÓN ESPACIAL

En una regresión espacial, se **incluye explícitamente el espacio o contexto geográfico en el modelo.**

- Cuando hay evidencia de que el espacio juega un rol importante en el proceso que queremos modelar
- “space as proxy”: la ubicación funciona como un proxy razonable para otros factores que no podemos pero debiéramos incluir en nuestro modelo.

- Podemos considerar dos tipos de procesos que original efectos espaciales:
  - **Heterogeneidad espacial (SH)**
  - **Dependencia espacial (SD)**

HETEREOGENEIDAD ESPACIAL (SH)

- Aparece cuando no podemos asegurar a ciencia cierta que el proceso que estamos modelando opera bajo las mismas “reglas” en toda la región geográfica de interés.

➔ hay efectos sobre la variable dependiente que están **intrínsicamente ligados a ubicaciones específicas.**

➔ Se considera las **diferencias sistemáticas as través del espacio**, sin recurrir a interdependencias.

Ej: valor de propiedades frente al mar vs. cualquier otra ubicación

HETEREOGENEIDAD ESPACIAL (SH)

¿Cómo incluimos SH en un modelo de regresión?

a) Efectos espaciales fijos: $\boldsymbol{Y = \alpha_r + \beta_1 X_1 + \beta_2 X_2 + \epsilon}$

- La constante $\boldsymbol{\alpha_r}$ puede variar para cada región ➔ hay diferencias de nivel en el resultado $\boldsymbol{Y}$ según la región.
- Se debe definir antes de realizar la regresión.
- La estimación es no-espacial, sigue el método OLS normal.

HETEREOGENEIDAD ESPACIAL (SH)

¿Cómo incluimos SH en un modelo de regresión?

b) Regímenes espaciales: en lugar de asumir que la variable dependiente se comporta uniformemente para todo el espacio, se asume que que **hay efectos sistemáticos que siguen un patrón espacial** y afectan su comportamiento.

- Se consideran diferencias de nivel en el resultado $\boldsymbol{Y}$ + efectos de la variables exógenas ($\boldsymbol{X_i}$) en el resultado.

- Tanto el término constante como los otros coeficientes pueden variar regionalmente : $\boldsymbol{Y} = \boldsymbol{\alpha_r} + \boldsymbol{\beta_{1r}}\, \boldsymbol{X_1} + \boldsymbol{\beta_{2r}}\, \boldsymbol{X_2} + \boldsymbol{\epsilon}$

Pysal – spreg (https://pysal.org/spreg/api.html)

Regimes Models

Regimes models are variants of spatial regression models which allow for structural instability in parameters. That means that these models allow different coefficient values in distinct subsets of the data.

`spreg.OLS_Regimes` (y, x, regimes[, w, ...])	Ordinary least squares with results and diagnostics.
`spreg.TSLS_Regimes` (y, x, yend, q, regimes[, ...])	Two stage least squares (2SLS) with regimes.
`spreg.ML_Lag_Regimes` (y, x, regimes[, w, ...])	ML estimation of the spatial lag model with regimes (note no consistency checks, diagnostics or constants added) [Ans88].
`spreg.ML_Error_Regimes` (y, x, regimes[, w, ...])	ML estimation of the spatial error model with regimes (note no consistency checks, diagnostics or constants added); [Ans88]
`spreg.GM_Lag_Regimes` (y, x, regimes[, yend, ...])	Spatial two stage least squares (S2SLS) with regimes; [Ans88]
`spreg.GM_Error_Regimes` (y, x, regimes, w[, ...])	GMM method for a spatial error model with regimes, with results and diagnostics; based on Kelejian and Prucha (1998, 1999) [KP98] [KP99].
`spreg.GM_Error_Het_Regimes` (y, x, regimes, w)	GMM method for a spatial error model with heteroskedasticity and regimes; based on Arraiz et al [ADKP10], following Anselin [Ans11].
`spreg.GM_Error_Hom_Regimes` (y, x, regimes, w)	GMM method for a spatial error model with homoskedasticity, with regimes, results and diagnostics; based on Drukker et al. (2013) [DEP13], following Anselin (2011) [Ans11].
`spreg.GM_Combo_Regimes` (y, x, regimes[, ...])	GMM method for a spatial lag and error model with regimes and endogenous variables, with results and diag-nostics; based on Kelejian and Prucha (1998, 1999) [KP98] [KP99].
`spreg.GM_Combo_Hom_Regimes` (y, x, regimes[, ...])	GMM method for a spatial lag and error model with homoskedasticity, regimes and endogenous variables, with results and diagnostics; based on Drukker et al. (2013) [DEP13], following Anselin (2011) [Ans11].
`spreg.GM_Combo_Het_Regimes` (y, x, regimes[, ...])	GMM method for a spatial lag and error model with heteroskedasticity, regimes and endogenous variables, with results and diagnostics; based on Arraiz et al [ADKP10], following Anselin [Ans11].
`spreg.GM_Endog_Error_Regimes` (y, x, yend, q, ...)	GMM method for a spatial error model with regimes and endogenous variables, with results and diagnostics; based on Kelejian and Prucha (1998, 1999) [KP98] [KP99].
`spreg.GM_Endog_Error_Hom_Regimes` (y, x, yend, ...)	GMM method for a spatial error model with homoskedasticity, regimes and endogenous variables.
`spreg.GM_Endog_Error_Het_Regimes` (y, x, yend, ...)	GMM method for a spatial error model with heteroskedasticity, regimes and endogenous variables, with re-sults and diagnostics; based on Arraiz et al [ADKP10], following Anselin [Ans11].
`spreg.Skater_reg` ([dissimilarity, affinity, ...])	Initialize the Skater_reg algorithm based on [AA21].

DEPENDENCIA ESPACIAL (SD)

En este caso los efectos espaciales están asociados a la **configuración espacial de las observaciones y sus características.**

- La relación entre los valores de las observaciones es función de la distancia entre ellas.

Ej: el valor de una propiedad depende de qué tan cerca esté de otras zonas más/menos valoradas.

- Solución es más compleja, ya que no se cumplen los principios de OLS.

DEPENDENCIA ESPACIAL (SD)

Efectos espaciales exógenos:

Para incorporar la dependencia espacial, se incluye el **lag espacial una de las variables dependientes**:

$$Y = \beta X + \gamma W X + \epsilon = \beta X + \gamma X_{lag} + \epsilon$$

$\boldsymbol{W}$: matriz de pesos espaciales

$X_{lag}^{i} = WX = \sum_j w_{ij} X_i$: valor promedio de X en la vecindad de la ubicación i

- Usualmente se aplica el lag espacial a aquellas variables que pensamos pueden afectar el resultado en una determinada ubicación.
- Ej: una propiedad con piscina, *en un sector con muchas piscinas*, puede tener mayor valor, porque son un atractivo visual.
- Como el efecto espacial es exógeno, se puede **utilizar OLS** para estimar los coeficientes.

DEPENDENCIA ESPACIAL (SD)

Modelos de lag espacial (efectos endógenos):

Para incorporar la dependencia espacial, se incluye el lag espacial de la variable $\boldsymbol{Y}$ en el lado derecho de la regresión:

$$\boldsymbol{Y} = \boldsymbol{\beta X} + \boldsymbol{\rho W Y} + \boldsymbol{\epsilon} = \boldsymbol{\beta X} + \boldsymbol{\rho Y_{lag}}$$

- El valor de $\boldsymbol{Y}$ en la vecindad de la ubicación $\boldsymbol{i}$ es un importante predictor del valor de $\boldsymbol{Y_i}$.

Ej: el valor de una propiedad está influenciado por el valor de las propiedades en el vecindario

- Esta especificación sí viola los principios de OLS, por lo tanto se requiere **otro método de estimación de los coeficientes:**
    - Máxima verosimilitud (ML)
    - IV- GMM
    - Métodos bayesianos

DEPENDENCIA ESPACIAL (SD)

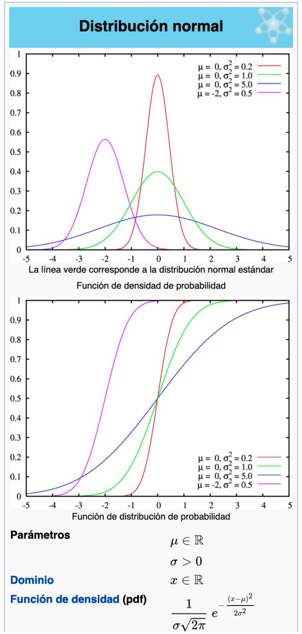
Modelos de error espacial (efectos endógenos):

Hay un efecto espacial en ciertas variables no modeladas

$$Y = \beta_1 X_1 + \beta_2 X_2 + \lambda W\epsilon + \epsilon$$

- En este caso, la dependencia espacial, más que ser consideradara como una posible explicación útil del fenómeno que queremos explicar, es más bien un "problema".
- Esta especificación sí viola los principios de OLS, por lo tanto se requiere **otro método de estimación de los coeficientes.**
    - Máxima verosimilitud (ML)
    - IV- GMM
    - Métodos bayesianos

Análisis de Regresión – Máxima Verosimilitud (ML)



- Consideremos un modelo típico de regresión lineal: $\boldsymbol{Y_i = \alpha + \beta X_i + \epsilon_i}$
- Se asume que los errores $\boldsymbol{\epsilon_i}$ siguen una distribución normal con media 0:

$$\boldsymbol{\epsilon_i \sim N(0, \sigma^2) \Rightarrow f(\epsilon_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\epsilon_i^2}{2\sigma^2}}}$$

$$\boldsymbol{\epsilon_i = Y_i - \alpha - \beta X_i \Rightarrow f(Y_i | \alpha, \beta, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(Y_i - \alpha - \beta X_i)^2}{2\sigma^2}}}$$

- La verosimilitud corresponde a la probabilidad de ocurrencia de una configuración de parámetros:

$$\mathcal{L}(\theta \mid x) = f_\theta(x),$$

- En nuestro caso, la función de verosimilitud de la variable dependiente $\boldsymbol{Y_i}$ para un conjunto de parámetros $(\boldsymbol{\alpha, \beta, \sigma})$ es: $\boldsymbol{L(\alpha, \beta, \sigma | Y_i) = f(Y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(Y_i - \alpha - \beta X_i)^2}{2\sigma^2}}}$

Análisis de Regresión – Máxima Verosimilitud (ML)

- Por lo tanto, la función de verosimilitud conjunta de las respuestas $\boldsymbol{Y_1}, \boldsymbol{Y_2}, \ldots \boldsymbol{Y_n}$ está dada por:

$$\boldsymbol{L(\alpha,\beta,\sigma|Y_1, Y_2, \ldots Y_n) = f(Y_1)\, f(Y_2) \ldots f(Y_n) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(Y_1-\alpha-\beta X_1)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(Y_2-\alpha-\beta X_2)^2}{2\sigma^2}} \ldots = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(Y_i-\alpha-\beta X_i)^2}{2\sigma^2}}}$$

- El método **ML determina los parámetros ajustados** (α, β, σ) **que maximizan la verosimilitud conjunta** $\boldsymbol{L(\alpha,\beta,\sigma|Y_1, Y_2, \ldots Y_n)}$:

$$\boldsymbol{L(\alpha,\beta,\sigma|Y_1, Y_2, \ldots Y_n) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(Y_i-\alpha-\beta X_i)^2}{2\sigma^2}}} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{\frac{\sum_i (Y_i-\alpha-\beta X_i)^2}{2\sigma^2}}$$

Método de Momentos Generalizados (GMM)

Cuando el método de ML no entrega buenos resultados a un determinado ajuste, es posible utilizar también el método de Momentos Generalizados (GMM):

Técnica econométrica genérica de estimación de parámetros de una ecuación de regresión, desarrollada como una extensión del método de los momentos.

Su aplicación es recomendada cuando hay sospechosa de problemas de endogeneidad entre las variables explicativas del modelo y el número de momentos es mayor que el número de parámetros a estimar.

https://github.com/earthlab/tutorials/blob/master/python/intro-to-spatial-regression.ipynb

http://darribas.org/gds_scipy16/ipynb_md/08_spatial_regression.html
https://sustainability-gis.readthedocs.io/en/latest/lessons/L4/spatial_regression.html