

UNIVERSIDAD AUTÓNOMA DE COAHUILA
CENTRO DE INVESTIGACIONES SOCIOECONÓMICAS

TESIS

TESIS
2012

Marco Antonio Pérez Mendez

“Análisis exploratorio espacial del índice
de marginación en México, 2010”



UNIVERSIDAD AUTÓNOMA DE COAHUILA
CENTRO DE INVESTIGACIONES SOCIOECONÓMICAS
MAESTRÍA EN ECONOMÍA REGIONAL

TESIS

**“Análisis exploratorio espacial del índice
de marginación en México, 2010 ”**

que se presenta como requisito parcial para obtener
el grado de Maestro en Economía Regional

Marco Antonio Pérez Mendez

Comité Evaluador:

Directora: Dra. Myrna L. Sastré Gutiérrez

Lector: Dr. Manuel Ángeles Villa

Lector: Dr. Luis Gutiérrez Flores

Saltillo, Coahuila
Marzo 2013

Índice general

1. Introducción	1
1.1. Justificación y Motivación	5
1.2. Contribución	12
1.2.1. Preguntas de Investigación y Objetivos	13
2. Revisión de Literatura	15
3. Referentes teórico-conceptuales	22
3.1. Configuración de las desigualdades socioeconómicas espaciales	23
3.2. Marco conceptual	26
3.2.1. Evolución histórica del concepto de marginación	26
3.2.2. Marginación en México	28
3.2.3. La marginación como un proceso de desigualdad espacial	30
4. Diseño de investigación	31
4.1. Medida operacional y unidad de análisis espacial	32
4.2. Metodología	35
4.2.1. Técnicas de análisis de datos espaciales	35
4.2.2. Análisis Exploratorio de Datos y Análisis Exploratorio de Datos Espaciales	37
4.2.3. Distribución espacial	39
4.2.4. Efectos espaciales	41

5. Minería de datos espaciales, análisis exploratorio espacial y geovisualización	51
5.1. Análisis univariado del índice de marginación	53
5.1.1. Distribución espacial	53
5.1.2. Autocorrelación espacial global	58
5.1.3. Autocorrelación espacial local	60
5.2. Análisis univariado de los componentes del índice	62
5.2.1. Distribución espacial	62
5.2.2. Autocorrelación espacial global	65
5.2.3. Autocorrelación espacial local	66
5.3. Hallazgos principales	67
5.4. Análisis de los casos de Chihuahua e Hidalgo	69
6. Resultados y conclusiones	73
6.1. Resultados	74
6.2. Conclusiones	79
Bibliografía	83

Índice de figuras

1.1. Gráfico de caja del índice de marginación 2010	6
1.2. Distribución geográfica del índice de marginación 2010 por Estado	7
1.3. Distribución geográfica del índice de marginación 2010 por municipio	7
1.4. Diagramas de dispersión del índice de marginación contra sus componentes . .	9
1.5. Grado de marginación	10
1.6. Gráficos de caja de los componentes del índice de marginación 2010	11
1.7. Distribución geográfica de los componentes del índice de marginación 2010 . . .	12
4.1. Dimensiones de la Marginación	32
4.2. Criterios de contigüidad	43
4.3. Matriz de contigüidad	45
5.1. Distribución espacial del índice de marginación	53
5.2. Cartograma del índice de marginación	54
5.3. Distribución espacial del grado de marginación	55
5.4. Mapa de percentiles del índice de marginación	56
5.5. <i>Boxmap</i> del índice de marginación	56
5.6. Mapa de desviaciones estándar del índice de marginación	57
5.7. Diagrama de dispersión de Moran del índice de marginación 2010	59
5.8. LISA <i>Cluster Map</i> del índice de marginación escalado	61
5.9. Distribución espacial de los componentes	62

5.10. Cartogramas de los componentes	63
5.11. <i>Boxmap</i> de los componentes	64
5.12. Digramas de dispersión de Moran de los componentes	66
5.13. LISA <i>cluster map</i> de los componentes	67
5.14. Distribución del índice en Chihuahua e Hidalgo	69
5.15. <i>Boxmap</i> del índice en Chihuahua e Hidalgo	70
5.16. I de Moran, Chihuahua e Hidalgo	71
5.17. LISA, Chihuahua e Hidalgo	72

Índice de cuadros

1.1. Índice de correlación	8
1.2. Estratificación del índice de marginación, 2010	10
4.1. Indicadores socioeconómicos	34
5.1. Municipios por tipo de aglomeración	61

Capítulo **1**

Introducción

Introducción

La marginación es un conjunto de elementos relacionados. En principio, es un fenómeno que refleja la existencia de desigualdades socioeconómicas entre individuos y agregados sociales, ya que muestra la vulnerabilidad de la población para satisfacer necesidades relacionadas con ciertos bienes y servicios. En otro sentido, la marginación cuenta con una identificación geográfica caracterizada por la diferenciación en el acceso al bienestar. En la presente investigación se plantea una diferencia conceptual. Dicha diferencia se da entre el concepto de marginación empleado en la investigación y el empleado tradicionalmente.

La concepción tradicional de la marginación la define como un fenómeno estructural y multidimensional. El carácter estructural se debe a que se gesta por medio de las condiciones históricas de un pueblo y define sus características por medio de las relaciones prevalecientes en su desarrollo (De~la Vega *et al*, 2010, p. 11). La multidimensionalidad emerge porque se expresa por medio de diferentes dimensiones que generan la situación de marginación relacionada con el acceso a bienes y servicios que son reflejo del desarrollo socioeconómico.

Por ejemplo, Ávila *et al* (2000, p. 11) consideran que “la marginación es un fenómeno estructural que se origina bajo un patrón histórico de desarrollo”. Por otro lado, para Cortés (2006, p. 75), “la marginación intenta dar cuenta del acceso diferencial de la población al disfrute de los beneficios del desarrollo”. Así, la concepción tradicional de la marginación pretende reflejar el grado de desigualdad socioeconómica de la población.

A pesar de que el estudio de la marginación “permite diferenciar unidades territoriales según las privaciones que padece su población” (Ávila *et al*, 2000, p. 11), no se han realizado estudios a nivel nacional que incorporen formalmente el espacio en el análisis de la marginación. Los realizados se basan en análisis estadísticos estándar o bien en el uso de herramientas de la econometría tradicional (Bistrain, 2008; Camberos y Bracamontes, 2007; Peña *et al*, 2007). En México, el indicador empleado en el estudio de la marginación es el índice construido por el Consejo Nacional de Población (CONAPO). Es un “indicador que

contribuye a identificar las disparidades territoriales que existen entre las entidades federativas y los municipios del país” (De~la Vega *et al*, 2010, p. 9).

Aunque el concepto tradicional de la marginación pretende mostrar el grado de las disparidades socioeconómicas regionales, no muestra la incorporación explícita del espacio. La investigación de la dimensión espacial de la marginación pudiera tener implicaciones que no se han explorado en los estudios tradicionales, ya que estos no versan en ningún sentido sobre los efectos que el espacio pudiera inducir. Autores como Camberos y Bracamontes (2007); Martínez y López (2009); Montes y Ortega (2006); Pérez (2010); Bistrain (2008); McVicar (2000) y Salcedo (2004), por citar algunos, realizaron sus análisis sin considerar el espacio y la influencia que el mismo podría tener en los resultados.

La relevancia del espacio en el estudio de la marginación se ha sugerido anteriormente, aunque no ha sido tratado con métodos espaciales explícitos para el caso de México. Por ejemplo, para Cortés (2006, pp. 73-79) “la marginación predica sobre espacios geográficos”, pues se refiere a localidades específicas. “Hace referencia a la manera como enraizan los frutos del desarrollo en el espacio” enfatizando las disparidades que generan los procesos socioeconómicos. Se refiere de manera específica a “agregados sociales espacialmente localizados”, de nuevo para referirse a su localización geográfica.

Con base en lo anterior, la marginación mide las diferencias del desarrollo socioeconómico observadas en diferentes unidades geográficas. Como muestra evidencia previa en la literatura (Rey y Sastré-Gutiérrez, 2010), la desigualdad socioeconómica y las disparidades regionales se reflejan en el desarrollo heterogéneo de las sociedades en el espacio. En este sentido, la dimensión espacial de la marginación puede brindar una perspectiva adicional que permita profundizar el entendimiento de las desigualdades generadas por el desarrollo heterogéneo de manera agregada.

Con base en ello, se incorpora el espacio formalmente en el análisis para explorar los posibles efectos que genera en la distribución de la marginación. Esto puede ofrecer una

perspectiva adicional que permita identificar la existencia de agrupaciones geográficas con valores similares en el índice. En la presente investigación se pretenden explorar, por medio de técnicas de análisis estadístico espacial, en general con Minería de Datos Espaciales (MDE) y en particular con el análisis exploratorio de datos espaciales (AEDE), la existencia de efectos espaciales en la distribución del índice de marginación e indagar si las desigualdades del desarrollo socioeconómico reflejadas en dicha distribución pudieran ser resultado de la existencia de dichos efectos. Este análisis puede servir de base para establecer hipótesis sobre las relaciones socioeconómicas en el espacio (Sastré-Gutiérrez, 2011) y las externalidades asociadas con la proximidad o la distancia (Lobao *et al*, 2007, p. 8).

La Minería de Datos Espaciales MDE y el AEDE son metodologías que permiten establecer hipótesis sobre las posibles relaciones espaciales y los hallazgos que se encuentran con la utilización de sus herramientas. Permite determinar, entre otras cosas, la existencia de aglomeraciones o *clusters* de marginación y detectar la existencia de patrones de distribución por medio de geovizualización con la finalidad de extraer información relevante en conjuntos de datos relativamente grandes.

El análisis exploratorio espacial (AEDE) permite corroborar estadísticamente el vínculo aparente entre diferentes áreas. Esto permite investigar si la distribución de la marginación es resultado de un proceso aleatorio o si por el contrario muestra indicios de un proceso bien definido que genera diferentes áreas con grados de marginación similar. Esto se consigue por medio de la utilización de instrumentos estadísticos del análisis espacial. El análisis exploratorio de datos espaciales se basa en un conjunto de herramientas que sirven para investigar la presencia de efectos espaciales, en particular, autocorrelación espacial y heterogeneidad espacial.

1.1. Justificación y Motivación

En particular, se sigue a Anselin (1999, pp. 68-69) cuando apunta que, si los fenómenos socioeconómicos “se basan en datos espaciales o cuando los fenómenos en consideración son espaciales por naturaleza, se requieren técnicas especializadas de estadística espacial”. Es importante determinar como las “interacciones individuales generan comportamientos colectivos o patrones de agregación”, ya que el “*pensar espacialmente*” permite complementar el estudio de la desigualdad.

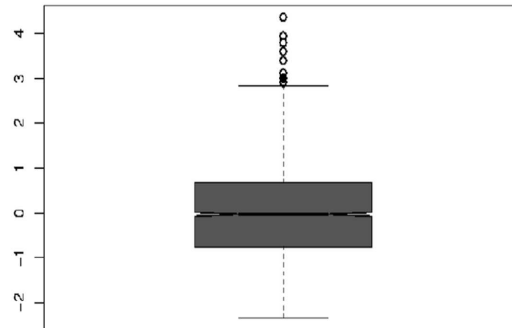
Para lograr tal objetivo, se requieren explorar tanto la distribución espacial de un determinado fenómeno, como la posible existencia de efectos espaciales. Un aspecto de suma importancia que se erige como punto de partida en el análisis espacial es la visualización del fenómeno en el mapa geográfico. En el caso del presente estudio, la geovizualización puede brindar una intuición sobre la distribución espacial y al mismo tiempo identificar áreas que resaltan la desigualdad socioeconómica reflejada en grados de marginación muy altos.

Un análisis estadístico tradicional del índice de marginación en México, puede servir de pauta para resaltar las diferencias entre los estudios tradicionales y aquellos que se basan en análisis espaciales explícitos¹. Con esto, se pretende contrastar las herramientas tradicionales y la necesidad de instrumentos especializados de análisis espacial.

Si se observa el gráfico de caja del índice de marginación, se aprecian varias cuestiones importantes. El gráfico de caja se construye por medio de cuartiles. La caja se ubica entre el primer y tercer cuartil. El segundo cuartil coincide con la mediana y muestra donde se ubica en la distribución estadística. Los puntos mas allá de los bigotes reflejan la existencia de valores atípicos. Al mismo tiempo, el gráfico de caja muestra la dispersión y la asimetría de la variable. Se aprecia en la figura 1.1 la existencia de valores atípicos y un leve sesgo de la distribución estadística.

¹Se emplea el índice de marginación en México para el año 2010 realizado por CONAPO.

Figura 1.1: Gráfico de caja del índice de marginación 2010



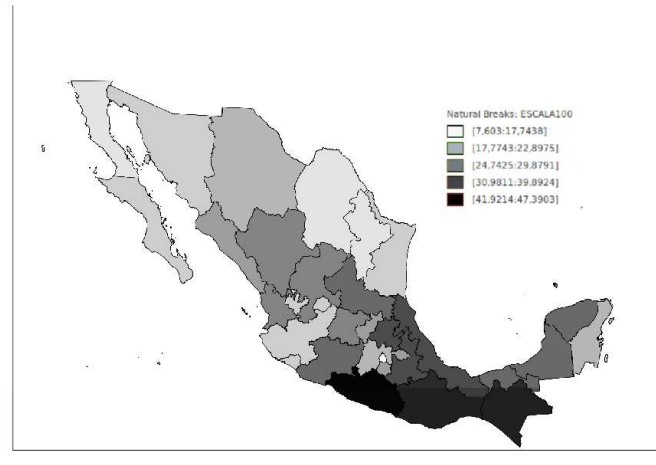
Fuente: Elaboración propia con base en el índice de marginación de México, CONAPO. 2010.

A pesar de que el gráfico de caja brinda información de gran utilidad, con un análisis de este tipo es imposible identificar geográficamente los valores atípicos. En cambio, por medio de Sistemas de Información Geográfica (SIG), se cuenta con una visualización espacial del índice. Los SIG son una fuente de información precisa y un marco dentro del cual el análisis exploratorio de datos se lleva a cabo (Goodchild, 1991; Anselin y Getis, 1992; Anselin, 1994).

EL uso de los SIG permite ubicar geográficamente a los valores atípicos tradicionales y observar diferentes aspectos de un fenómeno en el espacio. Una visualización de la distribución geográfica del índice de marginación a nivel estatal muestra la diferencias entre estados, lo que permite identificar grandes áreas con valores similares. Al mismo tiempo, se aprecia una franja de estados con bajo grado de marginación al norte y una con altos grados de marginación al sur de México. Al comparar la distribución a nivel estatal y municipal (figuras 1.2 y 1.3), se encuentra que dentro de los estados existe una gran desigualdad.

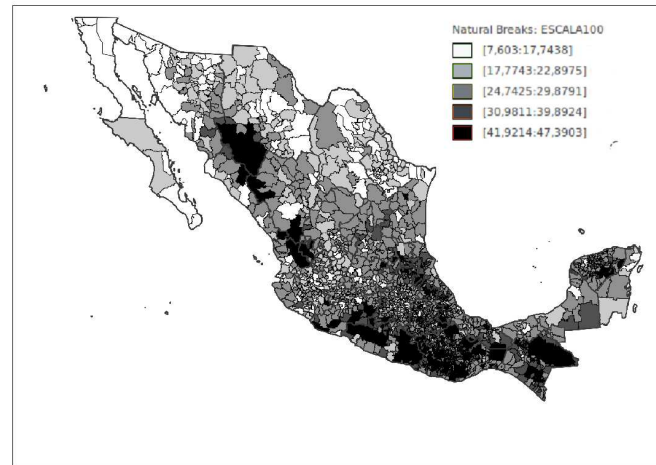
El uso de los SIG brinda un conjunto de información relacionada con un fenómeno. Ofrece la utilización de mapas, que en este caso, permiten ubicar geográficamente los municipios con altos grados de marginación. Las áreas con niveles similares se presentan como áreas de interés pues las desigualdades socioeconómicas se pudieran acentuar en ellas.

Figura 1.2: Distribución geográfica del índice de marginación 2010 por Estado



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010 y Sistemas de Información Geográfica

Figura 1.3: Distribución geográfica del índice de marginación 2010 por municipio



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010 y Sistemas de Información Geográfica

A continuación se estudia como se correlaciona el índice de marginación con los indicadores que lo componen, mostrando una fuerte correlación positiva entre ellos. Este resultado era previsible pues la construcción del índice se basa en cada uno de los indicadores. Sin embargo, lo que indican tanto el índice de correlación de Pearson presentado en el cuadro 1.1 como en la figura 1.4 es el grado en el que cada indicador incide en la medida global (De~la Vega *et al*, 2010, p. 330).

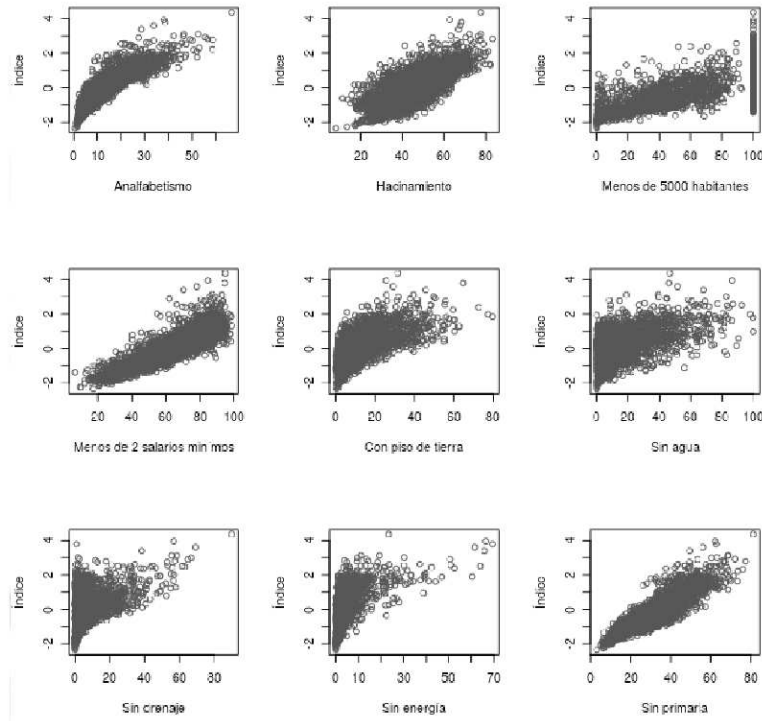
En el gráfico 1.4 se puede observar la correlación entre cada indicador socioeconómico y el índice global para cada unidad espacial. Al mismo tiempo, se observa el comportamiento del índice de marginación ante las variaciones en cada uno de los indicadores socioeconómicos. A pesar de eso, el diagrama de dispersión sigue sin aportar argumentos en términos espaciales pues carece de vínculo geográfico. Por lo que no es posible identificar si las unidades espaciales con indicadores similares se encuentran geográficamente cercanas.

Cuadro 1.1: Índice de correlación

	ÍNDICE
ANALFA	0.88
SINPRIM	0.88
SINDRENAJE	0.46
SINENERGIA	0.58
SINAGUA	0.58
HACINAMIENTO	0.72
PISOTIERRA	0.68
N5000HAB	0.64
N2SMIN	0.86

Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

Figura 1.4: Diagramas de dispersión del índice de marginación contra sus componentes



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

Para tratar de agrupar unidades con valores similares en el índice de marginación se emplea el grado de marginación, el cual representa las categorías en las cuales se agrupa dicho índice. El grado de marginación depende de la vulnerabilidad socioeconómica con la que cuenta la población. A medida que la vulnerabilidad aumenta dicho incremento se refleja en el grado de marginación.

La metodología utilizada para la definición de las categorías del grado de marginación es el método de estratificación óptima de Dalenius y Hodges (De la Vega *et al*, 2010, p. 331). Este método busca identificar estratos que minimicen la varianza al interior de los mismos y la maximice entre los grupos. En este sentido, pretende que las observaciones en cada estrato traten de ser homogéneas y que las observaciones entre cada grupo sean lo más heterogéneas

posible. Esto pretende diferenciar claramente las observaciones entre un grupo y otro.

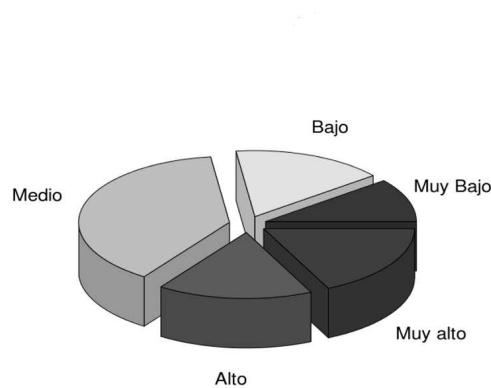
Por medio de dicha clasificación, se obtienen 5 estratos clasificados de la siguiente manera:

Cuadro 1.2: Estratificación del índice de marginación, 2010

Grado de marginación	Número de municipios	Límite del IM		Índice escalado	
		inferior	superior	inferior	superior
Muy bajo	262	[-2.3418	-1.2486]	[1.2109	13.5634]
Bajo	401	(-1.2486	-0.7021]	(13.5634	19.7382]
Medio	944	(-0.7021	0.3927]	(19.7382	32.1105]
Alto	408	(0.3927	0.9384]	(32.1105	38.2767]
Muy alto	441	(0.9384	4.3632]	(38.2767	76.9751]

Fuente: De la Vega *et al* (2010, p. 332)

Figura 1.5: Grado de marginación



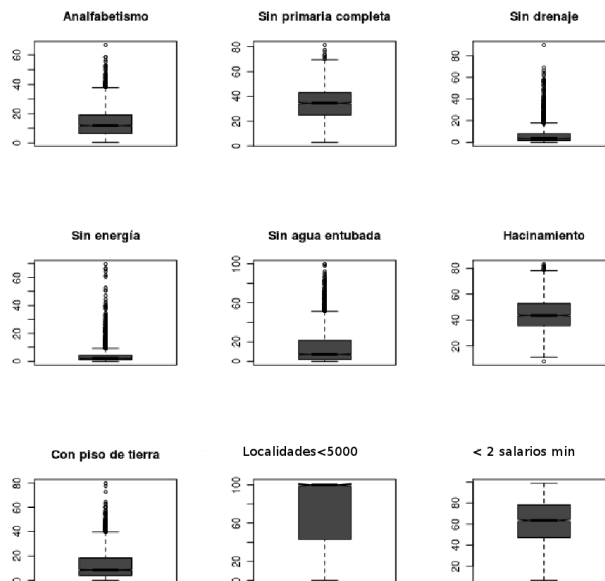
Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

El gráfico de pastel indica las categorías en las que se distribuye el índice y muestra la distribución porcentual de la variable pero, una vez más, carece de vínculo geográfico. Los

municipios con grado alto y muy alto de marginación representan el 34 % del total de municipios del país. Sin embargo, aunque se conoce su ubicación espacial por medio de los mapas realizados por CONAPO se desconoce si dicha ubicación influye de alguna manera en el grado que muestra cada observación.

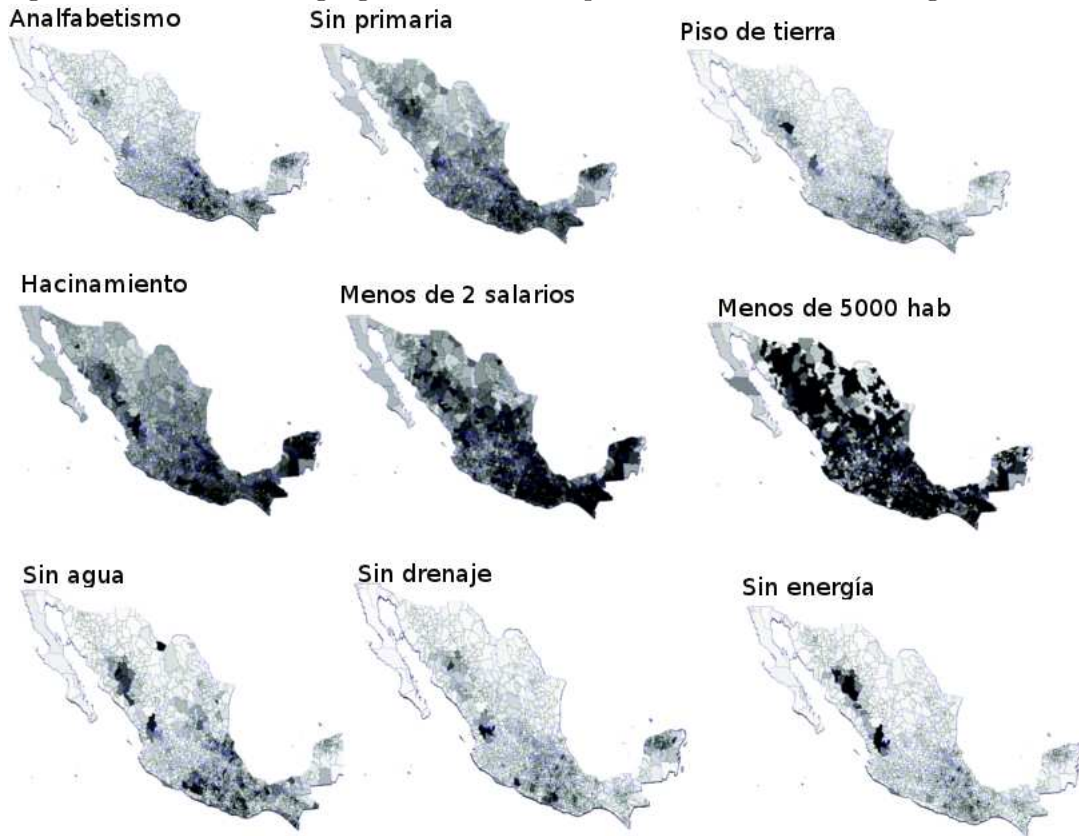
Al observar los gráficos de caja de los indicadores socioeconómicos se aprecia la existencia de valores atípicos y gran dispersión y asimetría en la mayoría de ellos. Es importante mencionar que dichos gráficos no dicen nada en términos espaciales. Entonces, se requiere el uso de sistemas de información geográfica para visualizar los valores atípicos. A continuación se muestran las distribuciones de los indicadores socioeconómicos, donde se aprecian puntos atípicos y su ubicación geográfica en el mapa de México.

Figura 1.6: Gráficos de caja de los componentes del índice de marginación 2010



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

Figura 1.7: Distribución geográfica de los componentes del índice de marginación 2010



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010 y Sistemas de Información Geográfica

1.2. Contribución

En el presente trabajo se pretende explorar la distribución espacial del índice de marginación y, con esto, desentrañar el grado en el que dicho fenómeno se manifiesta en la dimensión espacial. Esto permite contribuir con la literatura sobre la investigación de la marginación al incorporar de manera explícita el espacio. Al mismo tiempo, el enfoque y la metodología propuestas permiten identificar si la incorporación de la dimensión espacial pudiera influir en la distribución del índice de marginación.

Con dicha exploración se pretende observar si su distribución pudiera acentuar las desigualdades socioeconómicas como resultado de la posible asociación entre unidades espaciales. La investigación se llevará a cabo empleando la metodología de la Minería de Datos Espaciales (MDE) de la cual forma parte el Análisis Exploratorio de Datos Espaciales (AEDE), Dicha metodología, la MDE en general y el AEDE en particular, “presta atención en atributos espaciales y de asociación” (Anselin, 1998, p. 81) y sirve de base para extraer información relevante en conjuntos de datos de gran tamaño.

1.2.1. Preguntas de Investigación y Objetivos

1.2.1.1 Preguntas de Investigación

La presente investigación pretende dar respuesta a las siguientes preguntas de investigación: ¿La inclusión de la dimensión espacial muestra aspectos no explorados, como la existencia de dependencia espacial, en la distribución del índice de marginación y sus componentes en México para el año 2010? ¿La exploración por medio de MDE y AEDE arroja hallazgos y patrones interesantes al incluir la dimensión espacial en la distribución del índice de marginación y sus componentes?

1.2.1.2 Objetivo general

El objetivo del presente trabajo es investigar la existencia de patrones espaciales en la distribución del índice de marginación y sus componentes en México, para el año 2010.

1.2.1.3 Objetivos particulares

1. Explorar la presencia de efectos espaciales, en específico dependencia espacial, en el índice de marginación 2010 por medio de Minería de Datos Espaciales (MDE) y Análisis Exploratorio de Datos Espaciales (AEDE).
2. Identificar los efectos espaciales de cada uno de los indicadores que componen al índice de marginación 2010.

3. Contrastar espacialmente los patrones de distribución de los indicadores socioeconómicos y la medida global de marginación.
4. Investigar la presencia de autocorrelación espacial en áreas sugeridas, en una primer aproximación, por medio de MDE para un subconjunto de unidades espaciales

La presente investigación se organiza en seis capítulos. En el capítulo uno se muestra la motivación por medio de un contraste estadístico tradicional resaltando la necesidad de herramientas explícitas de análisis espacial. En el mismo capítulo se señala la contribución y el alcance de la investigación. En el capítulo dos se realiza la revisión de literatura donde se identifican 4 grupos de literatura. En el capítulo tres se plantean los referentes teórico-conceptuales, los cuales se dividen en dos secciones, Marco teórico espacial y Marco conceptual. En el capítulo cuatro se muestra el diseño de la investigación, se exhibe la medida operacional y la metodología empleada, resaltando la importancia del Análisis Exploratorio de Datos Espaciales. En el capítulo cinco, se lleva acabo la aplicación por medio de Minería de Datos Espaciales MDE y AEDE para el índice de marginación en México y sus componentes. Por último, en el capítulo seis se indican los resultados de la investigación y se ofrecen algunas sugerencias y conclusiones.

Capítulo **2**

Revisión de Literatura

La revisión de literatura se clasifica en cuatro grandes cuerpos. Primero, se basa en investigaciones sobre el fenómeno de la marginación que no consideran el estudio del espacio, se enfoca en estudios que utilizan al índice de marginación construido por CONAPO en el caso de México. En segundo lugar, en investigaciones que reproducen el índice de marginación a una escala diferente, ya sea a nivel de AGEB o bien de zona metropolitana. Tercero, en trabajos que muestran interés en la dimensión espacial de los fenómenos socioeconómicos pero no estudian dicha dimensión con herramientas explícitas de análisis espacial. Cuarto y último, en investigaciones que emplean el Análisis Exploratorio de Datos Espaciales o diversas técnicas de análisis espacial en el estudio de fenómenos socioeconómicos.

Algunos de los estudios de marginación revisados, se centran en las diferencias entre determinadas regiones o ciertos grupos sociales. Por ejemplo, Camberos y Bracamontes (2007), realizan una investigación para el estado de Sonora y muestran las diferencias socioeconómicas de doce regiones. Realizan un análisis de la política social en México y los esfuerzos para combatir la marginación. Establecen tendencias y proyecciones de corto y mediano plazo por medio de regresiones para 1970, 1980 y 1990, considerando como variable dependiente el índice de cada año y como dependientes los indicadores que lo componen. Realizan, además, un ejercicio de simulación de los años 1990 y 2000 para determinar las tendencias de largo plazo del índice de marginación en cada una de las regiones.

Martínez y López (2009) buscan mostrar la desigualdad social del estado de Chiapas por medio del índice de marginación y el índice de desarrollo humano, enfocando el análisis en los municipios con población indígena. Estudian las variaciones del índice en 2000 y 2005 con la finalidad de mostrar la incidencia de la marginación en la población indígena de Chiapas.

López *et al* (1996) utilizan el índice de marginación para construir una medida de bienestar social que les permita identificar el grado de bienestar de una población en relación con la morbilidad y el uso de los servicios de salud. Encuentran que el índice de bienestar social les permite, por medio de diferencias de morbilidad y uso de servicios de salud pública, la identificación de áreas sociodemográficas estratégicas. Concluyen con el argumento de la

existencia de una "Ley inversa de atención" que básicamente pregona que las localidades que muestran una necesidad mayor por servicios de salud pública, son las que menos las usan.

A nivel internacional se encuentran trabajos como el de McVicar (2000), quien realiza un estudio sobre la población joven y las políticas de inclusión social en el norte de Irlanda. A lo largo del documento, el autor argumenta que la situación de marginación del estrato de población que estudia, es resultado de procesos de exclusión laboral interrelacionado con desventajas sociales. Realiza un modelo probit para calcular los efectos marginales del desempleo en la población joven marginada.

Pratschke y Haase (2007) realizan una crítica a las medidas de desventaja social frecuentemente utilizadas, entre ellas, los índices de marginación contruidos por componentes principales. Proponen una metodología basada en un modelo de ecuaciones estructurales que permite comparar las dimensiones de privación social tanto en tiempo como en espacio para la república de Irlanda y muestran la articulación espacial de la desventaja social en 1991 y 2002.

Murie y Mustered (2004) identifican a la marginación como una de las formas de exclusión social en ciudades y vecindarios europeos. Estudian la exclusión social como resultado de las políticas de bienestar en seis países europeos y dos ciudades de cada país. De las mismas, se estudian dos vecindarios de cada ciudad, con el objetivo de resaltar los efectos de la exclusión social. En el análisis consideran, patrones de privación, cambios sociales, económicos, demográficos y la organización y operación del estado de bienestar por medio de una metodología cualitativa.

En la sección de trabajos que reproducen el índice de marginación se encuentran: Bistrain (2008), quien analiza la marginación de la población rural de México y construye indicadores de marginación por medio de cuatro técnicas diferentes. Con los indicadores analiza la situación socioeconómica de la población rural para 1990 y 2000. Construye índices de marginación rural por media aritmética, componentes principales, conglomerados y por discriminante. El autor concluye que la mejor técnica para estudiar a la población rural es el índice por media

aritmética, pues permite estudiar diferencias en tiempo y espacio, sin embargo, el estudio no analiza los efectos que el espacio genera en la marginación.

En los trabajos de Montes y Ortega (2006) y Salcedo (2004), se reproduce el índice de marginación empleando la misma técnica empleada por CONAPO, salvo que se realiza para una unidad geográfica mas desagregada. En el primero, se construye el índice de marginación para la zona metropolitana de Monterrey, utilizando datos a nivel de AGEBs por medio de componentes principales y se muestran las diferencias entre los municipios que forman la zona metropolitana. En el segundo, se realiza el mismo trabajo para el Distrito Federal.

En la tercer sección se encuentran trabajos que muestran interés en la dimensión espacial de los fenómenos socioeconómicos, aunque no se tratan de manera formal. En esta línea se encuentran trabajos como Pardo *et al* (2007) y López-Calva *et al* (2008) quienes se acercan al análisis espacial de la pobreza. Utilizan una metodología econométrica novedosa para optimizar las mediciones de la pobreza y realizan una visualización de la distribución de las variables en el mapa. La visualización se presenta en dichos artículos como el punto final del estudio.

González *et al* (2004) realizan una investigación que busca identificar las distribución territorial de la mortalidad infantil y su relación con algunos indicadores socioeconómicos, en específico, con el índice de marginación en la península de Yucatán. Consideran el análisis espacial como un análisis cartográfico que permite vincular los indicadores socioeconómicos con el índice de mortalidad infantil, mostrando aparentes patrones de distribución en dos momentos en el tiempo, 1990 y 2000. Sin embargo, y a pesar del interés que muestran en el análisis espacial, no realizan un análisis formal por medio de instrumentos explícitos, salvo el enfoque cartográfico.

En Mojarro *et al* (2003) se pretende mostrar el vínculo entre la marginación y la mortalidad materna a nivel municipal. Encontrando, como era de esperarse, una gran correlación entre ambos indicadores lo que permite la focalización de las políticas públicas enfocadas a

disminuir el rezago socioeconómico a nivel municipal. Al mismo tiempo muestra como las instituciones de salud pública deben incrementar la calidad de la atención hacia mujeres embarazadas o en lactancia ubicadas en localidades con un grado alto y muy alto de marginación.

Peña *et al* (2007) buscan analizar el patrón de consumo alimentario en Nuevo León por estrato de marginación. Los autores consideran que la alimentación debería formar parte del índice de marginación. Identifican diferencias importantes en los patrones de consumo alimentario de los diferentes estratos principalmente ocasionados por las variaciones en el ingreso. Consideran que con base en dichas diferencias es menester construir un índice alimentario que capture los patrones con base en el ingreso. El reto es incorporar en la construcción del índice de marginación una dimensión alimentaria con la finalidad de establecer la seguridad alimentaria en Nuevo León.

Una aportación que muestra interés en el aspecto espacial de la marginación y reproduce el índice a una escala diferentes es Solís (2002), En este trabajo, el autor se basa en la premisa de que las expresiones de la marginación dependen en gran medida de la escala en la que se presentan, por ejemplo el índice de CONAPO se construyen para estados y municipios. El punto que resalta este trabajo es que dichos índices son de poca utilidad para un entorno urbano. En función de eso, se plantea la construcción de un índice de marginación urbana que incorpore, además de las dimensiones empleadas tradicionalmente, una dimensión de género que capture la participación equitativa de la sociedad. EL resultado al que se llega es que el índice de marginación urbana captura fenómenos que no se aprecian en el indicador tradicional debido a las características socioeconómicas de la sociedad urbana.

En la cuarta sección de la revisión, se analiza la literatura sobre disparidades espaciales y la utilización de técnicas explícitas de análisis espacial. Dichas investigaciones socioeconómicas se relacionan con estudios de pobreza o desigualdad, o bien indicadores de ellas, utilizando el Análisis Exploratorio de Datos Espaciales.

Un estudio que muestra la necesidad de la incorporación explícita en la investigación de las disparidades regionales en México es, sin duda, Sastré-Gutiérrez y Rey (2008). Muestran el interés creciente en la inclusión de la dimensión espacial, realizan un estudio de la desigualdad regional por medio de la descomposición del índice de Theil con una adaptación planteada por Rey (2004). Basan el análisis inferencial en la técnica de permutación aleatoria condicional. Analizan información sobre el índice de Theil y sus componentes inter e intraregional para seis escenarios de regionalización y señalan como los escenarios son sensibles ante la agregación lo que induce la falacia ecológica. Concluyen con el argumento de que la desigualdad económica entre estados exacerba la polarización espacial formando clusters antagónicos.

Rey y Sastré-Gutiérrez (2010) analizan la existencia de efectos espaciales en los datos, con el objetivo de estudiar la desigualdad económica interregional en México durante el periodo 1940-2000. Emplean análisis exploratorio de datos espaciales para estudiar la evolución de la dependencia espacial y la desigualdad global entre estados, la heterogeneidad entre regiones y la sensibilidad de los resultados ante diferentes escenarios de regionalización.

Sánchez~Peña (2006) realiza un estudio de pobreza para la zona metropolitana de Guadalajara. La autora estudia y construye las medidas I de Moran e Indicadores Locales de Asociación Espacial (*LISA*²), realiza modelos de regresión espacial y regresión geográficamente ponderada identificando clusters de pobreza en la zona metropolitana de Guadalajara.

Jiménez~Montero *et al* (2004) realizan un estudio de la dinámica socioeconómica de la región del norte de Coclé en Panamá, basándose en el índice de desarrollo humano y utilizando la técnica de análisis exploratorio espacial en una parte de su investigación.

Hernández~Franco *et al* (2007), utilizan métodos exploratorios para calcular los segundos niveles de un modelo de regresión logit multinivel con pobreza de capacidades como variable dependiente y características del hogar como variables independientes.

²Local Indicators of Spatial Association.

O'Loughlin (1983) realiza un análisis de las desigualdades sociales de ciudades alemanas, estadounidenses y canadienses. Por medio de nueve indicadores de desigualdad social construye indicadores de privación, segregación y entropía. Estudia las distribuciones espaciales de los indicadores en las ciudades seleccionadas, los efectos de la desigualdad y la concentración espacial.

Anselin *et al* (2007) realizan un análisis exploratorio del riesgo infantil en Virginia, construyen una medida de riesgo infantil con base en indicadores de cuidados prenatales, bajo peso al nacer y mortalidad infantil. Utilizan la metodología de componentes principales para obtener una medida única y continua del riesgo infantil. Con base en ella, realizan la exploración, encontrando hallazgos interesantes. Construyen cinco matrices de peso espacial, las cuales emplean en la construcción de los indicadores de autocorrelación espacial. Detectan “*outliers*” espaciales y encuentran evidencia de la existencia de heterogeneidad espacial sistemática, pues los *clusters* se mantienen en el tiempo lo que muestra un grado creciente de polarización.

Capítulo **3**

Referentes teórico-conceptuales

3.1. Configuración de las desigualdades socioeconómicas espaciales

La presente sección sigue el enfoque socio-económico espacial, propuesto en Sastré-Gutiérrez (2011)³, en una estrategia de integración que parte de bases de análisis espacial (Isard, 1956; Isard *et al*, 1970; Goodchild y Janelle, 2004; Goodchild *et al*, 2000), y la corriente sociológica espacial (Tickamyer, 2000; Lobao *et al*, 2007), misma que enmarca el análisis socioeconómico espacial y las desigualdades en el espacio. El enfoque en Sastré-Gutiérrez (2011) ha sugerido, hasta ahora, la incidencia del espacio en variaciones regionales del ingreso, movilidad espacial y distribucional para el caso de México (Rey y Sastré-Gutiérrez, 2010; Sastré-Gutiérrez y Rey, 2013). Un enfoque similar, extendido a múltiples variables ha brindado evidencia de diversos factores socioeconómicos y escalas espaciales para otras regiones (Rey *et al*, 2011).

La presente investigación busca aportar evidencia empírica al enfoque anterior dentro del cuerpo disciplinario denominado “ciencia social espacialmente integrada” (*Spatially Integrated Social Science, SISS*), la cual enfatiza la importancia de las características espaciales en los fenómenos sociales. La *SISS* busca dar un tratamiento formal a procesos sociales que se presentan en la dimensión espacial. Esto se logra por medio de diferentes componentes del análisis espacial, como la visualización del fenómeno espacial, análisis de datos espaciales y visualización de información espacial (Goodchild y Janelle, 2004; Goodchild *et al*, 2000). El uso de la ciencia social espacialmente integrada, en el presente estudio, pretende explorar los fenómenos socioeconómicos con un entendimiento de los factores espaciales y mostrar como la profundización en el análisis de tales factores permite identificar hallazgos en el estudio de los fenómenos socioeconómico-espaciales.

En la perspectiva sociológica, existen por lo menos dos enfoques que incorporan el estudio del espacio en las relaciones sociales: *la sociedad en el espacio* y *el espacio en la sociedad*. La primera estudia como los procesos sociales trabajan sobre espacios determinados. Por su parte, la segunda trata las características del espacio en una sociedad específica y como las

³Laboratorio de Análisis Socioeconómico Espacial (Spatial SEALab)

diferencias entre distintos lugares generan disparidades sociales (Lobao *et al*, 2007, p. 11).

De acuerdo con este enfoque sociológico, el estudio de la desigualdad en el espacio se basa en cuatro premisas (Lobao *et al*, 2007, p. 10): 1) se considera que el espacio se relaciona de diferentes formas con el *estatus* socioeconómico, 2) el espacio es un canal que evidencia las desigualdades; 3) se reconoce que el propio espacio genera procesos de desigualdad y, 4) el espacio y la desigualdad están causalmente interrelacionados.

La estrategia empleada brinda un entendimiento de los fenómenos espaciales y su interrelación al poder diferenciarlos claramente en el estudio de la distribución y configuración de los procesos socioeconómicos espaciales. Entonces, el espacio se entiende como un concepto que permite teorizar relaciones sociales a diferentes escalas y contabilizar las externalidades asociadas con la distancia (Lobao *et al*, 2007, p. 8). “El espacio es un concepto que produce ruido en las relaciones sociales, por medio de la incorporación de la distancia en la modelación de las relaciones socioeconómicas. Por ejemplo, la pobreza, los logros académicos, las características demográficas y la estructura económica tienden a aglomerarse con niveles similares alrededor de unidades espaciales específicas” (Lobao *et al*, 2007, p. 9). Tales lugares proveen una ubicación que puede operar como un contenedor de fenómenos sociales determinado por factores que describen la estructura y los procesos socioeconómicos (Tickamyer, 2000, p. 806).

Por lo tanto, el espacio se intersecta con los fenómenos socioeconómicos de manera compleja, ya que es visto como un canal que limita o amplifica los efectos de los procesos socioeconómicos. Así, la dimensión espacial de los fenómenos socioeconómicos permite profundizar el entendimiento de las formas en las que se distribuyen dichos fenómenos (Lobao *et al*, 2007). Dicha forma es preponderante porque la localización influye en la calidad y cantidad de los bienes y servicios, tanto públicos como privados, a los que la población tiene acceso y de manera mas general en la calidad del desarrollo socioeconómico en el que se desenvuelven (Green, 2011, p. 2).

Así, los indicadores relacionados con dichos fenómenos generan un conjunto de diferencias

socioeconómicas que desembocan en la formación de la desigualdad espacial (Lobao *et al*, 2007, p. 3). Por tal motivo, Al incorporar factores socioeconómico-espaciales se pretende identificar como y por que la dimensión espacial contribuye con la desigualdad del desarrollo socioeconómico (Lobao *et al*, 2007, p. 3). Desde la perspectiva de la equidad, la desigualdad espacial genera áreas socialmente indeseables ya que contribuyen con la desigualdad social entre regiones. Incluso, la desigualdad espacial puede ser un factor de inestabilidad social si muestra diferencias significativas en el bienestar socioeconómico (Kim, 2008, p.2). Las zonas marginadas pueden ser el resultado de las diferencias espaciales en el desarrollo socioeconómico.

Por medio de la utilización del enfoque mencionado, se pretende aportar evidencia empírica a la Ciencia Social Espacialmente Integrada y explorar la distribución de la marginación como resultado de un proceso de desigualdad espacial. Esto se logra al estudiar la distribución espacial de los fenómenos socioeconómicos y la desigualdad espacial del desarrollo socioeconómico en un momento en el tiempo.

3.2. Marco conceptual

3.2.1. Evolución histórica del concepto de marginación

La marginación se define, tradicionalmente, como el proceso socioeconómico que coloca a una persona o a un grupo de personas en condiciones económicas, sociales, legales o políticas de inferioridad. El carácter multidimensional en los estudios tradicionales destaca a la marginación como un fenómeno que aqueja a ciertos grupos o sectores de la población impidiéndoles el goce de sus libertades, sean las mismas de carácter económico, político, social o legal. Las cuatro cuestiones se consideran, típicamente, como las dimensiones básicas de la marginación, lo que le da el carácter multidimensional.

De acuerdo con Campoy (1995, p. 68-69), la marginación tiene sus raíces en el trabajo realizado por Park (1928) quien subraya la importancia del hombre marginal relacionándolo con aquellos individuos migrantes desposeídos de los beneficios tradicionales. También señala que el trabajo de Stonequist (1935) amplía la importancia del hombre marginal relacionándolo con ciertos niveles de exclusión. Para la autora, la marginación exige la referencia de un cierto grupo social, al que se compara con los grupos marginados, ya que el primero necesariamente implica la existencia de ciertas pautas que permiten considerar al segundo grupo como marginado.

Hay un tránsito conceptual entre marginación, marginalidad, exclusión y pobreza y en muchas ocasiones no se hace diferencia entre dichos conceptos. El tránsito al que se hace referencia se relaciona con las formulaciones originales de los autores antes señalados. La diferencia conceptual entre dichas concepciones teóricas se puede apreciar en trabajos especializados como Campoy (1995), Cortés (2006) y Moreno~Jiménez (2001), quienes muestran el paso natural entre los diferentes conceptos hasta llegar al utilizado actualmente en la mayoría de los trabajos sobre marginación.

Las formulaciones iniciales relacionadas con la marginación la ubican o bien con grupos de migrantes en zonas conurbadas con bajos niveles de bienestar socioeconómico, o bien con

individuos a los que se aleja de dichos beneficios por alguna característica particular. Ya sea por baja instrucción académica, pertenencia a un grupo étnico, o simplemente, por contar con bajos recursos económicos.

Para los objetivos de la presente investigación, baste mencionar que la marginación y los conceptos antes mencionados se encuentran interrelacionados aunque no pretenden explicar el mismo fenómeno. Por ejemplo, la marginación se caracteriza por la ausencia de bienestar socioeconómico en ubicaciones particulares. Se diferencia de la exclusión por el hecho de que, en el primero, los agregados sociales en los que se enfoca “se encuentran al margen de las condiciones de desarrollo pero no como resultado de no ser parte del mismo, en cambio la exclusión implica que se está en una posición en la que se ha quedado fuera del sistema socioeconómico al que se hace referencia” (Bistrain, 2008, p. 18).

Marginación y pobreza también se usan indistintamente ya que son realidades que se retroalimentan aunque existen diferencias conceptuales puesto que la pobreza es un estado personal, mientras que la marginación se refiere a grupos bien definidos (Moreno~Jiménez, 2001, p. 79). Un ejemplo tradicional que permite diferenciar entre un concepto y otro es el considerar a individuos que por escasez de recursos económicos no tiene acceso a servicios de educación, por lo que es pobre y al mismo tiempo marginado. En cambio cuando el mismo individuo cuenta con los recursos necesarios para satisfacer sus necesidades de educación pero no lo hace por que en su lugar de residencia no existe la infraestructura para ello, entonces, estará marginado sin ser pobre (Bistrain, 2008, p. 15).

Para Moreno~Jiménez (2001, p. 63) existen ocho criterios que permiten identificar a los grupos marginados: 1) falta de recursos económicos, 2) falta de fuentes de empleo, 3) escasas redes sociales o familiares, 4) vivienda en condiciones precarias, 5) comportamientos no aceptados socialmente, 6) características personales no aceptadas socialmente, 7) escasa formación y 8) carencia de salud. Este espectro de factores contribuyen con una marginación sistemática de determinados grupos identificados con ordenamientos espaciales (Campoy, 1995, p. 81).

El estudio de la marginación se acerca al concepto actual con los desarrollos de Vekemans y la fundación del instituto Desal⁴ donde se pretendió estudiar y comprender el fenómeno de la marginalidad del mundo hispano. Remiten el problema de la marginación hasta la época de la conquista española (Campoy, 1995, p. 78) caracterizando el fenómeno con la realidad de América Latina a la luz de herramientas de carácter histórico y estructural.

3.2.2. Marginación en México

El estudio de la marginación ha sido utilizado al “definir estrategias y políticas sociales enfocadas a la dificultad estructural para propagar el progreso en los sectores que históricamente muestran una desigualdad persistente” (Comité técnico para la medición de la pobreza, 2002, p. 11). En México, el estudio de la marginación surge con los primeros esfuerzos para contabilizarla realizados por la COPLAMAR⁵. En 1977 se diseñó un programa que pretendía profundizar el conocimiento de las zonas marginadas y deprimidas del país (Coplamar, 1982, p. 9).

Los trabajos de la COPLAMAR se centraban en tres principios fundamentales: “ 1) el crecimiento económico no constituye el propósito del desarrollo sino un medio para alcanzarlo, 2) el desarrollo se expresa en el grado de satisfacción de las necesidades esenciales de la población y 3) la planeación debe partir de dichas necesidades y, en función de ellas, las características de la estructura productiva” (Coplamar, 1982, p. 10). En función de dichos principios se estudió el grado de las necesidades de alimentación, educación, vivienda y salud a nivel nacional, estatal y municipal. Se agruparon “municipios homogéneos en cuanto al nivel de acceso de su población a los mínimos de bienestar definidos en las necesidades esenciales y así conocer los diferentes grados de marginación en el interior de las entidades federativas” (Coplamar, 1982, p. 13).

Partiendo de un marco de referencia en el que se enfatiza la importancia de las desigual-

⁴Instituto para el Desarrollo Económico y Social de América Latina

⁵Coordinación General del PPlan Nacional de Zonas Deprimidas y Grupos Marginados

dades regionales, la COPLAMAR, argumenta que el concepto de marginación sirve para “caracterizar aquellos grupos que han quedado al margen de los beneficios del desarrollo nacional y de los beneficios de la riqueza generada, pero no necesariamente al margen de la generación de dicha riqueza ni mucho menos de las condiciones que la hacen posible” (Coplamar, 1982, p. 22). La técnica estadística adoptada por la COPLAMAR para el cálculo de los índices fue componentes principales, herramienta que se usa hoy en día para el mismo objetivo.

Posteriormente, el CONAPO retomó la misma herramienta para calcular los índices de marginación con base en los censos y conteos de población y vivienda. El primer estudio realizado por CONAPO se encuentra en el documento “Desigualdad Regional y Marginación Municipal en México, 1990”. Con la misma idea se realizaron los Índices de Marginación 1995 y 2000. Aunque en el año 2000 se incorporan los índices de marginación a nivel localidad y de marginación urbana. Siendo este un ejercicio recurrente en 2005 y 2010.

Los índices realizados por la COPLAMAR y por CONAPO no difieren sustancialmente, salvo en el número de indicadores que se toman como insumo para la construcción del índice. En esencia, los indicadores brindan la misma información aunque los objetivos de política difieren marginalmente. EL concepto de marginación empleado en la formulación de cada documento de investigación difiere en términos de forma, mas no de fondo, pues solamente se parafrasea el concepto tomado por la COPLAMAR.

El objetivo de la COPLAMAR era conocer la distribución espacial de la marginación, objetivo que se ha modificado por el CONAPO, ya que tan solo pretende identificar la distribución geográfica del fenómeno sin considerar las herramientas teóricas y estadísticas que permitan alcanzar el objetivo planteado por la COPLAMAR. El único esfuerzo realizado por el CONAPO se encuentra en el primer documento de 1990, incorporando un anexo titulado “dimensión espacial de la marginación en México” aunque solo se refiere a la distribución geográfica del fenómeno.

En ese sentido, es necesario caracterizar a la marginación como un fenómeno socioeconómi-

co espacial y profundizar su entendimiento por medio de instrumentos sofisticados del análisis espacial mostrando evidencia empírica.

3.2.3. La marginación como un proceso de desigualdad espacial

Considerando que la marginación se refiere a las unidades geográficas a las que aplica y permite clasificar a la población que vive e dichos ámbitos, entonces, la marginación comprende a toda la población que habita en zonas marginadas (Cortés, 2006, p. 79). Por otro lado, considerando que se toma el acceso de la población a los beneficios del desarrollo socioeconómico y se mide el grado de satisfacción de las necesidades de dichos individuos, se toman todas las unidades geográficas del país y toda la población que vive en ellas.

Con dicho telón de fondo, la marginación se traduce en un proceso de desigualdad espacial de carácter nacional por el hecho de existir diferencias espaciales significativas en el desarrollo socioeconómico de México. La marginación pudiera ser el resultado de las externalidades asociadas con la proximidad y la distancia reflejadas en la distribución del desarrollo socioeconómico. Esto impide que los indicadores socioeconómicos de vivienda, educación, ingresos monetarios y distribución de la población cuenten con un nivel de satisfacción adecuado mostrando diversos niveles de vulnerabilidad reflejados en la distribución del índice de marginación.

Capítulo **4**

Diseño de investigación

4.1. Medida operacional y unidad de análisis espacial

El índice de marginación es el resultado de la combinación de nueve indicadores que son procesados por medio de la técnica estadística de componentes principales. El índice es el primer componente resultado de la implementación de dicho análisis y se construye por medio de indicadores socioeconómicos a nivel municipal. Los mismos se agrupan en cuatro dimensiones (De~la Vega *et al*, 2010): vivienda, educación, ingresos monetarios y distribución de la población, y se construyen con información del Censo General de Población y Vivienda 2010, del Instituto Nacional de Estadística y Geografía (ver Figura 4.1).



Figura 4.1: Dimensiones de la Marginación

Fuente: Elaboración propia con base en De~la Vega *et al* (2010)

El presente estudio se realizará para México, tomando como unidad espacial de análisis a los 2,456 municipios considerados en los indicadores construidos por el Consejo Nacional de Población (CONAPO), tomando como medida operacional al índice de marginación 2010.

De acuerdo con De~la Vega *et al* (2010), el índice de marginación es una medida resumen en la que cada indicador I_{ij} refleja las vulnerabilidades de la población. Así, se tiene para cada unidad espacial un conjunto de indicadores I_{ij} , donde j se refiere al indicador socioeconómico e i a la unidad espacial. Se tienen 9 indicadores socioeconómicos georeferenciados con 2,456 municipios. La construcción de los indicadores se aprecia en la tabla 4.1

Para cada unidad espacial se cuenta con 9 indicadores, por lo que se tiene un conjunto de indicadores de la siguiente manera:

$$\begin{bmatrix} I_{1,1} & I_{1,2} & \cdots & I_{1,9} \\ I_{2,1} & I_{2,2} & \cdots & I_{2,9} \\ \vdots & \ddots & & \vdots \\ I_{2456,1} & & \cdots & I_{2456,9} \end{bmatrix}$$

Los cuales se integran en una medida única por medio del análisis de componentes principales por lo que se le asigna un indicador de marginación a cada unidad espacial: IM_i , donde $i = 1, \dots, 2456$. Además, se tiene un indicador georeferenciado con cada unidad espacial.

Cuadro 4.1: Indicadores socioeconómicos

Nombre del indicador	Fórmula	Donde:
I_{i1} = Población analfabeta	$I_{i1} = \frac{P_i^{anal}}{P_i^{15+} - NE_i^{alJfa}} \times 100$	P_i^{anal} = Población de 15 años o más analfabeta
		P_i^{15+} = Población de 15 años o más
		NE_i^{alJfa} = Población de 15 años o mas que no especificó su condición de alfabetismo
I_{i2} =Población sin primaria completa	$I_{i2} = \frac{P_i^{si} + P_i^{1-5}}{P_i^{15+} - NE_i^{ms}} \times 100$	P_i^{si} = Población de 15 años o más sin instrucción
		P_i^{1-5} = Población de 15 años o más que aprobó entre el primer y quinto año de primaria
		P_i^{15+} = Población de 15 años o más
		NE_i^{ms} = Población de 15 años o mas que no especificó nivel de instrucción
I_{i3} = Porcentaje de ocupantes en viviendas sin drenaje ni excusado	$I_{i3} = \frac{O_i^{sde}}{O_i^t - NE_i^{de}} \times 100$	O_i^{sde} = Ocupantes en viviendas sin drenaje
		O_i^t = Total de ocupantes en viviendas particulares
		NE_i^{de} = Número de ocupantes en viviendas que no especificaron drenaje
I_{i4} =Porcentaje de ocupantes en viviendas sin energía eléctrica	$I_{i4} = \frac{O_i^{see}}{O_i^t - NE_i^{ee}} \times 100$	O_i^{sde} = Ocupantes en viviendas sin energía eléctrica
		O_i^t =Total de ocupantes en viviendas particulares
		NE_i^{ee} = Número de ocupantes en viviendas que no especificaron energía eléctrica
I_{i5} =Porcentaje de ocupantes en viviendas sin disponibilidad de agua	$I_{i5} = \frac{O_i^{sa}}{O_i^t - NE_i^a} \times 100$	O_i^{sde} = Ocupantes en viviendas sin agua
		O_i^t = Total de ocupantes en viviendas particulares
		NE_i^a = Número de ocupantes en viviendas que no especificaron agua
I_{i6} =Porcentaje de viviendas con algún nivel de hacinamiento	$I_{i6} = \frac{V_i^h}{V_i^t - NE_i^h} \times 100$	V_i^h = viviendas con algún nivel de hacinamiento
		V_i^t =Total de viviendas particulares
		NE_i^h = Número de viviendas que no especificaron número de cuartos
I_{i7} =Porcentaje de ocupantes en viviendas con piso de tierra	$I_{i7} = \frac{O_i^{pt}}{O_i^t - NE_i^p} \times 100$	O_i^{pt} = Ocupantes en viviendas con piso de tierra
		O_i^t = Total de ocupantes en viviendas particulares
		NE_i^p = Número de ocupantes en viviendas que no especificaron el tipo de piso
I_{i8} =Porcentaje de población en localidades con menos de 5000 habitantes	$I_{i8} = \frac{P_i^{l < 5000}}{P_i^t} \times 100$	$P_i^{l < 5000}$ = Población que habita en localidades con menos de 5000 habitantes
		P_i^t = Población total
I_{i9} =Porcentaje de población ocupada con ingresos de hasta dos salarios mínimos	$I_{i8} = \frac{P_i^{l < 2sm}}{P_i^t} \times 100$	$P_i^{l < 2sm}$ =Población ocupada que no recibe ingresos o recibe hasta 2 salarios mínimos
		P_i^t =Población total

Fuente: Elaboración propia con base en (De la Vega *et al.*, 2010)

4.2. Metodología

4.2.1. Técnicas de análisis de datos espaciales

El trabajo empírico de muchos campos de la investigación científica involucra datos para los cuales la ubicación de las observaciones es un aspecto de suma importancia (Griffith, 2003, p. VII). Tal como se puede apreciar en el interés creciente sobre aspectos espaciales que muestran disciplinas como la economía, la sociología, la criminología y el urbanismo, entre otros. Para satisfacer la necesidad que genera la utilización de datos espaciales se han desarrollado herramientas que brindan un tratamiento adecuado de ellos.

EL análisis espacial se puede considerar como un estudio cuantitativo de fenómenos que se manifiestan en el espacio. Lo cual implica el estudio de la ubicación del fenómeno, áreas en las que se presenta, distancias, e interacciones expresadas por la primera ley de la geografía de Tobler (Anselin, 1989, p. 2). Dichos fenómenos se capturan por medio de datos espaciales, los cuales son la observación de una variable asociada a una localización en el espacio (Chasco~Yrigoyen, 2003, p. 17).

El análisis estadístico de datos espaciales es distinto del tradicional, principalmente por tres razones: 1) el patrón espacial de la ubicación, 2) la asociación espacial entre atributos observados en diferentes ubicaciones y 3) la variación del fenómeno observado ocasionada principalmente por la variación en la ubicación. Estos argumentos se vuelven los focos principales de la investigación (Griffith, 2003, p. VII). Considerando estos tres atributos de los datos espaciales, las herramientas estadísticas tradicionales no son eficaces pues no consideran los errores de medida ocasionados por ellos y generan problemas inferenciales.

El reconocimiento de los atributos de los datos espaciales ha generado el rápido crecimiento de un cuerpo sofisticado de técnicas especializadas. Estas se han desarrollado en diversos campos como la geoestadística, la estadística espacial y la econometría (Griffith, 2003, p. VII). Dichas técnicas son relevantes en la medida en la que en el estudio de los fenómenos se incorpora la localización de las observaciones. La localización se puede entender de diferentes

formas, principalmente en función del fenómeno de estudio.

Las observaciones espaciales se pueden representar por medio de tres tipos de datos: puntos, superficies y polígonos (Anselin, 1989, p. 2). Cada uno requiere un tratamiento estadístico distinto llamados análisis de patrones de puntos, geoestadística y análisis de polígonos (Anselin *et al*, 2006, S3), respectivamente. En el ámbito socioeconómico, los puntos, las líneas y los polígonos son las entidades fundamentales para representar el fenómeno espacial (Fischer, 2001, p. 18).

Los datos espaciales cuentan con un par de características de suma importancia, como lo son, la georeferenciación y la multidireccionalidad (Chasco~Yrigoyen, 2003, p. 18). La primera hace referencia a la posición absoluta o relativa de un fenómeno sobre el espacio y la segunda se centra en fenómenos espaciales como interacción, dependencia, externalidades, contagio, desbordamiento o replica (Chasco~Yrigoyen, 2003, p. 18). La multidireccionalidad también se entiende como las relaciones y vínculos que se presentan entre un punto en el espacio y otro ubicado en cualquier otro lugar.

Los datos espaciales se refieren a las unidades de análisis espacial. En la presente investigación las variables se relacionan con una estructura de polígonos, entonces, las herramientas de análisis se enfocan en los patrones espaciales de los atributos observados en ellos. Es importante resaltar que cuando se trabaja con polígonos, estos abarcan todo el espacio por medio de localizaciones discretas. Los atributos son datos fijos y en la mayoría de los casos el interés radica en la distribución de la variable asociada a cada localización.

De acuerdo con Anselin *et al* (2006) el análisis espacial se define ampliamente como un conjunto de tres componentes importantes: análisis exploratorio de datos espaciales, visualización y modelación espacial. Las líneas de separación entre los tres campos frecuentemente se desfiguran (Anselin *et al*, 2006, s3) pues dichos componentes se encuentran íntimamente interrelacionados.

4.2.2. Análisis Exploratorio de Datos y Análisis Exploratorio de Datos Espaciales

El presente trabajo busca adherirse, por medio del uso de técnicas de análisis espacial, al cuerpo disciplinario denominado “ciencia social espacialmente integrada” (Anselin *et al*, 2007, p. 13). Para ello, es necesario definir los datos espaciales como la observación de una variable asociada a una localización en el espacio geográfico. Los datos georeferenciados sirven para mejorar el entendimiento de la marginación y determinar la existencia de patrones de distribución espacial.

Cuando se trabaja con datos georeferenciados se deben utilizar herramientas que permitan detectar ciertas características espaciales en los mismos. Algunas de las características pudieran ser, tendencia, valores atípicos, esquemas de asociación, dependencia espacial y concentración espacial (Acevedo y Velásquez, 2008, p. 12). Las herramientas del Análisis Exploratorio de Datos Espaciales (AEDE) se utilizan para identificar relaciones sistemáticas entre variables cuando no existe un conocimiento claro sobre su distribución y su relación con el espacio geográfico (Acevedo y Velásquez, 2008, p. 13).

La metodología de la Minería de Datos Espaciales y el Análisis Exploratorio de Datos Espaciales (AEDE) será utilizada en el presente trabajo para realizar las primeras aproximaciones a la estructura de la información espacial. Permitirá caracterizar y clasificar diferentes situaciones socioeconómicas de fenómenos espaciales como la marginación. Siguiendo a Buzai y Baxendale (2008, p. 25) con este tipo de análisis es posible identificar clusters de exclusión social y grupos marginados y caracterizar las dinámica espacial que siguen dichos grupos. De la misma manera, es posible investigar si las variables cuentan con alguna estructura y si se distribuyen por medio de algún proceso de asignación en el espacio e identificar si hay variables que cuentan con un mayor peso relativo en ciertas áreas (Sánchez~Peña, 2006, p. 2).

La metodología se utiliza en una amplia gama de disciplinas que van desde el análisis de

políticas públicas, ciencias sociales, geología, marketing, entre otras (Lowe, 2003). Se trata de una extensión del análisis exploratorio de datos (AED), que ayuda a detectar e identificar de manera explícita las propiedades espaciales en datos georeferenciados. Surge por la necesidad de contar con técnicas adicionales en la detección de patrones espaciales en los datos, formulación de hipótesis *ex post* partiendo de la estructura estadístico-espacial, formulación de modelos espaciales y por último, se considera como una herramienta estadística importante al momento de vincular técnicas numéricas con fenómenos espaciales (Haining y Wise, 1997).

El análisis exploratorio de datos (AED) surge de la necesidad de “identificar relaciones sistemáticas en una variable cuando no existen expectativas claras sobre la naturaleza de esas relaciones” (Chasco~Yrigoyen, 2003, 28) y cuando no se tiene una hipótesis clara que permita realizar un contraste para su verificación. Así, el AED se convierte en la base del descubrimiento de patrones de comportamiento. Al mismo tiempo permite formular posibles hipótesis explicativas de los datos por medio de técnicas estadísticas (Chasco~Yrigoyen, 2003, p. 28).

Al introducir el carácter espacial en el análisis, el AED se convierte en análisis exploratorio de datos espaciales (AEDE) y sigue la misma lógica. Salvo que el interés radica en identificar la posible relación de una variable en un punto en el espacio con la misma variable ubicada en otro punto. Si no se tiene una hipótesis clara sobre la distribución de una variable en el espacio, el AEDE juega un papel de suma importancia al permitir identificar estructuras o regímenes espaciales.

Al igual que el AED, el AEDE utiliza de forma extensiva la visualización de diagramas que brindan cierta información que permite identificar patrones y características. Los sistemas de información geográfica sirven como punto de partida en el análisis espacial, al brindar una primer visualización de la distribución espacial y así, suponer la existencia de una estructura espacial o la existencia de aglomeraciones.

Las herramientas del AEDE, se agrupan en cuatro categorías: visualización de la distribución espacial, visualización de la asociación espacial, indicadores locales de asociación espacial

e indicadores multivariantes de asociación espacial (Anselin, 1998, p. 4).

4.2.3. Distribución espacial

La distribución espacial de un fenómeno busca determinar las condiciones y características de la ubicación de dicho fenómeno. Se basa en la asimetría de las relaciones espaciales, en la importancia de factores explicativos ubicados en otro lugar y en la importancia de la interdependencia espacial (Anselin, 1988, 7). En el AEDE, la visualización de la distribución espacial se fundamenta en elementos gráficos de la estadística clásica, utilizando a los mapas no sólo como herramientas de representación cartográfica (Chasco Yrigoyen, 2003, 37).

La base de la visualización de la distribución espacial se encuentra en el diseño de un dispositivo estadístico y gráfico, como la función de distribución acumulativa y el diagrama de caja en el caso de la estadística tradicional. En el AEDE, el objetivo es la visualización de la distribución de los valores observados en un subconjunto de ubicaciones en el espacio. El gráfico de caja espacial o “*box map*” es un dispositivo que caracteriza la distribución acumulativa para identificar los valores atípicos en el mapa (Anselin, 1998, 5).

La distribución espacial es el resultado de un conjunto de características de un determinado fenómeno y al mismo tiempo es el resultado de un determinado proceso. En el caso de los fenómenos socioeconómicos, la distribución espacial es el resultado de un proceso socioeconómico que determina las diferencias del fenómeno en diferentes ubicaciones (Bertaud y Malpezzi, 2003, 3). Por lo tanto, la distribución espacial es el resultado de un proceso socioeconómico espacial, el cual define una cierta estructura espacial en la que se refleja el fenómeno (Anselin, 1988, 42).

Al estudiar la distribución espacial de un fenómeno se requieren argumentos estadísticos y teóricos que permitan inferir cierta información sobre el conjunto de información que se pretende estudiar. Los procesos aleatorios implican la existencia de una variable aleatoria vinculada con una localización. Para poder inferir ciertas características sobre el proceso sub-

yacente se requiere un cierto grado de estabilidad en el proceso. Dicha estabilidad se obtiene asumiendo la existencia de estacionaridad, ergodicidad e isotropía (Anselin, 1988, 42).

La distribución de un fenómeno surge como parte de un proceso espacial. Un campo espacial aleatorio es un conjunto de valores aleatorios distribuidos en el espacio: $\{z(s) : S \in D\}$, donde: $z(s)$ = variable aleatoria localizada en s ; D = conjunto de todas las posibles localizaciones. Un proceso espacial aleatorio es el caso especial donde $Z(s)$ pertenece a una ubicación en el espacio (Anselin, 1988, 41). La noción de proceso espacial aleatorio se basa en la dependencia de valores observados en diferentes puntos en el espacio y se asume como resultado de un proceso subyacente desconocido (Anselin, 1988, 42).

La visualización de la distribución espacial del fenómeno se basa en herramientas geográficas y estadísticas. La geovisualización se muestra como el primer paso en el análisis espacial y sirve para observar la existencia aparente de ciertos patrones espaciales. También sirve como una primera aproximación hacia el estudio de los efectos espaciales, al brindar información visual que permita suponer la existencia de patrones espaciales.

La distribución se puede apreciar por medio del *boxplot* y el *boxmap*, los cuales son un “método de representación univariante que considera que los datos espaciales son realizaciones individuales de un proceso estocástico” (Chasco~Yrigoyen, 2003, 38). Una herramienta importante en la visualización de la distribución espacial se encuentra en los Sistemas de Información Geográfica (SIG). Los SIG permiten la visualización de las variables en el espacio. Muestran las diferencias espaciales de un fenómeno al vincular variables estadísticas con mapas geográficos. Al mismo tiempo se basan en la georelación de la estructura estadística en relación con la combinación de dos tipos de información: información estadística e información de la ubicación de dicho fenómeno (Anselin, 1992, 4).

Los sistemas de información geográfica sirven para representar las distribuciones de los fenómenos espaciales y las características que los describen, pues es claro que los procesos socioeconómicos no se presentan de manera homogénea en el espacio y no existen lugares que

no tengan nada en común (Goodchild, 1991, 196). Se presentan como una ventana al mundo y como un espejo de los procesos humanos para el aprendizaje y el razonamiento acerca de los fenómenos en el espacio (Goodchild, 1991, 198). Los SIG proveen los argumentos para una mejor comprensión de los fenómenos socioeconómicos (Goodchild, 1991, 196).

4.2.4. Efectos espaciales

Cuando se estudian fenómenos que se presentan en el espacio, los datos para su análisis se organizan por medio de unidades de observación espacial. Dichas observaciones se caracterizan por su ubicación absoluta o relativa en el espacio geográfico (Anselin, 1988, 11). Cuando se analizan este tipo de datos aparecen algunos efectos espaciales como la dependencia espacial y la heterogeneidad espacial (Moreno Serrano y Vayá, 2002, p. 84). La existencia de estos efectos implica la falta de independencia de las observaciones, la existencia de una estructura entre datos espaciales determinada por la localización y la existencia de inestabilidad estructural entre diferentes localidades (Chasco Yrigoyen, 2003, p. 13). Este tipo de datos se usan usualmente para el diseño y la planeación local y regional enfocadas al cálculo de impactos socioeconómicos (Anselin, 1988, p. 13).

4.2.4.1 Dependencia espacial

Siguiendo a Anselin (1988, p. 11), “la dependencia espacial se puede considerar como la existencia de una relación funcional entre lo que sucede en un punto en el espacio y lo que sucede en cualquier otro”. “La dependencia espacial está determinada por una noción de espacio relativo o localización relativa que realza el efecto de la distancia” (Chasco Yrigoyen, 2003, p. 9) y se debe a dos condiciones: errores de medida en las unidades espaciales contiguas y al fenómeno de la interacción espacial (Anselin, 1988, p. 11).

La dependencia espacial se puede expresar como un proceso espacial de la siguiente manera:

$$Y_i = f(y_1, y_2, \dots, y_N), \quad \forall i \in S$$

Mostrando como la existencia de un valor determinado en una ubicación depende de los valores de las unidades vecinas. La cual refleja la conexión en el espacio de diversos puntos ubicados en diferentes lugares. La conexión es el reflejo de la estructura, la que al mismo tiempo es el resultado de un proceso. La dependencia espacial necesita la determinación de las unidades espaciales que tienen una influencia en una unidad particular, lo que implica la comprensión de los vecinos y vecinos mas cercanos (Anselin, 1988, p. 16).

De acuerdo con Anselin (1988, p. 16) el entendimiento de la dependencia radica en la comprensión de los vecinos en el espacio. Lo cual se logra por:

$$P[x_i | x] = P[x_i | x_j]$$
$$x_j \quad \forall j \in J$$

Lo que significa que la probabilidad condicional del valor de una variable x_i ubicada en el punto i , depende del valor de x_j en otras localizaciones pertenecientes a un conjunto de vecinos J .

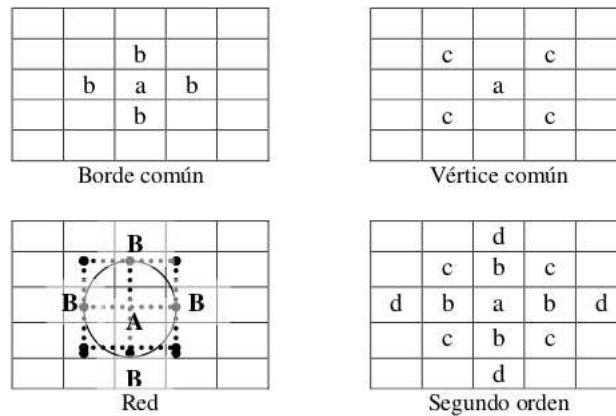
La concepción del vecino implica la incorporación de una estructura adicional al conjunto de datos. Esto se logra combinando la noción de dependencia estadística con una noción del espacio. Es decir, magnitudes estadísticas relacionadas con la distancia y la ubicación relativa (Anselin, 1988, p. 17). El conjunto de vecinos para cada unidad espacial, se representa por medio de una asociación espacial, la cual se captura por medio de la matriz de contigüidad.

La matriz de contigüidad captura la estructura subyacente de los vecinos expresada por valores 0 y 1, donde se le asigna el valor 1 si dos unidades espaciales comparten una frontera común mayor a cero, por lo que se consideran unidades contiguas y cero en cualquier otro caso. Los criterios de contigüidad se basan en el supuesto de la existencia de una estructura espacial directamente relacionada con la interacción de dos unidades de observación. Así, la elección de un vecino relevante se basa en aquellas observaciones que interactúan entre si por medio de la primera ley de la geografía de Tobler. Estas interacciones potenciales se identifican

por medio de una distancia relativa, una frontera común, una función de distancia inversa, o bien, un patrón de interacción definido *a priori* (Anselin, 1992).

Cuando el criterio de contigüidad se basa en frontera común mayor a cero, existen tres criterios. Estos son *rook*, *bishop* y *queen*, que significan respectivamente, frontera común, vértice común y la unión de ambos. Estos criterios son los mas utilizados en la literatura. Por ejemplo en Anselin (1998), en Anselin (1999), Anselin *et al* (2007), Caliber (2001), Mennis y Guo (2009), Le˜Gallo y Ertur (2003) y O’Loughlin (2002), por citar algunos.

Figura 4.2: Criterios de contigüidad



Fuente: Anselin (1988, 18)

Como se aprecia en la figura anterior, existen vecinos de orden superior, lo que se consigue introduciendo el concepto de rezago espacial. El cual, básicamente, se refiere al concepto de una unidad rezagada k unidades espaciales. En este concepto se vuelve a presentar el problema de la multidireccionalidad, pues el rezago puede ir hacia cualquier punto en el mapa.

El concepto de rezago espacial consiste en un promedio ponderado de los valores de una variable en las localizaciones vecinas con unos pesos y ponderaciones fijas y dadas de

forma exógena. El retardo espacial considera la suma ponderada de los valores que toma una variable en una unidad espacial en función de un determinado criterio de contigüidad (Chasco~Yrigoyen, 2003, p. 61). El rezago espacial se define como:

$$L^s x_i = \sum_j w_{ij} x_j \quad \forall j \in J$$

En notación matricial:

$$L^s x = W_s x$$

w_{ij} se refiere a la matriz de pesos espaciales y se refiere a una generalización de la matriz de contigüidades. El criterio de contigüidad por medio del cual se definen las interacciones se captura por medio de una matriz de interacciones. En la que los valores se definen como 1 si se cumple el criterio de contigüidad y 0 en cualquier otro caso. La matriz de interacciones espaciales δ_{ij} es simétrica (Chasco~Yrigoyen, 2003, p. 55).

donde: $\delta_{ij} = 1$, si se cumple el criterio de contigüidad

$\delta_{ij} = 0$ en cualquier otro caso

$\delta_{ij} = \delta_{ji}$

$\delta_{ii} = 0$, porque ninguna unidad espacial puede ser vecina de si misma.

Figura 4.3: Matriz de contigüidad

	1	2	3	
	4	5	6	
	7	8	9	

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

Fuente: Chasco (2003, 55)

La matriz de pesos es una generalización de la matriz de interacciones. La matriz \mathbf{W} permite la elección de un conjunto de ponderaciones que reflejen la interacción entre dos unidades espaciales. Define la estructura de interdependencias en un sistema regional y recoge el efecto de la ubicación i sobre la ubicación j a través de un peso o una ponderación w_{ij} (Chasco y Yrigoyen, 2003, p. 57).

La matriz de pesos \mathbf{W} sigue la misma lógica que la matriz de interacciones espaciales, salvo que la primera deja de ser una matriz binaria. Con esto, los pesos espaciales $w_{ij} \neq 0$ indican la existencia de interacción espacial con una determinada ponderación. La definición de la matriz de pesos \mathbf{W} depende en gran medida del tipo de fenómeno espacial que se quiera capturar y en la literatura se puede encontrar una gran diversidad de matrices de pesos, pues todas ellas responden a necesidades específicas.

Como ya se ha mencionado, las interacciones responden a la necesidad de reflejar una determinada estructura espacial, resultado de un proceso. Las interacciones son, al mismo tiempo, el reflejo de la existencia de dependencia espacial. Si existe dependencia espacial, entonces “una variable se encontrará espacialmente autocorrelacionada cuando los valores ob-

servados en un lugar determinado dependen tanto de factores externos como de los valores de la misma variable ubicada en diferentes lugares” (Chasco Yrigoyen, 2003, p. 9). Es necesario aclarar que la dependencia espacial y la autocorrelación espacial no son exactamente lo mismo. La diferencia estriba en que “la autocorrelación espacial es una forma de dependencia espacial definida estadísticamente a través de los primeros momentos de la distribución conjunta de una variable espacial” (Chasco Yrigoyen, 2003, p. 49).

4.2.4.1. Autocorrelación espacial

Una de las diferencias entre los conceptos de dependencia y autocorrelación espacial es el carácter operativo de la segunda. Esto se consigue al poder construir herramientas estadísticas que permitan medir su intensidad. El principio que subyace en el concepto de autocorrelación es la existencia de un estadístico para un conjunto de datos que se comporte bajo la hipótesis nula de no autocorrelación espacial. Esto implica que el espacio no importa o que la asignación de ciertos valores en ubicaciones particulares es irrelevante. Con esto, se consigue que las pruebas de autocorrelación espacial se basen en la magnitud de un indicador que combine el valor observado en cada ubicación con los valores de las ubicaciones vecinas (Anselin, 1992, p. 10).

De acuerdo con Griffith (2003, p. 3), “la autotocorrelación espacial es la existencia de correlación entre los valores de una misma variable que se atribuyen estrictamente a la proximidad de dichos valores en el espacio geográfico”. Las variables que reflejan fenómenos espaciales muestran algún tipo de organización en el espacio, por lo tanto, la inexistencia de autocorrelación espacial significa un fenómeno geográficamente aleatorio y paisajes caóticos ⁶ (Griffith, 2002, p. 3).

Se considera que esta característica se presenta cuando el estadístico para un patrón geográfico toma valores extremos comparados contra lo que se podría esperar bajo la hipótesis nula de inexistencia de autocorrelación espacial. Un aspecto de suma importancia en el diseño

⁶Por caóticos se refiere a la falta de procesos determinísticos en la distribución del fenómeno.

de una prueba de autocorrelación es, sin duda, la distribución estadística que siga la prueba, ya que la interpretación de los valores extremos depende de dicha distribución. Hay tres enfoques para determinar la distribución de una prueba, el supuesto de normalidad, el supuesto de aleatoriedad y el supuesto de permutación. (Anselin, 1992)

Hay dos razones por las que la medición de la autocorrelación espacial es importante, primero porque su existencia viola los supuestos clásicos de la estadística tradicional y segundo porque describe el patrón general a lo largo del paisaje geográfico (Griffith, 2002, p. 4). Una razón adicional surge por el hecho de que un gran número de variables empleadas en las ciencias sociales tienden a presentar autocorrelación espacial, ya sea positiva o negativa (Griffith, 2003, p. 5). Las pruebas de autocorrelación espacial se diseñan para cuantificar el grado de aglomeración de valores, ya sean similares o diferentes. Además permiten realizar inferencia estadística basadas en la distribución definida por uno de los tres enfoques antes mencionados. (Anselin, 1992).

El análisis de autocorrelación espacial cuenta con dos perspectivas, la global y la local. La primera se basa en un contraste diseñado para detectar la existencia de tendencia o estructuras espaciales generales en la distribución de una variable sobre un espacio geográfico completo. En la segunda, el contraste se define como la concentración de valores altos o bajos de una variable en relación con el valor medio de la misma en un lugar específico del espacio global analizado (Chasco Yrigoyen, 2003, p. 62).

4.2.4.1.1. Autocorrelación espacial global

Siguiendo a Anselin (1992), la medida de autocorrelación espacial más usada es el estadístico I de Moran, el cual se considera como una medida global.

$$I = \frac{N \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_i (x_i - \bar{x})^2} \quad (4.1)$$

donde: N = Número de observaciones de la variable x

w_{ij} = matriz de pesos espaciales

x_i = variable ubicada en i

x_j = variable ubicada en j

\bar{x} = media de la variable x

$S_0 = \sum_i \sum_j w_{ij}$, factor de normalización

En términos matriciales, se tiene:

$$I = \frac{N}{S_0} \frac{z'wz}{z'z} \quad (4.2)$$

Donde z_i son desviaciones de la media. El índice I de Moran muestra el grado de asociación entre el vector z y el vector wz , llamado vector de rezagos espaciales (LeñGallo y Ertur, 2003, p. 4). La interpretación del índice I de Moran no es directa, pues no se encuentra acotado entre el rango $[-1,1]$ (Griffith, 2003, p. 5). La interpretación de la prueba se basa en la media del propio indicador. $E[I] = \frac{-1}{N-1}$, valores superiores a la media indican autocorrelación espacial positiva y menores a la media indican autocorrelación espacial negativa. El rango del índice de Moran depende del conjunto de valores propios obtenidos de la forma cuadrática asociada al propio índice, tomando el máximo y el mínimo como valores extremos (DeñJong *et al*, 1984, p. 20). El conjunto de valores propios depende de la matriz de pesos y del proceso espacial específico que caracteriza a la matriz de contigüidad (Tiefelsdorf, 1998, p. 78)

4.2.4.1.2. Autocorrelación espacial local

En el análisis espacial, un aspecto de suma importancia es la identificación de patrones locales de asociación espacial. Existen un conjunto de indicadores que persiguen tal objetivo. Los indicadores locales de asociación espacial o LISA sirven para dos propósitos. Primero, se pueden interpretar como indicadores que capturan la falta de estacionaridad a nivel local, o bien la presencia de una estructura local; el segundo aspecto es que se puede utilizar para evaluar la influencia de ubicaciones individuales en la magnitud estadística global y permite

identificar valores extremos (Anselin, 1995, p. 1).

Los indicadores LISA permiten la descomposición de los indicadores globales como el I de Moran. De acuerdo con Anselin (1995), un indicador local de asociación espacial es aquel que satisface las siguientes condiciones:

- Para cada observación, los LISA brindan un indicio del grado de significancia del agrupamiento espacial de valores similares en torno a dicha observación.

$$Prob[L_i > \delta_i] \leq \alpha_i$$

Donde: δ_i = valor crítico

α_i = es un nivel de significancia o pseudo-significancia

- La suma de los LISA para cada observación es proporcional al indicador de asociación espacial global

$$\sum_i L_i = \gamma \Delta$$

Donde: L_i = indicador local de asociación espacial

, γ = factor de escala

Δ = indicador global de asociación espacial

El uso de los indicadores locales de asociación espacial responde a dos cuestiones. La existencia de inestabilidad estructural espacial, particularmente, cuando se trabaja con un gran número de observaciones espaciales y al hecho de que las herramientas para medir el grado de asociación espacial global ignora esta posible inestabilidad (Anselin, 1995, p. 2). Además es razonable pensar que el grado de autocorrelación no es uniforme en todo el mapa. Al contrario, varía de acuerdo con la localización por lo que se encuentran subregiones con un grado alto de autocorrelación espacial y otras con grado bajo (Lee y Wong, 2001, p. 168).

Un par de aspectos adicionales que motivan el diseño de los LISA es la existencia de patrones locales contrarios a la tendencia global y la necesidad de inferir la significancia estadística de un patrón de asociación espacial para una ubicación determinada (Anselin, 1995,

p. 3). También porque permiten identificar áreas en las que los valores de la variable son extremos o geográficamente homogéneos (Oliveau *et al*, 2005).

En general los indicadores locales de asociación espacial operan bajo la hipótesis nula de no autocorrelación espacial local. En particular, el índice de autocorrelación local de Moran, se define como (Anselin, 1995)

$$I_i = z_i \sum_j w_{ij} z_j \quad (4.3)$$

donde: z_i y z_j son desviaciones estandarizadas de la media con la forma:

$$z_i = \frac{(x_i - \bar{x})}{\delta}$$

donde δ es la desviación estándar de x_i .

El índice local de Moran para una unidad espacial i es la desviación de la media de i multiplicada por la suma de los productos de las desviaciones de la media de j por los pesos de los vecinos de i (Lee y Wong, 2001, p. 168). La interpretación del índice local se realiza contrastando los valores estandarizados contra el valor esperado (Lee y Wong, 2001, p. 170) y las pruebas de significancia se realizan con base en los primeros momentos de la distribución del propio índice. De acuerdo con Anselin (1995), el valor esperado del I local de Moran bajo la hipótesis nula de aleatoriedad es:

$$E[I_i] = \frac{-w_i}{n-1} \quad (4.4)$$

Capítulo **5**

Minería de datos espaciales, análisis
exploratorio espacial y geovisualización

El análisis exploratorio que se presenta a continuación se realiza empleando el *software* Geoda 3.2 y QuantumGIS. Primero se realiza un análisis univariado del índice de marginación, el cual contempla 3 apartados, distribución espacial empleando herramientas de geovisualización, autocorrelación espacial global y autocorrelación espacial local por medio de minería de datos espaciales y AEDE. Después se reproduce el mismo ejercicio para cada componente del índice.

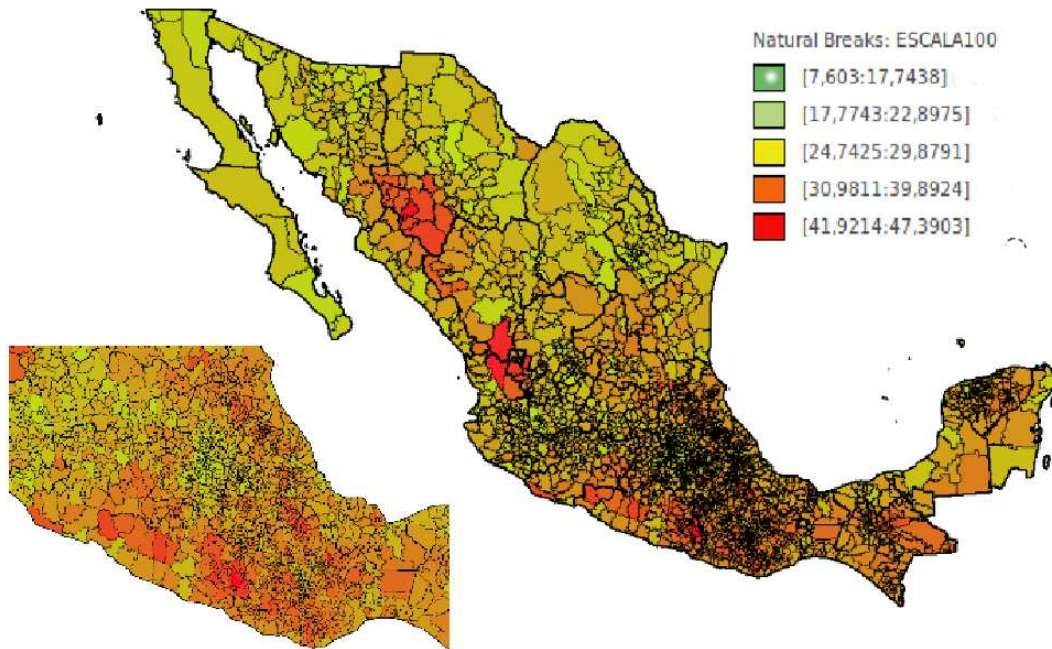
El marco de referencia de la minería de datos se emplea para contrastar y extraer información interesante en un conjunto de datos relativamente grande. En particular, la minería de datos espaciales permite extraer información relevante y patrones de comportamiento espacial en conjuntos de datos de gran tamaño. La minería de datos espaciales emplea diferentes tipos de algoritmos que permiten extraer información de dicho conjunto. Algunos ejemplos de técnicas son las redes neuronales, el *clustering* y las reglas de asociación (Zhu y Davidson, 2007; Chen *et al*, 2011; Shekhar *et al*, 2003). El análisis exploratorio de datos espaciales realizado en el presente capítulo se enmarca dentro de la metodología de la minería de datos espaciales, principalmente porque pretende extraer información relevante en un conjunto de datos espaciales de gran tamaño.

5.1. Análisis univariado del índice de marginación

5.1.1. Distribución espacial

La distribución espacial del índice de marginación se realiza por medio de un mapa degradado, donde el valor del índice va de menor a mayor, resaltando valores menores con verde y altos con rojo. En la figura 5.1 la intensidad del color se debe al sesgo estadístico que muestra la variable, tal como se aprecia en la figura 1.1. Esta primer geovisualización de la distribución del índice pareciera indicar que hay homogeneidad en la distribución de la variable.

Figura 5.1: Distribución espacial del índice de marginación

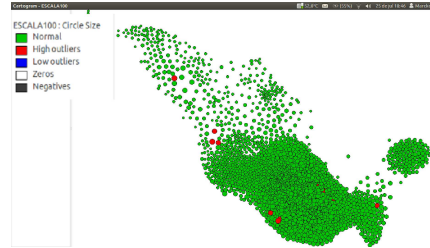


Fuente: Elaboración propia con base en el índice de marginación en México, en CONAPO, 2010

Por medio del uso del cartograma, el cual es una representación que permite identificar el valor de las unidades espaciales representadas por la variable en cuestión, se permite una apreciación de la distribución de la variable. Aunque en este caso, parece mostrar la falta de

desigualdades en la distribución.

Figura 5.2: Cartograma del índice de marginación

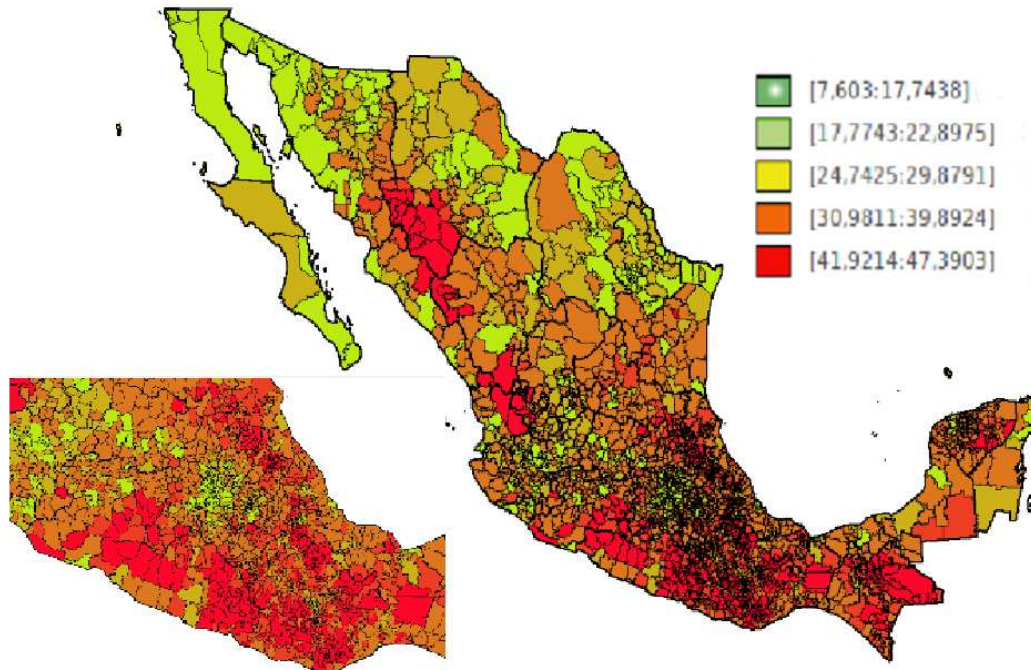


Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

En los ejercicios siguientes se mostrará como, a pesar de la primer impresión que muestra la distribución del índice, al profundizar el análisis con herramientas especializadas se encuentran patrones de distribución espacial bien definidos.

El índice de marginación se agrupa en estratos, los cuales representan los diferentes grados de marginación. Tal como se aprecia en la figura 1.5 y en la tabla 1.2. Se puede apreciar que los estratos son intervalos irregulares por lo que la representación del índice por cuantiles no es tan relevante si se compara con la distribución del grado de marginación, aunque brinda una perspectiva enriquecedora cuando se realiza *linking* con herramientas estadísticas.

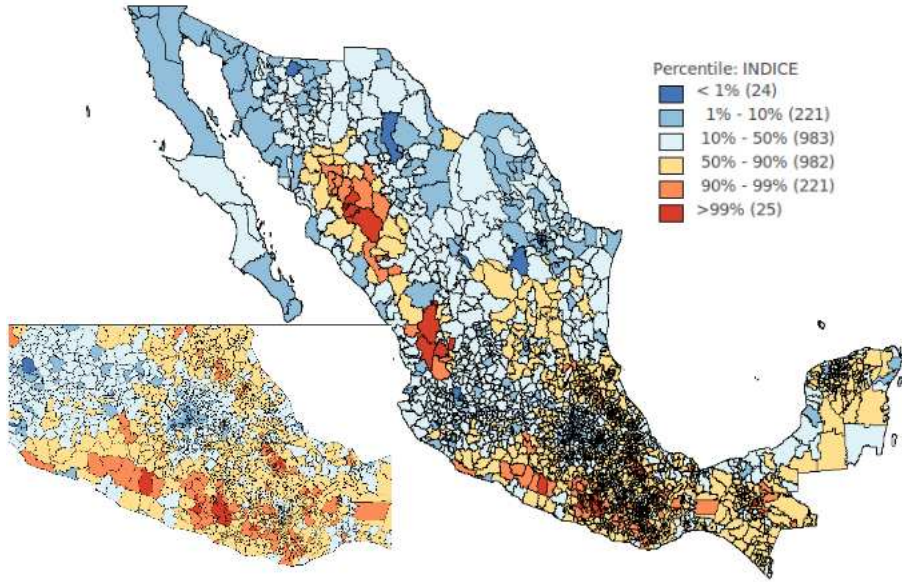
Figura 5.3: Distribución espacial del grado de marginación



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

Existen tres herramientas de análisis espacial que permiten detectar patrones en la distribución espacial de un fenómeno. Estas herramientas son: *box map*, mapa de percentiles y mapa de desviación estándar. El primero representa la distribución espacial de la variable en función de la distribución del diagrama de caja. Distribuye las unidades espaciales en cuartiles y ubica valores atípicos tradicionales.

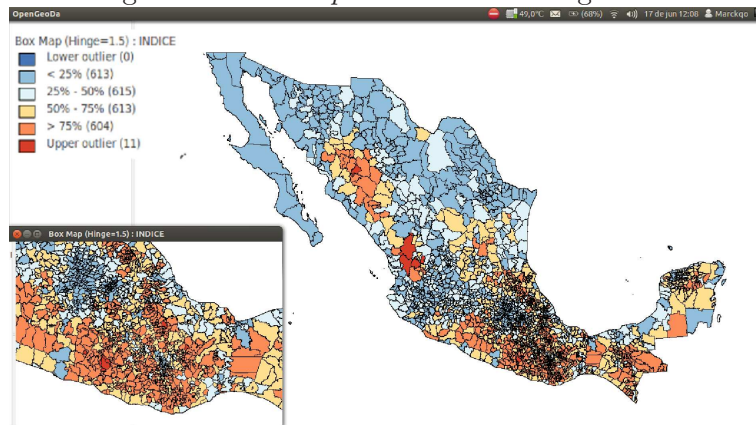
Figura 5.4: Mapa de percentiles del índice de marginación



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

El mapa de percentiles permite observar la distribución de la variable en grupos definidos.

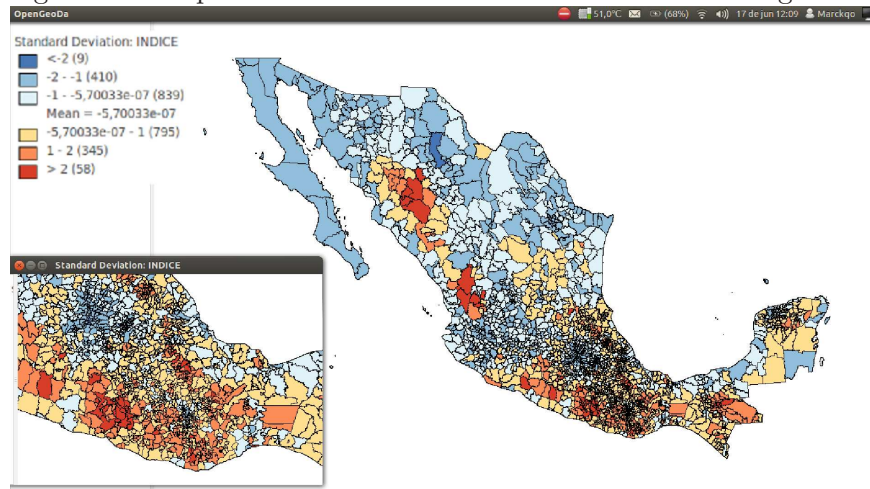
Figura 5.5: *Boxmap* del índice de marginación



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

EL *box map* es una herramienta de suma importancia ya que muestra la ubicación espacial de los cuartiles representados en el diagrama de caja, permite ubicar los valores atípicos.

Figura 5.6: Mapa de desviaciones estándar del índice de marginación



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

EL mapa de desviaciones estándar considera una representación a partir de desviaciones de la media. En función de la distribución estadística de la variable se toma cierto número de desviaciones.

Las tres herramientas recién expuestas muestran que la distribución de la variable no es casual, al contrario, existen agrupaciones bien definidas. Se aprecia una franja de municipios con valores bajos al norte de México. Se muestra una aglomeración al sur y sureste, aunque es importante resaltar que al norte existe una aglomeración con valores muy altos. Las herramientas utilizadas brindan una perspectiva global de la distribución espacial del índice de marginación.

5.1.2. Autocorrelación espacial global

Pare verificar si los patrones observados por las herramientas de distribución espacial son aleatorios se recurre al estadístico I de Moran para determinar la existencia de autocorrelación espacial global. Se busca determinar si la existencia de las aglomeraciones encontradas son resultado de la dependencia espacial con los vecinos o si es una distribución casual que se observa por azar.

$$I = \frac{N}{S_0} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (5.1)$$

$$I = 0,7026 \quad (5.2)$$

El criterio de decisión en el estadístico I de Moran depende del valor de la esperanza del propio estadístico, $E[I] = \frac{-1}{N-1}$, si $I > E[I]$ entonces existe autocorrelación espacial positiva y si es menor entonces existe autocorrelación espacial negativa.

$$I = 0,7026 \quad (5.3)$$

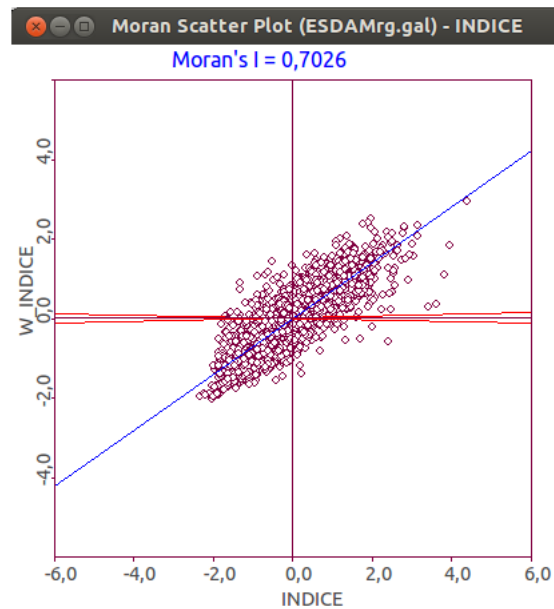
$$E[I] = -0,0004 \quad (5.4)$$

Este hallazgo es de suma importancia pues sugiere que el valor observado en una unidad espacial puede depender de manera positiva de los valores observados en los vecinos de primer orden bajo un criterio de contigüidad tipo *queen* y un *pseudo p* valor de 0.0001 obtenido por medio de 9,999 permutaciones. Este hallazgo debe ser tomado con precaución debido al número de unidades espaciales empleado.

Siguiendo a Anselin (1996), cuando se trabaja con una cantidad relativamente grande de observaciones espaciales, los indicadores de autocorrelación espacial no son tan eficientes pues resumen la distribución espacial completa en un solo número. Cuando el número de observaciones es grande puede ocurrir que el grado de asociación espacial entre las observaciones no muestre la posible existencia de inestabilidad o bien la falta de estacionariedad.

Para incorporar el argumento señalado en Anselin (1996), se diseñó una herramienta de gran utilidad, conocida como Diagrama de dispersión de Moran o *Moran Scatterplot*. El indicador I de Moran es equivalente a un coeficiente de regresión del rezago contra la variable. El I de Moran permite observar el grado de asociación entre la variable en desviaciones de la media y el rezago espacial por medio de un diagrama de dispersión (Anselin, 1996, p. 117). En este sentido, el *scatterplot* de Moran es una representación del promedio estandarizado de los vecinos contra la variable estandarizada con una línea de regresión con pendiente igual al I de Moran.

Figura 5.7: Diagrama de dispersión de Moran del índice de marginación 2010



Fuente: Elaboración propia con base en el índice de marginación en México,
CONAPO, 2010

Los cuadrantes I y III indican las observaciones que muestran autocorrelación espacial

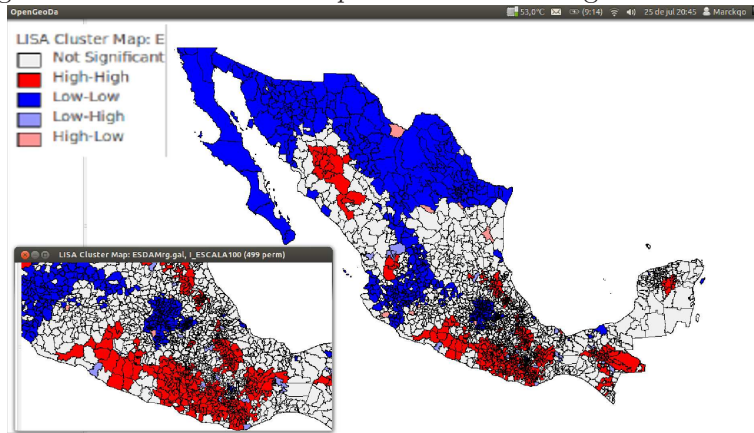
positiva, los cuadrantes II y IV muestran observaciones con autocorrelación espacial negativa. En este caso, el cuadrante II muestra la existencia de *outliers espaciales*, los cuales son observaciones con un valor bajo pero con vecinos con valores altos. El cuadrante IV muestra lo contrario, unidades con valores altos y vecinos con valores bajos. Los *outliers* espaciales son observaciones que contrastan no por su valor si no por la relación con sus vecinos.

La figura 5.7 sugiere la presencia de autocorrelación espacial positiva, ya que como se observa la mayoría de las unidades espaciales se encuentran en los cuadrantes II y III. Dichos cuadrantes sugieren la presencia de autocorrelación espacial positiva ya que, en ellos, se encuentran las unidades con vecinos que muestran valores similares. Los cuadrantes I y IV representan unidades con vecinos que muestran valores contrarios, aunque existen unidades que parecen mostrar dicho comportamiento, la gran mayoría se encuentra en los cuadrantes II y III. Este hallazgo se reafirma por la pendiente de la línea, dicha pendiente representa el I de Moran por lo que esta herramienta de visualización también sugiere la existencia de autocorrelación espacial positiva.

5.1.3. Autocorrelación espacial local

Para resaltar el hecho de que el patrón de asociación espacial puede no ser uniforme en todo el conjunto de datos, se recurre al análisis de autocorrelación espacial local. Dentro de las herramientas diseñadas para analizar autocorrelación espacial local se encuentran los Indicadores Locales de Asociación Espacial o LISA. Dentro de ellos se encuentran dos herramientas: *LISA Significance Map* y *LISA Cluster Map*. Ambas herramientas se encuentran íntimamente relacionadas ya que las dos muestran las unidades espaciales que cuentan con indicadores locales de Moran estadísticamente significativos al 95% de confianza. EL primero muestra todas las unidades espaciales estadísticamente significativas y el segundo diferencia entre *clusters* y *outliers espaciales*. A continuación se muestra el *cluster map* como herramienta de exploración ya que, dentro de MDE, dicho instrumento sirve para sugerir la existencia de patrones de distribución.

Figura 5.8: LISA *Cluster Map* del índice de marginación escalado



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

En la figura 5.8 se muestran los núcleos de los *clusters* representados por *high-high* (H-H) de color rojo como *clusters* con valores altos, *low-low* (L-L) con color azul, como *clusters* con valores bajos y las unidades marcadas como *high-low* (H-L) de color rosa y *low-high* (L-H) de color violeta representan *outliers espaciales*.

En la tabla 5.1, se muestra el número de municipios por tipo de aglomeración. Con base en los LISA, se encuentran 4 tipos de aglomeraciones, H-H, L-L, L-H y H-L. Los municipios que se muestran en las columnas H-H y L-L en la tabla 5.1 son aquellos en los que la exploración por medio de Minería de Datos espaciales (MDE) sugieren la presencia de autocorrelación espacial positiva. Las columnas H-L y L-H, sugieren la presencia de *outliers* espaciales.

Cuadro 5.1: Municipios por tipo de aglomeración

Municipios por tipo de aglomeración				
	High-High	Low-Low	Low-High	High-Low
Municipios	489	604	24	11

Fuente: Elaboración propia con base en el índice de marginación en México,

Conapo, 2010.

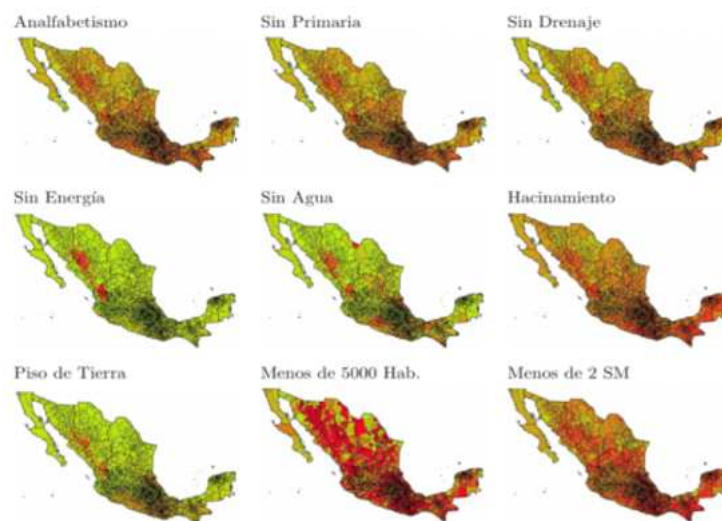
5.2. Análisis univariado de los componentes del índice

En la presente sección se reproduce el ejercicio anterior para cada uno de los componentes del índice de marginación 2010.

5.2.1. Distribución espacial

En la siguiente figura se muestra la distribución de los componentes. Se aprecian diferencias importantes que contrastan a la vista como las unidades con valores altos, aunque en algunos parece que brinda la misma impresión que la medida global. Resaltan las distribuciones de los componentes “localidades con menos de 5000 habitantes” y “habitantes con ingresos menores a dos salarios mínimos”, los cuales muestran valores muy altos en las unidades espaciales. Otra distribución que resalta es “localidades sin energía eléctrica” que muestra una aparente homogeneidad en el indicador aunque exhibe un grupo de municipios con valores muy altos en relación con el resto.

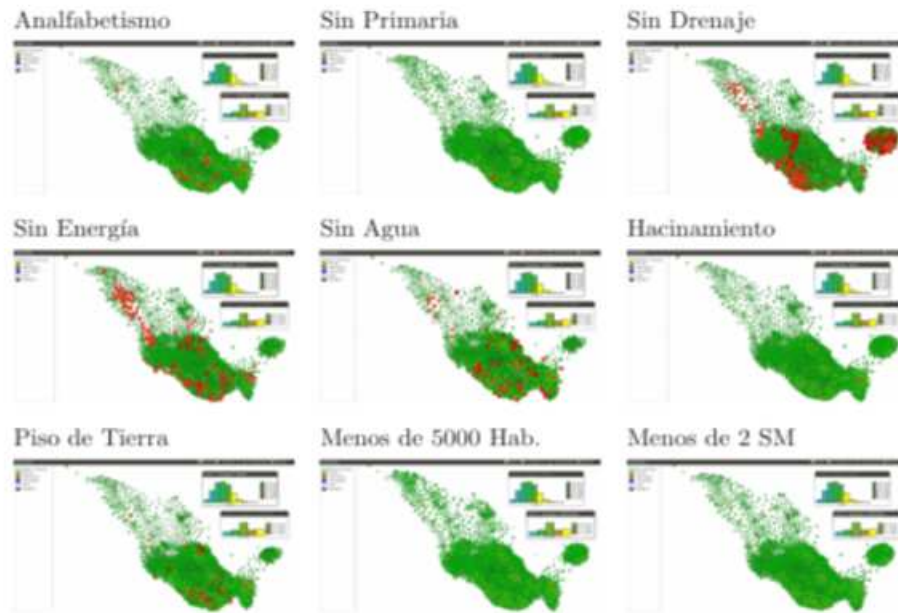
Figura 5.9: Distribución espacial de los componentes



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

A continuación se muestran los cartogramas de los componentes con un *linking* con el grado muy alto de marginación y el índice de marginación. Se aprecian variaciones importantes en relación con el nivel medio de las observaciones. Una característica importante en el indicador localidades sin drenaje muestra que en la península de Yucatan la gran mayoría de las localidades no cuenta con drenaje resultado del tipo de suelo de la zona, el cual impide la implementación de redes de drenaje en la zona. También se aprecian que los cartogramas muestran mejor las diferencias de las localidades sin energía y sin agua entubada.

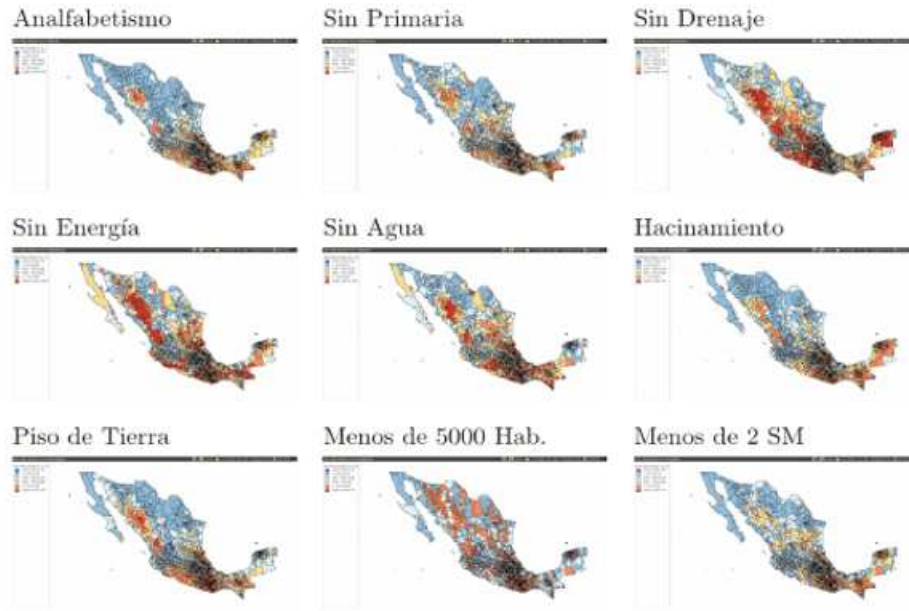
Figura 5.10: Cartogramas de los componentes



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

La distribución espacial de los componentes analizada por medio de los *box map* muestra diferencias importantes en cada uno de ellos. Esto permite estudiar la distribución espacial de los componentes estudiando agrupaciones en cuartiles.

Figura 5.11: *Boxmap* de los componentes



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

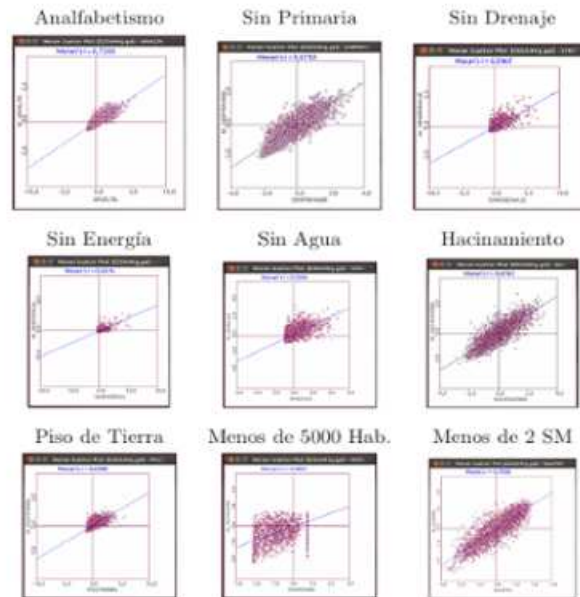
La comparación entre el *boxmap* de los componentes muestra patrones similares, salvo en algunos indicadores como “sin drenaje”, “sin energía” y “localidades con menos de 5000 habitantes”. EL componente “menos de dos salarios mínimos” parece ser el indicador con una distribución mas homogénea, aunque parece mostrar una distribución del tipo *norte-sur*.

5.2.2. Autocorrelación espacial global

Los indicadores I de Moran global para cada uno de los componentes sugieren la existencia de autocorrelación global positiva. Dichos indicadores oscilan entre 0.3832 para “menos de 5000 habitantes” y 0.7200 para “analfabetismo”, representando el I de Moran menor y mayor, respectivamente. Este hallazgo se corresponde con el I de Moran del índice de marginación, que también sugiere autocorrelación espacial positiva.

Los diagramas de dispersión de Moran muestran el grado de autocorrelación espacial en cada uno de los componentes del índice. Como resultado de ello, se aprecia que en los componentes “Sin drenaje”, “Sin energía”, “Sin agua” y “piso de tierra” se aglomeran en el cuadrante II, por lo que muestra la existencia de unidades espaciales con valores altos en los componentes rodeadas de unidades también con valores altos. Es interesante el diagrama del componente “Menos de 5000 habitantes” ya que parece mostrar aleatoriedad, sin embargo, como la mayoría de observaciones se encuentran en el cuadrante III, la pendiente de la regresión representada por el I de Moran sugiere la presencia de autocorrelación espacial positiva.

Figura 5.12: Digramas de dispersión de Moran de los componentes



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

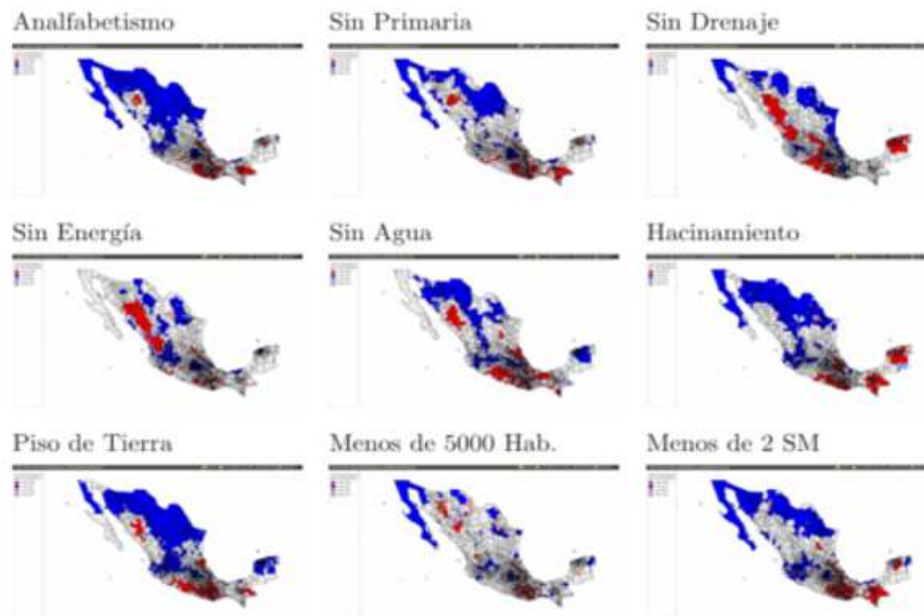
5.2.3. Autocorrelación espacial local

Los LISA muestran, al igual que los *boxmap*, patrones similares entre el índice de marginación y los componentes del mismo. Al igual que en la comparación de los *box map*, los indicadores con diferencias observables en la distribución son “sin energía” y “localidades con menos de 5000 habitantes”. En el primero, los núcleos de los clústeres disminuyen significativamente, aunque el núcleo del clúster *high-high* aumenta de tamaño considerablemente. En el segundo, disminuye el tamaño de los clusters aunque el patrón parece similar. Se aprecia, además, que los componentes “Analfabetismo”, “Hacinamiento” y “Piso de tierra”, parecen mostrar patrones de distribución similares al índice de marginación. Esto puede ser resultado del alto grado de correlación entre dichos componentes y el índice global, tal como se aprecia en el cuadro 1.1, en la sección 2.1 del capítulo 2.

En el *cluster map* de los componentes se aprecia un patrón global en el que se resalta al

norte un gran *cluster* de baja vulnerabilidad en los componentes. Por otro lado, es importante resaltar que la sierra Tarahumara, las Huastecas y los Estados de Guerrero, Oaxaca y Chiapas muestran *clusters* de alta vulnerabilidad en todos los componentes. Es de resaltar que la Sierra Tarahumara parece mostrar un patrón local de alta vulnerabilidad socioeconómica.

Figura 5.13: LISA *cluster map* de los componentes



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

5.3. Hallazgos principales

La exploración del índice de marginación en México por medio de MDE ha brindado un conjunto de hallazgos, los cuales se muestran como resultado de la inclusión explícita de la dimensión espacial. El primer hallazgo es un patrón sugerido de autocorrelación espacial positiva, tanto en el índice de marginación como en sus componentes. Este hallazgo es de suma importancia ya que sugiere que el índice de marginación en los municipios de México cuentan con un proceso de dependencia espacial positiva. Es decir, que los indicadores relacionados con las unidades espaciales estudiadas cuentan con una cierta influencia de los indicadores rela-

cionados con las unidades vecinas, esto bajo un criterio de contigüidad tipo *queen* de primer orden.

El segundo hallazgo se obtuvo por medio del análisis de autocorrelación espacial local. Por medio de los LISA se identificaron agrupaciones espaciales con valores similares, tanto en el índice de marginación como sus componentes. Se encontraron agrupaciones de baja marginación y baja vulnerabilidad socioeconómica. Estas agrupaciones se ubican al norte del país, principalmente en todos los estados fronterizos y en las zonas metropolitanas de Guadalajara y el Distrito Federal. El otro tipo de agrupaciones son las que muestran un nivel alto de marginación y alta vulnerabilidad socioeconómica. Estas agrupaciones se encuentran en el sur del país, en los estados de Guerrero, Chiapas y Oaxaca, en las haustecas y en la sierra tarahumara.

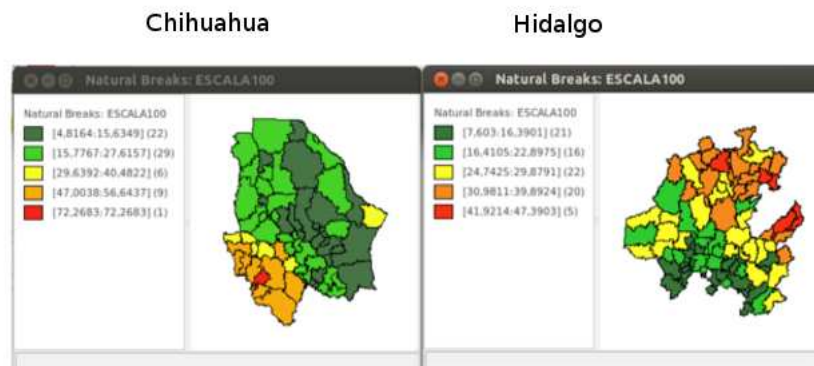
Los patrones sugeridos por MDE parecen mostrar que el proceso de dependencia espacial positiva se da tanto en el índice de marginación como en los componentes de dicho índice. El análisis previo se enmarca dentro de la metodología de Minería de Datos Espaciales y permite la exploración y la sugerencia de patrones. Debido al número de unidades espaciales empleadas, la inferencia estadística del I de Moran debe ser tomada con precaución aun cuando esta se realice por medio de 9,999 permutaciones espaciales, ya que el número de unidades espaciales empleado es demasiado grande. El análisis previo se muestra como una primera aproximación en la exploración del índice de marginación en México y sus componentes para el año 2010.

Para explorar si los patrones sugeridos por MDE en el índice de marginación se mantienen para una escala menor, se procederá, en la siguiente sección, a realizar una exploración del índice por medio de un subconjunto de unidades espaciales. Se hará un estudio de casos específicos ya que, a nivel nacional, la MDE tan solo brinda sugerencias de patrones, tanto para el índice como para los componentes.

5.4. Análisis de los casos de Chihuahua e Hidalgo

Utilizando los patrones sugeridos en la sección anterior, se seleccionaron dos entidades federativas para reproducir el ejercicio. La selección de las entidades se basó, además, en dos criterios. El primero fue el número de municipios contenidos en dichas entidades, ya que se seleccionaron entidades con un número intermedio de observaciones. El segundo se basa en el análisis de autocorrelación espacial local, en específico, los resultados obtenidos con los LISA. En estos se aprecia un *cluster* de marginación al sur de Chihuahua rodeado por un gran *cluster* de baja vulnerabilidad socioeconómica. Por lo tanto, se seleccionó Chihuahua para investigar si el comportamiento del *cluster* de marginación encontrado se reproduce a nivel estatal. Por otro lado, se seleccionó Hidalgo ya que se aprecia un *cluster* de marginación en la Huasteca aunque no se aprecia con claridad, ya que el tamaño de las unidades no lo permite, por tal motivo se reproduce el ejercicio para observar los resultados.

Figura 5.14: Distribución del índice en Chihuahua e Hidalgo

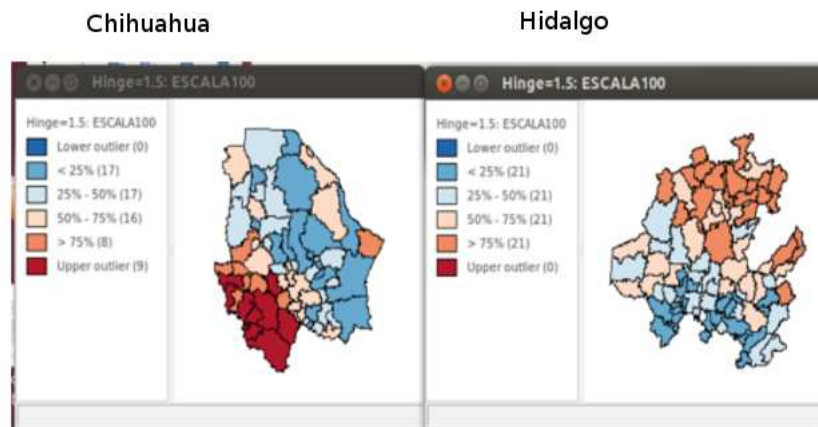


Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

La distribución del índice de marginación en Chihuahua muestra dos zonas. Al norte una

zona de baja marginación y al sur una zona de alta marginación. Esta distribución pudiera ser resultado de la cercanía con la frontera norte y las condiciones geográficas de la sierra tarahumara. En Hidalgo, se aprecia una distribución inversa, alta marginación al norte y baja al sur del Estado, distribución que pudiera ser resultado de la cercanía con el Distrito Federal y lo agreste de la sierra que atraviesa al Estado.

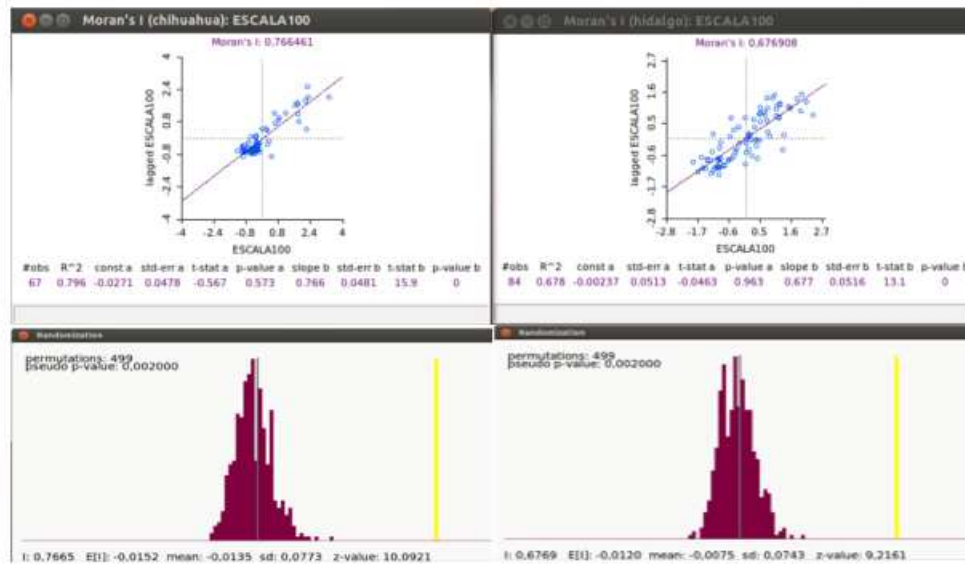
Figura 5.15: *Boxmap* del índice en Chihuahua e Hidalgo



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

El *boxmap* muestra que, para el caso de Chihuahua, la zona sur está representada por un conjunto de *outliers* ya que superan los cuartiles definidos en el Estado. Para el caso de Hidalgo, los municipios de la zona norte se ubican en el último cuartil, aunque no se presentan *outliers* todos los municipios con alta y muy alta marginación se ubican dentro del último cuartil.

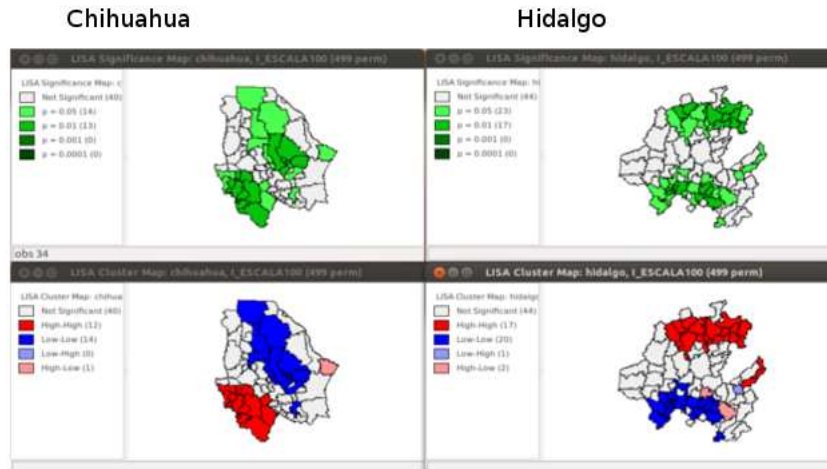
Figura 5.16: I de Moran, Chihuahua e Hidalgo



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

Para Chihuahua el I de Moran es de 0.76 con una esperanza de -0.0152, mientras que para Hidalgo es de 0.67 con una esperanza de -0.012. Al realizar el proceso de inferencia por medio de 499 permutaciones espaciales, se cuenta con un valor z de 10.09 y 9.21, respectivamente. Con este resultado, se rechaza la hipótesis nula de inexistencia de autocorrelación espacial a favor de autocorrelación espacial positiva con un *pseudo p* valor de 0.002.

Figura 5.17: LISA, Chihuahua e Hidalgo



Fuente: Elaboración propia con base en el índice de marginación en México, CONAPO, 2010

El análisis de autocorrelación espacial local coincide con el hallazgo encontrado a nivel nacional. Por medio de 499 permutaciones espaciales, se identifica un *cluster* de baja marginación al sur del Estado de Chihuahua contiguo a la gran área de baja marginación del norte. Para el caso de Hidalgo, se aprecia el mismo comportamiento en la distribución del índice, un *cluster* de alta marginación al norte y uno de baja marginación al sur, que corresponde con el gran *cluster* del Distrito Federal.

Capítulo **6**

Resultados y conclusiones

6.1. Resultados

El análisis univariado del índice de marginación por medio de MDE realizado en el capítulo anterior sugirió la existencia de patrones de distribución espacial. Dichos patrones pudieran ser el resultado de la autocorrelación espacial que muestra en su distribución. La posible existencia de autocorrelación espacial positiva se aprecia en el resultado del estadístico I de Moran y en el *pseudo p* valor de 0.0001, utilizando 9,999 permutaciones. La ecuación 5.3 muestra que la distribución de los valores observados en las unidades espaciales podrían encontrarse influenciadas por el valor de las unidades vecinas, bajo un criterio de contigüidad tipo *queen* de primer orden

EL diagrama de dispersión de Moran mostrado en el gráfico 5.7 indica el grado de dependencia espacial entre las unidades espaciales. En el gráfico 5.7 se aprecia que la gran mayoría de unidades espaciales presentan dependencia espacial positiva, pues se agrupan en los cuadrantes II y III. Estos corresponden a las unidades espaciales con valores altos rodeadas de unidades con valores también altos (cuadrante II) y unidades con valores bajos rodeadas de unidades con valores bajos (cuadrante III). Este resultado pudiera sugerir la no aleatoriedad del patrón espacial general en la distribución del índice de marginación 2010. Con base en dicho hallazgo, se refuerza la idea que la desigualdad socioeconómica espacial reflejada en la distribución del índice de marginación puede ser resultado de la falta de aleatoriedad en el patrón de distribución, aunque se debe considerar que los hallazgos encontrados son patrones sugeridos por las técnicas empleadas.

Tal como plantea Anselin (1995), al trabajar con conjuntos de datos relativamente grandes se viola el supuesto de estacionariedad y estabilidad estructural, particularmente en el análisis de asociación espacial. Al mismo tiempo, y debido a la primera ley de la geografía de Tobler, a medida que aumentan el tamaño del conjunto de datos, el indicador de asociación espacial global tiende a mostrar, de manera persistente, la existencia de autocorrelación espacial. Precisamente porque ignora la inestabilidad potencial en el conjunto de datos. Para reforzar la evidencia sobre el hallazgo de autocorrelación espacial, se recurre a las herramientas de

autocorrelación espacial local por medio de los LISA.

EL análisis de autocorrelación espacial local muestra de nuevo evidencia de agrupaciones con autocorrelación espacial positiva. EL gráfico 5.8 exhibe la existencia de *clusters* de marginación espacial. Tales *clusters* de marginación se agrupan en dos conjuntos, *high-high* y *low-low*, En el gráfico referido, se aprecia la existencia de *clusters* de baja marginación espacial al norte del país, en el bajío y en el centro del país. Es importante resaltar la existencia de un *cluster* de alta marginación espacial en la sierra tarahumara de Chihuahua el cual resalta con el gran *cluster* de baja marginación del norte del país. Al norte de Veracruz e Hidalgo y al sur del País se encuentran ubicados dos *clusters* más. El análisis LISA propone la existencia de *clusters* como resultado de la falta de aleatoriedad en la distribución del índice de marginación.

El gráfico 5.9 muestra diferencias en la distribución espacial de los componentes, aunque el gráfico 5.11 parece indicar patrones de distribución similares a los del índice global. El gráfico 5.10 muestra diferencias importantes en “Sin drenaje”, “Sin energía” y “Sin agua”. El análisis de autocorrelación espacial por medio de MDE sugiere la existencia de autocorrelación espacial positiva en todos los componentes del índice. La primera aproximación a la distribución espacial de los componentes por medio de MDE insinúa que la distribución de los componentes no es casual y es el resultado de un proceso de dependencia espacial positiva.

La figura 5.12 parece mostrar, además de la existencia de autocorrelación espacial positiva en todos los componentes, la existencia de unidades espaciales con mayor vulnerabilidad en los componentes “Sin drenaje”, “Analfabetismo”, “Piso de tierra”, “Sin energía” y “Sin agua”. En ellos, se aprecia un gran conjunto de unidades espaciales con niveles altos en los componentes rodeados de unidades con niveles altos. Este hallazgo es importante pues sugiere que la gran mayoría de las unidades espaciales presentan gran vulnerabilidad socioeconómica. Dichas unidades cuentan con indicadores altos y vecinos que también cuentan con valores altos en los indicadores mencionados.

En la figura 5.13 se aprecia que en los componentes mencionados existe una disminución

considerable en los *clusters low-low* lo que propone que los *clusters* de baja vulnerabilidad socioeconómica disminuyen en dichos indicadores, lo que agrava la situación de desigualdad socioeconómica. La figura a la que se hace referencia sugiere, también, la existencia de un patrón de distribución similar al del índice global. En todos se aprecia la existencia de *clusters low-low* al norte y *clusters high-high* al sur del país. Al igual que en la distribución del índice de marginación, la sierra tarahumara muestra indicios de vulnerabilidad socioeconómica, salvo en los componentes “Menos de dos salarios mínimos” y “Hacinamiento” donde no existe un *cluster high-high* de vulnerabilidad.

Es importante hacer mención especial del hallazgo encontrado en la Sierra Tarahumara. A pesar de que se observa un patrón global con un gran *cluster* de baja vulnerabilidad al norte y uno de alta al sur, también se aprecia, de manera local, un patrón de distribución espacial con un *cluster* de alta vulnerabilidad al sur del estado de Chihuahua. En la figura 5.13 se aprecia que en la mayoría de los componentes del índice, la Sierra Tarahumara muestra indicios de alta vulnerabilidad socioeconómica. El hallazgo del patrón de distribución espacial sugerido por medio de MDE pudiera ser resultado de la existencia de autocorrelación espacial local positiva en la distribución espacial del índice de marginación y sus componentes para el año 2010.

Este hallazgo se aprecia por medio de los *cluster map* de los componentes “Sin primaria”, “Sin drenaje” y “Sin energía”, indicadores relacionados, tradicionalmente, con las grandes aglomeraciones poblacionales. Dichos indicadores son el reflejo del grado de urbanización de pequeñas localidades y en la medida en la que su vulnerabilidad sea menor, el nivel de bienestar tiende a ser mayor. Por tal motivo, la exploración por medio de MDE parece mostrar que en la Sierra Tarahumara existe un *cluster* de alta marginación y gran vulnerabilidad socioeconómica.

Los resultados obtenidos por medio de MDE se muestran como una primer aproximación en la exploración espacial de la marginación y tan solo sugiere la posible existencia de patrones de distribución espacial. Este hallazgo se muestra como una sugerencia ya que el número de

unidades espaciales empleadas en el análisis de MDE limita la fuerza del indicador al corroborar estadísticamente la existencia de dichos patrones. Para lograr tal cometido, fue necesario reproducir el ejercicio a una escala menor. La corroboración estadística de los patrones sugeridos se realiza enseguida, por medio de una exploración de la distribución del índice de marginación, empleando un número menor de unidades espaciales.

La exploración de los casos de Chihuahua e Hidalgo por medio de AEDE permite corroborar estadísticamente los patrones identificados, los cuales coinciden con los obtenidos por medio de MDE. Para las dos entidades seleccionadas se aprecia que la distribución del índice, reflejada en la figura 5.14, parece reproducir el comportamiento obtenido en la exploración por medio de MDE. Para el caso de Chihuahua, existen unidades espaciales con valores muy bajos al norte y unidades con valores muy altos al sur del estado. En el caso de Hidalgo, el comportamiento es el contrario, aunque es el mismo que el sugerido por MDE.

Los *box map* mostrados en la figura 5.15 muestran que, para el caso de Chihuahua, las unidades espaciales ubicadas al sur del estado representan valores atípicos, ya que se encuentran por encima de los cuartiles definidos en la construcción de un diagrama de caja. En el caso de Hidalgo, aunque no se aprecian valores atípicos, el norte del estado se ubica dentro del último cuartil.

El análisis de autocorrelación espacial mostrado en la figura 5.16 corrobora la presencia de autocorrelación espacial positiva, ya que el I de Moran es de 0.76 y 0.67, para Chihuahua e Hidalgo, respectivamente. Los diagramas de dispersión mostrados en dicha figura aportan evidencia interesante. Los municipios del estado de Chihuahua muestran autocorrelación espacial positiva, ya que la mayoría de ellos se encuentran en los cuadrantes II y III del diagrama de dispersión. Este hallazgo muestra que los municipios con valores altos en el índice se encuentran rodeados de vecinos con valores altos (cuadrante II) y que los municipios con valores bajos en el índice se encuentran rodeados por vecinos con valores bajos (cuadrante III). Dicho hallazgo se obtiene por medio de un criterio de contigüidad tipo *queen* de primer orden. Para corroborar el hallazgo sugerido por el I de Moran, se realizó un proceso de inferencia

estadística basado en 499 permutaciones espaciales.

La inferencia estadística mostrada en la figura 5.16 muestra que el patrón de distribución espacial del índice de marginación para los estados de Chihuahua e Hidalgo no es casual. Al contrario, es resultado de un proceso de autocorrelación espacial positivo. Dicho patrón es estadísticamente significativo y cuenta con un *pseudo p* valor de 0.002. El análisis de autocorrelación espacial global mostrado en la figura 5.16 corrobora el patrón sugerido por MDE para los estados seleccionados.

La figura 5.16 corrobora que la distribución espacial del índice de marginación para los estados de Chihuahua e Hidalgo para el año 2010 es resultado de un proceso de autocorrelación espacial positivo y estadísticamente significativo, el patrón de distribución es congruente con el análisis exploratorio realizado a nivel nacional por medio de MDE.

La figura 5.17 muestra el análisis de autocorrelación espacial local. Se aprecian municipios con autocorrelación espacial positiva y que al mismo tiempo son estadísticamente significativos. En la figura referida, se muestran los municipios estadísticamente significativos por medio del LISA *significance map* y se diferencian entre *clusters* de alta y baja marginación por medio del LISA *cluster map*. En la figura 5.17 se muestran los *clusters* de alta y baja marginación para los estados seleccionados. Dichos *clusters* son estadísticamente significativos al 95 %.

Tal como sugirió el análisis exploratorio a nivel nacional por medio de MDE y al analizar las áreas por medio de AEDE. Los patrones de distribución espacial a nivel estatal en el índice de marginación confirma la existencia de *clusters* de marginación en la Sierra Tarahumara y en la Huasteca Hidalguense. La exploración del índice de marginación para los casos de Chihuahua e Hidalgo corrobora estadísticamente dichos patrones.

6.2. Conclusiones

La presente investigación es el resultado de la inclusión explícita de la dimensión espacial en el estudio de la marginación y del tratamiento formal del espacio en dicho fenómeno socioeconómico. Como resultado de lo anterior, se exploró la presencia de efectos espaciales en la distribución del índice de marginación, en particular autocorrelación espacial. La MDE y el AEDE permitieron investigar la existencia de patrones espaciales en la distribución del índice de marginación, explorar la presencia de efectos espaciales y contrastar los patrones de distribución del índice de marginación y sus componentes en México para el año 2010.

La presente investigación representa una visión panorámica en el estudio del espacio y su influencia en la configuración de fenómenos socioeconómicos. Los resultados que de ella se desprenden permiten sugerir la existencia de patrones espaciales en la distribución del índice de marginación y sus componentes. Dichos patrones pueden ser el resultado de la existencia de efectos espaciales.

La autocorrelación espacial positiva obtenida por medio de MDE en los indicadores estudiados permite suponer que la distribución espacial de la marginación en México no es casual. Al contrario, es resultado de la influencia que existe entre municipios vecinos. Al mismo tiempo, se planteó la posible existencia de *clusters* de alta marginación y alta vulnerabilidad socioeconómica, ya que los patrones de distribución espacial entre el índice y los componentes parecen comportarse de manera similar.

La exploración por medio de MDE se muestra como una primer aproximación en el estudio de la dimensión espacial de la marginación en México. Los resultados obtenidos por medio de MDE orientan y sugieren la presencia de patrones espaciales en la distribución de la marginación. Los patrones sugeridos indican que, tanto en el índice de marginación como en sus componentes, parece existir un patrón de distribución espacial que posiblemente se encuentra determinado por el proceso de autocorrelación espacial positiva capturado por el I de Moran.

El mismo análisis sugirió la presencia de *clusters* de alta marginación y gran vulnerabilidad socioeconómica, principalmente en la Sierra Tarahumara, en las Huastecas y en los estados del sur de la República. Este análisis propone la existencia de patrones de distribución espacial en el índice de marginación y sus componentes, al sugerir que los *clusters* pueden ser el resultado de un proceso de autocorrelación espacial local, ya que los valores observados en los municipios dentro de los *clusters* se encuentran relacionados con los valores observados en los vecinos.

La exploración por medio de MDE sugirió la existencia de patrones de distribución espacial. Se procedió a reproducir el ejercicio para dos entidades federativas, ya que el número de unidades espaciales empleado en MDE limita la fuerza del indicador de autocorrelación al intentar corroborar estadísticamente los patrones sugeridos. El AEDE realizado para los estados de Chihuahua e Hidalgo permitió corroborar estadísticamente los patrones sugeridos por MDE para dichos estados.

En la exploración de casos específicos se encontró la presencia de autocorrelación espacial positiva y estadísticamente significativa con un *pseudo p* valor de 0.002. Con lo que se corrobora que el patrón para los casos estudiados coincide con el sugerido por MDE. Además el patrón en los estados es significativo aportando evidencia de que la distribución del índice de marginación en los estados no es casual. Al contrario, es resultado de un proceso de autocorrelación espacial positivo en el que los municipios vecinos tienen una cierta influencia.

La existencia de *clusters* sugeridos por MDE coincide con los *clusters* obtenidos por AEDE para los estados seleccionados. El *cluster* de alta marginación de la Sierra Tarahumara y el de la Huasteca Hidalguense se corroboran estadísticamente lo que permite identificar aglomeraciones en las que la vulnerabilidad socioeconómica es mayor y requiere atención prioritaria. La corroboración de los patrones locales permite profundizar el entendimiento de la marginación y del proceso de distribución espacial que da como resultado un paisaje socioeconómico con diferentes niveles de vulnerabilidad socioeconómica.

Con base en lo anterior se muestra que se alcanzaron los objetivos planteados en la presente investigación al explorar, identificar y contrastar efectos espaciales y sugerir, a nivel nacional, y corroborar, a nivel estatal, patrones de distribución espacial. Los resultados muestran que la exploración de los efectos espaciales en la distribución de la marginación permite identificar como la inclusión de la dimensión espacial puede ayudar en la identificación de patrones interesantes.

Los principales hallazgos de la investigación son los sugeridos por MDE a nivel nacional y la exploración por AEDE para los estados de Chihuahua e Hidalgo. Entre los hallazgos principales se encuentran la posible existencia de autocorrelación espacial positiva en el índice y sus componentes a nivel nacional y la presencia de *clusters* de marginación por medio de MDE. Al contrastar la distribución del índice de marginación con la de sus componentes se encontró que los patrones sugeridos parecen ser similares. En los hallazgos descubiertos en los estados seleccionados se encuentra la corroboración de los patrones que coinciden con los sugeridos por MDE, la identificación de autocorrelación espacial positiva y estadísticamente significativa y la identificación de *clusters* de marginación estadísticamente significativos que se corresponden con lo sugerido por MDE.

La presente investigación ha buscado aportar evidencia empírica a la ciencia social espacialmente integrada (“*SISS*”) contribuyendo con la identificación de hallazgos en la dimensión espacial de la marginación. Con esto, la investigación contribuye con el estudio de fenómenos socioeconómicos analizados desde la perspectiva espacial. La investigación se encuentra en línea con las investigaciones realizadas en el Laboratorio de Análisis Socioeconómico-Espacial (Spatial SEALab⁷), en el cual se enfatiza la importancia de las características espaciales en fenómenos socioeconómicos.

Con base en los hallazgos identificados en la presente investigación se sugiere, como futuras investigaciones, explorar la existencia de heterogeneidad espacial en la distribución del índice y sus componentes, explorar la existencia de autocorrelación espacial multivariada para

⁷<http://www.spatial-sealab.org/>

identificar los vínculos espaciales entre diferentes indicadores. Por último, se sugiere realizar un análisis confirmatorio espacial del índice de marginación y sus componentes en México para el año 2010.

Bibliografía

- Acevedo, Ingrid y Velásquez, Ermilson (2008). «Algunos conceptos de la econometría espacial y el análisis exploratorio de datos espaciales». *Ecós de Economía*, Número 27, pp. 1–236.
- Anselin, Luc. (1988). *Spatial econometrics: methods and models*. Kluwer Academic Publishers.
- (1989). *What is Special about Spatial Data?: Alternative Perspectives on Spatial Data Analysis*. National Center for Geographic Information and Analysis Santa Barbara, CA.
- (1992). «Spatial data analysis with GIS: an introduction to application in the social sciences», National Center for Geographic Information & Analysis (US).
- Anselin, Luc (1994). «Exploratory spatial data analysis and geographic information systems». *New tools for spatial analysis*, 54.
- Anselin, Luc. (1995). «Local indicators of spatial association, LISA». *Geographical analysis*, 27(2), pp. 93–115.
- Anselin, Luc (1996). «The Moran scatterplot as an ESDA tool to assess local instability in spatial association». En: Scholten H. Fischer, M. y D Unwin (Eds.), *Spatial Analytical Perspectives on GIS in Environmental and Socio-Economic Sciences*, pp. 111–125. Taylor and Francis, London..
- Anselin, Luc. (1998). «Exploratory spatial data analysis in a geocomputational environment». *Geocomputation, a Primer*, Wiley, New York, pp. 77–94.

- (1999). «The future of spatial analysis in the social sciences». *Annals of GIS*, 5(2), pp. 67–76.
- Anselin, Luc y Getis, Arthur (1992). «Spatial statistical analysis and geographic information systems». *The Annals of Regional Science*, 26(1), pp. 19–33.
- Anselin, Luc.; Sridharan, S. y Gholston, S. (2007). «Using exploratory spatial data analysis to leverage social indicator databases: the discovery of interesting patterns». *Social Indicators Research*, 82(2), pp. 287–309.
- Anselin, Luc. *et al* (2006). «How (not) to lie with spatial statistics». *American journal of preventive medicine*, 30(2), p. 3.
- Ávila, José Luis; Fuentes, Carlos y rodolfo Tuirán (2000). «Índice de marginación 2000: Metodología». *CONAPO*, 196.
- Bertaud, A. y Malpezzi, S. (2003). «The spatial distribution of population in 48 world cities: Implications for economies in transition». *Disponible sur <http://www.alain-bertaud.com>*.
- Bistrain, César (2008). *Análisis de la marginación en la población rural de México, 1990-2000*. Tesis de maestría, FLACSO.
- Buzai, Gustavo D. y Baxendale, Claudia A. (2008). «Análisis exploratorio de datos espaciales educativos: aplicación a la ciudad de Luján». En: Antonio Moreno Jiménez; Gustavo D. Buzai (Coordinadores) (Ed.), *Análisis y planificación de servicios colectivos con Sistemas de Información Geográfica*, p. 158 páginas. UAM-UNLU, Madrid, [versión electrónica]^a edición.
- Caliber, Associates (2001). «Exploratory and Confirmatory Spatial Data Analysis Approaches to Studying the Correlates of Juvenile Violent Crimes,». *ICF International*, Volume II Final Report.
- Camberos, Mario y Bracamontes, Joaquín (2007). «Marginación y políticas de desarrollo social: un análisis regional para Sonora». *Problemas del Desarrollo*, 38(149), pp. 113–135.

- Campoy, M. L. (1995). «Marginación y pobreza». *Revista del Ministerio de Trabajo y Asuntos Sociales*, 2, p. 67.
- Chasco Yrigoyen, Coro (2003). *Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales*. Comunidad de Madrid. Tesis doctoral.
- Chen, J.; Chen, Y.; Yu, J. y Yang, Z. (2011). «Comparisons with spatial autocorrelation and spatial association rule mining». En: *Spatial Data Mining and Geographical Knowledge Services (ICSDM), 2011 IEEE International Conference on*, pp. 32–37. IEEE.
- Comité técnico para la medición de la pobreza (2002). «Medición de la pobreza: variantes metodológicas y estimación preliminar». *serie: documentos de investigación*.
- Coplamar (1982). *Necesidades esenciales en México: Geografía de la marginación*. volumen 5. Siglo XXI.
- Cortés, F. (2006). «Consideraciones sobre la marginación, la marginalidad, marginalidad económica y exclusión social». *Papeles de población*, 47, pp. 71–84.
- De Jong, P.; Sprenger, C. y Veen, F. (1984). «On extreme values of Moran's I and Geary's c». *Geographical Analysis*, 16(1), pp. 17–24.
- De la Vega, Sergio; Romo, Raúl y González, Ana (2010). «Índice de marginación por entidad federativa y municipio 2010». *CONAPO*, 332.
- Fischer, Manfred M (2001). *Spatial Analysis in Geography*. volumen 22, capítulo capítulo 2, pp. 14752–14758. Pergamon.
- González, R.M.M.; de Fuentes, A.G. y Montejano, M.D.C. (2004). «Mortalidad infantil y marginación en la península de Yucatán». *Investigaciones Geográficas (Mx)*, (054), pp. 140–163.
- Goodchild, M.F.; Anselin, L.; Appelbaum, R.P. y Harthorn, B.H. (2000). «Toward spatially integrated social science». *International Regional Science Review*, 23(2), pp. 139–159.
- Goodchild, M.F. y Janelle, D.G. (2004). *Spatially integrated social science*. Oxford University Press, USA.

- Goodchild, Michael. (1991). «Geographic Information systems». *Progress in human geography*, 15, No. 2, pp. 194–200.
- Green, Ann (2011). «Spatial inequality and skills in a changing economy».
- Griffith, D.A. (2002). «Spatial autocorrelation». En: *Encyclopedia of social measurement*, Association of American Geographers Washington, DC.
- (2003). *Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization*. Springer Verlag.
- Haining, Robert y Wise, Stephen (1997). «Exploratory Spatial Data Analysis». [electrónico] December 05, NCGIA Core Curriculum in GIScience, <http://www.ncgia.ucsb.edu/gisc/units/u128/u128.html>.
- Hernández Franco, Daniel; Corona, Mónica Orozco y Báez, Sirenia Vázquez (2007). «Métodos de focalización en la política social en México. Un estudio comparativo». [Versión electrónica].
- Isard, W. et al (1970). *General theory: social, political, economic and regional*. Cambridge, Mass./London: Mass. Inst. Technol..
- Isard, Walter (1956). *Location and space-economy*. The M. I. T. Press.
- Jiménez Montero, Manuel de J.; Dávila, Juan Pablo Martínez; Rosado, Octavio Ruiz y López, Felipe Gallardo (2004). «Dinámica socioeconómica en la región del norte de Coclé». *Revista Mexicana del Caribe*, vol. IX(número 017), pp. pp. 67–94. Universidad de Quintana Roo, Chetumal, México.
- Kim, S. (2008). «Spatial Inequality and Economic Development: Theories, Facts, and Policies». *Urbanization and Growth*, pp. 133–166.
- Le Gallo, J. y Ertur, C. (2003). «Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995». *Papers in regional science*, 82(2), pp. 175–201.
- Lee, J. y Wong, D.W.S. (2001). *Statistical analysis with ArcView GIS*. John Wiley New York.

- Lobao, Linda; Hooks, Gregory y Tickamyer, Ann (2007). *The sociology of spatial inequality*. State University of New York Press.
- López, H.O.D.; Pérez, H.J.S. y Guzmán, L.A.M. (1996). «Uso de un índice de bienestar social para la planificación de la salud a nivel municipal». *Salud Pública de México*, (004), pp. 257–267.
- López-Calva, L.F.; Méndez Martínez, Á.; Rascón Ramirez, E.G.; Rodríguez-Chamussy, L. y Székely Pardo, M. (2008). «El ingreso de los hogares en el mapa de México». *El Trimestre Económico*, 75(300), pp. 843–896.
- Lowe, Jonathan W. (2003). «Emerging tools and concepts of exploratory spatial data analysis». http://www.giswebsite.com/pubs/200307/nr200307_p1.html
- Martínez, Germán y López, Carlos (2009). «Marginación, pobreza y migración en México: el caso de la población indígena de Chiapas.» En: *Asociación de Estudios Latinoamericanos*, Rio de Janeiro, Brazil.
- McVicar, D. (2000). «Marginalized Young People and Social Inclusion Policy in Northern Ireland». *Regional Studies*, 34(9), pp. 883–888.
- Mennis, J. y Guo, D. (2009). «Spatial data mining and geographic knowledge discovery—An introduction». *Computers, Environment and Urban Systems*, 33(6), pp. 403–408.
- Mojarro, O.; Tapia, G.; Hernández, F. y Macín, I. (2003). «Mortalidad materna y marginación municipal». *La situación demográfica de México 2003*, pp. 133–141.
- Montes, V.J. y Ortega, E. (2006). «Ubicación de la marginación en el área metropolitana de Monterrey». *Tópicos selectos en políticas de bienestar social*, 2, p. 125.
- Moreno Jiménez, María del Pilar (2001). *Psicología de la marginación social: concepto, ámbitos y actuaciones*.
- Moreno Serrano, R. y Vayá, E. (2002). «Econometría espacial; nuevas técnicas para el análisis regional: una aplicación a las regiones europeas». *Investigaciones regionales*, (1), pp. 83–106.

- Murie, Alan y Mustered, Sako (2004). «Social exclusion and opportunity structures in european cities and neighbourhoods». *Urban studies*, Vol. 41(No. 8), pp. pp. 1441–1459.
- Oliveau, S.; Guilmoto, C. *et al* (2005). «Spatial correlation and demography.».
- O’Loughlin, J. (1983). «Spatial inequalities in western cities: a comparison of north american and german urban areas». *Social indicators research*, 13(2), pp. 185–212.
- (2002). «The Electoral Geography of Weimar Germany: Exploratory Spatial Data Analyses (ESDA) of Protestant Support for the Nazi Party». *Political Analysis*, 10(3), pp. 217–243.
- Pardo, M.S.; López-Calva, L.F.; Meléndez, Á.; Martínez, E.G. y Rodríguez-Chamussy, L. (2007). «Poniendo a la pobreza de ingresos ya la desigualdad en el mapa de México». *economía mexicana NUEVA ÉPOCA*, 16(2).
- Park, R.E. (1928). «Human migration and the marginal man». *American journal of sociology*, pp. 881–893.
- Pérez, Marco. (2010). «Un análisis de pobreza regional para la Huasteca Hidalguense». *Tiempo económico*, primer cuatrimestre(14), pp. 33–52.
- Peña, E.G.R.; Lozano, C.V.; Martínez, P.C.C.; García, G.S.; Rodríguez, L.G.G. y Castañón, L.N.B. (2007). «Índice de marginación y el patrón de consumo alimentario familiar de Nuevo León». *Papeles de Población*, (54), pp. 265–285.
- Pratschke, Jonathan y Haase, Trutz (2007). «Measurement of social disadvantage and its spatial articulation in the republic of ireland». *Regional Studies*, 41.6, pp. 719–734.
- Rey, Sergio (2004). «Spatial analysis of regional income inequality». En: Goodchild M. y Janelle D. (Eds.), *Spatially Integrated Social Science: Examples in Best Practice*, pp. 280–299. Oxford University Press.
- Rey, Sergio; Anselin, L.; Folch, D.C.; Arribas-Bel, D.; Sastré-Gutiérrez, M.L. y Interlante, L. (2011). «Measuring spatial dynamics in metropolitan areas». *Economic Development Quarterly*, 25(1), pp. 54–64.

- Rey, Sergio y Sastré-Gutiérrez, M. L. (2010). «Interregional inequality dynamics in Mexico». *Spatial Economic Analysis*, Vol. 5(No. 3), Routledge.
- Salcedo, Soledad (2004). «Perfil de la marginación en el Distrito Federal». En: Dirección General de Equidad y Desarrollo Social (Ed.), *Pobreza, desigualdad y marginación en la Ciudad de México*, capítulo 2. Orientaciones y críticas a los instrumentos de medición de la pobreza en la ciudad de México, pp. 56–66. SEDESOL.
- Sastré-Gutiérrez, M. L. (2011). «Spatial Socioeconomics: an integrated analytical framework». *Spatial Socio-Economics Laboratory (Spatial SEALab)*, Working Paper.
- Sastré-Gutiérrez, M. L. y Rey, S. (2008). «Polarización espacial y dinámicas de la desigualdad intrarregional en México». *Problemas del Desarrollo*, 39(155).
- Sastré-Gutiérrez, M. L. y Rey, S. J. (2013). «Space-Time Income Distribution in Mexico». *The Annals of GIS, Geographic Information Systems*, forthcoming.
- Shekhar, S.; Zhang, P.; Huang, Y. y Vatsavai, R.R. (2003). «Trends in spatial data mining». *Data mining: Next generation challenges and future directions*, pp. 357–379.
- Solís, P. (2002). «Marginación urbana». *Situación Demográfica de México*, pp. 113–120.
- Stonequist, E.V. (1935). «The problem of the marginal man». *American journal of sociology*, pp. 1–12.
- Sánchez Peña, Landy (2006). «Métodos para el Análisis Espacial. Una Aplicación al Estudio de la Geografía de la Pobreza». En: *Mesa 10: Métodos, Técnicas y Datos en la Investigación Demográfica en América Latina y El Caribe, II Congreso de la Asociación Latinoamericana de Población*, [Versión electrónica], Guadalajara, México.
- Tickamyer, A.R. (2000). «Space matters! Spatial inequality in future sociology». *Contemporary Sociology*, 29(6), pp. 805–813.
- Tiefelsdorf, M. (1998). *Modelling spatial processes: The identification and analysis of spatial relationships in regression residuals by means of Moran's I*. Springer.

Zhu, X. y Davidson, I. (2007). *Knowledge Discovery and Data Mining: Challenges and Realities*. Igi Global.