

Instituto Tecnológico y de Estudios Superiores de Occidente

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial
15018, publicado en el Diario Oficial de la Federación del 29 de noviembre de 1976.

Departamento de Matemáticas y Física
Maestría en Ciencia de Datos



Análisis espacial de los indicadores socioeconómicos del Índice de Marginación Social en el Área Metropolitana de Guadalajara 2020

Tesis que para obtener el **GRADO** de
Maestro en Ciencia de Datos

Un trabajo presentado por:
Iván Alí Tonatiuh González Novelo

Directores del trabajo:
Dr. Riemann Ruíz Cruz

Tlaquepaque, Jalisco, Diciembre, 2023

Instituto Tecnológico y de Estudios Superiores de Occidente

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial 15018, publicado en el Diario Oficial de la Federación del 29 de noviembre de 1976.

Departamento de Matemáticas y Física Formulario de aprobación de Maestría en Ciencia de Datos

Título de la Tesis: **Análisis espacial de los indicadores socioeconómicos del Índice de Marginación Social en el Área Metropolitana de Guadalajara 2020**

Autor: **Iván Alí Tonatiuh González Novelo**

Tesis aprobada para completar todos los requisitos de grado para la Maestría en Ciencias en Ciencia de Datos.

Director de Tesis, **Dr. Riemann Ruíz Cruz**

Co-Director de Tesis,

Sinodal, **Dr. Alma Nayeli Rodríguez Vázquez**

Sinodal, **Dr. Juan Diego Sánchez Torres**

Revisor Académico, **Dr. Riemann Ruíz Cruz**

Tlaquepaque, Jalisco, Diciembre, 2023

Análisis espacial de los indicadores socioeconómicos del Índice de Marginación Social en el Área Metropolitana de Guadalajara 2020

Iván Alí Tonatiuh González Novelo

Resumen

El presente trabajo busca analizar los indicadores que integran el Índice de Marginación Urbano propuesto por el Consejo Nacional de Población, con base en su relación geográfica. El análisis comprende el estudio de 11 indicadores socioeconómicos que forman el Índice de Marginación Urbano en el Área Metropolitana de Guadalajara (2020) mediante herramientas de econometría, como la autocorrelación espacial utilizando el método de la I de Moran, así como la obtención de los indicadores locales de asociación espacial (LISA). Tanto la I de Moran y el indicador Lisa son útiles para describir patrones espaciales de las áreas geoestadísticas básicas (AGEBs).

The present works aims to analyze the 11 Socio economical indicators that integrates the Urban Marginalization Index on Guadalajara Metropolitan Area (year 2020) using tools for econometrics, such as the spatial auto correlation using the method of Moran I's, as well as calculating the local indicators of spatial association (LISA) in order to describe the features and spatial patterns of the Basic Geo statistics Zones.

Palabras clave

- indicadores socioeconómicos
- análisis espacial
- autocorrelación espacial
- Índice global i de Moran
- Indicador Local de Asociación Espacial (LISA)

Agradecimientos

Agradecer a todos los que formaron parte de mi maestría tomaría demasiado tiempo.

Por lo tanto, mencionaré solamente al Top 7:

- 1.- **Alejandro Cerda Hernández**, no sólo por ayudarme en aspectos técnicos, sino por evitar que me rindiera recordándome que la vida se vive en titán.
- 2.- **Beatriz Novelo**, mi mamá, por todo.
- 3.- **Andrea Noyola**, mi esposa, por soportarme y a veces, hasta amarme.
- 4.- **Omar Pont Torres**, por ser mi compadre y financiar la maestría.
- 5.- **Roberto Henoc Yáñez**, por la inspiración.
- 6.- **Alondra Hernández Navarrete**, por su amistad.
- 7.- **Magnolia Lara Quintero**, por todo su apoyo y confianza.

Muchas gracias, camaradas
Grazie mille, compagni
Spasibo bolshoye tovarishchi

Alí

*Dedicada a la memoria de Mario Lebrao y
Karla Jiménez Haro*

Contenido

	Page
1	Introducción 17
2	Antecedentes 19
3	Justificación 23
4	Problema 25
5	Hipótesis 27
6	Objetivos 29
7	Novedad o aportación 31
8	Estado del Arte o de la Técnica 33
8.1	Estudios realizados por la CONAPO 33
8.2	Indicadores propuestos 35
8.3	La importancia del análisis espacial 36
9	Marco teórico o conceptual 41
9.1	Econometría Espacial 41
9.2	La geografía 41
9.2.1	Autocorrelación espacial 43
9.2.2	Marco Geográfico 43
9.2.3	Antes de los patrones espaciales 44
9.3	Diferenciación entre correlación y autocorrelación 44
9.3.1	Una variable en el espacio 45
9.3.2	Cercanía y lejanía 45
9.3.3	Vecindad 46
9.4	Los índices estadísticos de autocorrelación, criterios de vecindad y matriz de contigüidad 47
9.4.1	Índices de Autocorrelación 49
9.4.2	Matriz W 50
9.4.3	Matriz de pesos estandarizada 53
9.4.4	Los índices de autocorrelación global y local 54
9.4.5	Indicadores locales de Asociación Espacial (LISA) 58
9.5	Algoritmo de clasificación K vecinos más cercanos 62
9.5.1	Calcular KNN: métricas de distancia 63
9.5.2	Métricas de distancia 63
9.5.3	Calcule KNN: definiendo k 64

10	Desarrollo y Resultados	65
10.1	65
11	Conclusiones generales del trabajo	67
11.1	Conclusiones globales	67
11.2	Conclusiones específicas	68
	Bibliography	79

Lista de Figuras

	Page
1.1 El fenómeno de la marginación: Estudios previos demuestran que el acceso al agua es uno de los indicadores más relevantes en cuestión de rezago en una población.	17
2.1 Ejemplo de AGEB: Por su tamaño, esta AGEB se rodea por 17 AGEB vecinas	19
2.2 Índice Absoluto de Marginación: La CONAPO ha realizado estudios sobre la marginación que sirven de antecedentes a la presente investigación	20
2.3 AGEB al 2020: El AMG se conforma por 1,618 áreas geostadísticas básicas	21
7.1 Mapa de Grados de Marginación 2020: Usando los resultados de CONAPO, el mapa muestra los 5 grados marginación. El presente trabajo permite considerar su localización para obtener conclusiones sobre las condiciones sociales que agravan o disminuyen el rezago social.	32
8.1 Índice de calidad de vida (ICdV) de Mar del Plata, 2010: El estudio de Gordziejczuk muestra los niveles de calidad de vida y su patron espacial.	38
9.1 Econometría espacial: De la geografía y la economía surge una disciplina cuantitativa que contempla la localización espacial como variable para su análisis. . . .	42
9.2 Autocorrelación y Correlación: La autocorrelación es univariada, pero horizontal. La correlación requiere al menos dos variables en el mismo espacio (vertical) . . .	45
9.3 Patrón agrupado o clúster: Si la unidad de análisis se rodea por unidades similares a ella, ya sea con valores 1 o con valores 0, se considera que existe un patrón clúster.	47

9.4	Patrón disperso: Si la unidad de análisis se rodea por unidades diferentes a ella, se considera que existe un patrón disperso.	48
9.5	Patrón aleatorio: Si la unidad de análisis se rodea por un número balanceado de unidades similares y diferentes a ella, se considera que existe un patrón aleatorio.	48
9.6	Criterios base de primer orden: Si utilizamos el criterio de contigüidad de primer orden (considerando vecino solamente a los inmediatamente contiguos a la unidad de análisis) y siguiendo los movimientos clásicos del ajedrez.	51
9.7	Criterio de vecindad: Ejemplos de vecindad basados en el criterio distancia.	52
9.8	Matriz <i>W</i> : Ejemplo de matriz de contigüidad física.	52
9.9	Ejemplo de polígonos espaciales: La matriz anterior corresponde a los vecinos de la unidad de análisis etiquetada 7.	53
9.10	Ejemplo de matriz de pesos estandarizada equiponderada: La matriz suma 1 en cada fila.	54
9.11	Importancia de la base para la vecindad: Resultado del cálculo del <i>I</i> de Moran para un mismo fenómeno y marco geográfico considerando los criterios de vecindad Torre y Reina. Nota: se evidencia cómo el cambio de criterio de vecindad afecta sustancialmente el resultado del análisis para el mismo fenómeno y distribución de las unidades espaciales.	57
9.12	Cuadrantes de asociación espacial: los cuadrantes permiten identificar zonas de agrupamiento (hotspots y coldspots) y valores atípicos (spatial outliers) Fuente: basado en Siabato 2018b.	59
9.13	Representación de los cuadrantes de asociación espacial: Los colores intensos representan los clústeres y los opacos las zonas atípicas (spatial outliers).	60

Lista de Tablas

	Page
8.1 Indicadores sociodemográficos a nivel AGEB urbana, 2010. Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.	34
8.2 Indicadores sociodemográficos a nivel AGEB urbana, 2020. Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.	35
8.3 Indicadores propuestos por CONAPO: El Consejo Nacional de Población eligió los 11 indicadores que determinan la marginación social en 2020. Cabe destacar que es la primera vez que se considera la disponibilidad de internet y celular en estudios de IMU.	36
8.4 Dimensiones, variables e indicadores del índice de calidad de vida de Mar del Plata, 2010: El estudio de Gordziejczuk emplea 8 indicadores de calidad de vida. .	37
8.5 Clasificación propuesta para el índice I de Moran de los indicadores seleccionados: El estudio de Gordziejczuk clasifica los resultados de acuerdo a su nivel de significancia.	39
9.1 Criterios de autocorrelación espacial: Los patrones espaciales que encontramos en la teoría, se ha convenido considerarlos significativos cuando cumplen los criterios en la práctica.	58

1 *Introducción*

La marginación es un fenómeno multidimensional y estructural originado por el modelo de producción económica expresado en la desigual distribución del progreso, en la estructura productiva y en la exclusión de diversos grupos sociales, tanto del proceso como de los beneficios del desarrollo (CONAPO, 2011).

La marginación se asocia a la carencia de oportunidades sociales y a la ausencia de capacidades para adquirirlas o generarlas, pero también a privaciones e inaccesibilidad a bienes y servicios fundamentales para el bienestar.



Figura 1.1: El fenómeno de la marginación: Estudios previos demuestran que el acceso al agua es uno de los indicadores más relevantes en cuestión de rezago en una población.

En consecuencia, las comunidades marginadas enfrentan escenarios de elevada vulnerabilidad social cuya mitigación escapa del control personal o familiar, pues esas situaciones no son resultado de elecciones individuales sino de un modelo productivo que no brinda a todos las mismas oportunidades. Las desventajas ocasionadas por la marginación son acumulables, configurando escenarios cada vez más desfavorables.

La gestión de políticas públicas para la construcción de infraestructura en una ciudad en crecimiento como Guadalajara requiere conocer las necesidades de sus habitantes en materia de

educación, salud, vivienda y disponibilidad de bienes.

Estos cuatro aspectos son abordados de manera oficial en el Censo Nacional de Población que se realiza cada 10 años en México, y a partir de la información obtenida, el Consejo Nacional de Población (CONAPO) selecciona un número determinado de indicadores socioeconómicos que describen el grado de marginación social.

En 2010, los investigadores del CONAPO seleccionaron 10 indicadores socioeconómicos (2 relativos a la educación, 2 a la salud, 5 a la vivienda y 1 a la disponibilidad de bienes), mientras que más recientemente en 2020, se modificaron los indicadores y se eligieron 11 variables (2 para la educación, 1 a la salud, 5 a la vivienda y 3 a la disponibilidad de bienes).

Basado en los indicadores seleccionados se realizó el cálculo de un índice global de marginación utilizando una combinación lineal de estos (regresión lineal). Este Índice de Marginación Urbana (IMU) fue calculado a nivel de Área Geostadística Básica (AGEB) y permite sintetizar las carencias en un número simple que puede estratificarse y categorizarse en los siguientes 5 niveles: *Muy baja marginación*, *Baja marginación*, *Media marginación*, *Alta marginación* y *Muy alta marginación*.

La aportación del CONAPO con el IMU 2020 deriva en el presente trabajo, donde se pretende abordar los 11 indicadores socioeconómicos del IMU desde la perspectiva de la econometría, analizando espacialmente cada variable por separado para medir los patrones demográficos y su significancia.

2 Antecedentes

Basados en la organización geográfica de una comunidad, una Área Geostadística Básica (AGEB) urbana es un área geográfica ocupada por un conjunto de manzanas perfectamente delimitadas por calles, avenidas, andadores o cualquier otro rasgo de fácil identificación en el terreno y cuyo uso del suelo es principalmente habitacional, industrial, de servicios, comercial, etcétera, y sólo son asignadas a las localidades urbanas.

Ejemplo de una AGEB se muestra en la figura 8.5 a continuación:

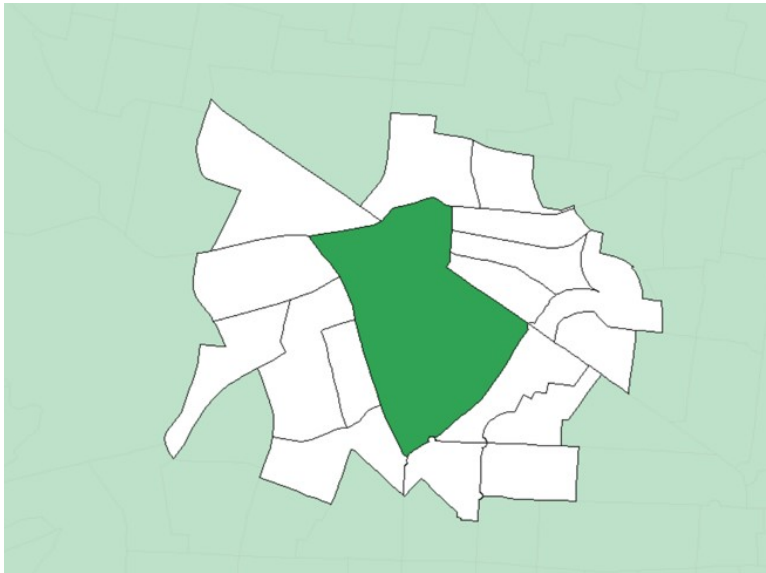


Figura 2.1: Ejemplo de AGEB: Por su tamaño, esta AGEB se rodea por 17 AGEB vecinas

Una localidad urbana es aquella que tiene una población mayor o igual a 2500 habitantes o que es cabecera municipal, independientemente del número de habitantes.

LA ESTIMACIÓN del índice de marginación urbana 2010 tiene como antecedente los ejercicios presentados en 2000 y 2005. El objetivo de estos ejercicios es construir una medida que refleje la situación de las AGEBs, permitiendo identificar al interior de las localidades y municipios de mayor tamaño, las zonas donde se presentan las mayores carencias sociales.



Figura 2.2: Índice Absoluto de Marginación: La CONAPO ha realizado estudios sobre la marginación que sirven de antecedentes a la presente investigación

Los tres ejercicios del índice de marginación urbana presentan diferencias tanto en el empleo de los indicadores socioeconómicos que intervienen en su estimación, como en los universos de estudio, debido a la dinámica que presentan las AGEBs en el país, pues conforme crecen las ciudades, más AGEBs adquieren la categoría de urbanas.

Aunque actualmente (2020) se consideran 4,326 AGEB urbanas en Jalisco, el análisis se realiza con las 1,618 que forman el Área Metropolitana de Guadalajara (integrada por los municipios de San Pedro Tlaquepaque, Tonalá, Zapopan, Tlajomulco de Zúñiga, El Salto, Juanacatlán, Ixtlahuacán de los Membrillos, Acatlán de Juárez, Zapotlanejo y el citado Guadalajara).

En el mapa mostrado en la figura 2.3 se muestran las AGEB actualizadas al 2020:

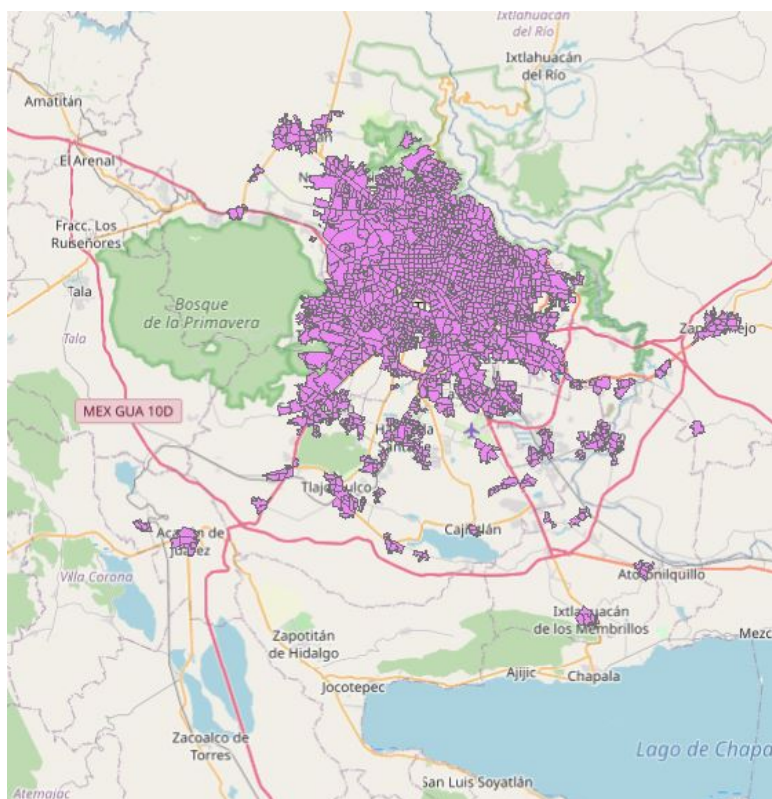


Figura 2.3: AGEB al 2020: El AMG se conforma por 1,618 áreas geostadísticas básicas

3 *Justificación*

El análisis realizado a nivel nacional abarca una problemática global del país. Sin embargo, las regiones, las entidades federativas y las ciudades, con su autonomía presentan condiciones sociales, culturales y económicas muy particulares debido a su autonomía.

La geografía que caracteriza al AMG como las elevaciones, bosques, ríos y lagos, genera la necesidad de analizar la marginación de forma específica.

Además, las técnicas y procesos de Ciencia de Datos proveen un enfoque novedoso al momento de responder las preguntas de desigualdad social.

4 *Problema*

Los municipios, localidades, colonias y AGEBs no se desarrollan económica-socialmente de manera continua y uniforme. El Estado de Jalisco cuenta con un área claramente urbanizada Área Metropolitana de Guadalajara (AMG) , y dos áreas urbanas muy notorias: Puerto Vallarta y Ciudad Guzmán.

El AMG requiere un análisis detallado para conocer las condiciones que lo han definido como urbano, así como las problemáticas derivadas de la falta de infraestructura, la concentración demográfica y la planeación desarticulada de los municipios en cuestión.

En resumen, el problema científico a resolver es describir las relaciones y patrones espaciales de los indicadores socioeconómicos que generan Alta, Media y Baja marginación en el AMG, para comprender las necesidades y carencias de los habitantes que permitan la creación de políticas públicas eficientes.

5 *Hipótesis*

La hipótesis general es que los indicadores cumplen ciertas condiciones de autocorrelación positiva.

Los 11 indicadores socio económicos que conforman el Índice de Marginación Urbana 2020 presentan una relación con la zona espacial en el que se manifiestan, debido a que el fenómeno del rezago social está ligado a ciertas localizaciones geográficas.

6 *Objetivos*

Generales

Describir la relación y patrones espaciales de los indicadores socioeconómicos de marginación urbana que caracterizan las AGEB de marginación muy alta, alta, media, baja y muy baja.

Específicos

- Analizar los 11 indicadores socioeconómicos del IMU2020 por separado para describir su autocorrelación espacial mediante el método de la I de Moran.
- Clasificar las AGEBs de acuerdo a su indicador local de asociación espacial, así como medir su significancia, generar clústeres y categorizar su relación como directa, indirecta o aleatoria.

7 *Novedad o aportación*

ESTE TRABAJO busca ahondar en las condiciones que describen la marginación urbana del Área Metropolitana de Guadalajara (AMG), para comprender los patrones geográficos de los 11 indicadores socioeconómicos, tanto global como localmente.

Utilizando un análisis de autocorrelación espacial, mediante métodos gráficos utilizando lenguaje de programación Python y el software de análisis espacial GeoDa¹, se busca crear una base de conocimiento sólida que además de contemplar las variables y atributos numéricos recabados en el Censo Nacional, contemple la localización espacial de dichas variables.

¹ <https://geodacenter.github.io/>

Con esta aportación, logramos emplear las técnicas de ciencia de datos para allanar el camino a futuras investigaciones que permitan tomar decisiones en política pública y predecir el crecimiento demográfico.

Lo anterior permitirá complementar los planes de desarrollo del Área Metropolitana y el Estado de Jalisco, para enfocar los esfuerzos y recursos en las zonas más rezagadas en materia de educación, salud, vivienda y disponibilidad de bienes.

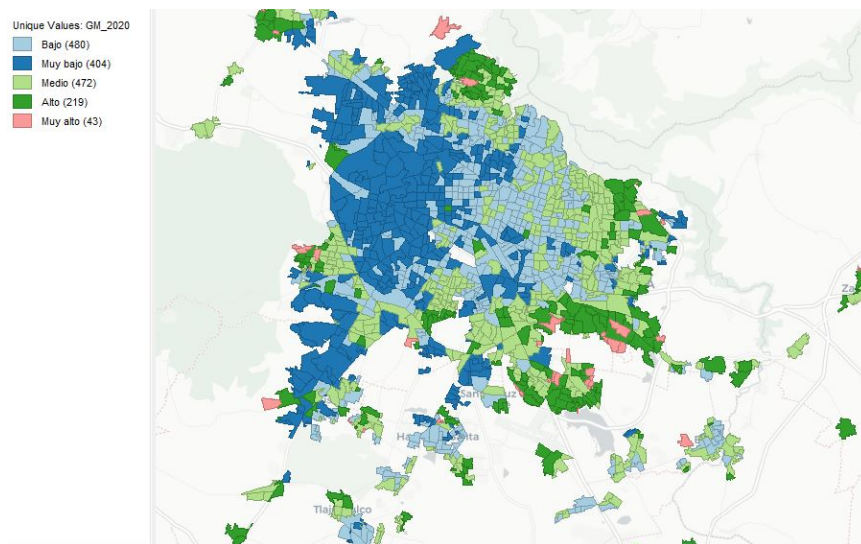


Figura 7.1: Mapa de Grados de Marginación 2020: Usando los resultados de CONAPO, el mapa muestra los 5 grados marginación. El presente trabajo permite considerar su localización para obtener conclusiones sobre las condiciones sociales que agravan o disminuyen el rezago social.

8 *Estado del Arte o de la Técnica*

A NIVEL NACIONAL, en México se han realizado estudios sobre la marginación y la pobreza, pero no desde la perspectiva del análisis espacial.

8.1 *Estudios realizados por la CONAPO*

El Consejo Nacional de Población (CONAPO) es una instancia gubernamental mexicana que tiene por objeto el diseño, operación y evaluación de las iniciativas públicas destinadas a regular el crecimiento de la población, los movimientos demográficos, así como la distribución de los habitantes de México en el territorio.

Como parte de la planeación demográfica del país, el CONAPO tiene como misión incluir a la población en los programas de desarrollo económico y social que se formulan dentro del sector gubernamental, y vincular sus objetivos a las necesidades que plantean los fenómenos demográficos, como la marginación.

El índice de marginación es una medida-resumen que permite diferenciar las Áreas Geoestadísticas Básicas (AGEBs) urbanas del país según el impacto global de las carencias que padece la población como resultado de la falta de acceso a la educación y la salud, la residencia en viviendas inadecuadas y la carencia de bienes.

Para su estimación el CONAPO utilizó como fuente de información los microdatos del Censo de Población y Vivienda 2020.

Por su parte, según la definición de INEGI, una AGEB urbana es un área geográfica ocupada por un conjunto de manzanas perfectamente delimitadas por calles, avenidas, andadores o cualquier otro rasgo de fácil identificación en el terreno y cuyo uso del suelo es principalmente habitacional, industrial, de servicios, comercial, etcétera, y sólo son asignadas al interior de las zonas urbanas que son aquellas con población mayor o igual a 2,500 habitantes y en las cabeceras municipales¹.

Con el índice de marginación es posible identificar las disparidades territoriales existentes en las regiones en un momento determinado,

¹ <https://www.inegi.org.mx/>

una cualidad que le ha conferido un valor relevante como herramienta analítica y operativa para la definición y focalización de políticas públicas, enfocadas al abatimiento de las carencias socio-económicas de la población mexicana.

Desde el año 1990, el CONAPO ha desarrollado una metodología para estimar la marginación social. Sin embargo, en el presente trabajo utilizaremos los datos más recientes del 2020, y los compararemos de ser necesario únicamente con los recabados una década antes, en el 2010.

Desde el 2010, cabe destacar que el Índice de Marginación Urbana (IMU) se ha elaborado mediante una combinación lineal (regresión lineal), utilizando los 10 indicadores como predictores para generar una recta con las siguientes características.

Término	Coefficiente
Constante	-1.70664
% Población de 6 a 14 que no asisten	0.02708
% Población de 15 años o más sin	0.00832
% Sin derechohabencia	0.00648
% Hijos fallecidos	0.04684
%Viviendas sin agua	0.0048
%Viviendas sin drenaje	0.00799
%Viviendas sin excusado	0.0054
%Viviendas con piso de tierra	0.01449
%Viviendas con hacinamiento	0.00852
%Viviendas sin refrigerador	0.00966
Resumen de Modelo	Resultado
S	0.01205
R-sq	0.9997
R-sq(pred)	0.9996

Tabla 8.1: Indicadores sociodemográficos a nivel AGEB urbana, 2010. Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

En la metodología del 2020 no se hace explícito el método de cálculo del IMU, basados en la metodología del 2010, sabemos que también se realiza una regresión lineal para calcular los pesos. El modelo lineal usado se muestra en la Tabla 8.2.

Término	Coefficiente
Constante	128
P6A14NAE	-0.1809
SBASC	-0.06603
PSDSS	-0.07427
OVSDE	-0.283
OVSEE	-0.3314
OVSAE	-0.07605
OVPT	-0.09899
OVSREF	-0.05435
OVHAC	-0.03023
OVSINT	-0.01281
OVSCCEL	-0.07259
Resumen del Modelo	Resultado
S	0
R-sq	100%
R-sq(pred)	100%

Tabla 8.2: Indicadores sociodemográficos a nivel AGEB urbana, 2020. Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

Para estimar los índices de marginación en sus diferentes desagregaciones, se utilizó el Método de Distancia de Pena Trapero (1977)², también conocido como método de medición de distancia P2 o método DP2, desarrollado para hacer comparaciones intertemporales y espaciales. Este método ha sido aplicado en investigaciones sobre calidad de vida en la Unión Europea y España.

El Índice de marginación urbana 2020 se desarrolló con indicadores socio-económicos del Censo de Población y Vivienda 2020, obtenidos a través de una petición de información al Laboratorio de Microdatos del INEGI. Gracias a ello se estimaron los indicadores que representan 11 formas de exclusión de la marginación en las dimensiones: educación, salud, vivienda y disponibilidad de bienes.

² Xesús B Pena Trapero. *Problemas de la medición del bienestar y conceptos afines: (una aplicación al caso español)*. Presidencia del Gobierno, Instituto Nacional de Estadística, 1977

8.2 Indicadores propuestos

Los 11 indicadores propuestos mantienen las bases teórico-conceptuales de los trabajos previos de marginación urbana (CONAPO, 2002, 2009 y 2012); no obstante, en esta actualización algunos indicadores cambiaron su diseño de cálculo.

Respecto al cálculo de los 11 indicadores se estableció a I_j^i como el indicador socioeconómico j , para la AGEB i , con $j = 1, 2, \dots, 11$ e $i = 1, 2, \dots, 50790$ para el ejercicio del 2020.

Se usó el criterio de selección establecido desde el ejercicio 2005, el cual establece que las AGEB urbanas susceptibles de entrar al análisis de marginación deben contar con al menos 20 viviendas particulares habitadas con información de ocupantes (CONAPO, 2009)

El índice de marginación es una de las principales herramientas analíticas y operativas para la definición y focalización de políticas y

	Indicadores socio-económicos
i1	Porcentaje de población de 6 a 14 que no asiste a la escuela.
i2	Porcentaje de población de 15 años o más sin educación básica.
i3	Porcentaje de población sin derechohabencia a los servicios de salud.
i4	Porcentaje de ocupantes en viviendas sin drenaje ni excusado.
i5	Porcentaje de ocupantes en viviendas sin energía eléctrica.
i6	Porcentaje de ocupantes en viviendas sin agua entubada.
i7	Porcentaje de ocupantes en viviendas con piso de tierra.
i8	Porcentaje de ocupantes en viviendas con hacinamiento.
i9	Porcentaje de ocupantes en viviendas sin refrigerador.
i10	Porcentaje de ocupantes en viviendas sin internet.
i11	Porcentaje de ocupantes en viviendas sin celular.

estrategias sociales que buscan abatir las carencias socio-económicas entre la población mexicana.

Sin embargo, es importante profundizar en los aspectos geográficos, estadísticos y espaciales de cada uno de los indicadores, para tener claridad sobre su significado y las relaciones espaciales que definen el fenómeno de la marginación urbana.

8.3 *La importancia del análisis espacial*

Con el fin de avanzar en la construcción de conocimiento sobre el fenómeno social urbano en una de las tres principales ciudades de México, a partir de la consideración de los indicadores de rezago social, en este trabajo se parte de la hipótesis de que el método de análisis de autocorrelación espacial ofrece las competencias necesarias para abordar la problemática planteada.

Al respecto podemos mencionar que dicho procedimiento está relacionado con la primera ley de la geografía de Tobler³, según la cual, en el espacio geográfico, “todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las cosas lejanas” (Tobler 1970, 236).

La motivación entonces responde al hecho de que el CONAPO en México trabaja igual desde 2010, e inspirado en estudios latinoamericanos, se decidió utilizar un enfoque distinto al incorporar la perspectiva geoestadística y técnicas de análisis espacial.

Los análisis basados en autocorrelación espacial permiten conocer el grado de agrupamiento, dispersión o aleatoriedad entre unidades espaciales vecinas, pudiéndose aplicar en escenarios univariado y bivariado.

Existe un estudio [(Gordziejczuk,2019)]⁴ donde se calcula la

Tabla 8.3: Indicadores propuestos por CONAPO: El Consejo Nacional de Población eligió los 11 indicadores que determinan la marginación social en 2020. Cabe destacar que es la primera vez que se considera la disponibilidad de internet y celular en estudios de IMU.

³ Harvey J Miller. Tobler's first law and spatial analysis. *Annals of the association of American geographers*, 94(2):284-289, 2004

⁴ Matías Adrián Gordziejczuk and Patricia Iris Lucero. Turismo y calidad de vida: un estudio de autocorrelación espacial aplicado a la ciudad de mar del plata, provincia de buenos aires, argentina. *Cuadernos de Geografía: Revista Colombiana de Geografía*, 28(1):23-42, 2019

autocorrelación espacial aplicada a la ciudad de Mar del Plata, en la provincia de Buenos Aires, Argentina.

En este artículo se indaga la existencia de autocorrelación espacial bivariada entre indicadores de turismo y calidad de vida, así como el grado de autocorrelación de cada variable. Las unidades espaciales de análisis son las 76 fracciones censales de Mar del Plata.

Los indicadores utilizados son las pernoctaciones hoteleras anuales, la cantidad de viviendas de presunto uso turístico, y un índice de calidad de vida compuesto por cuatro dimensiones y ocho indicadores obtenidos a partir del censo de población argentino del 2010.

Los resultados de la investigación muestran la configuración de dos zonas de autocorrelación espacial positiva dentro de la ciudad, lo que valida el supuesto de que existe una relación entre la existencia de equipamientos turísticos y una más alta calidad de vida de los residentes de las zonas dedicadas al turismo.

Así pues, el objetivo del trabajo de Gordziejczuk es indagar si existe autocorrelación espacial bivariada entre los indicadores de turismo y la calidad de vida.

Los indicadores que se utilizaron fueron los siguientes:

Dimensión	Variable	Indicador
Educación	Máximo nivel	% de población de 20 a 59 años con nivel secundario o polimodal completo.
	educativo logrado	% de población de 26 a 59 años con nivel universitario completo.
Saneamiento	Conexión de agua dentro de la vivienda	% de población en hogares con conexión de agua dentro de la vivienda.
	Conexión a cloaca o cámara séptica y pozo ciego	% de población en hogares con conexión a cloaca o con conexión a cámara séptica y pozo ciego.
Vivienda	Hacinamiento	% de población en hogares sin hacinamiento (dos o menos personas por cuarto).
	Calidad de los materiales (INMAT)	% de población en hogares con INMAT-1: materiales resistentes y sólidos en el piso y en el techo, con ciclrraso.
Actividad económica	Condición de actividad	% entre la población ocupada y la población de 14 años y más.
		% de jefes mayores de 64 años en situación de inactividad.

Tabla 8.4: Dimensiones, variables e indicadores del índice de calidad de vida de Mar del Plata, 2010: El estudio de Gordziejczuk emplea 8 indicadores de calidad de vida.

Habiendo definido las variables e indicadores correspondientes a cada dimensión de análisis, se crea una matriz de datos originales (mdo) para procesar la información. Todos los indicadores seleccionados corresponden a modalidades de beneficio, de tal manera que luego son correctamente resumidas en la medida sintética.

Los números índice definitivos son llevados a un sistema de

información geográfica (sig) —en este caso, Qgis—, realizando visualizaciones en mapas de las variables que describen la calidad de vida, los cuales permiten apreciar su distribución espacial.

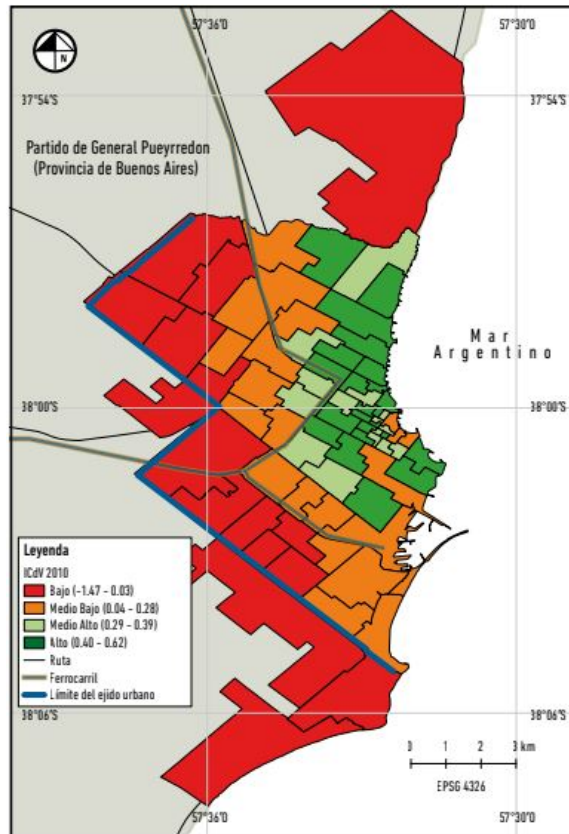


Figura 8.1: Índice de calidad de vida (ICdV) de Mar del Plata, 2010: El estudio de Gordziejczuk muestra los niveles de calidad de vida y su patrón espacial.

Respecto del turismo, se consideran los indicadores denominados pernoctaciones hoteleras anuales (PHA) en adelante, y cantidad de viviendas de presunto uso turístico en adelante (VPUT).

El conjunto de datos se transforma en una matriz de datos índice (mdi), con el fin de establecer la proporción de cada indicador por unidad espacial.

Posteriormente, dado que los datos originales pueden obedecer a unidades de medida diferentes (población, hogares, viviendas o superficies), se procede a la normalización de las variables y a la construcción de una matriz de datos estandarizados (mdz).

En cuanto a los resultados, los primeros tres análisis son univariados y se realizan para conocer el grado de autocorrelación espacial de cada variable consigo misma, estableciéndose para ello el siguiente orden: ICdV, PHA y VPUT. El cuarto, quinto y sexto examen son bivariados

y se ejecutan para indagar si existe autocorrelación espacial entre los indicadores de turismo y el de calidad de vida, así como también entre las dos variables turísticas (ICdV-PHA, ICdV-VPUT; PHA-VPUT).

Autocorrelación espacial	Índice I de Moran	Caso analizado	
		Univariado	Bivariado
No significativa	0.0000 – 0.2999		
Poco significativa	0.3000 – 0.4999		ICdV-PHA ICdV-VPUT
Medianamente significativa	0.5000 – 0.6999	ICdV-ICdV PHA-PHA	PHA-VPUT
Muysignificativa	0.7000 – 1.0000	VPUT-VPUT	

Tabla 8.5: Clasificación propuesta para el índice I de Moran de los indicadores seleccionados: El estudio de Gordziejczuk clasifica los resultados de acuerdo a su nivel de significancia.

Solo el ensayo univariado correspondiente al indicador de cantidad de VPUT ha logrado una medida global de autocorrelación entre unidades espaciales vecinas muy significativo.

Los otros dos análisis univariados y el bivariado entre PHA y VPUT arrojaron medidas medianamente significativas. Por su parte, el índice de calidad de vida (ICdV), en su análisis de autocorrelación respecto de los indicadores de turismo, ha alcanzado los valores menos significativos. En este sentido, vale destacar, en primer lugar, que ninguna prueba ha lanzado un índice I de Moran no significativo y, en segundo lugar, que, en relación con la implementación del método lisa sobre patrones locales, se ha logrado evidenciar la autocorrelación espacial negativa y la segregación urbana que reviste Mar del Plata. Si bien se tiene en cuenta que no pocas investigaciones llegan a la conclusión de la existencia de significativas disparidades intraurbanas en Mar del Plata, en este estudio se han podido relacionar los niveles de calidad de vida de la población local con algunos elementos que constituyen la estructura y funcionamiento de la actividad turística en la ciudad.

De esta manera, resulta evidente el hecho de que la variación de los fenómenos no se distribuye de forma aleatoria sobre el espacio geográfico y que existe cierta dependencia espacial entre las unidades contiguas, puesto que tanto los ensayos de autocorrelación espacial univariada como bivariada han respondido a un mismo patrón de distribución espacial.

Desde los sectores centrales hacia los periféricos, Mar del Plata transita de áreas centrales intensamente utilizadas con fines turísticos, que devienen en heterogéneas desde el punto de vista del bienestar social, a zonas urbanas de escaso o nulo interés turístico y valores de calidad de vida significativamente menores.

Asimismo, el estudio concluye que en medio se encuentran recortes espaciales de buena calidad de vida ubicados en proximidad al área turística tradicional.

9 Marco teórico o conceptual

En este capítulo se introducen los conceptos teóricos y matemáticos utilizados en este trabajo. En la sección 9.1, se incluyen los elementos para entender la importancia y la estructura de la econometría espacial.

9.1 Econometría Espacial

El origen de la econometría como disciplina surge en las primeras tres décadas del siglo XX, cuando el economista sueco Ragnar Frisch publicó el término por primera vez en 1926¹. En 1996, Maddala propuso que se definiera la econometría como: la aplicación de métodos estadísticos y matemáticos al análisis de datos económicos con el propósito de dar contenido empírico a las teorías económicas y verificarlas o refutarlas.

Más allá de la etimología, la econometría surge de la intersección entre la teoría económica, la economía matemática, la estadística económica y la estadística matemática. (fuentes) La teoría económica tiene un carácter fundamentalmente cualitativo, por lo que no proporciona medidas numéricas para la relación que puede existir entre las variables. Así destacaremos una aportación importante de la econometría al proporcionar valores numéricos que permiten cuantificar dicha relación.

¹ Ragnar Frisch, Olav Bjerkholt, and Ariane Dupont-Kieffer. *Problems and methods of econometrics: the Poincaré lectures of Ragnar Frisch 1933*. Routledge, 2009

9.2 La geografía

En sentido amplio, la geografía es la ciencia que estudia la superficie terrestre, las sociedades que la habitan y los territorios, paisajes, lugares o regiones que la forman al relacionarse entre sí. De ella se desprende la geografía humana, considerada la ciencia social centrada en el estudio de las sociedades y de sus territorios, cuyo objeto es el ser humano y sus reacciones con su entorno tanto en el aspecto estático de su organización, como en el dinámico de los cambios que experimentan. A su vez, la geografía humana engloba la geografía urbana, encargada del estudio de las ciudades y las regiones urbanas, su morfología (plano, estructura, edificación, sectores, procesos ecológicos), sus características socio-económicas, sus cambios y problemas.

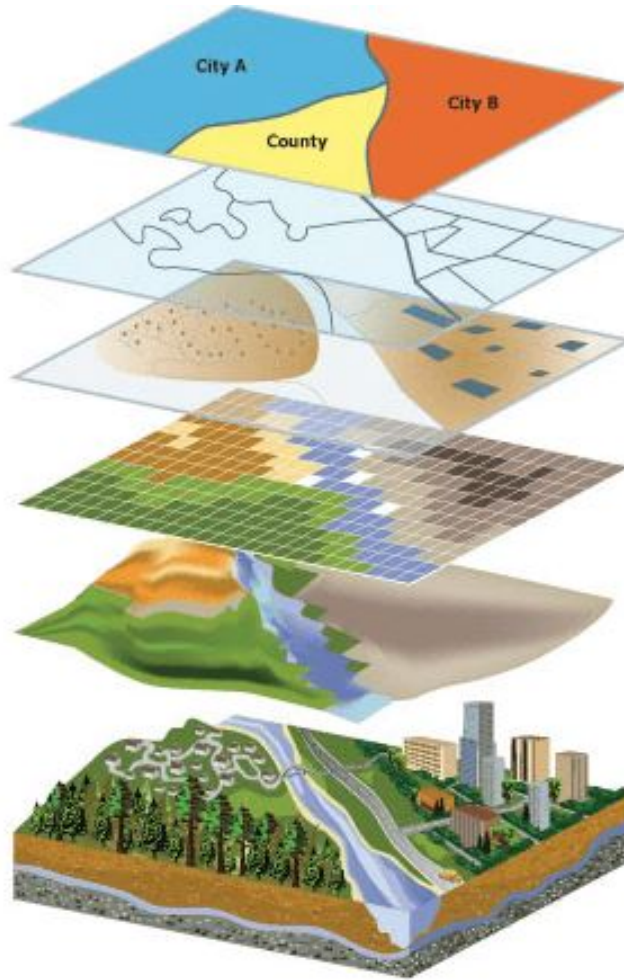


Figura 9.1: Econometría espacial: De la geografía y la economía surge una disciplina cuantitativa que contempla la localización espacial como variable para su análisis.

A partir de la década de los setenta, la geografía ha experimentado un cambio significativo gracias al desarrollo e implementación de un conjunto de metodologías de análisis espacial que han permitido afrontar, analizar y comprender problemas que anteriormente no se podían abordar. Entre los elementos que han aportado a este desarrollo resaltamos: (i) la econometría espacial; (ii) los aportes de Waldo Tobler con la primera ley de la geografía; y (iii) la formalización del concepto de autocorrelación espacial.

En 1970, paralelamente la econometría y la geografía encontraron una mutua conexión que derivó en lo que conocemos como econometría espacial, un campo especializado de la econometría general que considera específicamente los efectos espaciales (distribución y comportamiento a través del espacio geográfico) de las variables analizadas (fenómenos geográficos). A inicios de la década de los setenta, se aplicaron las técnicas para describir escenarios como el

urbanismo, los modelos de precios hedónicos y la geografía económica. El geógrafo belga Luc Anselin señaló que la econometría espacial es un “conjunto de técnicas que considera las debilidades metodológicas derivadas de involucrar explícitamente los efectos espaciales, en específico: la autocorrelación y la heterogeneidad espacial”. En 1989, Anselin también explica el análisis espacial como el estudio cuantitativo de fenómenos geográficos que se distribuyen en el espacio.

9.2.1 *Autocorrelación espacial*

El término autocorrelación espacial tiene su origen en el análisis de series temporales y, a su vez, en la noción de correlación de la estadística univariada. En pocas palabras, la autocorrelación mide el grado de asociación que una variable desarrolla a través de un espacio definido como marco geográfico, esto es, el espacio de análisis propiamente dicho.

9.2.2 *Marco Geográfico*

El marco geográfico estará delimitado por el problema que se requiere abordar y puede ser simple o compuesto. El marco es simple si no presenta discontinuidades en el espacio geográfico, por ejemplo, al analizar los niveles de producción de sorgo en un conjunto de predios continuos localizados en una zona rural; en caso contrario, el marco es compuesto, por ejemplo, cuando se analiza la distribución de la población en el archipiélago de 7107 islas que constituye la República de Filipinas.

Siabato, Willington, y Jhon Guzmán-Manrique (2019) enuncian: “La esencia de la autocorrelación es analizar la variabilidad de un fenómeno a través del espacio geográfico para determinar patrones espaciales⁴ y describir su comportamiento, es decir, puede ser entendida como el medio para comprender cómo se distribuye el fenómeno en el espacio analizado y en qué grado los elementos locales pueden verse afectados por sus vecinos. “ Resulta indudable que la primera ley de la geografía de Tobler refleja el sentido de la autocorrelación espacial al establecer que “todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las lejanas”⁵ (Tobler 1970, 236).

En el presente trabajo, las complejas dinámicas de las sociedades humanas deberán ser analizadas con cautela y rigor científico para acotar el significado de las relaciones en una zona y tiempo determinada, puesto que si existe una relación entre los elementos del fenómeno analizado en un marco geográfico (estructura espacial), sería posible identificar un patrón espacial.

9.2.3 *Antes de los patrones espaciales*

Siabato (2019) enfatiza que la estadística tradicional no contaba con las herramientas necesarias para incorporar los factores espaciales: “Inicialmente, los análisis no espaciales se enfocaban exclusivamente en los valores (medidas) de los fenómenos, sin tener en cuenta, de manera implícita, la localización espacial.” Entonces tenemos que el análisis espacial comprende dos tipos bien diferenciados de información: primero, los atributos de los fenómenos espaciales analizados, que incluyen medidas como el nivel de precipitación, la población desplazada, los niveles de contaminación sonora, caudal de los ríos, o también los atributos de variables cualitativas, tales como el uso del suelo o grupo religioso; segundo, que cada fenómeno espacial tiene una localización, que se puede caracterizar por su posición sobre un mapa, por sistemas coordenados o por múltiples referencias geográficas. El tercero de los tipos sería el tiempo, que en todo sentido puede tomarse como una dimensión más cuya distancia influye en el fenómeno, pero que en el presente trabajo no consideraremos por tratarse de un instante determinado en el que se tomó la captura de los atributos (2020).

9.3 *Diferenciación entre correlación y autocorrelación*

A partir de la aplicación y el desarrollo del análisis espacial y como producto de los exitosos métodos implementados en diversas investigaciones a través del tiempo, surge la confusión para distinguir (ya en el terreno de la econometría espacial) los conceptos de correlación espacial (que a su vez podría confundirse con la correlación a secas) y la autocorrelación espacial, que es un término fundamental de esta investigación. La correlación espacial se emplea generalmente en métodos como una medida cuantitativa de asociación entre dos variables. En este sentido, puede tratarse de variables cuya tendencia conjunta es *positiva*, cuando su relación es directa y ocurren simultáneamente creciendo o decreciendo; *negativa*, cuando su relación es inversa, y el incremento de una ocurre al decrecer la otra; o sin correlación, cuando podríamos hablar de que ninguna sigue tendencia con respecto a la otra o que son aleatorias entre sí.

La autocorrelación espacial que es el concepto más empleado en esta investigación se aplica principalmente para la determinación de patrones espaciales y en análisis locales y globales, los cuales serán profundizados a continuación. Cabe resaltar que la econometría espacial utiliza métodos que combinan ambos conceptos para complementarse, lo que permite responder preguntas univariadas y multivariadas.

En el afán por distinguir ambos términos, Getis de hecho aclara que “la autocorrelación espacial tiene un significado propio, mientras que las estadísticas de correlación se diseñaron para mostrar las relaciones entre variables, las estadísticas de autocorrelación se diseñaron para mostrar las correlaciones dentro de las variables, y la autocorrelación espacial muestra la correlación dentro de las variables a través del espacio.” (Siabato, 2019)

9.3.1 Una variable en el espacio

La Figura 9.2 pretende aclarar esta confusión enfocando nuestra atención en que la autocorrelación espacial se basa generalmente en una única variable, partiendo de la noción de correlación de la estadística univariada.

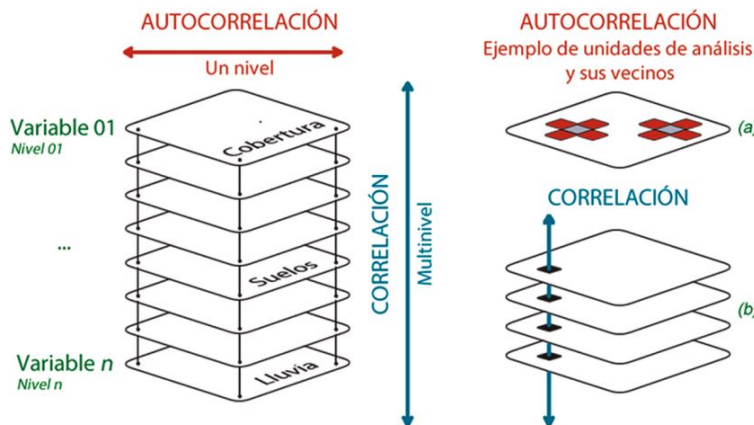


Figura 9.2: Autocorrelación y Correlación: La autocorrelación es univariada, pero horizontal. La correlación requiere al menos dos variables en el mismo espacio (vertical)

La diferencia fundamental entre estos conceptos es que al analizar la autocorrelación de un fenómeno geográfico no se consideran otras variables, sino los valores observados en las zonas vecinas de la variable analizada.

La autocorrelación describe el comportamiento de una única variable considerando un “plano horizontal” que está delimitado por el marco geográfico. Nótese entonces cómo la autocorrelación espacial considera relaciones entre valores de la misma variable, pero en diferentes localizaciones (horizontal), mientras la correlación considera diferentes pares de variables en la misma localización (vertical).

9.3.2 Cercanía y lejanía

Cuando aceptamos la primera ley de la geografía de Tobler, estamos empleando la cercanía y lejanía como base para medir los efectos

de cierto fenómeno. En un sentido general, sabemos que, si la distancia entre dos puntos o unidades geográficas es pequeña, estos se encuentran cerca. Como en cualquier otro problema matemático, tendremos que definir lo que consideramos una distancia pequeña, y termina siendo relativo al problema que buscamos resolver.

Miller (2004) presenta una rica discusión sobre estos conceptos y resalta además la importancia de la primera ley en el análisis espacial. Uno de los criterios de cercanía puede ser entonces definido por aquellas unidades geográficas o unidades de análisis que tienen contigüidad física, criterio comúnmente llamado “vecindad”. Sin embargo, las relaciones de vecindad se basan en diferentes criterios geográficos tales como: la existencia de contigüidad física; la definición de relaciones de distancia; o a través de diferentes criterios de análisis como las redes de transporte, inclusive puede ser expresada en términos temporales. Por tanto, se debe utilizar el término vecindad como criterio general y comprender que la contigüidad es solo uno de los posibles tipos de vecindad que se consideran en el análisis espacial.

9.3.3 *Vecindad*

Generalmente, la vecindad se establece a través de la matriz de contigüidad W , cuyos componentes denotaremos como w_{ij} y representan el criterio de vecindad entre las unidades de análisis i y j . Un ejemplo de la matriz de contigüidad se puede ver en la Figura 9.8. Esta matriz no es solamente un criterio fundamental para el análisis de la autocorrelación espacial, sino para gran parte de los métodos de análisis de los fenómenos geográficos. El criterio más simple de vecindad es la contigüidad física de primer orden, es decir, considerando exclusivamente los vecinos circundantes de la unidad de análisis.

Los fundamentos de este criterio fueron definidos por Moran (1948). Sin embargo, el análisis espacial puede presentar limitaciones para medir el grado de asociación entre unidades que no se encuentran contiguas, es decir, en escenarios en los que el fenómeno analizado se desarrolla en un marco geográfico compuesto. Ante este problema, Cliff y Ord, en sus estudios sobre autocorrelación espacial (1972, 1973) y procesos espaciales (1981), introducen el concepto de distancia entre unidades geográficas, desarrollado como marco de análisis en Getis y Ord (1992). Moreno Serrano y Vayá (2002, 86) presentan una definición concreta de estos conceptos y de la matriz de distancia inversa, desarrollada por Anselin (1980) con el objetivo de determinar la autocorrelación espacial para aquellas unidades de análisis que carecen de contigüidad física directa. Como su nombre lo indica, la asociación entre dos unidades espaciales decrece a medida que la distancia entre

ellas aumenta.

En general, durante las últimas cuatro décadas se han desarrollado múltiples métodos para determinar con mayor precisión el concepto de vecindad entre las zonas analizadas, cada uno de ellos orientado a casos de estudio particulares que luego se convierten en procesos de aplicación general. Es recomendable que los analistas espaciales consideren detenidamente el tipo de matriz de vecindad que utilizarán, porque las conclusiones que se pueden derivar del análisis realizado, si consideran un tipo u otro, pueden llegar a ser opuestas. Una elección inapropiada de la matriz puede conllevar resultados imprecisos, o poco relacionados con el fenómeno analizado, que pueden generar decisiones incorrectas y la implementación de políticas con impactos adversos. El analista debería plantearse y responder, con criterio y pleno conocimiento del fenómeno analizado, esta pregunta: ¿qué tipo de matriz ayudaría a describir y representar mejor el comportamiento del fenómeno geográfico que se está analizando?

En este punto, es claro que la autocorrelación espacial marca un punto de inflexión en la geografía cuantitativa (Cliff y Ord 1969), por lo que se establece como un pilar fundamental para el análisis de la distribución y variabilidad de los fenómenos geográficos (Getis 2007).

9.4 *Los índices estadísticos de autocorrelación, criterios de vecindad y matriz de contigüidad*

La autocorrelación espacial permite comprender la variación de un fenómeno en un marco geográfico de análisis. Si el fenómeno analizado tiende a agruparse en zonas uniformes, es decir, si tiende a conformar conglomerados o clústeres, entonces se evidencia la existencia de autocorrelación positiva.

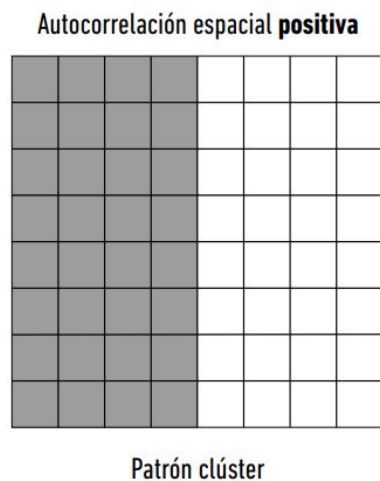


Figura 9.3: Patrón agrupado o clúster: Si la unidad de análisis se rodea por unidades similares a ella, ya sea con valores 1 o con valores 0, se considera que existe un patrón clúster.

Por el contrario, si las medidas de la variable en las unidades colindantes son disímiles, es decir, si el fenómeno tiende a estar disperso, entonces la autocorrelación espacial es negativa.

En este caso, si un atributo está presente en un determinado lugar, este tenderá a ser diferente ($\pm\sigma^2$) en los lugares vecinos.

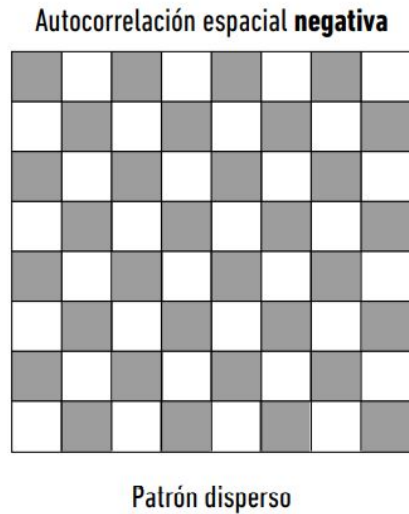


Figura 9.4: Patrón disperso: Si la unidad de análisis se rodea por unidades diferentes a ella, se considera que existe un patrón disperso.

Por último, cuando el fenómeno se comporta de forma aleatoria y no se identifica un comportamiento definido o estructurado, se dice que no existe autocorrelación espacial.

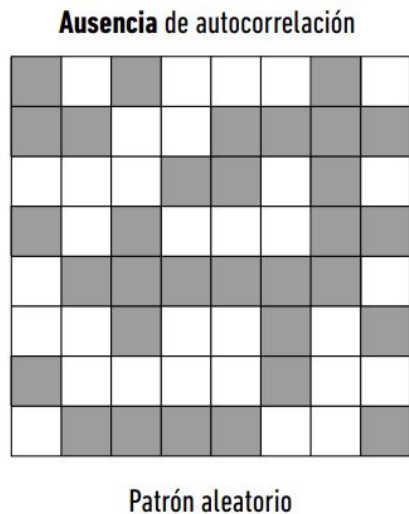


Figura 9.5: Patrón aleatorio: Si la unidad de análisis se rodea por un número balanceado de unidades similares y diferentes a ella, se considera que existe un patrón aleatorio.

En términos prácticos, este último caso implica que la presencia o ausencia de un atributo en un lugar determinado no influye, aparentemente, en la medida de dicho atributo en los lugares vecinos.

Entender este carácter aleatorio es fundamental, porque si imaginamos un atributo o variable, necesitaremos evaluar si su magnitud es resultado de la relación geográfica de cercanía o lejanía, o si dicha magnitud es aleatoria y podría darse de cualquier modo, si de manera artificial cambiáramos la localización de las unidades de análisis. Estos tres comportamientos caracterizan el fenómeno analizado en uno de los tres patrones espaciales básicos: *clúster*, *disperso* o *aleatorio*.

Si la autocorrelación espacial de los patrones mostrados en las Figuras 9.3, 9.4 y 9.5, se evaluaran a través del Índice de Moran, los resultados serían 0.86, -1 y -0.17 respectivamente (tomando como criterio rook o torre de grado 1).

La autocorrelación espacial se interpreta entonces como un índice estadístico descriptivo que permite medir las formas y las maneras como se distribuyen los fenómenos analizados en el espacio geográfico (Goodchild 1986). La autocorrelación espacial mide el grado en el que una variable geográfica está correlacionada con ella misma en dos puntos o zonas diferentes del área de estudio (x_i y x_{i+h} , donde h es la distancia entre los puntos), mide la similitud de la variable temática en un área determinada.

En este sentido, la autocorrelación espacial permite responder la pregunta: **¿cómo varía un fenómeno a través del espacio geográfico?**

9.4.1 Índices de Autocorrelación

La formalización de los primeros índices para medir la similitud de un fenómeno geográfico surge a finales de la década de 1940 e inicios de los cincuenta.

El *Índice de Moran*, uno de los más conocidos y extendidos, está fundamentado en los trabajos de Moran (1948) y Krishna Iyer (1949), y fue potenciado posteriormente por Geary (1954)¹¹, aunque su implementación en diferentes campos del saber, como en la economía regional, la economía urbana y los modelos de precios, requirió ajustes y complementos que permitieron su generalización y evolución dando paso a nuevos conceptos. Uno de los avances más importantes ha sido la capacidad de comprender no solamente las variaciones globales sino también las locales, esta idea marca uno de los hitos fundamentales en el desarrollo del análisis espacial puesto que los comportamientos globales suelen esconder aquellas especificidades locales a partir de las cuales se suelen resolver los problemas planteados; Anselin (1995) presenta el desarrollo y formalización de esta idea conocida como indicadores locales de asociación espacial —en adelante *LISA*, acrónimo de Local Indicators of Spatial Association.

De acuerdo con Getis (2007, 494), entre los índices que permiten analizar la autocorrelación espacial global y local se destacan:

- I de Moran (representación de covarianza global).
- c de Geary (representación de diferencias globales).
- G de Getis y Ord (representación multiplicativa global).
- I_i y c_i de Anselin (indicadores locales de asociación espacial –LISA–).
- O de Getis y Ord (representación local que considera la autocorrelación global).

Los índices de autocorrelación están determinados de forma general por la expresión de la Ecuación (9.1), donde w_{ij} es el componente que determina la relación geográfica y C_{ij} el que explica la variación temática del fenómeno analizado (Siabato 2018b).

$$\Gamma = \sum_{i=1}^n w_{ij} C_{ij} \quad (9.1)$$

La ecuación de gamma representa una técnica muy flexible que permite vincular diferentes criterios de asociación espacial a través de los distintos tipos de matriz W , como los presentados en la lista anterior. El cálculo de los índices se puede resumir en la sumatoria del producto de cada elemento w_{ij} y su correspondiente C_{ij} .

9.4.2 Matriz W

La definición de la matriz $W = [w_{ij}]$ utiliza dos criterios básicos de vecindad: la contigüidad física y la distancia. Estos criterios se clasifican como básicos porque son los más extendidos y los que suelen estar implementados en los programas informáticos de análisis espacial, sin embargo, no se deben excluir los criterios basados en funciones de análisis y en relaciones temporales. Para el primer caso, la matriz de contigüidad física se basa en criterios de contigüidad espacial de orden n , la Figura 9.6 muestra ejemplos básicos de primer y segundo orden para un marco geográfico dividido en unidades espaciales cuadradas y distribuidas uniformemente. Para determinar el primer orden, se responde a la pregunta ¿quiénes son mis vecinos? Para determinar las unidades asociadas de segundo orden, la pregunta sería ¿quiénes son los vecinos de mis vecinos? Este ejercicio se replica para cada orden $n + 1$ que se quiera determinar. Como se indica en la Figura 9.6, los criterios base corresponden a los movimientos clásicos del juego de ajedrez.

Por otro lado, el criterio de distancia establece quiénes son vecinos a partir de un límite máximo de vecindad determinado por una distancia d , lo importante en este escenario es definir claramente cómo se mide d . Es habitual considerar d entre los centroides de las unidades espaciales, pero también se pueden determinar otros criterios como los centros

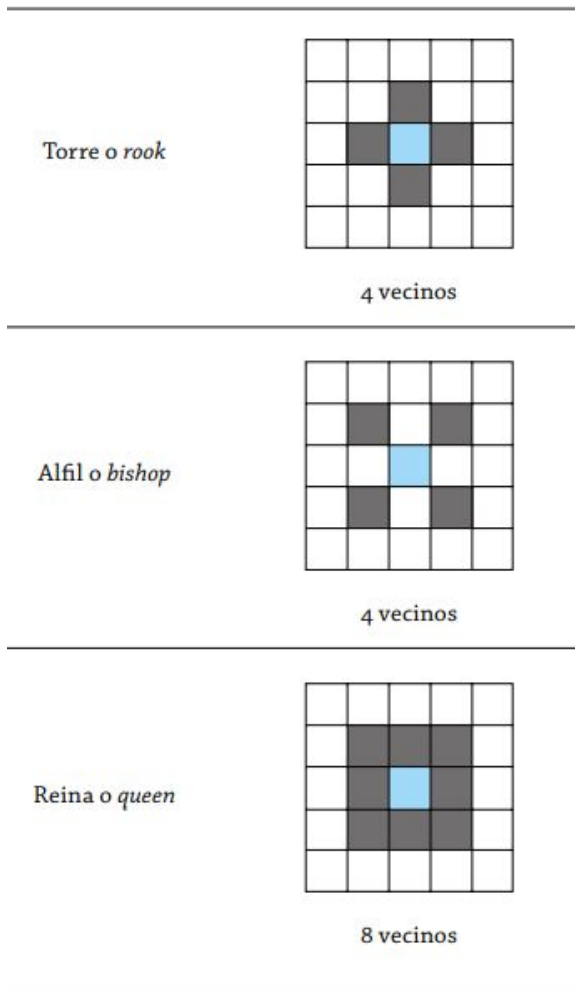


Figura 9.6: Criterios base de primer orden: Si utilizamos el criterio de contigüidad de primer orden (considerando vecino solamente a los inmediatamente contiguos a la unidad de análisis) y siguiendo los movimientos clásicos del ajedrez.

medios, los anillos de influencia, las relaciones de proporción perimetral o las proporciones entre líneas perimetrales no colindantes.

Por ejemplo, la Figura 9.7 muestra los vecinos de la unidad central (en azul) determinados a partir de distancias de $2000m$, $3000m$, $4000m$ y $5000m$ entre centroides. Las unidades de análisis del ejemplo son cuadrados de $2000m$ de lado. Es claro que al definir como parámetro de vecindad una distancia de $1000m$, no se encontrarían vecinos, la distancia mínima entre centroides es $2000m$. Los ejemplos muestran las claras diferencias que hay entre los criterios de vecindad.

Se debe destacar que la distancia agrupa todos los elementos, mientras que la contigüidad conforma anillos según el orden determinado. Estos criterios parecen contraintuitivos, por tanto, el analista espacial debe decidir si el análisis de autocorrelación espacial debe considerar el conjunto de vecinos o solamente las franjas (anillos) definidas por el orden n de contigüidad. No obstante, se debe

considerar que cada método puede tener modificaciones, por ejemplo, las matrices de contigüidad de orden dos y superiores pueden incluir los órdenes inferiores cambiando dramáticamente la definición de vecinos.

Ejemplos como los planteados en las Figuras 9.6 y 9.7 pueden ser representados para cada uno de los conceptos listados por Getis (2007, 494).

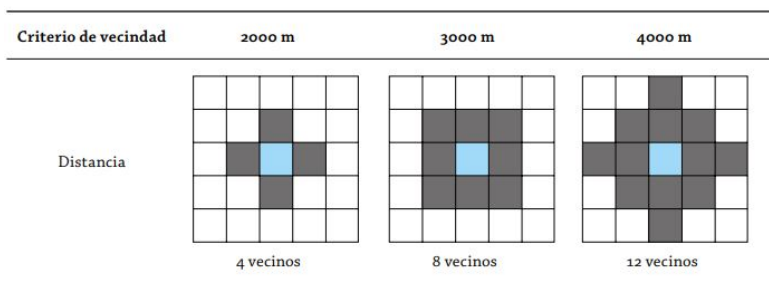


Figura 9.7: Criterio de vecindad: Ejemplos de vecindad basados en el criterio distancia.

W es una matriz cuadrada $n \times n$ (véase Ecuación 9.1), en la que n se corresponde con el número de unidades de análisis.

La Figura 9.8 muestra la definición de la matriz W para un marco geográfico irregular conformado por 10 unidades de análisis, considerando como criterio de vecindad Reina de primer orden (contigüidad física).

	1	2	3	4	5	6	7	8	9	10
1	0	1	1	0	0	0	0	0	0	0
2	1	0	1	0	1	0	0	0	0	0
3	1	1	0	1	1	1	0	0	0	0
4	0	0	1	0	0	1	1	1	0	0
5	0	1	1	0	0	1	0	0	0	0
6	0	0	1	1	1	0	1	0	1	0
7	0	0	0	1	0	1	0	1	1	0
8	0	0	0	1	0	0	1	0	1	1
9	0	0	0	0	0	1	1	1	0	1
10	0	0	0	0	0	0	0	1	1	0

Figura 9.8: Matriz W : Ejemplo de matriz de contigüidad física.

La Figura 9.9 muestra las unidades de análisis clasificadas por área representadas con una paleta divergente para facilitar su visualización.

En el ejemplo se definen los vecinos de la unidad de análisis 7, la Figura 9.9b resalta en tono oscuro sus cuatro vecinos: 4, 6, 8 y 9. Se podría pensar que la unidad 10 es también vecina, pero no es así.

Es importante no dejarse llevar por impresiones visuales y considerar las relaciones topológicas de las geometrías que se involucran en el

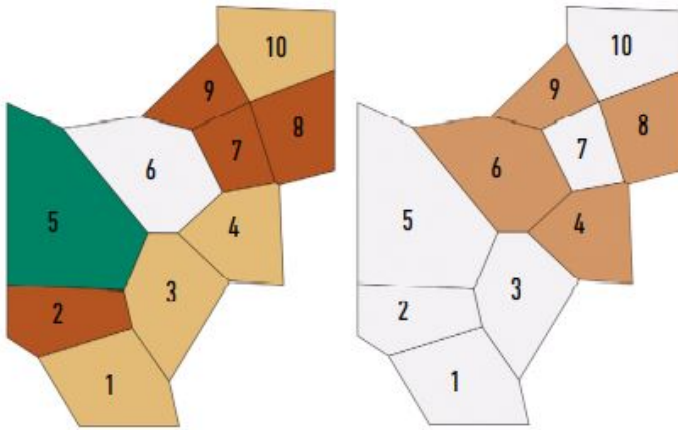


Figura 9.9: Ejemplo de polígonos espaciales: La matriz anterior corresponde a los vecinos de la unidad de análisis etiquetada 7.

análisis; en este caso, la unidad 7 no comparte nodos con la 10.

Una vez identificados los vecinos, se deben registrar en la matriz tal y como se muestra en la Figura 9.8. Nótese que en las intersecciones de W para cada unidad de análisis (es decir, la posición ij de la matriz) se registra 1 para identificar la colindancia y 0 para establecer su ausencia. Al repetir este ejercicio para cada unidad de análisis se completa la matriz.

Las filas de la matriz establecen los vectores de contigüidad y, evidentemente, el conjunto de vectores conforma la matriz, esta aclaración es importante porque los índices globales consideran la matriz, mientras que los locales utilizan los vectores para definir el índice de cada unidad. El vector de contigüidad determina la vecindad de cada unidad de análisis.

Algunas propiedades de la matriz de contigüidad resultan evidentes con este ejemplo: (1) su diagonal principal se conforma de ceros, esto implica que la unidad no es vecina de sí misma; (2) la matriz es simétrica; (3) el número de vecinos de cada unidad de análisis está determinado por la sumatoria de cada vector, es decir, el número de vecinos de la unidad i corresponde a la suma de la i -ésima fila de la matriz; (4) las unidades deben tener un identificador único para establecer las relaciones de forma unívoca.

9.4.3 Matriz de pesos estandarizada

Finalmente, se presenta el concepto de matriz de pesos estandarizada. A diferencia de la matriz de contigüidad, la matriz estandarizada no es simétrica, pues los pesos de cada vector varían en función del conjunto de vecinos (véase Figura 9.10).

La sumatoria de los vectores locales o de contigüidad (filas) debe ser

	1	2	3	4	5	6	7	8	9	10
1	0	0.50	0.50	0	0	0	0	0	0	0
2	0.33	0	0.33	0	0.33	0	0	0	0	0
3	0.20	0.20	0	0.20	0.20	0.20	0	0	0	0
4	0	0	0.25	0	0	0.25	0.25	0.25	0	0
5	0	0.33	0.33	0	0	0.33	0	0	0	0
6	0	0	0.2	0.20	0.20	0	0.25	0	0.20	0
7	0	0	0	0.25	0	0.25	0	0.25	0.25	0
8	0	0	0	0.25	0	0	0.25	0	0.25	0.25
9	0	0	0	0	0	0.25	0.25	0.25	0	0.25
10	0	0	0	0	0	0	0	0.50	0.50	0

Figura 9.10: Ejemplo de matriz de pesos estandarizada equiponderada: La matriz suma 1 en cada fila.

siempre 1. Por defecto los vectores de la matriz son equiponderados; sin embargo, resulta de gran utilidad modificar los pesos para priorizar las unidades de análisis que a priori se sabe que afectan en mayor grado la unidad central, ya sea por un criterio cuantitativo o considerando el criterio experto a partir de un conocimiento exhaustivo de la zona geográfica analizada.

Por ejemplo, la Figura 9.10 muestra pesos de 0.25 para los 4 vecinos de la unidad de análisis 7, valor derivado del criterio básico de equiponderancia: $(1 / \sum \text{número de vecinos})$.

En cualquier caso, se insiste en que la sumatoria de los pesos de cada vector siempre debe ser 1, de ahí su nombre, *row-standardized spatial weighted matrix*.

9.4.4 Los índices de autocorrelación global y local

Después de definir el concepto de autocorrelación espacial y enumerar algunos de los índices que permiten cuantificar el comportamiento espacial de los fenómenos geográficos, se presentan detalles básicos para su cálculo clasificados en dos tipos: globales y locales.

La diferencia principal entre ellos es que los índices globales evalúan asociaciones para el conjunto de valores del fenómeno, considerando todas las unidades de análisis como un bloque a través de la media global del fenómeno analizado (\bar{x}); además, los índices globales no son sensibles a escenarios en los que se presenta heterogeneidad en el comportamiento del fenómeno geográfico.

En este caso, se entiende la heterogeneidad como inestabilidad estructural y heteroscedasticidad.

La inestabilidad estructural implica la discontinuidad del comportamiento del fenómeno; por tanto, una función general podría describir perfectamente una zona, pero no aplicar a otras que presenten otros tipos de comportamientos.

La heteroscedasticidad implica que la dispersión o distribución de los errores a través de la zona analizada no es constante, es decir, se puede dar el caso en el que, para algunas zonas del área de estudio (marco geográfico), los residuales son mucho más variables que en otros. Como se puede observar, generalizar en este tipo de escenarios puede ser peligroso y llevar a conclusiones no sustentadas.

En contraste, los indicadores locales (*LISA*) sí consideran escenarios específicos (de ahí su nombre), por tanto, permiten identificar subzonas en las que se presenta agrupamiento o dispersión del fenómeno. Aunque el cálculo base de los índices *LISA* también considera la media global, estos indicadores asignan medidas de autocorrelación a cada unidad de análisis y, por tanto, permiten identificar individualmente el nivel de agrupamiento o dispersión de cada unidad con relación a sus vecinos o respecto de aquellas unidades que se hayan definido a través del criterio de vecindad. En este caso, los valores o medidas de las unidades de análisis suelen ser estandarizados para que puedan ser comparados entre sí.

I de Moran (Moran's I)

Es posible afirmar que tres de los índices globales más empleados por la comunidad académica son (i) el I de Moran, (ii) c de Geary y (iii) G de Getis y Ord.

Una característica del cómputo del *I de Moran* es que solo se tienen en cuenta los valores de las unidades de análisis determinadas a partir del criterio de vecindad, es decir, el valor de la unidad central no se considera para el cálculo; nótese que esto es consecuencia directa de la diagonal de la matriz de contigüidad, en la que se establece la restricción de que una unidad de análisis no es vecina de ella misma.

Como en la mayoría de los índices, su naturaleza global se deriva de comparar de forma directa los valores de cada unidad de análisis con la media global del fenómeno (\bar{x}), esto es, la media de todas las medidas realizadas en el marco geográfico, tal y como se muestra en el numerador de la Ecuación 9.2.

$$I = \frac{N \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum_i (x_i - \bar{x})^2} \quad (9.2)$$

Donde N es el número de unidades de análisis y $W = \sum_i \sum_j w_{ij}$ corresponde al número total de vecindades.

Considerando la figura 9.8 como referencia, se observa que (1) el número de unidades de análisis es $N = 10$; (2) el número total de vecindades, es decir, la sumatoria de los vecinos de cada unidad de análisis, es $W = 36$; (3) el producto cruz estará limitado por los ceros de la matriz 9.8, es claro que al resolver la expresión $\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})$, si el correspondiente w_{ij} es cero, el producto será siempre cero, este resultado es lógico si se piensa que solamente los valores vecinos sumarán al cálculo del valor I .

El dominio del *I de Moran* es $[-1, +1]$. Si se vinculan los patrones presentados anteriormente con los posibles valores del dominio, se evidencia autocorrelación espacial negativa cuando $I < 0$.

De hecho, al analizar una distribución perfecta ajedrezada como la representada en la Figura 2b a través de una matriz de contigüidad tipo Torre orden 1, resultará en un valor $I = -1$, indicando un comportamiento o patrón disperso del fenómeno. Si analizáramos el mismo fenómeno, pero considerando una matriz de contigüidad tipo Reina orden 1, se obtendría un $I = -0.0958$, indicando un comportamiento o patrón aleatorio. La Figura 9.11 refleja esta situación.

Esto muestra que el *I de Moran* no es una medida universal que describe unívocamente el comportamiento de las unidades de análisis, sino que depende del criterio de vecindad W seleccionado. La implementación del tipo de vecindad define el patrón espacial que informa el índice. La decisión sobre el tipo de matriz W que será

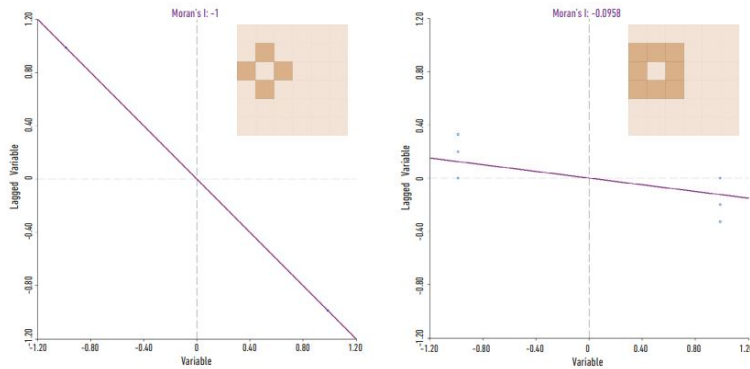


Figura 9.11: Importancia de la base para la vecindad: Resultado del cálculo del I de Moran para un mismo fenómeno y marco geográfico considerando los criterios de vecindad Torre y Reina. Nota: se evidencia cómo el cambio de criterio de vecindad afecta sustancialmente el resultado del análisis para el mismo fenómeno y distribución de las unidades espaciales.

aplicada debe estar entonces definida por la interacción de las unidades de análisis que mejor describa al fenómeno analizado. En cualquier caso, siempre será necesario el buen criterio del analista espacial para interpretar el índice apropiadamente y no llegar a conclusiones equivocadas. Se evidencia autocorrelación espacial positiva cuando $I > 0$. El patrón clúster de la Figura 9.3 refleja esta situación, los valores positivos de I muestran que las unidades de análisis vecinas tienden a ser similares. Continuando con el ejemplo, una distribución agrupada como la presentada en la Figura 9.3 y analizada con una matriz de pesos tipo Torre, resulta en un I de Moran = 0.8645.

En este caso, a diferencia del patrón disperso, al calcular I no se obtiene el límite superior del dominio; la razón principal para no obtener el valor teórico $I = 1$ es la franja vertical intermedia de blancos y negros, en donde claramente los vecinos son diferentes y no se conforman agrupamientos. Si este ejercicio se planteara con la matriz de contigüidad tipo Reina, el valor obtenido sería $I = 0.8093$.

Como se puede observar, cambiar el criterio de vecindad de Reina a Torre en escenarios agrupados no representa cambios significativos en la interpretación del índice; en los dos casos se concluye que el fenómeno está agrupado (patrón clúster). Otro aspecto interesante que permite resaltar este ejemplo es que, en las condiciones presentadas, valores del Índice de Moran superiores a 0.7 corresponden a fenómenos altamente agrupados en el espacio geográfico.

Finalmente, cuando $I = 0$ o computa valores cercanos a cero, el fenómeno se distribuye aleatoriamente (Figura 2c). No es posible asegurar de forma concluyente rangos de aleatoriedad en el dominio del índice, pues estos varían según la distribución de las unidades de análisis y las medidas del fenómeno. No obstante, como recomendación general empírica, se sugiere pensar en aleatoriedad del fenómeno cuando el índice se encuentra en el rango $-0.35 \leq I \leq 0.35$. Los valores mayores y menores deberían llevar a pensar en patrones tipo clúster ($I > 0.35$) o disperso ($I < -0.35$), siendo clúster para valores

positivos y disperso para valores negativos.

Patrones espaciales		
Autocorrelación espacial positiva	Autocorrelación espacial negativa	Ausencia de autocorrelación
Patrón clúster	Patrón disperso	Patrón aleatorio
<i>I de Moran teórico</i>		
$I > 0$	$I < 0$	$I = 0$
<i>I de Moran en la práctica</i>		
$I > 0.35$	$I < -0.35$	$-0.35 \leq I \leq 0.35$

Es importante aclarar que el ejercicio no es tan simple como se presenta, pues para responder la pregunta con criterio estadístico y un grado de confianza determinado se deben tener en cuenta elementos como, por ejemplo, el Z-score y la hipótesis de aleatoriedad espacial completa, en adelante, *csr*. Si bien el *I de Moran* nos presenta un valor que indica el grado de autocorrelación espacial para un conjunto de unidades de análisis considerando una variable específica, una pregunta lógica que surge es qué tan diferente de cero debe ser el índice para que realmente indique dispersión o agrupamiento, el Z-score y una prueba de hipótesis basada en la *csr* ayudan a responder esta pregunta y, además, a comprender si el resultado del análisis es estadísticamente significativo dado un nivel de confianza (Siabato 2018b).

El ejemplo presentado demuestra que, al aplicar un criterio Reina o Torre, o cualquier criterio de vecindad diferente, cambia el resultado del índice y, por tanto, la conclusión a la que el analista espacial llegaría sobre la distribución del fenómeno. En el primer ejemplo (Figura 9.4, el analista podría cambiar su reporte de análisis de disperso (Torre) a aleatorio (Reina); solo las condiciones del análisis y el conocimiento de la variable permitirían decir cuál es correcto. El impacto del criterio de vecindad en lattices² uniformes o regulares como el usado en el ejemplo es muy alto; no obstante, se debe aclarar que en lattices irregulares como el mostrado en la Figura 9.9 no se evidenciaría un impacto alto.

Un análisis de autocorrelación espacial exige que se configuren diferentes escenarios con distintos criterios de vecindad y, principalmente, interpretar los resultados con el conocimiento que se tenga de la realidad analizada, esto es, del fenómeno geográfico medido. Si el analista no conoce los pormenores de la variable o fenómeno, se debería apoyar en un experto en el tema y nunca interpretar resultados de variables geográficas que no comprende o desconoce.

Tabla 9.1: Criterios de autocorrelación espacial: Los patrones espaciales que encontramos en la teoría, se ha convenido considerarlos significativos cuando cumplen los criterios en la práctica.

² Lattice: Este término es común en el argot de los analistas espaciales, se refiere a la distribución de las unidades de análisis que conforman el marco geográfico. En el caso del ajedrez, los ejemplos se han presentado utilizando un lattice regular.

9.4.5 Indicadores locales de Asociación Espacial (LISA)

Anselin (1995, 94) define los indicadores locales de asociación espacial (LISA) como estadísticos que satisfacen dos requerimientos: (i) para

cada observación (unidad de análisis central), el *LISA* proporciona una medida del nivel de agrupamiento espacial significativo con relación a los valores que se localizan alrededor de esa observación (unidades vecinas); (ii) la suma de las medidas *LISA* de todas las observaciones es proporcional al indicador global de asociación espacial (autocorrelación espacial global), por lo que resultan útiles para medir la contribución de cada observación local al valor global.

Como se mencionó, las medidas globales presentan limitaciones que no permiten identificar las especificidades del fenómeno analizado, por ejemplo, no evidencian la presencia de conglomerados zonales o locales en el marco geográfico, y es precisamente en este escenario donde los índices locales resultan de gran interés para el análisis espacial. Los cuadrantes son la base del análisis de asociación espacial, una versión más completa y detallada de los cuadrantes es la que se presenta en la Figura 9.13.

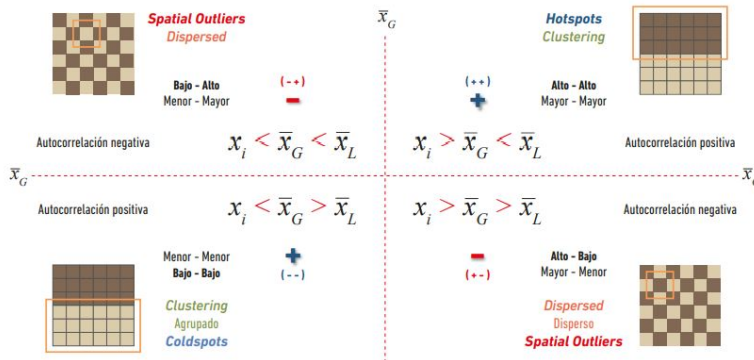


Figura 9.12: Cuadrantes de asociación espacial: los cuadrantes permiten identificar zonas de agrupamiento (hotspots y coldspots) y valores atípicos (spatial outliers) Fuente: basado en Siabato 2018b.

Como se puede observar en la Figura 9.13, los ejes están determinados por la media global (\bar{x}_G), por tanto, en los cuadrantes se representan valores mayores (a derecha) y menores (a izquierda) que la media del conjunto de medidas (muestra) involucradas en el análisis.

La abscisa representa las medidas x_i de las unidades de análisis centrales, mientras que en la ordenada se localizan los promedios de las unidades vecinas (\bar{x}_L).

Si estos valores son menores a la media global ($x_i < \bar{x}_G < \bar{x}_L$), se clasificarán como menor-menor o bajo-bajo. Por el contrario, si son mayores a la media global ($x_i > \bar{x}_G > \bar{x}_L$), se clasificarán como mayor-mayor o alto-alto. Al combinar estas opciones en los demás cuadrantes (ii y iv), resultan las opciones bajo-alto y alto-bajo respectivamente.

Nótese la operación de la ley de signos indicada en la figura 9.12, determinando positivo para los cuadrantes i y iii y negativo para ii y iv. Los cuadrantes y sus respectivas comparaciones son la base del análisis por asociación espacial, que permite identificar la presencia

de agrupaciones (clustering) o valores atípicos ("*spatial outliers*") en el fenómeno geográfico analizado. Las agrupaciones o clústeres pueden tener dos comportamientos: (i) la unidad central de análisis puede tener un valor por debajo de la media global y estar rodeada de vecinos que también están por debajo de la media, es decir, presenta un comportamiento bajo-bajo, en este escenario se configuran los llamados coldspots; (ii) en contraste, en el escenario alto-alto se determinan los hotspots.

Los cuadrantes bajo-alto y alto-bajo determinan la presencia de valores atípicos, que espacialmente se traducen como unidades de análisis centrales que están rodeadas de valores vecinos opuestos a su valor, ya sea por encima (cuadrante ii) o por debajo (cuadrante iv). Es claro que si, por ejemplo, una unidad productiva rural que genera 100 toneladas de sorgo al año está rodeada por otras unidades que producen una media de 500 toneladas en el mismo periodo, desde el punto de vista espacial este es un comportamiento atípico, no es normal que la producción de esta zona sea tan baja. En los términos presentados, este comportamiento sería clasificado como bajo-alto. Esto siempre y cuando la media de producción de toda la zona analizada (\bar{x}_G) sea por ejemplo 350 toneladas —o cualquier valor mayor a 100 toneladas—.

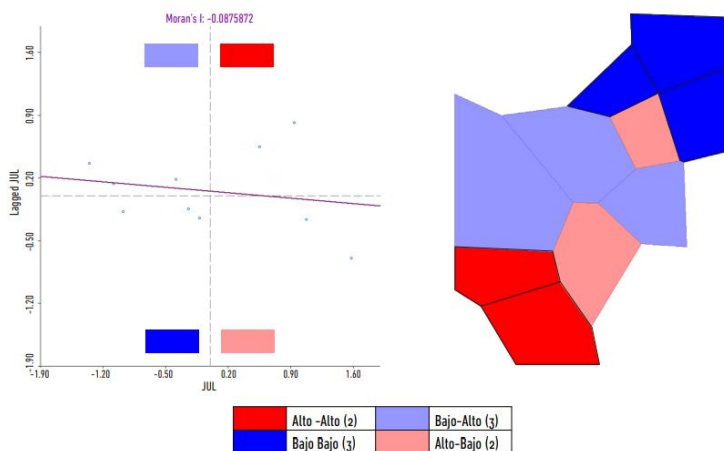


Figura 9.13: Representación de los cuadrantes de asociación espacial: Los colores intensos representan los clústeres y los opacos las zonas atípicas (spatial outliers).

La Figura 9.13 representa la asociación espacial de las unidades de análisis mostradas en la Figura 9.9 asignando valores de lluvia mensuales a cada unidad —julio de 2008, datos tomados de las estaciones meteorológicas del Instituto de Hidrología, Meteorología y Estudios Ambientales; las áreas representan polígonos de Voronoi tomando como semillas las estaciones—. Se puede observar que los ejes de la gráfica no están determinados por \bar{x}_G , sino localizados en 0, esto se debe a que los valores x_i representados están normalizados

considerando el valor medio \bar{x}_G y la desviación estándar. Nótese cómo los puntos representados de los cuadrantes se corresponden con el número de unidades de análisis categorizadas. Cada punto representa la medida estandarizada de la unidad de análisis (Z_i) frente al promedio de sus respectivos vecinos; por tanto, el gráfico de asociación espacial siempre tendrá tantos puntos como unidades de análisis componga el marco geográfico (10 en el ejemplo mostrado). Las expresiones de normalización para la construcción de la gráfica de asociación espacial son las que se presentan en la Ecuación 3.

$$Z_i = \frac{x_i - \bar{x}}{SD_x} \quad \bar{x}_L = \sum w_{ij} Z_i \quad (9.3)$$

Es fácil identificar cómo la Figura 9.13 revela la presencia de dos clústeres: un hotspot (rojo intenso) y un coldspot (azul intenso), sin excluir los comportamientos atípicos en la región central. El numerador de las expresiones matemáticas presentadas determina si el patrón es agrupado o disperso. O'Sullivan y Unwin (2010)³ y Siabato (2018b)⁴ señalan que un resultado positivo en la sumatoria implica que zonas con valores similares (altos o bajos) tienden a estar juntas.

Por el contrario, un resultado negativo implicaría que valores altos están cerca de valores bajos o viceversa. En ocasiones, los resultados positivos y negativos están balanceados tendiendo a cero, lo que indica un patrón aleatorio.

Desde un punto de vista más formal, se debe resaltar nuevamente el hecho de que no basta con localizar la relación de la unidad central con sus vecinos en uno de los cuadrantes. Para definir el grado de autocorrelación del fenómeno analizado, se deben implementar pruebas de hipótesis (Los autores indican que la prueba más común en este escenario es la prueba de límites de Durbin y Watson) (ej. *csr*) con el fin de validar si la distribución espacial del fenómeno se produce aleatoriamente o no. Solo de esta forma se puede identificar, con un nivel de confianza determinado, si hay presencia de agrupaciones o zonas atípicas geográficas.

Los programas informáticos para análisis espacial realizan pruebas estándar con grados de confianza del 90%, 95% y 99%. Si la hipótesis nula (H_0) de aleatoriedad espacial completa (*csr*) se rechaza, entonces el fenómeno tiene una probabilidad del 90%, 95% o 99% de no haberse generado aleatoriamente. En otras palabras, es "poco" probable (0.1; 0.05 ó 0.01) que el fenómeno sea aleatorio y, por tanto, será agrupado o disperso.

Considerando que la prueba de hipótesis *csr* es bilateral o de dos colas, el fenómeno se clasificará como agrupado si está a la derecha (positivo); en caso contrario, el fenómeno geográfico analizado presentará un patrón disperso. Esto será verdad sólo en el nivel

³ David O'Sullivan and David J Unwin. Geographic information analysis and spatial data. *Geographic information analysis*, pages 1–32, 2010

⁴ Wellington Siabato. Sobre la evolución de la información geográfica: las bodas de oro de los sig. *Cuadernos de Geografía: Revista Colombiana de Geografía*, 27(1):1–9, 2018

de confianza indicado. Este análisis aplica para escenarios locales y globales. La Figura 9.13 representa, además, una interesante relación de los índices global y local de Moran, la pendiente (m) obtenida a partir de la regresión lineal simple que describe la recta ($y = mx + b$) derivada de la nube de puntos, corresponde al índice de autocorrelación global. Siendo, por tanto, los puntos la representación del índice local y la recta del global.

El índice local de Moran (I_i de Moran) permite conocer la contribución de cada unidad de análisis a la variación global, es decir, permite medir cuánto aporta cada unidad a la autocorrelación global del marco geográfico analizado. Se debe recordar que el promedio de las medidas locales resulta en la medida global. Los análisis basados en el I_i ⁵ de Moran son los más intuitivos y sencillos, pues su interpretación sigue sin variación los conceptos presentados sobre asociación espacial, por tanto, permiten identificar tanto clústeres positivos y negativos como valores atípicos. De los LISA presentados en esta sección es el único que permite la categorización completa y sin excepciones de las observaciones en todos los cuadrantes (véase Figura 9.13).

La expresión que se computa para determinar las medidas individuales de cada unidad del lattice analizado es la mostrada en la Ecuación (9.4), nótese cómo los términos de la expresión se corresponden con los presentados en la Ecuación (9.3). Z_i es la medida estandarizada de la unidad de análisis, mientras la sumatoria ponderada aplica las contribuciones locales de cada vecino a través de la matriz de pesos estandarizada. De los cuatro indicadores, el I_i de Moran es el más extendido.

$$I_i = Z_i \sum w_{ij} Z_j \quad (9.4)$$

⁵ La I representa el índice global de Moran, mientras que la I_i hace referencia al índice local

9.5 Algoritmo de clasificación K vecinos más cercanos

La distancia entre unidades de análisis es una de las opciones más comunes para obtener la matriz de pesos W . Una alternativa para generar los pesos que se basa en distancias, pero evita el problema de las observaciones aisladas o islas, es mediante el algoritmo de cálculo de pesos por medio de K vecinos más cercanos.

El clasificador de K vecinos más cercanos (o más próximos, ya que se utiliza de manera indistinta) es uno de los métodos más utilizados de clasificación debido a su versatilidad y sus características, entre las que destaca la generalización y su facilidad de implementación. Aunque se le considera como un método simple, usualmente es capaz de equiparar y superar otros métodos más sofisticados. Por ejemplo, este método se ha empleado en visión por computadora, donde se ha aplicado de manera exitosa a una gran variedad de problemas, como

son el reconocimiento de rostros, estimación de poses articuladas y reconocimiento de carácter, entre otros.

En el software Geoda⁶, desarrollado por Luc Anselin, se explican las ventajas de este método en la documentación de github. En contraste con una banda de distancia, los vecinos más cercanos no implican una relación de simetría. El hecho de que la observación B es el vecino más cercano de A no implica que A es el vecino más cercano de B , puesto que puede existir un punto C que es más cercano a B que A .

⁶ <https://geodacenter.github.io/>

Asimismo, en el algoritmo que obtiene los pesos por K vecinos cercanos existe un potencial problema: la ocurrencia de empates, por ejemplo, cuando existe más de una locación u observación que tiene la misma distancia a la unidad de análisis. Esto suele ocurrir en casos donde los polígonos considerados son regulares, por ejemplo, un tablero de ajedrez, pero son improbables en estudios con geografías reales. De cualquier modo, existe un número de soluciones para desempatar, desde aleatoriamente seleccionar uno de los vecinos empatados, hasta incluir todos ellos. El algoritmo de clasificación de K vecinos más cercanos pertenece al dominio del aprendizaje no paramétrico y basado en instancias. No paramétrico dado que no hace suposiciones explícitas sobre la forma funcional de los datos, evitando una estimación pobre de la distribución subyacente de los datos. El aprendizaje basado en instancia debido a que no aprende explícitamente un modelo. En lugar de esto, memoriza las instancias de formación que posteriormente se utilizan en las fases de predicción.

9.5.1 *Calcular KNN: métricas de distancia*

En resumen, el objetivo del algoritmo del vecino más cercano es identificar los vecinos más cercanos de un punto de consulta dado, de modo que podamos asignar una etiqueta de clase a ese punto. Para hacer esto, KNN tiene algunos requisitos:

9.5.2 *Métricas de distancia*

Para determinar qué puntos de datos están más cerca de un punto de consulta determinado, será necesario calcular la distancia entre el punto de consulta y los otros puntos de datos. Estas métricas de distancia ayudan a formar límites de decisión, que dividen los puntos de consulta en diferentes regiones. Por lo general, verá límites de decisión visualizados con diagramas de Voronoi.⁷

Si bien hay varias medidas de distancia entre las que puede elegir, este trabajo solo mencionará las más utilizadas:

Distancia euclidiana ($p = 2$): Esta es la medida de distancia más utilizada y está limitada a vectores de valor real. Usando la fórmula a continuación, mide una línea recta entre el punto de consulta y el otro

⁷ Bárbaro Kenier Casanova Pérez and Carlos José Castillo Piñeiro. Solución informática para la localización de servicios utilizando diagramas de voronoi. B.S. thesis, Universidad de las Ciencias Informáticas. Facultad 3, 2015

punto que se mide.

$$d_{ij} = \sqrt{\sum_{t=1}^n (x_{it} - x_{jt})^2} \quad (9.5)$$

Distancia Manhattan ($p = 1$): Esta es también otra métrica de distancia popular, que mide el valor absoluto entre dos puntos. También se conoce como distancia de taxi o distancia de cuadra de la ciudad, ya que comúnmente se visualiza con una cuadrícula, que ilustra cómo se puede navegar de una dirección a otra a través de las calles de la ciudad.

$$d_{ij} = \sum_{t=1}^n |x_{it} - x_{jt}| \quad (9.6)$$

Distancia minkowski: Esta medida de distancia es la forma generalizada de las métricas de distancia Euclidiana y Manhattan. El parámetro, p , en la fórmula a continuación, permite la creación de otras métricas de distancia. La distancia euclidiana se representa mediante esta fórmula cuando p es igual a dos, y la distancia de Manhattan se denota con p igual a uno.

$$d_{ij} = \sqrt[p]{\sum_{t=1}^n (x_{it} - x_{jt})^p} \quad (9.7)$$

Distancia de hamming: Esta técnica se usa típicamente con vectores booleanos o de cadena, identificando los puntos donde los vectores no coinciden. Como resultado, también se la conoce como la métrica de superposición.

9.5.3 Calcule KNN: definiendo k

El valor k en el algoritmo k -NN define cuántos vecinos se verificarán para determinar la clasificación de un punto de consulta específico. Por ejemplo, si $k = 1$, la instancia se asignará a la misma clase que su vecino más cercano. Definir k puede ser un acto de equilibrio ya que diferentes valores pueden llevar a un ajuste excesivo o insuficiente. Los valores más bajos de k pueden tener una varianza alta, pero un sesgo bajo, y los valores más grandes de k pueden generar un sesgo alto y una varianza más baja. La elección de k dependerá en gran medida de los datos de entrada, ya que los datos con más valores atípicos o ruido probablemente funcionarán mejor con valores más altos de k . En general, se recomienda tener un número impar para k para evitar empates en la clasificación, y las tácticas de validación cruzada pueden ayudarlo a elegir la k óptima para su conjunto de datos.

10 *Desarrollo y Resultados*

10.1

Descripción del conjunto de datos. Como se mencionó en el 8.1 de Estado del Arte, los datos sobre el Índice de Marginación Urbana y sus 11 indicadores de rezago social organizados por AGEB se descargaron directamente del sitio oficial del CONAPO. En 2020, el archivo cuenta

11 Conclusiones generales del trabajo

11.1 Conclusiones globales

Los fenómenos sociales de una zona urbana son definidos por factores que pueden ser difícil de medir, sin embargo, las herramientas estadísticas de la econometría proporcionan una manera concreta para el cálculo de la autocorrelación espacial.

Al relacionar los factores con su área geoespacial, el presente trabajo permite confirmar el impacto de los indicadores en la propagación de la marginación en áreas específicas y delimitadas de la ciudad.

Aunque es necesario aclarar que no hay una afirmación sobre la causalidad, sino sobre correlación, siendo que las zonas de la ciudad con mayores carencias (indicadores relativamente altos), se encuentran en espacios geográficos aislados, en la periferia del AMG.

A partir de los indicadores propuestos por el CONAPO en 2022, el presente trabajo buscó conocer si la geolocalización de las Áreas Geostadísticas Básicas generaría un patrón de autocorrelación espacial en dichos indicadores.

La selección del criterio de vecindad y el número de vecinos fue crucial para poder comparar los resultados entre los indicadores, así como para garantizar el cálculo de los índices globales y locales.

Dichos índices permitieron describir la relación y patrones espaciales de los indicadores socioeconómicos, comprendiendo las características que comparten las AGEB en sus cinco categorías de marginación.

Las zonas de marginación muy alta se caracterizan geográficamente por estar en las orillas o periferia de la ciudad, así como por un alto porcentaje de carencia en servicios de vivienda como internet, telefonía celular, así como servicios básicos como agua potable y drenaje. Además, mantienen porcentajes relativamente altos en carencias educativas y hacinamiento.

Las zonas de marginación alta se caracterizan por un aislamiento menor que las anteriores, con porcentajes altos en servicios como

internet y telefonía celular, pero comparten un porcentaje elevado de hacinamiento.

Las zonas de marginación media se reparten uniformemente por el AMG, con porcentajes relativamente bajos en lo referente a escasez de servicios como internet y telefonía celular. En general, los servicios de vivienda se encuentran cubiertos y el nivel de hacinamiento es moderado.

Las zonas de marginación baja se concentran en zonas céntricas, donde los porcentajes de marginación son bajos en todos los sentidos.

Las zonas de marginación muy baja, sin embargo, se encuentran focalizadas en fraccionamientos o condominios de gran tamaño, ubicados en el norte y sur de la ciudad. En estos, los porcentajes de marginación son prácticamente nulos, ya que se cuenta con todos los servicios, acceso a la educación y a la salud.

11.2 Conclusiones específicas

Cada indicador socioeconómico localizado en su correspondiente área geoadministrativa básica (AGEB) responde a diversidad de factores que debido a la compleja naturaleza de los fenómenos sociales dificulta su estudio. Sin embargo, el presente trabajo permite concluir que algunos de los indicadores tienen un componente espacial relevante, reflejado en los índices globales y locales de autocorrelación espacial.

El porcentaje de población de 15 años o más sin educación básica (i2), el porcentaje de población sin derechohabiencia a los servicios de salud (i3), el porcentaje de ocupantes en viviendas con hacinamiento (i8) y el porcentaje de ocupantes en viviendas sin internet (i10) son los indicadores cuyos resultados en 2020 permiten afirmar que sus magnitudes son influenciadas por un factor espacial.

Por lo tanto, podría recomendarse a las instituciones de gobierno la planeación de políticas públicas que permitan atender las necesidades de educación, salud, vivienda e internet en las zonas donde se presenta un grado mayor de rezago social, enfocando los recursos y esfuerzos en municipios y colonias estratégicas para promover el desarrollo y bienestar de las comunidades marginadas.

Apéndice A: Código implementado

En esta sección se presenta el código implementado para obtener los resultados en el trabajo.

```
# -*- coding: utf-8 -*-
"""
Created on Fri Jun 24 13:09:30 2022

@author: IATGonzalez
"""

#%% 0. Carga de bibliotecas
# Graphics
import matplotlib.pyplot as plt
import seaborn
from pysal.viz import splot
from splot.esda import plot_moran
import contextily
# Analysis
import geopandas
import pandas
from pysal.explore import esda
from pysal.lib import weights
from numpy.random import seed

#%% 1. Carga del archivo con los datos de los 11 indicadores socioeconómicos del IMU

data_path = 'C:/Users/IATGonzalez/OneDrive - ITESO/MCD/LA BENDITA TESIS MARGINACION/Al 2020 TESIS/IMU_JALISCO'
ref = pandas.read_csv(data_path, index_col='CVE_AGEB',encoding='latin-1')
ref.info()

#dataframe auxiliar
refAMG=pandas.DataFrame(ref)

#Codigos de los municipios del área metropolitana de guadalajara
value=[2,39,44,51,70,97,98,101,120,124]
```

```
refAMG = refAMG[refAMG['MUN'].isin(value)]
```

```
#regresar al dataframe original
ref=refAMG
```

```
### 2.- Carga del archivo dbf con la geometría del Estado de Jalisco
```

```
lads = geopandas.read_file(
    "C:/Users/IATGonzalez/OneDrive - ITESO/MCD/LA BENDITA TESIS MARGINACION/14_SCINCE_zip/14/ageb_urb.dbf"
).set_index('CVEGEO')
lads.info()
```

```
### 3.- Generar un dataframe uniendo la información geométrica y geográfica con los datos de los indicadores
```

```
db = geopandas.GeoDataFrame(
    lads.join(
        ref[['OVSINT']]), crs=lads.crs
    ).to_crs(epsg=3857)[
        ['OID', 'OVSINT', 'geometry']
    ].dropna()
db.info()
```

```
### 4.- Mapa para relacionar la variable OVSINT en la geografía del área metropolitana de Guadalajara
```

```
f, ax = plt.subplots(1, figsize=(9, 9))
db.plot(
    column='OVSINT',
    cmap='viridis',
    scheme='quantiles',
    k=5,
    edgecolor='white',
    linewidth=0.,
    alpha=0.75,
    legend=True,
    legend_kwds={"loc": 2},
    ax=ax
)
contextily.add_basemap(
    ax,
```

```

    crs=db.crs,
    source=contextily.providers.Stamen.TerrainBackground
)
ax.set_axis_off()

### 5.- Generación de pesos basados en la geografía, en los centroides de los 9 vecinos. Estandarizados por

# Generate W from the GeoDataFrame
w = weights.KNN.from_dataframe(db, k=9)
# Row-standardization
w.transform = 'R'

### 6.- Generación de variable de rezago que promedia los 9 vecinos de cada observación

db['OVSINT_lag'] = weights.spatial_lag.lag_spatial(
    w, db['OVSINT']
)

### 7.- Generación de mapas con la variable de rezago

f, axs = plt.subplots(1, 2, figsize=(12, 6))
ax1, ax2 = axs

db.plot(
    column='OVSINT',
    cmap='viridis',
    scheme='quantiles',
    k=5,
    edgecolor='white',
    linewidth=0.,
    alpha=0.75,
    legend=True,
    ax=ax1
)
ax1.set_axis_off()
ax1.set_title("% OVSINT")
contextily.add_basemap(
    ax1,
    crs=db.crs,
    source=contextily.providers.Stamen.TerrainBackground,
)

```

```

db.plot(
    column='OVSINT_lag',
    cmap='viridis',
    scheme='quantiles',
    k=5,
    edgecolor='white',
    linewidth=0.,
    alpha=0.75,
    legend=True,
    ax=ax2
)
ax2.set_axis_off()
ax2.set_title("% OVSINT - Spatial Lag")
contextily.add_basemap(
    ax2,
    crs=db.crs,
    source=contextily.providers.Stamen.TerrainBackground,
)

plt.show()

### 8.- Estandarización restando la media y dividiendo la desviación estndar

db['OVSINT_std'] = ( db['OVSINT'] - db['OVSINT'].mean() )/(db['OVSINT'].std)
db['OVSINT_lag_std'] = ( db['OVSINT_lag'] - db['OVSINT_lag'].mean() )/(db['OVSINT'].std)

### 9.-Nube de datos de Moran

f, ax = plt.subplots(1, figsize=(6, 6))
seaborn.regplot(
    x='OVSINT_std',
    y='OVSINT_lag_std',
    ci=None,
    data=db,
    line_kws={'color':'r'})
)
ax.axvline(0, c='k', alpha=0.5)
ax.axhline(0, c='k', alpha=0.5)
ax.set_title('Moran Plot - % OVSINT')
plt.show()

### 10.- Cálculo de I de moran

w.transform = 'R'

```

```

moran = esda.moran.Moran(db['OVSINT'], w)
moran.I
moran.p_sim

### 11.- Gráfica de Moran
plot_moran(moran);
fig,ax=plot_moran(moran, zstandard=False,fitline_kwds=dict(color='#4393c3'));

plt.show()
### 12.- Otros índices complementarios

geary = esda.geary.Geary(db['OVSINT'], w)

geary.C

geary.p_sim

### 13.- Carga de bibliotecas para índice local LISA

#LOCAL

import matplotlib.pyplot as plt # Graphics
from matplotlib import colors
import seaborn # Graphics
import geopandas # Spatial data manipulation
import pandas # Tabular data manipulation
import rioxtarray # Surface data manipulation
import xarray # Surface data manipulation
from pysal.explore import esda # Exploratory Spatial analytics
from pysal.lib import weights # Spatial weights
import contextily # Background tiles

### 14.- Mapa de autocorrelación

# Set up figure and a single axis
f, ax = plt.subplots(1, figsize=(14, 14))
# Build choropleth
db.plot(
    column='OVSINT',
    cmap='viridis',
    scheme='quantiles',
    k=5,

```

```

    edgecolor='white',
    linewidth=0.,
    alpha=0.75,
    legend=True,
    legend_kwds=dict(loc=2),
    ax=ax
)
# Add basemap
contextily.add_basemap(
    ax,
    crs=db.crs,
    source=contextily.providers.CartoDB.VoyagerNoLabels
)
# Remove axes
ax.set_axis_off();

### 15.- Calculo de la variable de rezago

db['w_OVSINT'] = weights.spatial_lag.lag_spatial(w, db['OVSINT'])
### 16.- Estandarización de los valores de la unidad de análisis y la variable de rezago

db['OVSINT_std'] = ( db['OVSINT'] - db['OVSINT'].mean() )/db['OVSINT'].std()
db['w_OVSINT_std'] = ( db['w_OVSINT'] - db['OVSINT'].mean() )/db['OVSINT'].std()

### 17.- Gráfica de dispersión en seaborn para la distribución de los valores

# Setup the figure and axis
f, ax = plt.subplots(1, figsize=(6, 6))
# Plot values
seaborn.regplot(
    x='OVSINT_std', y='w_OVSINT_std', data=db, ci=None
);

### 18.- Nube de puntos con cuatro cuadrantes

# Setup the figure and axis
f, ax = plt.subplots(1, figsize=(6, 6))
# Plot values
seaborn.regplot(
    x='OVSINT_std', y='w_OVSINT_std', data=db, ci=None
)

```

```

# Add vertical and horizontal lines
plt.axvline(0, c='k', alpha=0.5)
plt.axhline(0, c='k', alpha=0.5)
# Add text labels for each quadrant
plt.text(20, 5, "HH", fontsize=25, c='r')
plt.text(12, -11, "HL", fontsize=25, c='r')
plt.text(-20, 8.0, "LH", fontsize=25, c='r')
plt.text(-25, -11.0, "LL", fontsize=25, c='r')
# Display
plt.show()

### 19.- Cálculo de Moran Local o LISA

lisa = esda.moran.Moran_Local(db['OVSINT'], w)

### 20.- Gráfica de alfombra o rug para la distribución de los valores

# Draw KDE line
ax = seaborn.kdeplot(lisa.Is)
# Add one small bar (rug) for each observation
# along horizontal axis
seaborn.rugplot(lisa.Is, ax=ax);
plt.xlim([-1,1])
plt.ylim([0,.2])

### 21.- SPLIT para generar los mapas de cluster y significancia.

from split import esda as esdaplot

# Set up figure and axes
f, axs = plt.subplots(nrows=1, ncols=2, figsize=(60, 60))
# Make the axes accessible with single indexing
#axs = axs.flatten()

        # Subplot 1 #
        # Choropleth of local statistics
# Grab first axis in the figure
ax = axs[0]
# Assign new column with local statistics on-the-fly
db.assign(
    Is=lisa.Is
# Plot choropleth of local statistics
).plot(

```

```

    column='Is',
    cmap='plasma',
    scheme='quantiles',
    k=5,
    edgecolor='white',
    linewidth=0.1,
    alpha=0.75,
    legend=True,
    ax=ax
)
plt.xlim([100,100])
plt.ylim([100,100])

        # Subplot 2 #
        # Quadrant categories
# Grab second axis of local statistics
ax = axs[1]
# Plot Quadrant colors (note to ensure all polygons are assigned a
# quadrant, we "trick" the function by setting significance level to
# 1 so all observations are treated as "significant" and thus assigned
# a quadrant color
esdaplot.lisa_cluster(lisa, db, p=1, ax=ax);

        # Subplot 3 #
        # Significance map
# Grab third axis of local statistics
ax = axs[2]
#
# Find out significant observations
labels = pandas.Series(
    1 * (lisa.p_sim < 0.05), # Assign 1 if significant, 0 otherwise
    index=db.index          # Use the index in the original data
# Recode 1 to "Significant and 0 to "Non-significant"
).map({1: 'Significant', 0: 'Non-Significant'})
# Assign labels to 'db' on the fly
db.assign(
    cl=labels
# Plot choropleth of (non-)significant areas
).plot(
    column='cl',
    categorical=True,
    k=2,
    cmap='Paired',
    linewidth=0.1,
    edgecolor='white',

```

```

    legend=True,
    ax=ax
)

        # Subplot 4 #
        # Cluster map
# Grab second axis of local statistics
ax = axs[3]
# Plot Quadrant colors In this case, we use a 5% significance
# level to select polygons as part of statistically significant
# clusters
esdaplot.lisa_cluster(lisa, db, p=0.05, ax=ax);

        # Figure styling #
# Set title to each subplot
for i, ax in enumerate(axs.flatten()):
    ax.set_axis_off()
    ax.set_title(
        [
            'Local Statistics',
            'Scatterplot Quadrant',
            'Statistical Significance',
            'Moran Cluster Map'
        ][i], y=0
    )
# Tight layout to minimise in-between white space
f.tight_layout()

# Display the figure
plt.show()

### 22.- Muestra de valores Q de LISA

lisa.q[:10]
### 23.- Conteo de valores LISA

counts = pandas.value_counts(lisa.q)
counts

### 24.- Condición de significancia para p-values

(lisa.p_sim < 0.05).sum() * 100 / len(lisa.p_sim)
### 25.- Construcción de tabla de significancia

```

```
# Assign pseudo P-values to 'db'
db['p-sim'] = lisa.p_sim
# '1' if significant (at 5% confidence level), '0' otherwise
sig = 1 * (lisa.p_sim < 0.05)
# Assign significance flag to 'db'
db['sig'] = sig
# Print top of the table to inspect
db[['sig', 'p-sim']].head()

### 26.- Revisión de columnas de significancia

# Print bottom of the table to inspect
db[['sig', 'p-sim']].tail()

### 27.- Categorizar las columnas de la base de datos segun significancia y su cluster

# Pick as part of a quadrant only significant polygons,
# assign '0' otherwise (Non-significant polygons)
spots = lisa.q * sig
# Mapping from value to name (as a dict)
spots_labels = {
    0: 'Non-Significant', 1: 'HH', 2: 'LH', 3: 'LL', 4: 'HL'
}
# Create column in 'db' with labels for each polygon
db['labels'] = pandas.Series(
    # First initialise a Series using values and 'db' index
    spots, index=db.index
# Then map each value to corresponding label based
# on the 'spots_labels' mapping
).map(spots_labels)
# Print top for inspection
db['labels'].head()

### 28.- Revisión de etiquetas y los valores en la columna de cluster.

db['labels'].value_counts()
```

Bibliography

Bárbaro Kenier Casanova Pérez and Carlos José Castillo Piñeiro. Solución informática para la localización de servicios utilizando diagramas de voronoi. B.S. thesis, Universidad de las Ciencias Informáticas. Facultad 3, 2015.

Ragnar Frisch, Olav Bjerkholt, and Ariane Dupont-Kieffer. *Problems and methods of econometrics: the Poincaré lectures of Ragnar Frisch 1933*. Routledge, 2009.

Matías Adrián Gordziejczuk and Patricia Iris Lucero. Turismo y calidad de vida: un estudio de autocorrelación espacial aplicado a la ciudad de mar del plata, provincia de buenos aires, argentina. *Cuadernos de Geografía: Revista Colombiana de Geografía*, 28(1):23–42, 2019.

Harvey J Miller. Tobler’s first law and spatial analysis. *Annals of the association of American geographers*, 94(2):284–289, 2004.

David O’Sullivan and David J Unwin. Geographic information analysis and spatial data. *Geographic information analysis*, pages 1–32, 2010.

Willington Siabato. Sobre la evolución de la información geográfica: las bodas de oro de los sig. *Cuadernos de Geografía: Revista Colombiana de Geografía*, 27(1):1–9, 2018.

Xesús B Pena Trapero. *Problemas de la medición del bienestar y conceptos afines:(una aplicación al caso español)*. Presidencia del Gobierno, Instituto Nacional de Estadística, 1977.